ПРОБЛЕМЫ УПРАВЛЕНИЯ

3.2005

Научно-технический журнал
6 номеров в год

УЧРЕДИТЕЛЬ
Институт проблем управления им. В.А. Трапезникова РАН

Издатель
ООО «СенСиДат-Контрол»
Ген. директор Н.Н. Кузнецова

Адрес редакции
117997, ГСП-7, Москва,
ул. Профсоюзная, д. 65, к. 104.
Тел./факс (095) 330-42-66,
тел. (095) 334-92-00

E-mail: datchik@ipu.ru
www.ipu.ru/period/pu

Оригинал-макет и электронная версия подготовлены ООО «ЭЛЕКТРОНИНФОРМ»

Отпечатано с готовых диапозитивов в типографии ГКС

Подписано в печать 11.04.2005 г.

Заказ № РВ305

Журнал зарегистрирован в Министерстве Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций

Свидетельство о регистрации ПИ №77-11963 от 06 марта 2002 г.

Журнал входит в Перечень ВАК (Бюл. — 2004. — № 3)

Подписные индексы:
81708 в каталоге Роспечати
38006 в объединенном каталоге «Пресса России»

СОДЕРЖАНИЕ

УДК 62-501.52

О ДОСТАТОЧНЫХ И НЕОБХОДИМЫХ УСЛОВИЯХ АСИМПТОТИЧЕСКОЙ УСТОЙЧИВОСТИ НЕЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМ

В. П. Жуков

Институт проблем управления им. В. А. Трапезникова, г. Москва

Получены новые достаточные и необходимые условия асимптотической устойчивости состояний равновесия автономных динамических объектов, описываемых системой нелинейных обыкновенных дифференциальных уравнений типа Коши произвольного порядка. Применение специального класса функций (вместо функций Ляпунова) позволило получить обращение теоремы об асимптотической устойчивости, имеющее ясный геометрический смысл и наглядность.

ВВЕДЕНИЕ

Обращению теоремы Ляпунова об асимптотической устойчивости посвящено много работ. Подробно этот вопрос рассматривался в монографии [1], где приведена обширная литература. Из нее видно, что указанное обращение достигается путем построения для каждой асимптотически устойчивой системы весьма сложных функций, удовлетворяющих требованиям теоремы. В данной работе получены необходимые и достаточные условия асимптотической устойчивости для автономных систем, которые отличаются от рассматриваемых ранее тем, что в них вместо класса функций Ляпунова применяется класс функций, зависящих от параметра. Благодаря этому для любой асимптотически устойчивой системы удается указать весьма простую в смысле геометрической интерпретации непрерывную функцию, удовлетворяющую новым необходимым и достаточным условиям. Обращение теоремы об асимптотической устойчивости приобретает ясный геометрический смысл и большую наглядность, что может облегчить нахождение новых используемых функций. В данной работе в качестве иллюстрации предлагаемого подхода приводятся результаты для автономных асимптотически устойчивых систем. В дальнейшем предполагается развитие аналогичного подхода для исследования на асимптотическую устойчивость нелинейных неавтономных динамических систем.

1. ОПИСАНИЕ КЛАССА ИСПОЛЬЗУЕМЫХ ФУНКЦИЙ

В рассмотрение вместо функций Ляпунова вводится класс функций $V(\mathbf{x}, \varepsilon)$, зависящих от параметра ε, что позволяет для каждой асимптотически устойчивой системы указать простую в смысле геометрической трактовки функцию $S(\mathbf{x}, \varepsilon)$; для большого класса асимптотически устойчивых систем можно указать еще более простую функцию $S(\mathbf{x})$. Подробно свойства этих простых функций будут рассмотрены при их построении в доказательстве необходимости условий сформулированной далее теоремы. Здесь же скажем, что геометрический смысл функции $S(\mathbf{x}, \varepsilon)$ состоит в том, что она означает расстояние от данной точки $\mathbf{x}$ до границы некоторой окрестности асимптотически устойчивой точки равновесия (эта окрестность определяется параметром ε), измеренное вдоль траектории; функция $S(\mathbf{x})$ имеет смысл расстояния от данной точки $\mathbf{x}$ до точки равновесия, измеренного вдоль траектории (длина положительной полутраектории). Функции $S(\mathbf{x})$ не могут применяться, например, в случае устойчивых фокусов (на фазовой плоскости), у которых длина положительных полутраекторий бесконечна (функция $S(\mathbf{x})$ при этом не имеет смысла); в этом случае применима функция $S(\mathbf{x}, \varepsilon)$.

Будем рассматривать непрерывные неотрицательные скалярные функции $V(\mathbf{x}, \varepsilon)$ векторного аргумента $\mathbf{x}$, заданные в некоторой области $G_V \subseteq G$, являющейся окрестностью точки равновесия $\mathbf{x}_* = 0$, и зависящие в общем случае от скалярного параметра $\varepsilon \in R^+$, где $R^+ = [0, \infty)$ — положительная полуось вещественной числовой оси. Говоря иначе, рассматриваются семейства скалярных функций по параметру ε (в случае, когда функция $V(\mathbf{x}, \varepsilon) = V(\mathbf{x})$ не зависит от ε, будем считать, что семейство состоит из одной этой функции). Класс рассматриваемых функций характеризуется следующими свойствами. При любом значении параметра $\varepsilon > 0$ существует такое множество $\delta_\varepsilon \subset G_V$, содержащее точку $\mathbf{x}_* = 0$, что $V(\mathbf{x}, \varepsilon) = 0$ при $\mathbf{x} \in \delta_\varepsilon$. При этом, если функция V явно зависит от $\varepsilon > 0$, то множество δ_ε является компактной окрестностью точки $\mathbf{x}_* = 0$. В случае функции $V(\mathbf{x})$, не зависящей от ε, множество δ_ε состоит лишь из точки $\mathbf{x}_* = 0$. Предполагается, что при $\varepsilon \to 0$ семейство δ_ε сходится к точке $\mathbf{x}_* = 0$ ($\dim\delta_\varepsilon = \max\limits_{\mathbf{x}_i, \mathbf{x}_j \in \delta_\varepsilon} \rho(\mathbf{x}_i, \mathbf{x}_j) \to 0$ при $\varepsilon \to 0$, где ρ — расстояние между точками x_i и x_j), т. е. для любого шара s с центром в этой точке существует такое число ε_s, что все множества δ_ε, соответствующие значениям $\varepsilon \leqslant \varepsilon_s$, принадлежат этому шару. При любом фиксированном ε функция $V(\mathbf{x}, \varepsilon)$ является положительной при любом $\mathbf{x} \in G_V \backslash \delta_\varepsilon$. Предполагается, что при любом $\varepsilon > 0$ производная $\dot V(\mathbf{x}, \varepsilon)$ в силу системы (1) — см. § 2 — на множестве $G_V \backslash \delta_\varepsilon$ непрерывна по $\mathbf{x}$ (соответственно, производная $\dot V(\mathbf{x})$ непрерывна вне точки $\mathbf{x}_* = 0$). Для этого функция $V(\mathbf{x}, \varepsilon)$ не обязательно должна быть непрерывно дифференцируемой на множестве $G_V \backslash \delta_\varepsilon$; соответствующим примером служит упомянутая выше функция $S(\mathbf{x}, \varepsilon)$, для которой $\dot S(\mathbf{x}, \varepsilon) = -|\mathbf{f}(\mathbf{x})|$, если точка $\mathbf{x}_* = 0$ асимптотически устойчива. Этих требований, налагаемых на производную $\dot V(\mathbf{x}, \varepsilon)$, достаточно для доказательства приведенной далее теоремы. Однако, учитывая ее применение для исследования конкретных систем, будем считать функции $V(\mathbf{x}, \varepsilon)$ непрерывно дифференцируемыми на $G_V \backslash \delta_\varepsilon$ (подробнее об этом далее).

Зависящую от параметра $\varepsilon > 0$ функцию $V(\mathbf{x}, \varepsilon)$ с указанными свойствами (включая сходимость множества δ_ε к точке $\mathbf{x}_* = 0$) будем называть δ_ε — определенно-положительной на множестве G_V. Функции Ляпунова $V(\mathbf{x})$ являются, очевидно, частным случаем δ_ε — определенно-положительных функций.

Характер класса функций $V(\mathbf{x}, \varepsilon)$, используемых в теореме о достаточных и необходимых условиях асимптотической устойчивости (см. § 2), существенно зависит от характера множеств δ_ε, используемых при построении для любой асимптотически устойчивой точки системы (1) — см. далее — функции $S(\mathbf{x}, \varepsilon)$, удовлетворяющей этой теореме и имеющей простой геометрический смысл. Удалось указать такие множества δ_ε, при которых для любой асимптотически устойчивой точки можно построить непрерывную функцию $S(\mathbf{x}, \varepsilon)$, что позволило сформулировать приведенную в этой работе теорему в терминах класса непрерывных функций $V(\mathbf{x}, \varepsilon)$ (непрерывные функции $S(\mathbf{x}, \varepsilon)$ и соответствующие им множества δ_ε будут описаны при доказательстве необходимости условий данной далее теоремы). При иных множествах δ_ε функция $S(\mathbf{x}, \varepsilon)$ для некоторых асимптотически устойчивых точек равновесия оказывается разрывной (разрывы имеют место при переходе с некоторой траектории на соседние к ней траектории, вдоль каждой траектории функция $S(\mathbf{x}, \varepsilon)$ является гладкой и имеющей, как увидим далее, производную $\dot S = -|\mathbf{f}(\mathbf{x})|$, $\mathbf{x} \in G_V \backslash \delta_\varepsilon$; эти разрывные функции имеют такой же простой геометрический смысл, как и непрерывные). Вследствие этого теорему для этого случая пришлось бы сформулировать либо в терминах разрывных функций $V(\mathbf{x}, \varepsilon)$, которые являются гладкими вдоль траекторий и разрывными при переходе с одной траектории на другие, либо в терминах разрывных функций более общего типа, имеющих разрывы и на траекториях. Но это существенно усложняет применение теоремы для исследования конкретных систем (1). Исходя из высказанных соображений, рассматривается именно класс непрерывных параметрических функций $V(\mathbf{x}, \varepsilon)$.

При доказательстве достаточности условий приведенной далее теоремы вопрос о вычислении производной $\dot V(\mathbf{x}, \varepsilon)$ в силу системы (1) не возникает, так как доказательство опирается лишь на факт ее существования, гарантированный условиями теоремы. При доказательстве необходимости этих условий вопрос о вычислении этой производной возникает и разрешается, как увидим далее, весьма просто благодаря тому, что для каждой асимптотически устойчивой точки равновесия мы указываем функцию $S(\mathbf{x}, \varepsilon)$ с производной в силу системы (1) в виде $\dot S = -|\mathbf{f}(\mathbf{x})|$, $\mathbf{x} \in G_V \backslash \delta_\varepsilon$, которая легко вычисляется. Применение достаточных условий теоремы для исследования конкретных систем вида (1) на асимптотическую устойчивость требует умения вычислять производную в силу системы от функций $V(\mathbf{x}, \varepsilon)$, выбираемых в качестве претендентов для удовлетворения условиям теоремы. В этом случае мы не можем считать, что $V(\mathbf{x}, \varepsilon) = S(\mathbf{x}, \varepsilon)$ и $\dot S = -|\mathbf{f}(\mathbf{x})|$, ибо это верно только, если заранее известно, что исследуемая точка равновесия асимптотически устойчива. Однако заранее это не известно. Поэтому при применении достаточных условий в такой неопределенной ситуации производную в силу системы следует вычислять по общей формуле $\dot V(\mathbf{x}, \varepsilon) = (\operatorname{grad} V(\mathbf{x}, \varepsilon), \mathbf{f}(\mathbf{x}))$, которая из-за необходимости вычисления градиента $V(\mathbf{x}, \varepsilon)$ требует гладкости функции $V(\mathbf{x}, \varepsilon)$ по $\mathbf{x}$. Исходя из сказанного, будем требовать гладкости функций $V(\mathbf{x}, \varepsilon)$ на множестве $G_V \backslash \delta_\varepsilon$ при каждом значении $\varepsilon > 0$. На границе $\Gamma(\delta_\varepsilon)$ множества δ_ε эта функция может быть негладкой. Заметим, что при $\mathbf{x} \in \operatorname{Int}\delta_\varepsilon$ производная $\dot V(\mathbf{x}, \varepsilon) = 0$ в силу того, что $V(\mathbf{x}, \varepsilon) = 0$ при $\mathbf{x} \in \delta_\varepsilon$.

2. ДОСТАТОЧНЫЕ И НЕОБХОДИМЫЕ УСЛОВИЯ АСИМПТОТИЧЕСКОЙ УСТОЙЧИВОСТИ

Рассматривается автономный динамический объект, описываемый системой обыкновенных нелинейных дифференциальных уравнений произвольного порядка

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \ \mathbf{x} \in R^n, \qquad (1)$$

имеющей изолированную точку равновесия $\mathbf{x}_* = 0$. Функция $\mathbf{f}(\mathbf{x})$ задана в некоторой области $G \subseteq R^n$, являющейся окрестностью точки $\mathbf{x}_* = 0$, непрерывна и обеспечивает единственность решения. Как правило, будем считать ее гладкой (класса C^1).

Теорема. *Точка равновесия* $\mathbf{x}_* = 0$ *системы* (1) *асимптотически устойчива тогда и только тогда, если существует такая* δ_ε *— определенно-положительная функция* $V(\mathbf{x}, \varepsilon)$, *что при любом* $\varepsilon > 0$ *ее полная производная по времени в силу системы* (1) *отрицательна на множестве* $G_V \backslash \delta_\varepsilon$.

Если функция $V(\mathbf{x}, \varepsilon) = V(\mathbf{x})$ не зависит от параметра, то получаем формулировку обращенной теоремы Ляпунова об асимптотической устойчивости. Справедливость такой формулировки доказана ранее в ряде работ обращением теоремы Ляпунова об асимптотической устойчивости. При доказательстве приведенной теоремы доказана, в частности, справедливость достаточности условий обращенной теоремы Ляпунова, но необходимость условий этой теоремы не доказана, ибо при доказательстве необходимости условий ставилась цель получить для каждой асимптотически устойчивой системы (1) функцию $V = S$, характеризующуюся определенной геометрической интерпретацией, что для некоторых асимптотических устойчивых систем можно сделать, используя лишь функции $S(\mathbf{x}, \varepsilon)$, явно зависящие от параметра ε, к которым функция Ляпунова $V(\mathbf{x})$ не относится.

При доказательстве необходимости условий теоремы (см. § 4) указаны требования, при выполнении которых для асимптотически устойчивой точки равновесия системы (1) на любом компакте $B \subset G_п$ ($G_п$ — область притяжения точки равновесия) существует определенно-положительная непрерывная функция $S(\mathbf{x})$, имеющая на множестве B непрерывную по $\mathbf{x}$ производную в силу системы $\dot{S} = -|\mathbf{f}(\mathbf{x})|$, а также являющаяся на множестве $B \backslash (\mathbf{x}_* = 0)$ непрерывно дифференцируемой по $\mathbf{x}$ (функция Ляпунова типа $S(\mathbf{x})$). Эти требования выполняются в классе линейных систем с постоянными коэффициентами (см. доказательство теоремы в § 4). Поэтому получаем

Следствие. *Для любой асимптотически устойчивой точки равновесия линейной системы с постоянными коэффициентами* (1) *на любом множестве* $B \subset G_п$ *существует непрерывная функция Ляпунова типа* $S(\mathbf{x})$, *имеющая на множестве* B *непрерывную по* $\mathbf{x}$ *производную в силу системы* (1) *в виде* $\dot{S} = -|\mathbf{f}(\mathbf{x})|$ *и непрерывно дифференцируемая на множестве* $B \backslash (\mathbf{x}_* = 0)$.

Таким образом, в случае исследования на асимптотическую устойчивость линейных систем с постоянными коэффициентами нет необходимости прибегать к поиску функций $S(\mathbf{x}, \varepsilon)$, явно зависящих от параметра. В этом случае целесообразнее искать более простые функции Ляпунова типа $S(\mathbf{x})$, характеризующиеся простым геометрическим смыслом.

3. ПРИМЕР

Рассмотрим конкретную нелинейную динамическую систему с асимптотически устойчивой точкой равновесия, для которой не существует функции Ляпунова типа $S(\mathbf{x})$, но существует функция $S(\mathbf{x}, \varepsilon)$, зависящая от параметра. Пусть система описывается в полярных координатах дифференциальными уравнениями $\dot{\rho} = -\rho^n$, $\dot{\theta} = 1$, где n — любое положительное число. Точка равновесия, соответствующая значению $\rho = 0$, является асимптотически устойчивым фокусом. За время $dt > 0$ функция $S(\rho, \varepsilon)$ получает отрицательное приращение

$$dS(\rho, \varepsilon) = -\sqrt{(\rho d\theta)^2 + (d\rho)^2} = -\sqrt{(\rho dt)^2 + (\rho^n dt)^2} =$$
$$= -\rho dt\sqrt{1 + \rho^{2(n-1)}} = \frac{d\rho}{\rho^{n-1}}\sqrt{1 + \rho^{2(n-1)}},$$
$$\text{где } \sqrt{(\rho d\theta)^2 + (d\rho)^2} = |dS(\rho, \varepsilon)|.$$

При $n = 1$ приращение $dS(\rho, \varepsilon) = d\rho$. Интегрируя правую часть этого выражения от окружности радиуса ρ до окружности радиуса ε, а левую от $S(\rho, \varepsilon)$ до 0, получаем зависящую от параметра ε функцию $S(\rho, \varepsilon) = \sqrt{2}(\rho - \varepsilon)$. Если интегрировать правую часть от окружности радиуса ρ до точки $\rho = 0$, то получим функцию Ляпунова $S(\rho) = \sqrt{2}\rho$. Заметим, что $S(\rho) = \lim\limits_{\varepsilon \to 0} S(\rho, \varepsilon)$.

Если же $n = 2$, то $dS(\rho, \varepsilon) = d\rho\sqrt{1 + 1/\rho^2}$ и существует функция $S(\rho, \varepsilon) = \sqrt{1 + \rho^2} - \sqrt{1 + \varepsilon^2} + \ln\dfrac{(1 + \sqrt{1 + \varepsilon^2})\rho}{(1 + \sqrt{1 + \rho^2})\varepsilon}$,

но в силу расходимости интеграла $\int\limits_\rho^0 \sqrt{1 + 1/\rho^2}\, d\rho$ не существует функции $S(\rho)$ с конечными значениями, что видно также из того, что $\lim\limits_{\varepsilon \to 0} S(\rho, \varepsilon) = \infty$. Таким образом, при $n = 2$ не существует функции Ляпунова типа $S(\rho)$, но существует зависящая от параметра функция $S(\rho, \varepsilon)$. Оценки показывают, что этот же факт имеет место при $n > 2$. Этот факт связан с тем (как показано в доказательстве теоремы), что при $n \geqslant 2$ длина любой положительной полутраектории, лежащей в сколь угодно малой окрестности точки равновесия, имеет бесконечное значение.

4. ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ

Достаточность. Покажем, что если условия теоремы выполнены, то точка $\mathbf{x}_* = 0$ асимптотически устойчива. Для этого сначала докажем обычную устойчивость этой точки, т. е. докажем, что при условиях теоремы для любой окрестности $\eta \subset G_V$ точки $\mathbf{x}_* = 0$ существует такая окрестность $q \subset \eta$ этой точки, что любое решение $\mathbf{x}(t)$ уравнения (1), начинающееся в момент времени t_0 в q, не покинет окрестность η при любом $t_0 > t$.

Пусть $\eta \subset G_V$ — любая окрестность точки $\mathbf{x}_* = 0$. Искомое множество $q \subset \eta$ построим с помощью следующего способа, который применим как в случае, когда теореме удовлетворяет функция $V(\mathbf{x}, \varepsilon)$, явно зависящая от параметра ε, так и в случае, когда ей удовлетворяет не зависящая от этого параметра функция $V(\mathbf{x})$. Рассмотрим для обоих этих случаев множества $X_\alpha = \{\mathbf{x}\colon V \leqslant \alpha \geqslant 0\}$, где $V = V(\mathbf{x}, \varepsilon)$ в первом случае и $V = V(\mathbf{x})$ во втором случае. При любом фиксированном $\varepsilon > 0$ в силу непрерывности функции V имеем: $X_\alpha \to \delta_\varepsilon$ при $\alpha \to 0$

(в случае функции $V(\mathbf{x})$ множество δ_ε состоит из точки $\mathbf{x}_* = 0$), т. е. для любого числа $\Delta > 0$ существует такое число $\alpha_\Delta > 0$, что при любом $\alpha \leqslant \alpha_\Delta$ $X_\alpha \subset \delta_{\Delta\varepsilon}$, где $\delta_{\Delta\varepsilon}$ — Δ — окрестность множества δ_ε, определяемая как $\delta_\varepsilon = \bigcup_{\mathbf{x} \in \delta_\varepsilon} \Delta(\mathbf{x})$, где $\Delta(\mathbf{x})$ — шар радиуса Δ с центром в точке $\mathbf{x}$. Действительно, на компактном множестве $g_{\Delta\varepsilon} = \overline{G_{1V}/\delta_{\Delta\varepsilon}}$ непрерывная положительная функция V достигает своей нижней грани $\inf\limits_{\mathbf{x} \in g_{\Delta\varepsilon}} V = C_{\Delta\varepsilon} > 0$ $(\bar{G}_{1V} \subset G_V,\ \delta_\varepsilon \subset G_{1V})$. Возьмем для любого $\Delta > 0$ положительное число $\alpha_\Delta < C_{\Delta\varepsilon}$ (например, $\alpha_\Delta = 0{,}5C_{\Delta\varepsilon}$). Очевидно, что при любом $\alpha < \alpha_\Delta$ $X_\alpha = \{\mathbf{x}\colon V \leqslant \alpha\} \subset \delta_\varepsilon$ (точки $\mathbf{x} \in X_\alpha$ не могут лежать в множестве $g_{\Delta\varepsilon}$, так как $\inf\limits_{\mathbf{x} \in g_{\Delta\varepsilon}} V = C_{\Delta\varepsilon} > \alpha_\Delta \geqslant \alpha$). Так как функции $V(\mathbf{x}, \varepsilon)$, $V(\mathbf{x})$ по условию являются непрерывными, то (по определению непрерывности этих функций в точке $\mathbf{x}_* = 0$) для любого $\alpha > 0$ существует такая окрестность S_α этой точки, что при $\mathbf{x} \in S_\alpha$ функция $V < \alpha$. Но это означает, что множество X_α содержит окрестность S_α и само является окрестностью точки $\mathbf{x}_* = 0$.

Исходя из изложенного и учитывая, что при $\varepsilon \to 0$ множество δ_ε сходится к точке $\mathbf{x}_* = 0$ и что множество η есть окрестность точки $\mathbf{x}_* = 0$ (точка $\mathbf{x}_* = 0$ входит в множество η с некоторым шаром), можно выбрать такие числа $\varepsilon(\eta) > 0$ и $\alpha(\eta) > 0$, чтобы множество $X_{\alpha(\eta)} = \{\mathbf{x}\colon V \leqslant \alpha(\eta)\} \subset \eta$. Множество $X_{\alpha(\eta)}$, являющееся окрестностью точки $\mathbf{x}_* = 0$, можно взять в качестве искомой окрестности $q \subset \eta$ этой точки. Действительно, любое решение $\mathbf{x}(t, \mathbf{x}_0)$, начинающееся при $t = t_0$ в множестве $X_{\alpha(\eta)}$, не может при $t > t_0$ его покинуть, ибо по условию теоремы $\dot{V} \leqslant 0$, и, следовательно, при $t > t_0$ значения функции не могут возрастать и поэтому для них имеем $V \leqslant \alpha(\eta)$. Следовательно, этим значениям соответствуют решения $\mathbf{x}(t, \mathbf{x}_0) \in X_{\alpha(\eta)}$. Заметим, что в случае, когда теореме удовлетворяет функция $V(\mathbf{x}, \varepsilon)$, явно зависящая от ε, в качестве множества $X_{\alpha(\eta)}$ можно взять множество $\delta_\varepsilon = \{\mathbf{x}\colon V(\mathbf{x}, \varepsilon) = 0\} \subset \eta$ — окрестность точки $\mathbf{x}_* = 0$; в случае функции $V(\mathbf{x})$ этого сделать нельзя, так как при этом множеством δ_ε является точка $\mathbf{x}_* = 0$, а не окрестность.

Так как $q = X_{\alpha(\eta)} \subset \eta$, то из изложенного следует, что при любом $\mathbf{x}_0 \in q = X_{\alpha(\eta)}$ решение $\mathbf{x}(t, \mathbf{x}_0) \subset \eta$. Таким образом, обычная устойчивость точки $\mathbf{x}_* = 0$ доказана.

Покажем, что для указанных решений $\mathbf{x}(t, \mathbf{x}_0) \subset \eta$ при условиях теоремы

$$\lim_{t \to \infty} \mathbf{x}(t, \mathbf{x}_0) = 0, \tag{2}$$

т. е. точка $\mathbf{x}_* = 0$ асимптотически устойчива. Условие (2) означает, что соответствующая решению $\mathbf{x}(t, \mathbf{x}_0)$ изображающая точка сходится к точке $\mathbf{x}_* = 0$ при $t \to 0$, т. е. в некоторый момент времени $t(q_1)$ попадает в любой шар $q_1 \subset \eta$ с центром в точке $\mathbf{x}_* = 0$ и при любом $t > t(q_1)$ остается в этом шаре. Покажем, что для любого выбранного шара q_1 существует такой момент времени $t(q_1)$. Для доказательства воспользуемся, в частности, тем, что точка $\mathbf{x}_* = 0$ обладает обычной устойчивостью. В соответствии с определением обычной устойчивости точки $\mathbf{x}_* = 0$ для любого шара q_1 существует такой открытый шар $s \subset q_1$, что если точка $\mathbf{x}(t, \mathbf{x}_0)$ попадает в некоторый момент времени $t(s)$ в шар s, то она при всех $t > t(s)$ будет оставаться в шаре q_1. Ясно, что если существует момент времени $t(s_1)$, то существует и момент времени $t(q_1)$; в частности за $t(q_1)$ можно принять $t(s)$. Остается доказать, что момент времени $t(s)$ существует, т. е. что при $t \to \infty$ $\mathbf{x}(t, \mathbf{x}_0)$ изменится от $\mathbf{x}(t_0)$ до $\mathbf{x}(t(s), \mathbf{x}_0) \in s$. Для этого выберем значение ε таким, чтобы соответствующее функции $V(\mathbf{x}, \varepsilon)$ множество δ_ε удовлетворяло условию $\delta_\varepsilon \subset s$. Но тогда имеем

$$\operatorname{supr}_{\mathbf{x} \in g_{s\eta}} \dot{V}(\mathbf{x}, \varepsilon) = c_\varepsilon < 0, \quad g_{s\eta} = \overline{\eta \backslash s}.$$

Действительно, непрерывная по условию отрицательная функция $\dot{V}(\mathbf{x}, \varepsilon)$ достигает на компактном множестве $g_{s\eta}$ своей точной верхней грани. Теперь предположим, что момент времени $t(s)$ не существует, т. е. точка $\mathbf{x}(t, \mathbf{x}_0)$ никогда не войдет в шар s, оставаясь все время в множестве $\overline{\eta \backslash s}$ (из множества η точка $\mathbf{x}(t, \mathbf{x}_0)$ выйти не может). Но это приводит к противоречию. Действительно, согласно формуле Лейбница—Ньютона, при сделанном предположении

$$V(\mathbf{x}(t, \mathbf{x}_0), \varepsilon) = V(\mathbf{x}_0, \varepsilon) + \int_{t_0}^{t} \dot{V}(\mathbf{x}(\tau, \mathbf{x}_0), \varepsilon) d\tau \leqslant V(\mathbf{x}_0, \varepsilon) + c_\varepsilon(t - t_0).$$

Так как $c_\varepsilon < 0$, то из приведенного соотношения видно, что при достаточно большом t значение функции $V(\mathbf{x}(t, \mathbf{x}_0), \varepsilon)$ становится любым отрицательным числом, чего не может быть, так как $V(\mathbf{x}, \varepsilon) \geqslant 0$ при $\mathbf{x} \in G_V$. Полученное противоречие доказывает, что точка $\mathbf{x}(t, \mathbf{x}_0)$ не может оставаться все время вне шара s, и, следовательно, существует момент времени $t(s)$, когда она в этот шар войдет. Но тогда, согласно приведенным выше рассуждениям, выполняется соотношение (2); т. е. точка $\mathbf{x}(t, \mathbf{x}_0)$ асимптотически устойчива. Достаточность условий теоремы доказана.

Сделаем необходимое для дальнейшего *замечание* о характере асимптотической устойчивости в случае автономных систем (1). Пусть $G_\text{п} \subset G$ — область притяжения точки $\mathbf{x}_* = 0$, т. е. множество всех таких точек $\mathbf{x}_0 \in G$, которым соответствуют решения $\mathbf{x}(t, \mathbf{x}_0)$) такие, что $\lim\limits_{t \to \infty} \mathbf{x}(t, \mathbf{x}_0) = 0$. Пусть множество $A \subset G_\text{п}$. Точка $\mathbf{x}_* = 0$ называется асимптотически устойчивой равномерно по t_0 и $\mathbf{x}_0 \in A$, если она асимптотически устойчива и для любой ее окрестности η можно указать такое число $T(\eta)$, что $\mathbf{x}(t, \mathbf{x}_0) \in \eta$ при $t \geqslant t_0 + T(\eta)$, каковы бы ни были t_0 и $\mathbf{x}_0 \in A$. Из соответствующей леммы в работе [1] следует, что в случае автономных систем (1) асимптотическая устойчивость точки $\mathbf{x}_* = 0$ обязательно равномерна по t_0 и $\mathbf{x}_0 \in B$, где B — любой компакт из $G_\text{п}$ (для неавтономных систем этот факт не имеет места; в этом случае для равномерности асимптотической устойчивости по t_0 и $\mathbf{x}_0$ система должна удовлетворять дополнительным требованиям).

Необходимость. Пусть точка равновесия $\mathbf{x}_* = 0$ автономной системы (1) асимптотически устойчива. Тогда она будет также асимптотически устойчивой равномерно по t_0 и $\mathbf{x}_0$ в указанном выше смысле. Известно [1, 2], что в этом случае существует функция Ляпунова $V(\mathbf{x})$, удовлетворяющая условиям теоремы Ляпунова об асимптотической устойчивости, а, следовательно, и условиям сформулированной выше теоремы. Таким образом, указанные теоремы являются обратимыми. Поэтому при доказательстве необходимости условий теоремы речь будет идти не об обратимости вообще, а об обратимости на основе применения специального класса функций $S(\mathbf{x}, \varepsilon)$, отличающихся от известных функций $V(\mathbf{x})$, обращающих теорему Ляпунова об асимптотической устойчивости, весьма простой геометрической трактовкой и, следовательно, наглядностью, в чем заключается их достоинство. Как будет показано далее, для очень широкого класса асимптотически устойчивых систем (1), не включающего в себя лишь достаточно экзотические системы, функции $S(\mathbf{x}, \varepsilon)$ могут не зависеть от параметра ε и являться функциями Ляпунова $V(\mathbf{x}) = S(\mathbf{x})$, обладая в то же время указанной геометрической наглядностью. Уже один этот факт показывает полезность новых необходимых и достаточных условий асимптотической устойчивости. Простота и наглядность геометрической интерпретации функций $S(\mathbf{x}, \varepsilon)$, $S(\mathbf{x})$ достигается ценой того, что в упомянутых выше экзотических случаях, которые будут описаны далее, приходится использовать зависящие от параметра функции $S(\mathbf{x}, \varepsilon)$. Именно этим объясняется, что теорема формулируется в общем виде в терминах функций $V(\mathbf{x}, \varepsilon)$, зависящих от параметра ε. Заметим, что при $\varepsilon > 0$ функции $V(\mathbf{x}, \varepsilon)$ и, в частности, функции $S(\mathbf{x}, \varepsilon)$, не являясь знакоопределенными, не могут удовлетворять условиям теоремы Ляпунова об асимптотической устойчивости, сформулированной в терминах функций $V(\mathbf{x})$. Поэтому появление нового класса функций $S(\mathbf{x}, \varepsilon)$ и $S(\mathbf{x})$ не может быть связано с теоремой Ляпунова. Для этого нужны новые необходимые и достаточные условия, которые и даны приведенной выше теоремой, сформулированной в терминах функций $V(\mathbf{x}, \varepsilon)$ и указывающей для каждой асимптотически устойчивой системы именно функции $S(\mathbf{x}, \varepsilon)$ или $S(\mathbf{x})$ с отмеченными выше свойствами.

Приступим непосредственно к построению для асимптотически устойчивых точек равновесия систем вида (1) функций $S(\mathbf{x}, \varepsilon)$ и $S(\mathbf{x})$, удовлетворяющих теореме. На любом компакте $B \subset G_{\text{п}}$ ($G_{\text{п}}$ — область притяжения), являющемся окрестностью точки $\mathbf{x}_* = 0$, для любой асимптотически устойчивой системы построим сначала функцию $S(\mathbf{x}, \varepsilon)$, явно зависящую от параметра ε. Сначала построим семейство $\{\delta_\varepsilon\}$ множеств $\delta_\varepsilon \subset B$, на которых $S(\mathbf{x}, \varepsilon) = 0$. Каждое из этих множеств должно по условию быть замкнутой окрестностью точки $\mathbf{x}_* = 0$, а при $\varepsilon \to 0$ множества семейства $\{\delta_\varepsilon\}$ должны сходиться к точке $\mathbf{x}_* = 0$. Кроме того, для целей построения непрерывной по $\mathbf{x}$ функции $S(\mathbf{x}, \varepsilon)$ мы построим множества δ_ε такими, чтобы они имели гладкую границу $\Gamma(\delta_\varepsilon)$, а решение $\mathbf{x}(t, \mathbf{x}_0)$, соответствующее любой точке $\mathbf{x}_0 \in \delta_\varepsilon$, сходясь при $t > t_0$ к точке $\mathbf{x}_* = 0$, не покидало бы δ_ε (т. е. множество δ_ε состоит из положительных полутраекторий, целиком лежащих в δ_ε, и является поэтому инвариантным множеством), а все траектории, входящие в δ_ε, пересекали бы границу $\Gamma(\delta_\varepsilon)$ трансверсально (т. е. не касаясь этой границы). Такие множества δ_ε можно построить, например, следующим образом. На основе известных результатов об обращении теоремы Ляпунова асимптотическая устойчивость точки равновесия автономной системы (1) является равномерной по t_0 и $\mathbf{x}_0 \in B \subset G_{\text{п}}$, и поэтому на любом компакте $B \subset G_{\text{п}}$ существует функция Ляпунова $V(\mathbf{x})$, удовлетворяющая этой теореме. Рассмотрим все множества $\delta_\varepsilon = \{\mathbf{x}\colon V(\mathbf{x}) \leqslant \varepsilon\} \subset B$, где ε — положительные числа. В силу свойств функций $V(\mathbf{x})$ (знакоопределенность, непрерывность) множества δ_ε при $\varepsilon \to 0$ сходятся к точке $\mathbf{x}_* = 0$ ($\lim\limits_{\varepsilon \to \infty} \dim\delta_\varepsilon = 0$; $\dim\delta_\varepsilon = \max\limits_{\mathbf{x}_i, \mathbf{x}_j \in \delta_\varepsilon} \rho(\mathbf{x}_i, \mathbf{x}_j)$). Каждое из множеств δ_ε является замкнутой окрестностью точки $\mathbf{x}_* = 0$ (это следует из непрерывности функции $V(\mathbf{x})$). Гладкость границы $\Gamma(\delta_\varepsilon)$ множества δ_ε следует из теоремы о неявной функции [3], примененной к уравнению $F(\mathbf{x}) = V(\mathbf{x}) - \varepsilon = 0$, неявно задающему $(n - 1)$-мерную гиперповерхность $\Gamma(\delta_\varepsilon)$. Действительно, функция $F(\mathbf{x})$ удовлетворяет всем условиям этой теоремы (она непрерывна вместе со своими частными производными $\partial F(\mathbf{x})/\partial x_i$, $i = 1, ..., n$ и т. д.). Поэтому задаваемая этой функцией неявно гиперповерхность $\Gamma(\delta_\varepsilon)$ является гладкой в некоторой окрестности любой точки $\mathbf{x} \in \Gamma(\delta_\varepsilon)$. Гладкость границы $\Gamma(\delta_\varepsilon)$ нужна для того, чтобы имело смысл понятие трансверсальности траекторий к этой границе. Сама же трансверсальность следует из определенной отрицательности производной в силу системы от функции Ляпунова ($\dot V(\mathbf{x}) < 0$ при любом $\mathbf{x} \in B$, кроме точки $\mathbf{x}_* = 0$; очевидно, что $\operatorname{supr}\limits_{\mathbf{x} \in b_\varepsilon} \dot V(\mathbf{x}) = c < 0$, где $b_\varepsilon = \overline{B_\varepsilon \backslash \delta_\varepsilon}$). Действительно, если $\dot V(\mathbf{x}) = (\operatorname{grad} V(\mathbf{x}), \mathbf{f}(\mathbf{x})) < 0$ при $\mathbf{x} \in \Gamma(\delta_\varepsilon)$, то это означает, что траектория (следовательно, вектор $f(\mathbf{x})$) не может быть касательной к гиперповерхности $\Gamma(\delta_\varepsilon)$, т. е. ортогональной к градиенту $V(\mathbf{x})$, который, очевидно, ортогонален к ней.

Функцию $S(\mathbf{x}, \varepsilon)$, заданную на любом компакте $B \subset G_{\text{п}}$, определим при каждом значении $\varepsilon > 0$ и соответствующем ему множестве δ_ε следующим образом. Любой точке $\mathbf{x} \in B$ поставим в соответствие расстояние $S(\mathbf{x}, \varepsilon)$ от этой точки $\mathbf{x}$ до множества δ_ε, измеренное вдоль траектории $T_{\mathbf{x}}$, проходящей через точку $\mathbf{x}$, и определяемое на всем множестве B как

$$S(\mathbf{x}, \varepsilon) = \min_{\mathbf{x}_1 \in \delta_\varepsilon \cap T_{\mathbf{x}}} \rho(\mathbf{x}, \mathbf{x}_1), \qquad (3)$$

где $\rho(\mathbf{x}, \mathbf{x}_1)$ — расстояние от данной точки $\mathbf{x} \in B$ до любой точки $\mathbf{x}_1 \in \delta_\varepsilon \cap T_{\mathbf{x}}$, измеренное вдоль траектории $T_{\mathbf{x}}$. В соответствии с этим определением $S(\mathbf{x}, \varepsilon) = 0$, если $\mathbf{x} \in \delta_\varepsilon$, и $S(\mathbf{x}, \varepsilon) > 0$, если $\mathbf{x} \in B\backslash\delta_\varepsilon$. Функция $S(\mathbf{x}, \varepsilon)$ при $\varepsilon > 0$ может быть выражена для $\mathbf{x} \in B\backslash\delta_\varepsilon$ следующим образом (далее решение $\mathbf{x}(t, \mathbf{x}_0)$ ради удобства будем обозначать $\tilde{\mathbf{x}}(t, \mathbf{x})$):

$$S(\mathbf{x}, \varepsilon) = \int\limits_0^{\tau_\varepsilon(\mathbf{x})} \dot S_+(\tilde{\mathbf{x}}(t, \mathbf{x}))dt, \qquad (4)$$

где $\tau_\varepsilon(\mathbf{x})$ — время, за которое изображающая точка, соответствующая решению $\tilde{\mathbf{x}}(t, \mathbf{x})$, начиная движение при $t = 0$ из точки $\mathbf{x} \in B$, достигает границы $\Gamma(\delta_\varepsilon)$ множества δ_ε (если $\mathbf{x} \notin \delta_\varepsilon$, то $\tau_\varepsilon(\mathbf{x}) > 0$); $\dot S_+(\tilde{\mathbf{x}}(t, \mathbf{x}))$ — скорость увеличения длины отрицательной полутраектории в точке $\tilde{\mathbf{x}}(t, \mathbf{x})$.

Производная $\dot S$ определяется выражением $\dot S_+(\tilde{\mathbf{x}}(t, \mathbf{x})) = |\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x}))|$. Его легко получить, учитывая, что $dS_+ = \sqrt{\sum\limits_{i=1}^{n} dx_i^2}$, $dx_i = f_i(\tilde{\mathbf{x}})dt$, $\sum\limits_{i=1}^{n} f_i^2 = |f|^2$. Производную $\dot S_+(\tilde{\mathbf{x}}(t, \mathbf{x}))$ нужно отличать от производной $\dot S(\tilde{\mathbf{x}}(t, \mathbf{x})) = -\dot S_+(\tilde{\mathbf{x}}(t, \mathbf{x})) = -|\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x}))|$, которая является скоростью уменьшения длины той части положительной полутраектории, которая лежит между точкой $\tilde{\mathbf{x}}(t, \mathbf{x})$ и границей $\Gamma(\delta_\varepsilon)$ множества δ_ε; эта длина задается формулой (4). Иначе говоря, $\dot S_+(\tilde{\mathbf{x}}(t, \mathbf{x}))$ является производной в силу системы от функции $S(\mathbf{x}, \varepsilon)$, задаваемой выражением (4). Возвращаясь к прежнему обозначению, можно записать

$$\dot S(\mathbf{x}, \varepsilon) = -|\mathbf{f}(\mathbf{x})|. \qquad (5)$$

Отметим, что в выражении (4) интегрирование ведется вдоль траектории решения $\tilde{\mathbf{x}}(t, \mathbf{x})$. Функция $S(\mathbf{x}, \varepsilon)$ при любом $\varepsilon > 0$ ограничена на множестве B. Это следует из оценки величины $\operatorname{supr}\limits_{\mathbf{x} \in B\backslash\delta_\varepsilon} S(\mathbf{x}, \varepsilon)$ на основе того, что точка $\mathbf{x}_* = 0$ асимптотически устойчива равномерно по t_0 и $\mathbf{x} \in B$. Действительно, $\operatorname{supr}\limits_{\mathbf{x} \in B\backslash\delta_\varepsilon} S(\mathbf{x}, \varepsilon) \leqslant T(B) \operatorname{supr}\limits_{\mathbf{x} \in B\backslash\delta_\varepsilon} |\mathbf{f}(\mathbf{x})| = c > 0$, ибо конечны оба сомножителя ($T(B)$ — время, за которое в силу равномерности по t_0, $\mathbf{x} \in B$ все точки $\mathbf{x} \in B\backslash\delta_\varepsilon$ перейдут в δ_ε при $t \to \infty$).

Для асимптотически устойчивых точек равновесия большого класса систем (1), удовлетворяющих определенным (описанным далее) требованиям, на любом компакте $B \subset G_{\text{п}}$ можно построить более простую (не зависящую от параметра) функцию $S(\mathbf{x}) = \rho(\mathbf{x}, \mathbf{x}_* = 0)$, которая получается из функции (3) путем предельного перехода:

$$S(\mathbf{x}) = \rho(\mathbf{x}, \mathbf{x}_* = 0) = \lim_{\varepsilon \to 0} \min_{\mathbf{x}_1 \in \delta_\varepsilon \cap T_\mathbf{x}} \rho(\mathbf{x}, \mathbf{x}_1) = \lim_{\varepsilon \to \infty} S(\mathbf{x}, \varepsilon), \quad (6)$$

где $\rho(\mathbf{x}, \mathbf{x}_* = 0)$ — расстояние от точки $\mathbf{x}$ до точки $\mathbf{x}_* = 0$ измеренное вдоль траектории T_x (т. е. длина положительной полутраектории). Этот предел всегда существует, ибо всегда существует предел положительной монотонной функции, которой является длина $S(\mathbf{x}, \varepsilon)$ участка траектории; он может быть как конечным, так и бесконечным. Из выражения (6) следует, что функция $S(\mathbf{x})$ является положительно определенной функцией ($S(\mathbf{x}_* = 0) = 0$, $S(\mathbf{x}) > 0$ при $\mathbf{x} \neq \mathbf{x}_* = 0$). Учитывая формулу (4), получаем для любого $\mathbf{x} \in B$

$$S(\mathbf{x}) = \int_0^\infty |\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x})|dt = \lim_{\varepsilon \to \infty} \int_0^{\tau_\varepsilon(\mathbf{x})} |\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x}))|dt. \quad (7)$$

Этот несобственный интеграл сходится (конечен) не всегда. У некоторых систем (1) правая часть $f(\mathbf{x})$ и соответствующий ей характер поведения траекторий в окрестности асимптотически устойчивой точки $\mathbf{x}_* = 0$ могут обусловливать расходимость этого интеграла. Для прояснения этого вопроса, запишем значение функции $S(\mathbf{x})$ в любой фиксированной точке $\mathbf{x} \in B$ иначе, а именно:

$$S(\mathbf{x}) = \int_0^\infty |\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x}))|dt = \int_0^{\tau_\varepsilon(\mathbf{x})} |\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x}))|dt + \int_{\tau_\varepsilon(\mathbf{x})}^\infty |\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x}))|dt =$$

$$= S(\mathbf{x}, \varepsilon) + \int_{\tau_\varepsilon(\mathbf{x})}^\infty |\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x}))|dt. \quad (8)$$

При этом значение ε (соответственно множество δ_ε) выбираем так, чтобы $\mathbf{x} \in B \backslash \delta_\varepsilon$. Как показано выше, функция $S(\mathbf{x}, \varepsilon)$ при любом $\varepsilon > 0$ ограничена на любом компакте $B \subset G_п$. Поэтому первый член $S(\mathbf{x}, \varepsilon)$ в правой части приведенного выражения конечен. Число ε можно взять любым (лишь бы $\mathbf{x} \in B \backslash \delta_\varepsilon$). Устремим его к нулю. От этого значение интеграла (7) не меняется (он либо конечен, либо расходится). Поэтому, учитывая конечность функции $S(\mathbf{x}, \varepsilon)$, интеграл (7) будет конечным (сходится), если при любом $\varepsilon > 0$ конечен интеграл, являющийся вторым членом в правой части выражения (8). Если же он расходится, то расходится и интеграл (7). Этот второй член является длиной положительной полутраектории, начинающейся на границе $\Gamma(\delta_\varepsilon)$ множества δ_ε, т. е. расстоянием $\rho(\mathbf{x} \in \Gamma(\delta_\varepsilon), \mathbf{x}_* = 0)$, измеренным вдоль траектории. Поэтому интеграл (7) сходится (конечен) и соответственно функция $S(\mathbf{x})$ ограничена на множестве B и удовлетворяет в этом отношении требованиям к классу функций $V(\mathbf{x}, \varepsilon)$ и $V(\mathbf{x})$, если выполняется условие, что при любом $\varepsilon > 0$

$$\operatorname{supr}_{\mathbf{x} \in \Gamma(\delta_\varepsilon)} \rho(\mathbf{x} \in \Gamma(\delta_\varepsilon), \mathbf{x}_* = 0) = c_\varepsilon < \infty$$

(очевидно, $c_\varepsilon > 0$). Из этого условия следует, что при любом $\varepsilon > 0$ $\operatorname{supr}_{\mathbf{x} \in \delta_\varepsilon} S(\mathbf{x}) = c_\varepsilon < \infty$.

Эти условия налагают требование на характер поведения траекторий в сколь угодно малой окрестности точки $\mathbf{x}_* = 0$, обеспечивающий, очевидно, сходимость интеграла (7) и ограниченность функции $S(\mathbf{x})$ на множестве B. Для некоторых асимптотически устойчивых точек равновесия систем вида (1), которые мы назвали ранее экзотическими, эти условия не выполняются (например, для всех нелинейных фокусов, задаваемых на фазовой плоскости в полярных координатах системой уравнений $\dot{\rho} = -\rho^n$, $\dot{\theta} = 1$, $n \geqslant 2$), соответственно тогда интеграл (7) расходится и функция $S(\mathbf{x})$ не ограничена на множестве B, ибо она при любом $\mathbf{x}$ (кроме точки $\mathbf{x}_* = 0$) принимает бесконечное значение (в точке $\mathbf{x}_* = 0$ она равна нулю). Работать с такими функциями в обычном смысле невозможно. В то же время функции $S(\mathbf{x}, \varepsilon)$, явно зависящие от параметра, и для этих экзотических случаев при $\varepsilon > 0$ являются ограниченными на множестве $B \subset G_п$. Именно из-за этих случаев, желая указать для каждой асимптотически устойчивой точки равновесия геометрически просто интерпретируемую функцию (3), обеспечивающую обращение теоремы об асимптотической устойчивости и имеющую конечные на множестве B значения, приходится использовать функции $S(\mathbf{x}, \varepsilon)$, зависящие от параметра и обнуляемые на множествах δ_ε, сходящихся к точке $\mathbf{x}_* = 0$. Соответственно поэтому приходится работать на классе функций $V(\mathbf{x}, \varepsilon)$, зависящих от параметра, и в этих терминах формулировать теорему.

Если для любого $\varepsilon > 0$ $\operatorname{supr}_{\mathbf{x} \in \delta_\varepsilon} S(\mathbf{x}, \varepsilon) = c_\varepsilon < \infty$ и, кроме того, при $\varepsilon \to 0$ (следовательно, $\dim\delta_\varepsilon \to 0$)

$$\lim_{\varepsilon \to \infty} \operatorname{supr}_{\mathbf{x} \in \delta_\varepsilon} S(\mathbf{x}, \varepsilon) = \lim_{\varepsilon \to \infty} c_\varepsilon = 0, \quad (9)$$

то функция $S(\mathbf{x})$ непрерывна в точке $\mathbf{x}_* = 0$, ибо условие (9) эквивалентно определению такой непрерывности ($\forall c_\varepsilon > 0\ \exists\ \delta_\varepsilon : \mathbf{x} \in \delta_\varepsilon \Rightarrow S(\mathbf{x}) < c_\varepsilon$) Далее будет показано, при каких условиях функция $S(\mathbf{x})$ непрерывна в каждой точке $\mathbf{x} \in B$. Заметим, что из выражения (6) следует, что всегда $S(\mathbf{x}_* = 0) = 0$, и, таким образом, функция $S(\mathbf{x})$ является определенно-положительной. Заметим также, что нарушение условия (9) означает, очевидно, разрывность функции $S(\mathbf{x})$ в точке $\mathbf{x}_* = 0$, ибо это условие является необходимым и достаточным условием непрерывности (как определение непрерывности).

Покажем теперь, что построенные функции $S(\mathbf{x}, \varepsilon)$ и $S(\mathbf{x})$ принадлежат классу функций $V(\mathbf{x}, \varepsilon)$, а также удовлетворяют условиям теоремы. Покажем сначала, что эти функции непрерывны на множестве $B \subset G_п$. Непрерывность функции $S(\mathbf{x}, \varepsilon)$ на $B \backslash \delta_\varepsilon$ следует из непрерывности по параметру $\mathbf{x}$ интеграла (4). Действительно, согласно теореме из курса математического анализа [3], обобщенной на векторный случай, интеграл (4) будет непрерывно зависеть от векторного параметра $\mathbf{x}$, если функция $|\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x}))|$ непрерывна по совокупности переменных $(t, \mathbf{x})$, а функция $\tau_\varepsilon(\mathbf{x})$ непрерывна по $\mathbf{x}$. Первое имеет место, так как правая часть системы (1) $\mathbf{f}(\tilde{\mathbf{x}})$ и, следовательно, $|\mathbf{f}(\tilde{\mathbf{x}})|$ непрерывно зависят от $\tilde{\mathbf{x}}$ по условию, а $\tilde{\mathbf{x}}$ непрерывно по совокупности переменных t, $\mathbf{x}$ согласно теореме о непрерывности решения по начальным условиям, примененной к расширенной системе $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, $\dot{\theta} = 1$. Непрерывность $\tau_\varepsilon(\mathbf{x})$ по $\mathbf{x}$ доказывается с помощью теоремы о неявной функции, которую мы применим к уже использовавшемуся ранее уравнению $F(\mathbf{x}) = V(\mathbf{x}) - \varepsilon = 0$, неявно задающему границу $\Gamma(\delta_\varepsilon)$ множества δ_ε. Вектор (точка) $\mathbf{x}_\Gamma \in \Gamma(\delta_\varepsilon)$, удовлетворяющий этому уравнению, задает функциональную связь $\mathbf{x}_\Gamma = \varphi(\mathbf{x}, \tau_\varepsilon(\mathbf{x}))$ между точкой $\mathbf{x}$, лежащей на траектории, проходящей через точку $\mathbf{x}_\Gamma$, и временем пробега $\tau_\varepsilon(\mathbf{x})$ изображающей точки $\tilde{\mathbf{x}}(t, \mathbf{x})$ от точки $\mathbf{x}$ до точки $\mathbf{x}_\Gamma$. В итоге получаем уравнение $F(\varphi(\mathbf{x}, \tau_\varepsilon(\mathbf{x}))) = 0$, неявно задающее зависимость $\tau_\varepsilon(\mathbf{x})$. По теореме о неявной функции эта зависимость будет непрерывной, если $\partial F(\mathbf{x}, \tau_\varepsilon)/\partial\tau_\varepsilon = \partial V(\mathbf{x}, \tau_\varepsilon)/\partial\tau_\varepsilon \neq 0$.

Это условие выполняется, так как $\partial V(\mathbf{x}, \tau_\varepsilon)/\partial\tau_\varepsilon = \dot{V} < 0$ (производная $\dot{V}$ в силу системы по условию теоремы отрицательна на множестве B/δ_ε). В итоге функция $S(\mathbf{x}, \varepsilon)$, задаваемая интегралом (4), непрерывна по $\mathbf{x}$ на множестве $B \backslash \delta_\varepsilon$. Непрерывность функции $S(\mathbf{x}, \varepsilon)$ по $\mathbf{x}$ на границе $\Gamma(\delta_\varepsilon)$ следует из того, что, согласно формуле (3), $S(\mathbf{x}, \varepsilon) \equiv 0$ на множестве δ_ε, а при $\mathbf{x} \in B\backslash\delta_\varepsilon$ непрерывная функция $S(\mathbf{x}, \varepsilon)$ задается выражением (4), из которого видно, что по мере уменьшения до нуля числа Δ, задающего окрестность $\delta_{\Delta\varepsilon}$ множества δ_ε и, следовательно, уменьшения до нуля величины $\sup\limits_{\mathbf{x} \in \delta_{\Delta\varepsilon}\backslash\delta_\varepsilon} \tau_\varepsilon(\mathbf{x})$, до нуля уменьшается и величина $\sup\limits_{\mathbf{x} \in \delta_{\Delta\varepsilon}\backslash\delta_\varepsilon} S(\mathbf{x}, \varepsilon)$. При этом имеет значение, что траектории трансверсальны к границе $\Gamma(\delta_\varepsilon)$. Таким образом, функция $S(\mathbf{x}, \varepsilon)$ непрерывна на границе $\Gamma(\delta_\varepsilon)$. Учитывая ее непрерывность на множествах $\mathrm{Int}\delta_\varepsilon$ и $B\backslash\delta_\varepsilon$, можно сделать вывод, что она непрерывна по $\mathbf{x}$ на множестве B.

Докажем теперь, что ограниченная на множестве B функция $S(\mathbf{x})$, задаваемая сходящимся интегралом (7), непрерывна по $\mathbf{x}$ на множестве B. Так как функция $S(\mathbf{x}, \varepsilon)$ непрерывна на множестве $B\backslash\delta_\varepsilon$ и считается, что справедливо соотношение (9), то выполняется определение непрерывности функции $S(\mathbf{x})$ в любой точке $\mathbf{x} \in B\backslash(\mathbf{x}_* = 0)$ (непрерывность в точке $\mathbf{x}_* = 0$ доказана ранее и следует из выражения (9)): для любого числа $\eta > 0$ существует такая окрестность δ точки $\mathbf{x}$, что при любом $\mathbf{x}_1 \in \delta$ $S(\mathbf{x}_1) - S(\mathbf{x}) < \eta$. Действительно, в соответствии с (9) можно выбрать такую малую δ_ε — окрестность точки $\mathbf{x}_* = 0$, чтобы $\sup\limits_{\mathbf{x} \in \delta_\varepsilon} S(\mathbf{x}) < \eta/2$. Кроме того, в силу непрерывности функции $S(\mathbf{x}, \varepsilon)$ можно выбрать такую окрестность δ точки $\mathbf{x}$, чтобы $S(\mathbf{x}_1, \varepsilon) - S(\mathbf{x}, \varepsilon) < \eta/2$ при любом $\mathbf{x}_1 \in \delta$. Тогда выполняется определение непрерывности функции $S(\mathbf{x})$ на множестве $B\backslash(\mathbf{x}_* = 0)$.

Непрерывность на множестве $B\backslash\delta_\varepsilon$ производных $\dot{S}(\mathbf{x}, \varepsilon)$ следует из выражения (5), так как по условию функция $\mathbf{f}(\mathbf{x})$ непрерывна по $\mathbf{x}$ на множестве $B\backslash\delta_\varepsilon$. На границе $\Gamma(\delta_\varepsilon)$ производная $\dot{S}(\mathbf{x}, \varepsilon)$ терпит разрыв, так как $\dot{S} \equiv 0$ на множестве $\mathrm{Int}\delta_\varepsilon$. Производная же $\dot{S}(\mathbf{x})$ непрерывна всюду на множестве B.

Перейдем к рассмотрению вопроса о непрерывной дифференцируемости (гладкости) по $\mathbf{x}$ ограниченных и непрерывных на множестве $B\backslash\delta_\varepsilon$ функций $S(\mathbf{x}, \varepsilon)$ и $S(\mathbf{x})$. Функция $S(\mathbf{x}, \varepsilon)$, согласно формуле (4), является собственным интегралом, зависящим от $\mathbf{x}$ как от параметра. На основании теоремы о дифференцируемости такого интеграла по параметру [3, 4] функция $S(\mathbf{x}, \varepsilon)$ является гладкой по $\mathbf{x}$ на множестве $B\backslash\delta_\varepsilon$ (т. е. ее частные производные $\partial S/\partial x_i$ непрерывны по x; x_i — компоненты вектора $\mathbf{x}$), если на множестве $B\backslash\delta_\varepsilon$ непрерывно дифференцируема по $\mathbf{x}$ скалярная функция $\tau_\varepsilon(\mathbf{x})$, а также непрерывна по $(t, \mathbf{x})$ векторная функция $\mathrm{grad}_x|\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x}))|$ (т. е. непрерывны по $(t, \mathbf{x})$ все частные производные $\partial|\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x})|/\partial x_i$, $i = 1, ..., n$). Выполнение первого условия следует из выполнения условий теоремы о неявной функции, примененной к уже рассматривавшемуся ранее уравнению $F(\mathbf{x}, \tau_\varepsilon(\mathbf{x})) = 0$. Выполнение второго условия вытекает из гладкости функции $\mathbf{f}(\tilde{\mathbf{x}})$ и теоремы о гладкой зависимости решения $\tilde{\mathbf{x}}$ от начальных условий $\mathbf{x}$. Таким образом, построенная функция $S(\mathbf{x}, \varepsilon)$ является гладкой по $\mathbf{x}$ на множестве $B\backslash\delta_\varepsilon$. Она не гладкая на границе $\Gamma(\delta_\varepsilon)$, ибо на множестве $\mathrm{Int}\delta_\varepsilon$ все производные $\partial S/\partial x_i \equiv 0$, а на множестве $\overline{B\backslash\delta_\varepsilon}$ функция $|\dot{S}(\mathbf{x}, \varepsilon)| = |(\mathrm{grad} S(\mathbf{x}, \varepsilon), \mathbf{f}(\mathbf{x}))| = |\mathbf{f}(\mathbf{x})|$ достигает нижней грани $\inf|\dot{S}(\mathbf{x}, \varepsilon)| = \inf|\mathbf{f}(\mathbf{x}))| = c > 0$, вследствие чего на множестве $\overline{B\backslash\delta_\varepsilon}$ $\inf|\mathrm{grad}S| \neq 0$, а поэтому хотя бы для одной производной $\partial S/\partial x_i$ на множестве $\overline{B\backslash\delta_\varepsilon}$ имеем $\inf\left|\dfrac{\partial S}{\partial x_i}\right| \neq 0$.

