

НБ МИФИ

621.039

К89

МОСКОВСКИЙ ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
ИНЖЕНЕРНО-ФИЗИЧЕСКИЙ ИНСТИТУТ



А. М. Кузьмин

МЕТОДЫ ОПТИМИЗАЦИИ
ЯДЕРНО-ЭНЕРГЕТИЧЕСКИХ
УСТАНОВОК

МОСКВА 1985

Министерство высшего и среднего специального
образования СССР

621.039

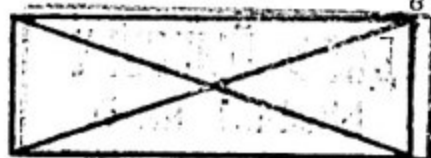
Московский ордена Трудового Красного Знамени
инженерно-физический институт

К89

А.М. Кузьмин

МЕТОДЫ ОПТИМИЗАЦИИ
ЯДЕРНО-ЭНЕРГЕТИЧЕСКИХ УСТАНОВОК

*Утверждено
редсоветом института
в качестве учебного пособия*



Москва 1985

Кузьмин А. М. Методы оптимизации ядерно-энергетических установок: Учебное пособие. — М.: Изд. МИФИ, 1985. — 80 с.

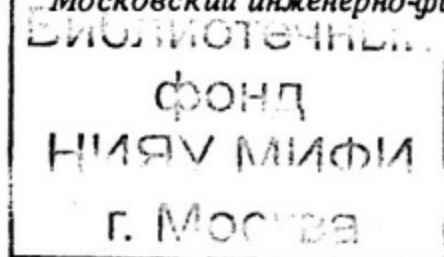
В пособии затрагивается широкий круг проблем, связанных с разработкой математических моделей физических процессов, формулировкой и решением оптимизационных задач. Рассматриваются способы получения коэффициентов чувствительности функционалов, необходимые условия оптимальности и алгоритмы метода линеаризации. Практическое использование алгоритмов иллюстрируется на примере решения задачи о выравнивании поля тепловыделения. Проводится анализ устойчивости и излагаются различные подходы к поиску оптимума в условиях неопределенностей в исходных данных.

Пособие предназначено для студентов старших курсов, аспирантов и инженеров, занимающихся автоматизацией расчетных оптимизационных исследований ядерно-энергетических установок.

Рецензенты:

В.В. Орлов, В.Д. Симонов

© Московский инженерно-физический институт, 1985 г.



ВВЕДЕНИЕ

В различных областях деятельности человека, в том числе при проектировании и эксплуатации ядерных реакторов и атомных электростанций, возникают разнообразные оптимизационные задачи. Оптимизационными принято называть такие задачи, в которых определяется наилучшее значение какой-либо характеристики при сохранении ограничений на ряд других характеристик, обеспечивающих безаварийную работу и необходимые технико-экономические условия эксплуатации установки. Успех в решении этих задач во многом зависит от умения согласовывать сложность математических моделей, выбранных для описания работы установки, с существующими методами оптимизации, возможностями ЭВМ, имеющимися в распоряжении исследователя, желаемой точностью и временем получения решения.

В последнее время большое внимание оптимизационным задачам уделяют в связи с возникновением нового направления в теории проектирования ядерно-энергетических установок (ЯЭУ), получившего название машинного проектирования [1]. Под машинным проектированием подразумевается процесс создания оптимальной в некотором смысле установки, системы, энергетического комплекса и т.п., базирующийся на разработке автоматизированной системы проектирования. Такая система предполагает тесное взаимодействие человека и ЭВМ с тем, чтобы наилучшим образом использовать их характерные особенности: быстродействие, большую память, высокую надежность и точность вычислений — для ЭВМ; интуицию, умение обобщать накопленный опыт и применять его для решения сходных задач, работать с нецифровыми данными и т.п. — для человека. Последнее предполагает наличие удобных средств общения человека и ЭВМ и соответствующее математическое обеспечение. В дальнейшем оптимизационные задачи, возникающие при проектировании установок, будем также называть задачами оптимального проектирования. В настоящее время имеются все необходимые предпосылки для успешного решения задач оптимального проектирования:

- созданы и имеются в достаточном количестве быстродействующие ЭВМ, обладающие большим объемом оперативной и внешней памяти;
- имеются самые современные и удобные средства общения человека с ЭВМ, такие, как перфораторы и быстродействующие печатающие устройства, дисплеи, графопостроители;
- создано необходимое математическое обеспечение, включающее разнообразные оптимизационные алгоритмы и диалоговые системы для обмена информацией между человеком и ЭВМ.

Из сказанного выше следует, что при решении оптимизационных задач приходится сталкиваться с широким кругом вопросов, начиная от формулировки задачи и описания протекающих в ЯЭУ процессов и кончая программированием и проведением расчетов на ЭВМ. Мы рассмотрим лишь те вопросы, которые касаются математических методов, используемых для автоматизации выбора наилучших характеристик ЯЭУ, при этом предполагаем, что читатель хорошо знаком со спецификой работы и способами описания процессов, протекающих в ядерных реакторах.

1. ФОРМУЛИРОВКА ЗАДАЧИ ОПТИМАЛЬНОГО ПРОЕКТИРОВАНИЯ

Придерживаясь установившейся в литературе терминологии, ядерно-энергетическую установку с реактором будем называть системой, а параметры u_1, u_2, \dots, u_n , за счет изменения которых в заданных пределах осуществляется выбор оптимальных технико-экономических характеристик, управляющими параметрами (или управлениями). Примером управлений могут служить размеры ядерного реактора, обогащение топлива, расход теплоносителя и т.п. Совокупность управлений удобно представлять в виде вектора управлений $\vec{u} = \{u_1, u_2, \dots, u_n\}$.

Различным значениям вектора управляющих параметров \vec{u} отвечают разные состояния нейтронно-физических, тепловых, гидродинамических процессов в реакторе и энергооборудовании, характеризуемые переменными состояниями $\varphi_l(x)$, $l=1, 2, \dots, p$. Переменные состояния определяются из решения некоторой системы нелинейных уравнений

$$\hat{A}_k(\vec{\varphi}, \vec{u}) = 0, \quad k=1, 2, \dots, p, \quad (1.1)$$

где $\vec{\varphi}(x)$ — вектор, составленный из компонент $\varphi_1(x), \varphi_2(x), \dots, \varphi_p(x)$, удовлетворяющих известным начальным и граничным условиям (x — совокупность пространственных, временных и др. координат); \hat{A}_k — операторы, выражающие вполне определенные действия над $\varphi(x)$ (интегрирование, дифференцирование и т.п.), если известны значения компонент вектора управлений \vec{u} . Совокупность уравнений (1.1), описывающих протекающие в системе процессы, принято называть математической моделью.

В качестве примера приведем систему уравнений, описывающих распределение нейтронных и изотопных полей в условно-критическом реакторе [2]:

$$-\hat{L}(\rho, \vec{u})\varphi + \frac{1}{\kappa_{эф}} \hat{Q}(\rho)\varphi = 0, \quad \varphi(x_g) = 0; \quad (1.2)$$

$$\hat{M}(\varphi)\rho = 0, \quad \rho(x_0) = f(\vec{u}), \quad (1.3)$$

где $\varphi(x)$ — распределение плотности потока нейтронов, удовлетворяющее граничному условию $\varphi(x_s) = 0$; $\rho(x)$ — совокупность ядерных концентраций, принимающих в начальный момент времени x_0 известные значения $\rho(x_0) = f(\vec{u})$; $\hat{L}(\rho, \vec{u})$, $\hat{Q}(\rho)$ — операторы уравнения переноса нейтронов, описывающие поглощение и замедление (\hat{L}) нейтронов, а также возникновение нейтронов (\hat{Q}) за счет деления ядер среды; $K_{эф}$ — эффективный коэффициент размножения нейтронов в реакторе; $\hat{M}(\varphi)$ — оператор уравнения выгорания, описывающий изменение со временем изотопного состава реактора.

Эти уравнения являются нелинейными, поскольку макроскопические сечения, входящие в операторы $\hat{L}(\rho, \vec{u})$ и $\hat{Q}(\rho)$, зависят от концентраций ядер, а скорость выгорания ядерного горючего, входящая в оператор $\hat{M}(\varphi)$, определяется плотностью потока нейтронов.

Обычно в качестве сравнительных характеристик состояний системы используются не переменные состояния $\vec{\varphi}(x)$, а функционалы $F_i(\vec{\varphi}, \vec{u})$, ставящие в соответствие переменным $\vec{\varphi}(x)$ (при известных управлениях \vec{u}) определенные числа F_i . Один из таких функционалов $F_0(\vec{\varphi}, \vec{u})$ может рассматриваться как критерий оптимальности, определяющий эффективность работы системы в целом, а на остальные — накладываться ограничения, выполнение которых обеспечивает безаварийную работу системы. Примерами таких функционалов могут служить коэффициент воспроизводства вторичного горючего, максимальная температура оболочки твэла, стоимость электроэнергии и т.п.

Теперь задачу выбора оптимальных проектных решений можно сформулировать в следующем виде.

Найти такие значения управлений u_j , удовлетворяющие неравенствам:

$$u_j^{\text{МНН}} \leq u_j \leq u_j^{\text{МАКС}}, \quad j = 1, 2, \dots, n, \quad (1.4)$$

при которых функционал $F_0(\vec{\varphi}, \vec{u})$ принимает оптимальное (минимальное или максимальное) значение (u) выполняются ограничения:

$$F_i(\vec{\varphi}, \vec{u}) \leq F_i^{\text{доп}}, \quad i = 1, 2, \dots, m, \quad (1.5)$$

для некоторых функционалов $F_i(\vec{\varphi}, \vec{u})$, переменные состояния $\vec{\varphi}(x)$ подчиняются уравнениям (1.1), а значения $u_j^{\text{МНН}}$, $u_j^{\text{МАКС}}$, $F_i^{\text{доп}}$ и других исходных данных известны.

Задачи в постановке (1.1) — (1.5) принято относить к классу детерминистических задач. Их отличает однозначный выбор критерия оптимальности $F_0(\vec{\varphi}, \vec{u})$ и полная определенность в значениях исходных данных, необходимых для описания состояния системы и ее характеристик. Однако при реальном проектировании такие задачи почти не встречаются. Как пра-

вило, имеются неопределенности как в выборе критерия оптимальности, так и в задании исходных данных.

Появление неопределенности в выборе критерия оптимальности F_0 связано с тем, что к проектируемой системе (или установке) предъявляются зачастую противоречивые требования, отражающие реальные условия ее работы в пределах энергосистемы страны. Например, можно понять желание проектировщиков создать атомную электростанцию, которая бы была максимально безопасна в эксплуатации и имела бы минимально возможную стоимость вырабатываемой электроэнергии.

Неопределенность исходных данных является объективным фактором, отражающим то обстоятельство, что наши знания об окружающем мире являются приближенными и неточными. Действительно, исходные данные, определяющие свойства используемых в реакторах материалов (микроскопические сечения взаимодействия, плотности материалов, теплотехнические константы и т.п.), получаются экспериментально и известны с некоторыми погрешностями. Кроме того, при проектировании приходится задаваться некоторыми параметрами, значения которых станут точно известны лишь после того, как установка будет изготовлена (параметры внешнего топливного цикла, режим эксплуатации и др.).

Совокупность всех исходных данных (включая те из них, которые полностью неопределены) обозначим через $\vec{\psi} = \{\psi_1, \psi_2, \dots, \psi_t\}$. В общем случае компоненты ψ_q вектора $\vec{\psi}$ могут принимать значения из некоторой (вообще говоря, неограниченной) области Ω , а от некоторых из них зависят как функционалы $F_i(\vec{\varphi}, \vec{u}, \vec{\psi})$ и операторы \hat{A}_k уравнений (1.1), так и допустимые значения $u_j^{\text{МНН}}(\vec{\psi})$, $u_j^{\text{МАКС}}(\vec{\psi})$, $F_i^{\text{дон}}(\vec{\psi})$.

В дальнейшем ограничимся рассмотрением лишь задач с неопределенными исходными данными, считая, например, что из числа нескольких критериев оптимальности выбран один определяющий, а на значения остальных критериев наложены ограничения вида (1.5). Такие задачи могут быть сформулированы следующим образом.

Найти такие значения управлений u_j , удовлетворяющих неравенствам

$$u_j^{\text{МНН}}(\vec{\psi}) \leq u_j \leq u_j^{\text{МАКС}}(\vec{\psi}), \quad j=1, 2, \dots, n, \quad (1.6)$$

при которых достигается наилучшее (например, минимальное) значение определяющего критерия:

$$F_0(\vec{\varphi}, \vec{u}, \vec{\psi}), \quad (1.7)$$

выполняются ограничения:

$$F_i(\vec{\varphi}, \vec{u}, \vec{\psi}) \leq F_i^{\text{дон}}(\vec{\psi}), \quad i=1, 2, \dots, m, \quad (1.8)$$

переменные состояния $\vec{\varphi}(x)$ находятся из решения уравнений

$$\hat{A}_k(\vec{\varphi}, \vec{u}, \vec{v}) = 0, \quad k=1, 2, \dots, p, \quad (1.9)$$

с известными граничными и начальными условиями, а параметры $v_q (q=1, 2, \dots, t)$ могут принимать значения из некоторой области Ω .

Задачи в постановке (1.6) – (1.9) принято называть задачами с неопределенными данными.

Решение приведенных выше задач является важным моментом в любой системе автоматизированного проектирования (САПР). При этом эффективность САПР во многом определяется тем, насколько удачно решены такие проблемы, как: создание математической модели проектируемой установки; автоматизация поиска оптимального управления; поиск наилучшего решения в условиях неопределенностей; рациональное разделение функций между ЭВМ и человеком.

Эти проблемы не имеют однозначного решения. В зависимости от условий разработки и функционирования САПР (от вида объекта проектирования, наличия той или иной вычислительной техники, времени проектирования и др.) они могут решаться по-разному. В данной работе сформулированы некоторые подходы к решению этих проблем, нашедшие применение при разработке САПР энергетических реакторов на быстрых нейтронах. Многие из них могут оказаться полезными при разработке САПР ядерных реакторов иного назначения.

2. МАТЕМАТИЧЕСКИЕ МОДЕЛИ

Математические модели, используемые при проектировании ЯЭУ, включают описание нейтронно-физических, теплофизических, гидродинамических, механических (расчеты на прочность) и др. процессов, происходящих в реакторе, теплообменнике и другом оборудовании АЭС. Поскольку наши знания об этих процессах не являются абсолютно точными, то разработка математической модели – процедура неформальная и во многом определяется опытом и интуицией разработчика, его умением выделять основные для рассматриваемой задачи проектирования процессы. Тем не менее можно сформулировать некоторые общие принципы, которых полезно придерживаться при создании математической модели и разработке отдельных расчетных модулей. Здесь и в дальнейшем под расчетным модулем понимается программа, составленная на одном из алгоритмических языков и позволяющая по известному алгоритму рассчитать одну или несколько характеристик проектируемой установки.

Требования, предъявляемые к точности математической модели, различны на разных этапах проектирования. На этапе эскизного проектиро-

вания, связанном с просмотром (хотя и целенаправленным) большого числа вариантов, приходится использовать упрощенное описание процессов, чтобы время счета на ЭВМ не было бы слишком большим. При этом следует учитывать, что упрощенные модели должны давать качественно правильное описание процессов, степень упрощения при описании различных процессов должна быть примерно одинакова.

Например, нет необходимости в знании детального распределения нейтронов в реакторе, если в формулировку задачи проектирования входят некоторые интегральные показатели (функционалы нейтронного поля).

На стадии создания технического проекта необходимо рассматривать все процессы и использовать возможно более точные их модели, чтобы иметь правильное представление о технико-экономических показателях установки. Усложнение расчетных моделей приводит к заметному росту времени счета на ЭВМ одного варианта. Однако на это идут, так как число просматриваемых вариантов не велико. Получаемые на этой стадии проектирования результаты могут в дальнейшем использоваться для корректировки приближенных моделей.

Из сказанного следует, что при разработке САПР ЯЭУ желательно удовлетворить следующим требованиям: система должна включать несколько расчетных модулей для описания одного и того же процесса и быть открытой, т.е. способной легко воспринимать новые расчетные модули, которые ранее в нее не входили; все расчетные модули должны быть согласованы по входным и выходным параметрам, включая их размерность; новые расчетные модули следует создавать, максимально используя возможности имеющихся в системе модулей.

Иметь несколько различных расчетных модулей для оценки одной и той же величины неудобно с точки зрения затрат памяти ЭВМ для их хранения и неизбежного увеличения времени на разработку САПР. В связи с этим часто предпочтение отдается таким расчетным модулям и алгоритмам, которые позволяют легко переходить от приближенного расчета к более точному. Большинство методов вычислительной математики допускает такой переход в определенных рамках, когда вид исходного уравнения не меняется. Если же уточнение расчетной модели связано с изменением вида уравнений, то обычно не удается избежать разработки различных расчетных модулей.

Рассмотрим некоторые часто используемые на практике алгоритмы, допускающие сравнительно простой переход от приближенного расчета к более точному. Проиллюстрируем их особенности на примере решения уравнения переноса нейтронов (1.2), записанного в многогрупповом диффузионном приближении

$$\nabla D^{(k)} \nabla \varphi^{(k)} - \Sigma^{(k)} \varphi^{(k)} + \sum_{j=1}^{k-1} \Sigma_d^{(j \rightarrow k)} \varphi^{(j)} + \frac{1}{k_{эф}} \chi^{(k)} \sum_{j=1}^m \nu_j^{(j)} \Sigma_f^{(j)} \varphi^{(j)} = 0, \quad (2.1)$$

где $\varphi^{(k)}(\vec{r})$ — плотность потока нейтронов k -й группы, удовлетворяющая граничному условию $\varphi^{(k)}(\vec{r}_s) = 0$; $D^{(k)}$ — коэффициент диффузии; $\sum_f^{(k)}$ — сумма сечений поглощения, упругого и неупругого рассеяния; $\sum_d^{(j \rightarrow k)}$ — сечение перевода нейтронов из группы j в группу k ; $\chi^{(k)}$ — спектр нейтронов деления; $\nu_f^{(k)} \sum_f^{(k)}$ — произведение сечения деления $\sum_f^{(k)}$ на среднее число нейтронов $\nu_f^{(k)}$, возникающих при делении ядер.

Будем считать, что решения уравнений (2.1) находятся с помощью метода итераций источника [2], на каждой итерации которого решаются последовательно, начиная с 1-й группы, уравнения вида:

$$\hat{L}\varphi(\vec{r}) = q(\vec{r}), \quad \varphi(\vec{r}_s) = 0, \quad (2.2)$$

где $\hat{L}\varphi(\vec{r}) = -\nabla D \nabla \varphi + \Sigma \varphi$, $q(\vec{r})$ — известное распределение источников, а индекс группы для простоты опущен. Отметим, что частным случаем уравнения (2.2) является известное уравнение теплопроводности [3], решение которого надо знать для оценки температур в реакторе.

2.1. Сеточные методы

В этих методах решение (2.2) сводится к решению системы линейных алгебраических уравнений с использованием следующей схемы рассуждений [4].

Вся область V изменения переменных $\vec{r} \in V$ разбивается на конечное число областей $\Delta V_{i,j,k}$ (таких, что $\sum_{i,j,k} \Delta V_{i,j,k} = V$) меньших размеров. Тем самым вводится в рассмотрение конечно-разностная сетка разбиения. Поступают тот или иной вид зависимости решения $\varphi(\vec{r})$ внутри области $\Delta V_{i,j,k}$ (обычно линейной с неизвестными значениями $\varphi_{i,j,k}$ в граничных точках области $\Delta V_{i,j,k}$). Путем интегрирования уравнения (2.2) внутри $\Delta V_{i,j,k}$ получают систему алгебраических (конечно-разностных) уравнений относительно неизвестных $\varphi_{i,j,k}$, которую затем решают с помощью известных методов линейной алгебры [5]. (Здесь i, j, k — индексы сетки разбиения вдоль определенным образом выбранных координатных осей.)

В случае одномерных геометрий приходят к системе трехточечных уравнений вида:

$$-a_{i-1} \varphi_{i-1} + d_i \varphi_i - c_{i+1} \varphi_{i+1} = q_i, \quad i = 1, 2, \dots, N, \quad (2.3)$$

которые решаются, используя метод прогонки [5].

В случае простых двумерных геометрий (например, конечного цилиндра с симметричными относительно оси свойствами) получают систему пятиточечных уравнений вида:

$$\begin{aligned}
 & -a_{i-1,j} \varphi_{i-1,j} - c_{i+1,j} \varphi_{i+1,j} - e_{i,j-1} \varphi_{i,j-1} - b_{i,j+1} \varphi_{i,j+1} + d_{i,j} \varphi_{i,j} = \\
 & = q_{i,j}, \quad i=1, 2, \dots, N, \quad j=1, 2, \dots, M, \quad (2.4)
 \end{aligned}$$

которую решают итерационно с помощью метода верхней релаксации, метода переменных направлений с определенным выбором ускоряющих параметров и др. [5].

Учитывая одно из сформулированных выше требований к САПР ЯЭУ, предпочтение имеет метод переменных направлений. Действительно, систему (2.4) можно переписать в виде:

$$-a_{i-1,j} \varphi_{i-1,j} + d_{i,j}^{(1)} \varphi_{i,j} - c_{i+1,j} \varphi_{i+1,j} = q_{i,j} + e_{i,j-1} \varphi_{i,j-1} + b_{i,j+1} \varphi_{i,j+1} - d_{i,j}^{(2)} \varphi_{i,j}; \quad (2.5)$$

$$-e_{i,j-1} \varphi_{i,j-1} + d_{i,j}^{(2)} \varphi_{i,j} - b_{i,j+1} \varphi_{i,j+1} = q_{i,j} + a_{i-1,j} \varphi_{i-1,j} + c_{i+1,j} \varphi_{i+1,j} - d_{i,j}^{(1)} \varphi_{i,j}, \quad (2.6)$$

где $d_{i,j} = d_{i,j}^{(1)} + d_{i,j}^{(2)}$.

Теперь легко организовать дополнительные (внутренние) итерации, решая последовательно системы (2.5) и (2.6) при сформированных (на основе значений $\varphi_{i,j}$ из предыдущей итерации) правых частях уравнений. При этом каждое из уравнений будет совпадать по виду с уравнениями (2.3), а системы уравнений (2.5) или (2.6) могут быть решены с помощью метода прогонки. Таким образом, имея расчетный модуль, реализующий на ЭВМ решение уравнений (2.3) для одномерного реактора, можно на его основе построить расчетный модуль для нахождения распределения нейтронов в двумерном реакторе.

В конечно-разностных методах переход от приближенного расчета $\varphi(\vec{r})$ к более точному происходит за счет увеличения числа узлов конечно-разностной сетки разбиения. Дальнейшего уточнения добиваются, увеличивая количество энергетических групп нейтронов.

2.2. Синтетические методы

В синтетических методах постулируется тот или иной вид зависимости потока нейтронов от координат. По степени детализации такой зависимости

их можно разделить на проекционные методы и методы итерационного синтеза.

В проекционных методах [6] искомое решение уравнений (2.2) ищется в виде разложения по заданной системе координатных функций $\psi_i(\vec{r})$:

$$\varphi(\vec{r}) = \sum_{i=1}^n \alpha_i \psi_i(\vec{r}), \quad (2.7)$$

удовлетворяющих граничному условию $\psi_i(\vec{r}_s) = 0$.

Неизвестные коэффициенты α_i находятся из решения системы линейных уравнений:

$$\sum_{i=1}^n \int_V \psi_j(\vec{r}) \hat{L} \psi_i(\vec{r}) d\vec{r} \alpha_i = \int_V q(\vec{r}) \psi_j(\vec{r}) d\vec{r}, \quad j=1,2,\dots,n. \quad (2.8)$$

В качестве координатных функций $\psi_i(\vec{r})$ часто берутся собственные функции какой-либо простой "родственной" задачи. Например, при нахождении потоков нейтронов в многозонном реакторе функции $\psi_i(\vec{r})$ определяются как первые n собственных функций задачи: $\Delta \psi_i(\vec{r}) + \lambda_i \psi_i(\vec{r}) = 0$, $\psi_i(\vec{r}_s) = 0$.

Точность расчета потока нейтронов повышается с ростом числа координатных функций. Одновременно растет размерность системы уравнений (2.8) и связанные с ее решением вычислительные затраты. Поэтому эффективность алгоритмов проекционных методов определяется, во-первых, тем, насколько хорошо небольшой набор функций $\psi_i(\vec{r})$ аппроксимирует искомое решение $\varphi(\vec{r})$, и, во-вторых, удастся ли до решения уравнений (2.8) получить и хранить в памяти ЭВМ в процессе всех итераций интегралы $\int_V \psi_j(\vec{r}) \hat{L} \psi_i(\vec{r}) d\vec{r}$.

Как показано в работе [7], проекционные методы могут с успехом применяться для получения детального хода потока нейтронов в сильно гетерогенных системах.

Методы итерационного синтеза [8] позволяют заменить решение уравнения (2.2) в частных производных итерационным решением системы обыкновенных дифференциальных уравнений. Проиллюстрируем эту особенность, предполагая, что искомая функция φ зависит от двух переменных $x \in V_x$ и $y \in V_y$.

Будем искать решение уравнения (2.2) в виде:

$$\varphi(x, y) = \sum_{i=1}^n f_i(x) \psi_i(y). \quad (2.9)$$

Тогда относительно неизвестных функций $f_i(x)$ и $\psi_i(y)$ можно получить системы нелинейных одномерных уравнений:

$$\sum_{i=1}^n \int_{V_y} \psi_j(y) \hat{L} f_i(x) \psi_i(y) dy = \int_{V_y} q(x, y) \psi_j(y) dy, \quad j=1,2,\dots,n; \quad (2.10)$$

$$\sum_{i=1}^n \int_{V_x} f_i(x) \hat{L} f_i(x) \psi_i(y) dx = \int_{V_x} q(x, y) f_i(x) dx, \quad j=1, 2, \dots, n, \quad (2.11)$$

которые решаются итерационно, придерживаясь на каждой итерации следующего порядка. Сначала задается (или берется из предыдущей итерации) ортогональная система функций $\psi_i(y)$ и определяются $f_i(x)$ из решения (2.10). Найденные $f_i(x)$ ортогонализируются и используются при формулировке (2.11). Затем из решения (2.11) находят новые значения функций $\psi_i(y)$, которые после ортогонализации подставляют в (2.10). При этом для получения $f_i(x)$ или $\psi_i(y)$ каждое из соответствующих уравнений приводят к конечно-разностному виду и решают последние с помощью метода прогонки.

В методах итерационного синтеза переход от приближенного расчета к более точному происходит, как и в проекционных методах, за счет увеличения числа слагаемых в (2.9). Однако в отличие от проекционных методов функции, по которым ведется разложение искомого решения, не задаются, а выбираются с учетом специфики рассматриваемого реактора. Поэтому в методах итерационного синтеза обычно удается ограничиться для получения достаточно точного решения небольшим числом слагаемых в разложении (2.9).

Замечания: 1) частным случаем рассмотренного метода является метод условного разделения переменных [9], к которому приходят в том случае, когда $n=1$, а разложение (2.9) записывают для каждой зоны реактора; 2) часто разложения вида (2.7) или (2.9) применяют для функций, удовлетворяющих уравнениям (2.1); в этом случае приходят к задачам на нахождение ведущего собственного числа и функции соответствующих уравнений (алгебраических или обыкновенных дифференциальных).

2.3. Полиномиальные аппроксимации функционалов

В заключение остановимся на одном способе построения математических моделей, которые могут найти применение на стадии выработки технических предложений или эскизном проектировании. Этот способ связан с построением полиномиальных аппроксимаций функционалов, т.е. с нахождением для каждого функционала $F(\vec{\varphi}, \vec{u})$ задачи (1.1) — (1.5) зависимостей вида:

$$F(\vec{\varphi}, \vec{u}) = F(\vec{\varphi}^0, \vec{u}^0) + \sum_{j=1}^n a_j (u_j - u_j^0) + \sum_{j=1}^n \sum_{k=1}^n b_{jk} (u_j - u_j^0)(u_k - u_k^0) + \dots, \quad (2.12)$$

где u_j^0 ($j = 1, 2, \dots, n$) — фиксированные значения компонент управления \vec{u} ; $\vec{\varphi}^0$ — соответствующее вектору \vec{u}^0 решение системы (1.1).

