Министерство образования Российской Федерации

САНКТ-ПЕТЕРБУРГСКИЙ

ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

В.Б. МЕЛАС

ЛОКАЛЬНО ОПТИМАЛЬНЫЕ ПЛАНЫ ЭКСПЕРИМЕНТА

Учебное пособие

Санкт-Петербург
Издательство СПбГТУ
1999

Предисловие

Учебное пособие посвящено исследованию локально-оптимальных планов для нелинейных по параметрам регрессионных моделей, в особенности, для дробно-рациональных моделей.

Основы математической теории оптимального планирования регрессионных экспериментов были заложены в работах Дж. Кифера в конце пятидесятых годов. В ее развитие существенный вклад внесли российские математики (см. Ермаков, Жиглявский, 1987; Мелас, 1999).

В этой теории под планом эксперимента понимается дискретная вероятностная мера. Оптимальным считается план, доставляющий экстремальное значение некоторой функции, заданной на множестве информационных матриц и имеющей строгий статистический смысл. Одной из таких функций является определитель матрицы.

В рамках этой теории получен ряд фундаментальных результатов для линейных по параметрам регрессионных моделей.

Нелинейные регрессионные модели находят широкое применение в эконометрике и других актуальных областях исследования. В последние годы им посвящено значительное число монографий и журнальных статей. Однако задачи оптимального планирования эксперимента для таких моделей до сих пор мало изучены, а локально-оптимальные планы рассматривались лишь в нескольких статьях и, в основном, для случаев, когда они могут быть построены в явном виде. Исключение представляет собой статья (Мелас, 1981), в которой точки локально оптимальных планов для экспоненциальной регрессии изучались как неявно заданные вектор-функции истинных значений нелинейно входящих параметров. Такой подход можно назвать функциональным подходом. В настоящем учебном пособии этот подход систематически развивается и применяется к исследованию дробно-рациональных моделей.

Пособие имеет следующую структуру.

В главе 1 излагаются основные факты, относящиеся к нелинейному регрессионному анализу и локально оптимальным планам эксперимента. Вторая глава содержит изложение разработанного автором учебного пособия функционального подхода. Третья глава содержит результаты по построению и исследованию локально-оптимальных планов для дробно-рациональной регрессии.

Работа поддержана Российским Фондом Фундаментальных исследований (грант No 98-01-00342).

Глава 1

Нелинейный регрессионный анализ

В настоящей главе мы изложим некоторые основные результаты, относящиеся к изучению нелинейных по параметрам моделей общего вида.

1.1 Нелинейная (по параметрам) регрессионная модель

Пусть N — заданное натуральное число. Пусть результаты эксперимента $y_1, \ldots, y_N \in \mathbb{R}^1$ описываются уравнением

$$y_i = \eta(x_j, \Theta) + \varepsilon_j, \tag{1.1}$$

где $x_1, \ldots, x_N \in \mathfrak{X}$, $\mathfrak{X}$ — некоторое множество, относительно которого мы сделаем некоторые предположения в дальнейшем,

$\eta(x, \Theta)$ — функция, известная с точностью до вектора параметров $\Theta = (\theta_1, \ldots, \theta_m)^T$,

ε_j — случайные ошибки — некоррелированные случайные величины такие, что

$$E\varepsilon_j = 0, \quad \mathcal{D}\varepsilon_j = \sigma^2. \tag{1.2}$$

Цель эксперимента — оценка вектора параметров Θ. Будем говорить, что параметр θ_j, $j \in 1 : m$ входит в модель (1.1) нелинейным образом, если при фиксированном x функция

$$\frac{\partial \eta(x, \Theta)}{\partial \theta_j}$$

существует и не является постоянной. Если эта функция является постоянной, то будем говорить, что соответствующий параметр входит в модель линейным образом. Функцию регрессии $\eta(x,\Theta)$ будем называть нелинейной (по параметрам), если хотя бы один параметр θ_j входит в модель (1.1) нелинейным образом.

Необходимо выбрать экспериментальные условия $x_1,\ldots,x_N$ и метод оценивания вектора параметров.

Введем два понятия плана эксперимента.

Под *точным планом эксперимента* мы будем понимать набор N точек, не обязательно различных, из заданного множества $\mathfrak{X}$. Под *приближенным планом эксперимента* будем понимать дискретную вероятностную меру, задаваемую таблицей

$$\xi = \{x_1,\ldots,x_n;\mu_1,\ldots,\mu_n\},$$

где $x_i \neq x_j$ $(i \neq j)$, $x_i \in \mathfrak{X}$, $\mu_i > 0$, $i,j = 1,\ldots,n$, $\sum_{i=1}^n \mu_i = 1$, n — произвольное натуральное число.

Понятие приближенного плана эксперимента было введено Дж. Кифером. Оно существенно облегчает математические исследования. При практическом использовании приближенных планов рекомендуют при условиях x_j проводить приблизительно $\mu_j N$ экспериментов, $j = 1,\ldots,n$.

Наилучший (в некотором точно определенном смысле) выбор плана — задача, которую мы рассмотрим в разделе 1.3.

При фиксированном (выбранном некоторым способом) плане в качестве метода оценивания будем использовать (нелинейный) метод наименьших квадратов (МНК), введенный Гауссом. По определению МНК-оценка есть вектор $\hat{\Theta}$, являющийся решением экстремальной задачи

$$\sum_{j=1}^{N} \left(\eta(x_j,\Theta) - y_j\right)^2 \to \min_{\Theta\in\mathbb{R}^m}.$$

МНК–оценки обладают замечательными асимптотическими свойствами, описанными в следующем разделе.

1.2 Асимптотические свойства МНК-оценок

Пусть Ω — ограниченное замкнутое множество в $\mathbb{R}^m$, $\mathfrak{X}$ — ограниченное замкнутое множество в $\mathbb{R}^k$, где k — некоторое натуральное число.

Пусть функция регрессии $\eta(x,\Theta)$ нелинейна по параметрам и определена при всех $x \in \mathfrak{X}$, $\Theta \in \Omega$. Через Θ_u будем обозначать истинное значение вектора параметров, т. е. такое значение Θ, при котором верна модель (1.1).

Далее под планом будет понимать приближенный план, если не оговорено противное. Для дискретных мер $\xi = \{x_1,\ldots,x_n;\mu_1,\ldots,\mu_n\}$ будем использовать

запись
$$\int_{\mathfrak{X}} g(x)\xi(dx) = \sum_{i=1}^{n} g(x_i)\mu_i,$$
где g — произвольная функция.

Введем следующие предположения:

а) функция $\eta(x, \Theta)$ непрерывна на $\mathfrak{X} \times \Omega$;

б) последовательность планов $\{\xi_N\}$ слабо сходится к плану ξ, т. е. для любой непрерывной функции $g(x)$ на $\mathfrak{X}$ при $N \to \infty$ имеет место соотношение
$$\int_{\mathfrak{X}} g(x)\xi_N(dx) \to \int_{\mathfrak{X}} g(x)\xi(dx);$$

в) величина
$$\int_{\mathfrak{X}} [\eta(x, \Theta) - \eta(x, \bar{\Theta})]^2 \xi(dx)$$
при $\bar{\Theta}$, $\Theta \subset \Omega$ равна нулю только при $\Theta = \bar{\Theta}$;

г) производные
$$\partial\eta/\partial\theta_i, \quad \partial^2\eta/\partial\theta_i\partial\theta_j, \quad i, j = 1, \ldots, m$$
существуют и непрерывны на $\mathfrak{X} \times \Omega$;

д) $\hat{\Theta}_u$ — внутренняя точка Ω и матрица
$$M(\xi, \Theta) = \int_{\mathfrak{X}} f(x, \Theta) f^T(x, \Theta)\xi(dx),$$
где
$$f^T(x, \Theta) = \left(\frac{\partial\eta(x, \Theta)}{\partial\theta_1}, \ldots, \frac{\partial\eta(x, \Theta)}{\partial\theta_m}\right)$$
невырождена при $\Theta = \Theta_u$.