Согласно выражению (7) функция $S(\mathbf{x})$ является несобственным интегралом, зависящим от параметра $\mathbf{x}$. Согласно теореме о дифференцируемости этого интеграла по параметру [3, 4], дающей достаточные условия, ограниченная и непрерывная на множестве B функция $S(\mathbf{x})$ является гладкой по $\mathbf{x}$ на множестве $B\backslash(\mathbf{x}_* = 0)$, если: 1) непрерывны на множестве $B\backslash(\mathbf{x}_* = 0)$ по совокупности переменных $(t, \mathbf{x})$ все частные производные $\partial|\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x}))|/\partial x_i$; 2) на множестве $B\backslash(\mathbf{x}_* = 0)$ существуют и сходятся равномерно по $\mathbf{x}$ все интегралы $\int\limits_0^\infty [\partial|\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x}))|/\partial x_i]dt$, $i = 1, ..., n$. Первое условие в силу гладкости функции $\mathbf{f}(\tilde{\mathbf{x}})$ выполняется. Вопрос о том, каким условиям должна удовлетворять правая часть системы (1), чтобы выполнялось второе условие, рассматривать здесь не будем (остается открытым вопрос о гладкости в общем случае непрерывной функции $S(\mathbf{x})$). Заметим только, что в случае асимптотически устойчивых линейных систем с постоянными коэффициентами все данные выше условия выполняются, и функции $S(\mathbf{x}, \varepsilon)$ и $S(\mathbf{x})$ являются гладкими вне множеств δ_ε (в точке $\mathbf{x}_* = 0$ функция $S(\mathbf{x})$ может быть негладкой). Выполнение этих условий на множестве $B\backslash(\mathbf{x}_* = 0)$ в линейном случае связано с тем, что как решение $\tilde{\mathbf{x}}(t, \mathbf{x})$, так и линейная по $\tilde{\mathbf{x}}$ функция $\mathbf{f}(\tilde{\mathbf{x}}(t, \mathbf{x}))$ мажорируются (покомпонентно) некоторой экспонентой e^{-at}, $a > 0$, в силу чего выполняются условие (9), обеспечивающее непрерывность функции $S(\mathbf{x})$ по $\mathbf{x}$, и все условия, обеспечивающие ее гладкость на множестве $B\backslash(\mathbf{x}_* = 0)$. Именно из этого вытекает справедливость следствия, сформулированного ранее.

Необходимость условий теоремы доказана.

5. О НАХОЖДЕНИИ ФУНКЦИЙ $S(\mathbf{x}, \varepsilon)$ И $S(\mathbf{x})$

Кратко рассмотрим вопрос нахождения функций $S(\mathbf{x}, \varepsilon)$ и $S(\mathbf{x})$. Пусть точка равновесия $\mathbf{x}_* = 0$ системы (1) исследуется на асимптотическую устойчивость. Если эта точка действительно асимптотически устойчива, хотя мы этого и не знаем, то соответствующая ей функция $S(\mathbf{x}, \varepsilon)$, как видно из формулы (4), при заданных на $f(\mathbf{x})$ условиях всегда существует; функция же $S(\mathbf{x})$, как мы знаем, существует при этом не всегда. Если эти функции являются на множестве $B\backslash\delta_\varepsilon$ гладкими ($B \subset G_п$ — компакт), то они могут быть определены на множестве $A\backslash\delta_\varepsilon$, где $A \subset B$ — некоторая компактная окрестность точки $\mathbf{x}_* = 0$, с помощью уравнения

$$(\mathrm{grad}\,V, \mathbf{f}(\mathbf{x})) = -|\mathbf{f}(\mathbf{x})|, \qquad (10)$$

являющегося линейным уравнением в частных производных относительно функции V. Здесь в левой части стоит производная в силу системы (1) от функции V, а в правой — функция от $\mathbf{x}$, которая является производной в силу системы (1) от искомых функций $S(\mathbf{x}, \varepsilon)$ или $S(\mathbf{x})$, если исследуемая система (1) является асимптотически устойчивой. Если при любом $\dim\delta_\varepsilon$ решение V этого уравнения существует (оно может существовать и в случае, если точка $\mathbf{x}_* = 0$ не является асимптотически устойчивой; при этом V не будет δ_ε — определенно-положительна) и является δ_ε — определенно-положительной функцией, то оно является функцией $S(\mathbf{x}, \varepsilon)$ или $S(\mathbf{x})$, а точка $\mathbf{x}_* = 0$ асимптотически устойчива. Наоборот, если известно, что точка $\mathbf{x}_* = 0$ асимптотически устойчива, то решение $V = S(\mathbf{x}, \varepsilon)$ уравнения (10) существует при любом $\dim\delta_\varepsilon$ и является δ_ε — определенно-положительной функцией.

При решении уравнения (10) на границе $\Gamma(\delta_\varepsilon)$ множества δ_ε должны быть заданы граничные условия: $V(\mathbf{x}, \varepsilon) = 0$ или $V(\mathbf{x}) = 0$ при $\mathbf{x} \in \Gamma(\delta_\varepsilon)$. Сами же множества δ_ε должны подбираться так, чтобы траектории системы (1) были трансверсальны к границе $\Gamma(\delta_\varepsilon)$. В случае достаточно простых систем (1) решение уравнения (10) может быть получено аналитически, в более сложных случаях — численными методами. Приведем для иллюстрации применение описанного способа для аналитического нахождения функций $S(\mathbf{x}, \varepsilon)$, $S(\mathbf{x})$ для простого нелинейного скалярного уравнения $\dot{x} = -x^n$, где n — любое нечетное целое положительное число. Здесь $f(x) = -x^n$, $|f(x) = |x^n|$, $\operatorname{grad} V = dV/dx$. В соответствии с уравнением (10) функция V вне множества (δ_ε) описывается уравнением $(dV/dx)(-x^n) = -|x^n|$ или $dV/dx = 1$ при положительных значениях x, лежащих вне множества δ_ε, и $dV/dx = -1$ при отрицательных значениях x, лежащих вне множества δ_ε. Если ищем функцию $V(x)$, то граничным является условие: $V(x) = 0$ при $x_* = 0$. В итоге, интегрируя полученные выражения для dV/dx в соответствующих пределах, учитывающих граничные условия, получаем функцию $V(x) = |x|$. Так как она определенно-положительна, то является функцией $S(x) = |x|$. Если же ищем функцию $V(x, \varepsilon)$, зависящую от параметра ε, то в качестве множеств δ_ε выбираем отрезки $\delta_\varepsilon = [-0{,}5\varepsilon, +0{,}5\varepsilon]$, а граничные условия имеют вид: $V(x, \varepsilon) = 0$ при $x = -0{,}5\varepsilon$ и $x = 0{,}5\varepsilon$. На отрезках δ_ε функция $V(x, \varepsilon) \equiv 0$ по определению функций $V(x, \varepsilon)$. После интегрирования выражений для производной dV/dx в соответствующих пределах получаем, что вне отрезков δ_ε функция $V(x, \varepsilon) = |x| -0{,}5\varepsilon$, так как $dV/dx = 1$ при $x > 0{,}5\varepsilon$ и $dV/dx = -1$ при $x < -0{,}5\varepsilon$. Так как полученная функция δ_ε — определенно-положительна, то она является функцией $S(\mathbf{x}, \varepsilon) = |x| - 0{,}5\varepsilon$.

Заметим, что вообще для любых нелинейных асимптотически устойчивых систем первого порядка $\dot{x} = f(x)$ уравнение (10) имеет вид $(dS/dx)f(x) = -|f(x)|$, что дает (с учетом знака $f(x)$ при $x > 0$ и $x < 0$): $S(x) = |x|$ при любом x и $S(x, \varepsilon) = |x| - 0{,}5\varepsilon$ при $|x| > 0{,}5\varepsilon$. Функция $S(\mathbf{x})$ является негладкой в точке $x_* = 0$.

Ранее иным способом были получены функции $S(\rho, \varepsilon) = \sqrt{2}(\rho - \varepsilon)$ и $S(\rho) = \sqrt{2}\rho$ для асимптотически устойчивой точки равновесия системы $\dot{\rho} = -\rho^n$, $\dot{\theta} = 1$. Этот результат может быть получен также как пример применения уравнения (10) для получения этих функций.

ЛИТЕРАТУРА

1. *Красовский Н. Н.* Некоторые задачи теории устойчивости движения. — М.: Физматгиз, 1959.
2. *Малкин Н. Г.* Теория устойчивости движения: — М.: Наука, 1966.
3. *Фихтенгольц Г. М.* Основы математического анализа. Т. 2. — М.: Наука, 1968.
4. *Шилов Г. Е.* Математический анализ. Функции нескольких вещественных переменных. — М.: Наука, 1972.

☎ *(095) 334-92-29*

❑

УДК 513.88

ИТЕРАЦИОННЫЙ АЛГОРИТМ ПОСТРОЕНИЯ ЦИКЛОВ НЕЛИНЕЙНЫХ АВТОНОМНЫХ СИСТЕМ
Ч. 1. Сходимость

И. Г. Исмаилов

Институт проблем управления им. В. А. Трапезникова, г. Москва

Предложен новый алгоритм приближенного построения цикла автономной нелинейной системы обыкновенных дифференциальных уравнений. Алгоритм обладает локальной сходимостью и эффективен в случае неустойчивых циклов.

ВВЕДЕНИЕ

Будем рассматривать задачу приближенного построения цикла автономной системы обыкновенных дифференциальных уравнений. Если некоторый цикл Γ орбитально асимптотически устойчив, то задача его приближенного построения, как правило, не вызывает затруднений: если начальное условие x_0 решения $p(t, x_0)$ системы достаточно близко к циклу Γ, то $\lim\limits_{t \to \infty} \inf\limits_{y \in \Gamma} |p(t, x_0) - y| = 0$. Таким образом, если отыскиваемый орбитально асимптотически устойчивый цикл достаточно хорошо локализован, то применением какого-либо метода численного интегрирования можно получить сколь угодно точное приближение к циклу. Однако во многих важных ситуациях (например, в задачах хаотической динамики) отыскиваемые циклы, как правило, неустойчивы. Задача приближенного построения таких циклов становится существенно сложнее. Один из приемов приближенного построения неустойчивых циклов автономных систем базируется на комбинации метода функционализации параметра [1] и какого-либо дискретизационного метода (метода механических квадратур, метода коллокации и др.) (см., например, работы [2, 3]).

Однако такой подход весьма трудоемок в вычислительном отношении и требует дополнительной информации об отличии от нуля топологического индекса отыскиваемого цикла. Поэтому представляет интерес разработка и исследование простых в применении итерационных алгоритмов приближенного построения неустойчивых циклов. Один из таких алгоритмов принадлежит Спарроу (C. Sparrow) [4]. Этот алгоритм базируется на методе Ньютона, поэтому реализация алгоритма Спарроу на каждом шаге требует обращения специальных матриц. Это обстоятельство существенно уменьшает диапазон применимости алгоритма и делает его неэффективным в вырожденных и плохо обусловленных задачах. В данной работе предлагается новый итерационный алгоритм приближенного построения циклов автономных систем, эффективный для неустойчивых периодических решений. Этот результат был опубликован в тезисах докладов [5—7].

1. ОСНОВНАЯ ТЕОРЕМА

Будем рассматривать систему обыкновенных дифференциальных уравнений

$$\frac{dx}{dt} = f(x), \quad x \in \mathbb{R}^N. \tag{1}$$

Пусть $\mathbb{R}$ — вещественная ось, $\mathbb{R}^N$ — евклидово пространство со скалярным произведением $\langle \cdot, \cdot \rangle$. Если $x \in \mathbb{R}^N$, $t \in \mathbb{R}$, то через $\{x, t\}$ обозначается пара в $\mathbb{R}^{N+1}$. Ниже через T обозначается операция транспонирования.

Предположим, что правая часть системы (1) непрерывно дифференцируема. Обозначим через $p(t, x)$ решение этой системы с начальным условием $p(0, x) = x$.

Зафиксируем некоторый вектор $a \in \mathbb{R}^N$, число $b \in \mathbb{R}$ и рассмотрим в пространстве пар $\{x, t\}$ систему уравнений

$$x = p(t, x), \tag{2}$$

$$\langle a, x \rangle = b. \tag{3}$$

Пусть пара $\{x^*, t^*\}$, $t^* \neq 0$ является решением системы (2), (3). Тогда x^* — это точка некоторого цикла системы (1), а t^* — его период. Следовательно, задача отыскания циклов системы (1) эквивалентна отысканию решений системы (2), (3).

Рассмотрим следующую итерационную процедуру

$$x_{k+1} = x_k - \gamma_k((I - V^T(t_k, x_k))(x_k - p(t_k, x_k)) + (\langle a, x_k \rangle - b)a), \tag{4}$$

$$t_{k+1} = t_k + \mu_k \langle f(p(t_k, x_k)), x_k - p(t_k, x_k) \rangle, \quad k = 0, 1, 2, \ldots, \tag{5}$$

где T — символ транспонирования, $V(x, t)$ ($V(0, x) = I$) — фундаментальная матрица линейной системы дифференциальных уравнений $\frac{dV}{dt} = f_x'(p(t, x))V$ а γ_k, μ_k — управляющие параметры итерационной процедуры (4), (5).

Предположим, что Г — изолированный цикл системы (1) с периодом T^*, $x^* \in \Gamma$ и

$$\langle a, x^*\rangle = b. \qquad (6)$$

Рассмотрим гиперплоскость π, определяемую уравнением (3). Если она трансверсальна циклу Г в точке x^*, то пара $\{x^*, T^*\}$ является изолированным решением системы (2), (3). Справедлива следующая

Теорема. *Пусть пара* $\{x^*, T^*\}$ — *изолированное решение системы уравнений*

$$(I - V^T(t, x))(x - p(t, x)) + (\langle a, x\rangle - b)a = 0 \qquad (7)$$

$$\langle f(p(t, x)), x - p(t, x)\rangle = 0, \qquad (8)$$

а управляющие параметры γ_k, μ_k *итерационной процедуры* (4), (5) *удовлетворяют неравенствам*

$$0 < \alpha_0 \leqslant \gamma_k \leqslant \beta_0 \leqslant 1, \qquad (9)$$

$$0 < \alpha_1 \leqslant \mu_k \leqslant \beta_1 \leqslant 1, \qquad (10)$$

где числа β_0 *и* β_1*достаточно малы. Пусть начальное приближение* $\{x_{0,}\ t_0\}$ *процедуры* (4), (5) *достаточно близко к паре* $\{x^*, T^*\}$. *Тогда последовательные приближения* $\{x_{k,}\ t_k\}$ *сходятся к паре* $\{x^*, t^*\}$:

$$\lim_{t\to\infty} (\|x_k - x^*\| + |t_k - T^*|) = 0. \qquad (11)$$

Доказательствотеоремы см. в Приложении.

2. ПРИМЕР ПРИМЕНЕНИЯ ТЕОРЕМЫ

В качестве приложения сформулированной теоремы рассмотрим многоконтурную автономную систему автоматического регулирования. Ее динамика описывается системой обыкновенных дифференциальных уравнений

$$\begin{cases} L_1\left(\frac{d}{dt}\right)x_1 = M_1\left(\frac{d}{dt}\right)f_1(x_1, ..., x_n) \\ \quad\quad\quad ... \\ L_n\left(\frac{d}{dt}\right)x_n = M_n\left(\frac{d}{dt}\right)f_n(x_1, ..., x_n). \end{cases} \qquad (12)$$

Здесь $L_i(p)$ и $M_i(p)$ многочлены:

$$L_i(p) = p^{l_i} + a_1^i p^{l_i - 1} + ... + a_{l_i}^i, \qquad (13)$$

$$M_i(p) = b_0^i p^m + b_1^i p^{m-1} + ... + b_m^i, \quad i = 1, ..., n, \qquad (14)$$

$p = \frac{d}{dt}$ — оператор дифференцирования по переменной t. Предполагается, что порядки многочленов (13) и (14) подчинены условию $l_i \geqslant m_i + 1$, $i = 1, ..., n$. Формально правые части дифференциальных уравнений определены лишь для достаточно гладких нелинейностей. Для определения решений системы (12) в случае произвольных непрерывных нелинейностей $f_i(x_1, ..., x_n)$, $i = 1, ..., n$ стандартным образом переходят к пространству состояний. В этом пространстве система (12) принимает вид:

$$\frac{dx}{dt} = f(x), \quad x \in \mathbb{R}^N, \qquad (15)$$

где функция $f(x)$ выражается через коэффициенты многочленов и функции f_i и при этом имеет такую же гладкость, как исходные нелинейности. Если функции f_i дифференцируемы m_i раз, то любое решение системы (15) будет решением системы (12). Таким образом, периодические колебания в многоконтурной автономной системе автоматического регулирования можно искать с помощью алгоритма (4), (5).

ЗАКЛЮЧЕНИЕ

Предположение о трансверсальности плоскости π искомому циклу Г не является ограничительным. Изменяя параметры a и b этого всегда можно добиться при условии некоторой начальной локализации. Равенства (7) и (8) фактически означают изолированность критической точки $\{x^*, T^*\}$ функции невязки $W(x, t)$, при этом отсутствует требование невырожденности минимума этой функции. Заметим, что задача поиска цикла сложнее в случае автономной системы, чем в случае T-периодической, где априори известен период. Предложенная схема дает как начальное приближение, так и значение периода искомого периодического решения системы (1). Схема может быть реализована с помощью любого профессионального пакета вычислительных алгоритмов, например Matlab или Mathematic.

ПРИЛОЖЕНИЕ

Доказательство теоремы. Рассмотрим в окрестности точки $\{x^*, t^*\} \in \mathbb{R}^{N+1}$ функцию

$$W(x, t) = 0{,}5(\|x - p(t, x)\|^2 + (\langle a, x\rangle + b)^2). \qquad (16)$$

Обозначим через ∇ оператор градиента по переменным $\{x, t\}$, через ∇_x — оператор градиента по x, а через ∇_t — производную по t. Непосредственный подсчет показывает, что

$$\nabla W(x, t) = \{I - V^T(t, x))(x - p(t, x)) + \\ + (\langle a, x\rangle - b)a, \langle -f(p(t, x)), x - p(t, x)\rangle|,$$

$$\nabla_x W(x, t) = (I - V^T(t, x))(x - p(t, x)) + (\langle a, x\rangle - b)a,$$

$$\nabla_t W(x, t) = -\langle f(p(t, x)), x - p(t, x)\rangle.$$

В силу выражения (6) точка $\{x^*, t^*\}$ является точкой минимума функции $W(x, t)$. А так как пара $\{x^*, T^*\}$ суть изолированное решение системы (7), (8), то она является изолированной критической точкой функции $W(x, t)$. Следовательно, существует шар $B \subset \mathbb{R}^{N+1}$ с центром $\{x^*, T^*\}$, в котором единственной критической точкой функции (16) и точкой абсолютного минимума будет $\{x^*, T^*\}$.

Положим $\varepsilon = \min\limits_{\{x,\, t\} \in \partial B} W(x, t)$, где ∂B — граница шара. Выберем положительное $\delta < \varepsilon$ таким образом, чтобы

$$\max_{\{x,\, t\} \in L(\delta)} \|\nabla W(x, t)\| < \min\{\|\{x, t\} - \{y, \tau\}\| : \{x, t\} \in L(\delta), \\ \{y, \tau\} \in \partial B \tag{17}$$

где $L(\delta)$— лебегово множество функции $W(x, t)$, т. е. $L(\delta) = \{\{x, t\} \in B : W(x, t) \leqslant \delta\}$.

Покажем теперь, что для любого начального приближения $\{x_{0,}\ t_0\} \in L(\delta)$последовательные приближения $\{x_{k,}\ t_k\}$ процедуры (4), (5) сходятся к паре $\{x^*, T^*\}$. Для этого покажем вначале, что из включения $\{x_{k,}\ t_k\} \in L(\delta)$ вытекают соотношения

$$\{x_{k+1,}\ t_{k+1}\} \in B, \tag{18}$$

$$W(x_{k+1}, t_{k+1}) - W(x_k, t_k) \leqslant -\nu\|\nabla W(x_k, t_k)\|^2, \tag{19}$$

где ν — достаточно малое положительное число.

В силу равенств (4), (5) и оценок (9), (10) имеем:

$$\|x_{k+1} - x_k\| \leqslant \|\nabla_x W(x_k, t_k)\|, \tag{20}$$

$$|t_{k+1} - t_k| \leqslant |\nabla_t W(x_k, t_k)|. \tag{21}$$

Из неравенств (17), (20) и (21) следует включение (18). Пусть числа K_1, K_2, L_1 и L_2 таковы, что для любых $\{x_1, t_1\}, \{x_2, t_2\} \in B$ выполнены неравенства

$$\|\nabla_x W(x_1, t_1) - \nabla_x W(x_2, t_2)\| \leqslant K_1\|x_1 - x_2\| + K_2|t_1 - t_2|,$$

$$\|\nabla_t W(x_1, t_1) - \nabla_t W(x_2, t_2)\| \leqslant L_1\|x_1 - x_2\| + L_2|t_1 - t_2|,$$

а константы β_0 и β_1 в неравенствах (9) и (10) таковы, что

$$\frac{1}{2}K_1\beta_0 + \frac{1}{4}(K_2 + L_1)\beta_1 < 1, \tag{22}$$

$$\frac{1}{2}L_2\beta_1 + \frac{1}{4}(K_2 + L_1)\beta_0 < 1. \tag{23}$$

Тогда

$$W(x_{k+1}, t_{k+1}) - W(x_k, t_k) = \int_0^1 \langle\nabla_x W(x_k + \tau(x_{k+1} - \\ - x_k), t_k + \tau(t_{k+1} - t_k)), x_{k+1} - x_k\rangle d\tau + \int_0^1 \nabla_t W(x_k + \\ + \tau(x_{k+1} - x_k), t_k + \tau(t_{k+1} - t_k)), x_{k+1} - t_k) d\tau = \\ = -\gamma_k\int_0^1 \langle\nabla_x W(x_k - \tau\gamma_k\nabla_x W(x_k, t_k), t_k - \tau\mu_k W(x_k, t_k)), \\ \nabla_x W(x_k, t_k)\rangle d\tau - \mu_k\int_0^1 \nabla_t W(x_k - \tau\gamma_k\nabla_x W(x_k, t_k), t_k - \\ - \tau\mu_k W(x_k, t_k)), \nabla_t W(x_k, t_k) d\tau = -\gamma_k\|\nabla_x W(x_k, t_k)\|^2 - \\ - \mu_k(\nabla_t W(x_k, t_k))^2 - \gamma_k\int_0^1 \langle\nabla_x W(x_k - \tau\gamma_k W(x_k, t_k), t_k - \\ - \tau\mu_k\nabla_t W(x_k, t_k)) - \nabla_x W(x_k, t_k), \nabla_x W(x_k, t_k)\rangle d\tau - \\ - \mu_k\int_0^1 \nabla_x W(x_k - \tau\gamma_k W(x_k, t_k), t_k - \tau\mu_k\nabla_t W(x_k, t_k)) - \\ - \nabla_t W(x_k, t_k) \cdot \nabla_t W(x_k, t_k) d\tau \leqslant \\ \leqslant -\gamma_k\|\nabla_x W(x_k, t_k)\|^2 - \mu_k(\nabla_t W(x_k, t_k))^2 + \\ + \frac{1}{2}\gamma_k(K_1\gamma_k\|\nabla_x W(x_k, t_k)\| + K_2\mu_2|\nabla_t W(x_k, t_k)|)\|\nabla_x W(x_k, t_k)\| + \\ + \frac{1}{2}\mu_k(L_1\gamma_k\|\nabla_x W(x_k, t_k)\| + L_2\mu_k|\nabla_t W(x_k, t_k)|)|\nabla_t W(x_k, t_k)\| \leqslant \\ \leqslant -\gamma_k(1 - \frac{1}{2}K_1\gamma_k - \frac{1}{4}(K_2 + L_1)\mu_k\|\nabla_x W(x_k, t_k)\|^2 - \\ - \mu_k(1 - \frac{1}{2}L_2\mu_k - \frac{1}{4}(K_2 + L_1)\gamma_k|\nabla_t W(x_k, t_k)\|^2 \leqslant \\ \leqslant -\alpha_0(1 - \frac{1}{2}K_1\beta_0 - \frac{1}{4}(K_2 + L_1)\beta_1\|\nabla_x W(x_k, t_k)\|^2 - \\ - \alpha_1(1 - \frac{1}{2}L_2\beta_1 - \frac{1}{4}(K_2 + L_1)\beta_0|\nabla_t W(x_k, t_k)|^2.$$

Из последней оценки и неравенств (22) и (23) вытекает неравенство (19), в котором

$$\nu = \min\{\alpha_0(1 - \frac{1}{2}K_1\beta_0 - \frac{1}{4}(K_2 + L_1)\beta_1, \\ \alpha_1(1 - \frac{1}{2}L_2\beta_1 - \frac{1}{4}(K_2 + L_1)\beta_0)\}.$$

Суммируя неравенства (19) по k от 0 до m, получаем

$$W(x_{m+1}, t_{m+1}) - W(x_0, t_0) \leqslant -\nu\sum_{\kappa=0}^{m}\|\nabla W(x_k, t_k)\|^2.$$

А так как справедливы включения (18), то из последнего неравенства следует, что ряд $\sum\limits_{\kappa=0}^{m}\|\nabla W(x_k, t_k)\|^2$ сходится. Но тогда $\lim\limits_{t\to\infty}\|\nabla W(x_k, t_k)\| = 0$. А так как $\{x^*, T^*\}$ — единственная критическая точка функции $W(x, t)$ на шаре B, то из последнего равенства следует сходимость (11). Теорема доказана.

ЛИТЕРАТУРА

1. *Бобылев Н. А., Красносельский М. А.* Функционализация параметра и теорема родственности для автономных систем // Дифференциальные уравнения. — 1970. — № 11. — С. 1946—1952.
2. *Бобылев Н. А.* К теории фактор-методов приближенного решения нелинейных задач // Доклады АН СССР. — 1972. — Т. 199, № 1. — С. 9—12.
3. *Бобылев Н. А.* Метод механических квадратур в задаче о периодических решениях // Успехи математических наук. — 1972. — Т. XXVII, вып. 4 (166). — С. 203—204.
4. *Sparrow C.* The Lorenz Equations: Bifurcations, Chaos and Strange Attractors. — New York: Springer, 1982.
5. *Бобылев Н. А., Исмаилов И. Г., Коровин С. К.* Об одном алгоритме построения предельных циклов в системах автоматического регулирования // IV междунар. семинар «Устойчивость и колебания нелинейных систем управления». — М., 1996. — С. 6.
6. *Ismailov I. G.* On the scheme of approximate construction of cycles of nonlinear systems // Fourth intern. conf. on «Control, automation, robotics and vision». — Singapore, 1996.
7. *Ismailov I. G.* On the approximate construction of cycles in automatic control systems // Fourth intern. symp. on «Method and Models in Automation and Robotics». — Poland, 1997.

☎ *(095) 334-79-00*

E-mail: Ilkham@ukr.net ❑

УДК 681.3.057.51-7.311.17

АНАЛИЗ ПОДХОДОВ К РЕШЕНИЮ ПРОБЛЕМЫ ПРАВИЛЬНОСТИ МАТЕМАТИЧЕСКИХ ЗНАНИЙ[1]

Т. Л. Гаврилова, А. С. Клещёв

Институт автоматики и процессов управления, г. Владивосток

Рассмотрены различные подходы к решению проблемы правильности математических знаний, предлагаемые математической практикой, математической и компьютерной логикой. Обсуждены критерии правильности математических знаний: универсальный, интуитивный, логический, формально-логический и компьютерный. Показано, что компьютерный критерий потенциально наиболее надёжен для обеспечения правильности математических знаний, а системы автоматизированного (человеко-машинного) доказательства теорем — наиболее перспективный путь его применения. Определены направления дальнейшего продвижения в решении указанной проблемы.

Доказательство является лишь временной помощью для ленивого разума. Жаждущие доказательства не способны воспринять доказываемую истину, независимо от того, насколько хорошо она доказана.

Лобсанг Рампа. Пещеры древних

ВВЕДЕНИЕ

Одна из важнейших проблем математики состоит в обеспечении правильности математических знаний. Принято считать, что математические знания более правильные, чем знания в естественных науках. Однако если сравнить требования к правильности математических знаний с современными требованиями к правильности, например, компьютерных программ, то положение в математике в этом отношении можно признать близким к катастрофическому. Представление о правильности математических знаний менялось с течением времени. Цель настоящей работы — предложить такой анализ современного состояния этой проблемы, из результатов которого были бы видны пути дальнейшего продвижения в ее решении. В качестве материала для анализа взяты в основном общеизвестные факты. При этом рассматриваются только основные идеи и опускаются все технические подробности.

[1] Работа выполнена в рамках Программы фундаментальных исследований № 16 Президиума РАН "Математическое моделирование и интеллектуальные системы" по проекту "Теоретические основы интеллектуальных систем, основанных на онтологиях, для интеллектуальной поддержки научных исследований" (№ 1002-251/П-16/021-387/110504-266) и Программы фундаментальных исследований № 16 ОЭММПУ РАН "Проблемы анализа и синтеза интегрированных технических и социальных систем управления" по проекту "Синтез интеллектуальных систем управления базами знаний и базами данных" (№ 1002-251/ОЭММПУ-16/080-387/190504-287).

1. ФОРМУЛИРОВКА ПРОБЛЕМЫ

Под математическими знаниями в данной работе понимается совокупность математических теорем. Д. Пойа так сформулировал основную проблему правильности математических знаний — «От математика требуется ответить на вопрос: данная теорема верна или неверна?» [1].

Из этой формулировки вытекают вопросы, по которым следует прийти к соглашению до того, как решать основную проблему:

— к какому миру относится теорема;

— какую форму имеет теорема;

— что значит, что теорема верна;

— как можно установить, что теорема верна?

К какому миру относится теорема? Теорема является утверждением, а всякое утверждение относится к некоторому миру. Будем называть мир, к которому относятся математические теоремы (о котором они нечто утверждают), математическим миром. Математический мир является коллекцией совокупностей математических объектов, причем объекты, входящие в одну и ту же совокупность, обладают общими для этой совокупности свойствами.

Поскольку математический мир является миром идей и существует только в сознании людей, имеющих математическое образование, он должен быть описан некоторым образом для того, чтобы все эти люди понимали математический мир одинаково. Описание математического мира состоит из описания основных неопределяемых понятий, формулировок аксиом об этих понятиях и определений производных понятий.

Каждому основному понятию в математическом мире соответствует некоторое значение — либо множество математических объектов, либо один объект; термин, обозначающий это понятие, является обозначением этого значения. Если значение основного понятия — множество объектов, то считается, что все эти объекты обладают свойствами, задаваемыми аксиомами, относящимися к этому основному понятию. Каждое определение вводит в математический мир производное понятие, значением которого может быть либо множество математических объектов, либо новый математический объект (функция, предикат и т. п.), и термин для обозначения этого значения. Если значение производного понятия — множество объектов, то оно является либо подмножеством уже существующего в математическом мире множества объектов (супермножества), либо строится из объектов уже существующих в мире множеств. В первом случае объекты подмножества наследуют все свойства объектов супермножества; кроме того, определение вводит дополнительные свойства, которыми обладают объекты подмножества, в отличие от остальных объектов супермножества. Во втором случае свойства элементов нового множества определяются способом построения этих элементов и свойствами объектов, из которых эти элементы строятся. Описание математического мира задает этот математический мир однозначно.

Законы математического мира называются теоремами. Формулировки теорем — это лишь интуитивные догадки математиков о свойствах математического мира, скрытых от их взоров. Поэтому теоремы не добавляют математическому миру новых свойств: они лишь явно (!) утверждают, что математический мир обладает теми свойствами, которыми его неявно (!) наделили аксиомы и определения. Теоремы образуют математические знания.

Какую форму имеет теорема? Описание математического мира (аксиомы и определения) и математические знания (теоремы) записываются на «математическом диалекте», который позволяет описывать неопределяемые понятия, формулировать утверждения (аксиомы и теоремы), определять производные понятия, а также вводить формальные способы записи. Математический диалект выполняет функции и языка, и метаязыка. Набор его неформальных оборотов не фиксирован, а наборы терминов и формальных способов записи постоянно расширяются средствами языка.

Что значит, что теорема верна? Теорема (догадка) является правильной (верной), если, обозрев весь математический мир, наблюдатель обнаруживает, что все совокупности объектов, о которых идет речь в теореме, — ее частные случаи — обладают теми свойствами, наличие которых у этих совокупностей утверждает теорема. Это определение «правильности» не специфично для математики, а является общенаучным. Математические знания правильны тогда и только тогда, когда правильны все входящие в них теоремы.

Как можно установить, что теорема правильна? Из ответа на предыдущий вопрос непосредственно вытекает «идеальный» критерий правильности математических знаний: чтобы установить, что теорема правильна, нужно установить, что правильны все частные случаи этой теоремы, причём то, что частный случай правилен, должно быть установлено непосредственно. Поэтому проблема установления правильности теоремы сводится к проблеме установления правильности всех ее частных случаев.

Если теорема имеет лишь конечное число частных случаев, то для установления правильности этой теоремы идеальный критерий может быть применён практически. Такие теоремы будем называть тривиальными. Для нетривиальных теорем идеальный критерий не позволяет установить их правильность. Однако именно нетривиальные теоремы представляют интерес для математики. Поэтому и возникает потребность в иных критериях правильности математических знаний.

2. РЕШЕНИЕ ПРОБЛЕМЫ В РАМКАХ МАТЕМАТИЧЕСКОЙ ПРАКТИКИ

На основе скудных исторических данных можно предположить, что до Евклида правильность математических знаний устанавливалась на основе «интуитивного» критерия: чтобы установить правильность нетривиальной теоремы, нужно понять, верна она или неверна. Если вернуться к эпиграфу к этой статье, то можно сказать, что «интуитивный критерий» специально предназначен для неленивого разума, поскольку требует восприятия истины, которая утверждается в теореме. Однако в наше время разум ленив, а применение этого критерия требует определённых умственных способностей и образования. Поэтому такой способ не рассматривается как приемлемый.

Евклид пользовался аристотелевой логикой для установления правильности математических знаний [2]. Тем самым в математическую практику был введен «логический» критерий правильности математических знаний: чтобы установить, что нетривиальная теорема правильна, нужно построить ее правильное интуитивное доказательство и установить правильность аксиом, на которых оно основано. Интуитивное доказательство считается правильным, если его правильность установлена на основе интуитивного критерия, т. е. чтобы установить правильность доказательства, нужно, как и в случае теорем, понять, правильно или неправильно это доказательство. Доказательства, правильность которых устанавливается на основе интуитивного критерия, получили название интуитивных. Что касается аксиом, то для классических аксиоматических систем применяется интуитивный критерий, а для современных — конвенциональный (аксиомы считаются правильными в силу соглашения между членами математического сообщества). Для ленивого разума логический критерий считается более приемлемым. В нем требование понимания правильности короткой формулировки теоремы заменяется требованием понимания правильности куда более длинного текста ее доказательства. Таким образом, логический критерий свел вопрос о правильности теорем к пониманию правильности аксиом и доказательств.

Интуитивное доказательство включает в себя основную и побочные линии рассуждения, прямые и обратные рассуждения, явные и повторно используемые, а также неявные модули, формальные преобразования, логические, специальные и содержательные рассуждения, выдвижение предположений и условные рассуждения, явные и неявные пропуски. Основная линия рассуждения ведет к доказываемой теореме, а побочные линии ее поддерживают. Прямые рассуждения ведут от уже доказанных утверждений, аксиом и определений к еще не доказанным утверждениям, а обратные рассуждения — наоборот («для доказательства данного утверждения достаточно доказать...»). Леммы могут рассматриваться как явно выделенные модули доказательства, причем некоторые из них используются повторно. Неявные модули вводятся оборотами типа «покажем, что...». Формальные преобразования выполняются над утверждениями, записанными формально (например, тождественные преобразования). Примером логического рассуждения служит замена доказательства теоремы о равносильности доказательством двух теорем, имеющих форму импликации; примером специального рассуждения — использование принципа полной математической индукции; примером содержательного — различные способы применения математических теорем. Предположения выдвигаются, например, при доказательстве теоремы, имеющей форму импликации («предположим, что верно условие; докажем, что в этом предположении имеет место заключение»); при этом используются условные рассуждения (в число доказанных утверждений включаются предположения). Явные пропуски указываются оборотами типа «легко видеть...», а неявные никак не указываются. Текст интуитивного доказательства представляет собой последовательность специальных предложений математического диалекта, которые формируются на основе правил рассуждения.

Умение доказывать теоремы вырабатывается на примерах (задачах на доказательство), поэтому список правил рассуждения, допустимых в математике, не фиксирован. При этом если правильность формальных преобразований может быть установлена с помощью компьютерной программы, то для других типов рассуждений это не так. Неправильность интуитивного доказательства может быть установлена либо полуформальными (обнаружение применения недопустимых правил рассуждения или неправильного применения допустимых), либо содержательными (построение опровергающего примера) методами.

Некоторые логики признают лишь полные интуитивные доказательства: «то, что предназначается быть доказательством, не может иметь никаких пробелов, никаких лазеек, никаких сомнений, иначе это не доказательство» [1]. Однако чем менее тривиальна теорема, тем, как правило, длиннее ее полное интуитивное доказательство и тем менее оно понятно. Но и чрезмерное сокращение интуитивного доказательства может негативно отразиться на его понятности, так как требует интеллектуальных усилий, направленных на восстановление пропусков в процессе его понимания.

Можно отметить целый ряд противоречий математической практики, связанных с применением логического критерия правильности математических знаний.

Стремление получать все более нетривиальные математические результаты ведет, как правило, к удлинению интуитивных доказательств. В математической практике интуитивные доказательства всегда строятся вручную. Поскольку форма интуитивных доказательств исключает возможность применения компьютеров для проверки их правильности, то единственным критерием правильности интуитивного доказательства, признаваемым в математической практике, служит мнение экспертов. При этом, однако, не существует общепризнанной процедуры экспертного подтверждения правильности опубликованных интуитивных доказательств; неизвестно, анализировали ли эксперты и, если да, то какие именно эксперты, правильность того или иного опубликованного доказательства. Лишь в случае, если кто-либо из читателей случайно или намеренно обнаружил в таком доказательстве ошибку или построил для доказанной теоремы опровергающий пример, этот факт делается достоянием математической общественности.

Из практики программирования известно, что любые тексты, создаваемые людьми, могут содержать ошибки, причем число таких ошибок пропорционально длине текстов, а коэффициент пропорциональности зависит от семантической сложности текста: чем сложнее текст, тем больше ошибок он содержит. Несомненно, что интуитивное доказательство является одним из наиболее сложных текстов с большим числом разнообразных внутренних связей между его частями. Поэтому чем длиннее доказательство, тем больше потенциальных ошибок оно содержит и тем менее оно понятно.

Стремление «сократить» длинное доказательство за счет опускания «очевидных» деталей может привести к тому, что «сокращённое» доказательство будет содержать ещё больше потенциальных ошибок и (или) станет ещё менее понятным. Пропуски в интуитивных доказательствах, указанные явно или неявно, являются неистощимым источником прямых ошибок и неточностей [3].

История математики знает огромное число случаев обнаружения ошибок в интуитивных доказательствах, которые, к сожалению, часто воспринимались математиками как исторические курьезы. Можно ожидать, что число необнаруженных ошибок в таких доказательствах значительно превосходит число обнаруженных. Поэтому, несмотря на усилия математиков, направленные на обеспечение правильности математических доказательств, и в настоящее время математика еще далека от того, чтобы быть «образцом достоверности и истинности» [4]. Она нуждается в таком критерии правильности математических знаний, применение которого позволило бы еще более сузить область применения интуитивного критерия.

3. РЕШЕНИЕ ПРОБЛЕМЫ В РАМКАХ МАТЕМАТИЧЕСКОЙ ЛОГИКИ

В конце XIX в. были обнаружены парадоксы, связанные с началом изучения современных аксиоматических систем. Возникший кризис доверия к математическим знаниям заставил математиков начать дискуссию об уточнении «правил игры» в математике и привел к тому, что, начиная с этого времени, математика стала заботиться не только о своих математических «основаниях», но и о своих логических основаниях [5].

Одним из направлений «борьбы с парадоксами» стал анализ тех средств, которые традиционно применялись для построения описания математического мира и математических доказательств (теория типов Рассела, программы конструктивной и финитной математики Брауэра и Гильберта).

Другим направлением «борьбы с парадоксами» было построение математических моделей математической практики и изучение их математическими методами в рамках возникшей тогда области математики — математической логики. В наиболее известной из таких моделей — исчислении предикатов первого порядка — моделями математического мира являются алгебраические системы, моделью математического диалекта — язык исчисления предикатов первого порядка, моделями описания математического мира — логические теории (множества нелогических аксиом), моделью правил рассуждения — исчисление предикатов первого порядка, моделью полных интуитивных доказательств — формальные доказательства.

Тем самым в рассмотрение был введен «формально-логический» критерий правильности математических знаний: чтобы установить, что нетривиальная теорема правильна, нужно построить для ее модели на языке исчисления предикатов первого порядка формальное доказательство, являющееся моделью ее полного интуитивного доказательства, и установить правильность нелогических аксиом, на моделях которых оно основано. Нелогические аксиомы считаются правильными, если их достоверность установлена на основе интуитивного или конвенционального критерия. Сравнительно с логическим критерием в этом критерии требование понимания правильности интуитивного доказательства заменяется требованием проверки соответствия конкретного формального доказательства (конечного математического объекта) определению понятия «формальное доказательство». Такая проверка может быть выполнена с помощью компьютерной программы. Для ленивого разума этот критерий наиболее приемлемый. Теорема Геделя о полноте исчисления предикатов первого порядка установила формальную адекватность модели объекту моделирования.

Обсудим теперь, исходя из существенных свойств математической практики, рассмотренных в § 2, вопрос о том, насколько соответствует этой практике исчисление предикатов первого порядка (далее — исчисление предикатов).

В языке исчисления предикатов зафиксированы следующие первичные понятия: множество (задается неявно), объект (элемент множества), функция (число аргументов которой задается неявно с помощью функциональных термов), предикат (число аргументов которого задается неявно с помощью атомных формул), имя (которое вводится неявно и связывается с объектом), пропозициональные связки и логические кванторы, переменная (вводимая с помощью кванторов), аппликация функции и аппликация предиката. Эти понятия разъясняются при описании языка, и с их использованием определяется его семантика. Среди этих первичных понятий нет даже понятия числа. Никакие другие первичные понятия не могут быть введены при описании тех или иных теорий. Понятно, что для математической практики этих основных понятий недостаточно.

Нелогические аксиомы являются предложениями языка исчисления предикатов. Они определяют свойства множеств и всех их элементов, а также всех объектов, функций и предикатов, упомянутых в теории с помощью имен. Исчисление предикатов не накладывает никаких ограничений (кроме синтаксических) на вид или смысл нелогических аксиом.

Для записи определений в языке исчисления предикатов нет специальной конструкции. Поэтому при записи теории все имена, упомянутые в ней, вводятся неявно, а определения синтаксически ничем не отличаются от аксиом и представляются с помощью искусственных приемов.

Язык исчисления предикатов является формальным языком, предназначенным для записи теорий (нелогических аксиом) и теорем. Он содержит фиксированный набор формальных аналогов для основных (увы, далеко не всех!) оборотов математического диалекта. Но он не позволяет включать в предложения математические выражения в той нотации, как это принято в математической практике. Считается очевидным, что любое математическое утверждение может быть адекватно переведено на язык исчисления предикатов и обратно, однако такой перевод может потребовать значительных интеллектуальных усилий.

Моделью математического мира, описанного на языке исчисления предикатов, является алгебраическая система (непустое множество элементарных, не имеющих внутренней структуры объектов, — область интерпретации, на которой заданы конечные совокупности функций и предикатов), для которой определяется интерпретация множества имен, входящих в теорию [3] (далее — интерпретация), состоящая из трех отображений. Первое из них сопоставляет каждому имени объекта некоторый элемент области интерпретации; второе сопоставляет каждому имени функции некоторую функцию на области интерпретации; третье каждому имени предиката сопоставляет некоторый предикат на области интерпретации. Каждое предложение языка исчисления предикатов есть некоторое утверждение об интерпретации входящих в него имен. Семантика языка определяет способ вычисления значения любого предложения относительно этой интерпретации.

Описание математического мира на языке исчисления предикатов обладает следующими свойствами. Одно и то же множество имен, входящих в теорию, может иметь несколько (как правило, бесконечно много) разных интерпретаций. Эти интерпретации могут отличаться как областями интерпретации, так и любым из трех вышеуказанных отображений. Все переменные, входящие в теорию, в каждой интерпретации имеют одну и ту же область значений — область этой интерпретации. Теория выделяет те интерпретации (модели теории), относительно которых все аксиомы этой теории являются истинными. Таким образом, теория описывает некоторое (как правило, бесконечное) множество своих моделей, т. е. единственность математического мира теряется. Эти особенности языка исчисления предикатов затрудняют понимание теорий, записанных на нем, а также перевод на этот язык математических утверждений.

Предложение языка исчисления предикатов называется теоремой теории, если оно не совпадает ни с одной из нелогических аксиом и истинно во всех моделях этой теории. Таким образом, в исчислении предикатов модель основной проблемы правильности математических знаний формулируется следующим образом: дана теория и предложение на языке исчисления предикатов; требуется установить, является это предложение теоремой этой теории или нет. Но в математической практике математический мир единственен, и поэтому важно установить, верна ли теорема относительно именно этого мира. Отсюда следует, что модель основной проблемы правильности математических знаний в исчислении предикатов не адекватна формулировке этой проблемы для математической практики.

Прямой способ установить, что предложение является теоремой теории (на основе семантики языка исчисления предикатов) — перебрать все возможные области интерпретации, для каждой из них просмотреть все модели теории с этой областью интерпретации, рассмотреть все частные случаи предложения (заменив для этого в предложении каждую переменную ее значением из области интерпретации) и вычислить значение каждого частного случая. В соответствии с теоремой Эрбрана можно не рассматривать все возможные области интерпретации, а ограничиться рассмотрением лишь одной области, известной как универсум Эрбрана. Универсум Эрбрана может быть конструктивно построен, однако его нельзя рассматривать в качестве естественной модели математического мира, понятной математику. Очевидно, что установить истинность предложения на основе семантики языка исчисления предикатов можно лишь в том случае, если множество моделей теории, областью интерпретации которых является универсум Эрбрана, конечно, и если либо это предложение не содержит переменных, либо универсум Эрбрана конечен (теория не содержит имен, значениями которых являются функции). Будем говорить, что в первом случае — теорема, а во втором случае — теория являются тривиальными. Для математики же интерес представляют лишь нетривиальные теоремы и теории.

Под формальным доказательством в исчислении предикатов понимается конечная последовательность предложений, каждое из которых есть либо нелогическая аксиома, либо результат применения логической аксиомы (исчисления предикатов), либо значение заключения правила вывода, значения посылок которого предшествуют этому предложению в доказательстве. Если A есть последнее предложение в доказательстве P, то P есть доказательство предложения A [6]. Формальное доказательство моделирует только полное доказательство, состоящее только из прямых рассуждений. Оно имеет, по меньшей мере, два преимущества перед интуитивным. Во-первых, оно использует точно определенный логический базис (логические аксиомы и правила вывода), правильность которого может анализироваться отдельно от любого конкретного доказательства. Во-вторых, оно допускает формальную проверку правильности, т. е. может быть написана компьютерная программа, анализирующая формальное доказательство и отвечающая на вопрос, правильно оно или нет.

В отличие от математической практики, где правила рассуждения, применяемые для построения интуитивных доказательств, явным образом не стандартизованы, различные варианты исчисления предикатов предлагают различные варианты такой стандартизации — различные наборы правил вывода и логических аксиом. Логические аксиомы, в отличие от нелогических, истинны по смыслу логических символов, т. е. они истинны в каждой интерпретации любого множества имен [6]. В исчислении предикатов рассматриваются только такие правила вывода, у каждого из которых заключение является логическим следствием посылок. Для описания выдвигаемых предположений, в отличие от правил рассуждения в математической практике, правила вывода исчисления предикатов не содержат никаких средств.

Обычно каждой пропозициональной связке и квантору соответствует свое правило вывода. Однако не всякое такое правило вывода формализует правило рассуждения, принятое в математической практике. С другой стороны, многие правила рассуждения, принятые в математической практике, не формализованы в исчислении предикатов. В этом отношении крайним случаем является вариант исчисления предикатов, где нет логических аксиом, а единственное правило вывода исчисления — принцип резолюции — формализует правило рассуждения, обычно не применяемое в математической практике.

Таким образом, исчисление предикатов мало соответствует математической практике по следующим причинам:

— множество основных понятий невелико и фиксировано, вводить новые основные понятия невозможно;

— для определения производных понятий нет естественной конструкции;

— формальные естественные аналоги имеются лишь для ограниченного набора оборотов математического диалекта, а формальные естественные аналоги для других оборотов вводить невозможно;

— математические выражения (в принятом в математической практике виде) включать в предложения нельзя;

— все переменные, входящие в теорию, в каждой интерпретации имеют одну и ту же область значений;

— нетривиальная теория задает бесконечное множество моделей математических миров;

— универсум Эрбрана не является естественной совокупностью объектов математического мира, понятной математику;

— формальное доказательство моделирует только прямые рассуждения;

— правила вывода и логические аксиомы формализуют только часть правил рассуждения, применяемых в математической практике;

— некоторые правила вывода исчисления предикатов формализуют такие правила рассуждения, которые не применяются в математической практике;

— для описания выдвигаемых предположений правила вывода не содержат никаких средств;

— формальное доказательство всегда является моделью полного интуитивного доказательства, модели сокращенного интуитивного доказательства не существует.

Установление четких правил игры при конструировании формальных доказательств и возможность формальной проверки их правильности в исчислении предикатов оборачиваются неестественностью этой модели сравнительно с математической практикой. Это вызывает значительные трудности при попытках применить эту модель в математической практике, а именно, переводить описание математического мира и математические знания на язык исчисления предикатов и строить формальные доказательства по интуитивным. Бедность языка исчисления предикатов ведет к тому, что сравнительно простые математические идеи часто выражаются громоздкими конструкциями. Бедность самого исчисления предикатов ведет к тому, что формальные доказательства по размеру значительно превосходят интуитивные и значительно отличаются от них в содержательном плане. Отсутствие математической модели сокращенных интуитивных доказательств затрудняет переход от формальных доказательств к интуитивным. Все эти особенности исчисления предикатов не позволили применять в математической практике формально-логический критерий достоверности математических знаний. Даже в работах по математической логике (!) для теорем строятся не формальные, а интуитивные доказательства.

4. РЕШЕНИЕ ПРОБЛЕМЫ В РАМКАХ КОМПЬЮТЕРНОЙ ЛОГИКИ

Одно из направлений в области искусственного интеллекта — разработка методов поиска правильных формальных доказательств с помощью компьютеров [7]. В рамках этого направления были предложены такие варианты исчисления предикатов первого порядка (принцип резолюции, обратный вывод), на основе которых разработаны компьютерные программы, формирующие правильное формальное доказательство по записи теоремы и нелогических аксиом на языке исчисления предикатов. Тем самым в рассмотрение был введен «компьютерный» критерий правильности математических знаний: чтобы установить, что нетривиальная теорема правильна, нужно построить для ее модели на языке исчисления предикатов компьютерное доказательство, являющееся моделью ее полного интуитивного доказательства, и установить правильность нелогических аксиом, на моделях которых оно основано. Отличие компьютерного критерия от формально-логического состоит в том, что компьютерное доказательство строится компьютером и поэтому более надежно, чем формальное доказательство, которое строится человеком. С помощью компьютера можно построить более сложные (длинные) доказательства. Однако чем сложнее (длиннее) доказательство, тем труднее его понять. Поэтому математики могут лишь верить в то, что такие доказательства правильны (правильно построены), но не могут понимать их [3].

Наиболее полно изучены системы автоматического доказательства теорем, в которых компьютерные доказательства строятся без участия человека [7]. Однако их применение столкнулось со значительными трудностями. Причины их — в полуразрешимости исчисления предикатов (если утверждение является теоремой теории, то его компьютерное доказательство может быть найдено за конечное число шагов, но если это утверждение не является теоремой, то поиск может продолжаться бесконечно), в высокой временной сложности алгоритмов автоматического доказательства теорем, в бедности языка исчисления предикатов (что затрудняет построение информационной базы для систем автоматического доказательства теорем), в отсутствии математической модели интуитивных доказательств (что затрудняет переход от компьютерных доказательств к интуитивным).

Другое направление конструирования компьютерных доказательств — автоматизированные системы, в которых поиском компьютерного доказательства управляет математик. Математик решает творческую задачу — сокращает сложность конструирования доказательств за счет своих дополнительных знаний, которые не заложены в программу. Компьютер выполняет всю рутинную работу и обеспечивает правильность результата. Такие программы разрабатываются в последнее время чаще всего в учебных целях [8—11]. Системы автоматизированного доказательства теорем не сталкиваются с проблемами вычислительной сложности, поскольку управление со стороны математика как раз и служит целям сокращения перебора при поиске доказательства. Однако такие системы, как правило, основаны на моделях математической логики, и математик, пользующийся ими, сталкивается с проблемами интеллектуальной сложности из-за неадекватности этих моделей математической практике. Основные причины трудностей применения таких систем — бедность исчисления предикатов (что затрудняет человеко-машинное конструирование компьютерных доказательств), а также, как и в случае систем автоматического доказательства теорем, бедность языка исчисления предикатов (что затрудняет построение информационной базы для систем автоматизированного доказательства теорем) и отсутствие математической модели интуитивных доказательств (что затрудняет переход от компьютерных доказательств к интуитивным).

5. ПУТИ ДАЛЬНЕЙШЕГО ПРОДВИЖЕНИЯ В РЕШЕНИИ ПРОБЛЕМЫ

Результаты проведенного анализа позволяют увидеть то направление, в котором необходимо двигаться при решении рассматриваемой проблемы. Прежде всего, в основу решения проблемы должен быть положен компьютерный критерий как обеспечивающий наибольшую правильность математических знаний. Далее, успех может обеспечить разработка систем автоматизированного доказательства теорем, поскольку только таким образом можно обойти проблемы вычислительной сложности — благодаря творческим способностям и дополнительным знаниям самих математиков. Наконец, следует разрабатывать более адекватные модели математической практики для этих систем.

В настоящее время ведется разработка такой модели математической практики [12, 13]. Она обладает следующими свойствами.

Модель математического диалекта двухуровневая — она содержит средства метаязыка для введения в язык формальных аналогов новых оборотов, неопределяемых понятий и способов записи. Расширяемый таким образом язык позволяет формулировать формальные математические и метаматематические утверждения и определения.

Объекты модели математического мира моделируют наиболее естественным образом математические объекты, включая их внутреннюю структуру. Формальное описание математического мира задает единственную его модель и имеет единственную интерпретацию.

Модель доказательства включает в себя все перечисленные выше виды рассуждений — прямые и обратные, логические, специальные и содержательные, выдвижения предположений и условные рассуждения. Наряду с полными доказательствами в модели присутствует и модель интуитивных (сокращенных) доказательств, а также модель объяснения сокращенных доказательств.

Разработка этой модели ведется с использованием последних достижений в области искусственного интеллекта — теории онтологий [14—16]. При проектировании системы автоматизированного доказательства теорем, основанной на такой модели, принята концепция компьютерных банков знаний, описанная в работах [17, 18].

ЛИТЕРАТУРА

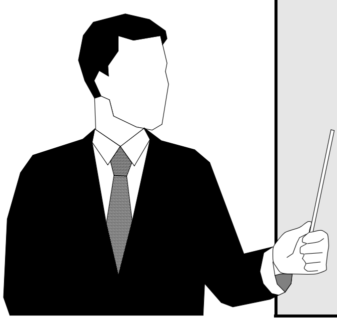
1. *Пойа Д.* Как решать задачу. — М.: Учпедгиз, 1961. — 207 с.
2. *Лорьер Ж.-Л.* Системы искусственного интеллекта. — М.: Мир, 1991. — 568 с.
3. *Непейвода Н. Н.* Прикладная логика. — Ижевск: Удмуртский ун-т, 1997. — 384 с.
4. *Гильберт Д., Бернайс П.* Основания математики. Логические исчисления и формализация арифметики. — М.: Наука, 1979. — 557 с.
5. *Крайзель Г.* Исследования по теории доказательств. — М.: Мир, 1981. — 289 с.
6. *Шёнфилд Дж.* Математическая логика. — М.: Наука, 1975. — 527 с.
7. *Чень Ч., Ли Р.* Математическая логика и автоматическое доказательство теорем. — М.: Наука, 1983. — 320 с.
8. *Peter B. Andrews, Matthew Bishop, Chad E. Brown, et al.* ETPS: A System to Help Students Write Formal Proofs // Research Report No. 03-002, April 2003 (Department of Mathematical Sciences. Carnegie Mellon University).
9. *Christoph Benzmuller and Volker Sorge.* A Blackboard Architecture for Guiding Interactive Proofs. F. Giunchiglia (Ed.): AIMSA'98, LNAI 1480, 1998. P. 102—114.
10. *Stephan Schmitt, Lori Lorigo, Christoph Kreitz, and Aleksey Nogin.* JProver: Integrating Connection-Based Theorem Proving into Interactive Proof Assistants. International Joint.
11. *Ulrich Endriss.* The Interactive Learning Environment WinKE for Teaching Deductive Reasoning. First International Congress on Tools for Teaching Logic. King's College, London — United Kingdom — September 6, 2000.
12. *Гаврилова Т. Л., Клещёв А. С.* Проблема конструирования правильных интуитивных доказательств в классических аксиоматических системах. Ч. 2. Конструирование полных доказательств. — Владивосток: ИАПУ ДВО РАН, 2003. — 34 с. (http://www.iacp.dvo.ru/es/publ/185_1.rtf).
13. *Гаврилова Т. Л., Клещёв А. С.* Проблема конструирования правильных интуитивных доказательств в классических аксиоматических системах. Ч. 3. Модель интуитивного доказательства. — Владивосток: ИАПУ ДВО РАН, 2003. — 21 с. (http://www.iacp.dvo.ru/es/publ/185_2.rtf).
14. *Клещёв А. С., Артемьева И. Л.* Математические модели онтологий предметных областей. Ч. 1. Существующие подходы к определению понятия «онтология»//НТИ. Сер. 2. — 2001. — № 2. — С. 20—27. (www.iacp.dvo.ru/es/publ/104_1.rtf).
15. *Клещёв А. С., Артемьева И. Л.* Математические модели онтологий предметных областей. Ч. 2. Компоненты модели//НТИ. Сер. 2. — 2001. — № 3. — С. 19—29. (www.iacp.dvo.ru/es/publ/104_2.rtf).
16. *Клещёв А. С., Артемьева И. Л.* Математические модели онтологий предметных областей. Ч. 3. Сравнение разных классов моделей онтологий//НТИ. Сер. 2. — 2001. — № 4. — С. 10—15. (www.iacp.dvo.ru/es/publ/104_3.rtf).
17. *Орлов В. А., Клещёв А. С.* Многоцелевой банк знаний. Ч. 1. Концепция и политика. — Владивосток: ИАПУ ДВО РАН, 2003. — 40 с. (www.iacp.dvo.ru/es/publ/186_1.rtf).
18. *Орлов В. А., Клещёв А. С.* Многоцелевой банк знаний. Ч. 2. Требования. — Владивосток: ИАПУ ДВО РАН, 2003. — 23 с. (www.iacp.dvo.ru/es/publ/186_2.rtf).