Очевидно, что разложения (2.12) справедливы для достаточно гладких функционалов, имеющих непрерывные функциональные производные по параметрам u_j (вплоть до q -го порядка, если в (2.12) ограничиваются слагаемыми порядка u_j^q) в пределах всей области Ω , где используется это разложение.

Для определения коэффициентов a_j , $\beta_{j,k}$ и др. можно использовать следующие методы.

Методы теории возмущений высших порядков [10]. В этом случае коэффициенты в разложении (2.12) имеют смысл функциональных производных:

$$a_j = \left. \frac{dF}{du_j} \right|_{W^0}, \quad \beta_{j,k} = \left. \frac{d^2F}{du_j du_k} \right|_{W^0} \quad \text{и т.д.}, \quad (2.13)$$

определяемых в точке $W^0 = (\vec{\varphi}^0, \vec{u}^0)$ с учетом зависимости переменных состояния $\vec{\varphi}(x)$ от управления \vec{u} через уравнения (1.1).

Расчет производных (2.13) выполняется с привлечением теории сопряженных функций [11]. Для получения коэффициентов a_j необходимо найти сопряженные функции $\varphi_l^+(x)$, $l = 1, 2, \dots, p$ из решения системы уравнений

$$\sum_{l=1}^p \left[\frac{\partial \hat{A}_l(\vec{\varphi}, \vec{u})}{\partial \varphi_k} \right]_{W^0}^+ \varphi_l^+(x) = - \left. \frac{\partial F}{\partial \varphi_k(x)} \right|_{W^0}, \quad k = 1, 2, \dots, p, \quad (2.14)$$

а при нахождении коэффициентов $\beta_{j,k}$ можно ограничиться решением n систем уравнений (здесь n — размерность вектора \vec{u}) относительно неизвест-

ных функций $\varphi_l^{(j)}(x) = \frac{d\varphi_l(x)}{du_j}$ ($j = 1, 2, \dots, n$):

$$\sum_{l=1}^p \left[\frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial \varphi_l} \right]_{W^0} \varphi_l^{(j)}(x) = - \left[\frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial u_j} \right]_{W^0}, \quad k = 1, 2, \dots, p, \quad (2.15)$$

где операторы $\left[\frac{\partial \hat{A}_l(\vec{\varphi}, \vec{u})}{\partial \varphi_k} \right]_{W^0}^+$ и $\left[\frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial \varphi_l} \right]_{W^0}$ связаны интегральным соотношением:

$$\sum_{l=1}^p \int \varphi_l^{(j)} \sum_{k=1}^p \left[\frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial \varphi_l} \right]_{W^0}^+ \varphi_k^+ dx = \sum_{k=1}^p \int \varphi_k^+ \sum_{l=1}^p \left[\frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial \varphi_l} \right]_{W^0} \varphi_l^{(j)} dx. \quad (2.16)$$

Равенство (2.16) носит название тождества Лагранжа для функций $\varphi_i^{(j)}$ и φ_k^+ и служит основой для построения оператора $\left[\frac{\partial \hat{A}_1(\vec{\varphi}, \vec{u})}{\partial \varphi_k} \right]_{w^0}^+$.

Отметим, что имеется определенная специфика в решении уравнений вида (2.14) и (2.15) в том случае, когда функции $\varphi_k(x)$ имеют смысл плотности потока нейтронов и определяются как собственные функции задачи (2.1). (Более подробно этот вопрос освещен в разделе 3.)

Методы теории планирования эксперимента [12]. Эти методы нашли широкое применение при построении математических моделей быстрых реакторов в работе [13].

Прежде чем перейти к описанию способа получения коэффициентов, введем степенные функции $g_l(\vec{u})$ вида:

$$g_1(\vec{u}) = 1, \quad g_{1+j}(\vec{u}) = u_j - u_j^0 \quad (j=1, 2, \dots, n),$$

$$g_{1+n+g(j,k)}(\vec{u}) = (u_j - u_j^0)(u_k - u_k^0), \quad \text{где } g(j,k) = \frac{j(j-1)}{2} + k, \\ 1 \leq k \leq j$$

и т.д. Запишем разложение (2.12) в виде:

$$F = (\vec{\alpha}, \vec{g}(\vec{u})), \quad (2.17)$$

где $\vec{g}(\vec{u})$ — вектор-функция с компонентами $g_l(\vec{u})$, $l=1, 2, \dots, s$; $\vec{\alpha}$ — вектор, составленный из неизвестных коэффициентов разложения α_l , $l=1, 2, \dots, s$.

Пусть в нашем распоряжении имеется программа для определения переменных состояния $\vec{\varphi}(x)$ и получения значения, функционала $F(\vec{\varphi}, \vec{u})$ при заданном управлении \vec{u} . Рассчитав по этой программе m значений $F^{(k)}$ ($k=1, 2, \dots, m$) функционала $F(\vec{\varphi}, \vec{u})$, соответствующих различным значениям $\vec{u}^{(k)} \in \Omega$, можно оценить коэффициенты α_l по формуле:

$$\vec{\alpha} = \hat{M}^{-1} \vec{y}, \quad (2.18)$$

где $\vec{y} = \sum_{k=1}^m F^{(k)} \vec{g}(\vec{u}^{(k)})$, \hat{M} — матрица с элементами $M_{j,l} = \sum_{k=1}^m g_j(\vec{u}^{(k)}) g_l(\vec{u}^{(k)})$;

Ω — заранее выбранная область, не выходящая за пределы той области, которая описывается неравенствами (1.4). Оценки коэффициентов α по (2.18) называют наилучшими линейными оценками, а матрицу \hat{M} — информационной матрицей.

Значения функционала F , получаемые по (2.17) (с коэффициентами, найденными из соотношения (2.18)) при управлениях $\vec{u}^{(k)}$, будут отли-

чаться от рассчитанных по программе значений $F^{(k)}$. Степень расхождения можно характеризовать дисперсией:

$$\Delta^2 = \frac{1}{m-s} \sum_{k=1}^m [F^{(k)} - (\vec{\alpha}, \vec{g}(\vec{u}^{(k)}))]^2.$$

Оказывается, что дисперсия Δ^2 зависит от способа размещения точек $\vec{u}^{(k)}$ в области Ω . В теории оптимального планирования экспериментов сформулировано несколько критериев и указаны соответствующие этим критериям способы размещения точек $\vec{u}^{(k)}$.

Для повышения точности расчета по (2.12) следует либо увеличивать порядок аппроксимации, либо сокращать размеры области Ω . Поскольку с ростом степени аппроксимации резко увеличиваются вычислительные затраты на получение коэффициентов разложения, то обычно в (2.12) ограничиваются членами порядка u_i^2 и выбирают не слишком большие размеры области Ω .

3. АВТОМАТИЗАЦИЯ ПОИСКА ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

Выбор подходящего алгоритма для автоматизации поиска наилучшего варианта проектируемой установки имеет важное значение при разработке САПР ЯЭУ и во многом определяет ее эффективность. Из общих соображений ясно, что такие алгоритмы и программы должны удовлетворять следующим требованиям:

- учитывать особенности задачи и соответствовать выбранной математической модели объекта;
- обеспечивать получение оптимального управления с заданной проектировщиком точностью;
- не требовать больших загрузок памяти ЭВМ и давать возможность найти решение за приемлемое время;
- обладать достаточной гибкостью в отношении вида оптимизируемого функционала, количества и типов управляющих параметров и ограничений.

Сформулированная ранее задача оптимального проектирования (в этом разделе будет рассматриваться задача в постановке (1.1) — (1.5)) относится к числу неклассических вариационных задач на условный экстремум. В практически важных случаях в таких задачах количество ограничений (1.5) достигает нескольких десятков, число управляющих параметров — несколько сотен, а функционалы $F_i(\vec{\varphi}, \vec{u})$ имеют сложную зависимость от управления \vec{u} . Поэтому наиболее подходящими для их решения, оказываются численные итерационные методы поиска оптимального управления [14] — [16].

В данном пособии ограничимся рассмотрением методов, в основе которых лежит построение минимизирующей $F_0(\vec{\varphi}, \vec{u})$ последовательности управлений $\vec{u}^{(0)}, \vec{u}^{(1)}, \dots, \vec{u}^{(s)}, \dots$, не нарушающих в первом приближении ограничений (1.4) и (1.5).

В этих методах переход от управления $\vec{u}^{(s)}$ в s -й итерации к управлению $\vec{u}^{(s+1)} = \vec{u}^{(s)} + \delta \vec{u}^{(s)}$ в $(s+1)$ -й итерации осуществляется с помощью однотипных (т.е. не зависящих от номера итераций) операций: а) линеаризации задачи в окрестности управления $\vec{u}^{(s)}$; б) нахождения вариации $\delta \vec{u}^{(s)}$ из решения некоторых вспомогательных задач; в) проверки точности выполнения необходимых условий оптимальности.

Подобная организация поиска оптимального управления не требует слишком больших загрузок памяти ЭВМ, поскольку информацию, необходимую для нахождения $\delta \vec{u}^{(s)}$, следует хранить лишь в течение одной итерации. Что касается времени решения задачи, то оно зависит от первоначально выбранного управления $\vec{u}^{(0)}$ и заданной точности ε получения оптимального решения. Имея определенный опыт в задании $\vec{u}^{(0)}$ и варьируя точность ε , можно добиться, чтобы решение было получено за желаемое время.

Вместе с тем указанные методы применимы к решению задач с достаточно гладкими функционалами. Большинство физических характеристик реактора удовлетворяют этому требованию. Исключение составляют функционалы, характеризующие максимальное (или минимальное) значение функции. В этом случае за счет увеличения числа ограничений (1.5) можно перейти к задачам с гладкими функционалами. Например, пусть имеется ограничение

$$\max_{r \in \Delta R} q(r) \leq q^{доп},$$

где ΔR — интервал непрерывности функции $q(r)$. Предположим, что на s -й итерации метода $\max_{r \in \Delta R} q(r) = q(r_0)$. Тогда, вводя в рассмотрение значения $q(r_l)$ в точках r_l ($l=1, 2, \dots, N$), решают на s -й итерации задачу с ограничениями: $q(r_l) \leq q^{доп}$, $l=0, 1, \dots, N$. При этом стараются расположить точки r_l так, чтобы они перекрыли область, где может оказаться $\max_{r \in \Delta R} q(r)$ на $(s+1)$ -й итерации.

3.1. Линеаризация задачи

Линеаризация задачи (1.1) — (1.5) связана с нахождением коэффициентов чувствительности $\mathcal{H}_j^{(i)}$ функционалов $F_i(\vec{\varphi}, \vec{u})$ к j -м компонентам управления:

$$\mathcal{H}_j^{(i)} = \lim_{\delta u_j \rightarrow 0} \frac{\Delta F_i}{\delta u_j}, \quad (3.1)$$

где ΔF_i — приращение функционала $F_i(\vec{\varphi}, \vec{u})$, вызванное изменением j -й компоненты управления u_j на величину δu_j . Из (3.1) следует, что коэффициенты $\mathcal{H}_j^{(i)}$ имеют смысл первых производных функционалов по параметрам u_j . Если такие производные существуют и являются непрерывными функциями управления, то в достаточно малой окрестности Ω_s управления $\vec{u}^{(s)}$ приращение ΔF_i будет с хорошей точностью совпадать с вариацией δF_i , определяемой линейным соотношением

$$\delta F_i = \sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j. \quad (3.2)$$

Здесь принято, что $\delta u_j = u_j - u_j^{(s)}$, а коэффициенты $\mathcal{H}_j^{(i)}$ определены в точке $W^{(s)} = (\vec{\varphi}^{(s)}, \vec{u}^{(s)})$. Напомним, что $\vec{\varphi}^{(s)}$ — решение (1.1) при управлении $\vec{u}^{(s)}$.

Знание коэффициентов чувствительности $\mathcal{H}_j^{(i)}$ позволяет сформулировать задачу на нахождение таких δu_j , чтобы:

$$F_0(W^{(s)}) + \sum_{j=1}^n \mathcal{H}_j^{(0)} \delta u_j \leq F_0(W^{(s)}); \quad (3.3)$$

$$\sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j \leq F_i^{don} - F_i(W^{(s)}), \quad i=1, 2, \dots, m; \quad (3.4)$$

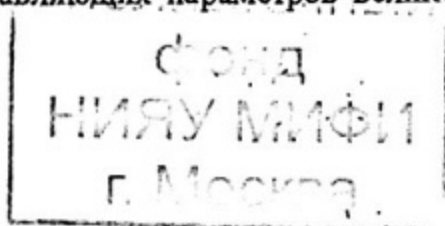
$$u_j^{min} \leq u_j^{(s)} + \delta u_j \leq u_j^{max}, \quad j=1, 2, \dots, n, \quad (3.5)$$

где $\delta \vec{u}$ — вектор с компонентами δu_j , $j=1, 2, \dots, n$. Задачу в постановке (3.3) — (3.5) будем называть линеаризованной по отношению к (1.1) — (1.5).

Проще всего коэффициенты чувствительности оценить численно по формуле:

$$\mathcal{H}_j^{(i)} \approx \frac{F_i(\vec{\varphi}', \vec{u}') - F_i(\vec{\varphi}'', \vec{u}'')}{\lambda \Delta u_j}, \quad (3.6)$$

где $F_i(\vec{\varphi}', \vec{u}')$ — значение функционала, рассчитанного при управлении \vec{u}' , в котором j -я компонента заменена на $u_j^{(s)} + \Delta u_j$; $F_i(\vec{\varphi}'', \vec{u}'')$ — значение того же функционала, полученного при управлении \vec{u}'' , в котором j -я компонента заменена на $u_j^{(s)} - \Delta u_j$. Однако такой способ требует больших затрат времени, если количество управляющих параметров велико. Кроме



того, может произойти потеря точности расчета, так как в числителе (3.6) стоит разность чисел, каждое из которых определяется с некоторой погрешностью (в случае приближенного решения уравнений (1.1) для переменных состояния $\vec{\varphi}'$ и $\vec{\varphi}''$). Поэтому часто предпочтение отдается способу расчета коэффициентов $\mathcal{H}_j^{(i)}$, основанному на использовании соотношений теории малых возмущений [11]. В этом случае общая схема расчета $\mathcal{H}_j^{(i)}$ выглядит следующим образом.

1) Запишем вариацию δF_i функционала F_i , ограничиваясь членами первого порядка малости относительно норм $\|\delta\vec{\varphi}\|$ и $\|\delta\vec{u}\|$:

$$\delta F_i = \sum_{j=1}^n \frac{\partial F_i}{\partial u_j} \Big|_{W^{(s)}} \delta u_j + \sum_{l=1}^p \int_V \frac{\partial F_i}{\partial \varphi_l(x)} \Big|_{W^{(s)}} \delta \varphi_l(x) dx, \quad (3.7)$$

где $\frac{\partial F_i}{\partial u_j} \Big|_{W^{(s)}}$ — частные производные, а $\frac{\partial F_i}{\partial \varphi_l(x)}$ — функциональные производные, рассчитываемые в точке $W^{(s)}$.

2) Проварируем уравнения (1.1) для переменных состояния:

$$\sum_{l=1}^p \frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial \varphi_l(x)} \Big|_{W^{(s)}} \delta \varphi_l(x) = - \sum_{j=1}^n \frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial u_j} \Big|_{W^{(s)}} \delta u_j, \quad k=1, 2, \dots, p, \quad (3.8)$$

предполагая существование и непрерывность производных $\frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial \varphi_l(x)} \Big|_{W^{(s)}}$ и $\frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial u_j} \Big|_{W^{(s)}}$ от операторов \hat{A}_k в точке $W^{(s)}$. Введем в рассмотрение сопряженные операторы $\left[\frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial \varphi_l} \right]_{W^{(s)}}^+$, определенные на классе функций $\varphi_k^+(x), k=1, 2, \dots, p$, удовлетворяющих интегральному соотношению (2.16). Выберем из этого класса функции, являющиеся решениями системы уравнений:

$$\sum_{k=1}^p \left[\frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial \varphi_l(x)} \right]_{W^{(s)}}^+ \varphi_k^+(x) = - \frac{\partial F_i}{\partial \varphi_l(x)} \Big|_{W^{(s)}}, \quad l=1, 2, \dots, p. \quad (3.9)$$

Отметим, если в число операторов \hat{A}_k входят дифференциальные операторы, то соотношение (2.16) выполняется лишь при определенных граничных условиях для функций $\varphi_k^+(x)$.

3) Умножая уравнения (3.8) на $\varphi_k^+(x)$, а (3.9) — на $\delta\varphi_l(x)$, интегрируя по области V и суммируя по всем индексам k и l , а затем вычитая из одного полученного при этом равенства другое, придем к соотношению:

$$\sum_{l=1}^p \int_V \left. \frac{\partial F_l}{\partial \varphi_l(x)} \right|_{W(s)} \delta\varphi_l(x) dx = \sum_{k=1}^p \int_V \varphi_k^+(x) \sum_{j=1}^n \left. \frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial u_j} \right|_{W(s)} \delta u_j dx.$$

4) Теперь вариацию (3.7) можно представить в виде:

$$\delta F_i = \sum_{j=1}^n \left\{ \left. \frac{\partial F_i}{\partial u_j} \right|_{W(s)} + \sum_{k=1}^p \int_V \varphi_k^+(x) \left. \frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial u_j} \right|_{W(s)} dx \right\} \delta u_j. \quad (3.10)$$

Из сравнения (3.2) и (3.10) получим для расчета коэффициентов чувствительности $\mathcal{H}_j^{(i)}$ формулу:

$$\left[\mathcal{H}_j^{(i)} = \left. \frac{\partial F_i}{\partial u_j} \right|_{W(s)} + \sum_{k=1}^p \int_V \varphi_k^+(x) \left. \frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial u_j} \right|_{W(s)} dx \right] \quad (3.11)$$

Центральным моментом в рассмотренной схеме является построение операторов $\left[\left. \frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial \varphi_l} \right]_{W(s)}^+$ и решение систем сопряженных уравнений

(3.9). При этом количество таких систем уравнений определяется числом функционалов $F_l(\vec{\varphi}, \vec{u})$ в задаче проектирования. Если m_1 — общее число таких функционалов, то расчеты коэффициентов $\mathcal{H}_j^{(i)}$ по (3.11) используют, когда $m_1 < n$. Когда $m_1 > n$ переходят к решению систем уравнений:

$$\sum_{l=1}^p \left. \frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial \varphi_l(x)} \right|_{W(s)} \varphi_l^{(j)}(x) = - \left. \frac{\partial \hat{A}_k(\vec{\varphi}, \vec{u})}{\partial u_j} \right|_{W(s)}, \quad (3.12)$$

каждая из которых получается из системы (3.8) при изменении лишь j -го управляющего параметра (очевидно, что: $\varphi_l^{(j)} = \frac{d\varphi_l(x)}{du_j}$). В этом случае:

$$\mathcal{H}_j^{(i)} = \left. \frac{\partial F_i}{\partial u_j} \right|_{W^{(s)}} + \sum_{z=1}^P \int_V \left. \frac{\partial F_i}{\partial \varphi_z(x)} \right|_{W^{(s)}} \varphi_z^{(j)}(x) dx. \quad (3.13)$$

Отметим, что в случае явной зависимости F_i от управления \vec{u} расчет коэффициентов $\mathcal{H}_j^{(i)}$ может быть выполнен с помощью известных формул дифференцирования, так как в этом случае $\mathcal{H}_j^{(i)} = \left. \frac{\partial F_i}{\partial u_j} \right|_{\vec{u}^{(s)}}$. Особенно простым оказывается такой расчет, когда построены полиномиальные аппроксимации функционалов (2.12).

3.2. Построение минимизирующей последовательности

Нахождение компонент $\delta u_j^{(s)}$, $j=1,2,\dots,n$ вектора $\delta \vec{u}^{(s)}$, удовлетворяющего условиям (3.3) – (3.5) линеаризованной задачи, может осуществляться разными методами [14] – [16]. В данном пособии ограничимся рассмотрением двух алгоритмов метода линеаризации, предполагая, что в задаче (1.1) – (1.5) все ограничения вида (1.5) имеют вид нестрогих неравенств. Это допущение не нарушает общности задачи, поскольку любое строгое равенство $F_{i_0}(\vec{\varphi}, \vec{u}) = F_{i_0}^{don}$ можно заменить двумя нестрогими неравенствами: $F_{i_0}(\vec{\varphi}, \vec{u}) \leq F_{i_0}^{don}$, $-F_{i_0}(\vec{\varphi}, \vec{u}) \leq -F_{i_0}^{don}$.

Первый алгоритм метода линеаризации предложен Р.П. Федоренко и использовался при решении разнообразных реакторных задач: об оптимальной остановке реактора [17], о выравнивании поля тепловыделения [18], о выборе оптимальных компоновок быстрых реакторов [9]. В этом алгоритме $\delta \vec{u}^{(s)}$ выбирается вдоль спроектированного на область (3.4), (3.5) направления быстрого убывания F_0 . Однако в силу линейности соотношений вида (3.2) перемещение по этому направлению ограничивается дополнительным требованием:

$$|\delta u_j^{(s)}| \leq \gamma_j^{(s)} |u_j^{(s)}|, \quad j=1,2,\dots,n, \quad (3.14)$$

где $\gamma_j^{(s)}$ – заданные положительные числа.

В зависимости от степени выполнения ограничений (1.5) различают два способа выбора $\delta \vec{u}^{(s)}$.

Пусть все ограничения (1.5) при $\vec{u} = \vec{u}^{(s)}$ выполнены с точностью до $\varepsilon > 0$, т.е. $F_i(W^{(s)}) - F_i^{don} \leq \varepsilon$, $i=1,2,\dots,m$. Тогда компоненты $\delta u_j^{(s)}$ определим, решая задачу линейного программирования:

$$\begin{aligned}
 & \min \sum_{j=1}^n \mathcal{H}_j^{(0)} \delta u_j^{(s)} ; \\
 & \sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j^{(s)} \leq F_i^{\text{дон}} - F_i(W^{(s)}), \quad i=1,2,\dots,m ; \\
 & S_j^- \leq \delta u_j^{(s)} \leq S_j^+, \quad j=1,2,\dots,n,
 \end{aligned}
 \quad \left. \vphantom{\sum_{j=1}^n} \right\} (3.15)$$

где

$$\begin{aligned}
 S_j^- &= \max(-\gamma_j^{(s)} |u_j^{(s)}|, u_j^{\text{мин}} - u_j^{(s)}) ; \\
 S_j^+ &= \min(\gamma_j^{(s)} |u_j^{(s)}|, u_j^{\text{макс}} - u_j^{(s)}) .
 \end{aligned}
 \quad \left. \vphantom{\sum_{j=1}^n} \right\} (3.16)$$

Для этого можно воспользоваться любым алгоритмом, изложенным в работе [19]. Один из таких алгоритмов описан в приложении 2.

Пусть степень нарушения некоторых из ограничений (1.5) превышает заданную точность ε . Например, при $\vec{u} = \vec{u}^{(s)}$ оказалось:

$$\begin{aligned}
 F_i(W^{(s)}) - F_i^{\text{дон}} &> \varepsilon, \quad i=1,2,\dots,l, \\
 F_i(W^{(s)}) - F_i^{\text{дон}} &\leq \varepsilon, \quad i=l+1,\dots,m.
 \end{aligned}$$

Такая ситуация типична на первых итерациях метода из-за неудачно выбранного начального управления $\vec{u}^{(0)}$.

Введем в рассмотрение неотрицательные невязки $\xi_i^{(s)}$, соответствующие невыполненным условиям $F_i(W^{(s)}) - F_i^{\text{дон}} - \xi_i^{(s)} \leq 0$, $i=1,2,\dots,l$, и определим компоненты $\delta u_j^{(s)}$ из решения следующей задачи линейного программирования:

$$\begin{aligned}
 & \min \left(\sum_{i=1}^l \xi_i^{(s)} M_i^{(s)} + \alpha_0 \sum_{j=1}^n \mathcal{H}_j^{(0)} \delta u_j^{(s)} \right) ; \\
 & \sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j^{(s)} - \xi_i^{(s)} \leq F_i^{\text{дон}} - F_i(W^{(s)}), \quad i=1,2,\dots,l ; \\
 & \sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j^{(s)} \leq F_i^{\text{дон}} - F_i(W^{(s)}), \quad i=l+1,\dots,m ; \\
 & S_j^- \leq \delta u_j^{(s)} \leq S_j^+, \quad j=1,2,\dots,n ; \\
 & \xi_i^{(s)} \geq 0, \quad i=1,2,\dots,l,
 \end{aligned}
 \quad \left. \vphantom{\sum_{j=1}^n} \right\} (3.17)$$

в которой числа S_j^- и S_j^+ находятся из условий (3.16). В зависимости от способа выбора чисел α_0 и $M_i^{(s)}$ различаем две стратегии поиска: 1-я стратегия — $\alpha_0 = 0$, $M_i^{(s)} = 1$, $i=1, \dots, l$. В этом случае при нахождении допустимого управления не контролируется изменение функционала F_0 (он может принять значение, заметно отличающееся от оптимального); 2-я стратегия — $\alpha_0 = 1$; $M_i^{(s)} > 1$ ($i=1, \dots, l$) — заданные штрафные множители. В этом случае при разумном выборе $M_i^{(s)}$ можно избежать роста F_0 в процессе выполнения нарушенных ограничений.

Значения штрафных множителей $M_i^{(s)}$ следует задавать, учитывая степень нарушения и важность соответствующих ограничений. Например, можно воспользоваться оценкой

$$M_i^{(s)} = \frac{F_i(W^{(s)}) - F_i^{don}}{\varepsilon^n} \beta_i,$$

где $n > 1$, $\beta_i > 0$ — заданные числа, выбираемые на основе опыта решения сходных задач.

Также трудно дать четкие рекомендации по выбору чисел $\gamma_j^{(s)}$. Обычно их задают, используя рекуррентное соотношение $\gamma_j^{(s)} = C^{(s)} \gamma_j^{(s-1)}$; $\gamma_j^{(1)} = \gamma_j^{(0)}$, где $\gamma_j^{(0)}$ — начальное приближение, $C^{(s)}$ — заданные коэффициенты: $C^{(s)} \geq 1$, если условия (1.5) не нарушены и значения $F_i(W^{(s)})$ хорошо совпадают с оценками $F_i(W^{(s)}) + \sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j^{(s-1)}$; $C^{(s)} < 1$, если условия (1.5) оказались нарушенными на s -й итерации.

Второй алгоритм метода линеаризации предложен Б.Н. Пшеничным и подробно изложен в работе [16]. Он отличается от первого алгоритма иным способом выбора $\delta \vec{u}^{(s)}$, учитывающим штраф за большие отклонения $\|\delta \vec{u}^{(s)}\|$ как при проектировании направления быстрее уменьшения F_0 на область (3.4), (3.5), так и при перемещении вдоль этого направления. При этом не вводятся дополнительные ограничения вида (3.14) на $\delta \vec{u}^{(s)}$.

На каждой s -й итерации вначале с помощью изложенного в приложении 4 метода решается задача квадратичного программирования:

$$\min \left(\varrho \sum_{j=1}^n \mathcal{H}_j^{(0)} v_j + \frac{1}{2} \sum_{j=1}^n v_j^2 \right); \quad (3.18)$$

$$\left. \begin{aligned} \sum_{j=1}^n \mathcal{H}_j^{(i)} v_j &\leq F_i^{don} - F_i(W^{(s)}), \quad i=1, 2, \dots, m; \\ u_j^{min} - u_j^{(s)} &\leq v_j \leq u_j^{max} - u_j^{(s)}, \quad j=1, 2, \dots, n; \end{aligned} \right\} (3.19)$$

где $\gamma = \begin{cases} 0 & \text{при поиске допустимого управления,} \\ 1 & \text{при минимизации } F_0. \end{cases}$

В результате находятся не только компоненты v_j вектора \vec{v} , определяющего направление изменения управления $\vec{u}^{(s)}$:

$$\delta \vec{u}^{(s)} = \vec{u}^{(s+1)} - \vec{u}^{(s)} = \alpha \vec{v}, \quad (3.20)$$

но и переменные $\lambda_i \geq 0$, $i = 1, 2, \dots, m+2n$ двойственной по отношению к (3.18), (3.19) задачи.