Пусть ξ_N имеет вид
$$\{x_1, \ldots, x_N; \frac{1}{N}, \ldots \frac{1}{N}\},$$
где x_i — не обязательно различные точки,
$$\hat{\Theta}_N = \arg\min_{\Theta\in\Omega} \sum_{i=1}^{N} (\eta(x_i, \Theta) - y_i)^2. \tag{1.3}$$

Теорема 1.2.1. *Если случайные ошибки одинаково распределены, результаты экспериментов описываются уравнением (1.1)–(1.2) и выполнены предположения а)–в), то последовательность МНК-оценок сильно состоятельна, т. е. при $N \to \infty$*

$$\hat{\Theta}_{(N)} \to \Theta_u$$

с вероятностью 1, где $\hat{\Theta}_N$ определено формулой (1.3). Если, кроме того, выполняются предположения г) и д), то при $N \to \infty$ последовательность распределений случайных векторов $\sqrt{N}(\hat{\Theta}_N - \Theta_u)$ сходится к нормальному распределению с нулевым вектором средних и дисперсионной матрицей $\sigma^2 M^{-1}(\xi, \Theta_u)$.

Доказательство этой теоремы можно найти в статье (Jennrich, 1969).

Матрица $M(\xi, \Theta_u)$ называется информационной матрицей (для нелинейных по параметрам регрессионных моделей).

1.3 Локально оптимальные планы эксперимента

Для линейных по параметрам регрессионных моделей наиболее распространенным критерием оптимальности плана эксперимента является D-критерий. Согласно этому критерию оптимальными считаются планы, максимизирующие определитель информационной матрицы. В статистическом смысле это означает, что МНК-оценки, построенные по результатам экспериментов в соответствии с D-оптимальным планом, будут иметь наименьший объем доверительного эллипсоида. Кроме того, согласно теореме эквивалентности Кифера–Вольфовица, максимальная по $x \in \mathfrak{X}$ дисперсия оценки значения функции в точке x, также будет минимальной. Подробнее об этом см. (Мелас, 1999, Введение; Ермаков, Жиглявский, 1987, гл. 2).

В случае нелинейных по параметрам моделей информационная матрица, как видно из предыдущего раздела, зависит от истинного значения вектора параметров, которое, разумеется, неизвестно исследователю. Предположим, однако, что известно некоторое приближение $\Theta^{(0)}$. Планы, максимизирующие определитель матрицы $M(\xi, \Theta)$ при $\Theta = \Theta^{(0)}$ будем называть *локально D-оптимальными.*

Эти планы зависят, вообще говоря, от вектора $\Theta^{(0)}$, хотя в некоторых случаях могут и не зависеть от него.

Линеаризируем модель в окрестности точки $\Theta = \Theta^{(0)}$:

$$\eta(x, \Theta) - \eta(x, \Theta^{(0)}) \approx (\Theta - \Theta^{(0)})^T f(x, \Theta^{(0)}).$$

Для линейной модели, представленной в правой части этого приближенного равенства, справедлива следующая теорема, которая будет служить одним из инструментов исследования.

Введем обозначения:

$$\begin{aligned}
f^T(x) &= f^T(x,\Theta^{(0)}) = \left(\frac{\partial\eta(x,\Theta^{(0)})}{\partial\theta_1},\ldots,\frac{\partial\eta(x,\Theta^{(0)})}{\partial\theta_m}\right),\\
M(\xi) &= M(\xi,\Theta^{(0)}) = \int_{\mathfrak{X}} f(x)f^T(x)\xi(dx) = \sum_{i=1}^{n} f(x_i)f^T(x_i)\mu_i,\\
d(x,\xi) &= f^T(x)M^{-1}(\xi)f(x).
\end{aligned}$$

Теорема 1.3.1. *План ξ^* является локально D-оптимальным для модели (1.1)-(1.2) тогда и только тогда, когда*

$$\max_{x\in\mathfrak{X}} d(x,\xi^*) = m.$$

Кроме того,

$$\max_{x\in\mathfrak{X}} d(x,\xi^*) = \inf_{\xi}\max_{x\in\mathfrak{X}} d(x,\xi),$$

где нижняя грань берется по всем (приближенным) планам эксперимента, и функция $d(x,\xi^*)$ достигает своего максимального значения во всех точках любого локально D-оптимального плана. Информационные матрицы всех локально D-оптимальных планов совпадают.

Теорема 1.3.1 является простой переформулировкой теоремы эквивалентности Кифера–Вольфовица, доказательство которой можно найти в книге (Ермаков, Жиглявский, 1987, с. 109).

Заметим, что при $n = m$ информационная матрица $M(\xi)$ принимает вид

$$M(\xi) = X^T\Lambda X,$$

где $\Lambda = \operatorname{diag}\{\mu_1,\ldots,\mu_m\}$, $X = (f_i(x_j))_{j,i=1}^m$, что легко проверить перемножением матриц в правой части. Поэтому

$$\det M(\xi) = (\det X)^2\mu_1\ldots\mu_m.$$

Заметим, что $\prod_{i=1}^m \mu_i \le \left(\sum_{i=1}^m \mu_i/m\right)^m$ в силу неравенства между средним арифметическим и средним геометрическим неотрицательных чисел. Поэтому для планов с числом точек, равным m, оптимальный выбор весов имеет вид $\mu_i = 1/m$, $i = 1,\ldots,m$. Для таких планов достаточно изучить задачу

$$(\det X)^2 \to \max_{x_1,\ldots,x_m\in\mathfrak{X}}.$$

Точные планы вида $\zeta = \{x_1, \dots, x_m\}$ будем называть *насыщенными.* Так как $\det X = 0$ если хотя бы две точки плана совпадают, то можно ограничиться исследованием насыщенных планов, для которых $x_i \neq x_j \ \ (i \neq j)$.

Если число точек в любом локально D-оптимальном плане равно числу параметров (а это имеет место для дробно-рациональной функции регрессии, как будет показано в главе 3), то можно ограничиться изучением насыщенных планов. Локально D-оптимальный (приближенный) план при этом можно получить добавлением равных весов к точкам насыщенного локально D-оптимального плана, т. е. плана локально D-оптимального в классе всех насыщенных планов.

В следующей главе излагается подход, позволяющий изучить зависимость точек и весов локально оптимальных планов от вектора $\Theta^{(0)}$.

Глава 2

Функциональный подход

Настоящая глава содержит изложение идей и результатов разработанного автором функционального подхода к исследованию локально оптимальных планов эксперимента для нелинейных по параметрам регрессионных моделей. Изложение ограничено случаем D-критерия оптимальности. Некоторые другие критерии оптимальности рассматриваются в книге (Мелас, 1999).

Идея функционального подхода заключается в изучении элементов (точек и весов) локально оптимальных планов как неявно заданных вектор-функций параметров, нелинейным образом входящих в рассматриваемую модель. Эти вектор-функции называются оптимальными план-функциями (раздел 3.1). В разделе 3.2 вводится основное векторное уравнение, определяющее оптимальные план-функции. Под вещественными аналитическими функциями понимаются, как обычно, вещественные функции, которые в некоторой окрестности любой точки заданного открытого множества могут быть разложены в (сходящиеся) ряды Тейлора. В случае, когда функция регрессии является вещественной аналитической по параметрам и по аргументу, развиваемый подход позволяет установить условия, при которых оптимальные план-функции также будут вещественными аналитическими (раздел 3.3).

По существу, результаты этого раздела, — глобальная версия известной теоремы о неявной функции.

В разделе 3.4 вводится представление для матрицы якобиана основного уравнения, которое позволяет установить невырожденность якобиана для некоторых классов регрессионных моделей. Результаты этого раздела ранее не публиковались.

В разделе 3.5 для вычисления коэффициентов рядов Тейлора неявных функций, вводятся рекуррентные формулы. Эти формулы применимы для случаев, когда якобиан соответствующей системы уравнений может обращаться в нуль в исходной точке. Кроме того, эти формулы ориентированы на использование пакетов символьной обработки данных Maple и Mathcad. Раздел является изложением работы (Мелас, Пепелышев, 1999).

2.1 Понятие оптимальной план-функции

Во многих случаях параметры, линейно входящие в исследуемую модель, не оказывают влияния на вид локально оптимальных планов. В настоящем разделе мы введем класс регрессионных функций, допускающих исключение таких параметров. Кроме того, мы введем важное понятие *оптимальной план-функции*, основанное на преобразовании задачи нахождения оптимальных планов в задачу нахождения экстремума функции от нескольких переменных (число которых фиксировано).