☎ *(4232) 31-40-01, 31-04-24*

E-mail: gavrilova@iacp.dvo.ru

E-mail: kleshev@iacp.dvo.ru ❑

Читайте в следующем номере



Копнин М.Ю., Кульба В.В., Микрин Е.А. *Структурно-технологический резерв и его использование для повышения устойчивости производственных систем*

Колемаев В.А., Бережной А.Е. *Моделирование смены технологического уклада*

Ведешенков В.А. *Способ самодиагностирования неоднородных цифровых систем*

Левин В.И. *Логическое моделирование разрывных функций*

Полетыкин А.Г. *Особенности разработки программного обеспечения сложных интегрированных АСУТП*

Бабушкина Н.А., Островская Л.А., Рыкова В.А. *и др. Моделирование эффективности действия противоопухолевых препаратов в сверхмалых дозах для оптимизации режимов их введения*

Абрамянц Т.Г., Маслов Е.П., Рубинович Е.Я. *Управление подвижными объектами в условиях искусственно организованной неполноты информации*

Научные *чтения памяти А.М. Петровского*

УДК 517.96+681.3

ПРИМЕНЕНИЕ НЕЧЕТКИХ МЕР И ИНТЕГРАЛОВ К ОПИСАНИЮ НЕЧЕТКИХ ДИНАМИЧЕСКИХ СИСТЕМ

С. Л. Блюмин, А. М. Шмырин

Липецкий государственный технический университет

Рассмотрен подход к учету нечеткости окрестностей по состоянию дискретно-временных систем на основе использования нечетких мер и интегралов.

ВВЕДЕНИЕ

Нечеткие меры и интегралы широко исследуются в современной прикладной математике, прежде всего, в связи с проблемами искусственного интеллекта [1]. Естественная область их определения и изучения, мотивируемая значительным числом приложений, находится в рамках элементарной теории конечных мер и интегралов на булеанах над полукольцами [2].

Нечеткие системы Вольтерра введены и исследованы в работе [3] в качестве примера реализации подхода к учету нечеткости окрестностей по состоянию дискретно-временных систем. В ней же обсуждается скалярная линейная нестационарная дискретная система Вольтерра

$$x[t] = \sum_{\tau = 0}^{t-1} \alpha[t; \tau] = \sum_{\tau \in T} \alpha[t; \tau]x[\tau], \quad (1)$$

где $t \in N = \{0, 1, ...\}$ — дискретное время, $x[t] \in R^n$ — вектор состояния системы, $\alpha[t; \tau]$ — коэффициенты. Скалярная нечеткая линейная по состояниям дискретная система типа Вольтерра описывается сопоставимым с выражением (1) уравнением

$$x[t] = \sum_{\tau = 0}^{t-1} \mu[t; \tau]x[\tau] = \sum_{\tau \in T} \mu[t; \tau]x[\tau], \quad (2)$$

где $\mu[t; \tau] \in [0, 1]$ — функции принадлежности.

Цель данной работы — показать, что нечеткие меры и интегралы являются подходящим математическим аппаратом для описания нечетких динамических систем. Систематическое изложение этого аппарата и его приложений к проблемам нечеткости и мягких вычислений можно найти в работе [4]. В последнее время нечеткие меры и интегралы интенсивно применяются в задачах агрегирования критериев при многокритериальном принятии решений [5—8].

1. НЕЧЕТКИЕ МЕРЫ

Понятие нечетких меры и интеграла было введено М. Сугено (см., например, работы [4—6]). Для современных приложений особый интерес эти понятия представляют применительно к конечным множествам. Пусть $T = MNX[t] = \{0, ..., t - 1\}$ — такое множество; при рассмотрении динамических систем оно наделено естественным порядком следования его элементов. Нечеткая мера на множестве T определяется как функция множества $m : 2^T \to [0, 1]$ (здесь 2^T — булеан или множество всех подмножеств множества T), обладающая свойствами $m(\varnothing) = 0$, $m(T) = 1$, $A \subseteq B \Rightarrow m(A) \leqslant m(B)$, где A, B — подмножества множества T.

Подчеркивается, что требование аддитивности (в обычном смысле, конкретный пример см. ниже), обязательно присутствующее в стандартных определениях мер, к нечетким мерам, вообще говоря, не предъявляется. Однако отметим, что, с позиций идемпотентной математики [9], при рассмотрении промежутка [0, 1] как полукольца с операциями max и min вместо базовых арифметических операций идемпотентная мера, как пример нечеткой меры, свойством аддитивности обладает (конкретные примеры см. ниже).

Прототипом нечетких мер является обычная вероятностная мера P — конкретный пример аддитивной (в обычном смысле) меры: если $A \cap B = \varnothing$, то $P(A \cup B) = P(A) + P(B)$. Конкретными примерами идемпотентно-аддитивных мер служат мера Π возможности, для которой $\Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}$, и мера N достоверности (необходимости), для которой $N(A \cup B) = \min\{N(A), N(B)\}$. Известен класс мер Сугено, характеризующихся параметром $\lambda > -1$ и обладающих свойством $S(A \cup B) = S(A) + S(B) + \lambda S(A)S(B)$ (при $\lambda = 0$ это обычные аддитивные меры), меры доверия и правдоподобия и др. (см., например, работы [4, 5]).

2. НЕЧЕТКИЙ ИНТЕГРАЛ СУГЕНО И СИСТЕМА ВОЛЬТЕРРА—СУГЕНО

Пусть $f: T \to [0, 1]$ — некоторая функция. Нечеткий интеграл Сугено от этой функции по нечеткой мере $m = m_t$ на множестве T определяется выражением

$$S(f, m, T) = \max_{\alpha \in [0, 1]}\{\min[\alpha, m_t, (F_\alpha)]\},$$

где $F_\alpha = \{t \in T : f[t] \geqslant \alpha\}$. Если $f = \mu$ — функция принадлежности некоторого нечеткого подмножества множества T, то F_α — множество уровня α этого подмножества. Другое представление этого интеграла можно получить в предположении, что функция f не возрастает на множестве T, так что при $t > s$ выполняется условие $f[t] \leqslant f[s]$ (для произвольной функции множество T может быть переупорядочено). В этом случае интеграл Сугено может быть представлен в виде

$$S(f, m, T) = \max_{\tau \in T}\{\min[m_t([0, \tau]), f[\tau]]\},$$

допускающем сопоставление с правой частью уравнения (2), описывающего линейные нечеткие дискретные системы типа Вольтерра. Выражения находятся в следующем соответствии: состоянию $x[\tau]$ соответствует функция $f[\tau]$, коэффициентам $m[t; \tau]$ соответствует мера $m_t([0, \tau])$, арифметическим операциям сложения и умножения соответствуют операции max и min. Отмеченное соответствие позволяет определить нечеткую дискретную максминную систему типа Вольтерра—Сугено уравнением

$$x[t] = \max_{\tau \in T}\{\min[m_t([0, \tau]), x[\tau]]\}.$$

Ее наиболее бросающиеся в глаза отличия от систем, рассмотренных в работе [3], — использование максминных операций, ограничение области значений состояния промежутком [0, 1] и его невозрастание.

3. НЕЧЕТКИЙ ИНТЕГРАЛ МАСЛОВА И СИСТЕМА ВОЛЬТЕРРА—МАСЛОВА

Другой тип нечеткого интеграла — идемпотентный интеграл Маслова [9] на полуполе $R_{\max}$ или *MAXPLUS*, в котором сложение заменяется операцией взятия максимума, а умножение — обычным арифметическим сложением — предполагает, что нечеткая идемпотентная мера определяется некоторой заданной на множестве T функцией $\varphi = \varphi_t$ в виде $m(A) = m_{t\varphi}(A) = \max\limits_{\tau \in A}\varphi_t[\tau]$ для всех $A \subseteq T$; сам идемпотентный интеграл определяется выражением $I(f, m, T) = \max\limits_{\tau \in T}\{\varphi_t[\tau] + f[\tau]\}$, сравнение которого с уравнением (2) позволяет определить нечеткую дискретную максплюсовую систему типа Вольтерра—Маслова уравнением

$$x[t] = \max_{\tau \in T}\{\varphi_t[\tau] + x[\tau]\},$$

характерное отличие которой также состоит в использовании модифицированных базовых операций.

4. НЕЧЕТКИЙ ИНТЕГРАЛ ШОКЕ И СИСТЕМА ВОЛЬТЕРРА—ШОКЕ

Наиболее близкое к рассмотренным ранее представление имеет интеграл Шоке по нечеткой мере [4—6]. Здесь, как и в случае интеграла Сугено, предполагается, что функция f не возрастает на множестве T, но область ее значений ограничена не промежутком [0, 1], а более широким промежутком $[0, \infty)$. Интеграл Шоке в базовых арифметических операциях определяется выражением $C(f, m, T) = \sum\limits_{\tau \in T}\{m_t([0, \tau]) - m_t([0, \tau - 1])\}f[\tau]$ и оказывается наиболее сопоставимым с правой частью уравнения (2): состоянию $x[\tau]$ соответствует функция $f[\tau]$, коэффициентам $\mu[t; \tau]$ соответствует приращение меры $m_t([0, \tau]) - m_t([0, \tau - 1])$; иных отмеченных выше отличий нет. Это позволяет определить нечеткую дискретную систему типа Вольтерра—Шоке уравнением

$$x[t] = \sum_{\tau \in T}\{m_t([0, \tau]) - m_t([0, \tau - 1])\}x[\tau].$$

Такая система допускает, например, представляющую определенный интерес интерпретацию условий устойчивости дискретных систем Вольтерра (1), приведенных в работе [3]: так как по определяющим свойствам нечеткой меры

$$\begin{aligned}\alpha_0 = \sup_{t \geq 0}\sum_{\tau = 0}^{\infty}|a[t; \tau]| = \sup_{t \geq 0}\sum_{\tau = 0}^{\infty}|m_t([0, \tau]) - m_t([0, \tau - 1])| = \\ = \sup_{t \geq 0}\sum_{\tau = 0}^{t-1}\{m_t([0, \tau]) - m_t([0, \tau - 1])\} = \\ = \sup_{t \geq 0}\{m_t(T) - m(\varnothing)\} = 1\end{aligned}$$

(при естественном соглашении $[0, -1] = \varnothing$), то система является устойчивой, но не асимптотически устойчивой. В работе [3] отмечено, что в современных приложениях нечетких систем классическое структурное свойство устойчивости далеко не самое важное.

Сопоставление введенных здесь нечетких систем с системами, рассмотренными в работе [3], показывает, что в уравнениях систем Вольтерра—Сугено и Вольтерра—Шоке величины $m_t([0, \tau])$ и $m_t([0, \tau]) - m_t([0, \tau - 1])$ допускают трактовку как значения функций принадлежности нечетких окрестностей по состояниям: из определяющих свойств нечеткой меры следует, что эти величины принимают значения в промежутке [0, 1].

Особого внимания требует предположение о невозрастании состояния введенных здесь систем на промежутке $[0, t - 1]$: если не придавать динамического смысла рассмотренным нечетким интегралам, то такое предположение можно считать не слишком обременительным, поскольку, как уже отмечалось, область определения можно соответствующим образом переупорядочить. Однако при динамической трактовке, как тоже отмечалось, оно наделено естественным порядком следования его элементов — моментов времени, и переупорядочение может затруднить динамическую интерпретацию. Здесь могут оказаться полезными соображения, используемые при исследовании нечетких конечно-аргументных систем [10], когда множество моментов времени формализуется при помощи конечной группы, в которой групповая операция не согласована с естественным временным порядком.

ЗАКЛЮЧЕНИЕ

Отметим, что нечеткие системы Вольтерра—Сугено и Вольтерра—Маслова не являются линейными в обычном смысле слова, но «линейны» над соответствующими идемпотентными полукольцами, подобно отмеченной выше аддитивности мер. Это одно из простейших в современной прикладной математике проявлений основной парадигмы идемпотентной математики [9], выражаемой идемпотентным принципом эвристического соответствия между важными, полезными и интересными конструкциями и результатами традиционной математики над числовыми полями и аналогичными конструкциями и результатами над идемпотентными полукольцами и полуполями. Идемпотентная математика, возникающая как двойник или «тень» традиционной, соотносится с последней примерно так же, как классическая физика с квантовой; переход от традиционной математики к идемпотентной трактуется как «деквантование» или «вторичное деквантование». Во многих отношениях идемпотентная математика проще традиционной; однако переход от традиционных понятий и результатов к их идемпотентным аналогам часто нетривиален. Некоторые соотношения между линейной и «линейной» алгебрами обсуждены в работе [1].

ЛИТЕРАТУРА

1. *Блюмин С. Л.* Математические проблемы искусственного интеллекта: регулярность по Дж. фон Нейману в линейной и «линейной» алгебрах // Системы управления и информационные технологии. — 2003. — № 1—2 (12). — С. 90—94.
2. *Блюмин С. Л.* Конечные меры и интегралы на булеанах над полукольцами // Материалы Воронеж. весенней мат. школы «Современные методы теории краевых задач. Понтрягинские чтения — XV» / ВГУ. — Воронеж, 2004. — С. 35.
3. *Блюмин С. Л., Шмырин А. М.* Нечеткие системы Вольтерра // Проблемы управления. — 2004. — № 4. — С. 75—78.
4. *Grabisch M., Murofushi M., Sugeno M.* Fuzzy Measures and Integrals — Theory and Applications. Studies in Fuzziness and Soft Computing. Vol. 40. — Heidelberg: Springer-Verlag, 2000. — 543 p.
5. *Yager R.* Criteria aggregations functions using fuzzy measures and the Choquet integral // Int. J. of Fuzzy Systems. — 1999. — Vol. 1, N 2. — P. 96—112.
6. *Ovchinnikov S.* Piecewise linear aggregation functions // Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems. — 2000. — Vol. 1, N 1. — P. 11—22.
7. *Klement E., Mesiar R., Pap E.* Measure-based aggregation operators // Fuzzy Sets and Systems. — 2004. — Vol. 142, N 1. — P. 3—14.
8. *Calvo T., Pradera A.* Double aggregation operators // Fuzzy Sets and Systems. — 2004. — Vol. 142, N 1. — P. 15—33.
9. *Литвинов Г. Л., Маслов В. П.* Идемпотентная математика // Докл. Воронеж. зимней мат. школы «Современный анализ и его приложения» / ВГУ. — Воронеж, 2000. — С. 20—21.
10. *Блюмин С. Л., Шмырин А. М.* Нечеткие окрестностные конечные системы // Материалы Воронеж. весенней мат. школы «Современные методы теории краевых задач. Понтрягинские чтения — XIV» / ВГУ. — Воронеж, 2003. — С. 26.

☎ *(0742) 32-81-33*

E-mail: amsh@lipetsk.ru ❑

III МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ ПО ПРОБЛЕМАМ УПРАВЛЕНИЯ – МКПУ III

Москва, 20 - 22 июня 2006 г.

На МКПУ III предполагается рассмотреть широкий круг вопросов, связанных с проблематикой современной теории управления: анализ и синтез систем управления, оптимальное управление и распределённые системы, адаптивные и робастные системы управления, стохастические системы управления, идентификация, системный анализ и теория систем, теория выбора и принятия решений, управление безопасностью сложных систем, системы распознавания образов и анализ данных, управление в медико-биологических, социально-экономических и организационных системах, активные системы, человеко-машинные системы, управление технологическими процессами и предприятиями, технические средства управления, системы логического управления, надёжность и техническая диагностика, вычислительные системы и сетевые технологии, искусственный интеллект, нейронные сети и системы управления, автоматизированное проектирование, управление транспортными потоками, управление в логистике, управление подвижными объектами.

Конференция состоится в Институте проблем управления им. В.А. Трапезникова РАН по адресу:
Москва, Профсоюзная ул., 65. *Официальные языки конференции – русский, английский.*

Заявки на участие в конференции принимаются по адресу:
117997 Москва, ГСП-7, Профсоюзная ул., 65, Институт проблем управления, лаб. № 44,
Оргкомитет III Международной конференции,
тел./факс (095) 334-89-69, e-mail: iccpripu@ipu.rssi.ru

Предварительная программа, правила оформления тезисов и сумма регистрационного взноса будут представлены на сайте Института проблем управления им. В.А. Трапезникова РАН (www.ipu.rssi.ru) с 15 мая 2005 г.

УДК 65.012

МЕТОД СЕТЕВОГО ПРОГРАММИРОВАНИЯ

В. Н. Бурков, И. В. Буркова, М. В. Попок, Т. И. Овчинникова

Институт проблем управления им. В.А. Трапезникова, г. Москва

Предложен новый подход к задачам дискретной оптимизации, названный методом сетевого программирования, в основу которого положена возможность представления функции многих переменных в виде суперпозиции более простых функций. Структура такой суперпозиции представляется в виде сети, входы которой соответствуют переменным, а выходы — функции. Показано, что если сеть является деревом, то решение задачи сводится к последовательному решению более простых задач. В общем случае предложено преобразовать сеть в дерево путем разделения вершин сети. Доказано, что решение задачи для преобразованной структуры дает нижнюю оценку для целевой функции исходной задачи (если решается задача минимизации). Метод проиллюстрирован на примере известной задачи о камнях.

ВВЕДЕНИЕ

Многие задачи дискретной оптимизации сводятся к следующей постановке: определить вектор $x = \{x_i\}$ с дискретными компонентами, минимизирующий аддитивную функцию

$$\varphi(x) = \sum_{i=1}^{n} \varphi_i(x_i) \tag{1}$$

при ограничении

$$f(x) \geqslant b. \tag{2}$$

Любая функция дискретных переменных допускает сетевое представление, такое, что вычисление значений функции сводится к последовательному вычислению значений более простых функций. В частности, любая функция дискретных переменных допускает дихотомическое представление, когда вычисление значения функции сводится к последовательному вычислению значений функций двух переменных. Так, функция $f(x) = f_0[f_1(x_1, x_2), f_2(x_2, x_3)]$ допускает дихотомическое представление (рис. 1). При этом функции f_0, f_1 и f_2 удобно представлять в матричном виде (рис. 2). Такое представление широко используется в методах комплексного оценивания программ развития предприятий, регионов, результатов деятельности подразделений, уровня безопасности объектов и др.

В работах [1, 2] доказаны теоремы о представлении непрерывных функций нескольких переменных суперпозициями непрерывных функций меньшего числа переменных (в частности, двух переменных). Так, например, любая непрерывная функция трех переменных представима в виде [2] $f(x_1, x_2, x_3) = h^1(x_1, \varphi_1(x_2, x_3)) + h^2(x_1, \varphi_2(x_2, x_3)) + h^3(x_1, \varphi_3(x_2, x_3))$. Ее сетевое представление приведено на рис. 3.

В сетевом виде можно представить и систему неравенств. Рассмотрим, например, систему неравенств

$$f_j(x) \leqslant b_j, \quad j = \overline{1, m}. \tag{3}$$

Без ограничения общности можно принять, что b_j — положительные и одинаковые числа, $b_j = b > 0$. В этом случае систему неравенств (3) можно заменить одним неравенством $f(x) \leqslant b$, где $f(x) = \max_j f_j(x)$.

Очевидно, что функция $f(x)$ допускает сетевое представление, если все функции f_j допускают такое представление.

В настоящей работе описывается новый метод решения задач дискретной оптимизации, использующий сетевое представление функции $f(x)$. Его естественно на-

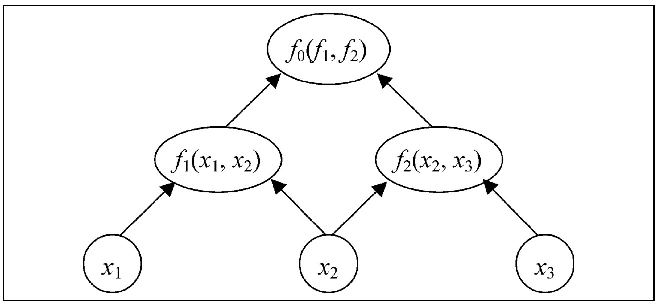


Рис. 1. Дихотомическое представление функции дискретных переменных

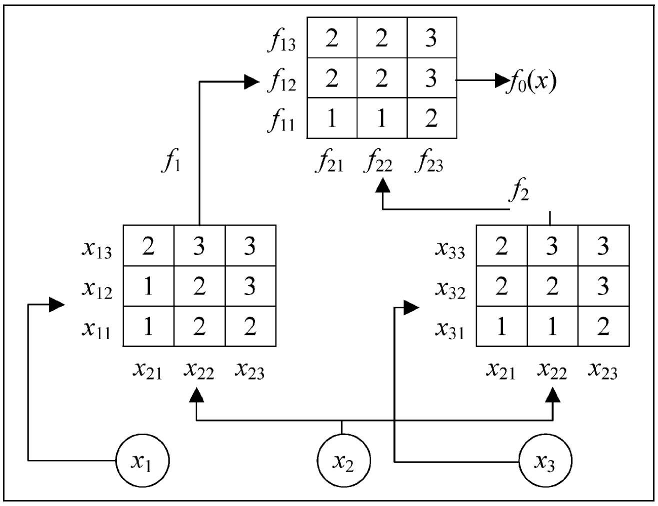


Рис. 2. Представление функций f_0, f_1 и f_2 в матричном виде

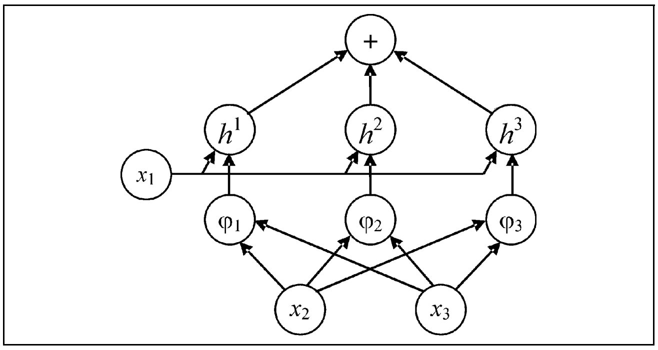


Рис. 3. Сетевое представление непрерывных функций трех переменных

звать методом сетевого программирования (в частном случае дихотомического представления получаем метод дихотомического программирования [3]).

1. СЕТЕВОЕ ПРЕДСТАВЛЕНИЕ ТИПА ДЕРЕВА

Рассмотрим случай, когда функция $f(x)$ допускает сетевое представление в виде дерева. Дадим описание метода сетевого программирования для задачи (1), (2). На рис. 4 приведен пример функции трех переменных, имеющей вид

$$f(x_1, x_2, x_3) = \varphi_0[f_1(x_1, x_2), x_3] = \varphi_0(y, x_3).$$

Значения функций $\varphi_i(x_i)$ даны в нижних половинах квадратов, соответствующих переменным x_1, x_2 и x_3. Идея метода состоит в следующем. Сначала решается задача минимизации функции двух переменных

$$\varphi_1(x_1) + \varphi_2(x_2)$$

при ограничении

$$f_1(x_1, x_2) \geqslant y,$$

соответствующая начальной вершине сетевого представления. Обозначим через $z(y)$ решение этой задачи в зависимости от y. Далее решаем задачу минимизации функции тоже двух переменных

$$z(y) + \varphi_3(x_3)$$

при ограничении

$$\varphi_0(y, x_3) \geqslant b,$$

соответствующую конечной вершине сетевого представления. Решение этой задачи определяет оптимальное решение исходной задачи.

Проиллюстрируем метод на примере рис. 4.

1 шаг. Рассматриваем нижнюю матрицу и для каждого элемента этой матрицы записываем в нижней половине соответствующей клетки сумму функций $\varphi_1(x_1)$ и $\varphi_2(x_2)$ для соответствующих значений x_1 и x_2. Так, например, клетке $(x_1, x_2) = (3, 2)$ соответствует сумма

$$\varphi_1(3) + \varphi_2(2) = 20 + 10 = 30.$$

Далее будем называть эту сумму затратами на достижение соответствующего состояния.

2 шаг. Из всех элементов матрицы, имеющих одно и то же значение $y = f_1(x_1, x_2)$, выбираем элемент с минимальной суммой $\varphi_1(x_1) + \varphi_2(x_2)$. Минимальную сумму записываем в нижнюю половину клетки, соответствующей этому значению y в верхней матрице. Так, например, значению $y = 3$ соответствуют 5 элементов нижней матрицы: (3; 2), (4; 2), (3; 3), (4; 3) и (2; 4). Из них элемент (3; 2) имеет минимальную сумму 30 (это число записано в нижней половине соответствующей клетки).

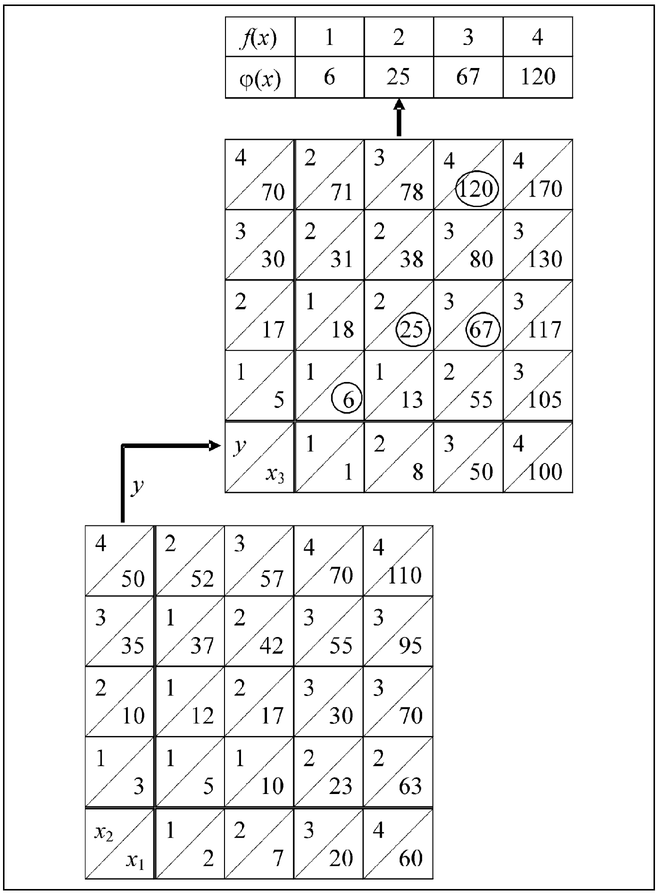


Рис. 4. Иллюстрация метода сетевого программирования

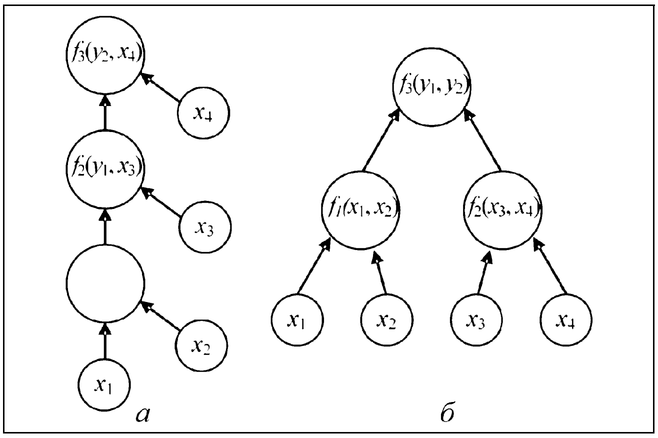


Рис. 5. Иллюстрация методов программирования:
а — динамического (ветвь дерева); *б* — дихотомического (произвольное дерево)

Поэтому в верхней матрице значению $y = 3$ соответствует число 30, записанное в нижней половине соответствующей клетки.

Далее шаги 1 и 2 повторяются для верхней матрицы. В результате для каждого значения $f(x)$ мы получаем минимальное значение $\varphi(x)$. На рис. 4 кружками выделены минимальные затраты.

Несложно обобщить описанный метод на случай произвольного сетевого представления функции $f(x)$ в виде дерева. Главное, чтобы задачи, соответствующие каждой вершине сетевого представления, имели эффективные методы решения. В случае дихотомического представления это всегда имеет место.

Заметим, что дихотомическое представление (см. рис. 4) имеет структуру в виде ветви дерева. В этом случае метод дихотомического программирования переходит в метод динамического программирования. Таким образом, метод дихотомического программирования в случае дихотомического представления в виде дерева является обобщением метода динамического программирования, расширяя круг задач, решаемых на основе данного подхода (рис. 5).

Если в методе динамического программирования решением задачи является путь в некоторой специальным образом построенной сети, то в методе дихотомического программирования решением задачи является частичное дерево в некотором специально построенном дереве. Соответственно, принцип оптимальности в методе дихотомического программирования можно сформулировать следующим образом: любое поддерево оптимального дерева должно быть оптимальным. Формально этот принцип оптимальности можно записать следующим образом: $\varphi_k(y) = \min\limits_{(i,j) \in p(y)} [\varphi_i(y_i) + \varphi_j(y_j)]$, где $p(y)$ — множество пар (i, j) таких, что $f_k(y_i, y_j) = y$.

2. ОБЩИЙ СЛУЧАЙ

Рассмотрим произвольное сетевое представление функции $f(x)$, задаваемое сетью, выходом которой является вершина, соответствующая функции $f(x)$, а входами — вершины, соответствующие переменным x_i, $i = \overline{1, n}$. Рассмотрим множество начальных вершин, которые не являются висячими, т. е. их степени исхода больше единицы. Разделим произвольным образом затраты $\varphi_i(x_i)$ на k_i частей, где k_i — число исходящих дуг. Фактически мы как бы разделили вершину i на k_i висячих вершин с соответствующей частью затрат. Далее применяем описанный выше алгоритм. При этом каждый раз, когда встречается вершина со степенью исхода больше единицы, мы делим затраты на соответствующее число частей. В результате применения алгоритма мы получим оптимальное решение для модифицированной сети. Однако это решение может не быть решением исходной задачи. Тем не менее, имеет место следующая

Теорема. *Полученное с помощью вышеописанного алгоритма решение дает нижнюю оценку оптимального решения исходной задачи.*

Доказательство. Заметим, что множество решений модифицированной сети содержит все решения исходной задачи. Эти решения имеют следующий вид. Если из вершины, соответствующей переменной x_{ik}, исходит хотя бы одна дуга полученного решения, то все дуги, исходящие из этой вершины, также принадлежат полученному решению. Отсюда следует, что полученное оптимальное решение модифицированной задачи дает нижнюю оценку для оптимального решения исходной задачи.

Пример. Рассмотрим сеть, представленную на рис. 1 и 2. На рис. 6 приведено решение соответствующей задачи. Затраты $\varphi_2(x_2)$ разделены на две части, поскольку переменная x_2 используется при вычислении функций и f_1, и f_2. В данном случае общие затраты, равные 8, 12 и 20 при значениях переменной x_2, равных 1, 2 и 3, соответственно, поделены пополам.

В каждой матрице кружками выделены клетки, соответствующие минимальным затратам на получение того или иного значения функций f_1, f_2 и f_0. В результате по-

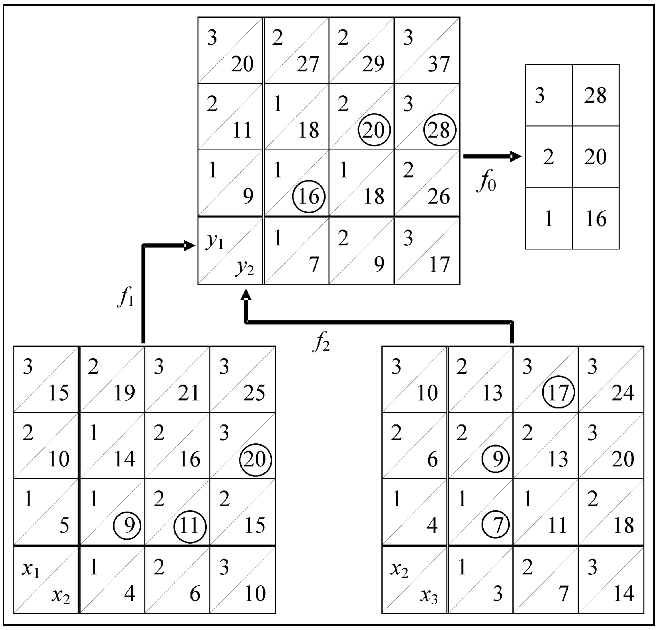


Рис. 6. Решение оценочной задачи методом сетевого программирования

лучены минимальные затраты $\varphi(f_0)$, требуемые для получения значений функции f_0. Если $f_0 = 1$, то $\varphi(1) = 16$, если $f_0 = 2$, то $\varphi(2) = 20$, если $f_0 = 3$, то $\varphi(3) = 28$.

Рассмотрим случай $f_0 = 2$. Ему соответствует оптимальное решение модифицированной задачи: $x_1 = 1$, $x_{21} = 2$, $x_{22} = 2$, $x_3 = 1$.

Здесь x_{21} соответствует значению x_2 в левой нижней матрице, а x_{22} — в правой нижней матрице. Поскольку оба значения $x_{21} = 2$, $x_{22} = 2$ вошли в оптимальное решение модифицированной задачи, то полученное решение является допустимым для исходной задачи, а значит, мы получили оптимальное решение исходной задачи.

Другая ситуация возникает в случае $f_0 = 3$. Оптимальное решение модифицированной задачи имеет вид: $x_1 = 1$, $x_{21} = 2$, $x_{22} = 3$, $x_3 = 2$ со значением затрат $\varphi_0 = 28$. Это решение не является допустимым для исходной задачи, поэтому значение $\varphi_0 = 28$ является нижней оценкой минимальных затрат для исходной задачи. Здесь возможны два варианта действий. Первый заключается в попытке улучшить нижнюю оценку, изменяя разбиение затрат $c_2 = \varphi_2(x_2)$ на две части — c_{21} и c_{22}. Очевидно, что для улучшения оценки следует часть c_{21} увеличить, а часть c_{22} уменьшить. Возьмем, например, $c_{21} = 10$, а $c_{22} = 2$. В этом случае оптимальное решение модифицированной задачи будет иметь вид: $x_1 = 1$, $x_{21} = 2$, $x_{22} = 2$, $x_3 = 3$ с затратами $\varphi_0 = 31$. Это решение является допустимым для исходной задачи, а значит, оптимальным. Однако изменение разбиения затрат на части может и не привести к получению допустимого решения для исходной задачи.

Второй вариант состоит в применении метода ветвей и границ. Разобьем множество всех решений исходной задачи на два подмножества: в первом из них $x_2 \leqslant 2$, а во втором $x_2 = 3$, и применим описанный выше алгоритм. Получим оценку снизу для первого подмножества. Получаем следующее решение: $x_1 = 1$, $x_{21} = 2$, $x_{22} = 2$, $x_3 = 3$ с затратами $\varphi_0 = 31$.

Получим оценку снизу для второго подмножества. Оптимальное решение модифицированной задачи имеет вид: $x_1 = 1$, $x_2 = 3$, $x_3 = 2$ со значением затрат $\varphi_0 = 32$.

Выбираем первое подмножество с минимальной оценкой. Поскольку полученное решение модифицированной задачи является допустимым для исходной задачи, то оно оптимально.

Рассмотрим на ряде задач построение оценочной задачи и метод ветвей и границ на основе полученной оценки.

3. ЗАДАЧА ЦЕЛОЧИСЛЕННОГО ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Рассмотрим следующую постановку задачи целочисленного линейного программирования. Определить целочисленный неотрицательный вектор $x = \{x_1, x_2, ..., x_4\}$, максимизирующий функцию

$$\varphi(x) = \sum_{i=1}^{n} c_i x_i \qquad (4)$$

при ограничениях

$$\sum_i a_{ij} x_i \leqslant b_j, \quad j = \overline{1, m}. \qquad (5)$$

Для построения оценочной задачи разделим c_i на m частей s_{ij} так, что

$$\sum_j s_{ij} = c_i, \quad i = \overline{1, n}, \qquad (6)$$

и рассмотрим m задач целочисленного линейного программирования следующего вида: определить целочисленный вектор x, максимизирующий функцию

$$s_j(x) = \sum_i s_{ij} x_i \qquad (7)$$

при ограничениях

$$\sum_i a_{ij} x_i \leqslant b_j. \qquad (8)$$

Обозначим через $\Phi_j(s_j)$ оптимальное решение j-й задачи ($s_j = \{s_{ij}\}$). Согласно теореме, величина

$$\Phi(s) = \sum_{j=1}^{m} \Phi_j(s_j) \qquad (9)$$

является оценкой сверху оптимального решения исходной задачи. Окончательно получаем следующую формулировку оценочной задачи: определить значения $\{s_{ij}\}$, $i = \overline{1, n}$, $j = \overline{1, m}$, максимизирующие величину $\Phi(s)$ при ограничениях (6). Оценочную задачу назовем двойственной к исходной задаче целочисленного линейного программирования. Обоснованием этого названия служит следующая интересная связь. Рассмотрим обычную задачу линейного программирования (4), (5) (без требования целочисленности). Для упрощения выводов примем, что все параметры системы ограничений — положительные числа. Заметим, что если не требовать целочисленности решений, то задача (7), (8) легко решается. Ее оптимальное решение:

$$x_{ij} = \begin{cases} b_j/a_{kj}, \text{ если } s_{kj}/a_{kj} = \max\limits_q (s_{qj}/a_{qj}), \\ 0, \text{ если } i \neq k. \end{cases}$$

Оптимальное значение величины $\Phi_j(s_j) = b_j \max\limits_q (s_{qj}/a_{qj})$.

Обозначим $y_j = \max\limits_q s_{qj}/a_{qj}$, $j = \overline{1, m}$. Заметим, что $y_j \geqslant s_{qj}/a_{qj}$ для всех q. Увеличим s_{qj} так, чтобы $s_{qj} = y_j a_{qj}$. Тогда оценочная задача (6), (9) запишется в следующем виде: определить значения $y_j \geqslant 0$, $j = \overline{1, m}$, минимизирующие функцию

$$B(y) = \sum_j b_j y_j$$

при ограничениях

$$\sum_j a_{ij} y_j \geqslant c_i, \quad i = \overline{1, n}.$$

Таким образом, в непрерывном случае оценочная задача становится двойственной задачей линейного программирования.

Рассмотрим на примере применение метода ветвей и границ для решения задачи целочисленного линейного программирования.

Пример. Определить значения $x_i = \{0;\ 1\}$, $i = \overline{1, n}$, максимизирующие функцию

$$\Phi(x) = 10x_1 + 8x_2 + 6x_3 + 7x_4$$

при ограничениях

$$6x_1 + 3x_2 + 2x_3 + 5x_4 \leqslant 11,$$
$$3x_1 + 5x_2 + 6x_3 + 3x_4 \leqslant 11.$$

Для построения оценочных задач возьмем $s_{i1} = a_{i1}$, $s_{i2} = c_i - a_{i1}$, $i = \overline{1, 4}$. Получаем две задачи о ранце.

Задача 1.

$$\varphi_1 = \max(6x_1 + 3x_2 + 2x_3 + 5x_4),$$
$$6x_1 + 3x_2 + 2x_3 + 5x_4 \leqslant 11.$$

Она имеет два решения: $x_1 = x_4 = 1$, $x_2 = x_3 = 0$ и $x_1 = x_2 = x_3 = 1$, $x_4 = 0$. В обоих случаях $\varphi_1 = 11$.

Задача 2.

$$\varphi_2 = \max(4x_1 + 5x_2 + 4x_3 + 2x_4),$$
$$3x_1 + 5x_2 + 6x_3 + 3x_4 \leqslant 11.$$

Ее решение: $x_1 = x_2 = x_4 = 1$, $x_3 = 0$, $\varphi_2 = 11$. Оценка сверху исходной задачи $\varphi_0 = \varphi_1 + \varphi_2 = 22$.

Для улучшения оценки увеличим s_{21} и s_{41} на единицу, уменьшив на единицу s_{22} и s_{42}. Получаем две новые оценочные задачи.

Задача 1.

$$\varphi_1 = \max(6x_1 + 4x_2 + 2x_3 + 6x_4),$$
$$6x_1 + 3x_2 + 2x_3 + 5x_4 \leqslant 11. \qquad (10)$$

Ее решения: $x_1 = x_4 = 1$, $x_2 = x_3 = 0$; $x_1 = x_2 = x_3 = 1$, $x_4 = 0$; $x_2 = x_3 = x_4 = 1$, $x_1 = 0$; $\varphi_1 = 12$.

Задача 2.

$$\varphi_2 = \max(4x_1 + 4x_2 + 4x_3 + 1x_4),$$
$$3x_1 + 5x_2 + 6x_3 + 3x_4 \leqslant 11. \qquad (11)$$

Ее решение: $x_1 = x_2 = x_4 = 1$, $x_3 = 0$; $\varphi_2 = 9$.

Оценка сверху исходной задачи уменьшилась на единицу: $\varphi_0 = \varphi_1 + \varphi_2 = 21$.

Применим метод ветвей и границ. Разобьем множество всех решений на два подмножества. В первом подмножестве $x_1 = 1$, а во втором $x_1 = 0$.

Оценим первое подмножество. Положив в ограничениях (10) и (11) $x_1 = 1$, получим следующие две задачи:

Задача 1.

$$\varphi_1 = \max(4x_2 + 2x_3 + 6x_4),$$
$$3x_2 + 2x_3 + 5x_4 \leqslant 5.$$

Ее решения: $x_4 = 1$, $x_2 = x_3 = 0$; $x_2 = x_3 = 1$, $x_4 = 0$; $\varphi_1 = 6$.

Задача 2.

$$\varphi_2 = \max(4x_2 + 4x_3 + 1x_4),$$
$$5x_2 + 6x_3 + 3x_4 \leqslant 8.$$

Ее решение: $x_2 = x_4 = 1$, $x_3 = 0$; $\varphi_2 = 5$.

Оценка сверху первого подмножества: $\varphi_0 = \varphi_1 + \varphi_2 + c_1 = 21$.

Оценим второе подмножество ($x_1 = 0$). Заметим, что при $x_1 = 0$ любое решение является допустимым для первой оценочной задачи. Поэтому достаточно решить вторую задачу, положив $s_{i2} = c_i$, $i = 2, 3, 4$. Ее решение $x_2 = x_3 = 1$, $x_4 = 0$ является оптимальным во втором подмножестве со значением целевой функции $\varphi_0 = 14$. Выбираем первое подмножество, имеющее большую оценку. Разбиваем первое подмножество на два. В одном из них $x_2 = 1$, а в другом $x_2 = 0$.

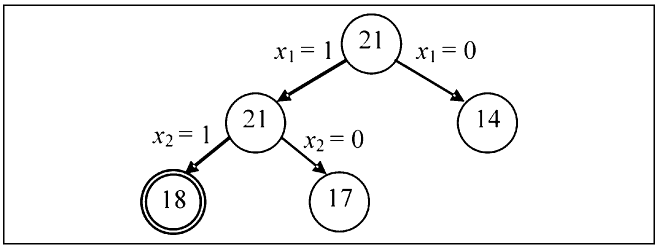


Рис. 7. Дерево ветвлений в задаче целочисленного программирования

Оценим первое подмножество ($x_2 = 1$). Рассматривая два ограничения

$$2x_3 + 5x_4 \leqslant 2,$$
$$6x_3 + 3x_4 \leqslant 3,$$

видим, что единственное решение $x_3 = x_4 = 0$, следовательно, оно оптимально в данном подмножестве со значением целевой функции $\varphi_0 = 18$.

Оценим второе подмножество ($x_2 = 0$). В данном случае достаточно сравнить два варианта: $x_3 = 1$, $x_4 = 0$; $\varphi_1 = 6$ и $x_3 = 0$, $x_4 = 1$; $\varphi_1 = 7$.

Оценка второго подмножества: $\varphi_0 = 10 + 7 = 17$. Ей соответствует оптимальное решение в этом подмножестве $x_1 = x_4 = 1$, $x_2 = x_3 = 0$ со значением целевой функции $\varphi_0 = 17$.

Выбираем первое подмножество, а следовательно, и оптимальное решение $x_1 = x_2 = 1$, $x_3 = x_4 = 0$; $\varphi_0 = 18$. Дерево ветвлений приведено на рис. 7.

Интересно сравнить описанный способ получения оценок для задач целочисленного линейного программирования с известными способами. В основном применяются два способа. В соответствии с первым из них решаются m задач о ранце с каждым ограничением отдельно. Очевидно, что наихудшее решение (по значению целевой функции) определяет оценку сверху исходной задачи. Для нашего примера имеем две задачи о ранце.

Задача 1.

$$\max(10x_1 + 8x_2 + 6x_3 + 7x_4),$$
$$6x_1 + 3x_2 = 2x_3 + 5x_4 \leqslant 11.$$

Ее решение $x_1 = x_2 = x_3 = 1$, $x_4 = 0$; $\varphi_1 = 24$.

Задача 2.

$$\max(10x_1 + 8x_2 + 6x_3 + 7x_4),$$
$$3x_1 + 5x_2 = 6x_3 + x_4 \leqslant 11.$$

Ее решение $x_1 = x_2 = x_4 = 1$, $x_3 = 0$; $\varphi_1 = 25$.

Таким образом, оценка сверху оптимального решения исходной задачи в данном случае равна 24, что больше чем оценка 21, полученная методом сетевого программирования. Более того, поскольку обе рассмотренные задачи о ранце получаются в методе сетевого программирования (первая — при $s_{i1} = c_i$, $s_{i2} = 0$, а вторая, наоборот, при $s_{i2} = c_i$, $s_{i1} = 0$), то можно утверждать, что применение метода сетевого программирования дает лучшие (или такие же) оценки. Согласно второму способу решается задача линейного программирования без требования целочисленности.

Рассмотрим простой

Пример. Определить значения $x_i = \{0; 1\}$, $i = 1, 2$, максимизирующие функцию

$$\varphi(x) = 42x_1 + 14x_2$$

при ограничениях

$$3x_1 + 4x_2 \leqslant 6,$$
$$5x_1 + 2x_2 \leqslant 6.$$

Получим оценку методом сетевого программирования. Для этого достаточно положить $s_{11} = 42$, $s_{21} = 14$, $s_{12} = s_{22} = 0$. Решение первой оценочной задачи $x_1 = 1$, $x_2 = 0$, $\varphi_1 = 42$ является допустимым для второй задачи и, следовательно, оптимально.

Получим оценку, решая нецелочисленную задачу линейного программирования. Ее решение $x_1 = 6/7$; $x_2 = 6/7$; $\varphi_0 = 48 > 42$. Как видим, оценка существенно хуже.

4. РЕШЕНИЕ ЗАДАЧИ О КАМНЯХ МЕТОДОМ СЕТЕВОГО ПРОГРАММИРОВАНИЯ

Рассмотрим постановку задачи о камнях. Имеется n камней разного веса. Требуется разбить их на m групп (куч) так, чтобы максимальный вес камней в группе был минимален. Задача о камнях имеет многочисленные варианты применения (равномерное распределение работ между исполнителями, функций по подразделениям организационной структуры и др.). Дадим ее формальную постановку.

Задача 1. Обозначим через a_i вес i-го камня, $x_{ij} = 1$, если камень i попал в j-ю кучку, $x_{ij} = 0$ — в противном случае. Суммарный вес камней в j-й группе $T_j = \sum_i a_i x_{ij}$.

Максимальный вес группы

$$T = \max_j \sum_i a_i x_{ij} \to \min. \qquad (12)$$

Поскольку каждый камень должен быть помещен только в одну группу, имеем ограничения:

$$\sum_j x_{ij} = 1, \quad i = \overline{1, n}. \qquad (13)$$

Задача заключается в минимизации функции (12) при ограничениях (13). Мы будем рассматривать вспомогательную задачу следующего вида.

Задача 2. Фиксируем допустимый вес каждой группы T и сформулируем следующую задачу: максимизировать сумму весов размещенных в ящики вместимостью T камней:

$$\Phi = \sum_{i,j} a_i x_{ij} \to \max \qquad (14)$$

при ограничениях (13) и

$$\sum_i a_i x_{ij} \leqslant T, \quad j = \overline{1, m}. \qquad (15)$$

Связь между задачами (12), (13) и (14), (15) очевидна. Минимальное значение T, при котором в оптимальном решении задачи 2 размещены все камни, определяет оптимальное решение задачи 1.

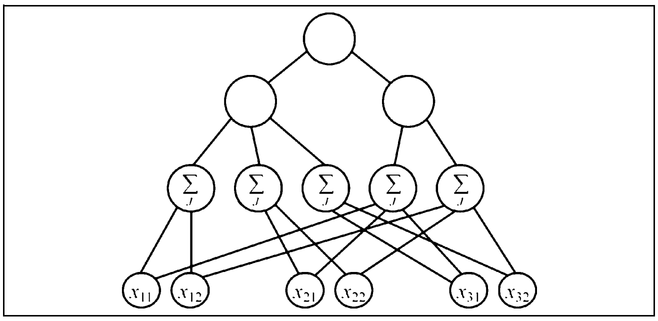


Рис. 8. Сетевое представление ограничений задачи о камнях

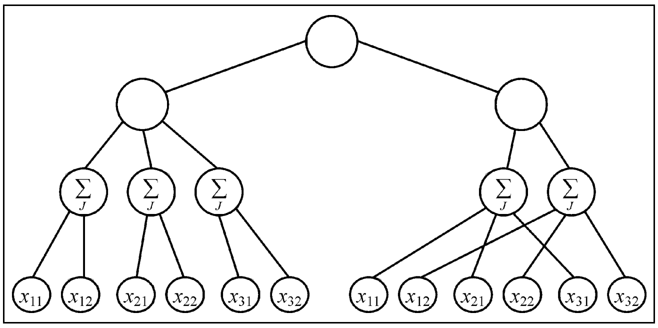


Рис. 9. Сетевое представление оценочной задачи

Сначала получим сетевое представление задачи 2. Оно в агрегированном виде представлено на рис. 8 для случая $n = 3$, $m = 2$.

Поскольку структура сетевого представления имеет вид сети, а не дерева, то для построения оценочной задачи разделяем каждую вершину нижнего уровня на две вершины. Преобразованная структура приведена на рис. 9. Все a_i также делим на две части u_{ij} и v_{ij} для каждой вершины нижнего уровня так, что

$$u_{ij} + v_{ij} = a_i \text{ для всех } i, j. \qquad (16)$$

Рассмотрим следующие две задачи.

Задача 1. Определить значения x_{ij}, максимизирующие функцию $\sum_{i,j} u_{ij} x_{ij}$ при ограничениях (13).

Задача 2. Максимизировать функцию $\sum_{i,j} v_{ij} x_{ij}$ при ограничениях (15).

Обозначим через $S_m(u)$ и $L_m(v)$ оптимальные решения первой и второй задач при заданных величинах u_{ij} и v_{ij}. Оценочная задача заключается в определении значений $\{u_{ij}\}$ и $\{v_{ij}\}$, минимизирующих функцию

$$F(u, v) = S_m(u) + L_m(v) \qquad (17)$$

при ограничениях (16).

Заметим, во-первых, что в оптимальных решениях первой и второй задач можно принять $u_{ij} = y_i$, $v_{ij} = a_i - y_i$, $j = \overline{1, m}$. Во-вторых, решение задачи 1 очевидно: $S_m(y) = \sum_i y_i$. В третьих, решение m задач 2 при заданных величинах $\{y_i\}$ сводится к решению одной задачи о ранце: определить значения $x_i = \overline{0, 1}$, максимизирующие функцию

$$\sum_i x_i(a_i - y_i) \tag{18}$$

при ограничении

$$\sum_i x_i a_i \leqslant T. \tag{19}$$

Решим задачу (18), (19) при $y_i = 0$, $i = \overline{1, n}$.

Обозначим через $Q = \{Q_j\}$ множество векторов x, удовлетворяющих ограничению (19) и упорядоченных по убыванию величин $M_j = \sum_{i \in Q_j} a_i$, $Y_j = \sum_{i \in Q_j} y_j$, а

$$Z = \max_j (M_j - Y_j).$$

Заметим, что при заданных $\{y_i\}$ величина Z определяет оптимальное решение каждой из m задач 2. Оценка (17) при этом имеет вид

$$F(y) = mZ + \sum_i y_i, \tag{20}$$

где $y_i \geqslant 0$ удовлетворяют неравенствам

$$\sum_{i \in Q_j} y_i + Z \geqslant M_j, \quad j = \overline{1, N}, \tag{21}$$

где N — число различных решений неравенства (19). Таким образом, оценочная задача свелась к определению значений $0 \leqslant y_i \leqslant a_i$, $i = \overline{1, n}$ и $0 \leqslant Z \leqslant M_j$, максимизирующих функцию (20) при ограничениях (21). Это обычная задача линейного программирования.

Фиксируем величину Z и определяем максимальный номер k такой, что $Z < M_k$. Рассматриваем следующую задачу линейного программирования: определить значения $0 \leqslant y_i \leqslant a_i$, $i = \overline{1, n}$, минимизирующие функцию

$$Y(Z) = \sum_i y_i,$$

при ограничениях (21), где $j = \overline{1, k}$. Двойственная задача имеет вид: определить значения $u_j \geqslant 0$, $j = \overline{1, k}$, максимизирующие функцию

$$\sum_{j=1}^{k} (M_j - Z)u_j$$

при ограничениях

$$\sum_{j \in R_i} u_j \leqslant 1, \quad i = \overline{1, n},$$

где R_i — множество групп j, содержащих камень i.

Обозначим через $Y_0(Z)$ минимальное значение $Y(Z)$. Оценочная задача сводится к минимизации функции одной переменной $Y_0(Z) + mZ \to \min$.

Берем $T_0 = A/m$, где $A = \sum_i a_i$, и решаем задачу 2. Если $\Phi_{\max}(T_0) < A$, то увеличиваем T_0 до T_1 так, чтобы появился хотя бы один новый вектор Q_j. Если $\Phi_{\max}(T_1) < A$, то продолжаем увеличивать T до тех пор, пока не получим величину T_k такую, что $\Phi_{\max}(T_k) \geqslant A$. Величина T_k является нижней оценкой для задачи 1. Далее можно применить метод ветвей и границ на основе полученной оценки.

Пример. Пусть $m = 3$ и имеется 7 камней весов a_i:

i	1	2	3	4	5	6	7
a_i	10	12	13	14	18	19	22

1 шаг. Имеем $A = 108$, $T_0 = 36$. Имеется только одно решение: $Q = (4, 7)$ со значением $M = 36$.

2 шаг. Увеличиваем T_0 до $T_1 = 37$. Имеются следующие решения: $Q_1 = (5, 6)$, $Q_2 = (4, 7)$, $Q_3 = (1, 2, 3)$, $Q_4 = (3, 7)$. Соответственно, $M_1 = 37$, $M_2 = 36$, $M_3 = 35$, $M_4 = 35$. Выпишем систему неравенств:

$$\begin{aligned}
y_5 + y_6 + z &\geqslant 37\\
y_4 + y_7 + z &\geqslant 36\\
y_1 + y_2 + y_3 + z &\geqslant 35\\
y_3 + y_7 + z &\geqslant 35
\end{aligned}$$

Имеем:

$$Z_0 = 37, \quad y_i = 0, \quad Y_0(37) = 111,$$
$$Z_1 = 36, \quad y_5 = 1, \quad Y_0(36) = 109,$$
$$Z_2 = 35, \quad y_5 = 2, \quad y_4 = 1, \quad Y_0(35) = 108.$$

Нетрудно показать, что дальнейшее уменьшение величины Z не приводит к уменьшению оценки. Поэтому оптимальное значение $Z_0 = Z_2 = 35$.

Берем $x_{51} = x_{61} = 1$, т. е. помещаем камни 5 и 6 в первую группу. Исключая эти камни, рассматриваем задачу меньшей размерности. Имеем для нее также $Y_0(35) = 71 = A - 37$. Получаем оптимальное решение: $x_{12} = x_{22} = x_{32} = 1$, $x_{43} = x_{73} = 1$ со значением $T_{\min} = 37$.

ЗАКЛЮЧЕНИЕ

Предложенный подход к решению задач дискретной оптимизации применим к любой задаче данного класса, поскольку любая функция дискретных переменных допускает сетевое представление. Если целевая функция не аддитивная, то вводом новой переменной всегда можно перейти к задаче с аддитивной целевой функцией, как это показано на примере задачи о камнях. Вопрос об эффективности применения метода сетевого программирования к решению конкретных задач дискретной оптимизации, конечно, требует дальнейших исследований. Заметим, что метод можно применить и к непрерывным задачам.

ЛИТЕРАТУРА

1. *Арнольд В. И.* О функциях трех переменных // Доклады АН СССР. — 1957. — № 2.
2. *Колмогоров А. Н.* О представлении непрерывных функций нескольких переменных суперпозициями непрерывных функций меньшего числа переменных // Доклады АН СССР. — 1956. — № 2.
3. *Бурков В. Н.*, *Буркова И. В.* Задачи дихотомической оптимизации // Материалы междунар. науч.-техн. конф. «Системные проблемы качества, математического моделирования, информационных и электронных технологий». — М., 2003. — С. 23—28.

☎ *(0742) 334-90-51*
E-mail: irbur27@mail.ru

❑

УДК 15:519.876

СТРУКТУРНО-ЦЕЛЕВОЙ АНАЛИЗ РАЗВИТИЯ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИТУАЦИЙ

В. И. Максимов

Институт проблем управления им. В. А. Трапезникова, г. Москва

Рассмотрены модели, предназначенные для анализа и моделирования проблем, возникающих в слабоструктурированных ситуациях. Представлен структурно-целевой анализ динамической модели управляемого развития ситуации. Рассмотрены методы нахождения векторных управлений, обеспечивающих целенаправленное развитие ситуации.

ВВЕДЕНИЕ

На рубеже тысячелетий внешняя среда для сложных социально-экономических объектов — геополитическая, экономическая, социальная и технологическая — все более приобретает свойства нестабильности и неопределенности. Нестабильность проявляется в том, что темпы изменения внешней среды растут, а неопределенность — в том, что возникающие ситуации все чаще становятся неузнаваемыми (совершенно новыми). В таких ситуациях управление развитием отрасли, округа, региона, крупной корпорации резко усложняется, причем прошлый опыт управления, пусть даже успешный, не всегда пригоден для разрешения новых проблемных ситуаций. Это приводит к увеличению вероятности принятия неверных стратегических решений по обеспечению целенаправленного развития социально-экономических ситуаций.

Возникла необходимость перехода от управления на основе прошлого опыта к стратегическому управлению, выявляющему внешние тенденции, риски, опасности и шансы, которые способны не только изменить сложившуюся ситуацию в настоящем, но и вызвать новые направления развития в будущем.

Учет и использование в своих интересах изменений, которые происходят во внешней среде, позволит экономить ограниченные ресурсы на развитие (инвестиции).