Перемещение вдоль направления \vec{v} определяется параметром α , принимающим значения из полуинтервала $0 < \alpha \leq 1$. Для получения численного значения α составляется вспомогательный функционал $\Phi_N(\vec{\varphi}, \vec{u})$:

$$\Phi_N(\vec{\varphi}, \vec{u}) = F_0(\vec{\varphi}, \vec{u}) + N P(\vec{\varphi}, \vec{u}), \quad (3.21)$$

где N — заданный штрафной множитель, а функционал

$$P(\vec{\varphi}, \vec{u}) = \max \left\{ 0, F_1(\vec{\varphi}, \vec{u}) - F_1^{доп}, F_2(\vec{\varphi}, \vec{u}) - F_2^{доп}, \dots, F_m(\vec{\varphi}, \vec{u}) - F_m^{доп}, \right. \\ \left. u_1 - u_1^{макс}, u_2 - u_2^{макс}, \dots, u_n - u_n^{макс}, -u_1 + u_1^{мин}, -u_2 + u_2^{мин}, \dots, -u_n + u_n^{мин} \right\}.$$

После этого в качестве α берется то наибольшее значение из полуинтервала $0 < \alpha \leq 1$, при котором начинает выполняться неравенство

$$\Phi_N(\vec{\varphi}^{(s+1)}, \vec{u}^{(s)} + \alpha \vec{v}) \leq \Phi_N(\vec{\varphi}^{(s)}, \vec{u}^{(s)}) - \xi \alpha (\vec{v}, \vec{v}). \quad (3.22)$$

Определением α и получением нового управления $\vec{u}^{(s+1)}$ по (3.20) заканчивается одна итерация. Итерационный счет продолжается до тех пор, пока с заданной точностью ε не будут выполнены необходимые условия оптимальности, что эквивалентно выполнению неравенства [16]:

$$\sum_{j=1}^n v_j^2 \leq \varepsilon. \quad (3.23)$$

Остановимся на вычислительных аспектах второго алгоритма.

1. В работе [16] получены следующие оценки для α и N :

$$0 < \alpha \leq \frac{1 - \xi}{(N+1)C}, \quad N > \sum_{i=1}^{m+2n} \lambda_i,$$

где $C > 0$ — некоторая константа. Отсюда видно, что выбор слишком большого штрафного множителя N приведет к получению малого значения α и

процесс сходимости замедлится. Поэтому рекомендуется при решении задачи (3.18), (3.19) оценивать двойственные переменные λ_i и в равенстве (3.21) задавать значение N не на много большим, чем $\sum_i \lambda_i$.

2. Наибольшее значение α , при котором выполняется неравенство (3.22), можно выбрать из последовательности чисел $1, \frac{1}{2}, \frac{1}{2^2}, \dots, \frac{1}{2^k}$. В работе [16] показано, что неравенство (3.22) всегда будет выполнено при некотором конечном значении $k = k_0$.

3. С точки зрения сходимости процесса в целом необходимо, чтобы точность решения задачи квадратичного программирования (3.18), (3.19) (соответствующей двойственной задаче) была выше точности ε выполнения неравенства (3.23) по крайней мере на порядок.

4. При выборе направления \vec{v} необходимо учитывать фактор разрешимости задачи квадратичного программирования. Если на какой-либо итерации s задача (3.18), (3.19) оказывается неразрешимой, то предлагается сократить число ограничений. При этом всегда в число условий (3.19) включают ограничения вида $u_j^{\text{МИН}} \leq u_j \leq u_j^{\text{МАКС}}$, $j = 1, 2, \dots, n$, а из числа ограничений вида $F_i(\vec{\varphi}, \vec{u}) \leq F_i^{\text{доп}}$, $i = 1, 2, \dots, m$ включают лишь те, для которых $F_i(\vec{\varphi}, \vec{u}) - F_i^{\text{доп}} \geq \rho(\vec{\varphi}, \vec{u}) - \delta$, где $\delta > 0$ — заданный параметр.

Если множество индексов выбранных таким образом ограничений обозначить через I_δ , то при решении двойственной задачи будут получены лишь $\lambda_i \geq 0$, $i \in I_\delta$. Поэтому в дальнейшем штрафной множитель N должен выбираться из условия $N > \sum_{i \in I_\delta} \lambda_i$.

Рекомендуется указанную ранжировку ограничений проводить, начиная с первой итерации (чтобы сократить размерность задачи (3.18), (3.19)) и принимая $\delta^{(0)} = \rho(\vec{\varphi}^{(0)}, \vec{u}^{(0)}) + \xi$, где ξ равно тому же значению, которое принималось в (3.22). В дальнейшем δ следует уменьшать в два раза всякий раз, когда задача (3.18), (3.19) оказывается неразрешимой.

сентябрь 6.10.2017

3.3. Необходимые условия оптимальности

Необходимые условия оптимальности формируют соотношения между переменными состояния $\vec{\varphi}(z)$ и управлением \vec{u} , которые обязаны выполняться, когда получено оптимальное управление \vec{u}^* . В ряде простых задач с их помощью удастся найти оптимальное решение. Однако в большинстве случаев их применение ограничивается проверкой на оптимальность управления, найденного одним из прямых методов, к числу которых относится изложенный выше метод линеаризации.

$$(1.5) \quad F_i(\vec{\varphi}, \vec{u}) \leq F_i^{don}, \quad i=1, 2, \dots, m$$

При получении необходимых условий оптимальности будем считать, что среди ограничений (1.5) первые l имеют вид нестрогих неравенств, а остальные $(m-l)$ — вид строгих равенств, т.е.

$$\left. \begin{aligned} F_i(\vec{\varphi}, \vec{u}) - F_i^{don} &\leq 0, \quad i=1, 2, \dots, l; \\ F_i(\vec{\varphi}, \vec{u}) - F_i^{don} &= 0, \quad i=l+1, \dots, m. \end{aligned} \right\} \quad (3.24)$$

Кроме того, ограничения (3.24) дополним неравенствами:

$$\left. \begin{aligned} u_j - u_j^{max} &\leq 0, \quad j=1, 2, \dots, n; \\ -u_j + u_j^{min} &\leq 0, \quad j=1, 2, \dots, n. \end{aligned} \right\} \quad (3.25)$$

Пусть \vec{u}^* — оптимальное управление. Тогда любое малое изменение \vec{u}^* до значения $\vec{u}' = \vec{u}^* + \delta \vec{u}$, принадлежащего допустимой области, приводит к тому, что $\delta F_0 > 0$. Поэтому пересечение множеств $\omega_0 \cap \omega_1 \cap \dots \cap \omega_m \cap \omega_{m+1} \cap \dots \cap \omega_{m+n}$ дает пустое множество. Здесь ω_0 — множество тех вариаций δu_j , для которых: $\sum_{j=1}^n \mathcal{H}_j^{(0)} \delta u_j < 0$; $\omega_i (i=1, \dots, m)$ — множества тех вариаций δu_j , для которых $\sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j \leq F_i^{don} - F_i(W^*)$; $\omega_{m+j} (j=1, \dots, n)$ — множества тех вариаций δu_j , для которых $S_j^- \leq \delta u_j \leq S_j^+$, где числа S_j^+ и S_j^- определяются из равенств (3.16). Это означает, что задача линейного программирования в постановке (3.15), сформулированная в окрестности точки оптимума, имеет нулевое решение:

$$\delta u_j = 0, \quad j=1, 2, \dots, n; \quad \sum_{j=1}^n \mathcal{H}_j^{(0)} \delta u_j = 0. \quad (3.26)$$

Перепишем задачу (3.15) в виде (опуская номер итерации "s"):

$$\left. \begin{aligned} \min f, \quad f &= \sum_{j=1}^n \mathcal{H}_j^{(0)} \delta u_j; \\ \sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j &\leq F_i^{don} - F_i(W^*), \quad i=1, 2, \dots, l; \\ \sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j &= 0, \quad i=l+1, \dots, m; \\ \delta u_j &\leq S_j^+, \quad j=1, 2, \dots, n; \\ -\delta u_j &\leq -S_j^-, \quad j=1, 2, \dots, n, \end{aligned} \right\} \quad (3.27)$$

где

$$S_j^+ \geq 0, \quad -S_j^- \geq 0, \quad F_i^{\text{доп}} - F(W^*) \geq 0, \quad (3.28)$$

причем строгие равенства имеют место при тех индексах $j \in J^{\text{МАКС}}$, $j \in J^{\text{МИН}}$ и $i \in I^{\text{доп}}$, при которых либо компоненты u_j^* , либо функционалы $F_i(W^*)$ принимают соответственно допустимые значения: $u_j^{\text{МАКС}}$, $u_j^{\text{МИН}}$, $F_i^{\text{доп}}$.

Известно, что любой задаче линейного программирования можно сопоставить так называемую двойственную задачу [19]. Этот вывод следует из общей теории двойственности для задач выпуклого программирования, частным случаем которых являются задачи вида (3.27). Придерживаясь общего правила получения двойственных задач, изложенного в приложе-

нии 1, заменим сначала каждое ограничение типа равенства $\sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j = 0$ на два эквивалентных неравенства: $\sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j \leq 0$, $-\sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j \leq 0$, $i=1, \dots, m$.

Затем составим функцию Лагранжа

$$\begin{aligned} \mathcal{L}(\delta \vec{u}, \vec{y}) = & \sum_{j=1}^n \mathcal{H}_j^{(0)} \delta u_j + \sum_{i=1}^l y_i \left[\sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j - F_i^{\text{доп}} + F_i(W^*) \right] + \sum_{j=1}^n y_{1+j} (\delta u_j - S_j^+) + \\ & + \sum_{j=1}^n y_{1+n+j} (-\delta u_j + S_j^-) + \sum_{i=1+1}^m y'_{i+2n} \sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j - \sum_{i=1+1}^m y''_{i+2n} \sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j \end{aligned}$$

с неотрицательными двойственными переменными $y_i \geq 0$, $i=1, 2, \dots, 1+2n$; $y'_{i+2n} \geq 0$, $y''_{i+2n} \geq 0$, $i=1+1, \dots, m$. Если функцию $\mathcal{L}(\delta \vec{u}, \vec{y})$ представить в виде:

$$\begin{aligned} \mathcal{L}(\delta \vec{u}, \vec{y}) = & \sum_{j=1}^n \left[\mathcal{H}_j^{(0)} + \sum_{i=1}^l \mathcal{H}_j^{(i)} y_i + y_{1+j} - y_{1+n+j} + \sum_{i=1+1}^m \mathcal{H}_j^{(i)} (y'_{i+2n} - y''_{i+2n}) \right] \delta u_j - \\ & - \sum_{i=1}^l [F_i^{\text{доп}} - F_i(W^*)] y_i - \sum_{j=1}^n S_j^+ y_{1+j} - \sum_{j=1}^n (-S_j^-) y_{1+n+j} \end{aligned}$$

и принять во внимание, что двойственная функция $g(\vec{y}) = \min \mathcal{L}(\delta \vec{u}, \vec{y})$ при любых значениях δu_j (так как в \mathcal{L} включены все ограничения на δu_j), то получим:

$$g(\vec{y}) = - \sum_{i=1}^l [F_i^{\text{дон}} - F_i(W^*)] y_i - \sum_{j=1}^n S_j^+ y_{l+j} - \sum_{j=1}^n (-S_j^-) y_{l+n+j}.$$

Здесь учтено, что $\min \mathcal{L}$ существует лишь в том случае, когда

$$\sum_{i=1}^l \mathcal{H}_j^{(i)} y_i + y_{l+j} - y_{l+n+j} + \sum_{i=l+1}^m \mathcal{H}_j^{(i)} y_{i+2n} = -\mathcal{H}_j^{(0)}, \quad j = 1, 2, \dots, n,$$

где переменные $y_{i+2n} = y'_{i+2n} - y''_{i+2n}$, $i = l+1, \dots, m$, и могут быть любого знака. Последнее равенство вместе с условиями неотрицательности $y_i \geq 0$, $i = 1, 2, \dots, l+2n$ описывают область допустимых значений функции $g(\vec{y})$. Окончательно приходим к следующей двойственной задаче:

$$\max g(\vec{y}), \quad g(\vec{y}) = - \sum_{i=1}^l [F_i^{\text{дон}} - F_i(W^*)] y_i - \sum_{j=1}^n S_j^+ y_{l+j} - \sum_{j=1}^n (-S_j^-) y_{l+n+j}; \quad (3.29)$$

$$\left. \begin{aligned} \sum_{i=1}^l \mathcal{H}_j^{(i)} y_i + y_{l+j} - y_{l+n+j} + \sum_{i=l+1}^m \mathcal{H}_j^{(i)} y_{i+2n} &= -\mathcal{H}_j^{(0)}, \quad j = 1, \dots, n, \\ y_i &\geq 0, \quad i = 1, 2, \dots, l, l+1, \dots, l+2n. \end{aligned} \right\} \quad (3.30)$$

Легко показать, что задача (3.27) является двойственной по отношению к задаче (3.29) — (3.30). Поэтому деление задач линейного программирования на исходную и двойственную является чисто условным. Часто их называют задачами двойственной пары.

Связь между решениями задач двойственной пары устанавливает теорема 6 (см. приложение 1). Используя эту теорему и учитывая равенства (3.26), необходимые условия оптимальности можно теперь сформулировать в следующей форме.

Если \vec{u}^* — оптимальное управление, то существует решение задачи линейного программирования (3.29) — (3.30):

$$\left. \begin{aligned} y_i = \lambda_i \geq 0, \quad i = 1, 2, \dots, l; \quad y_{i+2n} = \lambda_i, \quad i = l+1, \dots, m, \\ y_{l+j} = \mu_j \geq 0, \quad y_{l+n+j} = \nu_j \geq 0, \quad j = 1, 2, \dots, n, \end{aligned} \right\} \quad (3.31)$$

при котором линейная форма (3.29) принимает значение $\max g = 0$.

Это условие оптимальности запишем в другой форме. Примем для определенности, что в точке оптимума W^* :

$$\left. \begin{aligned} F_i(W^*) &= F_i^{\text{доп}}, \quad i \in I^{\text{доп}} \subset I, \\ u_j^* &= u_j^{\text{макс}}, \quad j \in J^{\text{макс}} \subset J; \quad u_j^* = u_j^{\text{мин}}, \quad j \in J^{\text{мин}} \subset J, \end{aligned} \right\} (3.32)$$

где I — множество индексов $i=1, 2, \dots, m$, J — множество индексов $j=1, 2, \dots, n$, а $J^{\text{макс}} \cap J^{\text{мин}} = \emptyset$. Поскольку переменные $y_i \geq 0$, $i=1, 2, \dots, 1+2n$, и при оптимальном управлении выполнены условия (3.28), то $\max q = 0$ лишь в том случае, когда

$$\lambda_i [F_i^{\text{доп}} - F_i(W^*)] = 0, \quad i \in I; \quad \mu_j S_j^+ = 0, \quad \nu_j S_j^- = 0, \quad j \in J. \quad (3.33)$$

При этом должны выполняться соотношения:

$$\left. \begin{aligned} \text{а) для } j \in J^{\text{макс}} : \quad S_j^+ &= u_j^{\text{макс}} - u_j^* = 0, \quad \mu_j \geq 0; \\ \text{для } j \notin J^{\text{макс}} : \quad S_j^+ &> 0, \quad \mu_j = 0; \\ \text{б) для } j \in J^{\text{мин}} : \quad -S_j^- &= u_j^* - u_j^{\text{мин}} = 0, \quad \nu_j \geq 0; \\ \text{для } j \notin J^{\text{мин}} : \quad -S_j^- &> 0, \quad \nu_j = 0. \end{aligned} \right\} (3.34)$$

Принимая во внимание (3.20) — (3.34), необходимые условия оптимальности (3.31) можно сформулировать следующим образом.

Для того чтобы управление \bar{u}^* было решением задачи (1.1) — (1.5), необходимо существование таких множителей λ_i , $i=1, 2, \dots, m$, μ_j и ν_j , $j=1, 2, \dots, n$, что:

$$\left. \begin{aligned} \lambda_i [F_i^{\text{доп}} - F_i(W^*)] &= 0, \quad \lambda_i \geq 0, \quad i=1, 2, \dots, m; \\ \mu_j [u_j^{\text{макс}} - u_j^*] &= 0, \quad \mu_j \geq 0, \quad j=1, 2, \dots, n; \\ \nu_j [u_j^* - u_j^{\text{мин}}] &= 0, \quad \nu_j \geq 0, \quad j=1, 2, \dots, n; \end{aligned} \right\} (3.35)$$

$$\mathcal{H}_j^{(0)} + \sum_{i=1}^m \lambda_i \mathcal{H}_j^{(i)} + \mu_j - \nu_j = 0, \quad j=1, 2, \dots, n, \quad (3.36)$$

где коэффициенты $\mathcal{H}_j^{(i)}$ определены в точке W^* . Параметры λ_i , $i=1, 2, \dots, m$, μ_j, ν_j , $j=1, 2, \dots, n$, определяемые из условий (3.35), (3.36) или получаемые

в результате решения задачи (3.29) – (3.30), будем называть множителями Лагранжа.

3.4. Выравнивание поля тепловыделения

Проиллюстрируем основные положения раздела 3 на примере решения задачи о выравнивании поля тепловыделения, часто встречающейся при проектировании энергетических ядерных реакторов. Интерес к подобной задаче вызван тем, что в реакторе с заданными размерами активной зоны выравнивание тепловыделения позволяет повысить мощность реактора, увеличить энергонапряженность топлива и, в конечном счете, снизить стоимость вырабатываемой на АЭС электроэнергии.

Формулировка задачи. Выравнивание тепловыделения происходит за счет перераспределения материалов в реакторе и может обеспечиваться разными способами за счет: изменения обогащения топлива, введения в активную зону поглотителей нейтронов, окружения активной зоны отражателем с хорошими замедляющими свойствами и т.д. Остановимся на выравнивании тепловыделения обогащением топлива, предполагая, что с изменением обогащения концентрация основного делящегося изотопа $\rho^{(a)}$ (например, урана-235) меняется так, что концентрация топлива $\rho_T = \rho^{(a)} + \rho^{(b)}$ остается постоянной (здесь через $\rho^{(b)}$ обозначена концентрация материала для воспроизводства ядерного горючего – урана-238). Для простоты ограничимся рассмотрением одномерного цилиндрического реактора типа БН-350 [20], состоящего из нескольких зон с разными обогащениями топлива. Будем также считать, что распределение нейтронов $\psi^{(k)}(r)$ описывается в многогрупповом диффузионном приближении системой уравнений (2.1).

Неравномерность поля тепловыделения $q(r)$ вдоль радиуса r на интервале $0 \leq r \leq R_{a.z.}$ ($R_{a.z.}$ – внешний радиус активной зоны) принято характеризовать коэффициентом K_r :

$$K_r = \frac{1}{\langle q \rangle} \max_{0 \leq r \leq R_{a.z.}} q(r), \quad (3.37)$$

где $q(r) = \sum_{k=1}^m \sum_f^{(k)} \psi^{(k)}(r)$, а среднее значение $\langle q \rangle = \frac{2}{R_{a.z.}^2} \int_0^{R_{a.z.}} q(r) r dr$.

Очевидно, что всегда $K_r \geq 1$, причем $K_r = 1$ лишь в том случае, когда распределение $q(r)$ постоянно при всех $r \leq R_{a.z.}$. Таким образом, задача о выравнивании тепловыделения эквивалентна задаче о минимизации коэф-

коэффициента неравномерности K_r . Если принять, что активная зона собирается из идентичных топливных кассет (отличающихся, быть может, лишь обогащением топлива), и учесть при этом, что с каждого твэла снимается мощность, не превышающая некоторого допустимого значения, то станет ясным, что минимизация K_r в реакторе с заданными внешними размерами активной зоны равносильна максимизации полной мощности реактора.

При формулировке задачи необходимо учесть, что задача о минимуме K_r относится к числу минимаксных задач и является существенно неклассической. Не исключено, что в многозонных реакторах минимум K_r будет достигаться в точке, где первая производная от K_r по управлениям (обогащениям топлива и толщинам зон) терпит разрыв. Проиллюстрируем это на примере реактора с двумя зонами разного обогащения, размеры которых заданы. Для этого рассмотрим коэффициенты $K_{r,i}$, характеризующие неравномерности тепловыделений в пределах i -х зон:

$$K_{r,i} = \frac{1}{\langle q \rangle} \max_{r \in \Delta R_i} q(r), \quad (3.38)$$

где ΔR_i — толщина i -й зоны.

Качественная зависимость коэффициентов $K_{r,i}$ ($i=1,2$) от обогащения топлива во второй зоне $X_2 = \rho_2^{(2)}/\rho_T$ представлена на рис. 1; обогащение топлива $X_1 = \rho_1^{(1)}/\rho_T$ при этом подбиралось так, чтобы выполнялось условие критичности: $K_{эф} = 1$. Очевидно, что зависимость коэффициента K_r от обогащения X_2 описывается огибающей $K_r = \max_i K_{r,i}$. Из рисунка видно, что в точке минимума K_r производная $\partial K_r / \partial X_2$ терпит разрыв. Поэтому при минимизации K_r с помощью метода линеаризации полезно ввести дополнительные ограничения

$$K_{r,i} - K_r \leq 0, \quad (3.39)$$

что позволяет избежать резкого возрастания K_r на одной итерации метода, когда управляющие параметры близки к оптимальным значениям. На рис. 2 проиллюстрирована работа первого алгоритма метода при заданной области изменения X_2 на 5-й итерации. Легко видеть, что ограничения (3.39) эквивалентны условиям $\max_{r \in \Delta R_i} q(r) \leq q^{дон}$, которые ранее рекомендовалось вво-

дить при рассмотрении функционалов, характеризующих максимальное значение функции (в данном случае $q(r)$).

Теперь задачу о выравнивании поля тепловыделения в реакторе с заданными внешними размерами можно сформулировать следующим образом.

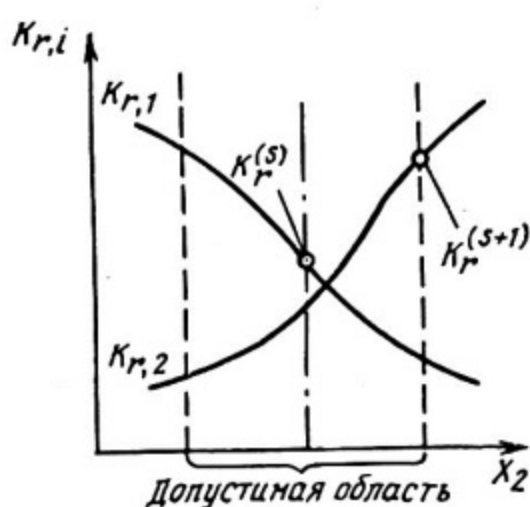


Рис. 1. Зависимости коэффициентов $K_{r,i}$ от обогащения X_2 топлива и изменение $K_r = \max K_{r,i}$ от S -й итерации к $(S+1)$ -й итерации в задаче без ограничения (3.39)

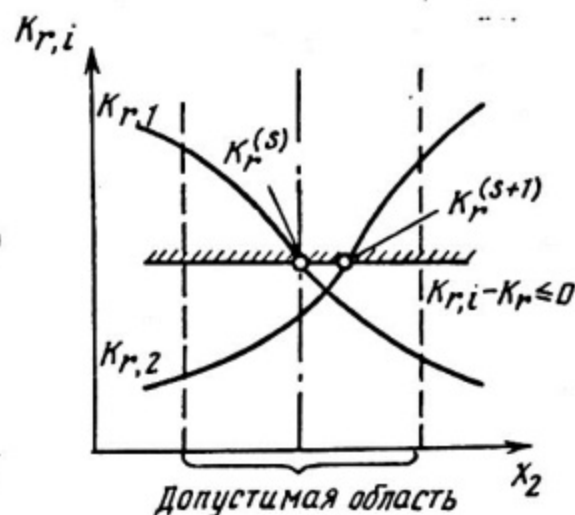


Рис. 2. Измерение $K_r = \max K_{r,i}$ от S -й итерации к $(S+1)$ -й итерации в задаче с ограничениями (3.39)

Найти такие значения концентраций $\rho_i^{(a)}$ и толщин ΔR_i зон, чтобы коэффициент неравномерности K_r (3.37) достигал наименьшего значения, распределения нейтронов $\Psi^{(k)}(r)$ определялись из решения (2.1), выполнялись ограничения:

$$K_{эф} = 1, \quad K_{r,i} - K_r \leq 0, \quad i=1, 2, \dots, n; \quad (3.40)$$

$$\sum_{i=1}^n \Delta R_i = R_{a.з.}, \quad 0 \leq \Delta R_i \leq R_{a.з.}, \quad 0 \leq \rho_i^{(a)} \leq \rho_T, \quad i=1, \dots, n, \quad (3.41)$$

а концентрация $\rho_i^{(b)} = \rho_T - \rho_i^{(a)}$. При этом $R_{a.з.}$, ρ_T , число зон n , микросечения элементов и концентрации всех материалов за исключением $\rho_i^{(a)}$ и $\rho_i^{(b)}$ известны.*

Концентрации $\rho_i^{(a)}$ и толщины зон ΔR_i ($i=1, 2, \dots, n$) играют роль управляющих параметров, условия (3.40), (3.41) определяют область допустимых значений управлений, а потоки нейтронов $\Psi^{(k)}(r)$, $k=1, 2, \dots, m$, имеют смысл переменных состояния.

Опыт решения задач с ограничениями показал целесообразность введения нормировки, т.е. использования безразмерных управлений и функционалов. Поэтому в данной задаче в качестве управлений полезно взять параметры:

* Предполагается, что в реакторе имеется n зон разного обогащения. Состав и размеры $(n+1)$ -й зоны (экрана) заданы.

$$u_j = \rho_j^{(a)} / \rho_T, \quad u_{j+n} = \Delta R_j / R_{a.ж.}, \quad j = 1, 2, \dots, n. \quad (3.42)$$

Тогда ограничения (3.41) примут вид:

$$\sum_{j=1}^n u_{j+n} = 1, \quad 0 \leq u_j \leq 1, \quad j = 1, 2, \dots, n, \quad (3.41a)$$

а концентрация $\rho_j^{(b)} = \rho_T (1 - u_j)$, $j = 1, 2, \dots, n$.

Линеаризация. При формулировке линеаризованной задачи вида (3.3) – (3.5) наиболее трудоемким является получение коэффициентов чувствительности функционалов $K_{ЭФ}$ и $K_{r,i}$. Эти функционалы определены на классе функций $\varphi^{(k)}(r)$, являющихся решениями (2.1) при всевозможных управлениях u_j (3.42), принадлежащих области (3.40), (3.41a). Кроме того, коэффициенты неравномерности $K_{r,i}$ (3.38) непосредственно зависят от u_j .