Предположим, что вектор параметров Θ имеет вид

$$\Theta^T = \left(\Theta_{(1)}^T \vdots \Theta_{(2)}^T\right),$$

где $\Theta_{(1)}^T = (\theta_1, \ldots, \theta_{m-k})$, $\Theta_{(2)}^T = (\theta_{m-k+1}, \ldots, \theta_m)$, причем параметры θ_i, $i = 1, \ldots, m-k$ входят в модель линейным образом, а параметры θ_{i+m-k}, $i = 1, \ldots, k$ входят в модель нелинейным образом. Пусть $m \geq 2k$. Обозначим $l = m - 2k$, $l \geq 0$. Рассмотрим функцию регрессии вида

$$\eta(x, \Theta) = \sum_{i=1}^{l} \theta_i \eta_i(x) + \sum_{i=l+1}^{m-k} \theta_i \eta_i(x, \theta_{i+k}).$$

Пусть

$$\xi = \{x_1, \ldots, x_n; \mu_1, \ldots, \mu_n\},$$

где n — некоторое натуральное число, — произвольный (приближенный) план эксперимента.

Тогда определитель информационной матрицы плана ξ принимает вид

$$\det M(\xi, \Theta) = \theta_{l+1}^2 \ldots \theta_{m-k}^2 \det \bar{M}(\xi, \Theta_{(2)}),$$

где

$$\bar{M}(\xi,\Theta_{(2)}) = \int_{\mathfrak{X}} f(x,\Theta_{(2)}) f^T(x,\Theta_{(2)}\xi(dx), \tag{2.1}$$

$$\begin{aligned}
&f^T(x,\Theta_{(2)}) = (f_1(x),\ldots,f_l(x), f_{l+1}(x,\theta_{m-k+1}),\ldots,\\
&f_{l+k}(x,\theta_m), f_{l+k+1}(x,\theta_{m-k+1}),\ldots,f_m(x,\theta_m)),\\
&f_i(x) = \eta_i(x), \quad i=1,\ldots,l, \qquad (2.2)\\
&f_{l+i}(x,\theta_{m-k+i}) = \eta_{l+i}(x,\theta_{m-k+i}), \quad i=1,\ldots,k,\\
&f_{l+k+i}(x,\theta_{m-k+i}) = \frac{\partial}{\partial\theta_{m-k+i}}\eta_{l+i}(x,\theta_{m-k+i}), \quad i=1,\ldots,k,\\
&\Theta = \Theta_u.
\end{aligned}$$

Пусть $\theta_{l+1}^{(0)},\ldots,\theta_{m-k}^{(0)} \neq 0$. Очевидно, что локально D-оптимальный план зависит только от вектора $\Theta_{(2)} = \Theta_{(2)}^{(0)}$.

Обозначим $z_1 = \theta_{m-k+1}^{(0)},\ldots,z_k = \theta_m^{(0)}$. Пусть $\tilde{Z} = \tilde{Z}_{(n)}$ ограниченное замкнутое множество в $\mathbb{R}^k$ такое, что при $Z = \operatorname{Int}\tilde{Z}$ для любого $z \in Z$ существует локально D-оптимальный план, включающий n различных точек.

Как показано в разделе 1.3, при $n = m$ веса точек в любом локально D-оптимальном плане имеют вид $\mu_i = 1/m$, $i = 1,2,\ldots,m$.

Дадим следующие определения.

Определение 2.1.1. *При $n > m$ оптимальный план-функцией будем называть вектор-функцию*

$$\tau^*(z) : Z \to V,$$

где

$$V = \{\tau = (x_1,\ldots,x_n,\mu_2,\ldots,\mu_n); x_i \in \mathfrak{X},\ \mu_i > 0,\ i=1,\ldots,m,\ \sum_{i=2}^n \mu_i < 1\}$$

такую, что при любом фиксированном $z = \Theta_{(2)}^{(0)} \in Z$ точки $x_1^* = \tau_1^*(z),\ldots,x_n^* = \tau_n^*(z)$ и веса $\mu_1^* = 1 - \sum_{i=2}^n \mu_i^*$, $\mu_i^* = \tau_{n+i}^*(z)$ образуют локально D-оптимальный план.

Определение 2.1.2. *При $n = m$ оптимальной план-функцией будем называть вектор-функцию*

$$\tau^*(z)\ :\ Z \to V,$$

где

$$V = \{\tau = (x_1,\ldots,x_m);\ x_i \in \mathfrak{X}\},$$

такую, что при любом фиксированном $z = \Theta_{(2)}^{(0)} \in Z$ *точки* $x_1^* = \tau_1^*(z), \ldots, x_m^* = \tau_m^*(z)$ *и веса* $\mu_i^* = 1/m$, $i = 1, \ldots, m$ *образуют локально* D*-оптимальный план.*

Понятие оптимальной план-функции введено в работе (Мелас, 1981) при изучении нелинейных по параметрам функций регрессии в виде алгебраической суммы экспонент. Оно является первым шагом к исследованию зависимости локально оптимальных планов от значений параметров, нелинейным образом входящих в исследуемую модель.

2.2 Основное уравнение

В настоящем разделе, исходя из необходимых условий экстремума функции нескольких переменных, мы введем уравнение, определяющее оптимальные план-функции неявным образом.

Заметим, что при любом фиксированном z вектор $\tau^*(z)$ является решением задачи

$$\varphi(\tau, z) \to \sup_{\tau \in V},$$

где

$$\varphi(\tau, z) = [\det \bar{M}(\xi, z)]^{1/m}. \tag{2.3}$$

Рассмотрим сначала частный случай, который нам потребуется в следующей главе.

Пусть $n = m$, $\mathfrak{X} = [r_1, r_2]$, где $r_1 < r_2$ — произвольные заданные вещественные числа. Так как все точки $x_i = \tau_i$ по определению различны, то без ограничения общности будем считать, что

$$r_1 \leq x_1 < x_2 < \ldots < x_m \leq r_2.$$

Пусть при $z \in Z$ точка x_1 любого локально D-оптимального плана совпадает с r_1, а $x_m < r_2$ для такого плана. В этом случае вектор τ и функцию $\varphi(\tau, z)$ переопределим следующим способом:

$$\tau = (\tau_1, \ldots, \tau_{m-1}) = (x_2, \ldots, x_m),$$
$$\varphi(\tau, z) = [\det \hat{M}(\zeta, z)]^{1/m},$$

где $\zeta = (x_1, \ldots, x_m)$, $x_1 = r_1$, $x_{i+1} = \tau_i$, $z = \Theta_{(2)}$,

$$\hat{M}(\zeta, z) = \sum_{i=1}^{m} f(x_i, \Theta_{(2)}) f^T(x_i, \Theta_{(2)}),$$

функция $f(x, \Theta_{(2)})$ определена формулой (2.2).

В силу принятых предположений при любом фиксированном $z \in Z$ максимум функции $\varphi(\tau, z)$, $\tau \in V$ достигается внутри множества V. Поэтому необходимым условием равенства $\tau = \tau^*(z)$ при любом фиксированном $z \in Z$ является равенство нулю производных

$$\frac{\partial}{\partial \tau_i}\varphi(\tau, z) = 0,\ i = 1, \ldots, m-1.$$

Обозначим

$$g_i = g_i(\tau, z) = \frac{\partial}{\partial \tau_i}\varphi(\tau, z),\ i = 1, \ldots, m-1,$$

$$g = (g_1, \ldots, g_{m-1})^T.$$

Тогда

$$g(\tau, z) = 0$$

при $\tau = \tau^*(z)$. Это уравнение будем называть *основным уравнением.* Оно позволяет свести задачу к исследованию неявных функций.

Однако, прежде чем перейти к такому исследованию, рассмотрим более общий случай. Пусть $\mathfrak{X}$ — ограниченное замкнутое множество в $\mathbb{R}^t$, $t > 1$. Предположим, что это множество имеет вид

$$\mathfrak{X} = \{x = (x_{(1)}, \ldots, x_{(t)}), \Phi_j(x) \le 0,\ j = 1, 2, \ldots, k\},$$

где $\Phi_j(x)$, $j = 1, \ldots, k$ — некоторые заданные непрерывно дифференцируемые функции.