Несмотря на значительное число публикаций по вопросам развития социально-экономических объектов (СЭО), проблема целеполагания на предпроектном этапе по-прежнему остается открытой, поскольку основные усилия в исследовании операций и теории принятия решений направлены на достижение уже определенной цели [1].

При анализе слабоструктурированных систем затруднен традиционный эконометрический (социометрический и т. п.) подход к анализу процессов для выработки комплексных (т. е. затрагивающих различные аспекты исследуемой системы) решений.

Предлагаемый подход ориентирован на качественный анализ сложных ситуаций, интерпретируемых как слабоструктурированные системы, характеризующиеся отсутствием точной количественной информации о происходящих в них процессах. Число переменных в таких ситуациях может измеряться десятками, и все они вплетены в паутину причин и следствий. Увидеть и осознать логику развития событий на таком многофакторном поле крайне трудно, и в то же время непрерывно приходится принимать решения о выборе тех или иных мер, способствующих развитию ситуации в нужном направлении.

Качественный анализ сложной ситуации предусматривает определение тенденций протекающих процессов, качественную оценку этих тенденций и выбор мер, способствующих их развитию в нужном направлении [2—5].

1. СТРУКТУРИЗАЦИЯ И СТАТИЧЕСКИЕ МОДЕЛИ СИТУАЦИЙ

Понятие ситуации включает в себя различные по природе и характеру процессы (социальные, экономические, политические и др.), характеризующие взаимодействие внутренней и внешней сред СЭО. Сложность выработки решений обусловлена тем, что эти процессы взаимосвязаны друг с другом, и для прогнозирования последствий решения необходимо учитывать эту сложную структуру взаимосвязей процессов. Положение осложняется тем, что, как правило, *достоверная количественная информация* об этих процессах отсутствует, и о них можно судить только по косвенным характеристикам. Это затрудняет построение количественных моделей развития процессов для принятия решений. Выход состоит в обращении к качественным моделям ситуации.

Возникающие проблемные ситуации (проблемы) можно разделить на два типа.

- *Хорошо структурированные*, для которых характерна четкая структура, определенные данные, известные источники информации, небольшие затраты на сбор информации.
- *Слабоструктурированные.* Во многих проблемных областях невозможно создание традиционных количественных формальных моделей. Для задач подобного типа характерны неопределенность, описание на качественном уровне, неоднозначность последствий решений поставленных проблем. Информационные технологии, обеспечивающие процесс передачи ЭВМ информации о таких проблемах, т. е. знаний и их обработке, носят название когнитивных (cognitive), иначе — познавательных.

Структуризация или концептуализация знаний характерна тем, что *разрабатывается структура полученных знаний* о предметной области, т. е. определяется список *основных понятий* о предметной области, выявляются отношения между понятиями, определяются связи данной предметной области с окружающим миром. Происходит разработка *неформального описания знаний* о предметной области, которую можно наглядно изобразить в виде графа, таблицы, текста и т. д. При формализации знаний аналитик-когнитолог выбирает один из способов формализации знаний, адекватный его представлению о предметной области (см. цветную вклейку).

Этап получения знаний имеет свои особенности, которые заключаются в том, что его можно разделить на более тонкие процессы (извлечения, приобретения, формирования), имеющие свою собственную специфику.

В процессе *извлечения (elicitation) знаний* происходит взаимодействие с экспертом — источником знаний, при котором становится ясным способ рассуждения специалистов при поддержке принятия решений и структура его представлений о предметной области. Процесс извлечения — это процедура, в которой аналитик, имеющий опыт в области системного анализа, математической логики и когнитивной психологии, создает *«скелетную» модель предметной области*, которая на последующих этапах будет наполнена конкретными сведениями об объектах предметной области.

Разработанная нами методология синтезирует *системный и когнитивный подходы.*

Цель когнитивной структуризации состоит в формировании и уточнении *гипотезы о функционировании исследуемого объекта*, рассматриваемого как сложная система, которая состоит из отдельных внутренних и внешних элементов, подсистем, взаимодействующих друг с другом, на основе структурной схемы причинно-следственных связей.

Содержательные выводы о тенденциях развития ситуации можно получить, пользуясь качественными оценками, сформулированными на концептуальном уровне.

В настоящей работе рассмотрены модели для решения слабоструктурированных проблем, построенные на основе *когнитивной карты ситуации.* Когнитивная карта представляет собой квадратную таблицу, в которой:

— строки и столбцы взаимно однозначно соответствуют базисным факторам, в терминах которых описывается ситуация;

— элемент, стоящий на пересечении i-й строки и j-го столбца, отражает факт непосредственного влияния i-го фактора на j-й фактор. Знак этого элемента отображает знак влияния (положительный или отрицательный), а модуль — силу такого влияния в соответствующей шкале.

Когнитивная карта является исходным статическим представлением (отображением) связей между факторами, существующими в исследуемой ситуации. Разрешение проблем целенаправленного развития, возникающих в слабоструктурированных ситуациях, требует построения динамической имитационной модели и получения на ее основе новых знаний о структуре и динамике исследуемой ситуации.

Модель сложной ситуации строится следующим образом.

- Для описания качественных значений базисных факторов выбирается набор соответствующих лингвистических переменных. Выбор градаций по значениям лингвистических переменных позволяет дать необходимую степень детализации — «слабо—средне—сильно» или, более подробно, «очень слабо—слабо—средне—сильно—очень сильно» и т. д. Каждой лингвистической переменной соответствует определенное число в шкале [0, 1], являющееся числовым эквивалентом этой переменной. Эти числовые эквиваленты назовем *качественными переменными.*
- На основе знаний об исследуемой ситуации составляется уравнение для каждого базисного фактора, включающее в себя набор всех факторов, непосредственно влияющих на этот фактор (этот набор выбирается по когнитивной карте ситуации) вместе с функциональными характеристиками соответствующих влияний.

Моделирование основано на сценарном подходе [2, 3, 6] и представляет собой процесс передачи управлений по вершинам графа модели.

Сценарий — набор тенденций, характеризующих саморазвитие ситуации в начальный момент, векторы целей развития, векторы управлений (управляющих воздействий); они характеризуют комплексы мероприятий, влияющих на развитие ситуации, и системы наблюдаемых параметров (факторов), иллюстрирующих тенденции результирующих процессов в ситуации.

Модель может исследоваться по трем основным направлениям:

- *прогноз саморазвития ситуации* (без внешнего воздействия на процессы в ситуации — ситуация развивается сама по себе);
- *прогноз развития ситуации с выбранным вектором управлений* (прямая задача);
- *синтез вектора управлений* для достижения необходимого направления развития ситуации (обратная задача).

Моделирование включает в себя определенную последовательность взаимосвязанных этапов:

— определение начальных условий, тенденций, характеризующих развитие ситуации на начальном этапе;

— задание целевых направлений (увеличение, уменьшение) и силы (слабо, сильно) изменения тенденций процессов в ситуации;

— выбор набора управляющих факторов, определение их силы и направленности воздействия на ситуацию;

— выбор вектора управлений (факторов) ситуацией, силу и направленность которых необходимо определить;

— выбор наблюдаемых факторов (индикаторов), характеризующих результирующее развитие ситуации, осуществляется в зависимости от целей анализа и желания пользователя;

— моделирование по одному из перечисленных ранее трёх направлений.

2. ИМИТАЦИОННЫЕ ДИНАМИЧЕСКИЕ КОГНИТИВНЫЕ МОДЕЛИ САМОРАЗВИТИЯ СИТУАЦИИ

С когнитивной картой ассоциируется графовая модель ситуации [7], анализ которой позволяет выявить структурные свойства ситуации. В основе модели лежит взвешенный орграф $G = (X, A)$, в котором X — множество вершин, взаимно однозначно соответствующее множеству базисных факторов, A — множество дуг, отражающих факт непосредственного влияния факторов. Каждая дуга, связывающая некоторый фактор x_i с некоторым фактором x_j, имеет вес a_{ij}, знак которого говорит о знаке влияния фактора x_i на фактор x_j, а модуль величины a_{ij} — о силе этого влияния. Таким образом, когнитивную карту можно рассматривать как матрицу смежности A_g графа $G = (X, A)$ [3, 8].

Пусть задан вектор начальных тенденций всех факторов $x(t) = (x_1(t), ..., x_n(t), x_{n+1}(t), ..., x_N(t))$ в момент $t = 0$. Динамика изменений его компонент во времени определяется видом модели, выбираемой для описания взаимодействия непосредственных влияний, одновременно приходящих на каждый базисный фактор от «соседних» факторов.

Модель саморазвития ситуации для приращений тенденций изменения факторов предназначена для исследования динамики *изменения тенденций* развития ситуаций, т. е. ее переменные отражают «приращение тенденций изменения фактора», численные значения которых определяются в качественной шкале $[-1, 1]$. В её основу положена гипотеза о том, что текущее приращение каждого фактора является взвешенной суммой текущих приращений воздействующих на него «соседних» факторов. Элементарный такт квантования по времени в модели равен 1. Обозначим через $p_i(t)$ значение (в шкале $[-1, 1]$) приращения i-го базисного фактора в момент t.

Модель для приращений тенденций изменения факторов имеет вид:

$$p_i(t+1) = \sum_{j \in I_i} a_{ij} p_j(t), \quad i = 1, ..., N, \tag{1}$$

где I_i — множество номеров факторов, от которых идут дуги (непосредственные влияния) к фактору x_i; a_{ij} — элементы матрицы $A = A_g^T$ (веса соответствующих дуг); N — количество факторов в когнитивной карте — модели ситуации; T — знак транспонирования.

В матричной форме *модель саморазвития ситуации* (1) выглядит так: $p(t + 1) = Ap(t)$, где $p(t) = (p_1(t), ..., p_N(t))^T$.

Модель саморазвития ситуации для значений тенденций изменения факторов получается из модели (1) путем представления приращения $p_i(t)$ как разности значений фактора x_i в соседние моменты времени: $p_i(t) = \Delta x_i(t) = x_i(t) - x_i(t - 1)$.

Переход от приращений факторов $p_i(t)$ к значениям факторов $x_i(t)$ позволяет исследовать значение любого фактора модели в различные моменты времени.

Тогда, учитывая выражение (1), значение фактора x_i в момент времени $(t + 1)$ определим как

$$x_i(t+1) = x_i(t) + \sum_{j \in I_i} a_{ij}(x_j(t) - x_j(t-1)), \quad i = 1, ..., N, \tag{2}$$

или в матричной форме

$$x(t+1) = (E_N + A)x(t) - Ax(t-1), \tag{3}$$

где $x(t) = (x_1(t), ..., x_N(t))^T$, E_N — единичная матрица порядка N.

Модифицированная модель саморазвития ситуации для значений тенденций изменения факторов отличается от модели (2) иным определением характера взаимодействия одновременных влияний «соседних» факторов на каждый фактор: алгебраическая сумма приращений факторов, входящая в выражение (2), заменяется средним арифметическим таких приращений, т. е.

$$x_i(t+1) = x_i(t) + \frac{1}{|I_i|} \sum_{j \in I_i} a_{ij}(x_j(t) - x_j(t-1)), \quad i = 1, ..., N,$$

где $|I_i|$ — мощность множества I_i.

Обобщенная модель саморазвития ситуации для значений тенденций изменения факторов. Описанные выше модели для значений факторов являются линейными, со всеми их достоинствами и недостатками. При более глубоком изучении ситуации может потребоваться учет того, что характер и сила влияний факторов друг на друга не постоянные, веса дуг могут менять знаки, а сила влияния — зависеть от текущих значений фактора. Учет подобных обстоятельств приводит к тому, что когнитивная карта ситуации превращается в функциональный граф, вес каждой дуги в котором есть, в общем случае, произвольная функция от текущих значений инцидентных ей факторов-вершин.

Одна из таких нелинейных моделей описывается следующими уравнениями:

$$x_i(t+1) = \frac{1}{\left|I_1^+(t)\right|} \sum_{j \in I_i^+(t)} a_{ij} x_j(t) + \frac{1}{\left|I_1^-(t)\right|} \sum_{j \in I_i^-(t)} a_{ij} x_j(t), \quad i = 1, ..., N. \tag{4}$$

Здесь $I_i^+(t)$, $I_i^-(t)$ — множества номеров факторов, от которых поступают на фактор x_i соответственно положительные и отрицательные непосредственные воздействия в момент t. Согласно уравнениям (4), результирующее значение представляет собой сумму средних арифметических в момент t как по положительным, так и по отрицательным воздействиям.

В силу нелинейности модель (4) обладает большим «богатством» фазового портрета, позволяющим исследовать ряд эффектов, свойственных сложным ситуациям.

В процессе развития модельной ситуации во времени на каждый фактор помимо непосредственных влияний от «соседних» факторов приходят также влияния от более «отдаленных» факторов, эти *опосредованные влияния* передаются через цепочки соответствующих факторов и соединяющих их дуг графа. Множество влияний, как непосредственных, так и опосредованных, которым подвержен каждый фактор в ситуации, описывается с помощью понятия *транзитивного замыкания* когнитивной карты ситуации [8], определяемого как сумма бесконечного ряда $E_N + A + A^2 + ... + A^t + ...$ по степеням матрицы A. Каждый элемент этого ряда характеризует прохождение путей длины t в графе, т. е. осуществление непосредственных, опосредованных через один фактор, через два фактора взаимовлияний.

Оценку суммы этого ряда можно получить только при условии *устойчивости матрицы смежности* A_g графа G. Тогда все элементы этого ряда стремятся к конечным пределам при неограниченном возрастании длины t.

Для нахождения *транзитивного замыкания* достаточно рассмотреть N членов в ряду по степеням матрицы A, где N — порядок матрицы A, т. е. число базисных факторов в когнитивной карте ситуации. Тогда транзитивное замыкание матрицы A оценивается матрицей:

$$Q = E_N + A + A^2 + ... + A^N \cong (E_N - A)^{-1}. \qquad (5)$$

Отсюда вытекает необходимость стабилизации графа G формального описания ситуации. Под *стабилизацией* графа G (или, что то же самое, под стабилизацией когнитивной карты ситуации) будем понимать такую «нормировку» матрицы A_g когнитивной карты, при которой все собственные значения результирующей матрицы содержатся внутри окружности единичного радиуса на комплексной плоскости. Это условие является необходимым и достаточным условием устойчивости линейной системы (т. е. ограниченности диапазонов значений её переменных), работающей в дискретном времени [3, 8].

Если исходная матрица A_g когнитивной карты неустойчива (т. е. некоторые из её собственных значений выходят из окружности единичного радиуса), то варианты её «нормировки» могут состоять в следующем.

Исходная матрица A_g умножается на так называемый стабилизирующий множитель $0 < k_{ст} < 1$, такой, что матрица $k_{ст}A_g$ будет устойчивой. При этом каждое собственное значение матрицы $k_{ст}A_g$ равно соответствующему собственному значению матрицы A_g, умноженному на $k_{ст}$. Согласно теории линейных систем время переходного процесса модели (т. е. время от подачи начальных условий на модель до момента перехода модели в установившееся состояние) тем меньше, чем меньше собственные значения матрицы $k_{ст}A_g$ модели, а следовательно — чем меньше величина $k_{ст}$. При малом $k_{ст}$ модель довольно быстро переходит в установившееся состояние и её динамика (временные изменения факторов) слабо выражена ввиду кратковременности переходного процесса.

Другой вариант стабилизации матрицы A_g состоит в делении каждого i-го столбца (или строки) матрицы A_g на число $(s_i + \varepsilon)$, где s_i — число ненулевых элементов i-го столбца (или строки), а ε — некоторое малое число, необходимость которого вызвана лишь наличием строк (или столбцов) в матрице A_g с одним ненулевым элементом, так что без числа ε результат деления такой строки (столбца) на s_i был бы равен 1, что недопустимо для выполнения условия устойчивости матрицы A_g.

Таким образом, транзитивное замыкание когнитивной карты описывает интегральные (т. е. непосредственные и опосредованные) влияния изменения одних факторов на изменения других факторов (т.е. связывает приращения факторов).

При этом (i, j)-й элемент матрицы Q транзитивного замыкания связывает знак исходного приращения Δx_i фактора x_i со знаком интегрального приращения Δx_j фактора x_j.

Другими словами, если в ситуации возмущается (получает приращение) только фактор x_i, то с учетом всех опосредованных влияний фактора x_i на фактор x_j знак результирующего приращения фактора x_j определяется как $\mathrm{sign}\Delta x_j = \mathrm{sign} q_{ij} \cdot \mathrm{sign}\Delta x_j$, где q_{ij} — элемент матрицы Q [8].

3. СТРУКТУРНО-ЦЕЛЕВОЙ АНАЛИЗ

Когнитивные имитационные модели, рассмотренные в § 2, предназначены для анализа и моделирования проблем, возникающих в слабоструктурированных ситуациях. Для разрешения проблемной ситуации необходимо понимание её структурных свойств (переходящих в структурные свойства модельного отображения проблемы) и динамических характеристик ситуации и самой проблемы, определяющих пути её разрешения.

Эти вопросы рассматриваются на примере модели для значений тенденций изменения факторов, определенной уравнением (3).

Из уравнения (3) следует (с учетом, что $X(t) = 0$ для $t < 0$)

$$X(t) = (E_N + A + A^2 + ... + A^t)X(0). \qquad (6)$$

Если управление развитием ситуации понимать как управление начальным состоянием $X(0)$, то при $t \geqslant N$ уравнение (6) при наличии управления принимает вид (с учетом выражения (5))

$$X(t) = QX(0) + QBU(0), \qquad (7)$$

где $U(0)$ — вектор управлений, B — (0, 1)-матрица, ненулевые элементы которой указывают на номера корректируемых координат начального состояния $X(0)$.

Из уравнения (7) следует, что влияния начальных условий и управлений на текущее состояние модели при больших t ($t \geqslant N$) опосредуются через матрицу Q транзитивного замыкания и зависят от её структурных свойств, которые рассматриваются в п. 3.2. Процедуры выбора управлений, обеспечивающих целенаправленное развитие ситуаций, также основанные на уравнении (7), описываются в п. 3.3.

3.1. Целеполагание при управляемом развитии ситуации

Гипотеза целеполагания [9—11]: лицо, принимающее решение, ЛПР (или аналитик), может указать, какое направление изменения фактора (наиболее значимых для него факторов) он рассматривает как желательное (благоприятное). Желательное направление изменения фактора x_i определяется показателем (оценкой) r_i, принимающим значение +1, если желательно увеличение значения фактора x_i (положительная динамика фактора), и −1, если желательно уменьшение фактора x_i (отрицательная динамика); $r_i = 0$, если затруднительно указать желательную динамику по фактору x_i.

Показатель r_i называется *оценкой динамики фактора* (ОДФ) x_i.

Цель развития ситуации описывается подмножеством целевых факторов когнитивной карты модели. Это означает, что вектор целей развития ситуации есть либо вектор значений целевых факторов (фиксированная цель), либо вектор направления изменения этих значений (нефиксированная цель).

Определение 1. *Фиксированная векторная цель включает в себя вектор Y^* целей вместе с заданными значениями тенденций изменения каждой из целей y_i^*.*

Фиксированная векторная цель — это точка в m-мерном пространстве тенденций целей, представляющая для пользователей вектор некоторых «идеальных» значений тенденций изменения целевых факторов.

Определение 2. *Нефиксированная векторная цель включает в себя вектор Y^* целей вместе с указанием благоприятных тенденций изменения его координат в соответствии с их оценками изменения факторов.*

Вектор благоприятных тенденций представляет собой вектор интересов ЛПР (аналитика). При этом на значения благоприятных изменений целевых факторов не накладывается ограничений (чем больше — тем лучше).

3.2. Структурно-целевой анализ управляемого развития ситуации

Анализ координат вектора целей на непротиворечивость. Обозначим через $Y = \{y_1, ..., y_m\}$ множество целевых факторов. Пусть $r(Y)$ — вектор желательной динамики целевых факторов (ОДФ).

Определение 3. *Вектор целей Y называется непротиворечивым, если*

$$r_i r_k = \operatorname{sign} q_{ik},\ \forall y_i,\ y_k \in Y, \tag{8}$$

где q_{ik} — элемент матрицы Q.

При выполнении условия (8) целевые факторы y_i и y_k согласованы, в противном случае — не согласованы.

Суть задания непротиворечивого вектора целей состоит в том, чтобы обеспечение желательного интегрального изменения одних целевых факторов не приводило бы к нежелательным интегральным изменениям других целевых факторов.

Анализ координат вектора управлений на согласованность с вектором целей. Для управляющих факторов понятие ОДФ неприменимо, поскольку аналитик изначально (до проведения структурно-целевого анализа) «не знает», как скажется изменение того или иного управляющего фактора на интегральном поведении целевых факторов.

Определение 4. *Вектор управлений согласован с вектором целей Y, если для каждой координаты вектора управлений $U = (u_1, ..., u_p)$ можно указать такой знак, что для результирующего знакового вектора* $\operatorname{sign} U$

$$r_s = \operatorname{sign} q_{st} \operatorname{sign} u_t,\ \ \forall u_t \in U, \forall y_s \in Y. \tag{9}$$

Суть согласованности управляющих факторов с вектором целей состоит в том, что при выполнении условия (9) всякое изменение управляющих факторов в соответствии с вектором $\operatorname{sign} U$ не вызовет изменения никакой координаты вектора целей Y в нежелательном направлении. Обозначим через $U^*(0)$ вектор управлений, знаки которых выбраны в соответствии с условием (9), а через $|U^*(0)|$ — вектор $U^*(0)$, в котором все координаты заменены их абсолютными значениями. Введенные понятия позволяют сформулировать следующее

Утверждение. *Если выбранный вектор целей Y непротиворечив и множество управляющих факторов согласовано с вектором целей, то возможен такой выбор вектора управлений U, для которого*

$$|U_1^*(0)| \leqslant |U_2^*(0)| \to Y(U_1^*(0)) \leqslant Y(U_2^*(0)), \tag{10}$$

где $Y(U_i^*(0))$ — вектор изменений целевых факторов, обусловленных подачей вектора $U_i^*(0)$, $i = (1, ..., m)$, управлений, т. е. свойство «доминирования» по модулям управлений переходит в свойство «доминирования» по результатам их воздействия на целевые факторы.

Другими словами, более «интенсивное» управление (с большими абсолютными значениями координат) вызовет более «интенсивные» изменения координат вектора целей в желательных направлениях.

Введенные определения полезны при анализе и моделировании ситуаций. Так, нарушение условий непротиворечивости выбранного вектора целей может помочь аналитику разобраться во взаимодействии целевых факторов и более «правильно», сообразуясь с ситуацией, задать вектор целей. Анализ вектора выбранных управляющих факторов на согласованность с вектором целей позволит отказаться от противоречивых управлений и, наоборот, активнее использовать «выигрышные» управляющие факторы, изменения которых в соответствии с подаваемыми на них управлениями приведут к большим благоприятным изменениям целевых факторов.

Анализ эффективности интегрального воздействия управляющих факторов на целевые факторы. Суть подачи управлений на ситуацию состоит в таком изменении управляющих факторов, чтобы их влияния на целевые факторы привели к благоприятным изменениям целевых факторов, т. е. к изменениям целевых факторов в направлении их ОДФ.

В связи с этим важно ответить на вопрос — какие из управляющих факторов являются более «действенными» по их интегральному влиянию на целевые факторы для получения положительного эффекта? Для ответа на этот вопрос полезно рассмотреть матрицу Q транзитивного замыкания когнитивной карты исследуемой ситуации и ОДФ целевых факторов.

Формально показатель эффективности $E(u_k)$ управляющего фактора u_k (т. е. максимальный положительный эффект от изменения фактора u_k) определяется как абсолютное значение суммы коэффициентов влияния данного управляющего фактора u_k на целевые факторы, умноженных на ОДФ целевых факторов, т. е.

$$E(u_k) = \left|\sum_{i=1}^{m} r_i q_{ki}\right|,$$

где r_i — ОДФ целевого фактора y_i, q_{ki} — элемент матрицы Q.

Действительно, максимальный положительный эффект Δy от подачи управления g_k на фактор u_k оценивается как

$$\Delta y = g_k \sum_{i=1}^{m} r_i q_{ik},$$

где знак воздействия g_k выбирается совпадающим со знаком суммы $\sum_{i=1}^{m} r_i q_{ik}$, а его значение равно 1.

Итак, структурно-целевой анализ динамической модели управляемого развития ситуации состоит из следующих этапов.

- *Анализ целей* (координат вектора целей) *на взаимную непротиворечивость* для ответа на вопрос — не является ли вектор целей (фиксированный или нефиксированный) противоречивым, т. е. не получится ли так, что достижение какой-либо из целей (координат в векторе заданных целей) будет препятствовать достижению других целей?
- *Проверка согласованности множества управляющих факторов с заданным вектором целей*, т. е. не получится ли так, что изменение значения какого-либо управляющего фактора (с помощью соответствующего управления) будет способствовать достижению некоторых подцелей в векторе цели и в то же время препятствовать достижению других подцелей вектора целей.
- *Оценка эффективности воздействия управляющих факторов* на все координаты вектора целей. Такая оценка полезна при выборе наиболее эффективных управляющих факторов, изменения которых с помощью выбираемых управлений обеспечат целенаправленное развитие ситуации.

Далее рассмотрим различные процедуры выбора управлений, обеспечивающих достижение фиксированных или нефиксированных целей.

3.3. Процедура нахождения вектора управлений для достижения заданной фиксированной цели

Прежде чем решать задачу поиска управлений (воздействий), обеспечивающих достижение фиксированной цели, необходимо проанализировать *саморазвитие ситуации* (без подачи управлений) из заданного начального состояния $X(0)$. Структура модели и начальное условие $X(0)$ могут оказаться такими, что её состояние при саморазвитии (лишь из-за взаимодействия факторов) будет смещаться «в сторону цели», что можно считать «*попутным ветром*» ситуации при достижении цели.

Динамика саморазвития состояния $x(t)$ модели описывается уравнением (3). Если текущие значения целевых факторов в векторе $x(t)$ удаляются от целевых значений, то считается, что состояние $x(t)$ удаляется от цели (что можно считать «*встречным ветром*» ситуации при достижении цели) и *необходима коррекция* развития процессов в модели с помощью тех или иных управлений.

Нахождение вектора управлений с помощью импульсных управлений, подаваемых в начальный момент времени. Фиксированная цель Y^* считается достигнутой из начального состояния $X(0)$, если установившиеся значения целевых факторов совпадают с их заданными значениями [3, 5, 8]. Установившийся целевой вектор $Y_{\text{уст}}$ определяется через матрицу транзитивного замыкания Q и начальное состояние $X(0)$ модели следующим образом:

$$Y_{\text{уст}} = CQX(0) + CQBU, \qquad (11)$$

где $CQX(0)$ — составляющая целевого вектора, обусловленная саморазвитием ситуации; $CQBU$ — составляющая целевого вектора, обусловленная подачей управлений; ненулевые элементы матриц C и B выделяют, соответственно, целевые и управляющие факторы в матрице A исходной когнитивной карты.

Фиксированная цель Y^* считается достигнутой, если $Y_{\text{уст}} = Y^*$, что с учетом выражения (10) дает $Y_{\text{уст}} = CQX(0) + CQBU = Y^*$.

Отсюда получаем уравнение для нахождения искомого вектора U управлений:

$$CQBU = Y^* - CQX(0). \qquad (12)$$

Вектор находится из уравнения (12) методом наименьших квадратов (МНК) [12, 13] и в случае, когда управляющих факторов больше, чем целевых (т. е. в случае $p > m$), МНК-решение имеет вид:

$$U^* = (CQB)^+(Y^* - CQX(0)) + (E_N - (CQB)^+CQB)h,$$

где $(CQB)^+$ — матрица, псевдообратная [13] матрице CQB; h — произвольный вектор размера p.

Таким образом, в случае $p > m$ (недоопределенная система уравнений (12)) имеется бесконечное множество допустимых МНК-решений, каждое из которых формально приводит к достижению цели [13]. Из множества этих решений выбирается так называемое нормальное решение вида

$$U^* = (CQB)^+(Y^* - CQX(0)), \qquad (13)$$

обладающее минимальной нормой среди всех решений вида (12).

В случае переопределенной ($p < m$) системы (12) существует единственное МНК-решение вида (13).

Замечание. МНК-процедура достижения фиксированной цели, минимизирующая сумму квадратов отклонений установившихся значений целевых факторов от соответствующих заданных значений, обладает двумя недостатками:

— координаты формального МНК-решения (13) могут выходить за пределы интервала $[-1, 1]$, определяющего допустимые области значений для управлений (тогда в программе вычислений они приравниваются к ближайшим граничным точкам интервала $[-1, 1]$);

— может оказаться, что для приближения текущих значений целевых факторов к заданным целевым значениям (т. е. чтобы попасть в «точку»), некоторые управления, согласно МНК-решению (13), следует изменять в направлениях, противоречащих предметным соображениям, вытекающим из исследуемой ситуации (подобные парадоксы исчезают при переходе к нефиксированным целям).

Нахождение вектора управлений с изменяющимися во времени координатами. В данном случае налагается требование, чтобы фиксированная цель Y^* была достигнута в заданный момент t^*. На основе этого требования определяется соответствующее матричное уравнение относительно векторов управлений $U(0)$, $U(1)$, $U(2)$, ..., $U(t^*)$, из которого искомая последовательность векторов управлений также находится как нормальное МНК-решение.

Нахождение вектора управлений путем покоординатного рассмотрения вектора целей. В данном методе каждая координата $y^*_{i_k}$ фиксированной цели Y^* рассматривается отдельно как скалярная подцель. Составляется уравнение, аналогичное уравнению (11), для вектора управлений, обеспечивающего достижение скалярной подцели $y^*_{i_k}$:

$$C_{i_k}(E_N - A)^{-1}Bg = y^*_{i_k} - C_{i_k}(E_N - A)^{-1}X(0), \qquad (14)$$

где C_{i_k} — i_k-я строка матрицы C.

Пусть g^k — нормальное МНК-решение уравнения (14). Рассматривается выпуклая оболочка (многогранник) U векторов g^1, ..., g^m вида

$$U = \alpha_1 g^1 + \alpha_2 g^2 + ... + \alpha_m g^m.$$

Показано [12], что коэффициенты α_i этой выпуклой оболочки, при которых вектор U обеспечивает максимальное приближение (в среднеквадратичном смысле) целевых координат установившегося состояния к фиксированной цели, могут быть найдены методом наименьших квадратов.

3.4. Процедура нахождения вектора управлений (решение обратной задачи) для достижения заданной нефиксированной цели

Пусть для установившихся значений целевых факторов справедливо соотношение (11).

Свойство аддитивности этого соотношения, т. е. независимости вкладов в изменения целевых факторов, обусловленные начальным состоянием (см. первое слагаемое в формуле (11)) и подаваемыми управляющими воздействиями (см. второе слагаемое), позволяет анализировать эти слагаемые раздельно [3]. Остановимся вначале на втором слагаемом.

Приращение целевого фактора y_i, обусловленное подачей вектора управлений U, представляется в виде

$$\Delta y_i = (CQBU)_i = \sum_{k=1}^{p} q_{ik}u_k, \qquad (15)$$

где q_{ik} — элементы матрицы Q. Благоприятное изменение вектора y_i означает, что оно происходит в направлении ОДФ r_i этого вектора. Тогда задача выбора вектора U состоит в том, чтобы максимизировать сумму

$$\sum_{i=1}^{m} r_i \Delta y_i \Rightarrow \max,$$

что с учетом выражения (15) запишется как

$$\sum_{i=1}^{m} r_i \sum_{k=1}^{p} \Delta q_{ik}u_k \Rightarrow \max.$$

После перегруппировки слагаемых исходная задача по выбору значений управлений представляется в виде:

$$\sum_{k=1}^{p}\left(\sum_{i=1}^{m} r_i q_{ik}\right)u_k \Rightarrow \max \qquad (16)$$

при условиях

$$-1 \leqslant u_k \leqslant 1, \qquad (17)$$

$$\sum_{k=1}^{p} U_k \leqslant P, \qquad (18)$$

где P — ограничения на число одновременно выбранных управлений.

Задача(16)—(18) решается методами линейного программирования. Её решение обозначим как $U^* = (u^*_1, u^*_2, ..., u^*_p)$.

Пусть $X(0)$ — вектор начальных тенденций в ситуации. Результирующий вектор Y^* наиболее благоприятных значений целевых факторов, достижимых из начального состояния $X(0)$ при использовании выбранного множества U управлений определяется как

$$Y^*_{\text{уст}} = CQX(0) + COQU^*.$$

Значения координат вектора $Y^*_{\text{уст}}$ — это наилучшие целевые значения, достижимые при начальном состоянии $X(0)$ и выбранном векторе управлений U таком, при котором достигается максимум суммы $\sum_{i=1}^{m} r_i \Delta y_i$, т. е. максимум положительного эффекта от применения выбранного вектора управлений U.

3.5. Предотвращение чрезвычайных ситуаций

Максимально положительный эффект (в плане благоприятного изменения всех целевых факторов) от применения возможных управлений определяется как

$$E = \sum_{i=1}^{m} r_i \sum_{k=1}^{p} q_{ik}u^*_k,$$

где r_i — ОДФ i-го целевого фактора, q_{ik} — элементы матрицы Q транзитивного замыкания исходной матрицы A когнитивной карты, $(u^*_1, u^*_2, ..., u^*_p)$ — решение задачи (16)—(18) линейного программирования.

При саморазвитии ситуации из состояния $X(0)$ установившееся состояние (при отсутствии управлений) определяется как $X^*_{\text{сам}} = Q\,X(0)$, а установившиеся значения целевых факторов — как $Y^*_{\text{сам}} = CQX(0)$.

Пусть S — интегральная оценка достижения цели развития ситуации:

$$S = \sum_{i=1}^{m} r_i y^*_{i\,\text{сам}} = R(Y)\,Y^*_{\text{сам}},$$

где $R(Y)$ — вектор ОДФ целевых факторов.

Состояние $Y^*_{\text{сам}}$ назовем чрезвычайной ситуацией (ЧС), если

— интегральная оценка достижения цели развития ситуации S отрицательна, т. е. саморазвитие ситуации не способствует достижению нефиксированной цели;

— никакой из векторов управлений в рамках ограничений R, подаваемых в момент $t = 0$, не позволяет добиться положительного эффекта в изменении целевых факторов.

Другими словами, $X^*_{\text{сам}}$ — чрезвычайная ситуация, если $S < 0$, $|S| > |E|$.

Состояние $X(0)$ в этом случае назовем *предвестником ЧС* $X^*_{\text{сам}}$.

В рамках модели предотвращение ЧС заключается в том, что помимо стремления к достижению цели необходимо следить за текущим состоянием модели. Если такое состояние становится предвестником ЧС, то для её предотвращения необходимо ослабить ресурсные ограничения и максимально усилить управления, препятствующие наступлению ЧС. Эти вопросы решаются путем предварительного моделирования саморазвития ситуации, а также рассмотрения сценариев с различными векторами управлений.

ЗАКЛЮЧЕНИЕ

Сформированные положения позволяют на формализованном уровне определить последовательность процедур по структуризации ситуации с учетом влияния вы-

Основные этапы технологии анализа и моделирования стратегического развития СЭО

Структуризация знаний о СЭО и внешней для него среды на основе PEST[1]-анализа и SWOT[2]-анализа

- Анализ исходной ситуации (“как есть”) вокруг СЭО с выделением базисных факторов, характеризующих экономические, политические и другие процессы, протекающие в СЭО и в его макроокружении и влияющие на развитие объекта.
- Определение факторов, характеризующих сильные и слабые стороны СЭО.
- Определение факторов, характеризующих возможности и угрозы со стороны внешней среды объекта.
- Построение проблемного поля СЭО.

Построение имитационной модели развития СЭО — формализация знаний, полученных на этапе когнитивной структуризации

- Определение и обоснование факторов.
- Установление и обоснование взаимосвязей между факторами.
- Построение графовой модели.

Структурно-целевой анализ

- Определение вектора целей развития СЭО (с учетом проблемного поля СЭО).
- Анализ непротиворечивости целей в векторе целей.
- Анализ непротиворечивости вектора целей и вектора управлений.
- Анализ эффективности интегрального влияния вектора управлений на достижимость вектора целей.

Сценарное исследование тенденций развития СЭО

- Задание сценариев исследования (на основе результатов структурно-целевого анализа).
- Выявление тенденций развития объекта в его макроокружении в условиях саморазвития и управляемого развития:
  — моделирование саморазвития ситуации в СЭО на основе экстраполяции начального состояния ситуации;
  — моделирование управляемого развития ситуации в СЭО, которое определяется выбираемой целью управления (вектором целей).
- Интерпретация результатов сценарного исследования.

Обоснование возможных сценариев развития и выработка рекомендаций

[1] *PEST* — *P* — political, *E* — economical, *S* — social, *T* — technological.
[2] *SWOT* — *Strengths* — сильные стороны, *Weaknesses* — слабые стороны, *Opportunities* — возможности, *Threats* — угрозы.

деленного «пограничного слоя» внешней среды, провести структурно-целевой анализ непротиворечивости в векторе целей, согласованности вектора управлений с вектором целей и начальными тенденциями развития СЭО.

Разработанные модели и структурно-целевой анализ реализованы в интеллектуальной аналитической технологии моделирования развития ситуации (см. таблицу) на базе компьютерных диалоговых комплексов, позволившей государственным организациям обосновывать цели развития СЭО в меняющихся ситуациях [14—20].

ЛИТЕРАТУРА

1. *Максимов В. И., Коврига С. В.* Целеполагание и стратегическое управление развитием сложных социально-экономических объектов в нестабильной внешней среде // Труды 5-ой Междунар. научн.-практ. конф. «Анализ систем на рубеже тысячелетий: теория и практика—2001» / ИПУ РАН, — М., 2001. — Т. 2. — С. 23—36.
2. *Максимов В. И., Качаев С. В., Корноушенко Е. К.* Управление сферами банковской деятельности // Банковские технологии. — 1999. — № 5—6. — С. 21—26.
3. *Максимов В. И., Корноушенко Е. К.* Аналитические основы применения когнитивного подхода при решении слабоструктурированных задач // Труды Ин-та пробл. упр. им. В. А. Трапезникова РАН. — М., 1999. — Т. II. — С. 95—109.
4. *Коврига С. В., Максимов В. И.* Технология когнитивного моделирования целенаправленного развития регионов РФ // Труды Ин-та пробл. упр. им. В. А. Трапезникова РАН. — М., 2000. — Т. XI.— С. 91—103.
5. *Максимов В. И.* Когнитивный анализ и управление развитием ситуаций // Материалы 1-й Междунар. конф. «Когнитивный анализ и управление развитием ситуаций» (CASC'2001) / ИПУ РАН. — М., — 20001. — Т. 2. — С. 10—21.
6. *Методы* формирования сценариев развития социально-экономических систем / В. В. Кульба, Д. А. Кононов, С. А. Косяченко, А. Н. Шубин. — М.: СИНТЕГ, 2004. — 296 с.
7. *Робертс Ф.* Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам. — М.: Наука, 1986.
8. *Корноушенко Е. К., Максимов В. И.* Управление ситуацией с использованием структурных свойств ее когнитивной карты // Тр. Ин-та пробл. упр. им. В. А. Трапезникова РАН. — М., — 2000. — Т. XI. — С. 85—90.
9. *Прангишвили И. В., Максимов В. И.* Разрешение проблемных ситуаций в период современной трансформации // Общество и экономика. — 2001. — № 11—12. — С. 42—70.
10. *Максимов В. И.* III Международная конференция «Когнитивный анализ и управление развитием ситуаций» CASC'2003 // Проблемы управления. — 2004. — № 1. — С. 66—69.
11. *Максимов В. И., Тер-Егиазарова Н. В.* IV Международная конференция «Когнитивный анализ и управление развитием ситуаций» CASC'2004 // Проблемы управления. — 2005. — № 1. — С. 83—87.
12. *Лоусон Ч., Хенсон Р.* Численное решение задач метода наименьших квадратов. — М.: Наука, 1986. — 232 с.
13. *Формулировка* и обсуждение ряда нестандартных задач и их математическое обеспечение в существующих пакетах / Алгоритмическое и программное обеспечение прикладного статистического анализа. — М.: Наука, 1980. — 424 с.
14. *Maximov V.* Cognitive Analysis and Situation Modelling // Proc. of the 8th IFAC Conference on «Social Stability: The Challenge of Technology Development» (SWIIS'01). Sept. 27—29, 2001, Vienna, Austria.
15. *Maximov V., Kornoushenko E.* Analitical Basics of Construction of Graph and Computer Models for Complicated Situations // Proc. of the 10th IFAC Symposium on Information Control Problems in Manufacturing (INCOM 2001). Sept. 20—22, 2001, Vienna, Austria.
16. *Maximov V., Kornoushenko E.* Mathematical Basics of Construction of Graph and Computer Models for Complicated Situations // Proc. of the IFAC Symposium on Modelling and Control of Economic Systems (SME 2001). Sept. 6—8, 2001, Klagenfurt, Austria.
17. *Maximov V., Kornoushenko E., Makarenko D.* Use of Cognitive Modelling for Analysis of Socio-Economic Processes and Estimation of Variants of the Regional Development // Proc. of the 10th IFAC Symposium on Information Control Problems in Manufacturing (INCOM 2001). Sept. 20—22, 2001. Vienna, Austria.
18. *Goal* Setting and Structute and Goal Analysis of Complex Systems and Situations / Z. Avdeeva, S. Kovriga, D. Makarenko, V. Maximov // Proc. of the 8th IFAC Symposium on Automated Systems Based on Human Skill and Knowledge, Sept., 2003, Göteborg, Sweden.
19. *Goal* Setting and Working out of the Strategy of Development of Socio-economic Objects / Z. Avdeeva, S. Kovriga, D. Makarenko, V. Maximov // Proc. of the IFAC Workshop on Technology and International Stability. July, 2003, Waterford, Ireland.
20. *Makarenko D., Avdeeva Z., Maximov V.* Cognitive Approach to Control of Socio-Economic Systems Security // Systems, Men & Cybernetics. Proc. of the 2004 IEEE International Conference. — Hague, Netherlands, 2004. — P. 899—903.

☎ *(095) 334-78-00*

E-mail: maxi@ipu.ru

❑

УДК 15:519.876

ПРИМЕНЕНИЕ СТРУКТУРНО-ЦЕЛЕВОГО АНАЛИЗА РАЗВИТИЯ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИТУАЦИЙ

В. И. Максимов, С. В. Коврига

Институт проблем управления им. В.А. Трапезникова, г. Москва

Показано применение структурно-целевого анализа для определения целей развития сложного социально-экономического объекта (региона) и выявления конфликтных областей между активными субъектами ситуации.

ВВЕДЕНИЕ

При целеполагании и формировании стратегии развития сложного социально-экономического объекта (СЭО) важная роль отводится *активным субъектам ситуации*[1] (АСС), т. е. субъектам, существенно влияющим на формирование целей развития СЭО через реализацию своих интересов и противодействие интересам других субъектов[2].

Повышение значимости учета человеческого фактора на этапе целеполагания также обусловлено тем, что для каждого АСС свойственны свои внутренние представления и знания (картины, модели мира) о ситуации. Эти представления позволяют ему ориентироваться, адаптироваться к проблемным условиям, ставить цели и, в конечном счете, принимать решения. Картина мира включает в себя набор убеждений, особенностей восприятия ценностных и практических установок субъекта, которыми он руководствуется в своей деятельности и влияет на ход развития ситуации. Различные представления и точки зрения, убеждения и интересы АСС приводят к разному видению направлений развития СЭО. Следовательно, могут ставиться и различные стратегические цели развития СЭО.

Структуры знания в мышлении субъекта (лица, принимающего решения), оказываются важнейшими элементами ситуации, неустранимыми из модели принятия решений [1], одной из основных составляющих которой является этап целеполагания.

Когнитивное картирование (построение когнитивной карты) является одним из способов создания «образа» ситуации, который существует у субъекта в неявной (невысказанной) форме, не осознанной им самим до конца [1—4]. При этом реконструируемый образ содержит «наиболее яркие, значимые признаки» для субъекта [1].

В рамках технологии когнитивного анализа и моделирования [5] статическая часть модели представляется *когнитивной коллективной картой* с мозаичной структурой [6], в которой агрегируются индивидуальные представления АСС, компетентных в различных предметных областях, связанных с развитием СЭО в изменяющейся внешней среде. В когнитивной карте отображаются непосредственные взаимовлияния между существенными (базисными) факторами, посредством которых создается образ ситуации, характеризующей собственно объект и его внешнее окружение.

Формально когнитивная карта представляет собой взвешенный ориентированный граф $G = (X, A)$, в котором X — множество вершин, взаимно однозначно соответствующих множеству базисных факторов ситуации, A — множество дуг, отражающих факт непосредственного влияния факторов [5]. Каждая дуга, связывающая некоторый фактор x_i с некоторым фактором x_j, имеет вес a_{ij}, знак которого говорит о знаке влияния фактора x_i на фактор x_j, а модуль величины a_{ij} — о силе этого влияния. Когнитивную карту можно также рассматривать как матрицу смежности A_g графа G.

Структурно-целевой анализ [7] когнитивной карты направлен на определение вектора целей развития сложного социально-экономического объекта и выявление конфликтных областей между АСС.

1. СТРУКТУРНО-ЦЕЛЕВОЙ АНАЛИЗ МОДЕЛИ РАЗВИТИЯ РЕГИОНА

При постановке целей развития региона руководству не всегда удается отследить, что определяемые им цели могут оказаться противоречивыми, т. е. достижение одних целей может негативно сказаться на достижимости других. На согласованность целей также могут влиять и выбираемые средства их достижения.

---

[1] Понятие ситуация включает в себя различные процессы (социальные, экономические, политические и др.), характеризующие взаимодействие внутренней и внешней сред СЭО.

[2] В качестве СЭО может рассматриваться регион, округ, транснациональная компания и др. Соответственно, к АСС можно отнести федеральное и региональное руководство, руководство компании и его окружение.

Это обусловлено, во-первых, тем, что при постановке целей затрагиваются интересы и установки различных АСС, вовлеченных в процесс целеполагания при формировании стратегии развития региона. Реальные стратегии развития — это результат согласования долгосрочных интересов различных АСС. Во-вторых, тем, что трудно учесть структуру опосредованных влияний между факторами, влияющих на достижение вектора целей, что становится источником скрытых противоречий в целях. Опосредованность влияния заключается в том, что складывающиеся тенденции развития ситуации в регионе могут способствовать достижимости одних целей, но при этом мешать достижению других целей или выбираемый способ достижения каких-то целей может препятствовать достижению остальных целей. И в том, и в другом случае это может носить скрытый характер, т. е. на этапе постановки целей предвидеть такие варианты очень сложно, а порой и невозможно.

Поэтому важно уже при целеполагании выявлять противоречия в векторе целей, которые могут быть вызваны либо собственными противоречиями в целях, либо выбираемыми способами достижения этих целей, что и позволяет выполнять в рамках технологии когнитивного моделирования *структурно-целевой анализ.*

На его основе определяются интегральные (непосредственные и все возможные опосредованные) влияния изменения одних факторов на изменения других факторов через *транзитивное замыкание* когнитивной карты [7], благодаря этому выявляются противоречия между факторами, образующими векторы целей и управлений.

Определение 1. *Вектор целей непротиворечив, если изменение любого целевого фактора в желаемом направлении не влечет за собой нежелательного изменения остальных целевых факторов в векторе целей* [7].

Определение 2. *Вектор управлений согласован с вектором целей, если изменение каждого из управляющих факторов в желаемом направлении не приводит к нежелательным изменениям целевых факторов в векторе целей* [7].

Нахождение противоречий в векторах целей и управлений базируется на введении показателя *оценка динамики фактора* (ОДФ) [7].

Анализируя ситуацию развития в регионе, можно выдвигать различные гипотезы о желательной динамике любого фактора модели. Для этого в когнитивную модель для каждого фактора вводится ОДФ. Если благоприятна положительная (отрицательная) динамика некоторого фактора, то этому фактору приписывается ОДФ, равная 1 (−1); если затруднительно дать оценку по фактору, то ОДФ полагается равной нулю. Задание вектора ОДФ по некоторому набору факторов модели отражает желательное изменение ситуации в регионе относительно выделенных факторов модели, т. е. введение показателей ОДФ позволяет получить оценку «благоприятности» того или иного состояния анализируемой ситуации.

«Образ» (модель) развития региона, реконструированный в когнитивной карте, включает в себя 38 базисных факторов, которые характеризуют ресурсную, демографическую, социальную, бюджетную, налоговую и экологическую сферы, региональную экономику и взаимоотношения с федеральным центром:

1. Ресурсный потенциал региона (1)
2. Возможность развития ресурсного потенциала (1)
3. Состояние основных фондов предприятий (1)
4. Трудовые ресурсы региона (1)
5. Уровень жизни населения региона (1)
6. Социальная напряженность в регионе (−1)
7. Доходы населения региона (1)
8. Уровень криминогенности в регионе (−1)
9. Уровень безработицы в регионе (−1)
10. Уровень развития социальной инфраструктуры (1)
11. Смертность в регионе (−1)
12. Рождаемость в регионе (1)
13. Миграция из региона (−1)
14. Экономический потенциал региона (1)
15. Инвестиционный климат региона (1)
16. Инвестиционный риск финансовых вложений в регион (−1)
17. Объем производства предприятий региона (1)
18. Спрос на продукцию, произведенную в регионе (1)
19. Прибыль (доходы) предприятий региона (1)
20. Издержки производства предприятий региона (−1)
21. Инвестиции в регион (1)
22. Уровень развития производственной инфраструктуры (1)
23. Уровень цен на потребительские товары в регионе (−1)
24. Доходы бюджета региона (1)
25. Расходы бюджета региона (−1)
26. Дефицит бюджета региона (−1)
27. Необходимость федеральной финансовой помощи региону (1)
28. Доходы от тендеров и приватизации (1)
29. Налоговые поступления в бюджет региона (1)
30. База налогообложения (1)
31. Собираемость налогов (1)
32. Налоговые льготы предприятиям региона (−1)
33. Неплатежи налогов в бюджет региона (−1)
34. Преференции региону (−1)
35. Зависимость региона от федерального центра (1)
36. Возможность проявления регионального сепаратизма (−1)
37. Экологическая безопасность среды обитания (1)
38. Условия жизнедеятельности коренного населения (1)

В скобках указаны значения ОДФ (1 или −1).

Пусть вектор целей развития региона представлен факторами № 2, 5, 6, 10, 13, 15, 17, 22, 35, 38. Интегральные взаимовлияния между целевыми факторами выделенного вектора целей развития региона представлены в табл. 1, которая является подматрицей M_Q матрицы транзитивного замыкания Q [7].

Каждое значение в подматрице M_Q, соответствующее значению $q_{ij} \in Q$ на пересечении i-й строки и j-го столбца матрицы Q, определяет интегральное влияние i-го целевого фактора на j-й целевой фактор через все цепочки опосредованных влияний и непосредственного влияния (i, j — порядковые номера факторов). Знак при q_{ij} определяет направление влияния: «+» — усиливающее, «−» — ослабляющее. Например, интегральное влияние целевого фактора № 17 «Объем производства предприятий региона» на целевой фактор № 15 «Инвестиционный климат региона» составляет +0,8 (очень сильное положительное влияние).

Определение непротиворечивости целевых факторов между собой заключается в сопоставлении ОДФ факторов и направлений интегральных влияний между ними.

Два целевых фактора i, j *непротиворечивы*, если знак произведения ОДФ i-го фактора на его интегральное влияние совпадает со знаком ОДФ j-го фактора. Например, в табл. 1 (см. первую строку) фактор № 2 «Возможность развития ресурсного потенциала» (ОДФ = 1) противоречив с фактором № 38 «Условия жизнедеятельности коренного населения» (ОДФ = 1), так как интегральное влияние первого на второй равно −0,2. Содержательно противоречие интерпретируется следующим образом: увеличение возможностей развития ресурсного потенциала региона (ОДФ = 1) способствует ухудшению условий жизнедеятельности коренного населения, а для последнего желательно улучшение (ОДФ = 1).

Наглядно результаты анализа непротиворечивости вектора целей по матрице Q представлены на диаграмме (рис. 1), где по осям координат расположены номера целевых факторов. Темный шар указывают на наличие противоречия между двумя целевыми факторами, светлые шары — на отсутствие противоречий, пустые места — на отсутствие интегральных влияний между факторами.

На диаграмме видно, что в целом вектор целей непротиворечив, за исключением фактора № 2 «Возможность развития ресурсного потенциала» и фактора № 38 «Условия жизнедеятельности коренного населения». В когнитивной карте структура взаимовлияний между этими факторами такова (см. цветную вклейку), что непосредственное отрицательное влияние первого фактора на второй лишь частично компенсируется из-за положительных опосредованных влияний между ними. Поэтому желательное изменение фактора «Возможность развития ресурсного потенциала» приводит к нежелательному изменению фактора «Условия жизнедеятельности коренного населения».

Один из возможных способов снятия этого противоречия заключается в выполнении двух условий:

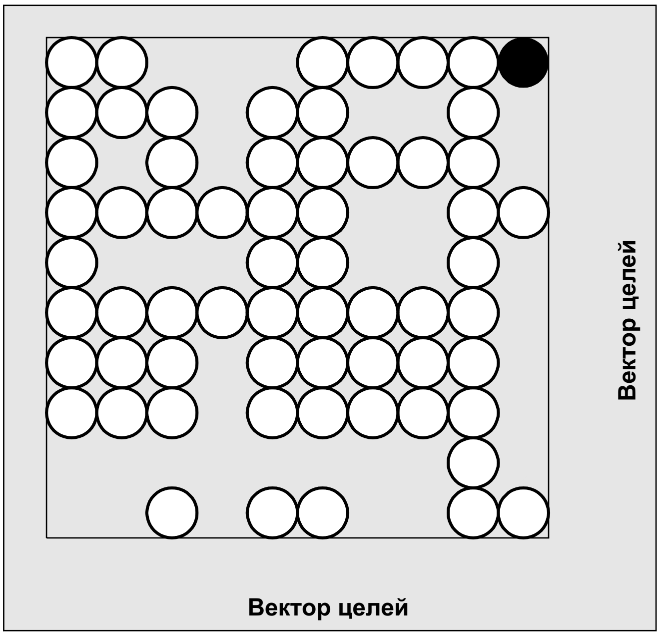


Рис. 1. Результаты анализа непротиворечивости вектора целей

1) нахождение такого *вектора управлений*, который был бы согласован с выбранным вектором целей (т. е. все управления обеспечивают благоприятную динамику целевых факторов), и при этом

2) интегральное влияние вектора управлений на целевой фактор «Условия жизнедеятельности коренного населения» должно превосходить интегральное влияние на целевой фактор «Возможность развития ресурсного потенциала».

Таблица 1

Интегральные взаимовлияния вектора целевых факторов

Вектор целей	ОДФ	Вектор целей									
		2	5	6	10	13	15	17	22	35	38
		1	1	−1	1	−1	1	1	1	−1	1
2. Возможность развития ресурсного потенциала региона	1	1	0,1	–	–	–	1	0,3	0,2	−0,7	**−0,2**
5. Уровень жизни населения региона	1	0,1	1	−0,5	–	−0,4	0,3	–	–	−0,3	–
6. Социальная напряженность в регионе	−1	−0,2	–	1	–	0,4	−0,5	−0,1	−0,1	0,5	–
10. Развитие социальной инфраструктуры	1	0,3	0,3	−0,2	1	−0,2	0,3	–	–	−0,3	0,2
13. Миграция из региона	−1	−0,1	–	–	–	1	−0,2	–	–	0,1	–
15. Инвестиционный климат региона	1	0,5	0,2	−0,3	0,2	−0,3	1	0,5	0,4	−0,8	–
17. Объем производства предприятий региона	1	0,4	0,5	−0,6	–	−0,6	0,8	1	0,2	−0,7	–
22. Уровень развития производственной инфраструктуры	1	0,5	0,2	−0,2	–	−0,3	0,8	0,5	1	−0,5	–
35. Зависимость региона от федерального центра	−1	–	–	–	–	–	–	–	–	1	–
38. Условия жизнедеятельности коренного населения	1	–	–	−0,5	–	−0,3	0,2	–	–	−0,2	1

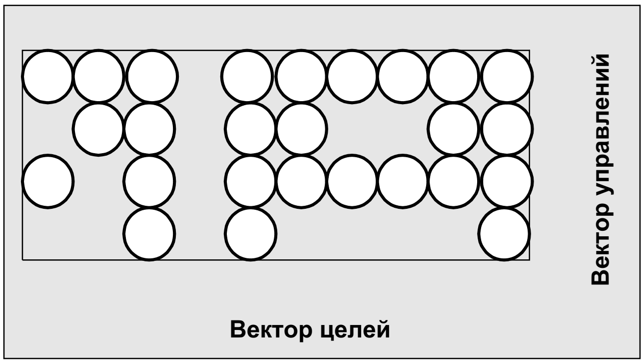


Рис. 2. Результаты анализа согласованности вектора управлений и вектора целей

Пусть в качестве вектора управлений задан вектор из четырех факторов: «Состояние основных фондов предприятий», «Доходы населения региона», «Инвестиции в регион», «Экологическая безопасность среды обитания».

Определение согласованности вектора управления с вектором целей заключается в сопоставлении ОДФ управляющих и целевых факторов и направлений интегральных влияний между ними. Управляющий фактор i согласован с целевым фактором j, если знак произведения ОДФ i-го фактора на его интегральное влияние ($q_{ij} \in Q$) совпадает со знаком ОДФ j-го фактора. В табл. 2 сведены интегральные влияния вектора управлений на вектор целей.

Выбранный вектор управлений согласован с вектором целей (см. табл. 2 и рис. 2) — условие (1) выполняется. Однако второе условие не обеспечивается, так как влияние вектора управлений на фактор № 2 «Возможность развития ресурсного потенциала» несколько превосходит влияние на целевой фактор № 38 «Условия жизнедеятельности коренного населения» (см. последнюю строку в табл. 2).

Поэтому, несмотря на то, что вектор управлений согласован с вектором целей, т. е. все управления обеспечивают благоприятную динамику целевых факторов, нейтрализовать противоречие между целевыми факторами № 2 и № 38 не удается. А это означает, что вектор управлений обеспечивает улучшение условий жизнедеятельности коренного населения, но из-за фактора «Возможность развития ресурсного потенциала» это улучшение будет сдерживаться.

Таким образом, структурно-целевой анализ дает возможность на этапе целеполагания формировать непротиворечивые векторы целей и согласованные с ними векторы управлений. Результаты анализа могут служить основой для сценарного моделирования развития ситуации в регионе [5].

2. ВЫЯВЛЕНИЕ КОНФЛИКТНЫХ ОБЛАСТЕЙ МЕЖДУ АКТИВНЫМИ СУБЪЕКТАМИ СИТУАЦИИ

Ситуация, в которой возникает предмет конфликта и выделяются противоборствующие стороны — субъекты или группы, имеющие определенные интересы относительно этого предмета, — называется конфликтной ситуацией [8]. Конфликтная ситуация возникает, когда реальные противоречия выражаются в противоборстве их носителей, которые стремятся помешать друг другу в достижении какой-либо цели, удовлетворении каких-либо интересов или изменить взгляды, мнения или позиции друг друга.

В технологии когнитивного анализа и моделирования [5] выявление конфликтных областей, характеризующих противоречия между АСС, опирается на структурно-целевой анализ [7].