Основные расчетные формулы для коэффициентов чувствительности получим, переписав (2.1) в виде:

$$\left\{ -\frac{1}{r} \frac{d}{dr} (r \vec{J}) - \hat{\Sigma}_{ad} \vec{\varphi} + \frac{1}{K_{ЭФ}} \vec{J} (\vec{\nu} \hat{\Sigma}_f \vec{\varphi}) = 0; \quad (3.43)$$

$$\left\{ \frac{d\vec{\varphi}}{dr} + \hat{\Sigma}_{tr} \vec{J} = 0, \quad (3.44)$$

где $\vec{\varphi}(r)$ – вектор-функция потока нейтронов с компонентами $\varphi^{(k)}(r)$, $k = 1, 2, \dots, m$;

$\vec{J}(r)$ – вектор-функция с компонентами $J^{(k)} = -\frac{1}{3 \hat{\Sigma}_{tr}^{(k)}} \frac{d\varphi^{(k)}}{dr}$, $k = 1, 2, \dots, m$,

имеющими смысл проекций тока нейтронов k -й группы на нормаль к внешней границе реактора; \vec{J} и $\vec{\nu} \hat{\Sigma}_f$ – векторы с компонентами $J^{(k)}$ и $\nu_f^{(k)} \hat{\Sigma}_f^{(k)}$

соответственно, $k = 1, 2, \dots, m$; $\hat{\Sigma}_{tr}$ и $\hat{\Sigma}_{ad}$ – матрицы вида:

$$\hat{\Sigma}_{tr} = \begin{pmatrix} \hat{\Sigma}_{tr}^{(1)} & 0 & 0 & \dots & 0 \\ 0 & \hat{\Sigma}_{tr}^{(2)} & 0 & \dots & 0 \\ 0 & 0 & \hat{\Sigma}_{tr}^{(3)} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \hat{\Sigma}_{tr}^{(m)} \end{pmatrix}; \quad \hat{\Sigma}_{ad} = \begin{pmatrix} \hat{\Sigma}_{ad}^{(1)} & 0 & 0 & \dots & 0 \\ -\hat{\Sigma}_d^{(1+2)} & \hat{\Sigma}_{ad}^{(2)} & 0 & \dots & 0 \\ -\hat{\Sigma}_d^{(1+3)} & -\hat{\Sigma}_d^{(2+3)} & \hat{\Sigma}_{ad}^{(3)} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ -\hat{\Sigma}_d^{(1+m)} & -\hat{\Sigma}_d^{(2+m)} & -\hat{\Sigma}_d^{(3+m)} & \dots & \hat{\Sigma}_a^{(m)} \end{pmatrix};$$

$(\nu \vec{\Sigma}_f, \vec{\Phi})$ – скалярное произведение векторов $\nu \vec{\Sigma}_f$ и $\vec{\Phi}$ в точке r .

В уравнениях (3.43), (3.44) макросечения $\Sigma_{tr}^{(k)}$, $\nu \Sigma_f^{(k)}$, $\Sigma_d^{(j+k)}$, $\Sigma_{ad}^{(k)}$ постоянны в пределах какой-либо зоны реактора, а функции $\vec{\Phi}(r)$ и $\vec{J}(r)$ непрерывны и удовлетворяют граничным условиям:

$$\vec{J}(0) = 0, \quad \vec{\Phi}(R) = 0, \quad (3.45)$$

где R – радиус экстраполированной границы реактора.

Следуя общей схеме получения коэффициентов чувствительности, изложенной в п. 3.2, проварируем (3.43) и (3.44) в окрестности известного (полученного на предыдущей итерации метода линеаризации) решения, ограничиваясь членами первого порядка малости относительно $\|\delta \vec{\Phi}\|$, $\|\delta \vec{J}\|$,

$\delta \Sigma_{tr}^{(k)}$, $\delta \Sigma_{ad}^{(k)}$, $\delta(\nu \Sigma_f^{(k)})$, $\delta K_{эф}$:

$$\begin{aligned} & -\frac{1}{r} \frac{d}{dr} (r \delta \vec{J}) - \hat{\Sigma}_{ad} \delta \vec{\Phi} + \frac{1}{K_{эф}} \vec{J} (\nu \vec{\Sigma}_f, \delta \vec{\Phi}) = \\ & = \delta \hat{\Sigma}_{ad} \vec{\Phi} - \frac{1}{K_{эф}} \vec{J} (\delta \nu \vec{\Sigma}_f, \vec{\Phi}) + \frac{\delta K_{эф}}{K_{эф}^2} \vec{J} (\nu \vec{\Sigma}_f, \vec{\Phi}), \end{aligned} \quad (3.46)$$

$$\frac{d \delta \vec{\Phi}}{dr} + 3 \hat{\Sigma}_{tr} \delta \vec{J} = -3 \delta \hat{\Sigma}_{tr} \vec{J}. \quad (3.47)$$

Умножая (3.46) скалярно на $\vec{\Phi}^+(r)$, а (3.47) – на $\vec{J}^+(r)$, интегрируя по объему цилиндрического реактора в пределах $0 \leq r \leq R$ и складывая, получим:

$$\begin{aligned} & \int_0^R (\delta \vec{J}, \frac{d \vec{\Phi}^+}{dr} + 3 \hat{\Sigma}_{tr}^T \vec{J}^+) r dr + \int_0^R (\delta \vec{\Phi}, -\frac{1}{r} \frac{d}{dr} (r \vec{J}^+) - \hat{\Sigma}_{ad}^T \vec{\Phi}^+ + \\ & + \frac{1}{K_{эф}} \nu \vec{\Sigma}_f (\vec{J}^+, \vec{\Phi}^+)) r dr = \frac{\delta K_{эф}}{K_{эф}^2} \int_0^R (\vec{J}^+, \vec{\Phi}^+) (\nu \vec{\Sigma}_f, \vec{\Phi}^+) r dr - \\ & - \int_0^R (\vec{J}^+, 3 \delta \hat{\Sigma}_{tr} \vec{J}^+) r dr + \int_0^R (\vec{\Phi}^+, \delta \hat{\Sigma}_{ad} \vec{\Phi}^+) r dr - \\ & - \frac{1}{K_{эф}} \int_0^R (\vec{J}^+, \vec{\Phi}^+) (\delta \nu \vec{\Sigma}_f, \vec{\Phi}^+) r dr, \end{aligned} \quad (3.48)$$

где $\hat{\Sigma}_{tr}^r$, $\hat{\Sigma}_{ad}^r$ — транспонированные по отношению к $\hat{\Sigma}_{tr}$ и $\hat{\Sigma}_{ad}$ матрицы. При получении интегрального равенства (3.48) принято, что $\vec{\varphi}^+(r)$ и $\vec{J}^+(r)$ непрерывны при всех $0 \leq r \leq R$ и удовлетворяют граничным условиям:

$$\vec{J}^+(0) = 0, \quad \vec{\varphi}^+(R) = 0. \quad (3.49)$$

Эти условия вместе с равенствами (3.45) и непрерывностью $\vec{\varphi}^+(r)$, $\vec{J}^+(r)$, $\delta\vec{\varphi}(r)$, $\delta\vec{J}(r)$ приводят к тому, что

$$\sum_{i=1}^{n+1} (\vec{\varphi}^+, r\delta\vec{J}) \Big|_{R_{i-1}}^{R_i} = 0, \quad \sum_{i=1}^{n+1} (r\vec{J}^+, \delta\vec{\varphi}) \Big|_{R_{i-1}}^{R_i} = 0,$$

где R_i — радиусы зон реактора, $i = 0, 1, \dots, n+1$; $R_0 = 0$, $R_n = R_{a.z.}$, $R_{n+1} = R$.

Равенство (3.48) позволяет построить известные соотношения [21] теории малых возмущений для $\delta K_{эф}$ и $\delta K_{r,i}$. Действительно, если в качестве непрерывных функций $\vec{\varphi}^+(r)$ и $\vec{J}^+(r)$, удовлетворяющих условиям (3.49), выбрать решения $\vec{\varphi}_0^+(r)$ и $\vec{J}_0^+(r)$ однородных уравнений:

$$\begin{cases} -\frac{1}{r} \frac{d}{dr} (r\vec{J}_0^+) - \hat{\Sigma}_{ad}^r \vec{\varphi}_0^+ + \frac{1}{K_{эф}} \nu \hat{\Sigma}_f (\vec{J}_0^+, \vec{\varphi}_0^+) = 0; \\ \frac{d\vec{\varphi}_0^+}{dr} + 3 \hat{\Sigma}_{tr}^r \vec{J}_0^+ = 0, \end{cases} \quad (3.50)$$

соответствующие ведущему собственному числу $K_{эф}$, то для вариации $\delta K_{эф}$ найдем

$$\delta K_{эф} = \frac{K_{эф}^2}{\text{ЦНД}} \left[\int_0^R (\vec{J}_0^+, 3\delta \hat{\Sigma}_{tr}^r \vec{J}) r dr - \int_0^R (\vec{\varphi}_0^+, \delta \hat{\Sigma}_{ad}^r \vec{\varphi}) r dr + \right. \\ \left. + \frac{1}{K_{эф}} \int_0^R (\vec{J}_0^+, \vec{\varphi}_0^+) (\delta \nu \hat{\Sigma}_f, \vec{\varphi}) r dr, \right] \quad (3.51)$$

где ЦНД = $\int_0^R (\vec{J}_0^+, \vec{\varphi}_0^+) (\nu \hat{\Sigma}_f, \vec{\varphi}) r dr$ — ценность нейтронов деления.

Для получения аналогичного соотношения для коэффициента неравномерности $K_{r,i}$ предварительно запишем вариацию $\delta K_{r,i}$, считая, что максимальное тепловыделение в i -й зоне достигается в точке радиусом r_i . Тогда выражение (3.38) для $K_{r,i}$ можно представить в виде обычного дробно-линейного функционала:

$$K_{r,i} = \frac{\int_0^R (\vec{Z}_{f,i}, \vec{\varphi}) \delta(r-r_i) r dr}{R_{a.z.}^2 \int_0^{R_{a.z.}} (\vec{\Sigma}_f, \vec{\varphi}) r dr}, \quad (3.52)$$

где через $\vec{\Sigma}_{f,i}$ обозначено значение вектора $\vec{\Sigma}_f$ (с компонентами $\Sigma_{f,i}^{(k)}$) в i -й зоне. После этого вариация $\delta K_{r,i}$ запишется в виде:

$$\frac{\delta K_{r,i}}{K_{r,i}} = \frac{\int_0^R (\delta \vec{\Sigma}_{f,i}, \vec{\Phi}) \delta(r-r_i) r dr}{\int_0^R (\vec{\Sigma}_{f,i}, \vec{\Phi}) \delta(r-r_i) r dr} - \frac{\int_0^{R_{a.z.}} (\delta \vec{\Sigma}_f, \vec{\Phi}) r dr}{\int_0^{R_{a.z.}} (\vec{\Sigma}_f, \vec{\Phi}) r dr} + \int_0^{R_{a.z.}} \left(\frac{\partial K_{r,i}}{\partial \vec{\Phi}(r)}, \delta \vec{\Phi}(r) \right) r dr, \quad (3.53)$$

где функциональная производная $\partial K_{r,i} / \partial \vec{\Phi}$ рассчитывается по формуле

$$\frac{\partial K_{r,i}}{\partial \vec{\Phi}(r)} = \frac{\vec{\Sigma}_{f,i} \delta(r-r_i)}{\int_0^R (\vec{\Sigma}_{f,i}, \vec{\Phi}) \delta(r-r_i) r dr} - \frac{\vec{\Sigma}_f(r) \eta_{a.z.}}{\int_0^{R_{a.z.}} (\vec{\Sigma}_f, \vec{\Phi}) r dr}; \quad \eta_{a.z.} = \begin{cases} 1, & r \leq R_{a.z.}, \\ 0, & r > R_{a.z.}. \end{cases} \quad (3.54)$$

Принимая теперь в качестве $\vec{\Phi}^+(r)$ и $\vec{J}^+(r)$ частные решения $\vec{\Phi}_i^+(r)$ и $\vec{J}_i^+(r)$ уравнений

$$\begin{cases} -\frac{1}{r} \frac{d}{dr} (r \vec{J}_i^+) - \hat{\Sigma}_{ad}^T \vec{\Phi}_i^+ + \frac{1}{K_{3\varphi}} \partial \vec{\Sigma}_f (\vec{J}_i^+, \vec{\Phi}_i^+) = -\frac{\partial K_{r,i}}{\partial \vec{\Phi}(r)}, \\ \frac{d \vec{\Phi}_i^+}{dr} + 3 \hat{\Sigma}_{tr}^T \vec{J}_i^+ = 0, \quad i=1, 2, \dots, n, \end{cases} \quad (3.55)$$

удовлетворяющие равенствам (3.49) и условию ортогональности:

$$\int_0^R (\vec{J}_i^+, \vec{\Phi}_i^+) (\partial \vec{\Sigma}_f, \vec{\Phi}) r dr = 0,$$

для вариации $\delta K_{r,i}$ окончательно получим:

$$\begin{aligned} \frac{\delta K_{r,i}}{K_{r,i}} &= \frac{\int_0^R (\delta \vec{\Sigma}_{f,i}, \vec{\Phi}) \delta(r-r_i) r dr}{\int_0^R (\vec{\Sigma}_{f,i}, \vec{\Phi}) \delta(r-r_i) r dr} - \frac{\int_0^{R_{a.z.}} (\delta \vec{\Sigma}_f, \vec{\Phi}) r dr}{\int_0^{R_{a.z.}} (\vec{\Sigma}_f, \vec{\Phi}) r dr} + \int_0^R (\vec{J}_i^+, 3 \hat{\Sigma}_{tr}^T \vec{J}) r dr - \\ &- \int_0^R (\vec{\Phi}_i^+, \delta \hat{\Sigma}_{ad}^T \vec{\Phi}) r dr + \frac{1}{K_{3\varphi}} \int_0^R (\vec{J}_i^+, \vec{\Phi}_i^+) (\delta \partial \vec{\Sigma}_f, \vec{\Phi}) r dr. \end{aligned} \quad (3.56)$$

Отметим, что уравнения (3.50) и (3.55) для ценностей нейтронов по отношению к $K_{эф}(\varphi_0^T)$ и $K_{r,i}(\vec{\varphi}_i^T)$ являются частным случаем уравнений (3.9).

Равенства (3.51) и (3.56) устанавливают линейную связь между вариациями функционалов $\delta K_{эф}$ и $\delta K_{r,i}$ и вариациями макросечений зон. С их помощью легко получить коэффициенты чувствительности к обогащениям топлива u_j и относительным размерам зон u_{n+j} , $j=1,2,\dots,n$, учитывая определения (3.1), (3.2) и дополнительные соотношения, связывающие вариации макросечений с вариациями обогащений и размеров зон.

1. При нахождении коэффициентов чувствительности $\mathcal{H}_j^{(i)}$ к обогащениям u_j ($i=0,1,\dots,n$, $j=1,2,\dots,n$) следует принять во внимание известные формулы, выражающие макросечения $\Sigma_{p,j}^{(\kappa)}$ через концентрации $\rho_j^{(l)}$ и соответствующие микросечения $\sigma_{p,l}^{(\kappa)}$ 1-х элементов в j -х зонах:

$$\Sigma_{p,j}^{(\kappa)} = \sum_l \sigma_{p,l}^{(\kappa)} \rho_j^{(l)},$$

где p — тип процесса взаимодействия ($p=c, f, tr$ и др.). Они позволяют матрицы $\hat{\Sigma}_{tr}$, $\hat{\Sigma}_{ad}$ и векторы $\vec{\Sigma}_f$, $\nu \vec{\Sigma}_f$ записать в виде: $\hat{\Sigma}_{p,j} = \sum_l \hat{\sigma}_{p,l} \rho_j^{(l)}$ ($p=tr, ad$); $\vec{\Sigma}_{f,j} = \sum_l \vec{\sigma}_{f,l} \rho_j^{(l)}$; $\nu \vec{\Sigma}_{f,j} = \sum_l \nu \vec{\sigma}_{f,l} \rho_j^{(l)}$. Здесь структура $\hat{\sigma}_{p,l}$, $\vec{\sigma}_{f,l}$ и $\nu \vec{\sigma}_{f,l}$ такая же, как и соответствующих им $\hat{\Sigma}_p$, $\vec{\Sigma}_f$ и $\nu \vec{\Sigma}_f$.

Учитывая теперь, что с изменением обогащения топлива u_j в j -й зоне ($j=1,2,\dots,n$) меняются лишь концентрации делящегося $\rho_j^{(d)}$ и воспроизводящего $\rho_j^{(b)}$ материалов в той же зоне так, что $\delta \rho_j^{(d)} = \rho_T \delta u_j$, $\delta \rho_j^{(b)} = -\rho_T \delta u_j$, будем иметь:

$$\left. \begin{aligned} \delta \hat{\Sigma}_{p,j} &= (\hat{\sigma}_p^{(d)} - \hat{\sigma}_p^{(b)}) \rho_T \delta u_j, \quad (p=tr, ad); \\ \delta \vec{\Sigma}_{f,j} &= (\vec{\sigma}_f^{(d)} - \vec{\sigma}_f^{(b)}) \rho_T \delta u_j, \quad \delta \nu \vec{\Sigma}_{f,j} = (\nu \vec{\sigma}_f^{(d)} - \nu \vec{\sigma}_f^{(b)}) \rho_T \delta u_j. \end{aligned} \right\} (3.57)$$

2. При получении коэффициентов чувствительности $\mathcal{H}_j^{(i)}$ к относительным размерам u_{n+j} ($i=1,\dots,n$, $j=1,2,\dots,n$) следует иметь в виду, что:

$$R_i = \sum_{j=1}^i \Delta R_j, \quad \delta R_i = \sum_{j=1}^i R_{a.z.} \delta u_{n+j}. \quad (3.58)$$

Кроме того, изменение радиуса R_i на величину δR_i (при неизменных значениях $R_1, \dots, R_{i-1}, R_{i+1}, \dots, R_n$) эквивалентно замене в пределах

слоя толщиной δR_i материала $(i+1)$ -й зоны на материал i -й зоны, если $\delta R_i > 0$, или замене материала i -й зоны на материал $(i+1)$ -й зоны, если $\delta R_i < 0$ (рис. 3). Поэтому слагаемые в (3.51), (3.56) при достаточно малом δR_i можно записать в виде:

$$\left. \begin{aligned} \int_0^R (\vec{J}^+, 3\delta \hat{\Sigma}_{tr} \vec{J}) r dr &= (\vec{J}^+(R_i), [3\hat{\Sigma}_{tr,i} - 3\hat{\Sigma}_{tr,i+1}] \vec{J}(R_i)) R_i \delta R_i; \\ \int_0^R (\vec{\Psi}^+, \delta \hat{\Sigma}_{ad} \vec{\Psi}) r dr &= (\vec{\Psi}^+(R_i), [\hat{\Sigma}_{ad,i} - \hat{\Sigma}_{ad,i+1}] \vec{\Psi}(R_i)) R_i \delta R_i; \\ \int_0^R (\vec{J}, \vec{\Psi}^+) (\delta \vec{\Sigma}_f, \vec{\Psi}) r dr &= (\vec{J}, \vec{\Psi}^+(R_i)) (\vec{\Sigma}_{f,i} - \vec{\Sigma}_{f,i+1}, \vec{\Psi}(R_i)) R_i \delta R_i; \\ \int_0^R (\delta \vec{\Sigma}_f, \vec{\Psi}) r dr &= (\vec{\Sigma}_{f,i} - \vec{\Sigma}_{f,i+1}, \vec{\Psi}(R_i)) R_i \delta R_i, \end{aligned} \right\} (3.59)$$

где δR_i связаны с δu_{n+j} ($j=1, 2, \dots, n$) равенствами (3.58).

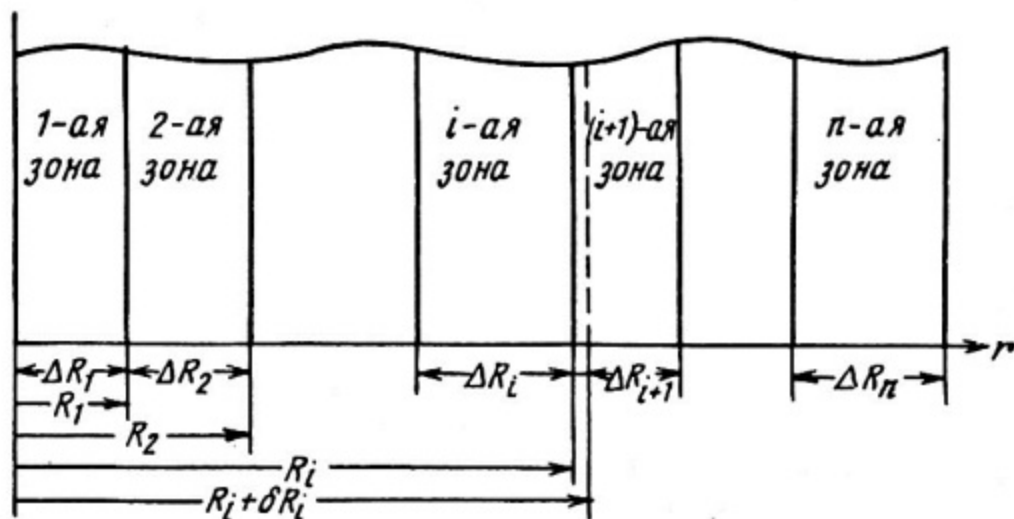


Рис. 3. Схема расположения зон и перемещения границы i -й зоны при замене материала $(i+1)$ -й зоны на материал i -й зоны в слое толщиной δR_i

Подставляя (3.57) и (3.59) в (3.51) и (3.56), получим для коэффициентов чувствительности $\mathcal{H}_j^{(i)}$ выражения:

$$\left. \begin{aligned}
 \mathcal{H}_j^{(0)} &= \frac{\kappa_{\text{эф}}^2}{\omega_{\text{нл}}} \rho_T h_j^{(0)} ; & \mathcal{H}_{n+j}^{(0)} &= \frac{\kappa_{\text{эф}}^2}{\omega_{\text{нл}}} R_{a.з.} h_{n+j}^{(0)} ; \\
 \mathcal{H}_j^{(i)} &= K_{r,i} \rho_T \left[\frac{(\vec{\sigma}_f^{(a)} - \vec{\sigma}_f^{(b)}, \vec{\varphi}(r_j)) \eta_j^{(i)}}{(\vec{\Sigma}_{f,j}, \vec{\varphi}(r_j))} - \frac{\int_{\Delta R_j} (\vec{\sigma}_f^{(a)} - \vec{\sigma}_f^{(b)}, \vec{\varphi}) r dr}{\int_0^{R_{a.з.}} (\vec{\Sigma}_f, \vec{\varphi}) r dr} + h_j^{(i)} \right] ; \\
 \mathcal{H}_{n+j}^{(i)} &= K_{r,i} R_{a.з.} \left[- \sum_{i=j}^{\pi} R_i \frac{(\vec{\Sigma}_{f,i} - \vec{\Sigma}_{f,i+1}, \vec{\varphi}(R_i))}{\int_0^{R_{a.з.}} (\vec{\Sigma}_f, \vec{\varphi}) r dr} + h_{n+j}^{(i)} \right] ,
 \end{aligned} \right\} (3.60)$$

где

$$\eta_j^{(i)} = \begin{cases} 1, & j=i \\ 0, & j \neq i \end{cases} \quad \begin{matrix} i=1, \dots, n \\ j=1, 2, \dots, n \end{matrix}$$

$$\begin{aligned}
 h_j^{(i)} &= \int_{\Delta R_j} \left\{ (\vec{J}_i^+, [3\hat{\sigma}_{tr}^{(a)} - 3\hat{\sigma}_{tr}^{(b)}] \vec{J}) - (\vec{\varphi}_i^+, [\hat{\sigma}_{ad}^{(a)} - \hat{\sigma}_{ad}^{(b)}] \vec{\varphi}) + \frac{(\vec{J}, \vec{\varphi}_i^+)}{\kappa_{\text{эф}}} (\nu \vec{\sigma}_f^{(a)} - \nu \hat{\sigma}_f^{(b)}, \vec{\varphi}) \right\} ; \\
 h_{n+j}^{(i)} &= \sum_{q=j}^{\pi} R_q \left\{ (\vec{J}_i^+(R_q), [3\hat{\Sigma}_{tr,q} - 3\hat{\Sigma}_{tr,q+1}] \vec{J}(R_q)) - (\vec{\varphi}_i^+(R_q), [\hat{\Sigma}_{ad,q} - \hat{\Sigma}_{ad,q+1}] \vec{\varphi}(R_q)) \right. \\
 &\quad \left. \times \frac{1}{\kappa_{\text{эф}}} (\vec{J}, \vec{\varphi}_i^+(R_q)) (\nu \vec{\Sigma}_{f,q} - \nu \vec{\Sigma}_{f,q+1}, \vec{\varphi}(R_q)) \right\} , \quad i=0, 1, \dots, n.
 \end{aligned}$$

Результаты решения. Численное решение задачи получено с помощью комплекса программ РОКБАР [9], в котором реализован первый алгоритм метода линеаризации. Расчеты проводились для охлаждаемого натрием быстрого реактора, состоящего из двух активных зон с топливом в виде смеси $UO_2 + PuO_2$ (воспроизводящий материал — ^{238}U , а делящийся материал — ^{239}Pu). Активные зоны окружены торцевыми и боковым экранами с воспроизводящим материалом в виде UO_2 . Потоки $\varphi^{(k)}(r)$ и ценности нейтронов $\varphi_i^{+(k)}(r)$ находились в двухгрупповом диффузионном приближении из решения конечно-разностных аналогов уравнений (3.43), (3.44), (3.50), (3.55), в которых сечение поглощения $\Sigma_a^{(k)}$ было заменено на эффективное сечение $\tilde{\Sigma}_a^{(k)}$:

$$\tilde{\Sigma}_a^{(k)} = \Sigma_a^{(k)} + \frac{\kappa^2}{3 \Sigma_{tr}^{(k)}} ,$$

где $\alpha^2 = \left(\frac{\lambda}{H_{a.z} + 2\delta_z} \right)^2$, $H_{a.z}$ — высота активной зоны, δ_z — эффективная добавка

горцевого экрана. Тем самым учитывалась утечка нейтронов в торцевые экраны.

Отметим, что использование $\sum_a^{(k)}$ вместо $\sum_a^{(k)}$ не меняет выражения для составляющих $h_{n+j}^{(i)}$ ($j=1, \dots, n$) коэффициентов чувствительности. Вместе с тем, в правой части формулы для составляющих $h_j^{(i)}$ ($j=1, \dots, n$) появляется

дополнительное слагаемое $(\vec{\varphi}_i^+, \frac{1}{3} \alpha^2 (\hat{\Sigma}_{tr}^{-1})^2 [\hat{b}_{tr}^{(d)} - \hat{b}_{tr}^{(b)}] \vec{\varphi})$. . Кроме

того, переход к конечно-разностным аналогам уравнений (3.43), (3.44) требует, строго говоря, построения сопряженных к ним уравнений. Однако можно показать, что сформулированные таким образом уравнения будут совпадать с теми, которые получатся из (3.50) или (3.55), если к ним применить ту же технику, которая использовалась при получении конечно-разностных аналогов уравнений (3.43), (3.44). В качестве управлений принимались параметры $u_1 = \rho_1^{(d)} / \rho_T$; $u_2 = \rho_2^{(d)} / \rho_1^{(d)}$, $u_3 = \Delta R_1 / R_{a.z}$, $u_4 = \Delta R_2 / R_{a.z}$. и вводилось дополнительное ограничение: $0,5 \leq u_3 \leq 1$. Двухгрупповые микросечения элементов, $R_{a.z}$, $H_{a.z}$ и δ_z были заданы и не менялись в процессе оптимизации.

Результаты итерационного поиска оптимума приведены на рис. 4, где сплошные линии построены из условия $K_r = const$, а пунктирные линии соответствуют спуску из двух разных точек: а) $u_2^{(0)} = 1,0$, $u_3^{(0)} = 0,5$; б) $u_2^{(0)} = 2,4$, $u_3^{(0)} = 0,8$. В обоих случаях на каждой итерации автоматически выбиралось такое направление спуска, которое способствует тому, чтобы наискорейшим образом попасть на гребень „ $a - \delta$ “, а затем вдоль гребня спуститься в оптимальную точку: $u_2^* = 1,47$, $u_3^* = 0,75$, соответствующую наименьшему значению $K_r^* = 1,19$. При этом коэффициенты неравномерности $K_{r,i} = K_r^*$ ($i=1, 2$). Последнее соотношение означает, что в рассмотренном реакторе при выравнивании поля тепловыделения обогащением топлива наименьшая неравномерность поля имеет место в том случае, когда максимальные тепловыделения $\max_{r \in \Delta R_i} q(r)$ в зонах равны.