Согласно методу неопределенных множителей Лагранжа (Фихтенгольц, 1966, с. 470) необходимым условием $\tau = \tau^*(z)$ является равенство нулю производных

$$\begin{array}{ll} \dfrac{\partial}{\partial \tau_{i(\nu)}} & \left[\varphi(\tau, z) - \sum_{j=1}^k \sum_{i=1}^n \lambda_{ij}\Phi_j(x_i)\right] = 0, \\ \dfrac{\partial}{\partial \tau_j} & \left[\varphi(\tau, z) - \sum_{j=1}^k \sum_{i=1}^n \lambda_{ij}\Phi_j(x_i)\right] = 0, \end{array}$$

где $i = 1, \ldots, n$, $\nu = 1, \ldots, t$, $j = nt + 1, \ldots, 2n(t+1) - 1$,

$$\begin{array}{rcl} \tau & = & (\tau_1, \ldots, \tau_n, \mu_2, \ldots, \mu_n) = (x_1, \ldots, x_n, \mu_2, \ldots, \mu_n), \\ \tau_i & = & (\tau_{i(1)}, \ldots, \tau_{i(t)}) = (x_{i(1)}, \ldots, x_{i(t)}),\ i = 1, \ldots, n, \end{array}$$

при некоторых значениях величин

$$\{\lambda_{ij} \ge 0;\ i = 1, \ldots, n,\ i = 1, \ldots, k\}.$$

Кроме того, для всех пар $(i,j) \in S = \{(i,j); \lambda_{ij} > 0\}$ должно выполняться равенство $\Phi_j(x_i) = 0$.

Пусть r число элементов множества S и каждой паре $(i,j) \in S$ сопоставлен номер $\nu = \nu(i,j)$, $\nu = 1, 2, \ldots, r$.

Обозначим

$$\begin{aligned}
\bar{\varphi}(\tau, \lambda, z) &= \varphi(\tau, z) - \sum_{j=1}^{k} \sum_{i=1}^{n} \lambda_{ij} \Phi_j(x_i), \\
\lambda_{ij} &= \lambda_\nu, \ \nu = \nu(i,j), \ \nu = 1, \ldots, r, \\
\lambda_{ij} &= 0, \quad (i,j) \notin S.
\end{aligned}$$

Пусть $g = (g_1, \ldots, g_s)^T$, $s = (t+1)n + r - 1$ — вектор производных

$$\begin{aligned}
&\frac{\partial}{\partial \tau_{i(j)}} \bar{\varphi}(\tau, \lambda, z), \ i = 1, \ldots, n, \ j = 1, \ldots, t, \\
&\frac{\partial}{\partial \tau_{nt+i}} \bar{\varphi}(\tau, \lambda, z), \ i = 1, \ldots, n-1, \\
&\frac{\partial}{\partial \lambda_\nu} \bar{\varphi}(\tau, \lambda, z) = \Phi_j(x_i), \ \nu = 1, \ldots, r.
\end{aligned}$$

Тогда необходимым условием равенства $\tau = \tau^*(z)$ является выполнение векторного равенства

$$g(\tau, \lambda, z) = 0.$$

Если при любом фиксированном $z \in Z$ величина r остается постоянной, то это уравнение определяет вектор-функцию $\tau^*(z)$ неявным образом.

Вернемся к рассмотренному ранее случаю $n = m$, $\mathfrak{X} = [r_1, r_2]$, $x_1 = r_1$ для любого локально D-оптимального плана.

Предположим, что функции $\eta_i(x)$, $i = 1, 2, \ldots, l$, $\eta_{l+j}(x, z_j)$, $j = 1, \ldots, k$ — вещественные аналитические функции при $x \in \operatorname{Int} \mathfrak{X}$, $z \in Z$. Тогда функция $\varphi(\tau, z)$, определенная формулой (2.3), является вещественной аналитической при $z \in Z$, $\tau \in \operatorname{Int} \mathfrak{X}^{m-1}$, так как определитель $\hat{M}(\zeta, z)$ представляется в виде алгебраической суммы произведений функций η_i и их производных.

Следовательно, функции $g_i(\tau, z) = \frac{\partial}{\partial \tau_i} \varphi(\tau, z)$, $i = 1, \ldots, m-1$ являются вещественными аналитическими при $\tau \in \operatorname{Int} \mathfrak{X}^{m-1}$, $z \in Z$.

Очевидно, что такие же выводы можно сделать и в общем случае, если предположить дополнительно, что функции $\Phi_j(x)$, $j = 1, \ldots, k$ являются вещественными аналитическими при $x \in \operatorname{Int} \mathfrak{X}$.

Кроме того предположим, что основное уравнение имеет единственное решение $\tau = \tau^*(z)$ при любом фиксированном $z \in Z$.

В следующем разделе мы изучим основное уравнение для функций g общего вида, удовлетворяющих сформулированным выше условиям.

2.3 Аналитичность неявных функций

Пусть m и k — любые натуральные числа, T и Z — ограниченные замкнутые множества в $\mathbb{R}^m$ и $\mathbb{R}^k$, соответственно, $\tilde{T} = \operatorname{int} T$, $\tilde{Z} = \operatorname{int} Z$, $\hat{T}$ — ограниченное замкнутое подмножество $\tilde{T}$. Пусть $g(\tau, z)$ — произвольная вещественная аналитическая вектор-функция при $\tau \in \tilde{T}$, $z \in \tilde{Z}$ такая, что при любом фиксированном $z \in Z$ уравнение

$$g(\tau, z) = 0 \tag{2.4}$$

при $\tau \in T$ имеет единственное решение $\tau = \tau(z) \in \hat{T}$.

Пусть $\tau^*(z)$ — вектор-функция, такая, что при фиксированном $z \in Z$ и $\tau = \tau^*(z)$ выполнено равенство (2.4).

В силу сделанного нами предположения эта функция определена единственным образом.

Обозначим

$$J(\tau, z) = \frac{\partial}{\partial \tau} g(\tau, z),$$

тогда $J(\tau, z)$ есть $m \times m$ матрица.

Предположим, что выполняется условие

$$(A) \qquad\qquad \det J(\tau^*(z), z) \neq 0 \quad z \in \tilde{Z}.$$

Имеет место следующая теорема.

Теорема 2.3.1. *(об аналитичности неявной функции). При выполнении сформулированных выше условий $\tau^*(z)$ есть вещественная аналитическая вектор-функция при $z \in \tilde{Z}$.*

Доказательство. Пусть W — любое компактное подмножество $\tilde{Z}$, $H = \hat{T} \times W$, $z_{(0)} \in W$, $\tau_{(0)} = \tau^*(z_{(0)})$. Тогда согласно нашему допущению $(\tau_{(0)}, z_{(0)}) \in H$.

Рассмотрим всевозможные точки $(\tau_{(0)}, z_{(0)})$ такого вида. По теореме о неявной функции (Бибиков, 1991, с. 173) для любой такой точки существует некоторая окрестность $U_{(0)}$, и окрестность точки $z_{(0)}$, которую обозначим через $W_{(0)} = W(z_{(0)})$, такие, что при $z \in W_{(0)}$ существует вещественная аналитическая вектор-функция $\tilde{\tau}(z)$ со свойствами: $\tilde{\tau}(z_{(0)}) = \tau_{(0)}$, $(\tau(z), z) \in U_{(0)}$ и $g(\tilde{\tau}(z), z) = 0$.

В силу единственности решения уравнения (2.4) при любом фиксированном $z \in \tilde{Z}$ эта вектор-функция совпадает с $\tau^*(z)$ при $z \in W_{(0)}$. Окрестности $W_{(0)} =$

$W(z_{(0)})$ образуют покрытие множества W. В силу компактности W из этого покрытия можно выделить конечное подпокрытие, таким образом

$$W \subset \sum_{i=1}^{L} W(z_{(0)}^{(i)}).$$

Следовательно, $\tau^*(z)$ является вещественной аналитической вектор-функцией на W, а значит, и на $\tilde{Z}$. □

В следующем разделе мы получим формулу для матрицы якобиана уравнения (2.4), которая облегчает проверку условия (A).

2.4 Якобиан основного уравнения

Изучим сначала матрицу якобиана основного уравнения для функций $\varphi(\tau, z)$ общего вида, представимых в виде минимума некоторой выпуклой функции.

Пусть m, k, t — произвольные натуральные числа, $T \subset \mathbb{R}^{m-1}, Z \subset \mathbb{R}^k, \mathfrak{A} \subset \mathbb{R}^t$ — произвольные открытые множества, причем $\mathfrak{A}$ — выпуклое.

Рассмотрим функцию $q(\tau, a, z), \tau \in T, a \in \mathfrak{A}, z \in Z$ со следующими свойствами:

(а) функция $q(\tau, a, z)$ дважды непрерывно дифференцируема по τ и a;

(б) функция $q(\tau, a, z)$ строго выпукла по a.