Если при структурно-целевом анализе не ограничиваться только заданием векторов целей и управлений, а задать вектор ОДФ для всех факторов модели, то в результате будут выявлены *области противоречий* на всей когнитивной карте. Суть противоречия между парой факторов состоит в том, что изменение одного фактора в желательном направлении (в соответствии с его ОДФ) приводит к нежелательному изменению второго фактора (т. е. к изменению в направлении, противоположном установленному для него ОДФ) через связывающую их структуру взаимовлияний в когнитивной карте.

Пусть множество факторов модели развития региона разбито на пять групп, каждая из которых определяет области интересы отдельных АСС: *I* — федеральное руководство, *II* — региональное руководство, *III* — региональные производители, *IV* — внешние инвесторы, *V* — население региона. Интересы каждой группы АСС отражаются посредством вектора ОДФ для факторов, входящих в группу.

Таблица 2

Интегральные взаимовлияния управляющих факторов на целевые

Вектор управлений	ОДФ	Вектор целей									
		2	5	6	10	13	15	17	22	35	38
		1	1	−1	1	−1	1	1	1	−1	1
3. Состояние основных фондов предприятий	1	0,7	0,4	−0,5	–	−0,5	1	1	0,3	−0,9	0,2
7. Доходы населения региона	1		0,7	−0,4	–	−0,3	0,2	–	–	−0,2	0,3
21. Инвестиции в регион	1	0,6	0,4	−0,4	0,3	−0,5	1	0,8	0,6	−0,8	0,1
37. Экологическая безопасность среды обитания	1	–	–	−0,2	–	−0,1	−0,2	–	–	0,1	0,5
Интегральное влияние всего вектора управлений на целевые факторы		**1,3**	1,5	−1,5	0,3	−1,4	1	0,8	0,9	−1,8	**1,1**

На диаграмме (см. цветную вклейку) область (*II*, *II*) — это область интересов регионального руководства; (*II*, *III*) и (*III*, *II*) — области интересов двух групп: регионального руководства и региональных производителей и т. д. Области с черными шарами указывают на области противоречий, которые являются потенциальными источниками зарождения конфликтной ситуации между группами АСС.

Определив конфликтные области АСС (области противоречий) на когнитивной карте, можно по-разному подходить к определению векторов целей и управлений при выработке стратегий развития региона.

Подход 1. Формирование векторов целей и управлений вне конфликтных областей интересов АСС. В этом случае любой выбранный вектор целей непротиворечив, а любой вектор управлений согласован с выбранным вектором целей. Тогда определение непротиворечивых векторов целей и согласованных с ними векторов управлений способствует достижению поставленных целей при условии, что заданные начальные тенденции развития ситуации в регионе не противодействуют достижению целей при выбранном векторе управлений. Если указанное условие не выполняется, то при моделировании требуется корректировать вектор управлений для нейтрализации негативных исходных тенденций.

Если же при моделировании не удается переломить существующие тенденции из-за выбранного вектора управлений, то либо изменяются целевые установки (вектор целей), либо пересматривается реконструированный образ ситуации с целью выявления новых существенных возможностей для управления развитием ситуации в регионе.

Подход 2. Формирование векторов целей и управлений в конфликтных областях интересов АСС. В этом случае возможны две стратегии, определяющие поведение при развитии *продуктивного и деструктивного конфликтов* [8], соответственно.

Продуктивный конфликт — это конфликт, когда конфликтующие стороны находят возможность его разрешения. Противоречия в интересах и связанных с ними целях снимаются благодаря компромиссам, позволяющим сторонам находить возможности разрешения конфликта. Такой конфликт становится источником позитивного развития. Деструктивный конфликт имеет тенденцию к расширению и эскалации и часто завершается победой одной из конфликтующих сторон. Поэтому главная задача изучения конфликта с целью его разрешения — это выявление факторов, определяющих развитие конфликта по продуктивному или деструктивному пути.

Если вектор целей задан в конфликтной области интересов АСС (есть противоречивые цели), то снятие каких-то противоречий в целях (в лучшем случае, всех противоречий) возможно из-за выбора вектора управлений, который согласован с вектором целей и который нейтрализует негативные начальные тенденции в развитии ситуации; т. е. при моделировании выбирается путь разрешения продуктивного конфликта, который дает возможность развивать ситуацию в регионе частично или в полном объеме в желаемом направлении.

Если при моделировании есть интерес в развитии деструктивного конфликта, чтобы достичь целей одной из конфликтующих сторон, то возможны два варианта.

- Вектор целей выбирается из конфликтной области интересов АСС, т. е. имеющиеся противоречия могут способствовать нежелательному изменению целей одной из конфликтующих сторон при достижении целей другой стороной. Тогда для обострения противоречий находится такой вектор управлений, который обеспечивает достижение целей одной из сторон.
- Если необходимо противодействовать достижению вектора целей (или отдельных его целей), в котором нет противоречий[3], то находится несогласованный с вектором целей вектор управлений из конфликтной области интересов АСС, который может мешать достижению целей.

В обоих вариантах необходимо также учитывать возможное негативное влияние начальных тенденций развития ситуации на достижение целей.

Независимо от выбранного подхода, определение вектора управлений может осуществляться из любых областей когнитивной карты при условии его согласованности с вектором целей, т. е. изменение любого управляющего фактора (в соответствии с его ОДФ) не приводит к нежелательному изменению целевых факторов. Однако отметим, что выбор вектора управлений может при моделировании привести к обострению существующих противоречий в областях, не связанных с заданным вектором целей; т. е. заданный вектор целей выбранных групп АСС будет достигаться, но при этом могут обостриться противоречия в целях, интересах других групп АСС. Это необходимо учитывать при выработке стратегий развития региона, потому что обострение противоречий между группами АСС, чьи интересы не учитываются в рассматриваемом векторе целей, может негативно сказываться на развитии региона даже при достижении заданных целей.

ЗАКЛЮЧЕНИЕ

При выработке стратегии развития социально-экономического объекта важную роль играют активные субъекты ситуации, которые оказывают существенное влияние на формирование целей развития региона через реализацию своих интересов и противодействие интересам других субъектов.

Поэтому на этапе целеполагания необходимо выявление конфликтных областей, характеризующих противоречия между активными субъектами ситуации, что позволяет

- при постановке целей учесть различные интересы, выявить противоречия в интересах и, по возможности, найти пути преодоления этих противоречий выбором стратегии развития продуктивного конфликта, либо обойти эти противоречия путем формирования целей и способов их реализации вне конфликтных областей;
- выявить неосуществимые стратегии развития региона путем определения областей противоречий, которые ведут к деструктивному конфликту.

[3] Вектор целей выбирается вне конфликтной области интересов АСС.

ЛИТЕРАТУРА

1. *Саламатов В. А., Таран Т. А.* Реконструкция субъективного образа социальной реальности // Новости искусственного интеллекта. — 1998. — № 3. — С. 142—154.
2. *Силов В. Б.* Принятие стратегических решений в нечеткой обстановке. — М.: ИНПРО-РЭС, 1995.
3. *Kelly G. A.* The Psychology of Personal Constructs. Vol. № 1: A Theory of Personality. — N.-Y.: Norton & Company, 1995.
4. *Робертс Ф.* Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам. — М.: Наука, 1986.
5. *Максимов В. И.* Когнитивный анализ и управление развитием ситуаций // Когнитивный анализ и управление развитием ситуаций (CASC'2001) / ИПУ РАН: Матер. 1-й Междунар. конф. — М., 2001. — Т. 2. — С. 10—21.
6. *Максимов В. И., Райков А. Н.* Коллективные когнитивные карты в системах принятия решений // Междунар. симп. «Рефлексивное управление» / Ин-т психологии РАН: Тез. докл. — М., 2000. — С. 86—88.
7. *Максимов В. И.* Структурно-целевой анализ развития социально-экономических ситуаций // В этом номере журнала — С. 30—38.
8. *Таран Т. А.* Анализ и моделирование когнитивных конфликтов // Когнитивный анализ и управление развитием ситуаций (CASC'2001) / ИПУ РАН: Тр. 2-й Междунар. конф. — М., 2002. — Т. 2. — С. 96—109.

☎ *(095) 334-78-00*

E-mail: maxi@ipu.ru ❑

УДК 65.012; 658.5

ФИНАНСОВО-ЭКОНОМИЧЕСКИЕ МЕТОДЫ КОНТРОЛЛИНГА В УПРАВЛЕНИИ ХОЗЯЙСТВЕННОЙ ДЕЯТЕЛЬНОСТЬЮ ИНТЕГРИРОВАННЫХ КОМПАНИЙ. Ч. II

А. В. Карибский, Д. Ю. Мишутин, Ю. Р. Шишорин

Институт проблем управления им. В. А. Трапезникова, г. Москва

Рассмотрены формализованные финансово-экономические методы контроллинга, реализуемые при управлении хозяйственной деятельностью интегрированных компаний. Дана общая формализованная постановка задачи оптимизации учетной политики и описаны принципы построения имитационных моделей бюджетирования. Рассмотрены методы решения задачи оптимизации бюджета и приведен практический пример ее численного решения.

ВВЕДЕНИЕ

В первой части настоящей работы описаны и исследованы задачи контроллинга, возникающие в процессе финансово-экономического планирования и управления хозяйственной деятельностью интегрированных компаний [1]. Для практической реализации их решений предложен комплекс взаимосвязанных оптимизационных и имитационных моделей, включающий в себя модели оптимизации амортизационной, налоговой и договорной политики (формирования учетной политики) и детализированные имитационные модели бюджетирования (формирования финансовых планов). Рассмотрим методы формализации и решения поставленных задач в рамках предложенного комплекса моделей.

1. МЕТОДЫ ФОРМАЛИЗОВАННОГО ОПИСАНИЯ ЗАДАЧИ ОПТИМИЗАЦИИ УЧЕТНОЙ ПОЛИТИКИ

Как отмечалось [1], содержательная постановка общей задачи оптимизации учетной политики заключается в следующем: необходимо выбрать такие способы амортизации основных средств и параметры списания их стоимости (сроки полезного использования, коэффициенты ускорения и др.), условия договорных отношений с контрагентами (способы оплаты, длительности отсрочек, скидки/надбавки к ценам и т. п.) и методы расчета налогов («по оплате» или «по отгрузке»), чтобы обеспечивался максимальный объём накопленного на всем горизонте планирования потока денежных средств от операционной деятельности (чистой прибыли и амортизации за вычетом инвестиций в прирост оборотного капитала) при выполнении ряда финансово-экономических ограничений на результаты текущей хозяйственной деятельности.

В агрегированном виде эта общая задача оптимизации может быть формализована следующим образом: найти вектора Z, K, T_a, X и Y, матрицы M, Δt_B и Δt_C и значение переменной z_h которые доставляют экстремум целевой функции $F(Z, K, T_a, \Delta t_B, \Delta t_C, M, z_h, X, Y)$ на заданном горизонте расчета T для всей совокупности объектов компании:

$$F(\cdot) \to \text{extr}, \qquad (1)$$

при ограничениях, описываемых следующей системой уравнений с учетом условий целочисленности и непрерывности искомых переменных:

$$C(Z, K, T_a, \Delta t_B, \Delta t_C, M, z_h, X, Y) = 0; \qquad (2)$$

$$Z,\ M,\ z_h \in \{0, 1\},\ X,\ Y \in [0, 1],\ K,\ T_a \geqslant 0,$$
$$-\infty < \Delta t_B,\ \ \Delta t_C < +\infty, \qquad (3)$$

где Z — двоичный n-мерный вектор, отражающий выбор метода амортизации для каждого из n объектов компании; z_h — двоичная переменная, описывающая выбор метода налоговой и договорной политики; M — двоичная $(n \times T)$ матрица признаков ввода объектов в эксплуатацию; K и T_a — непрерывные n-мерные вектора коэффициентов ускорения амортизации объектов основных средств и сроков полезного использования объектов, соответственно; X, Y — условия переключения с одного метода амортизации на другой, исходя из степени износа объектов; Δt_B и Δt_C — непрерывные $(n \times T)$ матрицы, отражающие сроки оборачиваемости расчетов с покупателями и поставщиками, причем положительные элементы соответствуют отгрузкам продукции (приобретению материалов) в кредит, а отрицательные — по предоплате.

Для детализации постановки задачи (1)—(3) рассмотрим конкретные составляющие и формализованные выражения целевой функции F и системы ограничений G (см. также работы [2, 3]).

Пусть n — число объектов компании, в отношении которых осуществляется выбор учетной политики; i — индекс объекта, $i = \overline{1, n}$; $I_{i\tau}$ — объём инвестиций в i-й объект в τ й момент времени, объём $I_{i\tau}$ отличен от нуля при $1 \leqslant \tau \leqslant \bar{\tau}_i$, $\bar{\tau}_i$ — длительность фазы инвестирования в i-й объект, и равен нулю при всех других τ; B_{it} — объём отгруженной продукции покупателю в t-й момент времени i-м объектом в стоимостном выражении, моменты оплаты которой могут меняться в зависимости от выбранных условий договоров; C_{it} — объём возникших на i-м объекте в t-й момент обязательств по оплате производственных издержек в стоимостном выражении, сроки погашения которых могут меняться в зависимости от условий договоров; I_i — объём инвестиций в i-й объект за всю фазу инвестирования; Δt_{bi} — отрезок времени, в течение которого i-й объект находится на балансе, но ещё не начал функционировать. Переменные задачи: вектор булевых переменных $Z = \{z_1, z_2, ..., z_i, ..., z_n\}$, $z_i = 1$, если выбран нелинейный метод амортизации и $z_i = 0$ — если линейный; T_{ai} – период амортизации i-го объекта; X_i — отношение остаточной стоимости объекта к первоначальной, при котором происходит переключение с нелинейного на линейный метод амортизации; Y_i — отношение остаточной стоимости объекта к первоначальной, при котором происходит переключения с линейного на нелинейный метод амортизации; вектор моментов начала эксплуатации i-го объекта компании $M_i = \{m_{1i}, m_{2i}, ..., m_{ti}, ..., m_{Ti}\}$, $i = \overline{1, n}$, причем $m_{ti} = 1$, если i-й объект начинает функционировать в t-й период, $m_{ti} = 0$ — в противном случае; булева переменная z_h, принимающая значение 1, если выбрана схема учета денежных средств «по отгрузке», и 0 — «по оплате»; непрерывные переменные: k_i — коэффициент ускорения амортизации по i-му объекту, Δt_{itB} — срок оборачиваемости (в месяцах) при расчетах с покупателями за продукцию (услуги), отгруженную с i-го объекта в t-й период, положительное значение соответствует отгрузке в кредит, отрицательное — отгрузке по предоплате; Δt_{itC} — срок оборачиваемости (в месяцах) при расчетах по затратам, понесенным i-м объектом в t-й период, положительное значение соответствует оплате в кредит, отрицательное — предоплате.

В качестве критерия задачи (1)—(3) рассматривается наиболее часто применяемый критерий — максимум чистого дисконтированного дохода NPV на инвестированный капитал:

$$F(\cdot) = NPV = \sum_{t=1}^{T} NCF_t \frac{1}{(1+\delta)^t} \to \max, \qquad (4)$$

где NCF_t — поток чистых платежей, генерируемых объектами компании в период t, δ — ставка дисконтирования. За расчетный период (шаг расчета) в задаче принимается месяц — период начисления амортизационных отчислений. Значение NCF_t рассчитывается по формуле:

$$NCF_t = \text{ЧП}_t + Am_t - I_t - I_{\text{об}\,t} + Ost_t, \qquad (5)$$

где $ЧП_t$ — чистая прибыль, Am_t — амортизационные отчисления, I_t — инвестиции в развитие компании, $I_{\text{об}\,t}$ — прирост оборотного капитала, Ost_t — остаточная стоимость полностью не самортизированных объектов, все в t-й период. Отметим, что в формуле (5) динамика инвестиций I_t определяется только выбором моментов начала эксплуатации объектов (вектором переменных M_i), тогда как остальные составляющие величины NCF_t зависят от переменных, определяющих выбор учетной политики. Подставив формулу (5) в выражение (4), получим целевую функцию в виде:

$$F(\cdot) = \sum_{t=1}^{T_p} \frac{ЧП_t + Am_t - I_t - I_{\text{об}\,t} + Ost_t}{(1+\delta)^t}, \qquad (6)$$

где T_p — продолжительность фазы функционирования объектов компании на рассматриваемом горизонте планирования T. Функционирование объектов закончится, когда наступит момент T^* амортизации последнего работающего объекта или будет достигнут горизонт планирования T, т. е. $T_p = \min(T^*, T)$, причем $T^* = \max\limits_{i = 1, ..., n}(t_{ci}(M_i) + T_{ai}) - 1$, где $t_{ci}(M_i)$ — момент начала функционирования i-го объекта, выражаемый через переменные m_{ti} следующим образом: $t_{ci}(M_i) = \sum_{t=1}^{T} tm_{ti} \leqslant$ $\leqslant T^c \leqslant T$, где T^c — заданный самый поздний момент начала функционирования объектов, причем $\sum_{t=1}^{T} m_{ti} = 1$.

Объёмы инвестиций в i-й объект задаются набором постоянных величин $I_{i\tau}$. С учетом влияния на динамику инвестиций моментов ввода объектов в эксплуатацию объём инвестиций в проект в t-й момент времени задается выражением:

$$I_t(M_i) = \sum_{i=1}^{n} I_{i(\tau_i - t_{ci}(M_i))}. \qquad (6')$$

Амортизационные отчисления Am_t в t-й период по всей компании рассчитываются как сумма амортизационных отчислений по всем объектам: $Am_t = \sum_{i=1}^{n} Am_{ti}(z_i, k_i,$ $t_{ci}(M_i), T_{ai}, X_i, Y_i)$.

В свою очередь, расчет амортизационных отчислений $Am_{ti}(\cdot)$ для i-го объекта в t-й период зависит от выбранного метода амортизации и возможностей переключения на другой метод в процессе функционирования объекта, т. е.

$$Am_{ti}(\cdot) = z_i Am_{ti}^{\text{нел}}(k_i, t_{ci}(M_i), T_{ai}, X_i) + \\ + (1 - z_i)\cdot Am_{ti}^{\text{лин}}(k_i, t_{ci}(M_i), T_{ai}, Y_i), \qquad (7)$$

где $Am_{ti}^{\text{нел}}(\cdot)$ и $Am_{ti}^{\text{лин}}(\cdot)$ — алгоритмически вычисляемые функции расчета амортизации по нелинейному и линейному методам с учетом возможного переключения с одного метода на другой. Функция $Am_{ti}^{\text{нел}}(\cdot)$ в выражении (7) представляется в виде:

$$Am_{ti}^{\text{нел}}(\cdot) = I_i P_{it}^{\text{нел}}(k_i, t_{ci}(M_i), T_{ai}, X_i)\cdot U_{it}(t_{ci}, T_{ai}), \qquad (8)$$

где множитель $U_{it}(t_{ci}(M_i), T_{ai})$ введен для того, чтобы начисления прекращались по истечению срока полезного использования i-го объекта:

$$U_{it} = \begin{cases} 1 \text{ при } t_{ci}(M_i) \le t \le t_{ci}(M_i) + T_{ai} - 1 \\ 0 \text{ при } t < t_{ci}(M_i) \text{ и } t > t_{ci}(M_i) + T_{ai} - 1. \end{cases} \qquad (9)$$

Переключение на линейный метод списания происходит после $\tilde{T}_i^{\text{нел}}$ месяцев функционирования i-го объекта, когда накопленная амортизация превысит $(1 - X_i)$-ю долю от первоначальной стоимости объекта. Исходя из этого, момент переключения $\tilde{T}_i^{\text{нел}}$ находится из решения следующей системы неравенств:

$$\begin{cases} \sum_{t=1}^{\tilde{T}_i^{\text{нел}}} \frac{I_i k_i}{T_{ai}}\cdot\left(1 - \frac{k_i}{T_{ai}}\right)^{t-1} \ge (1 - X_i)I_i \\ \sum_{t=1}^{\tilde{T}_i^{\text{нел}}-1} \frac{I_i k_i}{T_{ai}}\cdot\left(1 - \frac{k_i}{T_{ai}}\right)^{t-1} < (1 - X_i)I_i, \end{cases} \qquad (10)$$

которую можно упростить, явно выразив величину $\tilde{T}_i^{\text{нел}}$ через параметры и переменные задачи.

Параметр $P_{it}^{\text{нел}}(\cdot)$ в выражении (8), являющийся переменной нормой амортизации по отношению к первоначальной стоимости I_i, описывается следующей логической функцией:

$$P_{it}^{\text{нел}} = \begin{cases} k_i\cdot\frac{1}{T_{ai}}\left(1 - \frac{k_i}{T_{ai}}\right)^{t - t_{c_i}} \text{ при } t_{ci}(M_i) \le t \le t_{ci}(M_i) + \tilde{T}_i^{\text{нел}} \\ \text{(до подключения)} \\ \frac{Ost_{it}/I_i}{T_{ai} - \tilde{T}_i^{\text{нел}}} \text{ при } t > t_{ci}(M_i),\ t \ge t_{ci}(M_i) + \tilde{T}_i^{\text{нел}} \\ \text{(после подключения)} \end{cases} \qquad (11)$$

и обеспечивает начисление амортизационных отчислений по нелинейному методу до момента переключения. К моменту переключения остаточная стоимость i-го объекта

$$Ost_i = I_i - \sum_{t=1}^{\tilde{T}_i^{\text{нел}}} \frac{I_i k_i}{T_{ai}}\left(1 - \frac{k_i}{T_{ai}}\right)^{t-1} \qquad (12)$$

и будет затем списана за $T_{ai} - \tilde{T}_i^{\text{нел}}$ месяцев, оставшихся до конца срока амортизации, равными долями в размере $Am_{it}^{\text{нел}} = Ost_{it}/(T_{ai} - \tilde{T}_i^{\text{нел}})$.

Функция $Am_{ti}^{\text{лин}}(\cdot)$ в выражении (7) имеет вид:

$$Am_{ti}^{\text{лин}}(k_i, t_{ci}(M_i), T_{ai}, Y_i) = \\ = I_i k_i P_{it}^{\text{лин}}(k_i, t_{ci}(M_i), T_{ai}, Y_i)/T_{ai} \qquad (13)$$

где $I_i k_i/T_{ai}$ — амортизационные отчисления по линейной схеме с коэффициентом ускорения k_i, $P_{it}^{\text{лин}}(\cdot)$ — логическая функция, корректирующая норму равномерной амортизации и отражающая возможность переключения с линейной на нелинейную схему, когда накопленная амортизация по i-му объекту превысит $(1 - Y_i)$-ю долю от его первоначальной стоимости.

В момент переключения с линейного метода на нелинейный должны выполняться два условия:

$$\begin{cases} \sum_{t=1}^{\tilde{T}_i^{\text{лин}}} I_i k_i/T_{ai} \ge (1 - Y_i)I_i \\ \sum_{t=1}^{\tilde{T}_i^{\text{лин}}-1} I_i k_i/T_{ai} < (1 - Y_i)I_i, \end{cases} \qquad (14)$$

где $\tilde{T}_i^{\text{лин}}$ — число периодов, в течение которых спишется $(1 - Y_i)$-я часть основных средств i-го объекта, причем срок полезного использования объекта будет равен T_{ai}/k_i месяцев[1]. Переключение на нелинейную схему в последний T_{ai}-й месяц этого срока (такое возможно при $Y_i \approx 1$) экономически неэффективно, поскольку оста-

[1] Если это число получается дробным, то — в течение $[T_{ai}/k_i] + 1$ месяцев (здесь квадратные скобки означают целое от деления), причем в последний месяц отчисления равны остатку $(T_{ai}/k_i - [T_{ai}/k_i])I_i k_i/T_{ai}$.

точная стоимость будет очень мала и ее списание существенно не повлияет на показатели эффективности компании. Таким образом, функция $P_{it}^{\text{лин}}(\cdot)$ имеет вид:

$$P_{it}^{\text{лин}}(\cdot) = \begin{cases} P_{it}^{*}, \text{ если } [T_{ai}/k_i] \le \tilde{T}_i^{\text{лин}} \\ \text{(переключения не происходит)} \\ P_{it}^{*}, \text{ если } [T_{ai}/k_i] > \tilde{T}_i^{\text{лин}} \\ \text{(переключение происходит).} \end{cases} \quad (15)$$

При этом параметр P_{it}^{*} зависит от значений $t_{ci}(M_i)$, k_i и T_{ai}, и задается следующим логическим выражением:

$$P_{it}^{*}(t_{ci}(M_i), k_i, T_{ai}) = $$
$$= \begin{cases} 1 \text{ при } t_{ci}(M_i) \le t \le [T_{ai}/k_i] + t_{ci}(M_i) - 1 \\ T_{ai}/k_i - [T_{ai}/k_i] \text{ при } t = t_{ci}(M_i) + [T_{ai}/k_i] \\ 0 \text{ при } t < t_{ci};\ t > t_{ci}(Mi) + [T_{ai}/k_i], \end{cases}$$

которое означает, что ежемесячно начисляется амортизация в размере I_ik_i/T_{ai} (за исключением, быть может, последнего месяца, в котором списывается вся остаточная стоимость, т. е. когда коэффициент ускорения k_i не кратен периоду амортизации).

В случае возможности переключения параметр P_{it}^{**} в функции (15) зависит от $t_{ci}(M_i)$, k_i, T_{ai} и неявно от Y_i (через $\tilde{T}_i^{\text{лин}}$) и описывается логическим выражением

$$P_{it}^{**}(t_{ci}(M_i), k_i, T_{ai}, Y_i) = $$
$$= \begin{cases} 1 \text{ при } t_{ci}(M_i) \le t \le \tilde{T}_i^{\text{лин}} + t_{ci}(M_i) \\ \left(1 - \dfrac{k_i}{T_{ai} - T_i^{\text{лин}}}\right)^{t - t_{ci}(M_i) - \tilde{T}_i^{\text{лин}}} \cdot \left(1 - \dfrac{k_i}{T_{ai}}\tilde{T}_i^{\text{лин}}\right)\dfrac{T_{ai}}{T_{ai} - T_i^{\text{лин}}} \\ \qquad \text{при } \tilde{T}_i^{\text{лин}} + t_{ci}(M_i) < t \le T_{ai} + t_{ci}(M_i) \\ 0 \text{ при } t < t_{ci}(M_i);\ t \le T_{ai} + t_{ci}(M_i) + 1, \end{cases} \quad (16)$$

отражающем начисление амортизации в первые $\tilde{T}_i^{\text{лин}}$ месяцев по линейному методу без изменения нормы амортизации ($P_{it}^{**}(\cdot) = 1$). Далее происходит списание по нелинейному методу с коэффициентом ускорения k_i и нормой амортизации $1/(T_{ai} - \tilde{T}_i^{\text{лин}})$. При этом в качестве первоначальной стоимости основных средств для нелинейного метода берется остаточная стоимость на момент переключения. Это отражается в выражении (16) поправочным коэффициентом $(1 - k_i\tilde{T}_i^{\text{лин}}/T_{ai})$. По истечении срока полезного использования отчисления прекращаются ($P_{it}^{**}(\cdot) = 0$).

Суммируя изложенное, с учетом выражений (7), (8) и (13) получим

$$Am_t(\cdot) = \sum_{i=1}^{n} I_i(z_iP_{it}^{\text{нел}}(k_i, t_{ci}(M_i), T_{ai}, X_i)U_{it}(t_{ci}(M_i), T_{ai}) + $$
$$+ (1 - z_i)k_iP_{it}^{\text{лин}}(k_i, t_{ci}(M_i), T_{ai}, Y_i)/T_{ai}). \quad (17)$$

Рассмотрим теперь способы формализованного описания учетной политики компании (выбор метода расчета прибыли по отгрузке или оплате (z_h) и условий договоров (отсрочек платежей Δt_{itB}, Δt_{itC})) и ее влияние на эффективность функционирования. В общем виде чистая прибыль компании $ЧП_t$ в t-й период времени определяется по формуле:

$$ЧП_i = (R_t - Am_t - НИ_t)(1 - P_t), \quad (18)$$

где R_t — разница между притоком и оттоком денежных средств по всем действующим объектам компании; Am_t — амортизационные отчисления в этот период по этим объектам; $НИ_t$ — налог на имущество, все в t-й период; P_t — логическая функция ставки налога на прибыль, описываемая выражением:

$$P_t = \begin{cases} p \text{ при } (R_t - Am_t - НИ_t) \ge 0 \\ 0 \text{ при } (R_t - Am_t - НИ_t) < 0. \end{cases}$$

Выручка R_t рассчитывается по выражению (для упрощения предполагается, что в каждый период продукция полностью либо отгружается в кредит, либо по предоплате, такое же предположение относится и к оплате затрат, понесенных при производстве и реализации этой продукции): $R_t = \sum_{i=1}^{n} r_{it}$, где $r_{it} = B'_{it} - C'_{it}$, B'_{it} — поступление выручки в t-й период за ранее отгруженную продукцию, C'_{it} — отток средств в t-й период по оплате расходов, понесенных ранее при производстве продукции (без амортизации и налога на имущество) по i-му объекту[2]. Приток денежных средств (погашение обязательств) B'_{it} рассчитывается по выражению:

$$B'_{it} = (1 - z_h) \sum_{k = t_c}^{t_{ci} + T_{ai} - 1} B_{ik}(1 + s(\Delta t_{ikB}))\varphi(t, \Delta t_{ikB}) + $$
$$+ z_hB_{it}(1 + s(\Delta t_{itB})), \quad (19)$$

где B_{ik} — возникновение обязательств по оплате продукции, отгруженной в k-й период i-м объектом, $s(\Delta t_{ikB})$ — функция надбавок/скидок к цене продукции за ее отгрузку в кредит/предоплату, зависящая от отсрочек платежей Δt_{itB}, предоставленных потребителям в k-й период Δt_{itB} по i-му объекту; а схема поступления выручки описывается логической функцией

$$\varphi(t, \Delta t_{ikB}) = \begin{cases} 1 - |\{\Delta t_{ikB}\}| \text{ при } t = k + [\Delta t_{ikB}] \\ |\{\Delta t_{ikB}\}| \text{ при } t = k + [\Delta t_{ikB}] + 1 \\ 0 \text{ в других случаях.} \end{cases}$$

[2] При необходимости, потоки B_{it} и C_{it} можно представить в виде сумм потоков по отдельным договорам: $B_{it} = \Sigma B_{itj}$, $C_{it} = \Sigma C_{itl}$, где каждое слагаемое представляет собой объем средств, время поступления (уход со счета) которых можно изменять в некоторых пределах относительно t в соответствии с условиями конкретного договора (суммирование производится по всем j-м и l-м договорам на отгрузку продукции и получению сырья на t-й период по i-му объекту соответственно.) В этом случае для каждого договора дополнительно вводится своя переменная $\Delta t_{tij/l}$.

При выборе метода расчета «по оплате» ($z_h = 0$) наличие дробной части $\{\Delta t_{ikB}\}$ у функции $\varphi(t, \Delta t_{ikB})$ показывает долю отгрузки B_{ik}, выручка по которой должна поступить в $(k + [\Delta t_{ikB}] + 1)$-й месяц (квадратные скобки означают целую часть числа). Остальная часть выручки от продукции, отгруженной в k-й месяц, поступит в зависимости от знака Δt_{ikB} в предыдущий (следующий) $(k + [\Delta t_{ikB}])$-й месяц.

Отток средств C'_{it} описывается выражением, аналогичным выражению (19):

$$C'_{it} = (1 - z_h) \sum_{k = t_{ci}}^{t_{ci} + T_{ai} - 1} C_{ik}(1 + f(\Delta t_{ikC}))\varphi(t, \Delta t_{ikC}) + \\ + z_h C_{it}(1 + f(\Delta t_{itC})), \quad (20)$$

где C_{it} — возникновение обязательств компании по оплате расходов, понесенных при изготовлении продукции в t-й период i-м объектом (за вычетом амортизации и налога на имущество); $f(\Delta t_{ikC})$ — функция штрафа за просроченные платежи, зависящие от отсрочки платежей Δt_{ikC}, предоставленной i-му объекту в k-й период.

На практике функции надбавок/скидок $s(\Delta t_{itB})$ и штрафов $f(\Delta t_{itC})$, учитываемые в формулах (19) и (20), носят нелинейный характер и их вид может различаться для каждого i-го объекта. Для удобства изложения предполагается, что эти функции симметричны по аргументу, т. е. размер скидок за предоплату и отгрузку в кредит с отсрочкой на одно и то же число дней одинаков. Вид функций $s(\Delta t_{itB})$ и $f(\Delta t_{itC})$ приведен на рис. 1.

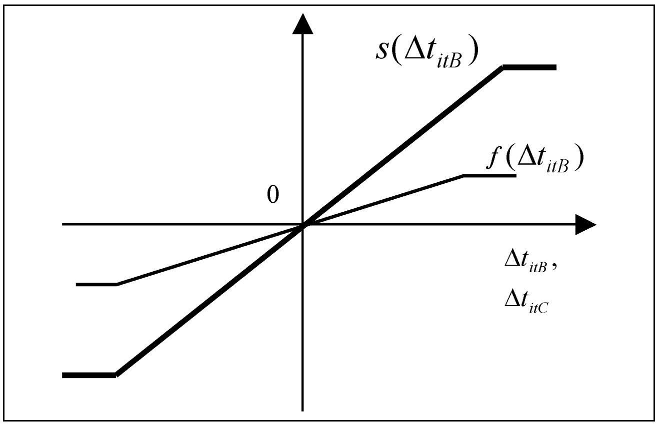


Рис. 1. Вид функций надбавок/скидок и штрафов

Условия договоров влияют на прирост оборотного капитала, входящего в расчет целевой функции (6). При выборе метода учета «по оплате» ($z_h = 0$) оборотный капитал компании не меняется ($I_{\text{об}\,t} = 0$). В случае выбора метода учета «по отгрузке» ($z_h = 1$) капитал изменится на величину:

$$I_{\text{об}\,t} = z_h \sum_{i=1}^{n} \{B_{it}(1 + s(\Delta t_{itB})) - C_{it}(1 + f(\Delta t_{itC})) - \\ - \sum_{k = t_{ci}}^{t_{ci} + T_{ai} - 1} [B_{ik}(1 + s(\Delta t_{ikB}))\varphi(t, \Delta t_{ikB}) - \\ - C_{ik}(1 + f(\Delta t_{ikC}))\varphi(t, \Delta t_{ikC})]\}. \quad (21)$$

Первые два члена под знаком суммирования по n объектам есть объём возникших в t-й период обязательств по оплате продукции покупателями и затрат, понесенных компанией, которые вошли в расчет прибыли. Третье слагаемое (сумма по периодам k) отражает погашение в t-м периоде ранее возникших обязательств, т. е. поступление выручки и оплату затрат.

Размер налога на имущество в уравнении (18) в t-й период определяется по формуле:

$$НИ_t = \alpha \sum_{i=1}^{n} НИ_i(z_i, k_i, t_{ci}(M_i), T_{ai}, X_i, Y_i, t) \times \\ \times P_i^{НИ}(t, t_{ci}(M_i), T_{ai}) U_i^{НИ}(t),$$

где α — годовая ставка налога на имущество; $НИ_i(z_i, k_i, t_{ci}(M_i), T_{ai}, X_i, Y_i, t)$ — налогооблагаемая база; $P_i^{НИ}(t, t_{ci}(M_i), T_{ai})$ — функция, ограничивающая период начисления налога на имущество по i-му объекту сроком его эксплуатации (полезного использования); $U_i^{НИ}(t)$ — функция, отражающая ежеквартальный характер начисления налога на имущество. При этом считается, что если объект к концу срока эксплуатации полностью не амортизировался, то он продается по остаточной стоимости и налог на имущество с этого момента не начисляется, т. е.

$$P_i^{НИ}(t, t_{ci}(M_i), T_{ai}) = \begin{cases} 1 \text{ при } t_{bi} \le t < t_{ci}(M_i) + T_{ai} \\ 0 \text{ при } t < t_{bi}, t \ge t_{ci}(M_i) + T_{ai}, \end{cases} \quad (21')$$

где $t_{bi} = t_{ci}(M_i) - \Delta t_{bi}$ — момент постановки на баланс i-го объекта. Функция $U_i^{НИ}(t)$ имеет вид:

$$U_i^{НИ}(t) = \begin{cases} 1 \text{ если } (t)\bmod 3 = 0 \\ 0 \text{ в противном случае,} \end{cases}$$

где mod3 — остаток от деления на 3. Как видно, функция $U_i^{НИ}(t)$ позволяет начислять налог на имущество ежеквартально, т. е. когда номер месяца горизонта планирования кратен трем. Налогооблагаемая база рассчитывается как среднеквартальная стоимость основных средств по формуле:

$$НИ_i(z_i, k_i, t_{ci}(M_i), T_{ai}, X_i, Y_i, t) = \\ = I_i - 0{,}5 \cdot \left(\sum_{\tau=1}^{t} Am_{i\tau}(\cdot) + \sum_{\tau=1}^{t-3} Am_{i\tau}(\cdot) \right). \quad (22)$$

Последнее слагаемое Ost_t, входящее в числитель целевой функции (6), представляет собой остаточную стоимость объектов, которые не полностью амортизировались по истечении сроков их полезного использования[3].

[3] Такое может произойти, если по каким-либо причинам не произошло переключения в нелинейной схеме списания.

Пусть остаточная стоимость i-го объекта определяется по формуле: $Ost_i = I_i - \sum_{t=1}^{T_a} Am_t(\cdot)$.

Введем функцию $L_{it}(t_{ci}(M_i),\ T_{ai})$, принимающую значение 1, если i-й объект закончил функционирование на $(t - 1)$-м шаге, и 0 в противном случае:

$$L_{it}(t_{ci}(M_i),\ T_{ai}) = \begin{cases} 1 \text{ при } t = t_{ci}(M_i) + T_{ai} \\ 0 \text{ при } t \neq t_{ci}(M_i) + T_{ai}. \end{cases}$$

Тогда суммарная остаточная стоимость объектов, прекративших функционирование на t-м шаге, будет определяться выражением:

$$Ost_t = \sum_{i=1}^{n} Ost_i \cdot L_{it}, \qquad (23)$$

где Ost_i рассчитывается по формуле (12).

2. ОБЩАЯ ПОСТАНОВКА ЗАДАЧИ ОПТИМИЗАЦИИ УЧЕТНОЙ ПОЛИТИКИ

С учетом формализованных выражений (4)—(23) общая постановка задачи оптимизации (1)—(3) конкретизируется следующим образом. После цепочки подстановок критерий задачи принимает вид:

$$F = \sum_{t=1}^{T^*} \Bigg\{ (1 - P_t(\cdot)) \cdot \sum_{i=1}^{n} [r_{it}(z_h,\ \Delta_{ikB},\ \Delta_{ikC}) -$$
$$- \alpha I_t P_i^{НИ}(t)\, U_i^{НИ}(t) \Bigg[1 - 0{,}5 \cdot \Bigg(\sum_{\tau=1}^{t} (z_i P_{i\tau}^{нел}(\cdot) U_{it} +$$
$$+ (1 - z_i) k_i P_{i\tau}^{лин}(\cdot)/T_{ai}) + \sum_{\tau=1}^{t-3} (z_i P_{i\tau}^{нел}(\cdot) U_{it} + (1 - z_i) \times$$
$$\times k_i P_{i\tau}^{лин}(\cdot)/T_{ai}) \Bigg) \Bigg] + P_t(\cdot) I_i \sum_{i=1}^{n} (z_i P_{i\tau}^{нел}(\cdot) U_{it} +$$
$$+ (1 - z_i) k_i P_{i\tau}^{лин}(\cdot)/T_{ai}) - I_t(M) - I_t^{об}(\cdot) +$$
$$+ \sum_{i=1}^{n} Ost_i(\cdot) L_{it}(\cdot) \Bigg\} \times \frac{1}{(1 + \delta)^t} \to \max,$$

где $I_t(M)$ определяется по выражению (6′), исходя из искомой динамики вложений денежных средств в объекты компании. В системе ограничений $G()$ задачи (1)—(3) выделяются следующие основные условия:

— ограниченности сверху и снизу коэффициентов ускорения в зависимости от выбираемого метода амортизации

$$z_i k_i + (1 - z_i) k_i - z_i \overline{K}_i^{нел} - (1 - z_i)\, \overline{K}_i^{лин} \leqslant 0,$$

$$z_i k_i + (1 - z_i) k_i - z_i \underline{K}_i^{нел} - (1 - z_i)\, \underline{K}_i^{лин} \geqslant 0,\ i = \overline{1, n},$$

где $\underline{K}_i^{лин}$, $\underline{K}_i^{нел}$ и $\overline{K}_i^{лин}$, $\overline{K}_i^{нел}$ — соответственно нижние и верхние границы переменных k_i;

— ограниченности снизу и сверху сроков полезного использования, причем нижняя граница не должна быть менее 12 месяцев, $12 \leqslant \underline{T_{ai}} \leqslant T_{ai} \leqslant \overline{T_{ai}}$, $i = \overline{1, n}$, где $\underline{T_{ai}}$ и $\overline{T_{ai}}$ — нижние и верхние границы переменных T_{ai};

— ограниченности сверху и снизу отсрочек платежей по оплате продукции и затрат $\underline{\Delta t}_{itB} \leqslant \Delta t_{itB} \leqslant \overline{\Delta t}_{itB}$, $\underline{\Delta t}_{itC} \leqslant \Delta t_{itC} \leqslant \overline{\Delta t}_{itC}$, $i = \overline{1, n}$; $t = \overline{1, T_p}$, где $\underline{\Delta t}_{itB}$, $\underline{\Delta t}_{itC}$ и $\overline{\Delta t}_{itB}$, $\overline{\Delta t}_{itC}$ — соответственно нижние и верхние границы переменных Δt_{itB} и Δt_{itC};

— ограниченности переменных, описывающих условия переключения с метода на метод $0 \leqslant Y_i \leqslant 1$, $0 \leqslant X_i \leqslant 1$ $i = \overline{1, n}$;

— обязательности ввода i-го объекта в эксплуатацию на горизонте планирования $\sum_{t=1}^{T} m_{ti} = 1$, $i = \overline{1, n}$;

— целочисленности переменных, отражающих ввод объектов в эксплуатацию, выбор метода амортизации и учетной политики $m_{ti} \in \{0; 1\}$, $i = \overline{1, n}$, $t = \overline{1, T}$; $z_i \in \{0; 1\}$, $i = \overline{1, n}$; $z_h \in \{0; 1\}$;

— неотрицательности коэффициентов ускорения $k_i > 0$, $i = \overline{1, n}$;

— переменные, описывающие задержки в платежах, могут принимать как неотрицательные, так и отрицательные значения $-\infty \leqslant \Delta t_{itB}, \Delta t_{itC} < +\infty$, $i = \overline{1, n}$; $t = \overline{1, T_p}$;

— общего вида (на прибыль, финансовые потоки, рентабельность и т. п.), сформированные с учетом соотношений (18)—(23) и более подробно рассмотренных в работе [3]: $G'(Z, K, \Delta t_B, \Delta t_C, M, z_h, X, Y, T_a) = 0$.

Таким образом, сформулированная задача относится к классу задач негладкой оптимизации с непрерывными и булевыми переменными и алгоритмически задаваемыми функциями и ограничениями. При этом булевы переменные входят в основном в кусочно-гладкие зависимости второго и третьего порядков, а непрерывные — в виде гладких степенных функций.

Получение глобального решения задач такого класса современными методами не гарантируется, а получение за приемлемое время решений, достаточно близких к оптимальным, не всегда возможно. Общий подход к решению задач такого класса заключается в применении методов оптимизационно-имитационного моделирования сложных систем, предполагающих построение взаимосвязанных оптимизационных (учитывающих аналитически задаваемые зависимости) и имитационных (отражающих алгоритмические зависимости) моделей и организацию итеративных процедур их взаимодействия [4, 5].

На практике общая задача оптимизации учетной политики естественным образом распадается на две взаимосвязанные последовательно решаемые подзадачи: формирования амортизационной политики (связанной, в первую очередь, с инвестиционным планированием) и формирования налоговой и договорной политики (влияющей на результаты текущей операционной деятельности). Формально такое разделение общей модели (1)—(3) осуществляется путем ее эвристической декомпозиции по переменным и ограничениям. Модели

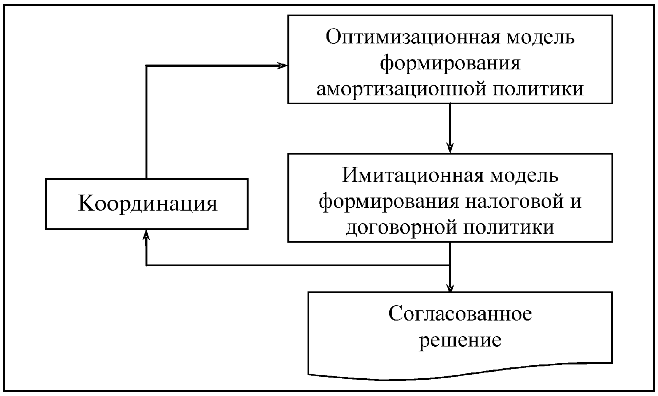


Рис. 2. Итерационная процедура взаимодействия оптимизационной и имитационной моделей

(меньшей размерности), полученные в результате декомпозиции, объединяются затем в последовательную итерационную процедуру с пошаговой координацией (рис. 2), которая обеспечивает формирование приемлемого для ЛПР решения общей задачи. Отметим, что в частных случаях решение задачи (1)—(3) в единой постановке поддается методами гладкой оптимизации (см. например, работу [3]).

В качестве примера рассмотрим формализованную постановку задачи оптимизации амортизационной политики, полученную в результате декомпозиции общей задачи. Пусть для определенности выбираемый метод амортизации для объекта не может быть изменен в течение всего срока его эксплуатации (т. е. отсутствует возможность оптимизации моментов переключения: $X = X_0 \leqslant 1$, $Y = 1$); сроки начала эксплуатации всех объектов заданы, а возможности изменения сроков оборачиваемости при расчетах с покупателями (поставщиками) и выбора учетной политики по оплате и отгрузке не рассматриваются, поскольку решаются в процессе имитационного моделирования (см. рис. 2). Это означает, что в общей постановке (1)—(3) фиксируются переменные z_h, Δt_B, Δt_C, M, X и Y и исключаются все связанные с ними ограничения. Тогда исходная задача (1)—(3) представляется в следующем компактном виде:

$$F(Z,\ K,\ T_a) \to \text{extr}; \qquad (24)$$

$$G(Z,\ K,\ T_a) = 0, \qquad (25)$$

где целевая функция (с учетом того, что выбор амортизационной политики влияет только на первые два слагаемых выражения (6)) приобретает вид:

$$F(z,\ k,\ T_a) = \sum_{T=1}^{T_p} \frac{ЧП_t + Am_t}{(1+\delta)^t} =$$

$$= \sum_{t=1}^{T_p} \frac{(R_t - НИ_t(k,\ T_a))\cdot(1 - P_t(\cdot)) + Am_t(k,\ T_a)P_t(\cdot)}{(1+\delta)^t},$$

причем, значения $НИ_t$ и Am_t рассчитывается по формулам (22) и (17) при фиксированных значениях переменных z_h, Δt_B, Δt_C, M и $X = X_0 \leqslant 1$, $Y = 1$. Группа ограничений (25) конкретизируется в виде следующей системы:

$$z_i k_i + (1 - z_i)k_i - z_i \overline{K}_i^{\text{нел}} - (1 - z_i)\,\overline{K}_i^{\text{лин}} \leqslant 0,$$

$$z_i k_i + (1 - z_i)k_i - \underline{K}_i^{\text{нел}} - (1 - z_i)\,\underline{K}_i^{\text{лин}} \geqslant 0,$$

$$\underline{T_{ai}} \leqslant T_{ai} \leqslant \overline{T_{ai}},$$

$$z_i \in \{0;\ 1\},\ i = \overline{1, n};\ t = \overline{1, T_p}.$$

Заметим, что на практике возможны различные содержательные постановки задачи (24), (25) — от выбора коэффициентов ускорения при заданных методах амортизации до определения всех необходимых составляющих амортизационной политики (методы амортизации, сроки полезного использования, коэффициенты ускорения и т. д.). Рассмотрим некоторые практические примеры ее численного решения.

3. ПРИМЕРЫ ЧИСЛЕННОГО РЕШЕНИЯ ЗАДАЧИ ОПТИМИЗАЦИИ АМОРТИЗАЦИОННОЙ ПОЛИТИКИ

Для иллюстрации полученных результатов рассмотрим два примера численного решения задачи (24), (25) оптимизации амортизационной политики. Пусть методы амортизации и сроки ввода в эксплуатацию для каждого объекта фиксированы, возможность их переключения не предусматривается (переменный вектор M_i и переменные Z_i, X_i и Y_i фиксируются для каждого объекта: $M_i = M_i^0$, $z_i = z_i^0$ и $X_i = 0{,}8$, $Y_i = 1$) и требуется определить только сроки полезного использования и коэффициенты ускоренного списания первоначальной стоимости объектов основных средств. Тогда задача оптимизации приобретает вид:

$$F(K,\ T_a) \to \text{extr}; \qquad (26)$$

$$G(K,\ T_a) = 0. \qquad (27)$$

Целевая функция (26) конкретизируется следующим образом:

$$F = \sum_{t=1}^{T^*}\Bigg(\sum_{i=1}^{N}\Bigg[B_{it} - C_{it} - I_t\cdot\Bigg\{1 - 0{,}5\cdot\sum_{\tau=1}^{t}(z_i^0 P_{i\tau}^{\text{нел}}(\cdot)U_{i\tau} +$$

$$+ (1 - z_i^0)k_i P_{i\tau}^{\text{лин}}(\cdot)/T_{ai}) + \sum_{\tau=1}^{t-3}(z_i^0 P_{i\tau}^{\text{нел}}(\cdot)U_{i\tau} +$$

$$+ (1 - z_i^0)k_i P_{i\tau}^{\text{лин}}(\cdot)/T_{ai})\Bigg)\Bigg\}\cdot P_i^{НИ}(t)\,U_i^{НИ}(t)\alpha\Bigg] \times$$

$$\times (1 - P_t(\cdot)) + P_t(\cdot)\sum_{i=1}^{n} I_i(z_i^0 P_{i\tau}^{\text{нел}}(\cdot)U_{it} +$$

$$+ (1 - z_i^0)k_i P_{i\tau}^{\text{лин}}(\cdot)/T_{ai})\Bigg)\cdot\frac{1}{(1+\delta)^t} \to \max, \qquad (28)$$

где функции $P_{i\tau}^{\text{нел}}$, $P_{i\tau}^{\text{лин}}(\cdot)$, U_{it} и $P_i^{НИ}(t)$ рассчитываются по выражениям (11), (15), (9) и (21′) с учетом, что $M_i = M_i^0$, $z_i = z_i^0$ и $X_i = 0{,}8$, $Y_i = 1$, а система ограничений

(27) (без учета ограничений общего вида) включает в себя неравенства

$$\underline{K_i} \leqslant k_i \leqslant \overline{K_i}, \quad \underline{T_{ai}} \leqslant T_{ai} \leqslant \overline{T_{ai}}, \quad i = \overline{1, n}. \qquad (29)$$

Видно, что задача (28), (29) относится к классу задач кусочно-квадратичной оптимизации с линейными ограничениями. Отметим, что каждую пару переменных k_i и T_{ai} в постановках (24), (25) и (26), (27) можно заменить на одну переменную kd_i. Это возможно, поскольку переменные k_i и T_{ai} входят в расчет величин $НИ_t$ и Am_t только в виде множителей k_i/T_{ai}. Тогда, заменив каждую пару неравенств (29) на одно неравенство

$$\underline{k}_i/\overline{T}_{ai} \leqslant kd_i \leqslant \overline{k}_i/\underline{T}_{ai}, \qquad (30)$$

получим кусочно-гладкую задачу (28), (30), эквивалентную задаче (28), (29), меньшей размерности и с многочленами относительно kd_i степени T^*.

В рамках такой (относительно простой) постановки удобно исследовать степень влияния решения (выбора коэффициентов ускорения и периода амортизации) на значение целевой функции. На рис. 3 (см. третью страницу обложки) для двух новых объектов компании представлен трехмерный график зависимости чистого дисконтированного дохода $NPV(kd_1, kd_2)$, рассчитанной по формуле (28) при условии, что объекты начинают функционировать одновременно, амортизируются по линейному методу ($z_1 = z_2 = 0$ и $Y_1 = Y_2 = 1$) и имеют достаточно высокую доходность. При этом переменные kd_1 и kd_2 варьируются в пределах от 0,01 до 0,05 с шагом 0,001 (что соответствует $1 \leqslant k \leqslant 4$; 80 мес. $\leqslant T_a \leqslant 100$ мес.). Значения переменных по оси абсцисс отображены в масштабе 1 : 100. Из рис. 3 видно, что достаточно высокая доходность объектов обеспечивает рост целевой функции с ростом kd_i до некоторых пределов. Однако применение очень высоких коэффициентов ускорения или резкое сокращение сроков полезного использования объектов приводит к снижению эффективности функционирования компании. Это связано с тем, что из-за высоких амортизационных отчислений в первые периоды работы объектов компания понесет такие большие убытки, что они не перекроются (с учетом дисконтирования) все возрастающими прибылями.

Представленная на рис. 3 ситуация отображает один из самых простых случаев задачи (26)—(27), когда целевая функция в критерии (28) имеет симметричную, «почти» гладкую и выпуклую структуру. Однако, в зависимости от параметров объектов, функционал в критерии (28) может иметь сложную структуру и в общем случае не является выпуклым. Такая ситуация проиллюстрирована на рис. 4 (см. третью страницу обложки), на котором представлена поверхность целевой функции для компании, состоящей из двух объектов, один из которых начинает функционировать на три года позже другого и амортизируется по нелинейному методу ($z_1 = 0$, $z_2 = 1$, $X_1 = 0,8$ и $Y_2 = 1$). Пределы варьирования переменных kd_1 и kd_2 оставлены теми же, что в предыдущем случае (от 0,01 до 0,05). При этом, как видно из рис. 4, наблюдается положительный скачок в изменении значения функционала при достижении высоких значений kd_1 и kd_2. Этот эффект связан с резким ростом амортизационных отчислений в момент переключения с нелинейного на линейный метод для второго объекта и их дополнительным приращением в последний месяц эксплуатации первого объекта, амортизируемого по линейному методу, которое совпало с моментом переключения.

Рассмотрим пример решения более сложной задачи типа (24), (25). Пусть для каждого из $N = 10$ взаимосвязанных объектов при фиксированных сроках их полезного использования $T_{ai} = T_{ai}^0$, $i = \overline{1, 10}$ требуется выбрать метод амортизации (линейный или нелинейный) и коэффициенты ускорения. При этом верхние границы $\overline{K}^{\text{нел}}$ и $\overline{K}^{\text{лин}}$ изменения коэффициентов ускорения для каждого объекта индивидуальны и меняются в зависимости от выбора линейного или нелинейного метода амортизации, а также доходности объектов $\bar{r}$, а нижние — одинаковы для всех и равны 1 ($\underline{K}^{\text{нел}} = \underline{K}^{\text{лин}} = 1$, что соответствует амортизации без ускорения). Оптимальная амортизационная политика формируется на 12-летнем горизонте T в месячном разрезе, причем $T_p = 138$ мес. Ставка дисконтирования δ принята на уровне 15 % годовых и применяется к оценке всех объектов. В соответствии с налоговым кодексом РФ нелинейная схема может переключаться на линейную в квартале, следующем за кварталом, в котором произошло списание 80 % стоимости объекта ($X_0 = 0,8$) [6]. Ставка налога на прибыль принята на уровне $p = 24\%$, а налога на имущество — 0,5 % в квартал из расчета $\alpha = 2$ % годовых. Каждый i-й объект характеризуется своей нормой доходности (прибыли) r_i, моментом начала эксплуатационной фазы t_{ci}, совпадающим с моментом постановки на баланс t_{bi} ($t_{ci} = t_{bi}$), и объёмом инвестиций I_i. Суммарная стоимость объектов компании составляет 662 млн. у. е. В качестве критерия оптимальности выступает максимизация суммарного чистого дисконтированного дохода, генерируемого всей совокупностью объектов:

$$NPV(z_i, k_i, i = \overline{1, 10}) \to \max, \qquad (31)$$

$$z_i k_i + (1 - z_i)k_i - z_i \overline{K}_i^{\text{нел}} - (1 - z_i)\overline{K}_i^{\text{лин}} \leqslant 0, \quad i = \overline{1, 10}; \qquad (32)$$

$$z_i k_i + (1 - z_i)k_i - z_i \underline{K}_i^{\text{нел}} - (1 - z_i)\underline{K}_i^{\text{лин}} \geqslant 0, \quad i = \overline{1, 10}; \qquad (33)$$

$$z_i \in \{0; 1\}, \quad i = \overline{1, 10}. \qquad (34)$$

Размерность сформированной негладкой частично-целочисленной задачи (31)—(34) составляет 20 переменных (из них 10 булевых) при 20-ти ограничениях. Входные параметры задачи при различных сочетаниях исходных приведены в табл. 1. Для решения задачи применялся эволюционный метод со следующими характеристиками: размер популяции — 100 элементов, вероятность мутации — 0,075, процедура отбора лучших решений из текущей популяции — равновероятный случайный выбор. Время поиска решения задач рассматриваемого класса, включая время генерации модели, существенно зависит от выбора начальной точки. Поэтому

Таблица 1

Входные данные задачи (26), (27)

i	$t_{bi} = t_{ci}$, мес.	Доходность r_i, % годовых			T_{ai}, мес.	d_i, % годовых	I_i, млн. у. е.	$\bar{K}_i^{лин}$		$\bar{K}_i^{нел}$
		средняя	низкая	высокая				1*	2**	
1	1	15,0	10,0	20,0	120	10	20	1,5	1,75	6
2	2	7,5	5,0	10,0	105	11,4	30	1,2	1,25	3
3	4	30,0	20,0	40,0	115	10,4	40	1,5	1,75	6
4	3	22,5	15,0	30,0	125	9,6	10	1,2	1,25	3
5	4	11,3	7,5	15,0	110	10,9	90	1,5	1,75	6
6	2	18,8	12,5	25,0	130	9,2	100	1,2	1,25	3
7	7	26,3	17,5	35,0	100	12,0	210	1,5	1,75	6
8	4	13,5	9,0	18,0	135	8,9	12	1,2	1,25	3
9	1	16,5	11,0	22,0	130	9,2	100	1,5	1,75	6
10	3	20,3	13,5	27,0	120	10,0	50	1,2	1,25	3
Среднее значение		19,9	13,3	26,5	114,7	10,6	Итого = 662	—		

* 1 — верхние границы при средней доходности.
** 2 — верхние границы при низкой и высокой доходностях.

процедура решения включала в себя также ряд последовательных принудительных рестартов, осуществляемых пользователем с достигнутых на предыдущем шагах рекордов. При этом на первом шаге процедуры за приемлемое время находилось достаточно хорошее первое рекордное решение, с которого затем производился рестарт, позволявший улучшить достигнутый рекорд и т. д. В качестве исходной начальной точки принималась точка, в которой все объекты амортизировались по линейному методу без ускорения ($z_i = 0$, $k_i = 1$, $i = \overline{1, 10}$).