4. УЧЕТ НЕОПРЕДЕЛЕННОСТИ ИСХОДНОЙ ИНФОРМАЦИИ

Большая часть исходной информации, используемой при проектировании ядерных реакторов, известна приближенно. Имеющиеся неточности в исходных данных обусловлены разными причинами. Это могут быть:

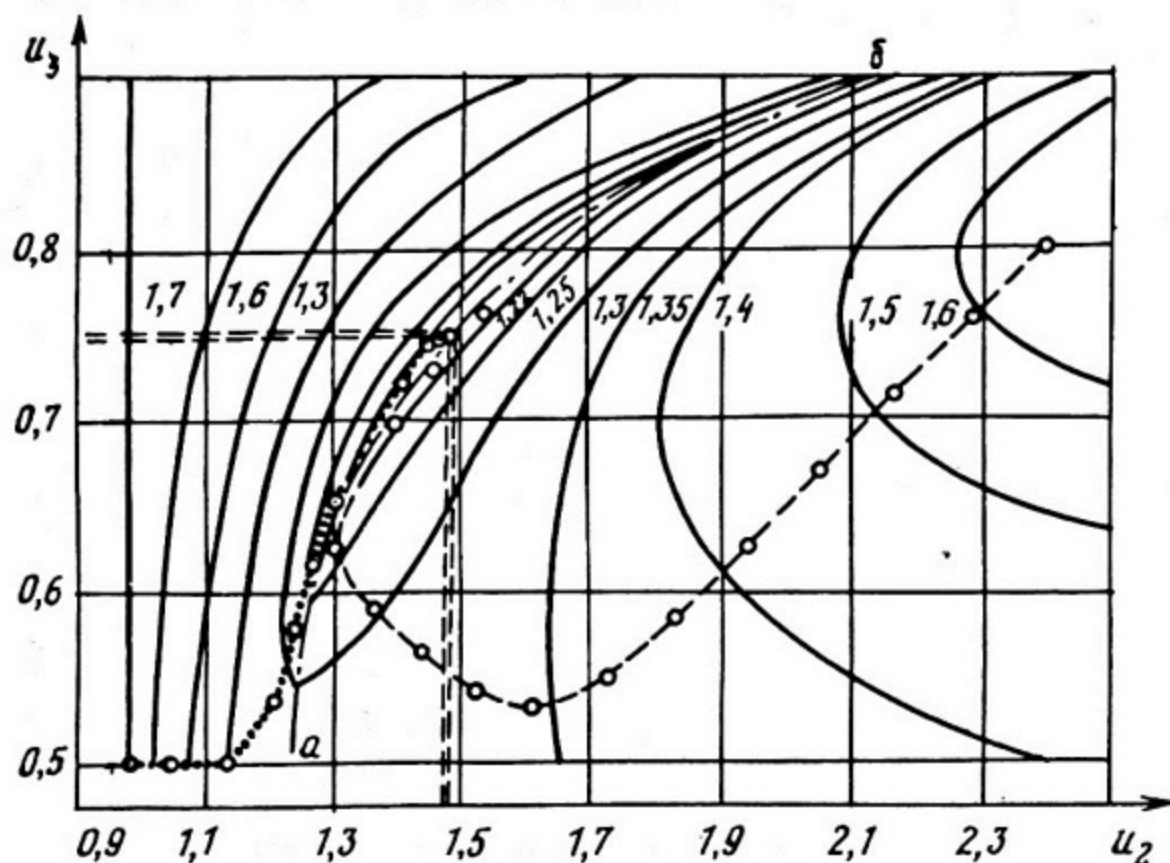


Рис. 4. Линии уровня $K_r = const$ и последовательный спуск из двух разных точек в задаче о минимуме коэффициента неравномерности K_r

экспериментальные ошибки, если говорить о данных, определяющих свойства материалов (ядерно-физические константы, теплофизические и прочностные свойства материалов и др.);

технологические неточности, связанные с изготовлением отдельных элементов конструкции, которые считаются в ряде случаев заданными и не выбираются в процессе проектирования (плотность топлива, диаметр и толщина оболочки твэлов и т.п.);

неопределенности в технико-экономических характеристиках реактора, АЭС и топливного цикла (стоимостные характеристики материалов и отдельных видов работ, параметры внешнего топливного цикла, режим эксплуатации и др.).

При создании САПР ЯЭУ важно не только знать, какие из исходных данных точно не известны, но и уметь математически описать имеющиеся неточности. С этой точки зрения приближенно известную исходную информацию можно разделить на две группы: 1) вероятно-определенные

данные, в том числе параметры, разброс значений которых поддается вероятностному описанию (например, известны функции распределения вероятностей); 2) неопределенные данные, для которых неточно известны или совсем отсутствуют вероятностные характеристики, а известны лишь диапазоны изменения.

Желание учесть влияние имеющихся в исходных данных неточностей на выбор оптимальных характеристик реактора приводит к рассмотрению задач в постановке (1.6) – (1.9). Теперь (по сравнению с задачей (1.1) – (1.5)) не будет существовать единственного оптимального значения критерия F_0 и управления \vec{u} . Каждому фиксированному набору значений исходных параметров $\vec{v} = \vec{v}'$ из области их возможных значений Ω будут соответствовать разные оптимальные значения $F_0 = F_0'$ и $\vec{u} = \vec{u}'$, т.е. можно говорить об отображении (путем решения соответствующих задач оптимального проектирования) области Ω на область неопределенности \mathcal{F} значений критерия F_0 и область U возможных оптимальных управлений \vec{u} .

Задачу проектирования в условиях неопределенности исходной информации можно было бы считать решенной, если бы были найдены области \mathcal{F} и U . Однако их построение, связанное с решением задач вида (1.1) – (1.5), при всех $\vec{v} \in \Omega$ практически невозможно и приходится вводить различные предположения и критерии отбора наиболее представительных значений \vec{v} . Некоторые из таких способов рассмотрим в данном разделе пособия.

4.1. Исследование устойчивости решения

Исследование устойчивости решения задачи оптимального проектирования связано с получением ответа на вопрос об изменении оптимума при малых отклонениях исходных данных от первоначально заданных (номинальных) значений \vec{v}^0 . Проведение подобного исследования важно с точки зрения выработки рекомендаций по уточнению, в первую очередь, тех из исходных данных, с изменением которых оптимальное значение F_0 сильно меняется. Степень влияния q -го исходного параметра на значение F_0 можно оценить, зная соответствующий коэффициент чувствительности

$$\alpha_q = \lim_{\delta v_q \rightarrow 0} \frac{\Delta F_0}{\delta v_q}, \quad (4.1)$$

где ΔF_0 – отклонение оптимального значения F_0 , вызванное изменением q -го параметра на величину δv_q .

Такие коэффициенты α_q , вообще говоря, разные при стремлении δv_q к нулю справа ($\alpha_q = \alpha_q^+$) или слева ($\alpha_q = \alpha_q^-$), можно было бы оценить,

сравнивая оптимальные значения F_0^* , полученные из решения задач вида (1.6) – (1.9) при v_q^0 и $v_q^0 + \delta v_q$. Однако такой способ потребовал бы больших затрат расчетного времени, поскольку обычно количество исходных данных велико. Ниже изложен другой подход, опирающийся на использование необходимых условий оптимальности и формализма теории малых возмущений.

Рассмотрим сначала случай, когда для каждого q -го параметра существуют единственные значения коэффициентов α_q (т.е. когда $\alpha_q^+ = \alpha_q^-$). Как будет видно из дальнейших рассуждений, такая ситуация в задачах с достаточно гладкими функционалами F_i возникает, когда решение задачи (1.6) – (1.9) при номинальных значениях v_q^0 соответствует изолированному оптимуму. Не ограничивая общности рассуждений, будем считать, что неравенства вида (1.6) на управляющие параметры u_i включены в число условий (1.8), общее количество которых пусть по-прежнему равно m . Для разностей вида $F_i(\vec{\varphi}, \vec{u}, \vec{v}) - F_i^{\text{дон}}(\vec{v})$ будем использовать прежнее обозначение $F_i(\vec{\varphi}, \vec{u}, \vec{v})$, так что условия (1.8) запишутся в виде:

$$F_i(\vec{\varphi}, \vec{u}, \vec{v}) \leq 0, \quad i = 1, 2, \dots, l; \quad (4.2)$$

$$F_i(\vec{\varphi}, \vec{u}, \vec{v}) = 0, \quad i = l+1, \dots, m. \quad (4.3)$$

Пусть при номинальном значении v^0 оптимальное управление \vec{u}^* и переменные состояния $\vec{\varphi}^*(x)$ такие, что:

$$F_i(\vec{\varphi}^*, \vec{u}^*, \vec{v}^0) < 0, \quad i = 1, 2, \dots, l_0;$$

$$F_i(\vec{\varphi}^*, \vec{u}^*, \vec{v}^0) = 0, \quad i = l_0+1, \dots, l, l+1, \dots, m.$$

Тогда необходимые условия оптимальности (3.26), (3.27) запишутся в виде:

$$\left. \begin{aligned} \mathcal{H}_j^{(0)} + \sum_{i=1}^m \lambda_i^0 \mathcal{H}_j^{(i)} &= 0, \quad j = 1, 2, \dots, n; \\ \lambda_i^0 F_i(\vec{w}^*) &= 0, \quad i = 1, 2, \dots, m; \\ \lambda_i^0 &\geq 0, \quad i = l_0+1, \dots, l, \end{aligned} \right\} \quad (4.4)$$

где через \vec{w}^* обозначена совокупность переменных $\vec{\varphi}^*(x), \vec{u}^*, \vec{v}^0$, а коэффициенты чувствительности $\mathcal{H}_j^{(i)}$ найдены с помощью известных соотношений теории малых возмущений в точке \vec{w}^* .

Предположим, что q -й компонент вектора \vec{v}^0 получил приращение δv_q . Обозначим через \vec{v}' вектор с компонентами: $\{v_1^0, \dots, v_{q-1}^0, v_q^0 + \delta v_q, \dots, v_n^0\}$.

$v_{q+1}^0, \dots, v_t^0 \}$, а через $\vec{u}' = \vec{u}^* + \delta \vec{u}$, $\vec{\varphi}'(x) = \vec{\varphi}^*(x) + \delta \vec{\varphi}(x)$, $\vec{F}_i' = \vec{F}_i(\vec{\omega}^*) + \Delta F_i$ — решение задачи (1.6) — (1.9) при $\vec{v} = \vec{v}'$. Этому решению соответствуют множители Лагранжа $\lambda_i' = \lambda_i^0 + \delta \lambda_i$, имеющие такие значения, что:

$$\lambda_i' F_i' = 0, \quad i = 1, 2, \dots, m. \quad (4.5)$$

Не исключено, что некоторые из ограничений (4.2), (4.3), которые при $\vec{v} = \vec{v}^0$ выполнялись как строгие равенства, при $\vec{v} = \vec{v}'$ примут вид строгих неравенств и наоборот. Так как в условиях изолированного оптимума вариации δF_i , $\delta \lambda_i$ и $\delta \vec{u}$ непрерывно зависят от δv_q [22], то с точностью до членов первого порядка малости имеем:

$$\delta F_i = h_q^{(i)} \delta v_q + \sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j, \quad i = 0, 1, \dots, m; \quad (4.6)$$

$$\lambda_i^0 [F_i(\vec{\omega}^*) + h_q^{(i)} \delta v_q + \sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j] + \delta \lambda_i F_i(\vec{\omega}^*) = 0, \quad i = 1, \dots, m, \quad (4.7)$$

где $h_q^{(i)}$ — коэффициенты чувствительности функционала F_i к q -му компоненту вектора \vec{v} , вычисленные в точке $\vec{\omega}^*$. Используя равенства (4.4) — (4.7), получим:

$$\delta F_0 = h_q^{(0)} \delta v_q + \sum_{i=1}^m \lambda_i^0 h_q^{(i)} \delta v_q + \sum_{i=1}^m \delta \lambda_i F_i(\vec{\omega}^*). \quad (4.8)$$

Все множество индексов $i = 1, 2, \dots, l$ разобьем на четыре подмножества I_0, I_1, I_2, I_3 так, как это указано в табл. 1.

Таблица 1

Множество индексов i	$i = 1, 2, \dots, l_0$		$i = l_0 + 1, \dots, l$	
	I_0	I_1	I_2	I_3
Вид ограничений, когда $\vec{v} = \vec{v}^0$	<	<	=	=
Вид ограничений, когда $\vec{v} = \vec{v}'$	<	=	=	<
Значения λ_i^0	0	0	≥ 0	≥ 0
Значения λ_i'	0	≥ 0	≥ 0	0

Тогда

$$\sum_{i=1}^m \delta \lambda_i F_i(\vec{w}^*) = \sum_{i=1}^l \delta \lambda_i F_i(\vec{w}^*) = \sum_{i \in I_1} \delta \lambda_i F_i(\vec{w}^*) = - \sum_{i \in I_1} \delta \lambda_i \delta F_i, \quad (4.9)$$

так как для $i \in I_0 \cup I_2 \cup I_3$ либо $\delta \lambda_i = 0$, либо $F_i(\vec{w}^*) = 0$, а при $i \in I_1$: $F_i(\vec{w}^*) = F_i' - \delta F_i = -\delta F_i$. Таким образом, последнее слагаемое в равенстве (4.8) имеет более высокий порядок малости, чем первые два.

Поделив обе части равенства (4.8) на δv_q^* и перейдя к пределу при $\delta v_q^* \rightarrow 0$, получим:

$$\alpha_q = h_q^{(0)} + \sum_{i=1}^m \lambda_i^0 h_q^{(i)}, \quad (4.10)$$

где учтено, что $\lambda_i^0 = 0$ при $i = 1, 2, \dots, l_0$. Отсюда, в частности, можно выяснить смысл множителей Лагранжа. Действительно, пусть в качестве q -го параметра берется допустимое значение $F_{i_0}^{don}$ для какого-либо функционала F_{i_0} . Тогда $h_q^{(0)} = 0$, $h_q^{(i)} = 0$ для всех $i \neq i_0$, $h_q^{(i_0)} = -1$ и, следовательно, $\alpha_q = -\lambda_{i_0}^0$.

Таким образом, множитель Лагранжа $\lambda_{i_0}^0$ численно равен взятому с противоположным знаком изменению оптимального значения критерия F_0 , вызванному единичным изменением ($\delta F_{i_0}^{don} = 1$) допустимого значения $F_{i_0}^{don}$ в условии $F_{i_0} \leq F_{i_0}^{don}$. Другими словами: $\lambda_{i_0}^0 = -\frac{\partial(\min F_0)}{\partial F_{i_0}^{don}}$. Оче-

видно, что соответствующая индексу i_0 двойственная переменная в (3.31)–(3.33) имеет тот же смысл.

Формула (4.10)* позволяет найти коэффициенты α_q , если известны множители Лагранжа λ_i^0 . Эти множители можно получить либо из условий (4.4), в которых для $(m - l_0)$ неотрицательных неизвестных λ_i^0 имеется n линейных неоднородных уравнений, либо как решение следующей двойственной задачи линейного программирования (см. п. 3.3):

$$\max q = -\min(-q), \quad q = \sum_{i=1}^m F_i(\vec{w}^*) y_i, \quad (4.11)$$

$$\left. \begin{aligned} \sum_{i=1}^m \mathcal{H}_j^{(i)} y_i &= -\mathcal{H}_j^{(0)}, \quad j = 1, 2, \dots, n; \\ y_i &\geq 0, \quad i = 1, 2, \dots, l, \end{aligned} \right\} \quad (4.12)$$

* Формула (4.10) была использована в работе [26] при оптимизации характеристик быстрых реакторов.

где коэффициенты $\mathcal{H}_i^{(i)}$ определены в точке \vec{w}^* и предполагаются известными. Здесь для двойственных переменных придерживаемся принятых в п. 3.3 обозначений, а задача (4.11), (4.12) сформулирована как задача поиска $\min(-q)$.

Остановимся на последнем способе, использование которого не вызывает трудностей, когда поиск оптимума осуществляется с помощью методов, базирующихся на последовательном решении задач линейного программирования.

При решении задачи (4.11), (4.12) возможны два случая: а) существует единственное неотрицательное решение $y_i = y_i^*$, $i = 1, 2, \dots, m$; б) существует бесконечное множество решений, образующее выпуклое многогранное множество G .

Случай, когда двойственная задача не имеет решений, исключается предположением о том, что \vec{w}^* — точка оптимума (и тогда обязаны существовать множители Лагранжа).

Пусть множество G является ограниченным. Тогда любой элемент $\vec{y}^* = \{y_1^*, y_2^*, \dots, y_m^*\}$ этого множества (в соответствии с известной теоремой о представлении выпуклого многогранного множества [19]) может быть получен как линейная комбинация:

$$\vec{y}^* = \sum_{q=1}^N \alpha_q \vec{y}^{(q)}, \quad (4.13)$$

где $\sum_{q=1}^N \alpha_q = 1$, $\alpha_q \geq 0$; $\vec{y}^{(q)} = \{y_1^{(q)}, y_2^{(q)}, \dots, y_m^{(q)}\}$ — вектора, определяющие вершины множества G ; N — общее число вершин множества G .

Для получения же вершин выпуклого многогранника G (в которых $\min(-q) = 0$) можно воспользоваться любым конечным методом линейного программирования.

Пусть выбран симплекс-метод и получен по схеме, описанной в приложении 2, один из оптимальных опорных планов $\{y_1^{(l)}, y_2^{(l)}, \dots, y_m^{(l)}\}$ задачи (4.11), (4.12), состоящий из J_C свободных и J_B базисных переменных, для которого линейная форма $q = -\sum_{l \in J_C} \gamma_l y_l^{(l)} = 0$ (так как $y_l^{(l)} = 0$ при $l \in J_C$).

Как известно, это имеет место, когда $\gamma_l \leq 0$ при всех $l \in J_C$. Поэтому возможны два случая: а) $\gamma_l < 0$ при всех $l \in J_C$; тогда множество G состоит из одного элемента; б) среди γ_l , $l \in J_C$ существует один или несколько нулевых коэффициентов; тогда возможно существование бесконечного числа элементов, образующих множество G .

Предположим, что $\gamma_s = 0$, где $s \in J_c$, и попытаемся перейти к новому опорному оптимальному плану, изменяя лишь компоненту $y_s^{(1)}$ и оставляя значения остальных небазисных компонент равными нулю. Тогда будут иметь место равенства (см. приложение 2, (П2.6)):

$$y_i^{(1)} = \beta_i - \alpha_{i,s} y_s^{(1)}, \quad i \in J_\delta. \quad (4.14)$$

В свою очередь здесь возможны два случая.

1. Среди $i \in J_\delta$ существуют такие i , для которых одновременно $\alpha_{i,s} > 0, \beta_i > 0$. При этом, выбирая

$$y_s^{(2)} = \min_{\substack{\alpha_{i,s} > 0 \\ \beta_i > 0}} \frac{\beta_i}{\alpha_{i,s}} = \frac{\beta_k}{\alpha_{k,s}},$$

построим новый опорный план $\{y_1^{(2)}, y_2^{(2)}, \dots, y_m^{(2)}\}$, доставляющий линейной форме q нулевое значение: $y_i^{(2)} = \beta_i - \alpha_{i,s} y_s^{(2)}$, $i \in J_\delta, i \neq k$; $y_k^{(2)} = 0$; $y_i^{(2)} = 0, i \in J_c, i \neq s$; $y_s^{(2)} = \frac{\beta_k}{\alpha_{k,s}}$. Дальнейшие выкладки

связаны с преобразованием ограничений и ничем не отличаются от тех, которые описаны в приложении 2.

2. Среди $i \in J_\delta$ нет таких i , для которых одновременно выполнялись бы неравенства $\alpha_{i,s} > 0, \beta_i > 0$. В этом случае необходимо вернуться к выбору какого-либо другого коэффициента $\gamma_{s'} = 0$ из числа $\gamma_l, l \in J_c$, пытаясь построить новый оптимальный опорный план за счет изменения компоненты $y_{s'}^{(1)}$. Таким образом можно найти все оптимальные опорные планы $\vec{y}^{(q)}, q = 1, 2, \dots, N$. Поскольку любой опорный план определяет вершину допустимой области (4.12), то найденные векторы $\vec{y}^{(q)}$ с компонентами $y_i^{(q)}, i = 1, 2, \dots, m$, позволяют, используя представление (4.13), построить множество G .

Как показано в работе [23], существование бесконечного числа множителей Лагранжа приводит к появлению различных значений α_q^+ и α_q^- для коэффициентов чувствительности критерия оптимальности F_0 . В общем случае, когда один и тот же параметр γ_q влияет на несколько функционалов $F_i, i = l_0 + 1, \dots, m$ (ограничения на которые в точке оптимума при $\vec{v} = \vec{v}^0$ принимают вид строгих равенств), эти коэффициенты находятся по формулам:

$$\left. \begin{aligned} \alpha_q^+ &= h_q^{(0)} + \max_{\vec{y}^{(p)} \in G} \sum_{i=1}^m y_i^{(p)} h_q^{(i)}; \\ \alpha_q^- &= h_q^{(0)} + \min_{\vec{y}^{(p)} \in G} \sum_{i=1}^m y_i^{(p)} h_q^{(i)}, \end{aligned} \right\} \quad (4.15)$$

где \max и \min берутся по всем оптимальным опорным планам $\bar{y}^{(p)} = \{y_i^{(p)}\}_{i=1,2,\dots,m}$
 $\bar{y}^{(p)} \in G$ $\bar{y}^{(p)} \in B$
 задачи (4.11), (4.12).

Равенства (4.15) получаются на основе теоремы 5.1 из работы [19] (с. 208), принимая во внимание смысл двойственных переменных и учитывая ограниченность множества B .

Изложенный выше способ нахождения коэффициентов чувствительности $\bar{\alpha}_q^+$ и $\bar{\alpha}_q^-$ проиллюстрируем на следующем примере. Рассмотрим задачу нелинейного программирования:

$$\left. \begin{aligned} \min F_0(u_1, u_2, v), \quad \text{где: } F_0 &= (u_1 - v^2)^2 + (u_2 - 2)^2; \\ F_1(u_1, u_2, v) &= u_1^2 + u_2^2 - v^2 \leq 0; \\ F_2(u_1, u_2, v) &= -3u_1 - 2u_2 + 1 \leq 0; \\ -\infty < u_1 &\leq \frac{v}{\sqrt{2}}; \quad -\infty < u_2 &\leq 1 \end{aligned} \right\} \quad (4.16)$$

при номинальном значении параметра $v^0 = \sqrt{2}$.

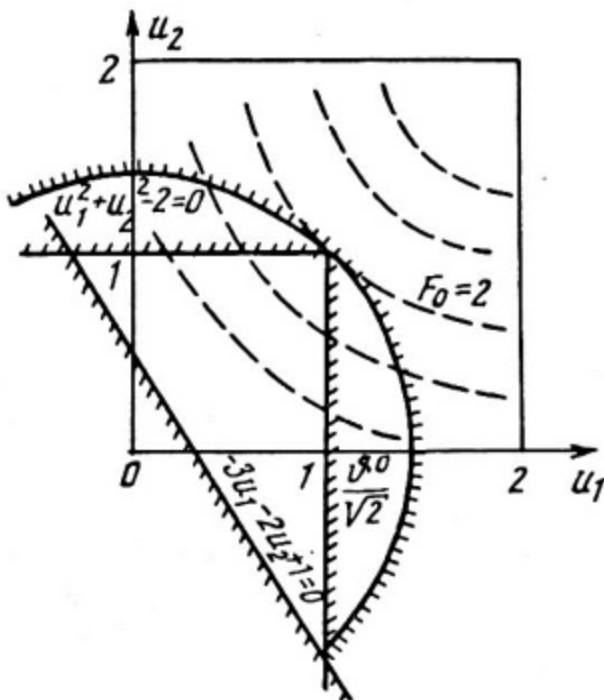


Рис. 5. Допустимая область и линии уровня (пунктир) функции $F_0(u_1, u_2)$ в задаче (4.16)

Область допустимых значений и линии уровня функции F_0 на плоскости переменных $u_1 - u_2$ представлены на рис. 5, из которого видно,

что при $\delta = \sqrt{2}$ $\min F_0 = 2$, $u_1^* = 1$, $u_2^* = 1$; причем F_7 , u_1 и u_2 достигают своих предельных значений.

В окрестности точки оптимума прямая (для вариаций δu_1 , δu_2) и двойственная (для переменных y_i , $i=1,2,\dots,6$) задачи запишутся:

$$\begin{aligned} \text{прямая задача} \\ \min f, \quad f = -2\delta u_1 - 2\delta u_2; \\ 2\delta u_1 + 2\delta u_2 \leq 0; \\ -3\delta u_1 - 2\delta u_2 \leq 4; \\ \delta u_1 \leq 0; \\ \delta u_2 \leq 0; \\ -\delta u_1 \leq 0,1; \\ -\delta u_2 \leq 0,1. \end{aligned}$$

$$\begin{aligned} \text{двойственная задача} \\ \max j = -\min(-q), \quad \text{где:} \\ q = -4y_2 - 0,1y_5 - 0,1y_6; \\ 2y_1 - 3y_2 + y_3 - y_5 = 2; \\ 2y_1 - 2y_2 + y_4 - y_6 = 2; \\ y_i \geq 0, \quad i=1,2,\dots,6. \end{aligned}$$

Здесь для определенности принято, что $S_1^- = -0,1$, $S_2^- = -0,1$.

Если условия типа равенства в двойственной задаче переписать в виде:

$$y_3 = 2 - (2y_1 - 3y_2 - y_5), \quad (4.17)$$

$$y_4 = 2 - (2y_1 - 2y_2 - y_6), \quad (4.18)$$

то легко получить исходный опорный план $\vec{y}^{(1)}$ с компонентами:

$$\left. \begin{aligned} y_3^{(1)} = y_4^{(1)} = 2 & \text{ — базисные переменные;} \\ y_1^{(1)} = y_2^{(1)} = y_5^{(1)} = y_6^{(1)} = 0 & \text{ — свободные переменные,} \end{aligned} \right\} \quad (4.19)$$

для которого линейная форма $q=0$ и представима в виде:

$$-q = 0 - (0 \cdot y_1 - 4y_2 - 0,1y_5 - 0,1y_6). \quad (4.20)$$

Поскольку в (4.20) все коэффициенты β_l при соответствующих свободных переменных неположительны, то опорный план $\vec{y}^{(1)}$ (4.19) является оптимальным. Одновременно из равенства нулю коэффициента β_1 при y_1 следует, что можно попытаться найти другой оптимальный план, переводя y_1 в разряд базисных переменных и по-прежнему считая $y_2 = y_5 = y_6 = 0$. Тогда:

$$y_3 = 2 - 2y_1; \quad y_4 = 2 - 2y_1. \quad (4.21)$$

Сопоставляя равенства (4.21) и (4.14), можно заключить, что $\beta_3 = \beta_4 = 2 > 0$,

$\alpha_{3,1} = \alpha_{4,1} = 2 > 0$. Поэтому $y_1^{(2)} = \min_{l=3,4} \frac{\beta_l}{\alpha_{l,1}} = 1$, и приходим к новому оптимальному плану $\vec{y}^{(2)}$ с компонентами:

$$\left. \begin{aligned} y_1^{(2)} = 1, y_4^{(2)} = 0 & \text{ — базисные переменные;} \\ y_2^{(2)} = y_3^{(2)} = y_5^{(2)} = y_6^{(2)} = 0 & \text{ — свободные переменные.} \end{aligned} \right\} \quad (4.22)$$

Выражая теперь y_1 из (4.18) через остальные неизвестные и подставляя в (4.17) и линейную форму (4.20), получим:

$$-q = 0 - (-4y_2 + 0y_3 - 0,1y_5 - 0,1y_6);$$

$$y_1 = 1 - (-1,5y_2 + 0,5y_3 - 0,5y_5);$$

$$y_4 = 0 - (y_2 - y_3 + y_5 - y_6).$$

Дальнейшие рассуждения не приведут к появлению других, отличных от $\vec{y}^{(1)}$ и $\vec{y}^{(2)}$, оптимальных планов. Таким образом, в рассматриваемой задаче (4.16) множество G множителей Лагранжа представляет собой отрезок прямой, заключенный между точками $\vec{y}^{(1)}$ и $\vec{y}^{(2)}$.