Пусть, далее, функция $\varphi(\tau, z)$ имеет вид

$$\varphi(\tau, z) = \min_{a \in \mathfrak{A}} q(\tau, a, z), \tag{2.5}$$

причем при любых $\tau \in T, z \in Z$ этот минимум достигается. Из выпуклости функции $q(\tau, a, z)$ по a вытекает, что этот минимум достигается на единственном векторе $a = \tilde{a} = \tilde{a}(\tau, z)$. Как известно (Демьянов, Малоземов, 1972) функция минимума дифференцируема, если минимум достигается на единственном векторе. Поэтому функция $\varphi(\tau, z)$ дважды непрерывно дифференцируема по τ. Обозначим $g(\tau, z) = \frac{\partial}{\partial \tau}\varphi(\tau, z)$ и рассмотрим уравнение

$$g(\tau, z) = 0. \tag{2.6}$$

Предположим, что это уравнение при любом $z \in Z$ имеет единственное решение $\tau = \tau^* = \tau^*(z)$. Изучим матрицу

$$J(\tau, z) = \frac{\partial}{\partial \tau} g(\tau, z) = \left(\frac{\partial^2}{\partial \tau_j \partial \tau_i} \varphi(\tau, z) \right)_{i,j=1}^{m-1} \tag{2.7}$$

при $\tau = \tau^*(z)$. Пусть $a^* = a^*(z) = \tilde{a}(\tau^*(z), z)$.

Рассмотрим следующие матрицы

$$E = \left(\frac{\partial^2}{\partial\tau_j\partial\tau_i} q(\tau, a, z)\right)_{i,j=1}^{m-1},$$

$$B = \left(\frac{\partial^2}{\partial\tau_j\partial a_i} q(\tau, a, z)\right)_{i,j=1}^{t,m-1}, \tag{2.8}$$

$$D = \left(\frac{\partial^2}{\partial a_j\partial a_i} q(\tau, a, z)\right)_{i,j=1}^{t}$$

при $\tau = \tau^*, a = a^*$. Из условия (б) вытекает, что матрица D положительно определена и, значит, существует обратная матрица D^{-1}.

Теорема 2.4.1. *При сформулированных выше условиях имеет место формула*

$$J(\tau^*(z), z) = E - B^T D^{-1} B. \tag{2.9}$$

„®Є § вҐ«мбвў®. В силу необходимых условий экстремума при любом фиксированном $z \in Z$ имеет место равенство

$$\frac{\partial}{\partial a} q(\tau, a, z) = 0 \tag{2.10}$$

при $\tau = \tau^* = \tau^*(z), a = a^* = a^*(z)$. Рассмотрим это векторное равенство при фиксированном z и произвольных τ, a как систему уравнений, неявным образом определяющую функцию $a(\tau)$ такую, что $a(\tau^*) = a^*$. Якобиан этой системы в точке (τ^*, a^*) равен $\det D \neq 0$. По теореме о неявной функции существует единственная непрерывная вектор-функция $a(\tau)$ такая, что $a(\tau^*) = a^*$ и заданная в некоторой окрестности точки τ^*. Эта функция непрерывно дифференцируема и

$$\left.\frac{\partial a(\tau)}{\partial \tau}\right|_{\tau=\tau^*} = -D^{-1}B. \tag{2.11}$$

Используя правило дифференцирования сложной функции, получаем

$$\begin{aligned}\left.\frac{\partial^2}{\partial\tau_j\partial\tau_i} q(\tau, a(\tau), z)\right|_{\tau=\tau^*} &= \left.\frac{\partial^2}{\partial\tau_j\partial\tau_i} q(\tau, a, z)\right|_{\tau=\tau^*, a=a^*} + \\ &+ \sum_{\nu=1}^{t} \left.\frac{\partial^2}{\partial a_\nu\partial\tau_i} q(\tau, a, z)\right|_{\tau=\tau^*, a=a^*} \frac{\partial a_\nu(\tau^*)}{\partial\tau_j}.\end{aligned}$$

Учитывая обозначения (2.8) и формулу (2.11), получаем

$$\left(\frac{\partial^2}{\partial\tau_j\partial\tau_i}q(\tau,a(\tau),z)\bigg|_{\tau=\tau^*}\right)_{i,j=1}^{m-1}=E-B^TD^{-1}B. \tag{2.12}$$

С другой стороны при любом фиксированном $z\in Z$ имеем

$$\varphi(\tau,z)=\min_{a\in\mathfrak{A}}q(\tau,a,z)=q(\tau,a(\tau),z) \tag{2.13}$$

при τ из некоторой окрестности точки $\tau^*=\tau^*(z)$. Дифференцируя это равенство дважды по τ, получаем

$$J(\tau^*(z),z)=J(\tau^*,z)=E-B^TD^{-1}B. \tag{2.14}$$

Применим полученный результат к функции $\varphi(\tau,z)$, определенной формулой (2.3).

Обозначим через $\mathcal{A}$ множество всех положительно определенных $m\times m$ матриц $A=(a_{ij})$ таких, что $a_{mm}=1$.

Лемма 2.4.1. *Для любых матриц $A,B\in\mathcal{A}$ выполняется неравенство*

$$(\det B)^{1/m}\le\frac{\operatorname{tr}AB}{m(\det A)^{1/m}} \tag{2.15}$$

причем равенство достигается для единственной матрицы $A=A^*\in\mathcal{A}$, где $A^*=const\,B^{-1}$.

„®Є § вҐ«мбвў®. Неравенство (2.15) есть специальный случай неравенства между средним арифметическим и средним геометрическим (Карлин, Стадден, 1976, с. 325).

Лемма 2.4.2. *Функция $\Psi(B)=(\det B)^{1/m}$ при $B\in\mathcal{A}$ строго вогнутая.*

„®Є § вҐ«мбвў®. В силу леммы 2.4.1 имеем

$$\Psi(B)=\min_{A\in\mathcal{A}}\frac{\operatorname{tr}AB}{m(\det A)^{1/m}}. \tag{2.16}$$

Следовательно, при $B_1,B_2\in\mathcal{A}$ выполняется

$$\Psi((B_1+B_2)/2)=\min_{A\in\mathcal{A}}\frac{\operatorname{tr}A(B_1+B_2)/2}{m(\det A)^{1/m}}\ge \tag{2.17}$$

$$\ge\frac{1}{2}\left(\min_{A\in\mathcal{A}}\frac{\operatorname{tr}AB_1}{m(\det A)^{1/m}}+\min_{A\in\mathcal{A}}\frac{\operatorname{tr}AB_2}{m(\det A)^{1/m}}\right)=(\Psi(B_1)+\Psi(B_2))/2, \tag{2.18}$$

причем равенство достигается только при $B_1=B_2$.

Лемма 2.4.3. *Функция*

$$\Psi(A) = \frac{\operatorname{tr} AB}{m(\det A)^{1/m}} \tag{2.19}$$

при $A \in \mathcal{A}$ и любой матрице $B \in \mathcal{A}$ строго выпуклая.

„®Є § вГ«мбвў®. Заметим, что

$$\frac{2}{\alpha+\beta} \leq \frac{1}{2}\left(\frac{1}{\alpha} + \frac{1}{\beta}\right), \tag{2.20}$$

для любых чисел $\alpha, \beta > 0$. В самом деле, это неравенство эквивалентно неравенству

$$(\alpha+\beta)^2 - 4\alpha\beta = (\alpha-\beta)^2 \geq 0. \tag{2.21}$$

Теперь утверждение леммы доказывается непосредственным вычислением с использованием леммы 2.4.2 и неравенства (2.20).