Результаты решения задачи на трех уровнях усредненной по совокупности объектов доходностей (среднем ($\bar{r}_1$ = 19,9 %), низком ($\bar{r}_2$ = 13,3 %) и высоком ($\bar{r}_3$ = 26,5%)) при средней норме линейной амортизации $\bar{d}$ = 10,6 % отражены в табл. 2. В таблице приведены: рекордное решение (z, k), значение функционала в рекордной точке и его прирост относительно значения в начальной точке; значение функционала в начальной точках; оценки решений при максимально допустимых значениях коэффициентов ускорения; число достигнутых промежуточных рекордов и потребовавшихся для их достижения рестартов процедуры поиска и эвристических корректировок промежуточных рекордов. Эвристические корректировки применялись в случае отсутствия улучшения рекорда в течение длительного времени (1000—1500 итераций генетического алгоритма). Так, из табл. 2 видно, что в компании, собирающейся эксплуатировать объекты средней доходности $\bar{r}_1$, для двух объектов (6-го и 7-го) выбран нелинейный метод с коэффициентами ускорения 2,995 и 3,047, лежащими между границами изменения коэффициентов, а для остальных — линейный метод с максимально возможными коэффициентами 1,2 и 1,5. Максимально достигнутый рекорд (574,56) улучшил начальное значение функционала (553,541) на 3,8%.

Динамика изменения входящих в расчет критериального показателя $NPV(\cdot)$ финансовых потоков (валового дохода R_t, чистой прибыли $ЧП_t$, амортизации Am_t, налога на имущество $НИ_t$ и т. д., см. формулу (28)) в рекордной точке отражена на рис. 5 (с м. третью страницу обложки). Видно, что в первые два года горизонта планирования компания несет убытки ($ЧП_t < 0$), которые тем не менее компенсируются на последующих шагах существенным приростом денежных потоков за счет высоких амортизационных отчислений.

В целом проведенные модельные расчеты показали, что оптимизация амортизационной политики для объектов со средней и высокой доходностями за 2—4 рестарта дает улучшение по функционалу до 5 % относительно решения в начальной точке.

Отметим, что на структуру и ход решения задачи существенное влияние оказывают как доходность объектов r_i, так и пределы изменения коэффициентов ускорения. Так, рост усредненной доходности объектов до уровня $\bar{r}_3$ и выше приводит к тому, что предпочтительным оказывается выбор нелинейных методов с максимально возможными коэффициентами ускорения, превышающими пределы их изменения для линейных методов, причем в этой ситуации могут потребоваться эвристические корректировки рекордов, достигнутых вблизи границы области определения задачи (серия расчетов с высокой доходностью, табл. 2). Падение усредненной доходности до уровня $\bar{r}_2$ ведет к выбору линейных методов с коэффициентами ускорения, лежащими между допустимыми границами их изменения и стремящимися к нижним границам при резком падении показателя $\bar{r}_2$ (серия расчетов с низкой доходностью, табл. 2).

Расчеты проводились в среде MS Excel 2000 с помощью оптимизационной платформы PSP фирмы «Frontline Systems, Inc.» (США) [7].

Как отмечалось ранее, целью решения задачи оптимизации амортизационной политики в частности и учетной политики в целом является определение условий и параметров процесса финансового планирования,

Таблица 2

Результаты решения задачи (26), (27)

	Доходность, %					
	19,9 (средняя)		13,3 (низкая)		26,5 (высокая)	
	z_i	k_i	z_i	k_i	z_i	k_i
Решение (z, k) (справа) и его параметры (внизу)	0	1,5	0	1,258	0	1,75
	0	1,2	0	1,107	1	3
	0	1,5	0	1,248	1	6
	0	1,2	0	1,199	1	3
	0	1,5	0	1,314	0	1,75
	1	2,99	0	1,243	1	3
	1	3,04	0	1,187	0	1,75
	0	1,2	0	1,249	1	3
	0	1,5	0	1,448	0	1,75
	0	1,2	0	1,236	1	3
Рекорд *NPV*, млн. у. е.	574,56		392,13		752,187	
NPV в начальной точке $(z, k) = (0, 1)$, млн. у. е.	553,541		383,127		720,239	
Прирост относительно начальной точки, %	3,80		2,35		4,44	
Оценки *NPV*, млн. у. е. при						
$(z, k) = (0, \underline{K})$	560,578		375,336		741,409	
$(z, k) = (1, \overline{K})$	573,178		385,663		746,746	
Число промежуточных рекордов	3		3		6	
Число рестартов	2		3		4	
Число эвристических корректировок	1		—		2	
Суммарное число итераций	11 868		9 205		4 404	

который осуществляется с помощью соответствующих имитационных моделей бюджетирования (см. рис. 3 в первой части статьи [1]).

ЗАКЛЮЧЕНИЕ

В новых экономических условиях развития и функционирования реального сектора экономики России особую актуальность приобретает применение современных подходов к планированию и управлению хозяйственной деятельностью, одним из которых является контроллинг. Эффективное решение финансово-экономических задач контроллинга — действенный инструмент повышения оперативности и качества принимаемых управленческих решений. Для его внедрения в повседневную практику плановых расчетов необходима дальнейшая разработка соответствующих методов математического моделирования, адекватных современным условиям экономического окружения, и информационных технологий их реализации. Цель авторов статьи заключалась в привлечении внимания специалистов к данной проблеме и обсуждении возможных путей ее решения.

ЛИТЕРАТУРА

1. *Карибский А. В., Мишутин Д. Ю., Шишорин Ю. Р.* Финансово-экономические методы контроллинга в управлении хозяйственной деятельностью интегрированных компаний. Ч. I // Проблемы управления. — 2004. — № 2. — С. 54—62.
2. *Мишутин Д. Ю., Шишорин Ю. Р.* Моделирование и оптимизация в процессе бюджетирования интегрированной Компании // Тр. института ИПУ РАН. — М., 2004. — Т. XIII. — С. 135—149.
3. *Баулин И. С., Мишутин Д. Ю.* Различные формы задач оптимизации учетной политики интегрированных компаний // Информация и экономика: Сб. науч. тр. / АлГУ. — Барнаул, 2004 (в печати).
4. *Карибский А. В., Филиппов В. А.* Оптимизационно-имитационный подход к решению задач планирования развития крупномасштабных систем // Синтез и проектирование многоуровневых систем управления: Тр. науч.-техн. конф. / АлГУ (Барнаул, 1982). — Барнаул, АлГУ, 1982. — Ч. 2.
5. *Цвиркун А. Д., Акинфиев В. К., Филиппов В.А.* Имитационное моделирование в задачах синтеза структуры сложных систем (оптимизационно-имитационный подход). — М.: Наука, 1985.
6. *Premium* Solver. Premium Solver Platform Version 5.0 for use with Microsoft Excel. Quick Overview [Электронный ресурс] / Frontline Systems, Inc. — Incline Village, Nevada (USA): Frontline Systems, Inc., 2002. — Электрон. текст. — Режим доступа: http://www.solver.com/PlatformV5Intro.pdf, для зарегистрированных пользователей. — Загл. с экрана.
7. *Мишутин Д. Ю.* Информационная технология бюджетирования в интегрированных компаниях // Информационная экономика и управление динамикой сложных систем: Сб. науч. тр. — М.-Барнаул: Бизнес-Юнитек, 2004. — С. 124—132.

☎ *(095) 334-90-41*

E-mail: ysh@petrocom.ipu.rssi.ru

mishutin@petrocom.ipu.rssi.ru ❑

УДК 331.109."78".005

МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ ИССЛЕДОВАНИЯ КОНФЛИКТНЫХ СИСТЕМ ПРИ НЕОПРЕДЕЛЕННОСТИ

Л. В. Жуковская

Институт проблем управления им. В. А. Трапезникова, г. Москва

Предложены методы и технология исследования конфликтных микросистем с учетом действия рыночной неопределенности, а также формализация гарантированного по риску решения, обеспечивающего устойчивость (равновесность) системы на микроуровне.

ВВЕДЕНИЕ

Экономическая теория всегда уделяла большое внимание исследованию фирм. В данной работе под фирмами понимаются коммерческие образования (компании, предприятия, организации и т. д.) независимо от их организационно-правовой формы, отраслевой принадлежности и масштабов деятельности. Различные направления экономической теории изучают фирмы, каждое в соответствии со своей методологией.

Институциональная экономическая теория объясняет обусловленность экономических явлений и процессов существующими в обществе институтами: правилами, принципами и нормами поведения. В институциональной экономической теории поведение хозяйствующих субъектов ориентировано на соблюдение институтов и стремление к прибыли. Понятие «стремление к прибыли» тождественно понятию «максимизация прибыли» в условиях постоянной неопределенности и отсутствии изменений.

С другой стороны, развитие фирм как микросистем имеет в своей основе процессы самоорганизации. Фирму можно рассматривать как сложную открытую (по отношению к экономике в целом) микросистему со своим жизненным циклом и чередованием равновесных и неравновесных состояний, периодов подъемов и спадов и т. п. Флуктуации в данной системе возникают под воздействием неопределенности.

Неопределенность можно определить как неполноту или неточность информации об условиях реализации выбранного решения (альтернативы). В экономических системах (как микро-, так и макро-) можно выделить следующие виды неопределенности:

- неопределенности в оценках состояния экономической системы;
- неопределенность относительно будущих флуктуаций как внешних, так и внутренних;
- неопределенности по поводу адекватности самой модели (например, верны ли значения ее параметров).

Экономическая микросистема, как правило, подвергается неожиданным трудно прогнозируемым действиям *как извне* (кроме перечисленных):

- проявление стихийных сил природы;
- различного рода техногенные аварии;
- изменения экономической политики;
- изменение правовой среды;
- возникающие из-за колебаний спроса и предложения скачки спроса на товары, изменение числа и номенклатуры поставок, изменение закупочных цен, срыв поставок;
- недобросовестность, некомпетентность или несостоятельность партнеров, контрдействия конкурентов, терроризм, рэкет;
- появление и внедрение новых технологий,

так и *изнутри:*

- поломка, выход из строя, физическое и моральное старение средств производства и объектов инфраструктуры;
- непредусмотренные дополнительные затраты (например, потери сырья и энергии при хранении и транспортировки продукции);
- несчастные случаи на производстве и болезни сотрудников;
- ошибки в управлении персоналом;
- ошибки в маркетинге;
- ошибки планирования и проектирования и т. п.

Неопределенности возникают также и в процессе принятия экономических решений:

- неопределенность в понимании целей управляемого процесса — эти цели отчасти носят субъективный характер, так как формулируются ЛПР (лицом, принимающим решения);
- при построении математической модели сложной открытой микросистемы, так как учесть все множество ограничений (на неконтролируемые и контролируемые параметры), в рамках которых протекает процесс, просто не представляется возможным (при сложившемся уровне и методах научного познания);

- связи между переменными процесса в виде дифференциальных и (или) алгебраических уравнений не всегда возможно представить адекватно самому процессу.

Флуктуации, возникающие под действием как внешних, так и внутренних неопределенностей, могут какое-то время компенсироваться с помощью механизма обратных связей, и система, несмотря на колебания (объема выпуска, положения на рынке, уровня инвестиций и т. д.), остается в границах своего «коридора развития», возвращаясь к равновесию. Но при достижении критических значений возмущающих параметров происходит резкое усиление флуктуаций. Они нарастают, охватывая всю систему, все ее блоки (производство, финансы, сбыт, снабжение, организацию труда и т. д.). Далее возникают два принципиально возможных варианта развития: *первый* — когда система прекращает свое существование, выходя за пределы «коридора своего развития» («поля возможностей»), *второй* — когда система выходит на новую траекторию, выбирая альтернативы в границах нового поля возможностей. При этом кардинально меняется структура, внутренние связи, цели, фактически все основные параметры функционирования фирмы, а сам процесс развития выходит на новый виток. Характеристики описанного процесса существенно зависят от действий управляющего центра фирмы, т. е. от решений, сознательно принимаемых высшим и средним управляющим звеном организации. Также следует учесть, что сложное переплетение неопределенного и сознательного в процессе развития фирмы может определяться, с одной стороны, интенциональным характером (с точки зрения соображений полезности) действий людей на микроуровне, а с другой стороны, — нелинейностью, например, зависимостью от индивидуальных и коллективных взаимодействий и, наконец, ограничением рамками институциональной структуры, которая, в свою очередь, является результатом спонтанного взаимодействия людей (неформальные нормы, правила, стереотипы) и сознательного нормотворчества, закрепленного законом [1].

По мнению Н. Н. Воробьёва, «многие явления, если подходить к ним с точки зрения теории игр, проясняются весьма естественно» [2].

МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ ИССЛЕДОВАНИЯ КОНФЛИКТНЫХ МИКРОСИСТЕМ ПРИ НЕОПРЕДЕЛЕННОСТИ

Рассмотрим следующий пример. Предположим, что на рынке сбыта присутствуют $N \geqslant 2$ конкурирующих предприятий, выпускающих однородный бесконечно-делимый товар. Себестоимость единицы товара для i-го предприятия равна $c_i > 0$, $i \in \mathbb{N} = \{1, 2, ..., N\}$. Если K — общее количество товара на рынке, то цена товара $p = \max\{a - Kb; 0\}$, где постоянная $a > 0$ есть цена при отсутствии товара, постоянная $b > 0$ — коэффициент эластичности. Производственные мощности предприятия считаются неограниченными в этих условиях, и они независимо друг от друга выбирают количество производимого ими товара. Весь произведенный товар продается по цене p. Предполагается, что себестоимость $c_i < a$, $i \in \mathbb{N} = \{1, 2, ..., N\}$. *Цель каждого предприятия состоит в том, чтобы при увеличении объема выпуска товара получить максимальную прибыль от его продажи.* При этом производители должны учитывать возможность появления на рынке импортного товара в размере y, причем о значении y известны лишь возможные границы его изменения.

Один из подходов к решению таких типичных для рынка задач сводится к построению и исследованию математических моделей. Конфликтность данной ситуации заключается в независимом (от других фирм, присутствующих на рынке) выборе каждым ЛПР «своих» стратегий. Конкретно для приведенной конфликтной ситуации такая модель представляется бескоалиционной игрой N лиц при неопределенности, математическая модель которой задается упорядоченным набором

$$\langle \mathbb{N}, \{X_i\}_{i \in \mathbb{N}}, Y, \{f_i(x, y)\}_{i \in \mathbb{N}} \rangle. \quad (1)$$

Поясним каждый из элементов набора (1).

Игроки. Множество порядковых номеров игроков $\mathbb{N} = \{1, 2, ..., N\}$. В экономических микросистемах роль игроков могут выполнять руководители высшего и среднего звена управления, т. е. те (в рамках институциональной структуры, в которых функционируют данные предприятия), кто может принимать решения и отслеживать их исполнение. В сформулированной выше задаче игроками являются руководители предприятий — продавцы товара в условиях совершенно конкурентного рынка.

Стратегии и неопределенности. Для любого i-го игрока (институционально) определены возможные действия, каждое из которых называется стратегией i-го игрока. В теории игр под стратегией понимается правило, по которому каждому состоянию информированности игрока ставится в соответствие то или иное действие (поведение) из возможных действий (поведений), допустимых при данной информации. В экономических микросистемах стратегией отдельного игрока может быть стратегия прибыли, стратегия роста, стратегии, связанные с внедрением новых технологий, премии, штрафы и другие поощрения и наказания. В рассматриваемом примере стратегией i-го игрока является увеличение объемов выпуска, т. е. количество x_i производимого товара на i-м предприятии (при постоянной отдаче в росте прибыли), а множество таких стратегий есть $X_i = [0, +\infty)$.

В духе синергетических представлений о развитии сложных систем, принцип открытости предполагает, что взаимодействие с внешней средой происходит для всех подсистем, образующих сложную экономическую систему. Элементы этой системы — фирмы, рассматриваемые как микросистемы, обладают открытостью. Это одна из характеристик, свойственных сложным системам. Таким образом, открытость микроэкономической системы требует, чтобы при принятии решений учитывались как внешние, так и внутренние факторы неопределенности (описанные выше), которые существенно влияют на исход игры. Причем зачастую о неопределенностях известны лишь границы изменений, а какие-либо статистические характеристики просто отсутствуют (именно такие неопределенности и рассматриваются в настоящей работе). Отметим, что учет неопределенностей является новым направлением в игровых задачах. Это направление — теория игр при неопределенности —

появилось буквально в последние годы и активно развивается в настоящее время. Практическая ценность таких исследований в том, что необходимость принимать решения, для которых не удается полностью учесть предопределяющие их условия, а также последующее их влияние, встречается в подавляющем большинстве областей экономики, экологии и социальных наук. При этом отказаться от принятия решения большей частью просто невозможно. В рассматриваемом примере «роль» неопределенности выполняет количество y импортируемого товара, неожиданно появившегося на рынке сбыта, а множество таких неопределенностей Y есть отрезок $[\alpha, \beta]$, где постоянные $\beta > \alpha \geqslant 0$.

Функции выигрыша. Игра происходит следующим образом. Каждый игрок, исходя из своих соображений (далее об этом подробно), выбирает и применяет свою стратегию $x_i \in X_i$, $i \in \mathbb{N}$. В результате образуется набор стратегий игроков $x = (x_1, ..., x_N) \in X = X_1 \times ... \times X_N$, называемый в теории игр *ситуацией.* Независимо от действий игроков реализуется неопределенность $y \in Y$ (может реализоваться любая неопределенность $y \in Y$). На полученных таким образом парах $(x, y) \in X \times Y$ определена функция выигрыша i-го игрока $f_i(x, y)$, значение которой в каждой конкретной паре (x, y) называется *выигрышем* или *исходом* для i-го игрока, $i \in \mathbb{N}$. Набор таких выигрышей – вектор-столбец $f(x, y) = (f_1(x, y), ..., f_N(x, y))$ — определяет исход всей игры. Не ограничивая общности, будем считать, что цель каждого игрока — добиться возможно большего своего выигрыша.

В рассматриваемой игре ситуацией будет набор количеств произведенных всеми предприятиями товаров — вектор $x = (x_1, ..., x_N) \in X \subset \mathbb{R}^N$ ($\mathbb{R}^N$ — евклидово N-мерное пространство), множество ситуаций $X = [0, +\infty)^N$ — декартово (прямое) произведение N положительных полуосей $\mathbb{R}_+ = [0, +\infty)$. Функция выигрыша i-го игрока

$$f_i(x, y) = \left[a - b\left(\sum_{i=1}^{N} x_i + y\right) - c_i\right]x_i, \quad i \in \mathbb{N}, \qquad (2)$$

есть его прибыль при производстве i-м предприятием товара в количестве x_i, $i \in \mathbb{N}$, и появлении на рынке импорта в размере y. В самом деле, в функции (2) учтено, что всего в ситуации $x = (x_1, ..., x_N)$ будет поставлено на рынок количество $K = \sum_{i \in N} x_i + y$ товара, который будет продаваться по зависящей от предложения K цене $p = \max\{a - bK, 0\}$; при $a - b\left(\sum_{i \in N} x_i + y\right) > 0$ доход i-го производителя от продажи будет $x_i\left[a - b\left(\sum_{i \in N} x_i + y\right)\right]$, а его издержки $x_i c_i$, и поэтому прибыль i-го предприятия может быть представлена в виде (2).

Варианты игры. В зависимости от возможности или невозможности совместных действий между фирмами-игроками (заключающихся в согласованном или несогласованном выборе ситуации) игры подразделяются на бескоалиционные, кооперативные и коалиционные.

В *бескоалиционном варианте* (в теории игр называемом просто «бескоалиционной игрой») каждая фирма выбирает свою стратегию самостоятельно с целью достичь, возможно, большего своего выигрыша (максимизация прибыли); т. е. рассматривается рыночная структура, которая характеризуется ценовой конкуренцией или конкуренцией за объем продаж между неспособными оказывать какое-либо влияние на рыночную цену фирмами. Таким образом, речь идет о конкурентных рынках.

Кооперативный вариант («кооперативная игра») — противоположный бескоалиционному. Здесь все игроки выбирают свои стратегии совместно и согласованно. Такой тип рыночной структуры характеризуется высокой степенью рыночной власти (способностью воздействовать на рыночную цену посредством регулирования объема предложения) и отсутствием конкуренции. Это — монопольная структура.

Коалиционный вариант («коалиционная игра») — все фирмы распределены на попарно непересекающиеся группы (коалиции), внутри которых кооперативный вариант, а между коалициями — бескоалиционный. С позиций экономической теории такой тип рыночной структуры характеризуется стратегическим взаимодействием немногих, обладающих рыночной властью и конкурирующих за объем продаж групп фирм, т. е. — олигополия.

Так как работа посвящена лишь бескоалиционному варианту игры, остановимся на нем более подробно. Впервые понятие равновесного решения бескоалиционной игры (без учета неопределенностей) предложил Дж. Нэш в 1951 г. [3]. Содержательный смысл равновесного по Нэшу решения (ситуации) состоит в том, что в отклонении от этой ситуации отдельный игрок не заинтересован, так как при таком отклонении его выигрыш увеличиться не может, но может и уменьшиться. Однако, наряду с достоинствами, равновесное по Нэшу решение обладает рядом негативных свойств. Поэтому в дальнейшем были введены и другие решения бескоалиционной игры (равновесие по Бержу, равновесие возражений и контрвозражений, активное равновесие), частично «снимающие» недостатки равновесия по Нэшу, но не учитывающие влияние неопределенностей. Впервые такой учет был произведен В. И. Жуковским и А. А. Чикрием [4] и развит на случай динамических игр в монографии В. И. Жуковского [5]. Предложенная там формализация гарантированного равновесия бескоалиционной игры при неопределенности стоит на стыке понятий равновесных решений (из теории бескоалиционных игр) и векторных минимумов (из теории многокритериальных задач), объединение которых основывается на принципе векторного максимина. Однако такой подход рассчитан на «катастрофу» (на реализацию «самой плохой» для всех игроков неопределенности), появление которой, как правило, маловероятно; и приводит к внутренней неустойчивости множества гарантированных равновесий. А с позиций синергетического подхода — к усилению флуктуаций и возможности выхода системы за пределы «коридора развития».

Избежать этих двух недостатков позволяет привлечение вместо принципа максимина соответствующим образом модифицированного принципа минимаксного сожаления. Этому вопросу и посвящена настоящая работа. В ней фактически применен принцип минимаксного сожаления (принцип Сэвиджа) к формализации гарантированного равновесия бескоалиционной игры при неопределенности.

Формализация бескоалиционной игры при неопределенности. Перейдем к описанию объекта исследования — общей математической модели конфликта, как правило, возникающего в экономике (например, при совершенной конкуренции производителей, продавцов, покупателей — см. приведенный пример). Итак, рассматриваем бескоалиционную игру N лиц при неопределенности

$$\Gamma = \langle \mathbb{N}, \{X_i\}_{i \in \mathbb{N}}, Y, \{f_i(x, y)\}_{i \in \mathbb{N}} \rangle, \tag{3}$$

в которой элементы множества $\mathbb{N} = \{1, 2, ..., N\}$ называются игроками (целое число $N \geqslant 2$); элементы x_i заданного множества $X_i \subseteq \mathbb{R}^{n_i}$ — стратегиями игрока i, $i \in \mathbb{N}$; упорядоченные наборы $x = (x_1, x_2, ..., x_N) \in X = X_1 \times X_2 \times ... \times X_N \subseteq \mathbb{R}^n$ — ситуациями ($n = n_1 + n_2 + ... + n_N$); неопределенности y являются элементами заданного множества $Y \subseteq \mathbb{R}^m$; каждая из функций $f_i(x, y) : X \times Y \to \mathbb{R}$ называется функцией выигрыша i-го игрока, а ее значение на конкретной паре $(x, y) \in X \times Y$ — выигрышем i-го игрока, $i \in \mathbb{N}$.

«Содержательно» процесс игры Γ состоит в одновременном и самостоятельном (независимом) выборе каждым игроком одной из своих стратегий $x_i \in X_i$, $i \in \mathbb{N}$, и, таким образом, складывается набор стратегий $(x_1, x_2, ..., x_N) = x$ — конкретная ситуация игры Γ. Независимо от действий игроков в игре реализуется некоторая (опять-таки конкретная) неопределенность $y \in Y$. Значение функции выигрыша $f_i(x, y)$ на сложившейся в результате паре (x, y) определяет исход игры для i-го игрока — его выигрыш. Как и ранее, будем считать, что цель каждого участника конфликта заключается в получении по возможности большего индивидуального выигрыша. При выборе своих стратегий игроки должны учитывать возможность появления в игре любой, заранее непредсказуемой неопределенности $y \in Y$.

После формализации математической модели конфликта основные задачи заключаются в выработке принципов оптимальности, установлении их реализуемости (т. е. установлении существования оптимальных в этом смысле решений) и реализации решений.

Под *решением игры при неопределенности* естественно понимать пару — ситуацию и векторную гарантию, которую игроки себе «обеспечивают» выбором своих стратегий из этой ситуации.

Под векторной гарантией игры (3) при ситуации $x^* \in X$ понимается [6, с. 41] вектор $f^* = (f_1^*, ..., f_N^*)$ такой, что существует неопределенность $y^* \in Y$, при которой $f_i(x^*, y^*) = f_i^*$, и для каждой неопределенности $y \in Y$ не могут одновременно выполняться строгие неравенства $f_i^* > f_i(x^*, y)$, $i \in \mathbb{N}$.

Таким образом, *векторная гарантия* — это такой набор выигрышей игроков, «порожденный» ситуацией x^*, что при реализации любой неопределенности $y \in Y$ появляющийся другой возможный в игре (3) набор выигрышей $(f_1(x^*, y), ..., f_N(x^*, y))$ не может быть меньше векторной гарантии одновременно по всем компонентам. Виды векторных гарантий для игры (3), их свойства и способы построения изложены в работе [6, с. 42—71].

Постановка задачи. Рассмотрим бескоалиционную игру (3) N лиц при неопределенности (в нормальной форме), где $\mathbb{N} = \{1, 2, ..., N\}$; стратегии i-го игрока $x_i \in X \subseteq \mathbb{R}^{n_i}$; ситуации

$$x = (x_1, ..., x_N) \in X = \prod_{i \in N} X_i \subseteq \mathbb{R}^n, \quad n = \sum_{i \in N} n_i;$$

неопределенности $y \in Y \subseteq \mathbb{R}^m$; на парах $(x, y) \in X \times Y$ определена функция выигрыша i-го игрока $f_i(x,y)$, $i \in \mathbb{N}$. В игре (3) игрок i стремится выбрать такую свою стратегию $x_i \in X_i$, чтобы достичь возможно большего своего выигрыша (значения $f_i(x,y)$) при условии, что все игроки одновременно выбирают свои стратегии и одновременно реализуется неопределенность $y \in Y$. Подчеркнем еще раз, что:

- выбор своих стратегий всеми игроками производится одновременно, независимо и самостоятельно (в этом состоит «бескоалиционный характер» игры (3)), в результате складывается ситуация $x = (x_1, x_2, ..., x_N) \in X$;
- при выборе своих стратегий игрок вынужден учитывать возможность реализации любой неопределенности $y \in Y$.

С целью формализации гарантированного решения игры (3) (с учетом подхода принципа минимаксного сожаления) каждой неопределенности $y \in Y$ и каждой ситуации $x = (x_1, x_2, ..., x_N) \in X$ поставим в соответствие *функцию риска i-го игрока*

$$\Phi_i(x, y) = \max_{z_i \in X} f_i(x \| z_i, y) - f_i(x, y), \quad i \in \mathbb{N}.$$

Функция $\Phi_i(x, y)$ численно оценивает «сожаление» (риск) i-го игрока о том, что при сложившихся неопределенности $y \in Y$ и стратегиях остальных игроков

$$x_{\mathbb{N}\setminus\{i\}} = (x_1, ..., x_{i-1}, x_{i+1}, ..., x_N) \in X_{\mathbb{N}\setminus\{i\}} = \prod_{j \in \mathbb{N}\setminus\{i\}} X_j$$

этот игрок применяет стратегию x_i, а не $x_i^{(i)} = \arg\max_{z_i \in X_i} f_i(x \| x_i, y)$, где $(x \| z_i) = (x_1, ..., x_{i-1}, z_i, x_{i+1}, ..., x_N)$.

Набор $(\Phi_1(x, y), ..., \Phi_N(x, y)) = \Phi(x, y)$ назовем *векторной функцией* риска в игре (3).

Это означает, что в реальной экономической ситуации риск характеризует отклонение полученного (реального) результата деятельности фирмы от ожидаемого (наилучшего). Например, риск недополучения прибыли характеризуется разницей между реально полученной и ожидаемой прибылью.

Естественно предполагать, что каждый игрок по возможности стремится уменьшить свой риск (значение своей функции риска), учитывая наибольшее «противодействие» неопределенности $y \in Y$. Так риск получает как количественную, так и качественную характеристики.

Далее, аналогично игре (3), рассматривается бескоалиционная игра N лиц при неопределенности

$$\langle \mathbb{N}, \{X_i\}_{i \in \mathbb{N}}, Y, \{F_i(x, y\} \rangle. \tag{4}$$

Водится понятие *гарантированного решения*, которое на содержательном уровне понимается как пара: ситуация (она выбирается игроками) и гарантированный векторный риск, который «обеспечат себе» игроки, следуя своим стратегиям в данной ситуации. В свою очередь, гарантированный векторный риск — это такой набор значений функций риска всех игроков («порожденных» данной ситуацией и специальной неопределенностью), что при реализации любой неопределенности, соответствующий ей векторный риск не может стать больше гарантированного одновременно по всем компонентам. Игроки путем выбора своих стратегий стремятся по возможности этот риск уменьшить (в векторном смысле). При построении пары (ситуация, неопределенность, реализующая гарантированный векторный риск) будем исходить из интуитивных представлений о выгодности, устойчивости и справедливости (как это диктуется общими положениями теории игр [7]).

Таким образом, содержательный смысл предлагаемого подхода состоит в следующем: при имеющейся информации о рынках каждый игрок, оценивая в игре качество своего функционирования, ориентируется не на (как принято) возможно большее значение функции выигрыша (прибыли), а стремится уменьшить разность между наибольшим и получившимся значениями функции выигрыша, т. е. уменьшить свой риск (сожаление), плату за отсутствие необходимой информации. Как раз отсутствие информации и проявилось в том, что игрок применил стратегию, которая не максимизирует его функцию выигрыша (прибыль) при реализовавшейся неопределенности, а минимизирует его функцию риска. В связи с этим исходная игра (3) заменяется на вспомогательную игру (4), отличающуюся тем, что функции выигрыша (прибыли) игроков заменяются на их «сожаления», т. е. на их функции риска. Игроки в такой вспомогательной игре формируют стратегии с целью возможно уменьшить значения своих функций риска.

В экономической теории *микроэкономические риски* можно определить как риски, появляющиеся в процессе потребительского выбора и риски конкретных фирм; т. е. риски при принятии экономических решений на уровне домашних хозяйств и фирм в заданных условиях, а также риски, возникающие при формировании этих условий в результате совместных действий игроков.

В рамках предлагаемого подхода микроэкономические риски определяются как разница между предполагаемым и реализовавшимся результатом деятельности (с учетом неопределенности); т. е. размер такого отклонения результата и есть мера микроэкономического риска.

В общем случае, с позиций синергетики, микрорешения отдельных экономических агентов (фирм), принимаемые на основе предлагаемого подхода, обеспечивают устойчивость (равновесность) системы на микроуровне, т. е. при наличии флуктуаций (отклонений) в системе не происходит усиления флуктуаций вследствие того, что нет ориентации на «катастрофу», а есть стремление достичь «самого хорошего, что может случиться».

Такая формализация гарантированного по риску решения оказалась, на наш взгляд, довольно перспективной. В частности, установлено [8], что если такое решение существует, то игроки, придерживаясь своих стратегий из ситуаций (входящих в решение), «обеспечат себе» нулевой вектор риска, какая бы неопределенность не реализовалась в данной системе. Так как по определению функция риска каждого игрока неотрицательна, но нулевой риск наиболее желателен для каждого игрока.

ЗАКЛЮЧЕНИЕ

Отметим, что в данной работе мы лишь коснулись совершенно нового подхода к принятию решений в бескоалиционных системах при неопределенности. Причем о неопределенностях известна лишь область возможных значений. Но уже полученные в работе [8] результаты (особенно касающиеся нулевого риска) представляются перспективными. В стороне остались кооперативный, коалиционный и иерархический варианты игры. Возможны и другие подходы к формализации рисков.

ЛИТЕРАТУРА

1. *Серегина С. Ф.* Роль государства в экономике. Синергетический подход. − М.: Дело и сервис, 2002.
2. *Воробьев Н. Н.* Основы теории игр. Бескоалиционные игры. — М.: Наука, 1984.
3. *Nash J. F.* Non-cooperative games // Ann. Math. — 1951. — Vol. 54. — P. 286—295.
4. *Жуковский В. И., Чикрий А. А.* Линейно квадратичные дифференциальные игры. — Киев: Наукова думка, 1994.
5. *Жуковский В. И.* Введение в дифференциальные игры при неопределенности. — М.: Междунар. НИИ проблем управления, 1997.
6. *Жуковский В. И.* Кооперативные игры при неопределенности и их приложения. — М.: УРСС, 1999.
7. *Воробьв Н. Н.* Теория игр для экономистов-кибернетиков. — М.: Наука, 1985. — С. 13.
8. *Жуковский В. И., Жуковская Л. В.* Риск в многокритериальных и конфликтных системах при неопределенности. — М.: Едиториал УРСС, 2004.

☎ *(095) 334-90-09*

❑

УДК 336.767

К ВЫБОРУ МЕТОДА ОЦЕНКИ ЭФФЕКТИВНОСТИ УПРАВЛЕНИЯ ИНВЕСТИЦИОННЫМ ПОРТФЕЛЕМ

Д. Ю. Голембиовский

Professional Risk Management International Association[1], *г. Москва*

Дан анализ методов оценки эффективности управления портфелем. Разработана система аксиом, предоставляющая возможность математического анализа методов оценки эффективности управления портфелями. Предложен метод оценки эффективности управления, основанный на расчете для каждого портфеля собственного эталонного портфеля, и разработана соответствующая программа.

ВВЕДЕНИЕ

Одной из актуальных задач управления в социально-экономических системах является финансовое управление и, в частности, управление инвестиционным портфелем. Инвестору денежных средств необходима оценка результатов труда менеджера, осуществляющего управление портфелем. В свою очередь, менеджеры портфелей заинтересованы в анализе эффективности различных стратегий управления. Помимо действий менеджера на результат управления оказывают влияние характер трансакций, осуществленных инвестором, и динамика рынка в период управления. В этих условиях оценка действий менеджера портфеля и выбранной стратегии управления представляет собой сложную проблему.

Ретроспективный анализ эффективности работы менеджеров или эффективности различных стратегий управления может быть осуществлен путем сравнения доходности соответствующих портфелей за один и тот же период времени. Оценка доходности портфеля не составляет труда, если на интервале оценивания не производились довложения или изъятия из портфеля активов или денежных средств. В этом случае доходность измеряется процентной ставкой за период:

$$r_p = \frac{A - A_0}{A_0} = \frac{A}{A_0} - 1 \qquad (1)$$

или годовой процентной ставкой:

$$r = (1 + r_p)^{365/T} - 1. \qquad (2)$$

Здесь A_0 — рыночная стоимость активов портфеля вместе с денежным остатком в начале рассматриваемого периода времени длительностью T, A — аналогичная характеристика портфеля в конце этого периода. Например, для портфеля облигаций величины A_0 и A включают в себя рыночную стоимость облигаций (по средневзвешенным ценам или ценам закрытия), сумму накопленного купонного дохода по всем купонным облигациям и денежный остаток портфеля в соответствующие моменты времени. В случае портфеля акций величина A учитывает также дивидендные платежи, поступившие в течение периода оценки. При этом дивиденд может остаться частью денежных средств портфеля или воплотиться в дополнительно приобретенные акции.

В том случае, когда в течение рассматриваемого периода производились довложения и (или) изъятия средств из портфеля, задача оценки доходности становится более сложной. В литературе (см., например, работы [1, 2]) рассматриваются следующие два основных метода ее решения.

1. МЕТОД ВНУТРЕННЕЙ ДОХОДНОСТИ И TIME-WEIGHTED МЕТОД

В общем случае балансовое уравнение финансовой операции имеет вид

$$\sum_{i=0}^{n} \frac{X_i}{(1 + r)^{(t_i - t_0)/365}} = 0, \qquad (3)$$

где $i = 0, ..., n$ — номера дней, когда осуществлялось проведение расходных или доходных платежей, $t_0, t_1, ..., t_n$ — соответствующие даты (заметим, что $t_n - t_0 = T$); $X_i \neq 0$ — размеры платежей. Значение X_i отрицательно в случае расходного платежа и положительно в случае доходного платежа. Решение уравнения (3) относительно ставки дисконтирования r называется *внутренней нормой доходности* или *эффективной ставкой* финансовой операции. Из показателя степени, в которую возводятся знаменатели слагаемых в уравнении (3), видно, что значение r характеризует годовые результаты финансовой операции. Если значение r умножить на 100, то полученная

[1] www.prmia.org

величина будет выражать доходность операции в годовых процентах.

При оценке доходности портфеля коэффициенты балансового уравнения определяются следующим образом:

$$X_0 = -A_0, \quad (4)$$

$$X_i = \Delta_i, \quad i = 1, ..., n - 1, \quad (5)$$

$$X_n = A_n + \Delta_n, \quad (6)$$

где A_0 — рыночная стоимость активов портфеля вместе с денежным остатком в день t_0, Δ_i — довложения (изъятия) средств, которые были осуществлены в день t_i (величина Δ_i отрицательна в случае довложения средств и положительна при их изъятии), A_n — сумма рыночной стоимости и денежного остатка портфеля на конец дня t_n.

Другой известный метод состоит в перемножении коэффициентов подорожания портфеля за те периоды времени, когда довложения или изъятия средств не производились. Пусть, например, A_{i-1} — стоимость портфеля в начале i-го такого интервала, A_i — стоимость портфеля в конце этого интервала, а Δ_i — сумма изъятых из портфеля денежных средств в конце этого интервала. (Если деньги добавлялись в портфель, то величина Δ_i будет отрицательной). Коэффициент подорожания портфеля за i-й период

$$k_i = (A_i + \Delta_i)/A_{i-1}. \quad (7)$$

Произведение всех коэффициентов (7) за рассматриваемый период, уменьшенное на единицу, дает оценку доходности портфеля:

$$r = \prod_{i=1}^{n} k_i - 1. \quad (8)$$

При пересчете на год получается соотношение

$$r = \left(\prod_{i=1}^{n} k_i\right)^{365/T} - 1. \quad (9)$$

В зарубежной литературе данную оценку принято называть доходностью, взвешенной по времени (time-weighted rate).

Нетрудно убедиться в том, что при отсутствии дополнительных вложений и изъятий средств из портфеля как расчет по формуле (5), так и решение уравнения (3) приводят к формуле (2). Оба описанных метода применимы для оценки доходности любых портфелей, активы которых обладают рыночной стоимостью. Это могут быть, например, уже упомянутые акции и облигации, а также паи взаимных фондов, валюты, драгоценные металлы, производные финансовые инструменты — фьючерсы, форварды и опционы.

Предположим, что нам требуется сравнить результаты работы двух менеджеров портфелей за некоторый период времени. Оценки доходности их портфелей в общем случае могут быть получены на основе уравнения (3) или соотношения (9). Лучшими признаются результаты того из менеджеров, оценка которого выше.

Каждому из данных методов присущи, однако, определенные недостатки. Рассмотрим следующий пример управления портфелем [1]. В начале квартала стоимость портфеля составляет \$50. В середине квартала рыночная стоимость его падает до \$25, после чего клиент дополнительно вкладывает в портфель \$25. В конце квартала рыночная стоимость портфеля становится равной \$100. Коэффициенты соответствующего балансового уравнения определяются соотношениями $X_0 = -50$, $X_1 = -25$ и $X_2 = 100$. Решение данного уравнения приводит к ставке внутренней доходности 291,46 % годовых.

Использование соотношения (9) дает значение годовой доходности 0 %. Действительно, $k_1 = (50 - 25)/50 = 0{,}5$; $k_2 = 100/50 = 2$, откуда $r = (0{,}5 \times 2)^{365/91} - 1 = 0$.

Последний результат часто считается более соответствующим цели — оценке эффективности управления портфелем, так как каждый доллар портфеля в первой половине квартала подешевел вдвое, а во второй половине — увеличил свою стоимость в два раза. Большое значение внутренней доходности (291,46 %) может быть обусловлено тем, что инвестор сделал дополнительное крупное вложение перед благоприятным изменением рыночной конъюнктуры; т. е. доходность в 291,46 % объясняется действиями клиента, а не менеджера.

Однако двукратный рост портфеля во второй половине квартала в равной мере может быть вызван не благоприятной рыночной конъюнктурой, а искусными действиями трейдера. Тогда значение доходности 291,46 %, по-видимому, более адекватно ситуации, чем значение 0. Основное качественное отличие между методами, использующими соответственно уравнение (3) и соотношения (8) и (9), состоит в том, что второй из них не учитывает динамику объема портфеля. Перемножая коэффициенты подорожания портфеля за периоды, когда внешних трансакций не производилось, рассчитывают получить «чистую» оценку мастерства менеджера. Вряд ли данная попытка абстрагироваться от внешних обстоятельств может быть признана удачной. Например, менеджер может превосходно управляться с небольшим портфелем и терпеть фиаско, когда объем его становится достаточным, чтобы оказать влияние на конъюнктуру рынка. В таблице представлен пример подобной ситуации. Оценка доходности портфеля, рассчитанная по формуле (8), составляет $(1{,}5 \times 1 \times 0{,}94 - 1) \times 100 = 40{,}87$ % за рассматриваемый период управления. При этом инвестор терпит убыток $108 - 100 - 10 = -2$ ед. Решение уравнения (3) дает оценку доходности $-5{,}57$ % годовых или $-0{,}89$ % за рассматриваемый период управления.

Часто для оценки эффективности управления сравнивают доходность портфеля с доходностью некоторого эталонного портфеля, характеризующего динамику соответствующего рынка в целом и адекватного целям инвестора денежных средств [3—12]. В частности, в качестве эталонов могут быть взяты портфели известных рыночных индексов — например, Dow Jones и индекс РТС для оценивания портфелей, соответственно, американ-

Пример управления несамофинансируемым портфелем

Дата	Стоимость портфеля вместе с денежным остатком, руб.	Дополнительные вложения, руб.	k_i
01.04.97	10	0	—
08.04.97	15	0	1,5
15.04.97	115	100	1
22.04.97	108	0	0,94

ских и российских акций; для оценки эффективности вложений в российские корпоративные облигации можно воспользоваться индексом RCBI (Russian Corporate Bond Index)[2]. Доходность эталонного портфеля может быть рассчитана на основе соотношений (1) и (2), где в качестве величин A_0 и A взяты соответствующие значения его стоимости. Разумный эталон может быть сконструирован и для других типов рынков. Например, для рынка валютообменных операций (FOREX) эталон формируется в виде некоторой корзины валют.

Работа менеджера оценивается положительно, если доходность его портфеля превышает доходность эталонного портфеля за соответствующий период времени. Результаты нескольких менеджеров могут быть ранжированы в соответствии с разностью оценок доходности их портфелей и соответствующей доходности эталонов.

Данному подходу, однако, также свойственны логические несоответствия. Предположим, что в рассмотренном примере из работы [1] доходность рыночного портфеля за квартал составила 0. Этой информации недостаточно, чтобы, например, оценить доходность портфеля в 291,46 %, рассчитанную по уравнению (3), как низкую или высокую. Все зависит от динамики рынка в течение квартала. Если в первой половине квартала стоимость рыночного портфеля сократилась вдвое, а во второй половине цены двукратно возросли, то данный результат соответствует работе «по рынку». В том случае, когда стоимость рыночного портфеля все время была постоянной, результат 291,46 % свидетельствует о более высоком мастерстве менеджера. Аналогичная проблема возникает и при оценке портфеля менеджера методом time-weighted. Доходность его (0) соответствует работе «по рынку» только в том случае, если рынок был «горизонтален» в течение всего квартала. В альтернативном же случае при данном графике трансакций менеджер, все время действуя «по рынку», должен был заработать больше! Таким образом, удовлетворительный метод оценки эффективности управления портфелем обязан учитывать динамику рынка, а не только значение рыночной эффективности за весь рассматриваемый период времени.

Если не удается построить эталонный портфель, адекватный инвестиционному стилю менеджеров, то при сравнении с эталоном учитывается различие рисков инвестиционного и эталонного портфелей [12—19]. Для этого применяются различные методы, основанные на модели CAPM (Capital Asset Price Model). Например, может вычисляться коэффициент «доходность—изменчивость» или коэффициент «доходность—разброс». Наше рассмотрение соответствует ситуации, когда адекватный эталонный портфель известен и нет необходимости в учете различия рисков.

Значительным недостатком балансовых уравнений считается возможность отсутствия решения или существования нескольких решений. В работе [20] утверждается, например, что применение уравнения (3) допустимо только в том случае, если все доходные платежи осуществляются после всех расходных платежей (тогда существование и единственность решения гарантированы). Более общее достаточное условие единственности решения уравнения (3) приводится в работе [2]. На практике, однако, ситуации, когда уравнение (3) имеет несколько решений, встречаются нечасто.

Заметим, что рассмотренные недостатки time-weighted метода несущественны при оценке доходности рынка. Действительно, относительно небольшие колебания объема рыночного портфеля, вызванные дополнительной эмиссией или погашением отдельных выпусков ценных бумаг, не изменяют инвестиционных возможностей участников. В связи с этим time-weighted метод целесообразно применять для оценки доходности эталонного рыночного портфеля, состав которого может динамически изменяться.

Что касается оценки эффективности управления портфелем инвестора, то здесь учет динамики объема портфеля является объективной необходимостью. Предлагаемый в работе метод оценивания основан на применении аппарата балансовых уравнений.

2. ТРЕБОВАНИЯ К МЕТОДУ ОЦЕНКИ ЭФФЕКТИВНОСТИ УПРАВЛЕНИЯ

Анализ изложенного выше позволяет формально определить аксиомы, которым должен удовлетворять метод оценки эффективности управления портфелем. Обозначим через $[t_0, t_n]$ период времени, для которого требуется получить оценку эффективности управления. Трансакции по формированию инвестиционного портфеля определяются соотношениями (4)—(6). Предположим, что доходность рыночного портфеля известна для любого интервала времени $[t_k, t_l] \subset [t_0, t_n]$, где t_k и t_l — моменты времени осуществления трансакций, $t_k < t_l$. Обозначим рыночную доходность за данный интервал $R(k, l)$. В соответствии с изложенным в § 1, доходность $R(k, l)$ вычисляется на основе time-weighted метода, поэтому

$$1 + R(k, l) = (1 + R(k, i))(1 + R(i, l)), \quad \forall k < i < l. \qquad (10)$$

При отсутствии доходных и расходных платежей внутри периода $[t_0, t_n]$ и нулевой доходности рынка оценка эффективности управления должна совпадать с соответствующим значением процентной ставки (2). Согласно введенной системе обозначений в этом случае $n = 1$, и интервал оценивания должен быть обозначен как $[t_0, t_1]$, а соответствующие трансакции X_0 и X_1. Таким образом, метод оценки эффективности управления портфелем должен отвечать требованию, которое выражает

Аксиома 1. *Если* $R(0, 1) = 0$, *то для интервала* $[t_0, t_1]$ *эффективность управления портфелем* $\rho = (-X_1/X_0)^{365/T} - 1$.

Здесь, как и ранее $T = t_1 - t_0$; минус в скобке обусловлен отрицательностью первоначальной трансакции X_0.

Как отмечалось в § 1, эффективность управления невозможно оценить без учета соответствующей динамики рынка. Следовательно, мера эффективности управления портфелем должна быть функцией трансакций $X_0, ..., X_n$ и введенной доходности рынка:

$$\rho = \Omega(X_0, ..., X_n, R(0, 1), R(0, 2), ..., R(n - 1, n)). \qquad (11)$$

«Провал» time-weighted метода для примера, представленного в таблице, связан с игнорированием реального размера полученной прибыли. Понятно, что наряду с размером прибыли удовлетворительная мера эф-

[2] http://www.micex.ru/stock/issue_4533.html

фективности управления портфелем должна учитывать и соответствующую конъюнктуру рынка. Однако в том случае, когда рыночная доходность была нулевой в течение всего рассматриваемого периода, естественно ожидать, что мера эффективности управления будет принимать отрицательное, нулевое или положительное значение в том случае, когда прибыль портфеля соответственно меньше нуля, равняется нулю или положительна. Заметим, что прибыль портфеля есть сумма введенных выше трансакций портфеля: $S = \sum_{i=0}^{n} X_i$.

Например, для портфеля (см. таблицу) $X_0 = -10$, $X_1 = -100$, $X_2 = 108$. Имеет место убыток $X_0 + X_1 + X_2 = -2$ ед.

Определим указанное соответствие формально.

Аксиома 2. *Если* $R(k, k+1) = 0$, $k = 0, ..., n-1$, *то*

$$S < 0 \Rightarrow \rho < 0,\ S = 0 \Rightarrow \rho = 0,\ S > 0 \Rightarrow \rho > 0.$$

Очевидно, что рассмотренный time-weighted метод не удовлетворяет аксиоме 2.

Для правильной оценки результатов работы менеджеров портфелей необходимо, чтобы в идентичных условиях большему значению прибыли отвечало большее значение меры эффективности управления. «Условия» работы менеджера в рамках рассматриваемой модели определяются размером начальной трансакции $-A_0$, последовательностью довложений — изъятий средств Δ_i, $i = 1, ..., n$ и функцией доходности рынка $R(k, l)$. Соответствующее требование к методу оценки эффективности можно формально записать следующим образом. Пусть заданы две ситуации управления портфелями на интервале времени $[t_0, t_n]$: $X_0^1 = -A_0^1$; $X_i^1 = \Delta_i^1$, $i = 1, ..., n-1$; $X_n^1 = A_n^1 + \Delta_n^1$ при доходности рынка $R^1(k, k+1)$, $k = 0, ..., n-1$, и $X_0^2 = -A_0^2$; $X_i^2 = \Delta_i^2$, $i = 1, ..., n-1$; $X_n^2 = A_n^2 + \Delta_n^2$ при доходности рынка $R^2(k, k+1)$, $k = 0, ..., n-1$.

Соответствующие прибыли портфелей $S^1 = \sum_{i=0}^{n} X_i^1$ и $S^2 = \sum_{i=0}^{n} X_i^2$.

Аксиома 3. *Если* $A_0^1 = A_0^2$, $\Delta_i^1 = \Delta_i^2$, $i = 1, ..., n$, *и* $R^1(k, k+1) = R^2(k, k+1)$, $k = 0, ..., n-1$, *то* $S^1 > S^2 \Rightarrow \rho^1 > \rho^2$.

Как видно, в условиях аксиомы 3 рассмотренные ситуации управления портфелями могут различаться только по параметрам A_n^1 и A_n^2. Заметим, что для рассмотренных в § 1 time-weighted метода и метода балансового уравнения сформулированная аксиома выполняется. Действительно, из $S^1 > S^2$ следует, что $A_n^1 > A_n^2$. В этих условиях оба метода обеспечат более высокую оценку для первого из портфелей.

Согласно аксиоме 2, при «горизонтальном» рынке нулевой прибыли портфеля должно соответствовать нулевое значение меры. Разумно в таком случае потребовать, чтобы менеджер, доходность портфеля которого все время соответствовала рыночной доходности, также получил бы в итоге нулевую оценку. Следовательно, если значение меры эффективности управления равно нулю на всех интервалах времени, где не производились изъятия или довложения средств, то за весь период управления также должна получаться нулевая оценка. Пусть $0 \leqslant i \leqslant n-1$ и $Y_i = -A_i$, $Y_{i+1} = A_{i+1} + \Delta_{i+1}$, где A_i и A_{i+1} — суммы денежного остатка и рыночной стоимости активов портфеля в соответствующие моменты времени. Мы приходим к следующему формальному условию.

Аксиома 4. *Если* $\rho = \Omega(Y_i, Y_{i+1}, R(i, i+1)) = 0$, $\forall i$: $0 \leqslant i \leqslant n-1$, *то* $\rho = \Omega(X_0, ..., X_n, R(0, 1), R(1, 2), ..., R(n-1, n)) = 0$.

Как было показано на примере в § 1, метод балансового уравнения может приводить к неоправданно высокой оценке действий менеджера, который работает «по рынку». Сформулированная аксиома исключает подобные коллизии.

3. ОПИСАНИЕ И СВОЙСТВА МЕТОДА «СОБСТВЕННОГО ЭТАЛОНА»

Суть предлагаемого метода состоит в расчете доходности, которую имел бы рассматриваемый портфель в том случае, если бы его менеджер все время работал «по рынку». Полученная величина используется в качестве базы сравнения.

Рассчитаем стоимость анализируемого портфеля вместе с денежным остатком на конец интервала $[t_0, t_n]$ при условии, что на всех отрезках этого периода, где не производилось изъятия или довложения средств, доходность портфеля совпадала с соответствующей доходностью рынка:

$$\bar{A}_n = (...((A_0(1 + R(0, 1)) - \Delta_1)(1 + R(1, 2)) - \Delta_2)...) \times$$
$$\times (1 + R(n-1, n)) - \Delta_n = A_0 \prod_{j=0}^{n-1}(1 + R(j, j+1)) -$$
$$- \sum_{i=1}^{n-1} \Delta_i \prod_{j=i}^{n-1}(1 + R(j, j+1)) - \Delta_n. \qquad (12)$$

Воспользовавшись соотношением (10), окончательно получим:

$$\bar{A}_n = A_0(1 + R(0, n)) - \sum_{i=1}^{n-1} \Delta_i(1 + R(i, n)) - \Delta_n. \qquad (13)$$

Пусть теперь $\bar{X}_n = \bar{A}_n + \Delta_n$. Обозначим через $IRR(X_0, ..., X_{n-1}, \bar{X}_n)$ внутреннюю доходность последовательности трансакций $X_0, ..., X_{n-1}, \bar{X}_n$. Аналогично, $IRR(X_0, ..., X_{n-1}, X_n)$ — внутренняя норма доходности последовательности $X_0, ..., X_{n-1}, X_n$. В соответствии с выражением (11) оценку эффективности управления портфелем предлагается вычислять следующим образом:

$$\rho = \Omega(X_0, ..., X_n, R(0, 1), R(1, 2), ..., R(n-1, n)) =$$
$$= IRR(X_0, ..., X_{n-1}, X_n) - IRR(X_0, ..., X_{n-1}, \bar{X}_n). \qquad (14)$$

Как отмечалось, в большинстве практических ситуаций балансовые уравнения имеют единственное решение. Будем считать в дальнейшем, что данное условие выполнено в отношении функций IRR, входящих в соотношение (14). Поскольку стоимость портфеля не может быть отрицательной величиной, то $X_0 = -A_0 < 0$, а $X_n = A_n + \Delta_n \geqslant 0$. При данных условиях справедлива следующая

Теорема. *Для меры эффективности управления, определенной соотношением* (14)*, выполняются аксиомы 1—4.*

Вернемся теперь к примеру управления портфелем из работы [1]. Присвоим нулевой номер первому дню квартала, номер 1 — дню, который соответствует его середине, и номер 2 — последнему дню квартала. Рассмотрим случай, когда стоимость рыночного портфеля в первой половине квартала сокращается вдвое, а к концу периода возвращается к исходному уровню. Тогда $R(0, 1) = -0{,}5$; $R(1, 2) = 1$. Рассчитывая по формуле (13), получим $\bar{A}_3 = \$100$. Тогда $\rho = IRR(-\$50, -\$25, \$100) - IRR(-\$50, -\$25, \$100) = 0$.

Пусть $R(0, 1) = 0$; $R(1, 2) = 0$. Тогда $\bar{A}_3 = \$75$, $\rho = IRR(-\$50, -\$25, \$100) - IRR(-\$50, -\$25, \$75) = 191{,}46\ \% - 0\ \% = 191{,}46\ \%$. В обоих случаях полученная оценка соответствует интуитивному представлению о результатах работы менеджера.

Традиционный подход к оценке эффективности управления состоит в сравнении доходности инвестиционных портфелей с доходностью эталонного рыночного портфеля за соответствующий период времени. Результаты данной работы свидетельствуют о том, что для адекватной оценки эффективности управления необходимо для каждого портфеля рассчитывать собственный эталон. Конечная стоимость предложенного эталонного портфеля зависит от довложений и изъятий денежных средств, осуществленных инвестором. Результаты работы менеджеров ранжируются в зависимости от разности между доходностью инвестиционного портфеля и доходностью «собственного эталона», соответствующего этому портфелю.

Понятно, однако, что предложенный в работе метод «собственного эталона» не единственный, удовлетворяющий сформулированным аксиомам.

4. ОПЫТ ПРАКТИЧЕСКОГО ПРИМЕНЕНИЯ МЕТОДА

На основе предложенного метода в 1996 г. был разработан программный комплекс «Treasure», позволяющий сравнивать эффективности управления портфелями ГКО и ОФЗ в условиях осуществления инвесторами дополнительных вложений и изъятий денежных средств. В течение двух лет (до краха ГКО в августе 1998 г.) данный комплекс находился в эксплуатации в КБ «Гута-Банк». На основе квартальных оценок эффективности управления определялась целесообразность уплаты и размер бонуса менеджеров портфелей. Внедрение данной системы привело к заметному повышению эффективности их работы. Если в начале периода эксплуатации программы менеджеры, как правило, сильно проигрывали своим эталонным портфелям, то спустя один—два квартала они стали в основном работать на уровне своих эталонов, иногда опережая их по итогам отчетного периода.

Пример отчета программы «Treasure» представлен на рис. 1. Каждая точка графиков отражает доходность портфеля инвестора или эталонного портфеля за соответствующий период времени. Доходность портфеля *1* выше доходности портфеля *2* за весь рассматриваемый период. Однако портфель *1* уступил соответствующему эталонному портфелю. Напротив, доходность портфеля *2* оказалась выше, чем доходность портфеля «эталон *2*». Таким образом, более высокая доходность портфеля *1* объясняется более удачным с позиции динамики рынка графиком внешних трансакций, а не мастерством менеджера.

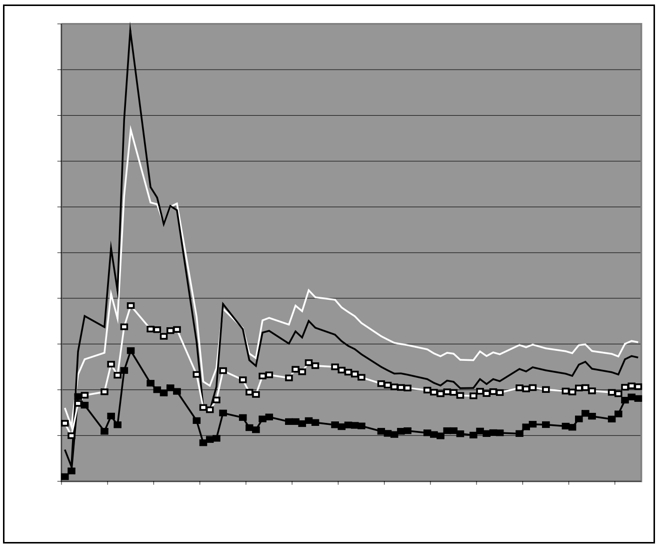


Рис. 1. Доходность за период, % годовых:
— эталон 1; — портфель 1; — эталон 2; — портфель 2

В результате внедрения программы получил количественное отражение ряд приемов управления портфелями. В частности, был подтвержден известный факт наличия в определенных пределах обратно пропорциональной зависимости между числом совершаемых операций и эффективностью управления портфелем. Менеджеры стали значительно более осмотрительны при принятии решений о перестройке портфеля.