Для расчета коэффициентов α^+ и α^- (индекс q для простоты опустим) следует воспользоваться общими формулами (4.15), так как параметр v^* влияет на значения двух функций, описывающих границу допустимой области в задаче (4.16). Учитывая, что в точке оптимума при $v^* = \sqrt{2}$, $h^{(0)} = -4v^*(c_7^* - v^{*2}) = 4\sqrt{2}$; $h^{(1)} = -2v^* = -2\sqrt{2}$; $h^{(3)} = -\frac{1}{\sqrt{2}}$; $h^{(2)} = h^{(4)} = h^{(5)} = h^{(6)} = 0$, и принимая во внимание значения (4.19) и (4.22) для компонент $y_i^{(1)}$ и $y_i^{(2)}$ оптимальных планов, получим:

$$\alpha^+ = 4\sqrt{2} + \max(-\sqrt{2}, -2\sqrt{2}) = 3\sqrt{2};$$

$$\alpha^- = 4\sqrt{2} + \min(-\sqrt{2}, -2\sqrt{2}) = 2\sqrt{2}.$$

4.2. Методы стохастического программирования

Методы стохастического программирования [24] применяются к задачам оптимального проектирования с вероятностно-определенными исходными данными. Можно выделить два направления в развитии этих методов. Первое из них связано с приведением задачи со случайными параметрами к детерминистической задаче. Второе направление основано на использовании вместо точных значений градиентов функций стохастических квазиградиентов, являющихся статистическими оценками этих векторов.

Детерминированные аналоги для задачи (1.6) – (1.9) можно построить, когда удастся вычислить либо математические ожидания $MF_i(\vec{\varphi}, \vec{u}, \vec{v})$, $MF_i^{\text{доп}}(\vec{v})$, $Mu_j^{\text{макс}}(\vec{v})$, $Mu_j^{\text{мин}}(\vec{v})$, присутствующих в задаче функционалов $F_i(\vec{\varphi}, \vec{u}, \vec{v})$ и функций $F_i^{\text{доп}}(\vec{v})$, $u_j^{\text{макс}}(\vec{v})$, $u_j^{\text{мин}}(\vec{v})$, либо найти вероятности $P\{F_i(\vec{\varphi}, \vec{u}, \vec{v}) \leq F_i^{\text{доп}}(\vec{v})\} \geq \gamma_i$, $P\{u_j \leq u_j^{\text{макс}}(\vec{v})\} \geq \alpha_j$, $P\{u_j \geq u_j^{\text{мин}}(\vec{v})\} \geq \beta_j$ ($i=1,2,\dots,m$; $j=1,2,\dots,n$) того, что каждое из ограничений вида (1.8), (1.6) будет выполнено с вероятностью, большей некоторых наперед заданных значений γ_i , α_j и β_j соответственно. В этом случае вместо задачи (1.6) – (1.9) рассматривают:

либо задачу вида

$$\left. \begin{aligned} \min MF_0(\vec{\varphi}, \vec{u}, \vec{v}) ; \\ MF_i(\vec{\varphi}, \vec{u}, \vec{v}) \leq MF_i^{\text{доп}}(\vec{v}), i=1,2,\dots,m ; \\ Mu_j^{\text{мин}}(\vec{v}) \leq u_j \leq Mu_j^{\text{макс}}(\vec{v}), j=1,2,\dots,n, \end{aligned} \right\} \quad (4.23)$$

либо задачу вида

$$\left. \begin{aligned} \max P\{F_0(\vec{\varphi}, \vec{u}, \vec{v}) \leq f_0\} ; \\ P\{F_i(\vec{\varphi}, \vec{u}, \vec{v}) \leq F_i^{\text{доп}}(\vec{v})\} \geq \gamma_i, i=1,2,\dots,m ; \\ P\{u_j \leq u_j^{\text{макс}}(\vec{v})\} \geq \alpha_j, j=1,2,\dots,n ; \\ P\{u_j \geq u_j^{\text{мин}}(\vec{v})\} \geq \beta_j, j=1,2,\dots,n, \end{aligned} \right\} \quad (4.24)$$

где f_0 – заданное число, а переменные состояния $\vec{\varphi}(x)$ связаны с искомым управлением \vec{u} и случайными параметрами \vec{v} уравнениями (1.9).

Каждая из этих задач ничем не отличается от ранее рассмотренных детерминистических задач и может решаться с помощью методов, изложенных в разделе 3.

К сожалению, переход к задачам (4.23) или (4.24) возможен лишь в простейших случаях. Например, когда ограничения (1.8) имеют вид

$$F_i(\vec{\varphi}, \vec{u}) \leq F_i^{\text{доп}}(\vec{v}), i=1,2,\dots,m, \quad (4.25)$$

или когда каждый из функционалов

$$F_i(\vec{\varphi}, \vec{u}, \vec{v}) = \sum_{l=1}^k a_{il}(\vec{v}) f_l(\vec{\varphi}, \vec{u}) + b_i(\vec{v}), i=0,1,\dots,m,$$

причем в уравнения (1.9) для переменных состояния $\vec{\psi}(x)$ не входят случайные параметры \vec{v} . (Решение некоторых задач с ограничениями вида (4.25) приведены в работе [13].)

Если существует математическое ожидание $\vec{v}^0 = M\vec{v}$ и диапазон изменения параметров \vec{v} невелик, то можно вначале при $\vec{v} = \vec{v}^0$ решить задачу вида (1.6) — (1.9), найти оптимальные значения критерия $F_0 = F_0^*$ и управления $\vec{u} = \vec{u}^*$ и рассчитать по (4.15) коэффициенты чувствительности α_q^+ и α_q^- . Пусть $\alpha_q^+ = \alpha_q^- = \alpha_q$. Тогда в узком диапазоне изменения компонент v_q вектора \vec{v} справедливо линейное соотношение

$$F_0 = F_0^* + \sum_{q=1}^t \alpha_q (v_q - v_q^0), \quad (4.26)$$

используя которое можно перейти к оценке неопределенности критерия F_0 .

Пусть известны среднеквадратические погрешности σ_q величин v_q и коэффициенты корреляции r_{qj} любой пары величин v_q и v_j . Учитывая линейную связь (4.26) и пользуясь известными соотношениями теории вероятностей для числовых характеристик случайных величин [25], имеем:

$$\sigma_0^* = \left[\sum_{q=1}^t \alpha_q^2 \sigma_q^2 + 2 \sum_{q < j} \alpha_q \alpha_j r_{qj} \sigma_q \sigma_j \right]^{\frac{1}{2}}, \quad (4.27)$$

где σ_0^* — среднеквадратическая погрешность критерия F_0 , а суммирование $\sum_{q < j}$ распространяется на все различные попарные комбинации величин v_1, v_2, \dots, v_t .

Наряду с оценкой неопределенности критерия F_0 по формуле (4.27) важное практическое значение имеет обратная задача, в которой по величине σ_0^* оцениваются допустимые погрешности σ_q исходных данных. Для этого необходимо задаться тем или иным значением σ_0^* . Естественно выбрать σ_0^* , исходя из достоверности сравнения оптимального значения F_0^* с допустимой величиной $F_{0, \text{доп}}$, которая, вообще говоря, тоже может быть известна с некоторой погрешностью.

Введем в рассмотрение среднюю вероятность P того, что $F_0^* \leq F_{0, \text{доп}}$:

$$P = \int_{-\infty}^{+\infty} f_0(x) \int_x^{\infty} f_{\text{доп}}(x') dx' dx = \int_{-\infty}^{+\infty} f_{\text{доп}}(x) \int_{-\infty}^x f_0(x') dx' dx, \quad (4.28)$$

где $f_0(x)$ и $f_{\partial on}(x)$ — плотности вероятностей распределения F_0 и $F_{0,\partial on}$ соответственно. Оценим вероятность P для случая, когда достаточно большое число параметров v_q влияет на изменения F_0 и $F_{0,\partial on}$. Тогда (независимо от законов распределения v_q) для $f_0(x)$ и $f_{\partial on}(x)$ можно принять нормальные законы:

$$f_0(x) = \frac{1}{\sigma_0^* \sqrt{2\pi}} \exp \left[-\frac{(x - F_0^*)^2}{2(\sigma_0^*)^2} \right];$$

$$f_{\partial on}(x) = \frac{1}{\sigma_{0,\partial on}^* \sqrt{2\pi}} \exp \left[-\frac{(x - F_{0,\partial on}^*)^2}{2(\sigma_{0,\partial on}^*)^2} \right].$$

1. Пусть $F_{0,\partial on}$ не является случайной величиной, тогда

$$f_{\partial on}(x) = \delta(x - F_{0,\partial on}^*), \quad P = \frac{1}{2} + \Phi \left(\frac{F_{0,\partial on}^* - F_0^*}{\sigma_0^*} \right), \quad (4.29)$$

где $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-v^2/2} dv$ — интеграл вероятностей.

Зная величины P , F_0^* и $F_{0,\partial on}^*$ и используя табличные значения функции $\Phi(z)$, можно найти аргумент $\omega_p = \frac{F_{0,\partial on}^* - F_0^*}{\sigma_0^*}$ и получить следующее условие для среднеквадратического отклонения σ_0^* :

$$\sigma_0^* \leq \frac{F_{0,\partial on}^* - F_0^*}{\omega_p}. \quad (4.30)$$

2. В общем случае (когда F_0 и $F_{0,\partial on}$ — случайные величины) удастся получить лишь приближенную оценку для P , раскладывая интегралы

$\int_x^\infty f_{\partial on}(x') dx'$ и $\int_{-\infty}^x f_0(x') dx'$ в степенные ряды в окрестности значений

$x = F_0^*$ и $x = F_{0,\partial on}^*$ соответственно. Если ограничиться тремя членами разложения, то получим:

$$P \approx \frac{1}{2} + \Phi \left(\frac{F_{0,\partial on}^* - F_0^*}{\sigma_{0,\partial on}^*} \right) - \frac{F_{0,\partial on}^* - F_0^*}{2} \cdot \left(\frac{\sigma_0^*}{\sigma_{0,\partial on}^*} \right)^2 f_{\partial on}(F_0^*); \quad (4.31)$$

$$P \approx \frac{1}{2} + \Phi \left(\frac{F_{0,доп}^* - F_0^*}{\sigma_0^*} \right) - \frac{F_{0,доп}^* - F_0^*}{2} \cdot \left(\frac{\sigma_{0,доп}^*}{\sigma_0^*} \right)^2 f_0(F_{0,доп}^*). \quad (4.32)$$

Выражения (4.32) и (4.31) дают близкие значения вероятности P и любое из них можно использовать для получения приближенных соотношений между σ_0^* и $\sigma_{0,доп}^*$ при заданной величине P . Например, если $\sigma_{0,доп}^*$ известна, то для σ_0^* получим условие (воспользовавшись соотношением (4.31)):

$$\left(\frac{\sigma_0^*}{\sigma_{0,доп}^*} \right)^2 \leq \frac{2}{F_{0,доп}^* - F_0^*} \cdot \frac{1}{f_{доп}(F_0^*)} \left[\frac{1}{2} + \Phi \left(\frac{F_{0,доп}^* - F_0^*}{\sigma_{0,доп}^*} \right) - P \right], \quad (4.33)$$

которое справедливо при выполнении неравенств

$$F_{0,доп}^* > F_0^* ; \quad \frac{1}{2} + \Phi \left(\frac{F_{0,доп}^* - F_0^*}{\sigma_{0,доп}^*} \right) > P.$$

Как только получено значение σ_0^* , можно перейти к выбору допустимых погрешностей исходных данных. Проще всего это сделать, предположив, что наименьшую погрешность должны иметь те исходные параметры ϑ_q , по отношению к которым наиболее чувствителен критерий F_0 [27], т.е. принять, что $\sigma_q = \frac{C}{|\mathcal{L}_q|}$, где $C > 0$ — неизвестный параметр, выбираемый так, чтобы удовлетворить условию (4.27). Например, если погрешности всех исходных данных $\vartheta_1, \vartheta_2, \dots, \vartheta_t$ не коррелируют (т.е. $r_{ij} = 0$), то:

$$C = \frac{\sigma_0^*}{\sqrt{t}} \quad \text{и} \quad \sigma_q = \frac{\sigma_0^*}{\sqrt{t} |\mathcal{L}_q|}. \quad (4.34)$$

Изложенный выше подход был использован в работе [27] для выработки требований к точности знания исходных данных в задачах, связанных с нахождением оптимальных компоновок быстрых реакторов.

Другой подход к решению задач (1.6) — (1.9) дает стохастический метод сокращения невязок [24]. Этот итерационный метод является обобщением метода стохастической аппроксимации на случай задач с ограничениями типа неравенства. Он интересен тем, что не требует знания аналитических выражений для математических ожиданий. Метод близок по своей идее к известным градиентным методам, но при выборе направления спуска на каждой итерации позволяет учесть стохастическую природу задачи.

Итерационную схему метода рассмотрим на примере решения стохастической задачи нелинейного программирования:

$$\left. \begin{aligned} \min MF_0(\vec{u}, \vec{v}); \\ MF_i(\vec{u}, \vec{v}) \leq 0, \quad i=1, 2, \dots, m; \\ \vec{u} \in U, \end{aligned} \right\} \quad (4.35)$$

в которой функции $MF_i(\vec{u}, \vec{v})$ имеют в ограниченной области U вторые производные.

Предварительно введем в рассмотрение операцию проектирования $\mathcal{P}_X(\vec{z})$ точки \vec{z} на множество X : $\mathcal{P}_X(\vec{z}) \in X$, $\|\vec{z} - \mathcal{P}_X(\vec{z})\| \leq \|\vec{y} - \vec{z}\|$ для любого $\vec{y} \in X$. Эта операция означает нахождение такой точки $\mathcal{P}_X(\vec{z}) \in X$, расстояние которой до фиксированной точки \vec{z} не превышает расстояния любой другой точки $\vec{y} \in X$ до той же точки \vec{z} . Очевидно, что в качестве $\mathcal{P}_X(\vec{z})$ можно принять решение следующей экстремальной задачи (при данном \vec{z}): $\min \|\vec{z} - \vec{y}\|$, $\vec{y} \in X$.

В частности, если множество X описывается неравенствами $a_i \leq y_i \leq b_i$, $i=1, 2, \dots, n$, то $\mathcal{P}_X(\vec{z})$ — вектор с компонентами

$$[\mathcal{P}_X(\vec{z})]_i = \begin{cases} a_i, & \text{если } z_i \leq a_i; \\ b_i, & \text{если } z_i \geq b_i; \\ z_i, & \text{если } a_i < z_i < b_i. \end{cases}$$

Составим функцию Лагранжа (см. (П.1.4) приложения 1): $\mathcal{L}(\vec{u}, \vec{v}, \vec{\lambda}) =$

$= F_0(\vec{u}, \vec{v}) + \sum_{i=1}^m \lambda_i F_i(\vec{u}, \vec{v})$, где $\vec{\lambda}$ — вектор с компонентами $\lambda_i \geq 0$, $i=1, 2, \dots, m$. Обозначим через $\vec{F}(\vec{u}, \vec{v})$ — вектор с компонентами $F_1(\vec{u}, \vec{v})$, $F_2(\vec{u}, \vec{v}), \dots, F_m(\vec{u}, \vec{v})$ и предположим, что значения $\vec{u}^{(s)}$ и $\vec{\lambda}^{(s)}$ на s -й итерации метода известны (заданы в случае $s=0$). Тогда значения $\vec{u}^{(s+1)}$ и $\vec{\lambda}^{(s+1)}$ в следующей $(s+1)$ -й итерации находятся из равенств:

$$\left. \begin{aligned} \vec{u}^{(s+1)} &= \mathcal{P}_U(\vec{u}^{(s)} - \rho_s \nabla_{\vec{u}} \mathcal{L}(\vec{u}^{(s)}, \vec{v}^{(s)}, \vec{\lambda}^{(s)})); \\ \vec{\lambda}^{(s+1)} &= \mathcal{P}_\Lambda(\vec{\lambda}^{(s)} + \rho_s \nabla_{\vec{\lambda}} \mathcal{L}(\vec{u}^{(s)}, \vec{v}^{(s)}, \vec{\lambda}^{(s)})), \end{aligned} \right\} \quad (4.36)$$

где компоненты $\xi_j^{(s)}$, $j=1, 2, \dots, n$ вектора $\xi^{(s)}$ имеют смысл статистических оценок производных от функции \mathcal{L} по управлениям u_j и определяются по формулам:

$$\xi_j^{(s)} = \frac{\mathcal{L}(\vec{u}^{(s)} + \Delta_s \vec{e}_j, \vec{\lambda}^{(s)}, \vec{v}^{(s,j)}) - \mathcal{L}(\vec{u}^{(s)}, \vec{\lambda}^{(s)}, \vec{v}^{(s,0)})}{\Delta_s}, \quad (4.37)$$

в которых \vec{e}_j — единичные орты направлений в n -мерном пространстве управлений; $\vec{v}^{(s,j)}$, $j=0,1,\dots,n$ — серия независимых по $s=0,1,\dots$ наблюдений состояния параметров \vec{v} , в частности можно взять $\vec{v}^{(s,0)} = \vec{v}^{(s,1)} = \dots = \vec{v}^{(s,n)} = \vec{v}^{(s)}$, $\gamma_s > 0$ — нормирующий множитель, который можно положить равным 1, если U — ограниченное множество.

Значения шага ρ_s и приращения Δ_s выбирают так, чтобы:

$$\rho_s \geq 0, \sum_{s=0}^{\infty} \rho_s = \infty, \sum_{s=0}^{\infty} (\rho_s |\Delta_s| + \rho_s^2) < \infty. \quad (4.38)$$

Например, можно принять: $\rho_s = \frac{\rho_0}{s}$ и $\Delta_s = \frac{\Delta_0}{s}$, $s=1,2,\dots$, где $\rho_0 > 0$ и Δ_0 — заданные константы. Отметим, что в методе стохастической аппроксимации нельзя регулировать шаг ρ_s так, как это делается в градиентных методах (уменьшать шаг, если минимизируемая функция цели возрастает), поскольку оценка производных по (4.37) не гарантирует монотонного уменьшения $F_0(\vec{u}, \vec{v})$ и $MF_0(\vec{u}, \vec{v})$. Поэтому шаг ρ_s предложено выбирать "программным" образом, т.е. из условия (4.38). Типичный характер изменения этих величин в процессе итераций изображен на рис. 6.

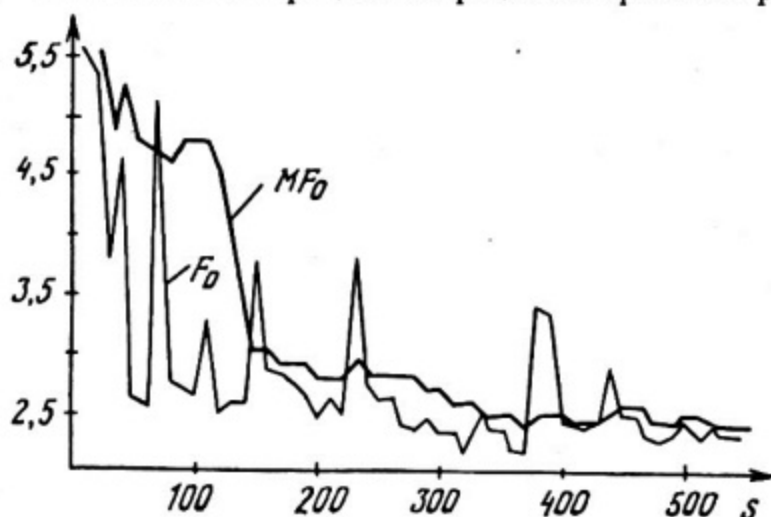


Рис. 6. Характер измерения F_0 и MF_0 в процессе итераций в стохастическом методе сокращения невязок

Рассмотренный алгоритм расчета $\xi_j^{(s)}$ легко обобщается на случай вариационных задач со случайными параметрами. Надо лишь иметь в виду, что в таких задачах стохастический квазиградиент $\xi_j^{(s)}$ должен вычис-

ляться (при независимых наблюдениях $\vec{v}^{(s,j)}$ для параметров \vec{v}) с учетом неявной зависимости функционалов $F_{\vec{v}}$ и переменных состояния $\varphi(x)$ от параметров \vec{v} .

4.3. Задачи с неопределенными исходными данными

Методы решения задач (1.6) – (1.9) с неопределенными исходными данными v_q , $q=1, \dots, t$, для которых неизвестны вероятностные законы распределения, а заданы лишь интервалы изменения:

$$v_q^{\text{МИН}} \leq v_q \leq v_q^{\text{МАКС}}, \quad q = 1, 2, \dots, t, \quad (4.39)$$

разработаны сравнительно слабо. Имеются рекомендации общего характера [28], предлагающие находить решение таких задач в два этапа. Вначале построить так называемую "платежную" матрицу, а затем с помощью различных критериев отобрать наиболее рациональные решения.

"Платежная" матрица (табл. 2) представляет из себя упорядоченный набор значений критерия оптимальности F_0 , соответствующих различным значениям параметров \vec{v} и допустимых управлений \vec{u} . Построению "платежной" матрицы предшествует выбор из числа неопределенных данных (4.39), таких, изменение которых оказывает наибольшее влияние на критерий оптимальности. Это можно сделать на основе анализа устойчивости (см. п. 4.1) по величине коэффициентов чувствительности α_q , полученных после решения задачи (1.6) – (1.9) при $v_q = \bar{v}_q = \frac{1}{2}(v_q^{\text{МАКС}} + v_q^{\text{МИН}})$.

Т а б л и ц а 2

Управления $\vec{u}^{(k)}$	Состояния параметров \vec{v}						$F_{0,k}^{\text{МАКС}}$
	$\vec{v}^{(1)}$	$\vec{v}^{(2)}$...	$\vec{v}^{(i)}$...	$\vec{v}^{(M)}$	
$\vec{u}^{(1)}$	$F_{0,1}^{(1)}$	$F_{0,1}^{(2)}$...	$F_{0,1}^{(i)}$...	$F_{0,1}^{(M)}$	$F_{0,1}^{\text{МАКС}}$
$\vec{u}^{(2)}$	$F_{0,2}^{(1)}$	$F_{0,2}^{(2)}$...	$F_{0,2}^{(i)}$...	$F_{0,2}^{(M)}$	$F_{0,2}^{\text{МАКС}}$
...
$\vec{u}^{(k)}$	$F_{0,k}^{(1)}$	$F_{0,k}^{(2)}$...	$F_{0,k}^{(i)}$...	$F_{0,k}^{(M)}$	$F_{0,k}^{\text{МАКС}}$
...
$\vec{u}^{(M)}$	$F_{0,M}^{(1)}$	$F_{0,M}^{(2)}$...	$F_{0,M}^{(i)}$...	$F_{0,M}^{(M)}$	$F_{0,M}^{\text{МАКС}}$
	$F_{0,\text{МИН}}^{(1)}$	$F_{0,\text{МИН}}^{(2)}$...	$F_{0,\text{МИН}}^{(i)}$...	$F_{0,\text{МИН}}^{(M)}$	

Пусть Q_N — множество индексов исходных данных, наиболее сильно влияющих на критерий F_0 . Хотя число N индексов, образующих множество Q_N , обычно невелико по сравнению с числом t , но рассматривать решения задач (1.6) — (1.9) при всевозможных v_q , $q \in Q_N$ не представляется возможным, так как эти параметры принимают непрерывный ряд значений внутри конечных интервалов (4.39). Поэтому ограничиваются конечным дискретным набором $\{\vec{v}^{(1)}, \vec{v}^{(2)}, \dots, \vec{v}^{(M)}\}$ значений параметров \vec{v} . При этом число M выбирают с учетом возможностей, имеющихся в распоряжении ЭВМ, а точки $\vec{v}^{(i)}$ стараются равномерно разместить внутри области, заданной неравенствами (4.39) (при $q \in Q_N$) в N -мерном пространстве. С этой целью область (4.39) сначала путем замены переменных $\frac{v_q - v_q^{\text{МИН}}}{v_q^{\text{МАКС}} - v_q^{\text{МИН}}} \rightarrow v_q$, $q \in Q_N$, превращают в N -мерный единичный куб. Затем координаты точек определяют, используя те или иные алгоритмы построения равномерно распределенных последовательностей.

Если остановиться на $ЛЛ_2$ последовательностях (которые в настоящее время считаются наилучшими из известных равномерно распределенных последовательностей), то алгоритм нахождения координат $v_j^{(i)}$ точек $\vec{v}^{(i)}$ в N -мерном единичном кубе выглядит следующим образом [29].

Пусть номер точки i записан в двоичной системе: $i = e_m e_{m-1} \dots e_2 e_1$. Тогда декартовы координаты $v_1^{(i)}, v_2^{(i)}, \dots, v_N^{(i)}$ точки i вычисляются по формуле $v_j^{(i)} = e_1 V_j^{(1)} * e_2 V_j^{(2)} * \dots * e_m V_j^{(m)}$, $j = 1, 2, \dots, N$, где $V_j^{(l)} = r_j^{(l)} 2^{-l}$ — направляющие числа, числители которых $r_j^{(l)}$ приведены в работе [29], а символ $*$ означает поразрядное сложение по модулю два в двоичной системе. Например: $7/16 * 5/32 = 11/32$, так как $7/16 = 0,0111$, $5/32 = 0,00101$, а "сумма" $0,0111 * 0,00101 = 0,01011$.

На рис. 7 представлено размещение точек в единичном квадрате для случаев, когда $M=7$ и $M=15$. Видно, что в обоих случаях проекции точек на каждую из координатных осей имеют вид "правильных" сеток (т.е. распределены равномерно), причем при введении новых точек не меняются координаты ранее выбранных точек.

Существуют и другие способы размещения заданного числа точек в N -мерном кубе. Например, в работе [30] предлагается координаты точек выбирать так, чтобы минимальное расстояние между точками было максимально. Для сравнения на рис. 8 представлено размещение семи точек в единичном квадрате, полученное с использованием этого дополнительного требования.

Определившись в значениях случайных параметров, перейдем к вычислению элементов "платежной" матрицы \hat{F}_0 (см. табл. 2). Для этого при каждом значении $\vec{v}^{(i)} = \{v_q^{(i)}\}$, $q \in Q_N$, $i=1, \dots, M$, находим решение задачи (1.6) —

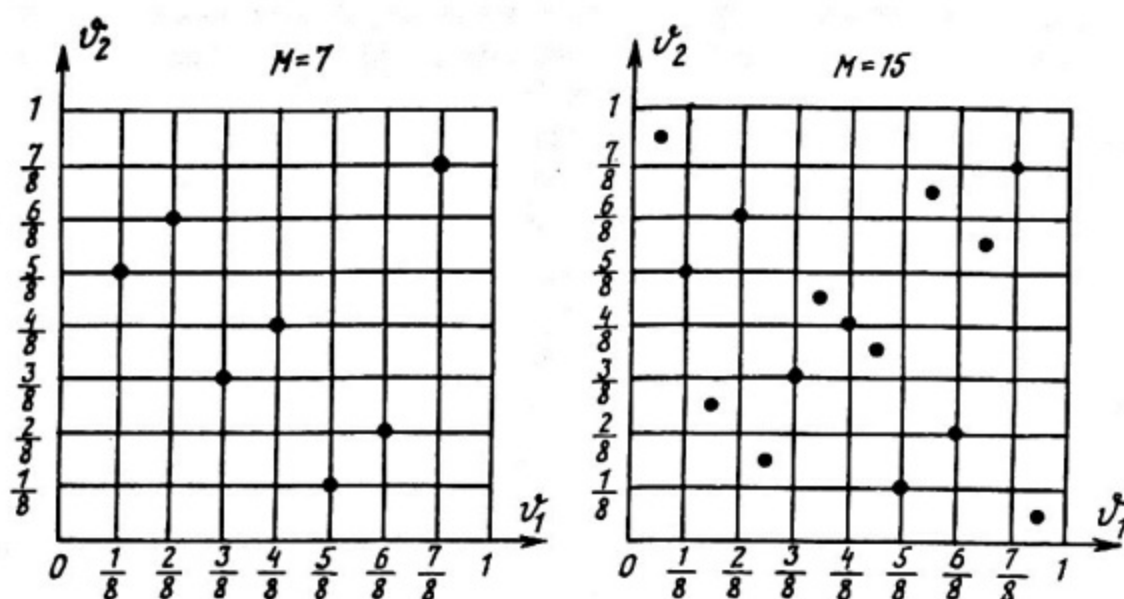


Рис. 7. Размещение точек в соответствии с ЛП τ последовательностью

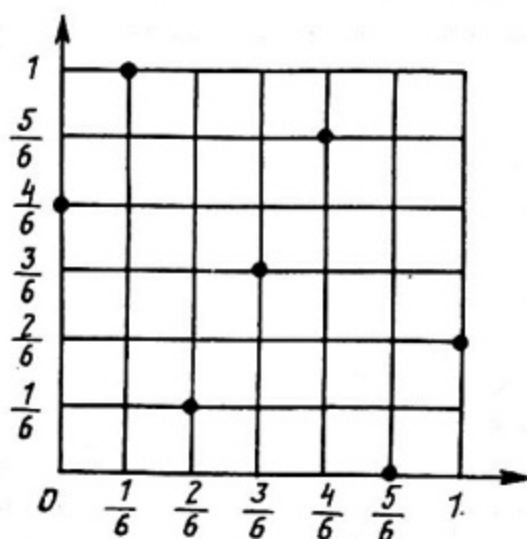


Рис. 8. Равномерное размещение семи точек в узлах двумерной сетки с 49-ю узлами

(1.9) с помощью любого из рассмотренных в разделе 3 методов и получаем оптимальные значения $F_{0,1}^{(1)}, F_{0,2}^{(2)}, \dots, F_{0,M}^{(M)}$ критерия F_0 и соответствующие управления $\vec{u}^{(1)}, \vec{u}^{(2)}, \dots, \vec{u}^{(M)}$. Затем каждое решение $\{F_{0,i_0}^{(i_0)}, \vec{u}^{(i_0)}\}$, полученное при $\vec{v} = \vec{v}^{(i_0)}$, адаптируем (т.е. приспособляем) к другим значениям $\vec{v}^{(i)}$, $i \neq i_0$. Поскольку управление $\vec{u}^{(i_0)}$ может не принадлежать допустимой области, если $\vec{v} \neq \vec{v}^{(i_0)}$, то необходимо иметь в распоряжении некоторые параметры $V_{k,l}^{(i)}$ (в общем случае L параметров для

каждого $\vec{v}^{(i)}$ и $\vec{u}^{(k)}$, за счет изменения которых можно было бы восстановить нарушенные ограничения (1.8).