Паре индексов (i, j), $i \leq j$, $i, j = 1, \ldots, m$, $(i, j) \neq (m, m)$ поставим в соответствие число $\nu = \nu(i, j)$:

$$\begin{aligned} \nu(1,1) &= 1, \nu(1,2) = 2, \ldots, \nu(2,2) = m+1, \ldots, \nu(m-1, m) = \\ &= t = m(m+1)/2 - 1. \end{aligned} \tag{2.22}$$

Для любого вектора $a \in \mathbb{R}^t$ определим матрицу $A(a)$ следующими соотношениями

$$a_{ji} = a_{ij} = a_{\nu(i,j)},\ a_{mm} = 1,\ i, j = 1, \ldots, m,\ i \leq j. \tag{2.23}$$

Определим множество $\mathfrak{A}$ следующим образом

$$\mathfrak{A} = \{a \in \mathbb{R}^t :\ A(a) \in \mathcal{A}\}. \tag{2.24}$$

Множество $\mathfrak{A}$, очевидно, открыто и выпукло в $\mathbb{R}^t$. Определим функцию

$$q(\tau, a, z) = (\det A(a))^{-1/m} tr\left(A(a)\overline{M}(\zeta, z)\right) m, \tag{2.25}$$

где матрица $\hat{M}(\zeta, z)$ определена формулой (2.2). Рассмотрим функцию $\varphi(\tau, z) = (\det \hat{M}(\zeta, z))^{1/m}$. В силу леммы 2.4.1 для этой функции справедлива формула (2.5). Все условия, определенные перед теоремой 2.4.1, очевидно, выполнены. Поэтому в силу теоремы 2.4.1.

$$J(\tau^*(z), z) = E - B^T D^{-1} B. \tag{2.26}$$

Обозначим $\psi(a) = (\det A(a))^{-1/m}$. Для матриц B и E непосредственным дифференцированием нетрудно проверить справедливость следующих формул.

$$\begin{aligned}
&E = diag\{E_{11}, \ldots, E_{m-1m-1}\},\\
&E_{ii} = \psi(a^*)\ \tfrac{\partial^2}{\partial x^2}(f^T(x)A(a^*)f(x))\Big|_{x=x^*_{i+1}}, i = 1, \ldots, m-1,\\
&A(a^*) = const\left(\hat{M}(\zeta^*, z)\right)^{-1},\\
&B = (b_{\nu k})^{t,m-1}_{\nu,k=1},\\
&b_{\nu k} = 2\psi(a^*)\ \tfrac{\partial}{\partial x}(f_i(x)f_j(x))\big|_{x=x^*_k}, \nu = \nu(i,j).
\end{aligned} \tag{2.27}$$

Заметим, что матрица $J = J(\tau^*(z), z)$ является отрицательно определенной, а значит и невырожденной, если выполняется хотя бы одно из следующих условий:

1) все диагональные элементы матрицы E отрицательны;

2) матрица B имеет полный ранг.

В самом деле, матрица $B^T D^{-1} B$ имеет вид SS^T и, значит, является неотрицательно определенной, а при матрице B полного ранга — положительно определенной. Так как $J = E - B^T D^{-1} B$, то J отрицательна определена при выполнении любого из условий 1)-2).

2.5 Степенные разложения неявных функций

Известно (Фихтенгольц, 1966, с. 460), что производные неявных функций могут быть вычислены рекуррентным образом с помощью метода неопределенных коэффициентов. В данном разделе, следуя работе (Мелас, Пепелышев, 1999), мы введем рекуррентные формулы, ориентированные на использование пакетов символьных вычислений Maple и Mathcad.

Пусть m и k — произвольные натуральные числа, $\tau = (\tau_1, \ldots, \tau_m) \in \mathbb{R}^m$, $z = (z_1, \ldots, z_k) \in \mathbb{R}^k$ и $g(\tau, z) = (g_1(\tau, z), \ldots, g_m(\tau, z))$ — вещественная аналитическая вектор-функция в некоторой окрестности U точки $(\tau_{(0)}, z_{(0)})$, такая, что $g(\tau_{(0)}, z_{(0)}) = 0$.

Для любого набора индексов $s = (s_1, \ldots, s_k)$, $s_i \geq 0$, $i = 1, \ldots, k$ и любой (скалярной, векторной или матричной) вещественной аналитической функции f обозначим

$$f_{(s)} = (f(z))_{(s)} = \frac{1}{s_1! \ldots s_k!} \frac{\partial^{s_1}}{\partial z_1^{s_1}} \cdots \frac{\partial^{s_k}}{\partial z_k^{s_k}} f(z)|_{z=z_{(0)}}.$$

Обозначим

$$S_t = \{s = (s_1, \dots, s_k); \quad s_i \geq 0, \quad \sum_{i=1}^{k} s_i = t\}$$

при $t = 0, 1, \dots$.

Пусть $l = (l_1, \dots, l_k)$,$l_i \geq 0$ — заданный вектор целых чисел.

Предположим, что функция g имеет вид

$$g(\tau, z) = (z_1 - z_{1(0)})^{l_1} \dots (z_k - z_{k(0)})^{l_k} \psi(z) \tilde{g}(\tau, z),$$

где $\psi(z)$ — однородный многочлен степени $p \geq 0$,

$$\psi(z) = \sum_{s \in S_p} a_{(s)} (z_1 - z_{1(0)})^{s_1} \dots (z_k - z_{k(0)})^{s_k},$$

такой, что $a_{(p,0,\dots,0)} \neq 0$, причем $\det \tilde{J}(\tau_{(0)}, z_{(0)}) \neq 0$, где

$$\tilde{J}(\tau, z) = \frac{\partial}{\partial \tau} \tilde{g}(\tau, z).$$

Рассмотрим уравнение

$$g(\tau, z) = 0.$$

При $z_i \neq z_j$, $(i \neq j)$, $z_i \neq z_{i(0)}$, $i, j = 1, \dots, k$ и при $\psi(z)$ таком, что $\psi(z) \neq 0$ при $z_i \neq z_j$, $i \neq j$, $z_i \neq z_{i(0)}$, $i, j = 1, \dots, k$, это уравнение эквивалентно уравнению

$$\tilde{g}(\tau, z) = 0.$$

Легко проверить, что функция $\tilde{g}(\tau, z)$ является вещественной аналитической при $(\tau, z) \in U$.

В силу теоремы о неявной функции (Бибиков, 1991, с. 173) существует вектор-функция $\tau^*(z)$, которая является вещественной аналитической вектор-функцией в некоторой окрестности H точки $\tau = \tau_{(0)}$, и такая, что $\tau^*(z_{(0)}) = \tau_{(0)}$, $(\tau^*(z), z) \in U$ и $\tilde{g}(\tau^*(z), z) = 0$, а значит и $g(\tau^*(z), z) = 0$ при $z \in H$. Эта вектор-функция в некоторой окрестности $H^* \subset H$ точки $z_{(0)}$ разлагается в (сходящийся) ряд Тейлора:

$$\tau^*(z) = \tau_{(0)} + \sum_{t=1}^{\infty} \sum_{s \in S_t} \tau^*_{(s)} (z_1 - z_{1(0)})^{s_1} \dots (z_k - z_{k(0)})^{s_k}.$$

Отметим, что при $k = 1$ множество S_t состоит из одного элемента $s = s_t = t$.

Построим рекуррентные формулы для вычисления коэффициентов $\tau^*_{(s)}$, $s \in S_t$, $t = 1, 2, \ldots$.

Пусть I — произвольное множество индексов вида $s = (s_1, \ldots, s_k)$, $s_i \geq 0$, $i = 1, \ldots, k$.

Обозначим

$$\tau_{<I>}(z) = \sum_{s \in I} \tau_{(s)}(z_1 - z_{1(0)})^{s_1} \ldots (z_k - z_{k(0)})^{s_k},$$

$$J(\tau, z) = \frac{\partial}{\partial \tau} g(\tau, z), \quad J_{(l)} = \left(J(\tau_{(0)}, z)\right)_{(l)}.$$

Пусть сначала $p = 0$.

Введем множество индексов

$$I_n = U^n_{t=0} S_t.$$

Тогда имеет место следующее утверждение.

Предложение 2.5.1. *При сформулированных выше условиях имеют место следующие формулы*

$$\tau^*_{(s)} = -J^{-1}_{(l)} g(\tau^*_{<I>}(z), z)_{(s+l)}$$

при $s \in S_t$, $I = I_{t-1}$, $t = 1, 2, \ldots$

Доказательство этого и следующего предложения будет дано в конце раздела.

С помощью предложения 2.5.1 можно построить рекуррентный алгоритм вычисления коэффициентов разложения функции $\tau^*(z)$ в ряд Тейлора при $p = 0$ и произвольном $l = (l_1, \ldots, l_k)$, $l_i \geq 0$, $i = 1, \ldots, k$.

Алгоритм 1. На шаге t ($t = 1, 2, \ldots$) вычисляем все коэффициенты с индексами из S_t по формуле

$$\tau_{(s)} = -J^{-1}_{(l)} \left(g(\tau_{<I_{t-1}>}(z), z\right)_{(s+l)}.$$

Из Предложения 2.5.1 вытекает, что $\tau_{(s)} = \tau^*_{(s)}$, $s \in S_t$, $t = 1, 2, \ldots$.