ЗАКЛЮЧЕНИЕ

В статье рассмотрен подход к оценке эффективности управления инвестиционным портфелем без учета понесенного риска. Такой подход применяют, когда риски ограничены с помощью позиционных лимитов на отдельные ценные бумаги и, возможно, на стоимость портфеля в целом. Менеджеру позволяется занимать любые позиции в пределах установленных лимитов. Таким образом, размер риска портфеля может различаться, но не превышает определенного верхнего предела. Одна из допустимых структур портфеля соответствует эталонному рыночному портфелю. Для учета влияния произведенных в течение периода управления трансакций следует пользоваться предложенным в работе методом собственного эталона.

Существует группа методов оценки эффективности управления, учитывающих размер риска портфеля в течение рассматриваемого периода времени. При этом для каждого портфеля рассчитывается показатель эффективности, который сопоставляет полученную доходность с оценкой понесенного риска. Первые показатели доходности с учетом риска были предложены в работах [22—25]. Рассматриваемые показатели рассчитываются для анализируемого портфеля и для эталонного рыночного портфеля и сравниваются. В том случае, когда инвестор производил изъятия и дополнительные вложения средств в портфель в периоде управления, целесообразно рассмотреть возможность использования результатов настоящей работы для формирования адекватного условиям инвестирования эталонного портфеля.

Еще одно направление возможных дальнейших исследований состоит в построении и анализе класса альтернативных мер эффективности управления инвестиционным портфелем, которые удовлетворяют сформулированным аксиомам.

ПРИЛОЖЕНИЕ

Доказательство теоремы. Факт выполнения аксиомы 1 для предлагаемого метода оценки эффективности управления портфелем непосредственно следует из ее условий и соотношения (14). Рассмотрим доказательство выполнения аксиомы 2.

Сумма членов балансового уравнения носит название приведенной стоимости платежей (Net Present Value). Она является функцией последовательности трансакций и ставки дисконтирования r: $NPV(X_0, ..., X_n, r)$. Нетрудно убедиться [21], что

$$X_0 = \lim_{r\to\infty}(NPV(X_0, ..., X_n, r)), \qquad (15)$$

$$X_n = \lim_{r\to -1}((1+r)^{(t_n - t_0)/365}NPV(X_0, ..., X_n, r)), \qquad (16)$$

$$S = NPV(X_0, ..., X_n, 0). \qquad (17)$$

Из условия аксиомы 2 следует, что $R(k, l) = 0$, $\forall k, l$: $[t_k, t_l] \subset [t_0, t_n]$.

Подставляя нулевое значение доходности рынка в выражение (13), получаем

$$\bar{A}_n = A_0 - \sum_{i=1}^{n}\Delta_i \Leftrightarrow -A_0 + \sum_{i=1}^{n}\Delta_i + \bar{A}_n = 0.$$

Тогда прибыль рассматриваемой последовательности трансакций

$$S = \sum_{i=1}^{n-1}X_i + \bar{X}_n = -A_0 + \sum_{i=1}^{n-1}\Delta_i + \bar{A}_n + \Delta_n = 0.$$

Из полученного результата, равенства (17) и условия единственности решения балансового уравнения следует, что $IRR(X_0, ..., X_{n-1}, \bar{X}_n) = 0$. Таким образом, в условиях аксиомы 2 всегда $\rho = IRR(X_0, ..., X_{n-1}, X_n)$.

Аналогично, если $S = \sum_{i=1}^{n}X_i = 0$, то из равенства (17) и единственности IRR вытекает, что $IRR(X_0, ..., X_{n-1}, X_n) = 0$ и, таким образом, $\rho = 0$.

Пусть теперь $S = NPV(X_0, .., X_n, 0) > 0$. Так как $X_0 = -A_0 < 0$, то из равенства (15) вытекает, что $NPV(X_0, ..., X_n, r) < 0$ при достаточно больших r. Тогда факт $\rho > 0$ является следствием непрерывности функции $NPV(X_0, ..., X_n, r)$ и единственности решения балансового уравнения. Данная ситуация иллюстрируется рис. 2.

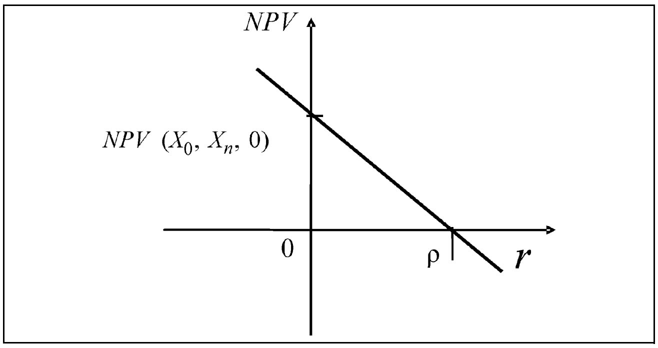


Рис. 2. Случай положительной прибыли

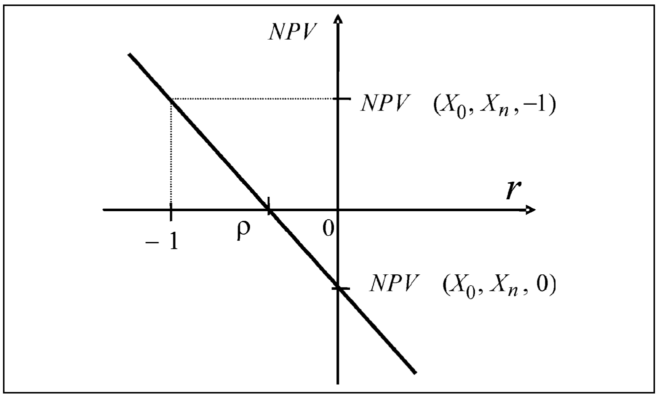


Рис. 3. Отрицательная прибыль

Аналогично рассматривается случай $S < 0$. Так как $X_n \geqslant 0$, то из равенства (16) следует, что при r, достаточно близких к -1, выполняется неравенство $NPV(X_0, ..., X_n, r) \geqslant 0$. Из единственности IRR и непрерывности функции NPV вытекает, что $\rho < 0$. Соответствующая иллюстрация дана на рис. 3.

Перейдем к аксиоме 3. Поскольку $X_i^1 = X_i^2$, $i = 0, ..., n-1$, и $S^1 > S^2$, то $X_n^1 > X_n^2$. Пусть $r^1 = IRR(X_0^1, ..., X_n^1)$ и $r^2 = IRR(X_0^2, ..., X_n^2)$. Из неравенств $X_n^1, X_n^2 \geqslant 0$ следует, что $NPV = (X_0^1, ..., X_n^1, r^2) > NPV(X_0^2, ..., X_n^2, r^2) = 0$. Так как $NPV = (X_0^1, ..., X_n^1, r) < 0$ при достаточно больших r (см. соотношение (15)), то из непрерывности функции NPV следует $r^1 > r^2$ (рис. 4).

Выражения для меры эффективности управления для двух рассматриваемых случаев управления портфелем имеют вид:

$$\rho^1 = IRR(X_0^1, ..., X_{n-1}^1, X_n^1) - IRR(X_0^1, ..., X_{n-1}^1, \bar{X}_n^1) = r^1 - IRR(X_0^1, ..., X_{n-1}^1, \bar{X}_n^1) \qquad (18)$$

и

$$\rho^2 = IRR(X_0^2, ..., X_{n-1}^2, X_n^2) - IRR(X_0^2, ..., X_{n-1}^2, \bar{X}_n^2) = r^2 - IRR(X_0^2, ..., X_{n-1}^2, \bar{X}_n^2). \qquad (19)$$

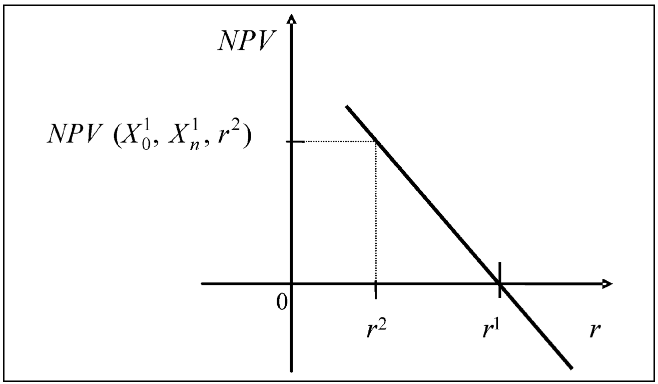


Рис. 4. Иллюстрация к доказательству справедливости аксиомы 3

Из условия аксиомы 3 вытекает, что стоимости портфелей $\bar{A}_n^1$ и $\bar{A}_n^2$, рассчитанные по соотношению (13), одинаковы. Поэтому $\bar{X}_n^1 = \bar{A}_n^1 + \Delta_n^1 = \bar{A}_n^2 + \Delta_n^2 = \bar{X}_n^2$ и $IRR(X_0^1, ..., X_{n-1}^1, \bar{X}_n^1) = IRR(X_0^2, ..., X_{n-1}^2, X_n^2)$. Из сравнения правых частей соотношений (18) и (19) получаем, что $\rho^1 > \rho^2$.

Для доказательства выполнения аксиомы 4 рассмотрим вначале интервал времени до первой трансакции $[t_0, t_1]$. По условию аксиомы $\rho = \Omega(Y_0, Y_1, R) = IRR(Y_0, Y_1) - IRR(Y_0, \bar{Y}_1) = 0$, где $Y_0 = -A_0$, $Y_1 = A_1 + \Delta_1$, $\bar{Y}_1 = A_0(1 + R(0, 1)) + \Delta_1$. Следовательно, $Y_1 = \bar{Y}_1 = A_0(1 + R(0, 1)) + \Delta_1$. Отсюда

$$A_1 = A_0(1 + R(0, 1)). \quad (20)$$

Для интервала $[t_1, t_2]$ справедливо $\rho = \Omega(Y_1, Y_2, R) = IRR(Y_1, Y_2) - IRR(Y_1, \bar{Y}_2) = 0$. Теперь $Y_1 = -A_1$, $Y_2 = A_2 + \Delta_2$, $\bar{Y}_2 = A_1(1 + R(1, 2)) + \Delta_2$. Поэтому $Y_2 = \bar{Y}_2 = A_1(1 + R(1, 2)) + \Delta_2$ и $A_2 = A_1(1 + R(1, 2))$. Подставляя в последнее равенство соотношение (20), получаем $A_2 = A_0(1 + R(0, 1))(1 + R(1, 2))$.

Продолжая рассуждать подобным образом, можно убедиться в том, что выражение для расчета A_n не отличается от выражения (12). Поэтому $A_n = \bar{A}_n$. Отсюда

$$IRR(X_0, ..., X_{n-1}, X_n) = IRR(X_0, ..., X_{n-1}, \bar{X}_n) \text{ и}$$
$$\rho = IRR(X_0, ..., X_{n-1}, X_n) - IRR(X_0, ..., X_{n-1}, \bar{X}_n) = 0.$$

Теорема доказана.

ЛИТЕРАТУРА

1. *Шарп У., Александер Г., Бэйли Д.* Инвестиции. — М.: ИНФРА-М, 1997.
2. *Kellison S. G.* The theory of interest. — Boston: Irwin, 1991.
3. *Roll R.* Performance evaluation and benchmark errors (I) // Journal of Portfolio Management. — 1980. — Vol. 6, N 4. — P. 5—12.
4. *Roll R.* Performance evaluation and benchmark errors (II) // Ibid. — 1980. — Vol. 7, N 2. — P. 17—22.
5. *Brinson G., Diermeier J. J., Schlarbaum G. G.* A composite portfolio benchmark for pension plans // Financial Analysts Journal. — 1986. — Vol. 42, N 2. — P. 15—24.
6. *Bailey J. V., Richards T. M., Tierney D. E.* Benchmark portfolios and the manager/plan sponsor relationship // Journal of Corporate Finance. — 1988. — Vol. 4, N 4. — P. 25—32.
7. *Rennie E. P., Cowhey T. J.* The successful use benchmark portfolios: a case study // Financial analysts journal. — 1990. — Vol. 46, N 5. — P. 18—26.
8. *Bailey J. V.* Are manager universes acceptable performance benchmarks? // Journal of Portfolio Management. — 1992. — Vol. 18, N 3. — P. 9—13.
9. *Bailey J. V.* Evaluating benchmark Quality // Financial Analysts Journal. — 1992. — Vol. 48, N 3. — P. 33—39.
10. *Coggin T. D., Fabozzi F. J., Rahman S.* The investment performance of U.S. equity pension fund managers: an empirical investigation // Journal of Finance. — 1993. — Vol. 48, N 3. — P. 1039—1055.
11. *Daniel K.*, et al. Measuring mutual fund performance with characteristic-based benchmarks // Ibid. — 1997. — Vol. 52. — P. 1035—1058.
12. *Haight G., Morrell S.* The Analysis of Portfolio Management: An Institutional Guide to Assessing and Analyzing Pension Fund, Endowment, Foundation and Trust Investment Performance. — N.-Y.: McGraw-Hill Trade, 1997.
13. *Treinor J. L.* How to rate management to investment funds // Harvard Business Review. — 1965. — Vol. 43, N 1. — P. 63—75.
14. *Sharpe W. F.* Mutual fund performance // Journal of Business. — Vol. 39, N 1. — P. 119—138.
15. *Jensen M. C.* The performance of mutual funds in the period 1945 — 1964 // Journal of Finance. — 1968. — Vol. 23, N 2. — P. 389—416.
16. *Jensen M. C.* Risk, the pricing of capital assets, and the evaluation of investment portfolios // Journal of Business. — 1969. — Vol. 42, N 2. — P. 167—185.
17. *Fama E. F., Kenneth R. F.* Common risk factors in the returns on bonds and stocks // Journal of Financial Economics. — 1993. — Vol. 33. — P. 3—53.
18. *Carhart M. M.* On persistence in mutual fund performance // Journal of Finance. — 1997. — Vol. 52. — P. 57—82.
19. *Hallahan T. A.* The information content of portfolio performance history and persistence in fund performance: An examination of rollover funds // Accounting and Finance. — 1999. — Vol. 39, N 3. — P. 255—274.
20. *Ованесов А., Четвериков В.* Поток платежей. Будьте осторожны с усредненными характеристиками! // Рынок ценных бумаг. — 1996. — № 21. — С. 39—42.
21. *Promislow S. D.* A new approach to the theory of interest // Trans. of the Society of Actuaries. — 1980. — Vol. 32. — P. 53—92.
22. *Treynor J. L.* How to rate management of investment funds // Harvard Business Review. — 1965. — Vol. 43, N 1. — P. 63—75.
23. *Sharpe W. F.* Mutual fund performance // Journal of Business. — 1966. — Vol. 39, N 1. — P. 119—138.
24. *Jensen M. C.* The performance of mutual funds in the period 1945 — 1964 // Journal of Finance. — 1968. — Vol. 23, N 2. — P. 389—416.
25. *Jensen M. C.* Risk, the pricing of capital assets, and the evaluation of investment portfolios // Journal of Business. — 1969. — Vol. 42, N 2. — P. 167—185.

☎ *(095) 937-07-37, доб. 26-83*
E-mail: d.golembiovsky@zenit.ru ❑

УДК 517.977.5:629.78

АДАПТИВНЫЕ ДЕКОМПОЗИРУЮЩИЕ АЛГОРИТМЫ УПРАВЛЕНИЯ ПОЛУАКТИВНОЙ СВЯЗКОЙ МЕХАНИЧЕСКИХ СИСТЕМ[1]

В. М. Суханов, Е. М. Фирсова

Институт проблем управления им. В. А. Трапезникова, г. Москва

Рассмотрена задача формирования адаптивных алгоритмов управления, обеспечивающих декомпозицию модели космического роботизированного модуля (КРМ), являющегося многосвязной нестационарной механической системой. Предложена методика синтеза алгоритма перестройки параметров регулятора на основе принципов беспоисковой адаптации с эталонной моделью, обеспечивающего желаемую динамику функционирования подсистем модуля. Исследована возможность демпфирования упругих колебаний транспортируемого модулем груза путем нестандартного применения штатных приводов манипулятора КРМ.

ВВЕДЕНИЕ

Рассматриваемый в работе космический робототехнический модуль (КРМ) является свободнолетающим маневрирующим транспортным средством, способным одновременно решать задачи поиска, захвата, транспортировки и установки полезного (не обязательно жесткого) груза в окрестности пилотируемой орбитальной станции. Как видно из рис. 1, механическая система КРМ представляет собой связку нескольких подсистем, состоящую из жесткого несущего тела (корпуса) и шарнирно присоединенной к нему подсистемы носимых тел, включающей в себя один или несколько трехзвенных манипуляторов с концевым схватом, удерживающим нежесткий полезный груз (Г), который рассматривается как третий компонент связки. Идеализированная модель манипулятора может быть определена в виде системы шарнирно связанных между собой жестких звеньев длиной r_1 и r_2. Образованную указанным способом механическую структуру для краткости обозначим КРМ-Г. Кроме того, на рис. 1 обозначено: CXY — базовая система координат, ox_0y_0 — система координат, связанная с корпусом КРМ; $oX'Y'$ — базовая система координат, смещенная в центр масс («o») несущего тела КРМ. Остальные обозначения на рисунке определены в тексте статьи ниже.

Полученная в работе [1] математическая модель свободнолетающего космического робототехнического модуля в режиме транспортировки нежесткого груза имеет вид

$$A(q)\ddot{q} + H\dot{q} + Bq = F_u(t) - \sum_{s=1}^{n}[\dot{q}^T D_k(q)\dot{q}]e_k, \quad (1)$$

где $q = (q_1 = X_0,\ q_2 = Y_0,\ q_3 = \vartheta,\ q_4 = \alpha_1,\ q_5 = \alpha_2,\ q_6 = \alpha_3,\ q_7 = \lambda_3,\ q_8 = \lambda_4)$ — вектор обобщенных координат КРМ-Г; первые три компоненты вектора q, т. е. сово-

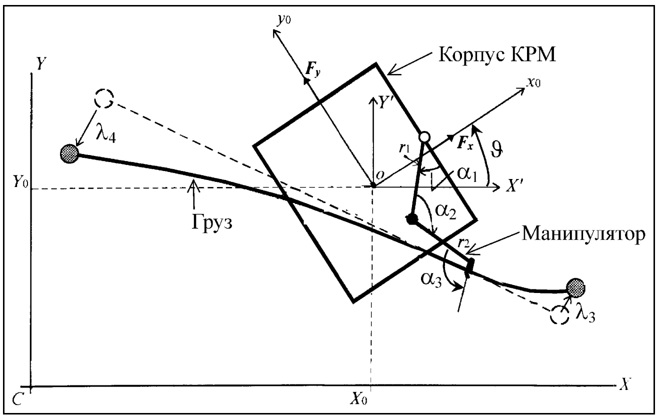


Рис. 1. Конфигурация связки «КРМ — нежесткий груз»

[1] Работа выполнена при финансовой поддержке РФФИ (проект 03-01-00062) и Программы фундаментальных исследований № 16 Отделения ЭММПУ РАН.

купность $q^0 = (X_0, Y_0, \vartheta)$, будем рассматривать как независимые обобщенные координаты, определяющие положение несущего тела (корпуса робота) в инерциальной системе координат CXY. Остальные пять компонент $q_4, ..., q_8$, которые обозначим $q^{\text{М}-\Gamma} = (\alpha_1, \alpha_2, \alpha_3, \lambda_3, \lambda_4)$, определяют положение носимых тел относительно осей, связанных с несущим телом модуля; $A(q)$ — квадратная $(n \times n)$ симметричная матрица коэффициентов инерции, являющихся функциями обобщенных координат; H, B — постоянные $(n \times n)$ матрицы, элементы которых определяются известным образом на основе диссипативной и потенциальной функций, зависящих чаще всего от физических свойств транспортируемого нежесткого груза; $F_u(t)$ — вектор управляющих сил; второе слагаемое в правой части уравнения (1) — матрица обобщенных кориолисовых и центробежных сил, порождаемых относительными (вращательными и поступательными) движениями несущего и носимых тел КРМ-Г; e_k — n-мерный единичный вектор с k-й ненулевой строкой.

Нелинейные дифференциальные уравнения (1) в рамках исследования плоского движения КРМ-Г являются наиболее общими и пригодны для описания большинства фаз (режимов) функционирования КРМ. В частности, модель (1) описывает траекторное и угловое движения связки КРМ-Г под действием реактивных сил F_x, F_y и моментов M_ϑ, M_{α_j} (создаваемых исполнительными органами системы ориентации и приводами звеньев манипулятора, соответственно) в плоскости системы координат CXY, связанной с орбитальной станцией и формируемой в пространстве ее радиотехническими средствами. Математическая модель (1) связки КРМ-Г представляет собой систему нелинейных уравнений с переменными коэффициентами и характеризуется наличием межсистемных связей, что существенно осложняет качественное решение задач управления на множестве режимов работы КРМ.

В связи с высокими требованиями к точности и безопасности функционирования КРМ вблизи поверхности орбитальной станции в условиях нестационарности и неопределенности изменения параметров КРМ в данной работе решается задача декомпозиции полной модели связки КРМ-Г на автономные подсистемы и обеспечения желаемой динамики изменения координат связки КРМ-Г на основе методов адаптивного управления.

Теория декомпозиции управляемых систем на отдельные подсистемы с последующим синтезом управлений для локальных подсистем является предметом исследований в течение ряда десятилетий [2]. В обширном списке публикаций по теории декомпозиции можно выделить два направления. В основу первого направления [3, 4], впервые сформулированного И. Н. Вознесенским, положена идея декомпозиции многосвязной системы путем формирования специальных компенсирующих обратных связей. В большинстве работ, относящихся к этому направлению, алгоритмы синтеза декомпозирующих обратных связей строятся на основе численных процедур, применение которых сопряжено с необходимостью иметь точную информацию о структуре и параметрах математических моделей управляемых систем, что делает такие процедуры малоэффективными.

Принципиально иной подход к решению задачи декомпозиции, осуществляемый не с помощью компенсирующих обратных связей, а за счет управляющих сигналов, формируемых с помощью нетрадиционных алгоритмов, учитывающих физические особенности объектов управления, развит в работе [5]. Такого типа алгоритмы придают системам свойства слабой чувствительности к изменениям параметров объекта управления, что делает их привлекательными для решения конкретных задач.

Особо актуальным как в теоретическом, так и в прикладном аспекте для осуществления декомпозиции нестационарных нелинейных систем управления, к которым относится и рассматриваемая в данной работе система управления свободнолетающим космическим роботизированным модулем, является развитие методов беспоисковой адаптации с эталонной моделью. В соответствии с этим в работе рассматривается задача формирования адаптивной системы управления КРМ-Г, которая позволяет реализовать декомпозицию системы на автономные подсистемы с желаемыми динамическими характеристиками движения.

1. УПРОЩЕННАЯ МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДИНАМИКИ КРМ-Г

Важнейшее условие безопасного функционирования роботизированного модуля вблизи поверхности орбитальной станции состоит в обеспечении предельно малых скоростей его перемещения при выполнении операций обслуживания. Выполнение этого требования позволяет упростить исходную модель (1), исключив из нее члены, содержащие произведения скоростей обобщенных координат, т. е. кориолисовы и центробежные силы как величины второго порядка малости. Кроме того, матрица H в уравнениях (1) из-за пренебрежимо малого естественного демпфирования при колебаниях упругого груза в условиях космоса считается нулевой ($H(h_{ik}) = [0]$).

В силу принятых допущений математическая модель (1) движения связки КРМ-Г принимает более простой вид

$$A(q)\ddot{q} + Bq = F_u(t), \qquad (2)$$

где $q = (q^0, q^m, q^r)^T$ — вектор обобщенных координат КРМ, в котором по сравнению с моделью (1) подвектор координат носимых тел разбит на два вектора, один из которых $q^{\text{М}} = (\alpha_1, \alpha_2, \alpha_2)^T$ определяет конфигурацию манипулятора в осях корпуса КРМ, а второй, $q^\Gamma = (q_7 = \lambda_3, q_8 = \lambda_4)^T$ — упругие смещения концевых масс нежесткого груза относительно его исходного (недеформированного) состояния, определенного на рис. 1 пунктирным отображением; B — постоянная (8×8) матрица упругости вида

$$B = \begin{bmatrix} [0] & [0] \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} b_4 & 0 \\ 0 & b_8 \end{bmatrix} \end{bmatrix}; \qquad (3)$$

$F_u(t) = (F_{u1}, ..., F_{un})^T$, $n = 8$ — вектор-столбец управляющих сил. В рассматриваемой постановке задачи полезный груз, будучи пассивным элементом связки, не имеет собственного управления, что дает основание положить F_{u7}, $F_{u8} \equiv 0$. Коэффициенты $a_{ik}(q)$, являющиеся элементами матрицы $A(q)$, не одинаковы для различных режимов функционирования КРМ при обслуживании орбитальной станции; их выражения и основные характеристики этих режимов приведены в работе [1].

Переменные, определяемые векторами $q^0 = (X_0, Y_0, \vartheta)$, $q^{М} = (\alpha_1, \alpha_2, \alpha_2)^T$, $q^{Г} = (\lambda_3, \lambda_4)^T$, следует рассматривать как группы координат подсистем КРМ-Г, отличающихся как их функциональным назначением, так и по способам управления ими. Особенностью объектов, описываемых уравнениями (2), является взаимовлияние движений по различным группам координат, проявляющееся через недиагональные элементы матрицы $A(q)$, т. е. через коэффициенты $a_{ik}(q)$, $i \neq j$, что существенно осложняет управление движением различных подсистем КРМ-Г, снижая точность и безопасность работы модуля вблизи орбитальной станции. Далее для краткости переменные элементы матрицы $A(q)$ будем записывать в виде $a_{ij}(q) \square a_{ij}$.

2. ДЕКОМПОЗИЦИЯ СВЯЗКИ КРМ-Г НА АВТОНОМНЫЕ ПОДСИСТЕМЫ

Синтез структуры адаптивной системы управления проводится в два этапа. На первом из них в предположении, что коэффициенты математической модели объекта известны и постоянны (гипотеза квазистационарности), формируется структура основного контура, т. е. определяется базовый алгоритм управления, обеспечивающий достижение цели управления на рассматриваемом режиме функционирования КРМ; определяется количество и места введения перестраиваемых параметров регулятора. На втором этапе осуществляется синтез алгоритмов адаптации, обеспечивающих выполнение требуемого качества управления для любого $t > 0$ [6]. Для определенности рассмотрим режим, введенный в работе [1] как фаза транспортировки груза. В этом режиме математическая модель движения системы КРМ-Г (2) с учетом структуры матриц (3) и $A(q)$ может быть переписана в виде следующей системы уравнений:

$$
\begin{gathered}
a_{ii}\ddot{q}_i + \sum_{k=1,\,k\neq i}^{8} a_{ik}\ddot{q}_k = F_{u_i}(t), \quad i = \overline{1,6} \\
a_{77}\ddot{q}_7 + b_7 q_7 + \sum_{k=1,\,k\neq 7}^{8} a_{7k}\ddot{q}_k = 0 \qquad (4)\\
a_{88}(q)\ddot{q}_8 + b_8 q_8 + \sum_{k=1,\,k\neq 8}^{8} a_{8k}\ddot{q}_k = 0,
\end{gathered}
$$

где $a_{78} = a_{87} = 0$. Коэффициенты a_{ik} характеризуют взаимовлияние координат связки КРМ-Г.

Для придания свойства автономности подсистемам управления координатами q^0, $q^{М}$, $q^{Г}$, обеспечения желаемой динамики изменения координат q^0, $q^{М}$ и стабилизации колебаний упругого груза $q^{Г}$ с помощью манипулятора, сформируем закон управления в виде

$$
F_{ui}^{Add2} = K_{1i}\dot{q}_i + K_{2i}q_i + \sum_{k=1,\,k\neq i}^{8} k_{ik}\hat{\ddot{q}}_k, \quad i = \overline{1,6}, \qquad (5)
$$

где $\hat{\ddot{q}}_k$ — оценки вторых производных от обобщенных координат, полученные с помощью наблюдающего устройства, вопросы реализации которого пока не обсуждаются. Предполагается, что оценки $\hat{\ddot{q}}_k$ получены с достаточной степенью точности, вследствие чего $\hat{\ddot{q}}_k \cong \ddot{q}_k$.

С учетом выражения (5) и в предположении, что $a_{ii}(q) \neq 0$, уравнения движения (4) по координатам q_i, $i = \overline{1,6}$, в замкнутой системе примут следующий вид

$$
\begin{gathered}
\ddot{q}_i + (K_{2i}\tilde{a}_{ii})q_i + (K_{1i}\tilde{a}_{ii})\dot{q}_i + \sum_{k=1,\,k\neq i}^{8} (\tilde{a}_{ik}\ddot{q}_k + k_{ik}\hat{\ddot{q}}_k) = \\
= (K_{i0}\tilde{a}_{ii})F_{u_i}(t), \qquad (6)
\end{gathered}
$$

где $\tilde{a}_{ik}(q) = \dfrac{a_{ik}(q)}{a_{ii}(q)}$, $\tilde{a}_{ik}(q) = \dfrac{1}{a_{ii}(q)}$, $\hat{\ddot{q}}_k \cong \ddot{q}_k$.

Желаемую динамику изменения координат $q_i(t)$ определим в виде полученной из уравнений (4) следующей системы независимых уравнений второго порядка:

$$
\begin{gathered}
\ddot{q}_i^0 + \mu_{1i}\dot{q}_i^0 + \mu_{2i}q_i^0 = F_{ui}(t), \\
\mu_{1i}, \mu_{2i} = \mathrm{const} > 0, \quad i = \overline{1,6}. \qquad (7)
\end{gathered}
$$

Принципиальный вопрос синтеза основного контура беспоисковой адаптивной системы заключается в отыскании условий, обеспечивающих тождественность уравнений движения координат $q_i(t)$ замкнутой системы и уравнений желаемых движений координат $q_i^0(t)$.

Необходимым и достаточным условием тождественности уравнений (6) и уравнений (7) является выполнение следующих соотношений:

$$
\tilde{a}_{ik}(q) + k_{ik} = 0,\ i = \overline{1,6},\ k = \overline{1,8}, \qquad (8)
$$

$$
(K_{1i}\tilde{a}_{ii}) = \mu_{1i},\ (K_{2i}\tilde{a}_{ii}) = \mu_{2i},\ (K_{0i}\tilde{a}_{ii}) = 1. \qquad (9)
$$

Действительно, подставляя эти соотношения в уравнения (6), получим систему независимых уравнений (7). Это доказывает, что условия (8) являются условиями автономности координат $q_i(t)$, а условия (9) обеспечивают желаемые динамические характеристики движений координат $q_i(t)$, т. е. система уравнений (6) декомпозируется на $n = 6$ автономных уравнений движения (7) с заданными динамическими характеристиками.

Таким образом показано, что на основе предложенного закона управления (5) и реализации условий (8) и (9) имеется возможность сформировать структуру регулятора, определить количество и места введения перестраиваемых параметров, целенаправленное изменение которых в соответствии с требованиями к качеству про-

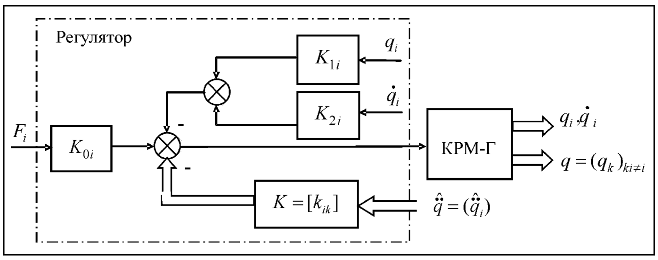


Рис. 2. Структурная схема основного контура беспоисковой адаптивной системы

цессов управления обеспечивает декомпозицию связки КРМ-Г на автономные подсистемы по координатам q_i.

Структурная схема основного контура беспоисковой адаптивной системы, реализующая алгоритм (5), приведена на рис. 2, из которого видно, что обратные перекрестные связи с коэффициентами $k_{ik} \in K$, $i \neq k$, обеспечивают компенсацию взаимовлияния между каналами управления координатами q_i и q_k, а обратные связи с коэффициентами K_{1i}, K_{2i} и коэффициент K_{0i} в прямой цепи обеспечивают желаемую динамку изменения координат $q_i(t)$.

3. СИНТЕЗ АДАПТИВНЫХ БЕСПОИСКОВЫХ АЛГОРИТМОВ, ОБЕСПЕЧИВАЮЩИХ АВТОНОМНОСТЬ И ЖЕЛАЕМУЮ ДИНАМИКУ ФУНКЦИОНИРОВАНИЯ КРМ-Г

Для синтеза алгоритмов беспоисковой перестройки коэффициентов регулятора применяется процедура, основанная на прямом методе Ляпунова [6]. Физически реализуемые эталонные модели автономного движения КРМ-Г по координатам q_i в рассматриваемом режиме транспортировки упругого груза описываются уравнениями

$$\ddot{q}_{i\text{М}} + \mu_{1i}\dot{q}_{i\text{М}} + \mu_{2i}q_{i\text{М}} = F_{ui}(t), \quad i = \overline{1, 6}. \tag{10}$$

Для синтеза алгоритмов беспоисковой перестройки коэффициентов регулятора рассмотрим уравнение, описывающее изменение координат q_i в замкнутой системе

$$\ddot{q}_i + (K_{1i}\tilde{a}_{ii})\dot{q}_i + (K_{2i}\tilde{a}_{ii})q_i + \sum_{k=1,\, k\neq i}^{8}[\tilde{a}_{ik}\ddot{q}_k + k_{ik}\hat{\ddot{q}}_k] = (K_{0i}\tilde{a}_{ii})F_{u_i}(t), \tag{11}$$

где $k_{ik}(t)$, $K_{1i}(t)$, $K_{2i}(t)$ и $K_{0i}(t)$ — коэффициенты, перестраиваемые по алгоритмам беспоисковой адаптации, обеспечивающим выполнение соотношений (8) и (9) для любого $t > 0$.

На режиме, отличном от номинального, уравнение (11) примет вид:

$$\ddot{q}_i + (K_{1i} + \Delta K_{1i})[\tilde{a}_{ii} + \Delta\tilde{a}_{ii}]\dot{q}_i + (K_{2i} + \Delta K_{2i})[\tilde{a}_{ii} + \Delta\tilde{a}_{ii}]q_i + \sum_{k=1,\, k\neq i}^{8}[\tilde{a}_{ii}(q) + \Delta\tilde{a}_{ik}(q)]\ddot{q}_k + \sum_{k=1,\, k\neq i}^{8}(k_{ik} + \Delta k_{ik})\hat{\ddot{q}}_k = (K_{0i} + \Delta K_{0i})[\tilde{a}_{ii}(q) + \Delta\tilde{a}_{ik}(q)]F_{u_i}(t), \tag{12}$$

где $\Delta\tilde{a}_{ik}(q)$ и $\tilde{a}_{ii}(q)$ — параметрические возмущения, $\Delta k_{ik}(t)$, $\Delta K_{1i}(t)$, $\Delta K_{2i}(t)$ и $\Delta K_{0i}(t)$ — приращения перестраиваемых коэффициентов регулятора.

Пренебрегая произведениями $\Delta\tilde{a}_{ii}\Delta K_{0i}$, $\Delta\tilde{a}_{ii}\Delta K_{1i}$, $\Delta\tilde{a}_{ii}\Delta K_{2i}$, $\Delta\tilde{a}_{ik}\Delta k_{ik}$ как величинами второго порядка малости, из сравнения уравнений (10) и (12) с учетом условий (8) и (9) получим уравнение относительно координатной ошибки ε_i:

$$\ddot{\varepsilon}_i + \mu_{1i}\dot{\varepsilon}_i + \mu_{2i}\varepsilon_i = Y_{1i}\dot{q}_i + Y_{2i}q_i + Z_{0i}F_{ui} + \sum_{k=1,\, k\neq i}^{8} y_{ik}\hat{\ddot{q}}_k, \tag{13}$$

где $Z_{0i} = k_{0i}\Delta\tilde{a}_{ii}\Delta k_{0i}$; $y_{ik} = k_{ik}\Delta\tilde{a}_{ik} + \tilde{a}_{ik}\Delta k_{ik}$; $Y_{1i} = K_{1i}\Delta\tilde{a}_{ii} + \tilde{a}_{ii}\Delta K_{1i}$; Z_{0i}, Y_{1i}, Y_{2i}, y_{ik} — параметрические рассогласования; $\varepsilon_i = q - q_{\text{М}}$ — отклонения координат системы от модели (ошибки).

Обозначив $\varepsilon_i = x_{1i}$, $\dot{\varepsilon}_i = x_{2i}$, представим уравнение (13) в матричной форме

$$\dot{x}_i = A_i x_i + \rho_i, \tag{14}$$

где $x_i = \begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix}$, $A_i = \begin{bmatrix} 0 & 1 \\ -\mu_{2i} & -\mu_{1i} \end{bmatrix}$, $\rho_i = \begin{bmatrix} 0 \\ \rho_{2i} \end{bmatrix}$, $\rho_{2i} = Y_{1i}\dot{q}_i + Y_{2i}q_i + Z_{0i}F_{ui} + \sum\limits_{k=1,\, k\neq i}^{8} y_{ik}\hat{\ddot{q}}_k$.

Зададим алгоритмы адаптации в неявном виде:

$$\frac{d}{dt}\Delta k_{ik} = \Psi_{ik}, \quad \frac{d}{dt}\Delta K_{1i} = \Psi_{1i}, \quad \frac{d}{dt}\Delta K_{2i} = \Psi_{2i}, \quad \frac{d}{dt}\Delta K_{0i} = \Psi_{0i}. \tag{15}$$

С учетом принятых обозначений получим:

$$\dot{y}_{ik} = \psi_{ik} + r_{\text{ik}}, \quad \dot{Y}_{1i} = \Psi_{1i} + r_{1i}; \quad \dot{Y}_{2i} = \Psi_{2i} + r_{2i}, \quad \dot{Z}_{0i} = \Psi_{0i} + r_{ii},$$

где ψ_{ik}, Ψ_{1i}, Ψ_{2i}, Ψ_{0i} — искомые алгоритмы адаптации; $r_{1i} = \frac{d}{dt}(K_{1i}\Delta\tilde{a}_{ii})$, $r_{2i} = \frac{d}{dt}(K_{2i}\Delta\tilde{a}_{ii})$ — скорости изменения параметрических возмущений.

Для рассматриваемого класса объектов можно считать, что параметрические возмущения при $t \geqslant t_0$ неопределенные, но постоянные, т. е.

$$r_{ik} = 0, \quad r_{0i} = 0, \quad r_{1i} = 0, \quad r_{2i} = 0. \tag{16}$$

Из уравнения (14) с учетом условий (15) и (16) получим систему:

$$\dot{x}_i = A_i x_i + \rho_i, \ \dot{Y}_{ik} = \Psi_{yik}, \ \dot{Y}_i = \Psi_Y, \ \dot{Z}_i = \Psi_Z, \tag{17}$$

где $Y_{ik} = \begin{bmatrix} 0 \\ y_{ik} \end{bmatrix}$, $Y_i = \begin{bmatrix} Y_{1i} \\ Y_{2i} \end{bmatrix}$, $Z_i = \begin{bmatrix} 0 \\ Z_{0i} \end{bmatrix}$, $\Psi_{yik} = \begin{bmatrix} 0 \\ \Psi_{yik} \end{bmatrix}$, $\Psi_Y = \begin{bmatrix} \Psi_{1i} \\ \Psi_{2i} \end{bmatrix}$, $\Psi_Z = \begin{bmatrix} 0 \\ \Psi_z \end{bmatrix}$.

Идеальной работе контуров адаптации соответствуют тождества

$$\varepsilon_i^{(\nu)} \equiv 0, \quad (\nu = 0, 1), \quad y_{ik} = 0, \quad Y_{1i} = 0, \quad Y_{2i} = 0, \quad Z_{0i} = 0. \tag{18}$$

Если движение, определяемое системой (17), ограничить пространством $\{x_i, Y_{ik}, Y_i, Z_i\}$, то выражения (18) сводятся к тождествам вида

$$x_i \equiv 0, \quad Y_{ik} \equiv 0, \quad Y_i \equiv 0, \quad Z_i \equiv 0, \tag{19}$$

которые являются нулевым решением матричного уравнения (14).

Определим алгоритмы перестройки параметров регулятора из условия устойчивости нулевого решения (19) уравнения (14). Для этого, воспользовавшись вторым

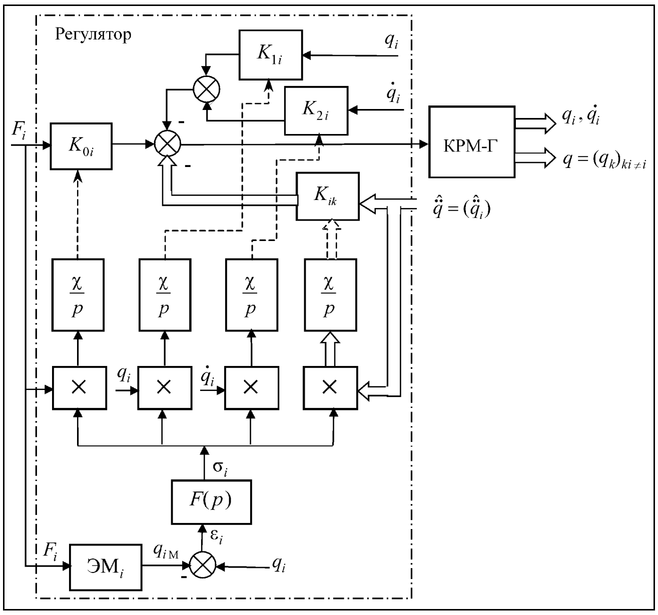


Рис. 3. Структурная схема декомпозирующей адаптивной системы с эталонной моделью *ЭМ$_i$*

методом Ляпунова [6], выберем квадратичную форму V_i в следующем виде:

$$V_i = \chi x_i^T P_i x_i + \sum_{k=1,\, k \ne i}^{8} Y_{ik}^T E Y_{ik} + Y_i^T E_2 Y_i + Z_i^T E_3 Z_i, \quad (20)$$

где P_i — симметричная (2×2) матрица; E_1, E_2 и E_3 — единичные (2×2) матрицы; $\chi = \text{const} > 0$.

Воспользовавшись уравнением (14), запишем производную функции Ляпунова

$$\dot V_i = \chi x_i^T Q_i x_i + 2\xi x_i^T P_i \rho_i + 2 \sum_{k=1,\, k \ne i}^{8} Y_{ik}^T E \Psi_{yik} + 2 Y_i^T E_2 \Psi_Y + 2 Z_i^T E_3 \Psi_Z, \quad (21)$$

где Q_i — отрицательно определенная матрица (2×2) вида $Q_i = A_i^T P_i + P_i A_i$. Матрица A_i является неособой матрицей с отрицательными действительными частями корней характеристического уравнения. В этом случае отрицательно определенной квадратичной форме $x_i^T Q_i x_i$ в выражении (21) соответствует заведомо положительно определенная квадратичная форма $x_i^T P_i x_i$ в уравнении (20). Алгоритмы адаптации ищутся из условия неположительности величины

$$\chi x_i^T P_i \rho_i + \sum_{k=1,\, k \ne i}^{8} Y_{ik}^T E_1 \psi_{ik} + Y_i^T E_2 \Psi_Y + Z_i^T E_3 \Psi_z \leqslant 0. \quad (22)$$

Поскольку при выполнении условия (22) имеет место $\dot V_i < 0$, а функция V_i при этом является положительно определенной, то нулевое решение (19) уравнения (14) устойчиво.

Раскрывая неравенство (22), в результате получим:

$$\chi\sigma_i \left(Y_{1i} \dot q_i + Y_{2i} q_i + Z_{0i} F_{ui} + \sum_{k=1,\, k \ne i}^{8} y_{ik} \hat{\hat q}_k \right) + \left(Y_{1i} \Psi_{1i} + Y_{2i} \Psi_{2i} + Z_{0i} \Psi_{0i} + \sum_{k=1,\, k \ne i}^{8} y_{ik} \psi_{ik} \right) \leqslant 0, \quad (23)$$

где $\sigma_i = p_{22} \dot\varepsilon_i + p_{21} \varepsilon_i$, p_{21} и p_{22} — элементы матрицы P, которые определяются при решении матричного уравнения $A_i^T P_i + P_i A_i = Q_i$.

Неравенство (23) выполняется, если алгоритмы адаптации принять в виде

$$\frac{d}{dt} \Delta k_{ik} = \psi_{ik} = \chi \sigma_i \hat{\hat q}_i, \quad \frac{d}{dt} \Delta K_{1i} = \Psi_{1i} = \chi \sigma_i \dot q_i,$$

$$\frac{d}{dt} \Delta K_{2i} = \Psi_{2i} = \chi \sigma_i q_i, \quad \frac{d}{dt} \Delta K_{0i} = \Psi_{0i} = \chi \sigma_i F_{ui}, \quad (24)$$

что теоретически завершает задачу синтеза беспоисковой адаптивной системы с эталонной моделью.

Структурная схема синтезированной таким образом системы, которая реализует автономность и желаемое качество движения по координатам q_i в соответствии с законом управления (5) и алгоритмами беспоисковой перестройки параметров регулятора (24), приведена на рис. 3, где пунктирными стрелками выделены цепи настройки соответствующих коэффициентов регулятора.

4. ДЕМПФИРОВАНИЕ УПРУГИХ КОЛЕБАНИЙ ГРУЗА С ПОМОЩЬЮ ПРИВОДОВ МАНИПУЛЯТОРА

Декомпозиция связки трех механических систем КРМ-Г на автономные подсистемы с желаемой динамикой позволяет решить задачу демпфирования упругих колебаний транспортируемого пассивного полезного груза с помощью управляющих воздействий, создаваемых исполнительными органами манипулятора.

Закон управления (5) и алгоритмы беспоисковой перестройки параметров регулятора (24) позволяют декомпозировать движение связки КРМ-Г на автономные подсистемы, описываемые уравнениями

$$\ddot q_i + \mu_{1i} \dot q_i + \mu_{2i} q_i = F_{ui}(t), \ \mu_{1i}, \mu_{2i} = \text{const} > 0, \quad i = \overline{1, 6}, \quad (25)$$

$$a_{77}(q) \ddot q_7 + b_7 q_7 + \sum_{i=1}^{6} a_{7i} \ddot q_i = 0,$$

$$a_{88}(q) \ddot q_8 + b_8 q_8 + \sum_{i=1}^{6} a_{8i} \ddot q_i = 0, \quad (26)$$

где $q = (q_1 = X_0,\ q_2 = Y_0,\ q_3 = \vartheta,\ q_4 = \alpha_1,\ q_5 = \alpha_2,\ q_6 = \alpha_3,\ q_7 = \lambda_3,\ q_8 = \lambda_4)^T$ — вектор обобщенных координат связки КРМ-Г. Напомним, что координаты X_0, Y_0 и ϑ определяют положение корпуса робота в базовой системе координат, координаты α_1, α_2 и α_3 — положение звеньев манипулятора относительно корпуса КРМ; координаты $q_7 = \lambda_3$ и $q_8 = \lambda_4$ — упругие смещения нежесткого полезного груза относительно его недеформированного состояния, заданного в системе координат схвата манипулятора; $F_{ui}(t)$, $i = \overline{1, 6}$ — элементы вектора-столбца управляющих сил, причем $F_{u7}(t) = 0$, $F_{u8}(t) = 0$, в силу пассивности транспортируемого груза.

Из уравнений (26) видно, что движения по координатам $q_7 = \lambda_3$ и $q_8 = \lambda_4$ представляют собой незатухающие колебания.

Ставится задача формирования демпфирующей составляющей в уравнениях (26) в условиях отсутствия собственных управляющих воздействий по координатам $q_7 = \lambda_3$ и $q_8 = \lambda_4$. Предлагается следующий подход к решению этой задачи.

Допустим, что алгоритм управления (5) и алгоритмы беспоисковой адаптации (24) обеспечивают выполнение условий декомпозиции связки КРМ-Г на автономные подсистемы. При отсутствии упругости в шарнирах и звеньях манипулятора КРМ в уравнениях (25) можно положить μ_{1i}, $\mu_{2i} \cong 0$. Это позволяет записать:

$$\ddot{q}_i = F_{ui}(t), \quad i = \overline{1, 6}. \tag{27}$$

Подставляя соотношения (27) в уравнения колебаний груза (26), получим

$$a_{77}(q)\ddot{q}_7 + b_7 q_7 = -\sum_{i=1}^{6} a_{7i}F_{ui}(t),$$
$$a_{88}(q)\ddot{q}_8 + b_8 q_8 = -\sum_{i=1}^{6} a_{8i}F_{ui}(t). \tag{28}$$

Перепишем уравнения (28) следующим образом:

$$a_{77}(q)\ddot{q}_7 + b_7 q_7 = -\sum_{i=1}^{3} a_{7i}F_{ui}(t) - \sum_{i=4}^{6} a_{7i}F_{ui}(t),$$
$$a_{88}(q)\ddot{q}_8 + b_8 q_8 = -\sum_{i=1}^{3} a_{8i}F_{ui}(t) - \sum_{i=4}^{6} a_{7i}F_{ui}(t).$$

Предполагая, что $a_{jj} \neq 0$, и вводя обозначения $\tilde{b}_j = b_j a_{jj}^{-1}$, $\tilde{a}_{ji} = a_{ji}a_{jj}^{-1}$, $i = \overline{1, 6}$; $j = 7, 8$, перепишем последние уравнения в общем виде

$$\ddot{q}_j + \tilde{b}_j q_j = -(F_{\text{Т}} + F_{\text{М}})_j, \tag{29}$$

где $F_{\text{Т}j} = -\sum_{1}^{3} \tilde{a}_{ji}F_{ui}(t)$ — управляющие воздействия по координатам X_0, Y_0 и ϑ, определяющим траекторное и угловое движение связки КРМ-Г; $F_{\text{М}j} = -\sum_{4}^{6} \tilde{a}_{ji}F_{ui}(t)$ — силовые воздействия, создаваемые приводами манипулятора.

В режиме транспортировки имеется возможность переопределить управляющие силы приводов манипулятора $F_{\text{М}j}$ в соответствии с задачей стабилизации колебаний упругого груза, формируя воздействия $F_{\text{М}j}$ из условия введения в уравнения движения (29) демпфирующей компоненты по модальным координатам $q_7 = \lambda_3$, $q_8 = \lambda_4$:

$$(F_{\text{Т}} + F_{\text{М}})_j = 2\xi_j\omega_j\dot{q}_j. \tag{30}$$

Из условия (30) получим алгоритм формирования управляющих воздействий $(F_{\text{М}})_j$ в виде

$$(F_{\text{М}})_j = 2\xi_j\omega_j\dot{q}_j - (F_{\text{Т}})_j, \quad \xi_j = \text{const} > 0, \quad \omega_j = \sqrt{\tilde{b}_j}. \tag{31}$$

Подставляя соотношения (31) в уравнения (29), получим уравнения, описывающие поведение координат упругого груза $q_7 = \lambda_3$, $q_8 = \lambda_4$ под действием сформированного указанным выше способом управления

$$\ddot{q}_j = 2\xi_j\omega_j\dot{q}_j + \omega_j^2 q_j = 0, \quad j = 7, 8.$$

Это — уравнения колебательного звена с затуханием, что позволяет считать решение задачи демпфирования упругих колебаний пассивного груза с помощью приводов манипулятора завершенной.

ЗАКЛЮЧЕНИЕ

Полученные результаты показывают, что применение принципа беспоисковой адаптации теоретически позволяет решить задачу высокоточного и безопасного управления нестационарной нелинейной механической системой типа «управляемая платформа — манипулятор — упругий груз», минуя процедуру определения нестационарных параметров математической модели объекта управления.

Последующие исследования должны быть направлены на выявление условий практической реализуемости предложенных алгоритмов адаптации, позволяющих декомпозировать математическую модель рассмотренного в работе роботизированного космического модуля на ряд независимых подсистем управления с желаемой динамикой.

ЛИТЕРАТУРА

1. *Рутковский В. Ю., Суханов В. М.* Динамическая модель свободнолетающего космического робототехнического модуля // Автоматика и телемеханика. — 2000. — № 5.
2. *Шильяк Д.* Децентрализованное управление сложными системами. — М.: Мир, 1994.
3. *Вознесенский И. Н.* О регулировании машин с большим числом регулируемых параметров // Автоматика и телемеханика. — 1938. — № 4.
4. *Крутько П. Д.* Аналитическое решение задачи Вознесенского для стационарных и нестационарных линейных систем // Изв. РАН. Теория и системы управления. — 1995. — № 4.
5. *Крутько П. Д., Черноусько Ф. Л.* Декомпозирующие алгоритмы управления движением нелинейных динамических систем // Изв. РАН. Теория и системы управления. — 2001. — № 4.
6. *Петров Б. Н., Рутковский В. Ю., Крутова И. Н., Земляков С. Д.* Принципы построения и проектирования самонастраивающихся систем управления. — М.: Машиностроение, 1972.

☎ *(095) 334-87-79*

E-mail: suhv@ipu.ru ❑

50 ЛЕТ В НАУКЕ
(к 75-летию Ивери Варламовича Прангишвили)

В истории компьютеризации нашего Отечества известно немало драматических, а иногда и анекдотических ситуаций, однако имели место и героические эпизоды, когда, например, в 1972 г. отечественная научная идея в области вычислительной техники была признана перспективной законодателями в этой сфере — специалистами корпорации CDC (Control Data Corporation) — лидерами в области высокопроизводительных ЭВМ. Причем признание было основано на мнении 15-ти специалистов во главе с вице-президентом фирмы после детального ознакомления с идеей, представленной доктором технических наук, профессором И. В. Прангишвили в виде проекта вычислительной системы с перестраиваемой структурой (ПС). Результатом двухнедельного изучения проекта был отчет американских специалистов, заканчивающийся предложением о совместной разработке фирмой CDC и советской стороной вычислительного комплекса на базе разработок Института проблем управления, а также строительства в СССР завода по производству вычислительных систем. С этим предложением и отчетом американских специалистов вице-президент фирмы CDC обратился к министру СССР К. Н. Рудневу. Возможно, что реализация этих предложений могла бы серьезно способствовать развитию отечественной вычислительной техники в стране. Но, к сожалению, открывшимся захватывающим перспективам не суждено было сбыться — как раз в это время в США был принят закон Джексона — Веника, запрещающий фирмам США такого рода деятельность (этот закон не отменен и по сей день).

В этом году исполняется 75 лет со дня рождения Ивери Варламовича Прангишвили и 50 лет его научной деятельности. Он активно работает и как ученый, и как директор Института проблем управления РАН им. В. А. Трапезникова, но юбиляр не пожелал каких-либо публикаций официального характера. Посовещавшись, мы — его коллеги, друзья, ученики и соратники — решили поделиться своими воспоминаниями о наиболее ярких эпизодах его (и нашей) научной биографии (в формате заочного круглого стола, проводимого Ведущим от редакции).

Ведущий. Из биографии[1] Ивери Варламовича известно, что еще в школе он проявил незаурядные способности в математике, был победителем городских и республиканских олимпиад. На выпускном экзамене комиссия, состоявшая из преподавателей школы и представителей вузов, рекомендовала ему поступить на механико-математический факультет, однако Ивери предпочел быть инженером и заниматься практической работой. Вероятно, сказалось влияние отца — Варлам Павлович Прангишвили был горным инженером, работал начальником шахт в Ткибули, Ткварчели, Бзыби, а затем главным инженером «Грузвзрывпрома». В 1949 г. Ивери Варламович поступил в Грузинский политехнический институт, который окончил в 1952 г. по специальности «электрические станции, сети и системы». По окончании института ему как сталинскому стипендиату предложили аспирантуру. Однако молодой человек решил практически освоить профессию инженера и, чтобы пройти её различные этапы, попросился на распределении в проектную организацию — Тбилисское отделение Гидроэнергопроекта, где работал инженером-проектировщиком, затем перешел в центральную лабораторию Грузэнерго на должность инженера, позже стал старшим инженером. Работал с системами телемеханики электростанций, разбросанных по всей Грузии. Эти системы только что начали внедрять для управления электрическими сетями с центрального диспетчерского пункта, размещенного в Тбилиси. Хорошо освоил телемеханику, понял ее слабые стороны и решил заниматься развитием и усовершенствованием систем телеуправления, телеизмерения и телеконтроля; появились идеи, как практически это сделать, но требовалось научное обоснование. Из литературы узнал, что научными основами телемеханики занимаются д-р техн. наук, профессор М.А. Гаврилов и его лаборатория в Институте автоматики и телемеханики (ИАТ — так тогда назывался наш Институт)

[1] Материалы биографического характера заимствованы из автобиографии И. В. Прангишвили и книги: Институту проблем управления им. В. А. Трапезникова — 65 лет. — М.: ИПУ, 2004. — 424 с.

АН СССР. Со свойственной ему непосредственностью связался по телефону с профессором М. А. Гавриловым, сказал, что хотел бы поступить к нему в аспирантуру по телемеханике, и получил предварительное согласие. Основной экзамен по специальности принимали М. А. Гаврилов и Я. З. Цыпкин и оценили его ответы на «отлично» (с некоторым удивлением, так как они редко ставили «отлично»). В аспирантуре его научным руководителем назначили М. А. Гаврилова.

Ивери Варламович аспирантский период вспоминает с огромным удовлетворением, особенно отмечает научную атмосферу конференций, семинаров, на которых нередко возникали споры и столкновения различных научных школ. В пятидесятые годы наш Институт стал очень сильным академическим учреждением, не имеющим конкурентов по вопросам фундаментальных исследований в области автоматики и телемеханики. Именно поэтому далеко не всем желающим удавалось попасть в число его сотрудников. Чем вам, коллегам того времени, помнится его аспирантская пора?

П. П. Пархоменко *(однокашник Ивери Варламовича по аспирантуре, ныне чл.-корр. РАН).* Присущее Ивери обаяние, незаурядные способности и редкостная работоспособность, помноженные на кавказскую доброжелательность и вежливость, практически сразу снискали уважение, а затем и любовь всех сотрудников лаборатории № 3 к молодому умному и интересному новому аспиранту Михаила Александровича Гаврилова (МАГ'а — как мы все уже тогда его называли).

В те годы проблемы бесконтактной техники, в том числе телемеханики, были новыми и перспективными. Именно этой области исследований были посвящены учеба и работа аспиранта Ивери Варламовича Прангишвили. Научный руководитель был глубоко удовлетворен своим аспирантом, высоко ценил результаты его исследований, продолжавших то направление в науке, молодым родителем и затем заботливым отцом которому был М. А. Гаврилов.

В. А. Жожикашвили *(«микрошеф» Ивери Варламовича по аспирантуре, позже Главный конструктор первой отечественной системы массового обслуживания СИРЕНА, сегодня — д-р техн. наук, профессор, зав. лабораторией № 17).* Ивери Варламович поступил в аспирантуру ИАТ в 1955 году. Он был молчаливым человеком с красивой улыбкой и очень дружелюбным, и его хорошо встретили в коллективе. В это время в лаб. № 3 была группа, которая занималась бесконтактной телемеханикой. МАГ определил Ивери Варламовича в эту группу. На пути к бесконтактной телемеханике группа встретилась с серьезными трудностями, связанными с тем, что не удавалось разработать бесконтактные выходные реле. Ивери Варламовичу удалось впервые создать такое реле на основе магнитно-транзисторной конструкции и колебательного контура. Он же разработал теорию устойчивости таких реле. Это был первый серьезный вклад молодого ученого в науку. Реле стало настолько популярным в среде специалистов, что его стали называть «реле Прангишвили».