Объединим эти параметры в векторы $\vec{V}_k^{(i)} = \{V_{k,1}^{(i)}, V_{k,2}^{(i)}, \dots, V_{k,L}^{(i)}\}$, $i = 1, 2, \dots, M$; $k = 1, 2, \dots, M$, и назовем параметрами адаптации. Очевидно, что восстановление ограничений (1.8) приведет к изменению функционала F_0 от значения $F_{0,i}^{(i)}$ до величины $F_{0,k}^{(i)}$:

$$F_{0,k}^{(i)} = F_{0,i}^{(i)} \pm \Delta F_{0,k}^{(i)}, \quad i = 1, 2, \dots, M, \quad k = 1, 2, \dots, M, \quad (4.40)$$

причем $\Delta F_{0,i}^{(i)} = 0$, $i = 1, 2, \dots, M$.

Вопрос о выборе адаптирующих векторов $\vec{V}_k^{(i)}$ и нахождении величин $\Delta F_{0,k}^{(i)}$ не является тривиальным. Действительно, выполнить условия (1.8) можно при различных значениях $\vec{V}_k^{(i)}$, что делает задачу адаптации неоднозначной. Чтобы исключить эту неоднозначность, необходимо сформулировать дополнительные критерии, зависящие от цели проектирования и смысла оптимизируемого функционала F_0 . Так, если F_0 имеет смысл затрат на создание проектируемой установки, то нужно выбирать $\vec{V}_k^{(i)}$, исходя из минимума F_0 при ограничениях (1.6) – (1.9). В других случаях может оказаться полезным выбор векторов $\vec{V}_k^{(i)}$, исходя из минимума $\|\vec{V}_k^{(i)} - \vec{V}_i^{(i)}\|$, где под $\vec{V}_i^{(i)}$ подразумеваются те значения параметров адаптации, которые принимались при нахождении $F_{0,i}^{(i)}$.

Однако в любом случае приходится решать новую оптимизационную задачу. Избежать этого можно при условии, что выбор $\vec{V}_k^{(i)}$ происходит на основе того же критерия оптимальности F_0 , который использовался при получении управления $\vec{u}^{(i)}$, а степень нарушения ограничений (1.8) невелика. Тогда применимы соотношения (4.15) для оценки изменений $\Delta F_{0,k}^{(i)}$ с помощью соответствующих коэффициентов чувствительности, рассчитанных при $\vec{v} = \vec{v}^{(i)}$, $\vec{u} = \vec{u}^{(i)}$.

Изложенный выше способ нахождения элементов "платежной" матрицы \hat{F}_0 приемлем, если полученное множество управлений $\vec{u}^{(k)}$, $k = 1, 2, \dots, M$, будет достаточно полным. Полнота выявления допустимых управлений зависит от свойств функционала F_0 и выбора множества состояний параметров \vec{v} . Если нет уверенности в достаточной представительности множества найденных управлений, то рекомендуется заранее задавать допустимые управления $\vec{u}^{(k)}$, $k = 1, 2, \dots, M_0$. Соответствующие значения $\vec{u}^{(k)}$ можно выбирать также, как дискретные значения \vec{v} . Но здесь имеется определенная трудность, связанная с необходимостью выполнения нетривиальных ограничений (1.8). Преодолевается это затруднение либо за счет введения параметров адаптации, либо путем исключения из рассмотрения тех управлений, при которых ограничения (1.8) нарушаются. Дальнейшие расчеты элементов матрицы \hat{F}_0 очевидны. В этом случае в матрице \hat{F}_0 число строк может не совпадать с числом столбцов (когда $M_0 \neq M$).

Перейдем теперь к рассмотрению некоторых критериев выбора рациональных решений. Выбор производится на основе анализа элементов $F_{0,k}^{(i)}$ "платежной" матрицы \hat{F}_0 . Для определенности будем считать, что матрица \hat{F}_0 имеет размерность $M \times M$.

Критерий Вальда (минимаксных затрат) выявляет наилучшее решение при наиболее неблагоприятном сочетании параметров \vec{D} . В соответствии с этим критерием находится такое управление \vec{u}_B , при котором

$$\min_k (\max_i F_{0,k}^{(i)}) = \min_k F_{0,k}^{\text{МАКС}} \longrightarrow \vec{u}_B. \quad (4.41)$$

Очевидно, что управление \vec{u}_B является крайне осторожным или, как говорят, "консервативным", и можно ожидать снижения критерия F_0 , если действовать с определенным риском.

Критерий Сэвиджа (минимаксного риска) можно использовать, если построена матрица рисков \hat{R}_0 с элементами $R_{0,k}^{(i)}$. Эта матрица строится на основе "платежной" матрицы \hat{F}_0 следующим образом.

Для каждого состояния $\vec{D}^{(i)}$ параметров \vec{D} определяется минимальное значение $F_{0,\text{МИН}}^{(i)} = \min_k F_{0,k}^{(i)}$ (см. нижнюю строку в табл. 2). Значение риска $R_{0,k}^{(i)}$ определяется в виде:

$$R_{0,k}^{(i)} = F_{0,k}^{(i)} - F_{0,\text{МИН}}^{(i)} = F_{0,k}^{(i)} - \min_k F_{0,k}^{(i)}, \quad (4.42)$$

т.е. как разность между значением критерия $F_{0,k}^{(i)}$, полученным при $\vec{u} = \vec{u}^{(k)}$, $\vec{D} = \vec{D}^{(i)}$, и наилучшим значением критерия $F_{0,\text{МИН}}^{(i)}$, соответствующим состоянию параметров $\vec{D}^{(i)}$.

Принцип Сэвиджа состоит в том, чтобы не допустить слишком больших значений рисков $R_{0,k}^{(i)}$. Поэтому выбор наилучшего по этому критерию управления \vec{u}_C осуществляется путем минимизации максимально возможного при всех состояниях $\vec{D}^{(i)}$ риска:

$$\min_k R_{0,k}^{\text{МАКС}} = \min_k (\max_i R_{0,k}^{(i)}) \longrightarrow \vec{u}_C. \quad (4.43)$$

Этот критерий так же, как критерий Вальда, ориентируется на наиболее неблагоприятное состояние параметров \vec{D} и является "консервативным".

Критерий "пессимизма—оптимизма" Гурвица позволяет учесть как наилучшие, так и наихудшие сочетания параметров \vec{D} . Для этого при каждом $\vec{u} = \vec{u}^{(k)}$ с весом показателя α усредняются максимально возможное значение $F_{0,k}^{\text{МАКС}} = \max_i F_{0,k}^{(i)}$ критерия F_0 с минимально возможным значением $F_{0,k}^{\text{МИН}} = \min_i F_{0,k}^{(i)}$, т.е. составляется сумма $\alpha F_{0,k}^{\text{МАКС}} +$

$+ (1-\alpha) F_{0,k}^{\text{МИН}}$, где $0 \leq \alpha \leq 1$. Затем выбирается такое управление \vec{u}_r , при котором

$$\min_k [\alpha F_{0,k}^{\text{МАКС}} + (1-\alpha) F_{0,k}^{\text{МИН}}] \rightarrow \vec{u}_r. \quad (4.44)$$

В частности, при $\alpha = 1$ условие (4.44) переходит в критерий Вальда (критерий "пессимизма"), а при $\alpha = 0$ — в критерий "крайнего оптимизма".

Недостаток критерия Гурвица в том, что значение параметра α приходится выбирать из субъективных соображений. Обычно принимают величину α в интервале между 0,5 и 1, причем тем ближе к 1, чем более серьезны последствия от незнания состояния параметров \vec{v} .

Критерий Лапласа (недостаточного основания) исходит из того, что любые сочетания параметров \vec{v}_q внутри интервалов (4.39) равновероятны. Он выделяет такое управление \vec{u}_l , при котором достигается минимум среднего арифметического значения $\langle F_{0,k} \rangle$:

$$\min_k \langle F_{0,k} \rangle = \min_k \frac{1}{M} \sum_{i=1}^M F_{0,k}^{(i)} \rightarrow \vec{u}_l. \quad (4.45)$$

Недостаток этого критерия в том, что он ориентируется на равномерный закон распределения параметров \vec{v} . Такое предположение, как правило, не соответствует действительности.

Часто бывает трудно отдать предпочтение какому-либо одному из рассмотренных выше критериев. Поэтому поступают следующим образом. С помощью критериев Вальда, Сэвиджа, Гурвица и Лапласа выделяют рациональные решения: \vec{u}_B , \vec{u}_C , \vec{u}_r , \vec{u}_l . Если все решения совпадают, то задачу можно считать решенной. В противном случае найденные решения следует подвергнуть дополнительному анализу, исходя из каких-либо других критериев.

В заключение отметим, что обычно встречаются задачи проектирования, в которых имеются как неопределенные, так и случайные данные с известными законами распределения. В таком случае используют комбинированные подходы, рассматривая серии задач с фиксированными наборами неопределенных данных и решая каждую из них с помощью методов стохастического программирования.

ОБЩИЕ СВЕДЕНИЯ ПО НЕЛИНЕЙНОМУ И ВЫПУКЛОМУ
ПРОГРАММИРОВАНИЮ

Задача (1.1) – (1.5) в отсутствии уравнений для переменных состояния $\vec{\psi}(x)$ превращается в общую задачу нелинейного программирования:

$$\min F_0(\vec{u}); \quad (\text{П.1.1})$$

$$F_i(\vec{u}) \leq F_i^{\text{доп}}, \quad i=1,2,\dots,m; \quad (\text{П.1.2})$$

$$\vec{u} \in U, \quad (\text{П.1.3})$$

где вектор \vec{u} по-прежнему содержит n компонент u_1, u_2, \dots, u_n ; U – множество точек в n -мерном пространстве, описываемое неравенствами (1.4), а любое ограничение типа равенства ($F = F^{\text{доп}}$) заменено на два неравенства ($F \leq F^{\text{доп}}, -F \leq -F^{\text{доп}}$) и включено в число условий (П.1.2).

Из теории нелинейного программирования [22] известно, что решение задачи (П.1.1) – (П.1.3) связано с отысканием седловой точки функции Лагранжа

$$\mathcal{L}(\vec{u}, \vec{\lambda}) = F_0(\vec{u}) + \sum_{i=1}^m \lambda_i [F_i(\vec{u}) - F_i^{\text{доп}}] \quad (\text{П.1.4})$$

в области $\vec{u} \in U, \vec{\lambda} \in \Lambda$. Здесь через Λ обозначено множество во всех неотрицательных значений множителей Лагранжа λ_i ($\lambda_i \geq 0, i=1,2,\dots,m$).

ОПРЕДЕЛЕНИЕ 1. Седловой точкой функции Лагранжа $\mathcal{L}(\vec{u}, \vec{\lambda})$ называется пара векторов ($\vec{u}^* \in U, \vec{\lambda}^* \in \Lambda$), для которых

$$\mathcal{L}(\vec{u}^*, \vec{\lambda}) \leq \mathcal{L}(\vec{u}^*, \vec{\lambda}^*) \leq \mathcal{L}(\vec{u}, \vec{\lambda}^*) \quad (\text{П.1.5})$$

при всех $\vec{u} \in U, \vec{\lambda} \in \Lambda$. Это же можно записать так:

$$\mathcal{L}(\vec{u}^*, \vec{\lambda}^*) = \max_{\vec{\lambda} \in \Lambda} \min_{\vec{u} \in U} \mathcal{L}(\vec{u}, \vec{\lambda}) = \min_{\vec{u} \in U} \max_{\vec{\lambda} \in \Lambda} \mathcal{L}(\vec{u}, \vec{\lambda}). \quad (\text{П.1.6})$$

Связь между координатами ($\vec{u}^*, \vec{\lambda}^*$) седловой точки и решением задачи (П.1.1) – (П.1.3) устанавливает теорема 1.

Если $(\vec{u}^*, \vec{\lambda}^*)$ – седловая точка функции Лагранжа $\mathcal{L}(\vec{u}, \vec{\lambda})$ в области $\vec{u} \in U$, $\vec{\lambda} \in \Lambda$, то \vec{u}^* – оптимальное решение задачи (П.1.1) – (П.1.3).

Обратное утверждение, вообще говоря, справедливо лишь при определенных условиях (см. теорему 5).

Необходимые условия оптимальности для задачи (П.1.1) – (П.1.3) с непрерывно дифференцируемыми функциями $F_i(\vec{u})$ формулируются теоремой 2.

Если \vec{u}^* – решение задачи (П.1.1) – (П.1.3), то существуют не все равные нулю множители λ_i^* , $i=0, 1, \dots, m$, такие что:

$$\left. \begin{aligned} \lambda_i^* &\geq 0, \quad i=0, 1, \dots, m; \\ \lambda_i^* [F_i(\vec{u}^*) - F_i^{\text{доп}}] &= 0, \quad i=1, 2, \dots, m, \end{aligned} \right\} \quad (\text{П.1.7})$$

и при всех $\vec{u} \in U$ выполняется неравенство

$$\left(\sum_{i=0}^m \lambda_i^* \nabla F_i(\vec{u}^*), \vec{u} - \vec{u}^* \right) \geq 0, \quad (\text{П.1.8})$$

где $\nabla F_i(\vec{u}^*)$ – градиент функции $F_i(\vec{u})$, вычисленный в точке \vec{u}^* .

Когда в задаче (П.1.1) – (П.1.3) функции $F_i(\vec{u})$, $i=0, 1, \dots, m$, и множество U выпуклы, приходим к задаче выпуклого программирования.

ОПРЕДЕЛЕНИЕ 2. Множество U является выпуклым, если из условия: $\vec{u}^{(1)} \in U$, $\vec{u}^{(2)} \in U$ вытекает, что:

$$\alpha \vec{u}^{(1)} + (1-\alpha) \vec{u}^{(2)} \in U, \quad (\text{П.1.9})$$

где $0 \leq \alpha \leq 1$.

ОПРЕДЕЛЕНИЕ 3. Функция $F(\vec{u})$, определенная на выпуклом множестве U , называется выпуклой, если для любых $\vec{u}^{(1)} \in U$, $\vec{u}^{(2)} \in U$ и всех $0 \leq \alpha \leq 1$

$$F(\alpha \vec{u}^{(1)} + (1-\alpha) \vec{u}^{(2)}) \leq \alpha F(\vec{u}^{(1)}) + (1-\alpha) F(\vec{u}^{(2)}). \quad (\text{П.1.10})$$

На рис. П.1 изображена выпуклая функция $F(u)$ одного аргумента. Видно, что выпуклая функция не обязательно непрерывно дифференцируема.

Частными случаями задач выпуклого программирования являются:

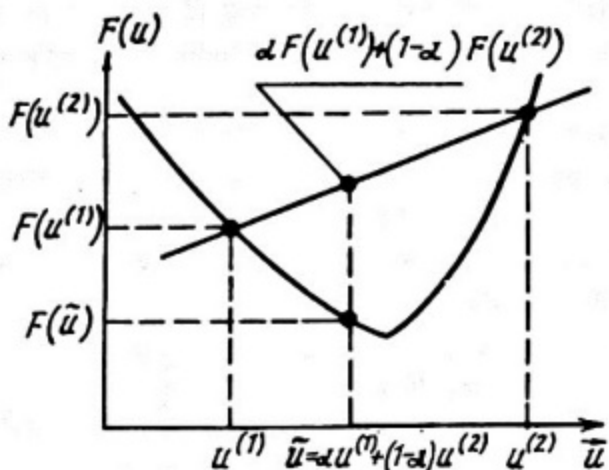


Рис. П.1. Выпуклая функция $F(u)$ одного аргумента

а) задача линейного программирования

$$\left. \begin{aligned} \min (\vec{c}, \vec{u}) ; \\ (\vec{a}_i, \vec{u}) \leq F_i^{\text{доп}}, \quad i = 1, 2, \dots, m ; \\ \vec{u}^{\text{мин}} \leq \vec{u} \leq \vec{u}^{\text{макс}}, \end{aligned} \right\} \quad (\text{П.1.11})$$

где \vec{c} , \vec{a}_i , $\vec{u}^{\text{мин}}$, $\vec{u}^{\text{макс}}$ — заданные вектора;

б) задача квадратичного программирования

$$\left. \begin{aligned} \min \left[\frac{1}{2} (\vec{u}, \hat{c} \vec{u}) + (\vec{d}, \vec{u}) \right] ; \\ (\vec{a}_i, \vec{u}) \leq F_i^{\text{доп}}, \quad i = 1, 2, \dots, m ; \\ \vec{u}^{\text{мин}} \leq \vec{u} \leq \vec{u}^{\text{макс}}, \end{aligned} \right\} \quad (\text{П.1.12})$$

если матрица \hat{c} положительно определена.

Замечательное свойство задач выпуклого программирования устанавливает теорема 3 (о единственности минимума).

Каждый локальный минимум \vec{u}^* задачи выпуклого программирования является ее глобальным минимумом.

Кроме того, для задач выпуклого программирования развита теория двойственности, позволяющая установить связь решения прямой задачи (П.1.1)–(П.1.3) с решением так называемой двойственной задачи.

Введем в рассмотрение двойственную функцию

$$q(\vec{\lambda}) = \min_{\vec{u} \in U} \mathcal{L}(\vec{u}, \vec{\lambda}), \quad (\text{П.1.13})$$

определенную на множестве Λ таких неотрицательных значений $\lambda_i \geq 0$ ($i = 1, 2, \dots, m$), при которых существует минимум функций Лагранжа $\mathcal{L}(\vec{u}, \vec{\lambda})$, т.е.:

$$\Lambda = \left\{ \lambda_i \geq 0, i = 1, 2, \dots, m, \text{ и таких, что существует } \min_{\vec{u} \in U} \mathcal{L} \right\}. \quad (\text{П.1.14})$$

Тогда двойственная задача запишется в виде:

$$\max_{\vec{\lambda} \in \Lambda} g(\vec{\lambda}). \quad (\text{П.1.15})$$

Знание двойственной функции $g(\vec{\lambda})$ позволяет получить нижнюю оценку для $F_0(\vec{u})$, используя теорему 4.

Для всех \vec{u} , удовлетворяющих условиям задачи (П.1.1) – (П.1.3), и для всех $\vec{\lambda} \in \Lambda$ (П.1.14)

$$g(\vec{\lambda}) \leq F_0(\vec{u}). \quad (\text{П.1.16})$$

ОПРЕДЕЛЕНИЕ 4. Говорят, что ограничения (П.1.2) задачи (П.1.1) – (П.1.3) удовлетворяют условию Слейтера, если существует такой вектор $\vec{u}' \in U$, что $F_i(\vec{u}') < F_i^{\text{доп}}$, $i = 1, 2, \dots, m$.

Важность этого условия обычно связана с тем, что множество Λ множителей Лагранжа оказывается конечным. При этом имеет место теорема 5.

Если \vec{u}^* – решение задачи выпуклого программирования (П.1.1) – (П.1.3) и ограничения (П.1.2) удовлетворяют условию Слейтера, то существует вектор $\vec{\lambda}^*$ с компонентами $\lambda_i^* \geq 0$, $i = 1, 2, \dots, m$, который является решением двойственной задачи (П.1.15) и при этом:

$$g(\vec{\lambda}^*) = F_0(\vec{u}^*).$$

Отметим, что для задач линейного программирования существует более сильное утверждение, которое формулируется теоремой 6.

Если одна из задач двойственной пары (прямая (П.1.11) и двойственная к ней (П.1.15)) имеет решение, то другая задача также разрешима; при этом для любых оптимальных решений этих задач имеет место равенство $\min F_0(\vec{u}) = \max g(\vec{\lambda})$.

**СИМПЛЕКС-МЕТОД РЕШЕНИЯ ЗАДАЧИ
ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ**

Симплекс-метод — один из методов решения задач линейного программирования, встречающихся в ранее рассмотренных схемах нахождения $\delta \bar{u}^{(s)}$. Вместе с тем он имеет самостоятельное значение как метод решения задач оптимального проектирования с линейной целевой функцией и линейными ограничениями.

Рассмотрим общую схему этого метода, считая, что любая задача линейного программирования приведена к каноническому виду:

$$\min f, \quad f = \sum_{j=1}^{n_0} c_j x_j ; \quad (\text{П.2.1})$$

$$\left. \begin{aligned} \sum_{j=1}^{n_0} a_{ij} x_j &= b_i, \quad i=1,2,\dots,m_0, \\ x_j &\geq 0, \quad j=1,2,\dots,n_0, \end{aligned} \right\} \quad (\text{П.2.2})$$

где c_j, a_{ij}, b_i — известные коэффициенты.

В применении к (3.15) это можно сделать, вводя неотрицательные неизвестные x_j и заменяя любое ограничение типа неравенства на ограничение типа равенства. В результате вместо двойных неравенств $S_j^- \leq \delta u_j^{(s)} \leq S_j^+$ получим равенства:

$$\left. \begin{aligned} \delta u_j^{(s)} &= x_{n+j} + S_j^- \\ x_{n+j} + x_j &= S_j^+ - S_j^- \end{aligned} \right\} \quad j=1,2,\dots,n,$$

а условия: $\sum_{j=1}^n \mathcal{H}_j^{(i)} \delta u_j^{(s)} \leq F_i^{\text{дон}} - F_i(W^{(s)})$ приведем к виду:

$$\sum_{j=1}^n \mathcal{H}_j^{(i)} x_{n+j} + x_{2n+i} = F_i^{\text{дон}} - F_i(W^{(s)}) - \sum_{j=1}^n \mathcal{H}_j^{(i)} S_j^-.$$

Симплекс-метод является итерационным методом, который за конечное число итераций позволяет найти решение задачи (П.2.1), (П.2.2).

Рассмотрим одну из итераций, начинающуюся с того момента, когда какие-либо t неизвестных ($t \leq m_0$) выражены через остальные так, что свободные члены этих выражений неотрицательны. Оставляя пока в стороне вопрос о возможности такого преобразования, запишем условия (П.2.2) в виде

$$\left. \begin{aligned} x_i &= \beta_i - \sum_{l \in J_c} \alpha_{i,l} x_l, \quad i \in J_\delta; \\ \beta_i &\geq 0, \quad i \in J_\delta, \end{aligned} \right\} \quad (\text{П.2.3})$$

а линейную форму (П.2.1) в виде

$$f = \gamma_0 - \sum_{l \in J_c} \gamma_l x_l. \quad (\text{П.2.4})$$

Здесь: J_δ — множество t индексов из набора $j = 1, 2, \dots, n_0$; J_c — множество остальных $n_0 - t$ индексов из того же набора. Заметим, что неизвестные $x_i, i \in J_\delta$ принято называть базисными переменными, а неизвестные $x_i, i \in J_c$ — свободными, или небазисными переменными.

Из условий (П.2.3) легко получить одно из допустимых решений:

$$\left. \begin{aligned} x_i' &= \beta_i, \quad i \in J_\delta; \\ x_i' &= 0, \quad i \in J_c, \end{aligned} \right\} \quad (\text{П.2.5})$$

при котором линейная форма (П.2.4) примет значение $f' = \gamma_0$. Посмотрим, можно ли от решения (П.2.5) перейти к какому-либо другому допустимому решению x_i'' , $i = 1, 2, \dots, n_0$, при котором линейная форма f примет значение $f'' \leq f'$. С этой целью анализируем коэффициенты $\gamma_l, l \in J_c$.

Возможны два случая:

1) $\gamma_l \leq 0$ при всех $l \in J_c$. Решение (П.2.5) является оптимальным решением задачи (П.2.1), (П.2.2);

2) среди $\gamma_l, l \in J_c$, имеется хотя бы один положительный коэффициент. Пусть это будет $\gamma_q > 0$, где q — один из индексов множества J_c . Тогда можно попытаться уменьшить значение f , увеличивая x_q и не меняя остальных небазисных переменных.

Поскольку все $x_l = 0$ при $l \in J_c$ и $l \neq q$, то равенства (П.2.3) можно записать в виде:

$$x_i = \beta_i - \alpha_{i,q} x_q, \quad i \in J_\delta. \quad (\text{П.2.6})$$

Здесь также возможны два случая:

2а) $\alpha_{i,q} \leq 0$ при всех $i \in J_\delta$. Можно неограниченно увеличивать x_q , не нарушая условий (П.2.2); при этом $f \rightarrow -\infty$. (Заметим, что в первом алгоритме метода линейаризации всегда рассматриваются задачи линейного программирования с ограниченными областями изменения $\delta u_j^{(s)}$. Поэтому случай (2а) при решении подобных задач не встречается);

2б) по крайней мере один или несколько коэффициентов $\alpha_{s,q} > 0$, где $s \in J_\delta$. Тогда x_q можно увеличивать лишь до значения x_q'' :

$$x_q'' = \min_{\alpha_{s,q} > 0} \frac{\beta_s}{\alpha_{s,q}} = \frac{\beta_k}{\alpha_{k,q}}, \quad (\text{П.2.7})$$

где $k \in J_\delta$ — тот индекс, при котором достигается минимум отношения (П.2.7). В этом случае получим новое допустимое решение $x_i'', i=1, \dots, n_0$:

$$\left. \begin{aligned} x_i'' &= \beta_i - \alpha_{i,q} x_q'', \quad i \in J_\delta, \quad i \neq k; \quad x_k'' = 0; \\ x_i'' &= 0, \quad i \in J_c, \quad i \neq q; \quad x_q'' = \frac{\beta_k}{\alpha_{k,q}}, \end{aligned} \right\} \quad (\text{П.2.8})$$

при котором линейная форма f равна f'' : $f'' = \gamma_0 - \gamma_q x_q'' \leq f'$.

Теперь преобразуем систему равенств (П.2.3), выразив из уравнения с номером k переменную x_q через остальные переменные:

$$x_q = \frac{\beta_k}{\alpha_{k,q}} - \frac{x_k}{\alpha_{k,q}} - \sum_{\substack{l \in J_c \\ l \neq k}} \frac{\alpha_{k,l}}{\alpha_{k,q}} x_l. \quad (\text{П.2.9})$$

Подставляя x_q в виде (П.2.9) в остальные уравнения (П.2.3), приведем их к виду:

$$x_i = \beta_i' - \sum_{l \in J_c'} \alpha_{i,l}' x_l, \quad i \in J_\delta', \quad (\text{П.2.10})$$

где $J_\delta' = (J_\delta/k) \cup q$ — совокупность индексов новых базисных переменных; $J_c' = (J_c/q) \cup k$ — совокупность индексов новых небазисных переменных. При этом $\beta_i' \geq 0, i \in J_\delta'$, что легко показать, используя (П.2.7).