Рассмотрим теперь случай $p > 0$. Определим множество индексов

$$\hat{S}_t = \{s = (s_1, \ldots, s_k); s_i \geq 0, i = 1, \ldots, k, s_1 + 2\sum_{i=2}^{k} s_i = t\}.$$

Пусть $\hat{I}_n = U^n_{t=0} \hat{S}_t$, $q = (p, 0, \ldots, 0)$.

Пусть $J_{(l+q)} = \left(J(\tau_{(0)}, z)\right)_{(l+q)}$.

Предложение 2.5.2. *При $p > 0$ имеют место формулы*

$$\tau^*_{(s)} = -J^{-1}_{(l+q)} g\left(\tau^*_{<I>}(z), z\right)_{(s+l+q)}$$

при $s \in \hat{S}_t$, $I = \hat{I}_{t-1}$, $t = 1, 2, \ldots$.

Алгоритм вычисления коэффициентов $\tau^*_{(s)}$ получаем из Алгоритма 1 заменой множества S_t на $\hat{S}_t$.

Алгоритм 2. *При $p > 0$ на шаге t $(t = 1, 2, \ldots)$ вычисляем все коэффициенты с индексами $s \in \hat{S}_t$ по формуле*

$$\tau_{(s)} = -J^{-1}_{(l+q)} g\left(\tau_{<\hat{I}_{t-1}>}(z), z\right)_{(s+l+q)}.$$

Из Предложения 2.5.2 вытекает, что $\tau_{(s)} = \tau^*_{(s)}$, $s \in \hat{S}_t$, $t = 1, 2, \ldots, s$. Отметим, что при $k > 1$ в Алгоритме 2 на каждом шаге вычисляется приблизительно в k раз меньшее число коэффициентов, чем в Алгоритме 1, а при $k = 1$ эти алгоритмы совпадают.

Перейдем к доказательству предложений 2.5.1 и 2.5.2.

Доказательство предложения 2.5.1. Пусть $\tau(z)$ — произвольная вещественная аналитическая в некоторой окрестности точки $z_{(0)}$ вектор- функция. Рассмотрим следующий вспомогательный результат.

Лемма 2.5.1. *При сформулированных в начале раздела условиях и при $p = 0$, $l = 0$ справедливы равенства:*

$$\frac{\partial^n}{\partial z^n}\left[g\left(\tau_{(n)}(z), z\right) - g\left(\tau(z), z\right)\right]|_{z=z_{(0)}} = 0,$$

при $k = 1$ и $\tau_{(n)}(z) = \sum_{i=1}^{n} \tau_{(i)} z^i + \tau_{(0)}$, $n = 1, 2, \ldots$,

$$\frac{\partial^n}{\partial z_1^{s_1} \ldots \partial z_k^{s_k}}\left[g\left(\tau_{<I>}(z), z\right) - g\left(\tau(z), z\right)\right]|_{z=z_{(0)}} = 0,$$

при $k \geq 1$, $s \in S_n$, где $I = I_n$, $n = 1, 2, \ldots$.

Доказательство. Пусть сначала $k = 1$.

Заметим, что справедливо равенство

$$\frac{\partial}{\partial z} g\left(\tau(z), z\right) = \frac{\partial}{\partial \tau} g(\tau, z)|_{\tau=\tau(z)} \times \tau'(z) + \frac{\partial}{\partial z} g(\tau, z)|_{\tau=\tau(z)}.$$

В самом деле, по определению производной и в силу формулы дифференцирования сложной функции, имеем

$$\frac{\partial}{\partial z}g(\tau(z),z)=\lim_{\Delta\to 0}\frac{g(\tau(z+\Delta),z+\Delta)-g(\tau(z),z)}{\Delta}=$$
$$=\frac{\partial}{\partial\tau}g(\tau,z)|_{\tau=\tau(z)}\tau^{'}(z)+\frac{\partial}{\partial z}g(\tau,z)|_{\tau=\tau(z)}.$$

Используя это равенство, получим

$$\frac{\partial^n}{\partial z^n}g(\tau(z),z)=\frac{\partial}{\partial\tau}g(\tau(z),z)\tau^{(n)}(z)+\frac{\partial^n}{\partial z^n}g(\tau,z)|_{\tau=\tau(z)}+$$
$$+\sum_{i,j=1}^{m}\frac{\partial^2}{\partial\tau_i\partial\tau_j}g\left(\tau(z),z\right)\sum_{t=1}^{n-1}\tau_i^{(t)}(z)\tau_j^{(n-t)}(z)+\ldots+$$
$$+\sum_{i_1,\ldots,i_n=1}^{m}\frac{\partial^n}{\partial\tau_{i_1}\ldots\partial\tau_{i_n}}g\left(\tau(z),z\right)\tau^{'}_{i_1}(z)\ldots\tau^{'}_{i_n}(z).$$

Переходя к пределу при $z\to z_{(0)}$, получаем

$$\begin{aligned}&\frac{\partial^n}{\partial z^n}g(\tau(z),z)|_{z=z_{(0)}}=\\&=n!J_{(0)}\tau_{(n)}+\frac{\partial^n}{\partial z^n}g(\tau_{(0)},z_{(0)})+\ldots+\\&+\sum_{i_1,\ldots,i_n=1}^{m}\frac{\partial^n}{\partial\tau_{i_1}\ldots\partial\tau_{i_n}}g(\tau_{(0)},z_{(0)})\tau_{i_1(1)}\ldots\tau_{i_n(1)}i_1!\ldots i_n!,\end{aligned}\tag{2.28}$$

где правая часть зависит только от $\tau_{(0)},\ldots,\tau_{(n)}$ и не зависит от $\tau_{(n+1)},\ldots$. Следовательно,

$$\frac{\partial^n}{\partial z^n}g\left(\tau(z),z\right)|_{z=z_{(0)}}=\frac{\partial^n}{\partial z^n}g\left(\tau_{(n)}(z),z\right)|_{z=z_{(0)}}.$$

Доказательство леммы для случая $k>1$ проводится аналогичным образом.

Перейдем к доказательству **Предложения 2.5.1**. Пусть $k=1$, $l=0$. Заметим, что в правой части равенства (2.28) только первое слагаемое зависит от $\tau_{(n)}$, а остальные зависят только от τ_t, $t\le n-1$. Так как $g(\tau^*(z),z)\equiv 0$ при $z\in H$, то

$$-\frac{\partial^n}{\partial z^n}g\left(\tau^*_{(n-1)}(z),z\right)|_{z=z_0}=n!J_{(0)}\tau^*_{(n)}.$$

При $k>1$, $l=0$, рассуждения проводятся аналогичным образом.

В случае $l\neq 0$ формулы проверяются непосредственным вычислением производной с учетом леммы 2.5.1. □

Очевидно, что при $p = 0$ формулы из предложения 1 остаются верными, если заменить в них S_t на $\hat{S}_t$ и I_t на $\hat{I}_t$.

Рассмотрим теперь случай $p > 0$.

Доказательство Предложения 2.5.2. Пусть сначала $l = 0$. Заметим, что

$$(g(\tau^*_{<I>}(z), z))_{(s+q)} = \sum_{w+v=s+q} a_{(w)} \tilde{g}(\tau^*_{<I>}(z), z)_{(v)} \qquad (2.29)$$

для любого множества индексов I.

При $w = q$, единственным вектором v, таким, что $w + v = s + q$ является вектор s. Пусть $s \in \hat{S}_n$, $I = \hat{I}_n$. Заметим, что при $w \neq q$ любой вектор v, такой что $w + v = s + q$, принадлежит множеству $\hat{S}_t$, $t \leq n - 1$. Отсюда правая часть равенства (2.29) имеет вид

$$a_{(q)} \tilde{g} \left(\tau^*_{<\tilde{I}_n>}(z), z \right)_{(s)}.$$

Непосредственным вычислением можно проверить, что $J_{(q)} = a_{(q)} \tilde{J}_{(0)}$. Отсюда вытекает справедливость предложения 2.5.2 при $l = 0$. Для произвольного l справедливость формул проверяется непосредственным вычислением. $\square$

В работе (Мелас, Пепелышев, 1999) с помощью описанных здесь алгоритмов построены разложения в ряд Тейлора точек насыщенных локально D-оптимальных планов для функции регрессии вида

$$\eta(x, \Theta) = \sum_{i=1}^{k} \theta_i \exp\left(-\theta_{i+k} x\right),$$

где $\Theta = (\theta_1, \ldots, \theta_{2k})^T$, $\theta_i \neq 0$, $\theta_{i+k} > 0$ $(i = 1, \ldots, k)$, $x \in [0, \infty)$.