Для телемеханики того времени была характерна передача на расстояние только одной компоненты (по принципу «включен — отключен»). Ивери Варламович вывел телемеханику на уровень передачи программ вместо простых команд и назвал эту технику «телеавтоматикой». Это был прорыв в неизведанную область.

Е. В. Бабичева *(сотрудница лаб. № 3 в те годы, позже — бессменный зам. зав. лабораторией № 31 — детища Ивери Варламовича).* Молодого красивого с небольшими усами грузина лаборатория № 3 (равно, как и весь Институт) встретила положительно. Освоив теорию релейно-контактных схем, он применил ее в практической разработке телеавтоматического устройства. Разработанное на новых принципах и новейших по тому времени элементах — ферритах — устройство оказалось весьма востребованным в решении ряда практических задач.

В. Д. Малюгин *(первый аспирант Ивери Варламовича, ныне д-р техн. наук, профессор, зав. Отделом аспирантуры и докторантуры Института).* Я пришел в ИАТ уже после того, как Ивери Варламович окончил аспирантуру, но в то время активно пересказывался анекдот «Как И. В. Прангишвили не стал чемпионом Москвы».

Поступив в аспирантуру, И. В. Прангишвили решил и дальше заниматься полюбившимся видом спорта — борьбой. Нашел секцию, приступил к тренировкам и как результативный спортсмен был направлен на сборы в Подмосковье. Отсутствие аспиранта Прангишвили заметил его руководитель МАГ. Когда он обнаружил, что аспиранта нет и второй, и третий день, то решительно написал служебную записку в аспирантуру об исключении своего подшефного. Руководство затормозило процесс исключения, а сотрудники лаборатории приложили большие усилия, чтобы передать о надвигающемся бедствии на сборы борцов. Прангишвили очень быстро появился в лаборатории, но со спортом пришлось проститься, как и с реальной надеждой на звание абсолютного чемпиона Москвы в тяжелом весе. Так приходят в науку.

Ведущий. Как прошла защита и что потом?

П. П. Пархоменко. Защита диссертации Ивери Варламовича прошла «без сучка и задоринки», внедрение результатов протекало активно. Достаточно упомянуть автоматизацию штаба ракетных войск и автоматизацию Гороховецкого полигона.

Е. В. Бабичева. Кандидатскую диссертацию Ивери Варламович защитил в 1959 году, имея 10 опубликованных работ и внедрение разработанных систем в Московском и Киевском военных округах. А дальше была работа над рядом систем для военных. Особенный интерес к этим телеавтоматическим устройствам проявило сначала командование танковыми подразделениями Московского военного округа, а потом и Управление танковыми войсками Министерства обороны СССР для использования при управлении мишенями на танковых полигонах во время обучения и маневров. Устройство для полигона представляло собой датчик псевдослучайных чисел, в котором сигнал возникал при совпадении сигналов от трех кольцевых регистров, число элементов в которых выражается простыми числами. Сигналы с помощью системы телеуправления передавались на исполнительные пункты, установленные в блиндажах полигона и воздействующие, в свою очередь, на подвижные мишени — цели. Последние, в непредсказуемые моменты времени, возникали перед глазами наводчика-танкиста (и иногда поражались). На полигоне нас опекали (т. е. помогали) кто-нибудь из полковников, часто довольно остроумных. Например: «Ивери Варламович, почему Вы ничего не делаете?» Ответ: «Я думаю». «А как я проверю, что Вы думаете?». И тому подобное.

В. А. Жожикашвили. Работами Ивери Варламовича заинтересовалось Министерство обороны, конкретно, маршал ракетных войск К. С. Москаленко, который несколько раз посетил ИАТ. Разработки Ивери Варламовича оказались столь полезны для армии, что К. С. Москаленко выделил для него квартиру в Москве на Ленинском проспекте. После защиты кандидатской диссертации директор Института академик В. А. Трапезников посоветовал Ивери Варламовичу остаться работать в Институте. Директор был очень увлеченным человеком, любил новые перспективные разработки и поддерживал наши исследования. В 1964 г. также по инициативе В. А. Трапезникова Ивери Варламовичу было предложено из его небольшой группы, но с уже определившейся тематикой, создать лабораторию. К этому времени (1964 г.) в издательстве «Наука» была опубликована первая в мировой практике монография под названием «Бесконтактные элементы и системы телемеханики», значительная часть которой описывает оригинальные разработки Ивери Варламовича, а именно, полупроводниковые реле и телеавтоматические системы.

Е. В. Бабичева. В 1964 г. была создана новая лаборатория под руководством уже непосредственно И. В. Прангишвили. Флагом лаборатории были «однородные структуры и микроэлектроника». Перед этим группа сотрудников вместе с Ивери Варламовичем находилась в составе лаборатории № 17, руководимой В. А. Жожикашвили. В связи с различием в научных направлениях двух коллективов дирекция посчитала обоснованным создание самостоятельной отдельной лаборатории под руководством И. В. Прангишвили, предложившим к тому времени принципиально новый подход к реализации логических, управляющих и вычислительных структур на основе однородных перестраиваемых структур. В тематику лаборатории входила разработка теории и инженерных методик создания телеавтоматических систем и однородных структур для управляющих вычислительных систем. Созданные полупроводниковые телеавтоматические системы нашли применение для управления производственными процессами обогатительных фабрик. Для управления промышленными объектами наша лаборатория, совместно с лабораториями С. М. Доманицкого и Б. С. Сотскова, создала серию полупроводниковых элементов «логика Т», которая была серийно освоена промышленностью и широко применялась. В лаборатории разрабатывалась элементная база и системы логического управления для корабельных энергетических установок, которые совместно с ленинградским НПО «Аврора» внедрялись в различных морских изделиях.

Б. П. Петрухин *(в то время один из ведущих сотрудников лаборатории чл.-корр. АН СССР Б. С. Сотскова, в настоящее время её заведующий).* В начале 1960-х гг. в качестве альтернативы релейно-контактным элементам для нижнего уровня систем управления промышленными объектами ведущие мировые фирмы начали разрабатывать так называемые бесконтактные логические и функциональные элементы. Одновременно в ИАТ начались разработки полупроводниковых (транзисторных) логических и функциональных элементов (от лаборатории Б. С. Сотскова ответственным исполнителем был я). Были проведены расчеты оптимальных параметров элементов в целях обеспечения серийнопригодности и максимальной надежности элементов. В Институте была выпущена опытная партия элементов, которая прошла опытную эксплуатацию и зарекомендовала себя с наилучшей стороны. В результате (уже совместно с ВНИИ «Электропривод») была создана Единая серия логических и функциональных элементов ЭТ («логика Т»), которая была принята Государственной комиссией в июле 1966 г. и запущена в серийное производство на Калининском заводе электроаппаратуры. Для продвижения серии в различные отрасли промышленности были организованы семинары и школы по обучению специалистов.

В результате системы управления на базе «логики Т» получили широкое распространение в различных отраслях промышленности — металлургии, станкостроении, машиностроении и др. Институт был награжден дипломом первой степени ВДНХ, а разработчики удостоены медалей.

А. А. Амбарцумян *(в 1970-е гг. аспирант МАГ'а, ныне — д-р техн. наук, профессор МГТУ им. Баумана, зав. лабораторией № 3 — той самой, в которой начинал Ивери Варламович).* В 1999 г. мне, в поисках хозяйственного договора по разработке АСУТП, довелось побывать на Московском мелькомбинате № 3. Главный инженер комбината показал мне функционально устаревшую, но действующую 30 лет на элеваторе систему автоматизированного управления. Каково же было мое удивление, когда я обнаружил, что действующая на комбинате система автоматизации выполнена на базе «логики Т». Вот что означает сделать для своего времени по-настоящему наукоемкую систему.

Н. А. Абрамова *(в 1960-е гг. аспирантка Ивери Варламовича, сегодня — д-р техн. наук, зав. сектором лаб. № 31).* После поступления в аспирантуру вначале меня направили к В. А. Жожикашвили, а тот, очень любезно поговорив несколько минут, уверенно сказал: «К Прангишвили!». Долгое время ИАТ видела только под тремя углами зрения: библиотека (там были удивительные и прекрасные сотрудники, помощь, теплоту и науку от них считаю неоценимой и сейчас), ученые советы (это было ослепительно) и культурно-просветительская деятельность (одна только встреча с И. Бродским чего стоит...). А вот Ивери Варламович с первой встречи поразил своей увлеченностью, рассказывал об однородных структурах, показывал фотографию того, что со временем приобрело название «chip», а позднее и «чип»... По-настоящему, по сути работы, моей первой статьей был реферат при поступлении в аспирантуру. Ивери Варламович дал мне тему — интегральные схемы с «картинкой» (название тогда у нас в стране было в новинку), и дал срок, кажется, 10 дней или две недели — найти на эту тему, что смогу. Я знала единственное место, где «находят», — библиотеку Ленина, где я привыкла писать школьные сочинения по литературе. Как мне удалось найти то «почти ничего» на английском языке, что было в библиотеке, где я провела все эти дни практически от открытия до закрытия, на тему, о которой еще вчера не слышала и затруднялась даже объяснить, что мне нужно, до сих пор не вполне понимаю. Но результатом, как мне кажется, шеф остался очень доволен.

Ведущий. Однородные среды стали флагом 31-й лаборатории почти на 10 лет; в 1968 г. Ивери Варламович защитил докторскую диссертацию, теоретическая часть которой основана на этих идеях, в 1969 г. получил звание профессора и именно из этих идей вышли параллельные системы. Как это было?

В. В. Игнатущенко *(ученик и первый соратник Ивери Варламовича по параллельным вычислительным системам, ныне д-р техн. наук, профессор МФТИ, зав. лабораторией № 4).* В начале 1960-х гг. ведущей тенденцией в создании электронных средств автоматики и вычислительной техники являлась *минимизация* структур таких средств, *максимальное упрощение* их схем — в целях обеспечения надежности непрерывно усложняющихся электронных систем. «Упрощение схемы» — было общепринятым правилом. Вот почему заведующий кафедрой вычислительной техники МАИ, чл.-корр. АН СССР Б. С. Сотсков поделился с одним из своих потенциальных дипломников мыслью: «Вот у нас в ИАТ'е есть один сумасшедший грузин, который твердит не о простоте, а об избыточности...». Эта идея, парадоксальная для того времени, потенциально заманчивая, но абсолютно ничем не подтвержденная, привела к тому, что «микроскопическую» группу И. В. Прангишвили «неожиданно» перевели в лабораторию директора Института, который дал Прангишвили карт-бланш на два года поисковых работ. Так началась эпоха систем с перестраиваемой структурой — ПС (в простонародье — «Пранги-систем»).

А в это время бурное развитие микроэлектроники в мире привело в начале 1960-х гг. к выдвижению и развитию концепции однородных микроэлектронных логических и вычислительных структур, состоящих из однотипных функциональных элементов с одинаковыми связями между ними (Минник, Канадей, Беркс...). Таким образом, ИАТ, где эта концепция разрабатывалась под руководством И. В. Прангишвили, стал одним из пионеров и лидеров в разработке этой новой проблематики.

Через два года в активе молодых ученых были теоретически обоснованные принципы построения однородных решающих полей, авторское свидетельство, макетные реализации однородных структур, публикации в научных журналах и доклад на престижном международном конгрессе. Более того, результаты теоретических исследований и практических (пусть только макетных, но сотворенных «собственными руками») разработок Института легли в основу первой в мире монографии «Однородные структуры» (авторы И. В. Прангишвили, Н. А. Абрамова, Е. В. Бабичева, В. В. Игнатущенко), целиком посвященной систематизированному изложению проблематики однородных перестраиваемых структур. Резонанс оказался столь заметен, что эта монография успешно экспонировалась на Всемирной выставке ЭКСПО—68.

Э. А. Трахтенгерц *(д-р техн. наук, профессор, соратник Ивери Варламовича по параллельным вычислительным системам).* Основная идея систем с перестраиваемой структурой, разрабатываемых под руководством И. В. Прангишвили, заключалась в том, чтобы перестроить, адаптировать цифровую вычислительную среду под параллельные вычисления для решаемой задачи. Это было чрезвычайно интересное направление, во многом предвосхитившее идеи, которые были реализованы на последующих этапах развития вычислительной техники. Но одна, возможно, центральная идея могла быть реализована на уже имеющейся в тот момент элементной базе. Эта идея заключалась в организации параллельных вычислений.

В то время как это не кажется странным сейчас, большинство специалистов (математиков, программистов, электронщиков) считали, что эффективно реализовать параллельные вычисления на цифровых вычислительных машинах либо невозможно вообще, либо крайне сложно и то только для узкого класса задач. Поэтому к идее параллельных вычислений отношение у них было, как теперь говорят, неоднозначное.

Ведущий. Попытки создания вычислительной техники для управления были в Институте и до Ивери Варламовича. Это как-то повлияло на работы по ПС?

Э. А. Трахтенгерц. В Институте под руководством профессора Б. Я. Когана была создана серия аналоговых вычислительных машин и разработаны методы расчета траекторий движения подвижных объектов и применения АВМ в качестве элементов систем управления этими объектами. Это направление[2] сыграло существенную роль в развитии Института, но привело, как иногда бывает, к недооценке начавшейся бурно развиваться на Западе и в СССР цифровой вычислительной техники. Такая позиция привела к запаздыванию в оснащении Института цифровыми вычислительными машинами (впоследствии Институт был оснащен ими очень хорошо).

Однако уже в 1960-х гг. стало понятно, что аналоговая вычислительная техника начинает вытесняться цифровой, поэтому была предпринята попытка использовать и сохранить опыт и достижения Института в области АВМ. Для этого совместно с югославским Институтом им. М. Пупина начала разрабатываться гибридная вычислительная система ГВС-100, состоявшая из аналоговой и цифровой вычислительных машин, объединенных аналого-цифровыми и цифроаналоговыми преобразователями. Опытный образец этой системы был создан, он был оснащен необходимым для экспериментальных работ программным обеспечением и применялся для моделирования систем управления сложными объектами, но дальнейшего развития это направление не получило. Независимо от достоинств и недостатков созданной гибридной системы, стала очевидна ее неконкурентоспособность с интенсивно развивающейся цифровой техникой.

Отметим, что работа над ГВС-100 позволила специалистам Института приобрести опыт разработки цифровых вычислительных систем и системного программного обеспечения, который вскоре был востребован.

Ведущий. В 1970 г. И. В. Прангишвили был назначен заместителем директора Института (к этому времени уже ИПУ), он получил прямой выход на Минприбор, как это повлияло на работу?

Э. А. Трахтенгерц. Я думаю, прежде всего, это проявилось в том, что Ивери Варламовичу удалось найти

---

[2] За разработку первых АВМ группе сотрудников Института в 1951 г. была присуждена Государственная премия СССР. На Международной выставке в Брюсселе установка ЭМУ-8 отмечена Большим призом (1958 г.). Установка ЭМУ-10 отмечена дипломом ВДНХ (1963 г.).

сторонников его идей и в Минприборе. Видимо, поэтому в начале семидесятых годов министр приборостроения, средств автоматизации и систем управления К. Н. Руднев поставил перед дирекцией Института задачу: создать на медленной отечественной элементной базе управляющую быстродействующую вычислительную цифровую машину. Эта задача была связана с тем, что уже в 1960-х гг. в СССР начало остро ощущаться отставание нашей вычислительной техники от мирового уровня. В то время оно было вызвано двумя факторами: плохим качеством, низким быстродействием элементной базы и слабостью программного обеспечения.

Для решения этой задачи дирекция собрала специалистов Института и предложила им принять участие в работе. Предложения, поступившие от них, были самые разные. Но, как оказалось впоследствии, самыми интересными и перспективными были предложения, основанные на идеях ПС, — создание вычислительной системы, реализующей параллельные вычисления и адаптирующей свои ресурсы к требованиям решаемой задачи. Параллелизм вычислений и адаптация ресурсов должны были компенсировать низкое быстродействие элементной базы и обеспечить необходимую скорость вычислений.

Над реализацией еще не очень ясных (в то время) идей создания параллельных цифровых комплексов начали усиленно работать под руководством профессора И. В. Прангишвили сотрудники его лаборатории: В. В. Игнатущенко, В. Д. Малюгин, Ю. С. Затуливетер, И. Л. Медведев и другие, а также присоединившиеся к ним сотрудники программистской группы, в которую входили С. Я. Виленкин, Л. Н. Горинович, Э. А. Трахтенгерц и др.

В. В. Игнатущенко. А дальше начинается почти фантастическая для тех лет история. В 1972 г., на фоне общего потепления советско-американских отношений, интерес к работам ИПУ, уже известным в мире, проявил один из западных лидеров супервычислений, корпорация CDC. Она предложила сотрудничество на паритетных началах: CDC брала на себя строительство заводов по производству новых дисковых носителей, современной элементной базы и трех вычислительных центров коллективного пользования. При этом ответственность за разработку многопроцессорной вычислительной системы с перестраиваемой структурой полностью ложилась на советских специалистов.

Этим заманчивым перспективам не суждено было сбыться. В США изменилась политическая ситуация, потепление сменилось похолоданием, и совместные работы были закрыты, фактически не начавшись. Однако энтузиазм ведущей западной фирмы по поводу наших разработок не прошел незамеченным в советских министерских креслах, и мы *самостоятельно* пошли творить...

В. Д. Малюгин. Идея однородных машин, благодаря рассказам И. В. Прангишвили, овладела сотрудниками лаборатории. Руками умельцев (И. П. Егоров, А. А. Ельтищев) был построен красивый макет однородной среды. Макет понравился, о коллективе услышали, был создан менее яркий, но близкий к жизни и технике новый макет уже однородной ЭВМ (главный разработчик М. А. Ускач). Дирекции новый макет понравился еще больше. Началась разработка настоящей, оригинальной ЭВМ ПС-300. Основным соисполнителем стало НПО «Элва» (г. Тбилиси). Машина появилась тогда же, когда появились первые микропроцессоры (1976 г.).

В. В. Игнатущенко. После первой «пробы пера» (микро-ЭВМ ПС-300) дальнейшее развитие нетрадиционных принципов динамической перестраиваемости вычислительных средств привело к разработке высокопроизводительных многопроцессорных вычислительных систем (МВС) с перестраиваемой структурой — как с одним, так и со многими потоками команд и данных.

Оригинальный принцип *перестраиваемости* МВС серии ПС заключается в способности МВС к динамическому перераспределению параллельных ресурсов каждого типа (устройств управления, процессорных элементов, памяти, устройств ввода-вывода) между и задачами, и (или) их параллельными фрагментами, и (или) параллельными командами фрагментов — перераспределению, осуществляемому операционной системой или *аппаратурными средствами* по указаниям в программе или автоматически (путем анализа процесса выполнения программ) в соответствии с *текущими требованиями* задач, их фрагментов и команд на ресурсы.

Другие принципы построения МВС серии ПС, широко реализуемые и в современных многопроцессорных вычислительных комплексах, включают в себя: параллелизм организации вычислительных процессов на нескольких уровнях — задач, параллельных фрагментов каждой задачи, параллельных векторных и скалярных задач каждого фрагмента; иерархию управления вычислительными процессами, децентрализацию управления вычислениями и обменами информацией; модульность и регулярность структуры системы; использование специальных программных и *аппаратурных* средств распараллеливания и конвейеризации как вычислений, так и управления ими.

Э. А. Трахтенгерц. Теоретические разработки и создание программных моделей разрабатываемой системы продвигались настолько успешно, что когда фирма CDC обратилась в Комитет по науке и технике СССР с предложением о сотрудничестве, разработчики системы ПС под руководством И.В. Прангишвили смогли сформулировать предложения, заинтересовавшие американцев. Однако, как известно, изменившаяся политическая ситуация все прервала. Тем не менее, визит вице-президента фирмы CDC к министру и отчет группы специалистов этой фирмы «не пропали даром». Министр поверил в идеи Института и приказал северодонецкому НПО «Импульс», входившему в систему Министерства, рассмотреть вопрос о создании систем ПС.

Интересно, что некоторые американские влиятельные газеты, узнав о предложениях фирмы CDC, обвинили ее в том, что она «продалась большевикам», а некоторые наши инстанции обвиняли нас в разглашении наших секретов американцам.

Ведущий. Создание реальных ПС в Северодонецке — как это удалось?

Э. А. Трахтенгерц. В конце 1960-х — начале 1970-х гг. в Советском Союзе предпринимались значительные усилия по преодолению кризиса в вычислительной технике. Были созданы вполне современные (по архитектуре систем того времени) вычислительные машины «Урал», «Минск», «Днепр», М-220, БЭСМ-4 и другие, которые, несмотря на отсталость элементной базы и слабость программного обеспечения, имели неплохую перспективу дальнейшего развития. В это же время в Америке фирмой IBM была выпущена ЭВМ IBM-360, которая быстро завоевала мировой рынок компьютеров. Советским правительством было принято решение закрыть производство вычислительных машин оригинальных отечественных конструкций и перейти на копирование американского семейства вычислительных систем IBM-360, которые начали производиться под маркой ЕС. Таким образом в стране были погублены сильные коллективы разработчиков оригинальной вычислительной техники. Они были переориентированы на копирование западных образцов. В русле этого направления северодонецкое НПО «Импульс» быстро и удачно скопировало модель малой машины фирмы «Hewlett Packard» и начала ее производство под маркой М-6000. Голодный советский рынок жадно поглощал эту систему.

В этих условиях предложение по созданию отечественной оригинальной вычислительной системы, да еще основанной на параллельных вычислениях, и в ИПУ, и в НПО «Импульс» встречалось по-разному. Одни отнеслись к этому скептически, другие — с большим интересом. Такие же противоречивые чувства это предложение вызывало и у специалистов других организаций, не участвующих в проекте.

Тем не менее, дирекция и специалисты НПО «Импульс» взялись за разработку этих систем. Руководил разработкой зам. директора Института проблем управления И. В. Прангишвили. С самого начала работы она велась по двум следующим направлениям.

- Создание векторного процессора с единым потоком команд, использующего в качестве управляющей системы серийные малые вычислительные машины М-6000 (в дальнейшем СМ-1 и СМ-2). Эти малые компьютеры производились НПО «Импульс». Векторные процессоры предназначались для быстрой параллельной обработки больших массивов однородной информации, были очень эффективны для решения достаточно широкого класса задач, связанных, главным образом, с обработкой сигналов (поэтому их называют также сигнальными процессорами). Это задачи геологоразведки, распознавания, радиолокации, медицинской диагностики и др.
- Создание многопроцессорного комплекса с общей оперативной памятью, множественным потоком команд и специальными процессорами: векторными и ассоциативными. Эти комплексы предназначались для решения широкого класса задач, где требовались высокая производительность и перестраиваемость вычислительной системы.

Системы первого направления были технологически проще (состояли из большого числа одинаковых процессорных элементов от 64 до 264) и могли использовать в качестве управляющей серийную малую вычислительную машину (М-6000, СМ-1). Два этих качества плюс высокая производительность предопределили выбор этих систем как первоочередных. Основные усилия с самого начала были сосредоточены на их разработке и производстве.

Эти усилия увенчались большим успехом, в результате упорного труда коллективов ИПУ и НПО «Импульс» были преодолены большие трудности в разработке оригинальной отечественной высокопроизводительной машины на базе медленных отечественных элементов. Она получила наименование ПС-2000. Таким образом, была выполнена задача, поставленная Министерством перед Институтом проблем управления. Первая серийная машина была выпущена в 1982 г., и до 1990 г. было произведено 242 ЭВМ ПС-2000. Они были установлены и работали (а многие работают до сих пор) на различных исследовательских, промышленных и оборонных объектах, имеющих большой пакет хорошо распараллеливаемых задач.

На базе ПС-2000 был создан промышленный экспедиционный вычислительный комплекс обработки данных сейсморазведки, комплексы обработки гидроакустической информации, обработки данных телеметрии и т. п. В развитии линии ПС-2000 проводились разработки высокопараллельной архитектуры, предназначенной для реализации на сверхбольших интегральных схемах (СБИС). Архитектура строилась в расчете на масштабируемое многопроцессорное исполнение кристаллов СБИС и была сбалансирована на всех уровнях.

При разработке вычислительных систем второго направления, получивших название ПС-3000, которая началась значительно позже разработки системы ПС-2000, реализовался оригинальный принцип перестраиваемости вычислительной техники. Перестраиваемость заключалась в способности многопроцессорного комплекса к динамическому перераспределению вычислительных ресурсов каждого типа между задачами и (или) их фрагментами. Такое перераспределение осуществлялось как программными, так и аппаратными средствами в соответствии с текущими, заранее непредсказуемыми требованиями задач и их фрагментов. Для системы ПС-3000 была разработана оригинальная операционная система, осуществляющая управление несколькими универсальными процессорами и работающими под их управлением спецпроцессорами, был создан транслятор с языка Фортран, расширенный параллельными операциями, и совместно с венгерской организацией транслятор с языка программирования ADA. Системы ПС-3000 как в проектировании, так и в производстве были значительно сложнее систем ПС-2000, первые машины были выпущены только во второй половине 1980-х гг., их доработка и дальнейший выпуск были прекращены в связи с изменениями, происходившими в стране.

М. С. Шкабардня *(министр СССР в 1980—1989 гг., д-р техн. наук, профессор, Герой Социалистического Труда, лауреат Государственной премии СССР).* В 1970—1980-е гг. И. В. Прангишвили неоднократно выступал на НТС и коллегиях Минприбора с докладами о разработанных в Институте под его руководством теоретических основах и принципах построения нового класса высокопроизводительных многопроцессорных проблемно ориентированных управляющих вычислительных систем с перестраиваемой структурой. На этой базе предприятиями отрасли были разработаны и освоены в серийном производстве УВК серии ПС-2000, ЭГВК ПС-2000, -2100, -3000. Эти комплексы из 64 и 128 одинаковых процессоров были первыми в стране высокопроизводительными машинами, обеспечивающими производительность до 200 млн. операций в секунду и нашли широкое применение в геофизике, в цифровой обработке сигналов в гидроакустике, обработке сигналов, идущих от космических станций, в управлении быстрыми процессами в исследовательских задачах и в спецтехнике. Комплексы ПС-2000 и ПС-2100 были лучшими в стране по показателю «производительность/стоимость».

Конкретно с 1981 по 1988 г. северодонецким НПО «Импульс» было изготовлено более 180 ЭГВК ПС-2000 и свыше 240 мультипроцессоров ПС-2000. Эти комплексы использовались в различных отраслях народного хозяйства. Так, ЭГВК ПС-2000 широко применялся в геофизике для обработки данных сейсморазведки нефти и газа. В целом в Мингео СССР использовалось около 90 ЭГВК ПС-2000. С 1986 г. ЭГВК ПС-2000 нашли применение в центре управления космическими полетами для обработки телеметрической информации, поступающей со спутников и космической станции «Мир». Они также эффективно применялись при решении задач ядерной физики в НИИЭФ (г. Саров) и НИТИ (г. Сосновый Бор Ленинградской обл.).

Н. А. Абрамова. Еще одним немаловажным итогом периода ОС/ПС является создание вокруг Ивери Варламовича новых научных коллективов: в ИПУ — лабораторий С. Я. Виленкина, Э. А. Трахтенгерца; вне Института — школы. Организация школ по ОС/ПС знаменовала собой формирование общей научной и человеческой среды, формирование длительных связей, которые для многих перешли в дружбу с «соратниками».

Первая волна, которой предшествовала 1-я Всесоюзная конференция по однородным средам в Академгородке Новосибирска (1968 г.), — это тесные связи и сотрудничество со школой д-ра техн. наук, профессора Э. В. Евреинова. Это — Новосибирское море, жаркие споры об «идеальных» однородных средах, все и всех понимающий В. П. Чистов…

Вторая волна — школы в п. Мозжинке, неразделимо переплетшиеся со школами М. А. Гаврилова не только по месту и времени, но и по людям; Л. Я. Розенблюм и А. Я. Макаревский, в разгар вечернего веселья предлагающие свою новую теорию однородных сред; от новосибирских связей осталась только О. Л. Бандман; прекрасное дружеское общение, однородные среды — по-прежнему «игрушечные», «красивые» (в терминологии д-ра техн. наук, профессора В. П. Чистова).

Третья волна — еще сохраняется человеческая среда второй волны, но мы уже работаем над проектами ПС; типичная школа этой волны — 1979 г., Москва — Всесоюзная школа-семинар по однородным вычислительным структурам и малым ЭВМ. Речь идет уже о проблемах создания ЭВМ на однородных перестраиваемых структурах, построения систем из микропроцессоров, матобеспечения микро-ЭВМ и другим вопросам, хотя доклады по «красивым» ОС еще остаются. Наконец, последняя волна, которая заканчивается в 1989 г., — это школы с тематикой, расширяющейся в проблематику ЭВМ.

А. А. Амбарцумян *(соратник Ивери Варламовича по атомному проекту — АСУТП АЭС).* К середине 1980-х в Институте все острее ощущается отставание в оснащенности компьютерной техникой. Во всем мире на оснащение научной и инженерно-технической работы мощным потоком поступают персоналки, а мы должны довольствоваться ДВК и СМ различных модификаций; ВЦ располагает машиной ЕС-1045 (оттого, что у других и этого нет, нам не легче). Для работ по САПР арендуем ночное время у ВЦ крупных предприятий оборонки. А тут еще перестройка, расширяются контакты с наукой Запада. В некоторых лабораториях появляются ввезенные с Запада для совместных работ персоналки, по мощности соизмеримые с техникой ВЦ. Завлабы идут к Ивери Варламовичу и просят, просят, а некоторые и требуют. Ивери Варламович уже 15 лет зам. директора, но фактически более 10-ти лет «тянет» Институт в одиночку. Он отлично познал механизмы «раздачи слонов» в социалистической экономике и поэтому понимает, что только участие в крупном, значимом на уровне Кремля проекте поможет переоснастить Институт, да и всколыхнуть многие лаборатории в «общем деле».

В 1986 г. случилась авария на Чернобыльской АЭС. Среди группы специалистов обоснованно сложилось мнение, что одна из причин катастрофы в примитивности (архаизме) применяемых системотехнических решений в АСУТП, особенно в части систем управления безопасностью. И тогда академик А. П. Александров вспомнил, что есть институт академика В. А. Трапезникова — ИПУ РАН, что есть опыт работы с этим институтом при создании АСУТП для подобных объектов в оборонной технике. Академики обменялись мнениями и достигли согласия. В Минатомэнерго считали, что дело не в системотехнике АСУТП (она полностью определялась одним из их институтов), а в плохих технических средствах, которые поставляет Минприбор (да и верхний уровень делают в ЦНИИКА), и поэтому пришли к мнению, что пусть и систему полностью поставляет Минприбор. Министр СССР М. С. Шкабардня переговорил с Ивери Варламовичем, предложил возглавить разработку перспективной АСУТП АЭС. Ивери Варламович был хорошо осведомлен о технике, поставляемой для АСУТП АЭС (значительная ее часть производилась в НПО «Элва» и НПО «Импульс», с которыми было тесное сотрудничество по ПС). Но он также и понимал — как трудно вести работу, в которой задействовано более сотни организаций различных ведомств, имеющих опыт работы в этой области, у них сложились отношения, они считают себя лучшими, а плохо потому, что у смежника что-то не так. Да и все ниши в этой проблеме уже заняты, без соответствующих полномочий лучше не браться. Соответствующие полномочия были обещаны, и работа закипела.

После первого анализа состояния дел (ЦНИИКА уже несколько лет вел эту работу) мы с удивлением обнаружили, что в прорабатываемом проекте перспективной системы существенно меняется только верхний уровень (информационная часть системы и средства поддержки оператора). В то же время нижний уровень, на котором сосредоточены все защиты, измерительные каналы, автоматические регуляторы и осуществляется функциональное групповое управление оборудованием и который составляет более 80% аппаратуры — остается неизменным. Стало ясно, что необходимо менять концепцию управления и всей системы. Лаборатория № 3 взяла на себя координацию работ ИПУ по концепции, а группа (вскоре преобразованная в новую лаб. № 40) Ф. Ф. Пащенко *(ныне д-р техн. наук, профессор, ученый секретарь Института, зав. лабораторией)* взяла на себя работу по подготовке новой Программы ГКНТ СССР по созданию перспективной АСУТП для АЭС.

Работа над Программой отнимала много времени: бесконечные совещания, согласования, отчеты, справки, посещения АЭС, визиты в инстанции и смежные Институты (в работы по программе были вовлечены более 200 организаций из 9-ти союзных министерств, а позже и около 100 организаций стран — членов СЭВ). Завершилась эта работа (1987 г.) Постановлением Правительства, в котором утверждалась Программа, Институту поручили возглавить разработки и внедрение АСУТП атомных электростанций, а Ивери Варламович был назначен Генеральным конструктором СССР по АСУТП атомных электростанций. В Постановлении была специальная часть, благодаря которой Ивери Варламовичу в 1988 — 1989 гг. удалось оснастить практически весь Институт современными зарубежными персоналками и другой оргтехникой.

А. Н. Шубин *(профессор, заместитель директора)*. В это же время (1987 г.) Ивери Варламович был избран и утверждён в должности директора Института проблем управления АН СССР и Минприбора СССР. Впервые выборы проходили на альтернативной основе, и за Ивери Варламовича проголосовало подавляющее большинство сотрудников Института.

А. А. Амбарцумян. Над концепцией управления в АСУТП работали сотрудники лабораторий: А. А. Амбарцумяна (отв. исполнители А. И. Потехин, Б. А. Лаговиер), Б. Г. Волика (отв. исполнитель Н. В. Лубков), В. И. Уткина, М. А. Розенблата (отв. исполнитель Н. Э. Менгазетдинов), И. В. Прангишвили (отв. исполнители: В. В. Игнатущенко, М. А. Зуенков), Ф. Ф. Пащенко. Основные положения концепции предусматривали новации на всех уровнях управления: на верхнем предлагалась система информационной поддержки на основе экспертной системы и базы знаний, на нижнем — новые типы регуляторов и распределенное логическое управление. Более того, в части развития индивидуального уровня управления приводами (исполнительными автоматами) наши предложения опередили технические решения западных фирм. В плане системотехнических решений была сформулирована концепция создания распределенной отказобезопасной системы управления, отвечающей требованиям МАГАТЭ по основным параметрам безопасности и уровню автоматизации. Аналитические расчеты подтвердили эффективность технических решений по основным системным показателям: динамическим характеристикам и надежности.

Одно из важнейших положений концепции заключалось в обосновании необходимости создания нового типа технических средств автоматизации — средств программируемой автоматики с параллельной структурой (СПА-ПС). Основные технические идеи, положенные в основу СПА-ПС, продолжают линию «больших» ПС и заключаются в распределенности, специализации и контролируемости процессов обработки и коммуникации[3].

Концепция была достаточно революционна для своего времени, особенно в части применения программируемых средств на уровне управления приводами, и здесь, я думаю, нам удалось её продвинуть только благодаря Ивери Варламовичу. Его талант просто (как он любит повторять — «на пальчиках») объяснять сложнейшие вещи как управления, так и схемотехнических приемов обеспечения требуемой надежности, буквально продавливал как НТС, так и административные инстанции. Концепция была принята на НТС Минатомэнерго, Минприбора и положена в основу проекта АСУТП Башкирской АЭС. Однако в силу политических причин (катастрофа в государстве и развал экономики) реализовать концепцию в полной мере не удалось, хотя все основные положения проверены в макетах и использованы в ряде проектов.

Работы по АСУТП АЭС продолжаются и сегодня. В настоящее время Институт является головным по созданию систем управления верхнего блочного уровня АЭС. Институтом разработаны принципы управления, созданы оригинальное программное обеспечение, технический и рабочий проекты системы управления верхнего уровня для строящейся атомной станции в Бушере (Иран), а затем для АЭС в Кудам-Кулане (Индия).

А. Н. Шубин. Плодотворную научную деятельность Ивери Варламович всегда совмещает с большой и ответственной научно-организационной и общественной работой. В этой части его деятельности (был ли он депутатом городского или районного Совета, членом других общественных организаций города или Института) следует дать самую высокую оценку ответственности за порученное дело, его отношению к заботам и нуждам простых людей, которые постоянно обращались к нему за помощью. В настоящее время Ивери Варламович продолжает совмещать свою напряженную научную работу с большой научно-организационной и общественной работой. Он является председателем Учёного совета Института проблем управления. С 1992 г. он — член Бюро Отделения проблем машиностроения, механики и процессов управления, а с 2002 г. — член Бюро Отделения энергетики, машиностроения, механики и процессов управления Российской академии наук. С 1995 г. возглавляет Научный совет Отделения РАН по теории управляемых процессов и автоматизации. Является заместителем председателя Национального комитета по автоматическому управлению, главным редактором журналов «Проблемы управления» и «Датчики и системы», а также председателем редакционной коллегии журнала «Автоматизация в промышленности» и членом редколлегий ряда центральных научных журналов. Под научным руководством И. В. Прангишвили защищено более 30 докторских и кандидатских диссертаций, он также ведёт активную преподавательскую работу.

За большие достижения в научной и производственной деятельности И. В. Прангишвили награждён двумя орденами Трудового Красного Знамени, орденом Дружбы, орденом Чести и многими медалями.

И. В. Прангишвили — действительный член нескольких зарубежных академий, вице-президент Международной инженерной академии и вице-президент

[3] Позже (1994—1997 гг.) СПА-ПС были освоены в серийном производстве в АО «НПО "Автоматика"» (г. Омск). Средства СПА-ПС сертифицированы органами Госстандарта РФ в качестве средств измерений и на соответствие требованиям ГОСТ по безопасности. С 1997 г. СПА-ПС используются в проектах распределенных и сосредоточенных автоматизированных систем, важных для безопасности, для объектов атомной и тепловой энергетики, газонефтедобывающей и перерабатывающей промышленности, металлургии и других производств.

Академии наук Грузии. Кроме того, Ивери Варламович — президент общества «Грузины в России».

С. В. Полянский *(многие годы один из профсоюзных лидеров Института)*. Длительное время Ивери Варламович был депутатом Моссовета. Деятельность на этом поприще еще раз подтвердила его жизненное правило — внимательно относиться к людям, оказывать им поддержку и помощь. Помимо непосредственной работы в Моссовете, Ивери Варламович участвовал в работе Черемушкинского райсовета (сессии, прием населения, контроль за выполнением наказов избирателей и т. д.). Ивери Варламович был избран депутатом от избирательного округа, расположенного вблизи станции метро «Беляево», где проживают многие сотрудники Института. Их бытовые, жизненные заботы, связанные с работой торговли, медицинским обслуживанием, работой школ, кинотеатра, транспорта, охраной порядка, условиями проживания, благоустройством территории и т. п., стали заботой Ивери Варламовича. В этот период было сделано много хорошего для жителей микрорайона, включая и наших сотрудников. По существу, Институт стал шефом микрорайона, установил связи с руководителями основных ключевых звеньев микрорайона, согласовал и координировал их деятельность в соответствии с планами развития и выполнения наказов избирателей. Будучи депутатом Моссовета, Ивери Варламович тратил много времени, энергии, проявлял настойчивость, выступая с обоснованными ходатайствами перед различными инстанциями по вопросам, с которыми к нему обращались сотрудники Института. Результатами его работы были улучшение жилищных условий, строительство автостоянок и гаражей для значительного числа сотрудников Института.

Е. В. Бабичева. В научно-организационном плане Ивери Варламович всегда ответственен, например, в течение 13 лет руководил секцией «Вычислительные системы» в Научном совете по комплексной проблеме «Кибернетика» АН СССР, возглавляемом академиком А. И. Бергом. В Совете постоянно обсуждались новые научные направления, предлагаемые основными разработчиками вычислительной техники, в частности, вычислительные системы, построенные на основе однородных микроэлектронных структур. Это направление, получившее широкое развитие и практическую реализацию, активно поддерживал А. И. Берг. Помимо научной деятельности, Институт совместно с Советом по проблеме «Кибернетика» проводил по тематике секции школы-семинары, привлекавшие большое число участников.

Ведущий. Вот уже около 15 лет мы в новой стране, с другим общественным строем. Тяжело всем социальным институтам страны, в том числе и науке. Как Ивери Варламовичу удается в этой ситуации сохранить Институт, и, самое важное, атмосферу творчества и чувство взаимного доверия и уважения в коллективе?

А. Н. Шубин. Это очень серьезная тема. Коротко отмечу два главных принципа, которым, как мне кажется, следует Ивери Варламович и требует их соблюдения от нас: открытость и прозрачность. Он всегда *открыт* ко всем предложениям, пожеланиям и т. д., с которыми к нему приходят сотрудники, но внимательно и вдумчиво их анализирует в беседах с коллегами, моделируя все возможные последствия. *Прозрачность* в решениях, влияющих на судьбу как отдельного сотрудника, так и Института в целом: еженедельно проводятся директорские совещания, с откровенным обсуждением всех оперативных экономических и хозяйственных вопросов. Ежемесячно он проводит заседания Ученого совета, на которых обязательно хотя бы один научный доклад по актуальной, перспективной тематике и его очередное сообщение «О текущем моменте». В них он докладывает, как дела с финансированием, откуда и какие поступления и как предполагается ими распорядиться. В свойственном ему импровизационном стиле информирует коллектив обо всех тенденциях «вверху», о том, что требуют от нас (как академической организации) и как мы предполагаем на это реагировать, с какой целью и какие структурные или кадровые изменения намечаются в Институте.

А. В. Толстых *(канд. техн. наук, в 1970 — 1980 гг. сотрудник Минприбора, в настоящее время — Генеральный директор ассоциации «Рост», заведующий научно-внедренческим отделом Института)*. Я с Институтом и его директором хорошо знаком еще со времен Минприбора, да и кандидатскую делал в ИПУ у В. Н. Буркова, но наиболее тесное сотрудничество началось уже в новое время, с созданием ассоциации «Рост». Институт мы пригласили в ассоциацию для научной поддержки и обоснования наших усилий по продвижению продукции отечественного приборостроения в системы различного назначения. Ивери Варламович, наряду с поддержкой, постоянно требовал от меня вовлечения в работы ассоциации лабораторий Института. Результатом совместных работ явилось создание современной системы производственного экологического мониторинга (ПЭМ) объектов уничтожения химического оружия (УХО) в России. В системе ПЭМ информационно-аналитический центр разрабатывался совместно с лабораториями В.Г. Лебедева и Ю. С. Леговича, а оптимизация в проекте системы номенклатуры приборно-технических средств осуществлялась на основе экономико-математических методов, разработанных в лаборатории В. Н. Буркова. В 2002 г. система ПЭМ была внедрена на объекте УХО в п. Горный[4]. В 2002 г. Ивери Варламович выдвинул идею создания в Институте научно-внедренческих отделов (НВО). Задачи НВО — вовлечение (без привлечения академических средств) в реальные проекты сотрудников Института с их научными результатами. Возможно, успех проекта системы ПЭМ послужил основанием Ивери Варламовичу поручить мне создание одного из таких отделов — НВО «Информационно-управляющие системы мониторинга». В настоящее время отдел совместно с лабораториями ведет тиражирование систем ПЭМ на трех новых объектах УХО. Сегодня, уже в составе Института, давление Ивери Варламовича в целях вовлечения лабораторий в новые проекты отдела многократно возросло. Это и доклады на Ученом совете

[4] Ввод в действие в декабре 2002 г. объекта УХО в п. Горный Саратовской области позволил России уже в апреле 2003 г. отчитаться перед мировым сообществом о выполнении первого этапа Международной конвенции — уничтожении 1% накопленных запасов химического оружия в нашей стране.

Института, и сообщения на директорских совещаниях и т. п., но, в случае возникновения у нас трудностей с продвижением каких-либо проектов, Ивери Варламович оперативно приглашает к себе специалистов и мы коллективно «куем аргументы». Вот именно таким образом в январе 2005 г. мы подготовили заявку на разработку Системы управления безопасностью объектов УХО, выполнение которой предполагает значительное расширение состава участников от ИПУ.

Ведущий. Научная работа Ивери Варламовича сегодня, в чем основная суть, как она видится коллегам?

Н. А. Абрамова. После долгого расхождения в части научных интересов моя новая «встреча» с Ивери Варламовичем произошла на почве темы «Разработка методов исследования крупномасштабных слабо формализуемых задач в социально-экономических, организационных и экологических системах». Оказалось, что систематические методы творческого переноса знаний, по которым у меня к тому времени были неплохие обобщающие результаты, очень тесно перекликаются с теми методами, которыми естественно пользуется Ивери Варламович. Продукт «новой встречи» — вторая коллективная монография — «Поиск подходов к решению проблем». Это была первая книга, в которой Ивери Варламович представил свое новое научное направление «Системные закономерности». По моему мнению, он пришел к ним в поисках объяснений поведения сложных, слабо структурированных, слабо определенных систем и ситуаций, в поисках возможности оценить принимаемые управленческие решения самого высокого уровня. Вновь (как и тогда, когда он закладывал направление ОС) он сталкивается с непониманием, недоумениями, прямым неприятием. Правда, теперь все это протекает в мягких и скрытых формах — многие любят Ивери Варламовича, да и положение его на иерархической лестнице существенно изменилось. Я интерпретирую найденные Ивери Варламовичем закономерности как виды, или общие модели, поведения и организации естественных систем и обусловленные этими моделями свойства, которые достаточно типичны (достаточно часто встречаются), чтобы использовать их в качестве возможных объяснительных моделей при решении задач анализа естественных систем или в качестве прототипов при создании новых систем и решении задач управления имеющимися системами.

Ведущий. Разрешите зачитать текст одного из выступлений Ивери Варламовича на Ученом совете, в котором, как мне кажется, он объясняет свое видение актуальности системных закономерностей именно сегодня:

«В период перестройки и реформирования страны интерес к науке со стороны государства резко упал. Наука, в том числе и фундаментальная, попала в хаос рыночной экономики. Фундаментальная наука, как наука о будущем, требует серьезной государственной поддержки. Для реализации результатов фундаментальной науки необходима инновационная инфраструктура: институты прикладных исследований, конкретизирующие результаты фундаментальных исследований в реальных проектах, производство, реализующее эти проекты в промышленности, но главное, спрос со стороны промышленности на новую технику и технологию.

Современная структура экономики, ориентированная на производство сырья, не предвещает хороших перспектив. Конкуренцию на мировых рынках в эпоху глобализации может обеспечить не сырьевая, а другая структура, ориентированная на наукоемкую высокотехнологичную продукцию и экономику знаний. Тогда спрос на науку, в том числе и на фундаментальную, возрастет.

Сегодня, по моему мнению, успехи деятельности организации, фирмы, корпорации, регионов и страны в целом на 70 — 80% определяются эффективностью систем управления. Богатство страны в большей степени зависит не от природных ресурсов, а от эффективности управления. Примеры: Япония, Финляндия, Гонконг, Сингапур, Германия и другие страны, у которых почти нет сырьевых ресурсов, а достойный уровень жизни обеспечивается эффективностью управления.

В России и странах СНГ низкая эффективность управления, поэтому, несмотря на наличие богатых сырьевых ресурсов, уровень и качество жизни значительно уступает развитым странам.

Одной из главных причин неэффективного управления в организационных, социальных, экономических и других структурах состоит в игнорировании достижений науки в области управления. Слабо применяются системный или целостный подход в управлении, мягкое резонансное управление, принципы «золотого сечения» в управлении, рефлексивное управление и другие достижения науки управления. Применение научных методов повышения эффективности управления позволит создать более устойчивую, стабильную и гармоничную систему управления различными структурами общества.

Сегодня наблюдается низкая управляемость на всех уровнях и отсутствие самоорганизации и саморазвития социально-экономических и организационных систем. Проводимые административные реформы, пенсионные реформы, реформы ЖКХ, медицины, образования и другие осуществляются без соответствующего научного обоснования и поэтому неэффективны и вызывают протестные настроения.

Наука управления должна подсказать, какие и как проводить реформы, чтобы народ не пострадал и принял бы их благожелательно.

Системный подход и системные закономерности, которыми я занимаюсь последние 10 лет, показывают, что когда отсутствует системный подход и нарушаются системные закономерности, тогда эффективность управления резко падает. Эти исследования позволяют выявить научные методы и механизмы повышения эффективности управления в сложных системах различной природы».

Ведущий. В заключение с удовольствием сообщаю, что участники заочного круглого стола и редакция журнала от всей души поздравляют Ивери Варламовича с двойным юбилеем и желают ему здоровья, счастья и многих лет активной творческой жизни!

Записал и провел круглый стол
А. А. Амбарцумян

☎ *(095) 334-87-89*

E-mail: ambar@ipu.ru ❑

ПЕРВАЯ РОССИЙСКАЯ КОНФЕРЕНЦИЯ ПО КОГНИТИВНОЙ НАУКЕ

С 9 по 13 октября 2004 г. в Казани под эгидой Казанского государственного университета им. В. И. Ульянова-Ленина состоялась Первая российская конференция по когнитивной науке.

Ее проведение поддержали Институт языкознания РАН, Факультет психологии МГУ им. М. В. Ломоносова, Филологический факультет Санкт-Петербургского университета, Московский городской психолого-педагогический университет, Психологический институт Российской академии образования, Федеральный центр патологии речи и нейрореабилитации, Cognitive Science Society, Inc., International Cognitive Linguistics Association, Slavic Cognitive Linguistics Association, Российская ассоциация искусственного интеллекта, Российское психологическое общество, Российская ассоциация лингвистов-когнитологов.

Председатель Программного комитета — *Б. М. Величковский* (Дрезденский университет и Федеральный центр патологии речи и нейрореабилитации, Москва), заместители председателя — *А. А. Кибрик* (Институт языкознания РАН и МГУ, Москва) и *Т. В. Черниговская* (СПбГУ, Санкт-Петербург).

Объявленная организаторами цель конференции — это создание совместного форума представителей разных наук, исследующих познание и его эволюцию, интеллект, мышление, восприятие, сознание, представление и приобретение знаний, язык как средство познания и коммуникации, мозговые механизмы познания, эмоций и сложных форм поведения. В мировом масштабе такие исследования объединяются под флагом *когнитивной науки* — одной из ведущих в прикладных и фундаментальных исследованиях XXI в. Организаторы конференции заявили о своем стремлении к формированию в России сообщества ученых-когнитологов, способных общаться между собой на едином языке.

На конференции было представлено около 150 докладов, включая пленарные и стендовые. Имея формальный статус российской, конференция фактически была международной — на ней был представлен ряд докладов и приглашенных лекций из США, Германии, Франции, Финляндии, Китая, Японии, Грузии; английский язык был одним из рабочих языков конференции. Важная черта конференции — высокий уровень научных требований к докладам (соответствующий уровню публикаций в академических изданиях). Активные дискуссии на заседаниях и кулуарные обсуждения свидетельствовали о возможности продуктивного взаимодействия представителей разных наук. В завершение конференции было принято решение о подготовке к организации Российской ассоциации когнитивной науки и ее вхождению в европейскую ассоциацию.

Попытаемся вкратце охарактеризовать прошедшую конференцию с позиции специалистов в области науки управления, и так или иначе заинтересованных в связях с когнитивной наукой. Предполагается, что здесь важно оценить, насколько существенен интеграционный процесс, анонсированный организаторами конференции и развитию которого она послужила, для развития науки управления, насколько целесообразно принять в нем участие. Понятно, что такого рода оценки не могут не быть субъективными и поверхностными, создавая лишь предпосылки более глубокого изучения материалов конференции[1] заинтересованными специалистами.

Основные научные направления, представленные на конференции, — это когнитивная психология, когнитивная лингвистика, нейронауки, искусственный интеллект.

Среди представленных докладов наряду с работами, лежащими в рамках этих направлений, существенное место занимали доклады междисциплинарного характера: по обсуждаемым научным проблемам, применяемым методам и используемым идеям, по ориентации изложения на слушателей «чужих» направлений.

Б. М. Величковский в своем обзорном докладе «Когнитивная наука вчера, сегодня, завтра» рассмотрел тенденции развития когнитивных исследований в течение последних десятилетий и представил свой сценарий их ожидаемого развития. В качестве существенной тенденции выделена нейрокогнитивная парадигма исследований, которая, в противовес предшествующей компьютерной метафоре, ориентирована на нейрофизиологические данные. Основной акцент в представленном сценарии развития когнитивной науки сделан на перспективах ее применения для решения прикладных проблем. Одной из таких областей применения, по оценке докладчика, может стать создание разнообразных когнитивных технических систем, а также медицина и образование. При этом любая техническая система может считаться когнитивной в той мере, в которой она учитывает знания, намерения и функциональные состояния пользователя. Представлены первые примеры таких систем, в частности, системы, основанные на учете фокусировки внимания человека.

С исследованиями Б. М. Величковского, относящимися к фокусировке внимания и искусственным механизмам его применения, перекликаются идеи *Х. Риттера* (Германия), которые представлены в двух докладах: «Искусственное внимание как основа для когнитивных роботов» и «Гиперболические самоорганизующиеся карты: согласование экранных представлений данных с человеческим вниманием» (на англ. языке). В частности, во втором докладе продемонстрирован оригинальный способ управляемого вниманием просмотра визуальных карт, которые создаются в результате извлечения данных из текста (data mining), например, карт понятий. При фокусировке внимания в определенной точке просматриваемой карты происходит зрительное преувеличение (гиперболизация) концептуального окружения точки внимания.

В докладе *А. Е. Кибрика* «Лингвистическая реконструкция когнитивной структуры», ориентированном на специалистов смежных областей, представлены ключевые идеи современного когнитивного подхода к языку: постулат об исходной когнитивной мотивированности языковой формы (в той мере, в которой эта форма мотивирована, она «отражает» стоящую за ней когнитивную структуру) и идея целенаправленной реконструкции (выявления) когнитивных структур по данным внешней языковой формы. На наглядных примерах продемонстрирована техника реконструкции глубинных когнитивных структур, стоящих за внешней языковой формой в ситуациях кажущейся «аномальности» языко-

[1] Первая российская конференция по когнитивной науке: Тез. докл. — КГУ, Казань, 2004. — 302 с.

вой формы, которая не поддается интерпретации с точки зрения обычной «наивной логики». Это такие случаи, когда (а) некоторая языковая форма «как бы» обозначает противоположные сущности, (б) разные языковые формы «как бы» обозначают одну и ту же сущность».

Доклад *Н. П. Бехтеревой*, *С. Г. Данько и С. В. Медведева* «Нейрофизиологические исследования творческого мышления: пути, результаты, проблемы» лежит в русле еще только зарождающегося направления психофизиологии творчества, имеющего целью выявление мозговой организации творческих процессов. Исследование стало возможным благодаря достигнутой в последнее десятилетие и продолжающей развиваться интеграции когнитивной психологии, нейропсихологии и когнитивной нейрофизиологии, благодаря прогрессу сложных современных технологий исследования деятельности мозга. Полученные на сегодня результаты интерпретируются авторами как свидетельство наличия многоуровневой системы обеспечения творческих процессов человека, существенными элементами которой являются интенсивные локальные процессы в некоторых зонах коры головного мозга.

Если рассматривать общие междисциплинарные доклады и доклады по когнитивной психологии с позиций современной науки управления, можно сделать следующие выводы.

- Интеллектуальная деятельность людей, в которой они выступают и как субъекты, и как объекты управления, и как создатели систем управления, сегодня не рассматривается в качестве ожидаемой области приложения знаний для когнитивной науки; исследования по когнитивным аспектам такой деятельности практически не представлены на конференции. Но именно эта деятельность в последние 10 – 20 лет занимает все большее место в науке управления в качестве объекта теоретического моделирования и практических приложений. При этом для представителей прикладных наук проблемой является отсутствие внутренних средств верификации своих теоретических моделей, о чем говорил в дискуссии *О. П. Кузнецов* в связи с моделями искусственного интеллекта.
- Знания о человеке, полезные и даже существенные для специалистов по управлению, искусственному интеллекту, информатике, были представлены в ряде докладов. Однако их ценность далеко не всегда столь очевидна, как это было, например, в случае докладов Х. Риттера.

Показателен доклад психологов *М. А. Холодной* и *И. С. Кострикиной* «Когнитивные и метакогнитивные предпосылки интеллектуальной компетентности». В нем отражены результаты экспериментального исследования зависимости успешности в профессионально ориентированных видах интеллектуальной деятельности от различных показателей, принятых в психологии для характеристики интеллектуальной компетентности (иными словами, интеллектуальных ресурсов) человека. При этом считается, что эксперты — это лица, интеллектуально компетентные в конкретной предметной области, т. е. опытные, знающие, принимающие эффективные решения.

Уместно сопоставить психологическую точку зрения на зависимость успешности эксперта (как его «выходной» характеристики) от его интеллектуальных ресурсов с парадигмой, широко принятой в науке управления. Согласно этой парадигме, эксперт (или иной специалист, которого естественно рассматривать как интеллектуально компетентное лицо в своей профессиональной деятельности) — это человек, который может представить или сформировать информацию, необходимую для последующего применения формальных методов при решении практических задач. Эта информация может включать в себя оценки, позволяющие снять неопределенность; структурирование своего видения проблемной ситуации или представлений о содержании и целях деятельности людей, их стратегий и предпочтений и т. д.

Интерпретируя структуру и содержание знаний и представлений, навязываемые человеку (эксперту или другому интеллектуально компетентному лицу) как его интеллектуальные ресурсы, нетрудно понять, что в рамках рассматриваемой парадигмы аспект влияния таких ресурсов на успешность почти не рассматривается. Иначе говоря, как правило, остается открытым вопрос о том, приведут ли предлагаемые человеку интеллектуальные ресурсы к более успешному решению практических задач. Более того, отсутствуют нормы, толкающие разработчика того или иного экспертного метода к каким-либо проверкам и подтверждениям успешности метода в прикладной деятельности, которые выходили бы за рамки рассуждений «здравого смысла», «естественности» и т. п. Тем более, отсутствуют нормы, определяющие приемлемые формы таких проверок и подтверждения успешности.

Коротко говоря, психологический взгляд на экспертов толкает к пересмотру принятой в науке управления парадигмы. С другой стороны, довольно очевидно, что знания, которыми обладают специалисты по управлению, имеющие опыт практического применения своих знаний, гораздо богаче, чем очень общие теоретические схемы, которыми по необходимости пользуются психологи. Целесообразность интеграции на основе междисциплинарных исследований представляется очевидной.

Проведенный анализ результатов конференции и обсуждение со специалистами Института проблем управления, которые так или иначе касаются когнитивных аспектов управления, и, более широко, человеческих факторов в управлении, позволил сделать следующие организационные выводы.

- Целесообразно способствовать интеграции и развитию теоретических подходов в рамках науки управления, способствующих более эффективному решению практических задач управления в условиях действия человеческих факторов. Одним из организующих начал и центров кристаллизации может стать журнал «Проблемы управления». Выход в свет данного номера, в котором опубликованы две статьи по когнитивному анализу и настоящее сообщение, быть может, станет отправной точкой этого процесса. В дальнейшем возможны появления в журнале соответствующей рубрики и публикация тематических подборок статей.
- Участие в интеграционном процессе, активизированном организаторами конференции и ассоциации по когнитивной науке, представителей науки управления полезно не только для развития этой науки, но и для когнитивных исследований как таковых. Целесообразно продолжить консультации с организаторами конференции и ассоциации по когнитивной науке для поиска более эффективных форм взаимодействия, начиная с расширения основных направлений деятельности планируемой ассоциации.

Н. А. Абрамова

☎ *(095) 334-90-61*

E-mail: abramova@ipu.ru ❑

CONTENTS & ABSTRACTS