Получением (П.2.10) заканчивается одна итерация симплекс-метода. Итерационный счет продолжается до тех пор, пока одна из итераций не завершится случаем 1.

Вернемся теперь к вопросу о приведении системы (П.2.2) к виду (П.2.3). Будем считать, что все коэффициенты $b_i \geq 0$, $i=1, \dots, m_0$ (если это не так, то обе части соответствующего равенства в системе (П.2.2) надо умножить на -1). Введем в рассмотрение новые переменные $z_i = b_i - \sum_{j=1}^{n_0} a_{ij} x_j$, $i=1, 2, \dots, m_0$, имеющие смысл невязок в (П.2.2), если заданы какие-либо значения x_j , $j=1, 2, \dots, n_0$.

Сформулируем новую задачу линейного программирования:

$$\left. \begin{aligned} \min h, \quad h &= \sum_{i=1}^{m_0} z_i ; \\ z_i &= b_i - \sum_{j=1}^{n_0} a_{ij} x_j, \quad i=1, 2, \dots, m_0 ; \\ z_i &\geq 0, \quad i=1, 2, \dots, m_0 ; \\ x_j &\geq 0, \quad j=1, 2, \dots, n_0 . \end{aligned} \right\} \quad (\text{П.2.11})$$

В этой задаче одно из допустимых решений легко получить, положив: $z_i = b_i$, $i=1, 2, \dots, m_0$; $x_j = 0$, $j=1, 2, \dots, n_0$. Поэтому следующие допустимые и оптимальное решения можно найти, используя изложенную выше итерационную схему.

При решении задачи (П.2.11) возможны два случая:

1) $\min h \neq 0$. Это означает, что не существует допустимого решения задачи (П.2.1), (П.2.2);

2) $\min h = 0$, что возможно, когда все $z_i = 0$, $i=1, 2, \dots, m_0$. Вспоминая ранее рассмотренную схему, можно заключить, что все z_i при этом окажутся в разряде небазисных переменных. Поэтому равенства в (П.2.11) примут вид:

$$x_i = \beta_i - \sum_{l \in J_c} \alpha_{il} x_l - \sum_{j=1}^{m_0} \xi_{ij} z_j, \quad i \in J_d,$$

от которых легко перейти к равенствам (П.2.3), исключив сумму $\sum_{j=1}^{m_0} \xi_{ij} z_j$.

При этом будет найдено также допустимое решение (П.2.5) задачи. (П.2.1), (П.2.2).

МЕТОД СОПРЯЖЕННЫХ НАПРАВЛЕНИЙ

Метод сопряженных направлений (или сопряженных градиентов) применим для минимизации непрерывно дифференцируемых функций $F(\vec{u})$ при отсутствии ограничений на компоненты u_1, u_2, \dots, u_n вектора \vec{u} . Он относится к классу итерационных методов, использующих знание лишь первых производных $\partial F / \partial u_j$ (или градиентов ∇F) функции $F(\vec{u})$. Рассмотрим особенности и алгоритм метода на примере минимизации квадратичной формы

$$F(\vec{u}) = \frac{1}{2} (\vec{u}, \hat{C} \vec{u}) + (\vec{d}, \vec{u}) + a \quad (\text{П.3.1})$$

с положительно определенной симметричной матрицей \hat{C} . Как известно [15], в этом случае метод сопряженных направлений становится конечным и позволяет получить $\min F$ не более чем за n итераций (n — число неизвестных).

На каждой k -й итерации метода переход от вектора $\vec{u}^{(k)}$ к вектору $\vec{u}^{(k+1)}$ осуществляется вдоль направления $\vec{V}^{(k)}$:

$$\vec{u}^{(k+1)} = \vec{u}^{(k)} + \alpha_k \vec{V}^{(k)}, \quad (\text{П.3.2})$$

выбираемого в виде линейной комбинации антиградиента $-\nabla F(\vec{u}^{(k)})$ и предыдущего направления $\vec{V}^{(k-1)}$:

$$\vec{V}^{(k)} = -\nabla F(\vec{u}^{(k)}) + \beta_k \vec{V}^{(k-1)}. \quad (\text{П.3.3})$$

Лишь на первой итерации $\vec{V}^{(0)} = -\nabla F(\vec{u}^{(0)})$, где $\vec{u}^{(0)}$ — начальный вектор. При этом коэффициент β_k выбирается так, чтобы направления $\vec{V}^{(k-1)}$ и $\vec{V}^{(k)}$ были сопряжены по отношению к матрице \hat{C} , т.е. из равенства

$$(\vec{V}^{(k)}, \hat{C} \vec{V}^{(k-1)}) = 0, \quad (\text{П.3.4})$$

а коэффициент α_k находится из условия

$$F(\vec{u}^{(k)} + \alpha_k \vec{V}^{(k)}) = \min_{\alpha} F(\vec{u}^{(k)} + \alpha \vec{V}^{(k)}). \quad (\text{П.3.5})$$

Поскольку $F(\vec{u}^{(k)} + \alpha \vec{V}^{(k)}) = \frac{1}{2} \alpha^2 (\vec{V}^{(k)}, \hat{C} \vec{V}^{(k)}) + \alpha (\nabla F(\vec{u}^{(k)}, \vec{V}^{(k)}) + F(\vec{u}^{(k)})$, то при определении α_k возможны случаи:

а) если $(\vec{V}^{(k)}, \hat{C} \vec{V}^{(k)}) \neq 0$, то

$$\alpha_k = - \frac{(\vec{V}^{(k)}, \nabla F(\vec{u}^{(k)}))}{(\vec{V}^{(k)}, \hat{C} \vec{V}^{(k)})}, \quad (\text{П.3.6})$$

б) если $(\vec{V}^{(k)}, \hat{C} \vec{V}^{(k)}) = 0$, а $(\vec{V}^{(k)}, \nabla F(\vec{u}^{(k)})) \neq 0$, то

$$\alpha_k = +\infty \quad (\text{П.3.7})$$

и минимальное значение $F(\vec{u})$ не достигается. В этом случае квадратичная форма неограничена снизу.

Критерием выхода из итераций служит выполнение с желательной точностью равенства

$$\nabla F(\vec{u}^*) = 0. \quad (\text{П.3.8})$$

Принятый в методе способ построения сопряженных направлений $\vec{V}^{(k)}$, $k=0, 1, \dots$, приводит к тому, что векторы $\nabla F(\vec{u}^{(k)})$ и $\nabla F(\vec{u}^{(k-1)})$ оказываются ортогональными. В результате расчет по (П.3.6) всегда дает

$$\alpha_k > 0, \quad (\vec{V}^{(k)}, \nabla F(\vec{u}^{(k)})) \leq 0, \quad k=0, 1, \dots \quad (\text{П.3.9})$$

Схема решения задачи на отыскание безусловного минимума функции $F(\vec{u})$ (П.3.1) представлена на рис. П.2. При этом использованы следующие обозначения: k — номер итерации, $\vec{u}^{(0)}$ — первоначально заданный вектор \vec{u} , $\|\nabla F(\vec{u}^{(k)})\| = \sqrt{\sum_{j=1}^n (\partial F / \partial u_j |_{\vec{u}^{(k)}})^2}$ — норма вектора $\nabla F(\vec{u}^{(k)})$.

Замечания.

1. Вычислительный опыт показывает, что для плохо обусловленных матриц \hat{C} (т.е. таких невырожденных матриц, у которых отношение максимального собственного числа к минимальному очень велико) в рассмотренном выше итерационном процессе накапливается погрешность вычислений. Поэтому, если $\|\nabla F(\vec{u}^{(k)})\|$ не равна нулю с достаточной точностью ϵ , то необходимо повторить вычисления, беря \vec{u}^* в качестве начального вектора.

2. Представленный на рис. П.2 алгоритм применим к минимизации неквадратичных непрерывно дифференцируемых функций при следующих изменениях:

а) коэффициент α_k следует определять из условия (П.3.5), используя какой-либо метод одномерного поиска [15];

б) расчет коэффициентов β_k ($k > 0$) проводить по формуле:

$$\beta_k = \frac{(\nabla F(\vec{u}^{(k)}), \nabla F(\vec{u}^{(k)}) - \nabla F(\vec{u}^{(k-1)}))}{(\nabla F(\vec{u}^{(k-1)}), \nabla F(\vec{u}^{(k-1)}))} \quad (\text{П.3.10})$$

Кроме того, после проведения π итераций рекомендуется следующий цикл из π итераций выполнять, принимая $\vec{u}^{(\pi+1)}$ в качестве начального вектора и полагая на первой итерации этого цикла $k=0$.

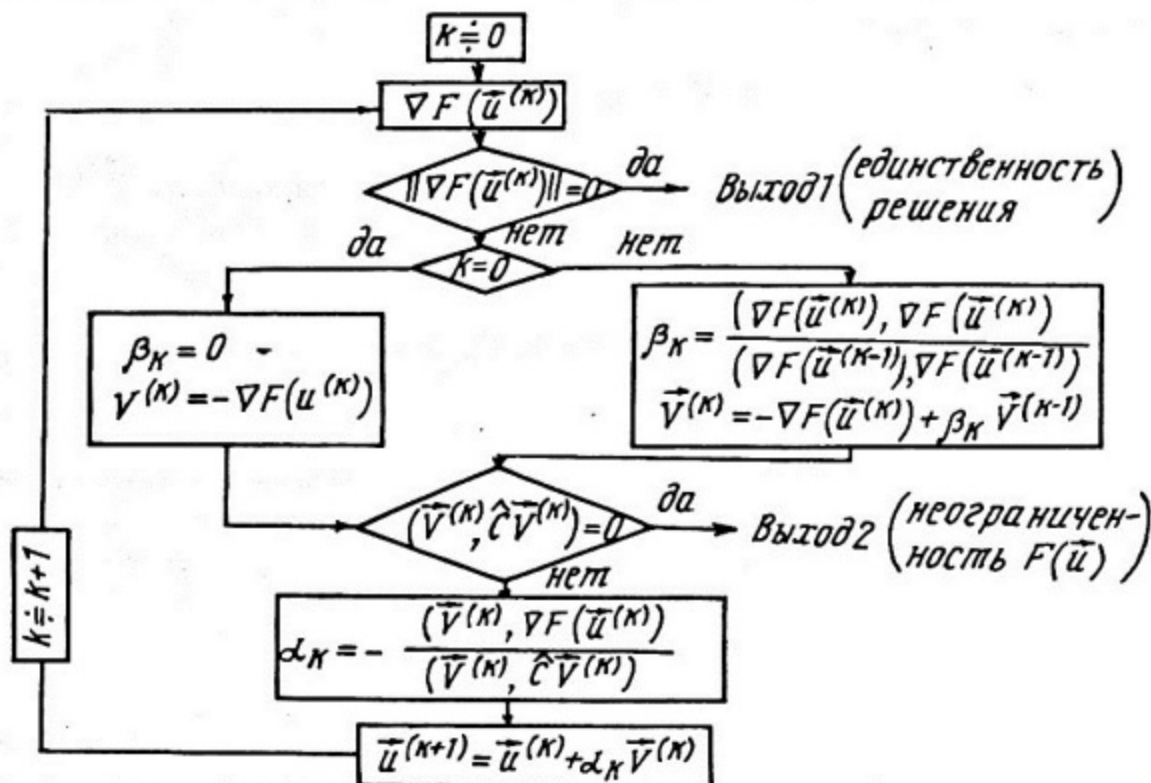


Рис. П.2. Схема безусловной минимизации по методу сопряженных направлений

ДВОЙСТВЕННЫЙ МЕТОД РЕШЕНИЯ ЗАДАЧИ
КВАДРАТИЧНОГО ПРОГРАММИРОВАНИЯ

Сформулированную в разд. 3 задачу квадратичного программирования (3.18), (3.19) можно записать в виде:

$$\left. \begin{aligned} \min f(\vec{x}), \text{ где } f(\vec{x}) &= \frac{1}{2} (\vec{x}, \vec{x}) + (\vec{h}^{(0)}, \vec{x}); \\ (\vec{h}^{(i)}, \vec{x}) &\leq b_i, \quad i = 1, 2, \dots, m_0, \end{aligned} \right\} \quad (\text{П.4.1})$$

если принять, что $m_0 = 2n + m$, \vec{x} — n -мерный вектор с неизвестными компонентами $x_j = v_j$, $j = 1, 2, \dots, n$, а величины b_i ($i = 1, 2, \dots, m_0$) и компоненты $\vec{h}_j^{(i)}$ векторов $\vec{h}^{(i)}$ связаны с коэффициентами задачи (3.18), (3.19) соотношениями:

$$\begin{aligned} h_j^{(0)} &= H_j^{(0)}, \quad j = 1, 2, \dots, n; \\ h_j^{(i)} &= H_j^{(i)}, \quad b_i = F_i^{\text{дон}} - F_i(W^{(s)}), \quad i = 1, 2, \dots, m; \end{aligned}$$

$$h_j^{(i)} = \begin{cases} 1, & j = i - m, \\ 0, & j \neq i - m, \end{cases} \quad b_i = u_{i-m}^{\text{макс}} - u_{i-m}^{(s)}, \quad i = m+1, \dots, m+n;$$

$$h_j^{(i)} = \begin{cases} -1, & j = i - m - n, \\ 0, & j \neq i - m - n, \end{cases} \quad b_i = u_{i-m-n}^{(s)} - u_{i-m-n}^{\text{мин}}, \quad i = m+n+1, \dots, m+2n.$$

Вопрос о разрешимости задачи (П.4.1) формулируется следующей теоремой [16].

В задаче квадратичного программирования минимум либо достигается в некоторой точке, либо функция $f(x)$ неограничена снизу.

Рассмотрим один из методов [16] решения задачи (П.4.1), использующий основные соотношения теории двойственности между неизвестными x_j и двойственными переменными λ_i , $i = 1, 2, \dots, m_0$. Предпочтение этому методу отдается потому, что, во-первых, он позволяет ограничиться решением более простой по сравнению с (П.4.1) задачей, и, во-вторых, наряду с неизвестными x_j позволяет найти двойственные переменные λ_i , необходимые для оценки параметра N в методе линеаризации.

Для формулировки двойственной по отношению к (П.4.1) задачи воспользуемся правилом, изложенным в приложении 1. Введем в рассмотрение неотрицательные двойственные переменные $\lambda_i \geq 0$, $i=1, 2, \dots, m_0$ и составим функцию Лагранжа:

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = \frac{1}{2}(\vec{x}, \vec{x}) + (\vec{h}^{(0)}, \vec{x}) + \sum_{i=1}^{m_0} \lambda_i [(\vec{h}^{(i)}, \vec{x}) - b_i]. \quad (\text{П.4.2})$$

Поскольку в функцию $\mathcal{L}(\vec{x}, \vec{\lambda})$ включены все ограничения задачи (П.4.1), то для построения двойственной функции минимум $\mathcal{L}(\vec{x}, \vec{\lambda})$ следует искать среди всех $\vec{x} \in E_n$, т.е. из условий $\partial \mathcal{L} / \partial x_j = 0$, $j=1, 2, \dots, n$. Эти условия приводят к равенству

$$\vec{x} = -\vec{h}^{(0)} - \sum_{i=1}^{m_0} \lambda_i \vec{h}^{(i)}, \quad (\text{П.4.3})$$

устанавливающему связь между неизвестными x_j и λ_i . Используя (П.4.3), получим:

$$\min_{\vec{x}} \mathcal{L}(\vec{x}, \vec{\lambda}) = -\frac{1}{2} \left(\sum_{i=1}^{m_0} \vec{h}^{(i)} \lambda_i, \sum_{i=1}^{m_0} \vec{h}^{(i)} \lambda_i \right) - \sum_{i=1}^{m_0} (\vec{h}^{(0)}, \vec{h}^{(i)}) \lambda_i - \frac{1}{2} (\vec{h}^{(0)}, \vec{h}^{(0)}) + \sum_{i=1}^{m_0} b_i \lambda_i. \quad (\text{П.4.4})$$

Теперь двойственную задачу можно записать в виде:

$$\min_{\vec{\lambda} \geq 0} g(\vec{\lambda}), \quad \text{где} \quad g(\vec{\lambda}) = \frac{1}{2}(\vec{\lambda}, \hat{C}\vec{\lambda}) + (\vec{d}, \vec{\lambda}) + a, \quad (\text{П.4.5})$$

где $\vec{\lambda}$ — m_0 -мерный вектор с компонентами $\lambda_i \geq 0$, $i=1, 2, \dots, m_0$; \hat{C} — симметричная положительно определенная матрица размерности $m_0 \times m_0$ с элементами $C_{ik} = (\vec{h}^{(i)}, \vec{h}^{(k)})$; \vec{d} — вектор с компонентами $d_i = (\vec{h}^{(0)}, \vec{h}^{(i)}) + b_i$; $a = -\frac{1}{2}(\vec{h}^{(0)}, \vec{h}^{(0)})$.

Таким образом, вместо (П.4.1) приходим к задаче квадратичного программирования (П.4.5) с простыми ограничениями, которая решается с помощью следующего двухслойного итерационного процесса. Ниже изложена последовательность рассуждений для одной из внешних итераций.

Шаг 1 любой внешней итерации начинается с проверки оптимальности заданного или полученного на предыдущей итерации вектора $\vec{\lambda}^{(0)}$. Очевидно, что $\vec{\lambda}^{(0)}$ будет решением задачи (П.4.5), если:

$$\left. \begin{aligned} \frac{\partial q}{\partial \lambda_i} \Big|_{\vec{\lambda}^{(0)}} = 0 & \text{ при тех } i \in I, \text{ для которых } \lambda_i^{(0)} > 0, \\ \frac{\partial q}{\partial \lambda_i} \Big|_{\vec{\lambda}^{(0)}} \geq 0 & \text{ для тех } i \in I, \text{ при которых } \lambda_i^{(0)} = 0, \end{aligned} \right\} \quad (\text{П.4.6})$$

где I — множество индексов $i=1, 2, \dots, m_0$. В противном случае можно понизить $q(\vec{\lambda})$, переходя к шагу 2.

Шаг 2 внешней итерации заключается в разбиении множества индексов I на два подмножества $I_a^{(0)}$ и $I_p^{(0)}$. В подмножество активных переменных $I_a^{(0)}$ включаются те индексы $i \in I^{(0)}$, для которых:

$$\left. \begin{aligned} \text{либо } \lambda_i^{(0)} > 0, \\ \text{либо } \lambda_i^{(0)} = 0, \text{ но } \frac{\partial q}{\partial \lambda_i} \Big|_{\vec{\lambda}^{(0)}} < 0. \end{aligned} \right\} \quad (\text{П.4.7})$$

Остальные индексы относятся к подмножеству пассивных переменных:
 $I_p^{(0)} = I^{(0)} \setminus I_a^{(0)}$.

Шаг 3 внешней итерации включает внутренние итерации по индексу $k=0, 1, \dots, q$ (где $q \leq m_0$) и связан с минимизацией функции $q(\vec{\lambda})$ по переменным $\lambda_i \geq 0$, $i \in I_a^{(k)}$ с помощью изложенного в приложении 3 метода сопряженных направлений. Поскольку расчет по (П.3.2) $\lambda_i^{(k+1)} = \lambda_i^{(k)} + \alpha_k V_i^{(k)}$ может привести к отрицательным значениям тех i -х компонент $\lambda_i^{(k+1)}$, $i \in I_a^{(k)}$, для которых $V_i^{(k)} < 0$, то предварительно находится значение

$$\bar{\alpha}_k = \min_{\substack{i \in I_a^{(k)} \\ V_i^{(k)} < 0}} \left(-\frac{\lambda_i^{(k)}}{V_i^{(k)}} \right) = -\frac{\lambda_l^{(k)}}{V_l^{(k)}}, \quad l \in L^{(k)} \subset I_a^{(k)}, \quad (\text{П.4.8})$$

где $L^{(k)}$ — множество тех индексов, при которых достигается указанный минимум. После этого вычисляется $\lambda_i^{(k+1)}$:

$$\left. \begin{aligned} \lambda_i^{(k+1)} &= \lambda_i^{(k)} + \alpha_k V_i^{(k)}, & \text{если } \alpha_k < \bar{\alpha}_k; \\ \lambda_i^{(k+1)} &= \lambda_i^{(k)} + \bar{\alpha}_k V_i^{(k)}, & \text{если } \alpha_k \geq \bar{\alpha}_k. \end{aligned} \right\} \quad (\text{П.4.9})$$

В последнем случае (при $\alpha_k \geq \bar{\alpha}_k$) $\lambda_i^{(k+1)} = 0$, $i \in L^{(k)}$ и происходит перестройка подмножеств $I_a^{(k)}$ и $I_p^{(k)}$ в соответствии с равенствами:

$$I_a^{(k+1)} = I_a^{(k)} \setminus L^{(k)}, \quad I_p^{(k+1)} = I_p^{(k)} \cup L^{(k)}. \quad (\text{П.4.10})$$

Процесс итераций в методе сопряженных направлений прекращается в двух случаях. В первом случае при каком-то $I_a^{(k)}$ будет получен с заданной точностью ϵ $\min q(\vec{\lambda})$ по переменным λ_i , $i \in I_a^{(k)}$ (выход 1 на рис. П.2). Тогда полученное решение принимается за исходное и происходит возврат к шагу 1 алгоритма. В другом случае (выход 2 на рис. П.2) будет установлена неограниченность квадратичной формы $q(\vec{\lambda})$ и решение задачи (П.4.5) прекращается. В соответствии с известными теоремами двойственности это означает неограниченность снизу функции $f(\vec{x})$.

Сходимость изложенного выше итерационного процесса решения задачи (П.4.5) обусловлена тем, что функция $q(\vec{\lambda})$ в процессе итераций постоянно уменьшается, а подмножество индексов $I_p^{(k)}$ расширяется. Поэтому после конечного числа итераций либо будет получено неотрицательное решение $\vec{\lambda}^*$, удовлетворяющее условиям (П.4.6), либо будет установлена неограниченность снизу функции $q(\vec{\lambda})$. В случае конечных значений λ_i^* , $i = 1, 2, \dots, m_0$ решение задачи (П.4.1) находится по формуле (П.4.3).

ИСПОЛЬЗОВАННАЯ ЛИТЕРАТУРА

1. Вычислительные методы выбора оптимальных проектных решений./Под ред. В.С. Михалевича – Киев: Наукова думка, 1977.
2. Фейнберг С.М., Шихов С.Б., Троянский В.Б. Теория ядерных реакторов. Т. 1. – М.: Атомиздат, 1978.
3. Михеев М.А., Михеева И.М. Основы теплопередачи. – М.: Энергия, 1973.
4. Шишков Л.К. Методы решения диффузионных уравнений двумерного ядерного реактора. – М.: Атомиздат, 1976.
5. Фаддеев Д.К., Фаддеева В.Н. Вычислительные методы линейной алгебры. – М.-Л.: Физматгиз, 1963.
6. Марчук Г.И., Лебедев В.И. Численные методы в теории переноса нейтронов. – М.: Атомиздат, 1981.
7. Земсков Е.А., Исакова Л.Я. О применении прямых методов в нейтронно-физическом расчете реакторов. – Атомная энергия, 1982, т. 52, вып. 5, с. 316–320.
8. Хромов В.В. Итерационный синтез уравнений переноса нейтронов. – В сб.: Физика ядерных реакторов. Вып. 3. М., 1973, с. 5–12.
9. Хромов В.В., Кузьмин А.М., Орлов В.В. Метод последовательной линеаризации в задачах оптимизации реакторов на быстрых нейтронах. – М.: Атомиздат, 1978.
10. Хромов В.В. Вариационный формализм полиномиальной аппроксимации функционалов возмущенного реактора. – В сб.: Физика ядерных реакторов. Вып. 6. М., 1978, с. 25–30.
11. Марчук Г.И. Методы вычислительной математики. – М.: Наука, 1977.
12. Налимов В.В., Чернова Н.А. Статистические методы планирования эксперимента. – М.: Наука, 1973.
13. Усынин Г.Б., Карабасов А.С., Чирков В.А. Оптимизационные модели реакторов на быстрых нейтронах. – М.: Атомиздат, 1981.
14. Федоренко Р.П. Приближенное решение задач оптимального управления. – М.: Наука, 1978.
15. Моисеев Н.Н., Иванилова Ю.П., Столярова Е.М. Методы оптимизации. – М.: Наука, 1978.
16. Пшеничный Б.Н. Метод линеаризации. – М.: Наука, 1983.
17. Артамкин В.Н. и др. Оптимальный режим остановки реактора. – Атомная энергия, 1964, т. 17, вып. 3, с. 189–193.
18. Кузьмин А.М., Силаев Ю.В., Орлов В.В., Хромов В.В. Оптимизационные исследования реакторов на быстрых нейтронах – Атомная энергия, 1971, т. 31, вып. 2, с. 83–88.
19. Юдин Д.Б., Гольштейн Е.Г. Линейное программирование. – М.: Физматгиз, 1963.
20. Казачковский О.Д. Развитие работ по быстрым реакторам в СССР. – Атомная энергия, 1974, т. 36, вып. 6, с. 444–449.

21. Зарицкий С.М. О некоторых возможных формулах теории возмущений. – В сб.: Физика ядерных реакторов. Вып. 1. М., 1968, с. 39–54.
22. Фиакко А., Мак–Кормик Г. Нелинейное программирование. – М.: Наука, 1972.
23. Кузьмин А.М. Исследование устойчивости решения оптимизационных задач с ограничениями. – В сб.: Физика ядерных реакторов. Вып. 7. М., 1978, с. 28–40.
24. Ермолов Ю.М. Методы стохастического программирования. – М.: Наука, 1976.
25. Вентцель Е.С. Теория вероятностей. – М.: Наука, 1969.
26. Кузьмин А.М. Коэффициенты чувствительности оптимизируемых характеристик АЭС к технологическим параметрам в задачах с ограничениями. – Атомная энергия, 1975, т. 39, вып. 4, с. 269–272.
27. Качалов В.А., Кузьмин А.М. Допустимые погрешности исходных данных в оптимизационных задачах. – В сб.: Физика ядерных реакторов. Вып. 6. М., 1978, с. 48–56.
28. Методические положения по выполнению оптимизационных (технико-экономических) расчетов в энергетике при неоднозначности исходной информации./ Под ред. акад. Л.А. Мелентьева. – Москва–Иркутск, 1977.
29. Соболев И.М., Статников Р.Б. Выбор оптимальных параметров в задачах со многими критериями. – М.: Наука, 1981.
30. Беляев Б.И. и др. Отбор заданного числа точек, равномерно расположенных в n -мерном кубе. – В сб.: Учет неопределенности исходной информации. Иркутск, 1973, с. 78–83.

СОДЕРЖАНИЕ

Введение	3
1. Формулировка задачи оптимального проектирования	4
2. Математические модели	7
2.1. Сеточные методы	9
2.2. Синтетические методы	10
2.3. Полиномиальные аппроксимации функционалов	12
3. Автоматизация поиска оптимального управления	15
3.1. Линеаризация задачи	16
3.2. Построение минимизирующей последовательности	20
3.3. Необходимые условия оптимальности	24
3.4. Выравнивание поля тепловыделения	29
4. Учет неопределенности исходной информации	39
4.1. Исследование устойчивости решения	41
4.2. Методы стохастического программирования	49
4.3. Задачи с неопределенными исходными данными	56
Приложение 1. Общие сведения по нелинейному и выпуклому програм- мированию	62
Приложение 2. Симплекс-метод решения задачи линейного програм- мирования	66
Приложение 3. Метод сопряженных направлений	70
Приложение 4. Двойственный метод решения задачи квадратичного программирования	73
Использованная литература	77

Кузьмин Анатолий Михайлович

МЕТОДЫ ОПТИМИЗАЦИИ
ЯДЕРНО-ЭНЕРГЕТИЧЕСКИХ УСТАНОВОК

Редактор Кочубей Е.Н.
Техн. редактор Н.М. Воронцова
Корректор П.Н. Маркина

Тем. план 1985 г., поз. 12 (87)

Л.-97539

Формат 60x84 1/16

Тираж 200 экз.

Подписано в печать 11/х - 1985г.

Объем 5 п.л.

Цена 30 коп.

Заказ 755

Уч.-изд.л. 4,75

Изд. № 039-1

Типография МИФИ, Каширское шоссе, 31