В следующей главе алгоритм 1 применяется для дробно-рациональной функции регрессии.

Глава 3

Дробно-рациональная модель

Дробно-рациональные функции образуют важный класс функций, который может быть использован, в частности, для аппроксимации произвольной непрерывной функции. Дробно-рациональная аппроксимация часто предпочтительнее полиномиальной, так как содержит, как правило, значительно меньшее число параметров.

Настоящая глава посвящена построению и исследованию локально D-оптимальных планов для дробно-рациональной регрессионной модели. В разделе 3.1 описываются различные способы задания этой модели. В разделе 3.2 показано, что существует единственный локально D-оптимальный план и число его точек равно числу оцениваемых параметров. В следующем разделе предложен алгебраический метод нахождения локально D-оптимальных планов, основанный на развитии идеи Стильтьеса. С помощью этого метода можно найти в явном виде локально D-оптимальные планы для некоторых специально выбранных значений параметров, что образует основу применения функционального подхода. В заключительном разделе построены степенные разложения оптимальных план-функций для дробно-рациональной модели, в числителе которой — многочлен второй степени, а в знаменателе — многочлен третьей степени. Для более простых моделей оптимальные планы найдены в явном виде в предыдущем разделе.

Результаты, изложенные в этой главе, публикуются впервые.

3.1 Описание модели

Пусть $\Theta_1 = (\theta_1, \ldots, \theta_{m-k})^T \in \mathbb{R}^{m-k}$, $\Theta_2 = (\theta_{m-k+1}, \ldots, \theta_m)^T \in \mathbb{R}^k$, $P(x) = P(x, \Theta_1)$ и $Q(x) = Q(x, \Theta_2)$ — многочлены одной переменной

$$P(x) = \sum_{j=1}^{m-k} \theta_j x^{j-1},\ Q(x) = \sum_{j=1}^{k} \theta_{j+m-k} x^{j-1} + x^k.$$

Пусть $\Theta = (\theta_1, \ldots, \theta_m)^T$ и Ω — некоторое заданное ограниченное подмножество $\mathbb{R}^m$ такое, что дробь

$$\eta_1(x, \Theta) = P(x, \Theta_1)/Q(x, \Theta_2) \tag{3.1}$$

несократима при $\Theta \in \Omega$. Пусть $\mathfrak{X} \subset \mathbb{R}^1$, $\mathfrak{X}$ —ограниченное замкнутое множество, не содержащее нулей многочлена $Q(x, \Theta_2)$ при $\Theta \in \Omega$.

Под дробно-рациональной функцией регрессии будем понимать функцию $\eta_1(x, \Theta)$, заданную формулой (3.1), где $\Theta \in \Omega, x \in \mathfrak{X}$, $\theta_1, \ldots, \theta_m$ — оцениваемые параметры.

Для этой функции и модели (1.1)–(1.2) мы в следующем разделе докажем, что локально D-оптимальный план существует, единственен и сосредоточен в m точках с равными весами. Кроме того, мы получим формулу для определителя информационной матрицы плана, сосредоточенного в m точках. Более подробные результаты будут получены при следующих дополнительных ограничениях:
(a) все корни многочлена $Q(x) = Q(x, \Theta_2)$ при $\Theta \in \Omega$ отрицательны и различны

$$Q(x) = \prod_{j=1}^{k} (x + \gamma_j),\ \gamma_1 > \ldots > \gamma_k > 0$$

и $\mathfrak{X} = [0, d],\ d > 0$;
(b) $m - k > k$;
(c) $m = 2k$.

Нетрудно проверить, что при выполнении условия (a) правую часть равенства (3.1) можно записать в виде

$$\eta_2(x, \tilde{\Theta}) = \sum_{j=1}^{l} \mathrm{tg}_j x^{j-1} + \sum l_{j=1}^{k} \frac{\mathrm{tg}_{j+l}}{x + \mathrm{tg}_{j+l+k}}, \tag{3.2}$$

где $l = m - 2k$ при $m > 2k$ и $l = 0$ при $m \le 2k$, $\mathrm{tg}_{j+l+k} = \gamma_j$, $j = 1, \ldots, k$. Так как дробь в правой части (3.1) предполагается несократимой, то, очевидно,

$\mathrm{tg}_{j+l} \neq 0$, $j = 1, \ldots, k$. Кроме того, очевидно, что при $m < 2k$ между tg_{j+l} существуют $m - 2k$ линейных связей. При $m \geq 2k$ функции регресии (3.1) и (3.2) являются эквивалентными в следующем смысле.

Лемма 3.1.1. *Пусть выполнено условие (a) и $m \geq 2k$. Тогда любой локально D-оптимальный план для модели (1.1)–(1.2) с функцией регрессии вида (3.1) является локально D-оптимальным планом для этой модели с функцией регрессии (3.2), где $\tilde{\Theta}$ таково, что $\eta_1(x, \Theta) = \eta_2(x, \tilde{\Theta})$ и наоборот.*

Доказательство этой леммы будет дано в следующем разделе.

3.2 Число точек локально D-оптимального плана

Перепишем функцию (3.2) опуская знак волны у θ :

$$\eta(x, \Theta) = \sum_{j=1}^{l} \theta_j x^{j-1} + \sum_{j=1}^{k} \frac{\theta_{j+l}}{x + \theta_{j+l+k}}. \tag{3.3}$$

Положим $\mathfrak{X} = [0, d]$, где d — некоторое положительное число, $\Theta \in \tilde{\Omega}, \tilde{\Omega} = \{\Theta : \eta(x, \Theta) = \eta_1(x, \overline{\Theta})\ \ \overline{\Theta} \in \Omega\}$. Заметим, что информационная матрица для модели (1.1)–(1.2) с фунцией регрессии вида (3.3) совпадает с информационной матрицей для линейной по параметрам функции регрессии

$$\beta^T f(x),$$

где $\beta^T = (\beta_1, \ldots, \beta_m)$ — вектор оцениваемых параметров, $f(x) = f(x, \Theta) = (f_1(x, \Theta), \ldots, f_m(x, \Theta))^T$, Θ фиксировано и

$$\begin{array}{rcl} f_1(x) & \equiv & 1, \ldots, f_l(x) = x^{l-1}, \\ f_{l+1}(x) & = & 1/(x + \theta_{l+k+1}), \ldots, f_{l+k}(x) = 1/(x + \theta_{l+2k}), \\ f_{l+k+1}(x) & = & -\theta_{l+1}/(x + \theta_{l+k+1})^2, \ldots, f_m(x) = \\ & = & f_{l+2k}(x) = -\theta_{l+k}/(x + \theta_{l+2k})^2, \quad i = 1, \ldots, k. \end{array} \tag{3.4}$$

Заметим, что при $\Theta \in \tilde{\Omega}$ выполняется соотношение $\theta_{j+l} \neq 0$, $j = 1, \ldots, k$. Кроме того, в силу предположения (a) $\theta_{1+l+k} > \ldots > \theta_{k+l+k} > 0$. Заметим, что функции (3.4) линейно независимы. Более того, справедлив следующий результат. Пусть

$$\xi = \{x_1, \ldots, x_m, \mu_1, \ldots, \mu_m\} \tag{3.5}$$

есть план эксперимента, $x_1 < \ldots < x_m$, $\mu_j > 0$, $j = 1, \ldots, m$.

Теорема 3.2.1. *Определитель информационной матрицы плана ξ, определенного формулой (3.5), для функций регрессии (3.1) и (3.3) имеет вид*

$$\det M(\xi, \Theta) = \prod_{i=1}^{m} \mu_i \left(C(\Theta) \prod_{1 \le i < j \le m} (x_j - x_i) / \prod_{i=1}^{m} Q^2(x_i) \right)^2, \tag{3.6}$$

где $C(\Theta)$ не зависит от ξ, причем для функции регрессии (3.3)

$$\begin{aligned} Q(x) &= \prod_{j=1}^{k} (x + \theta_{j+l+k}), \\ C(\Theta) &= const \prod_{j=1}^{k} \theta_{j+l} \prod_{1 \le i < j \le m} (\theta_{j+l+k} - \theta_{i+l+k})^4. \end{aligned} \tag{3.7}$$

„®Є § вҐ«мбвў®. Пусть функция регрессии имеет вид (3.3). Рассмотрим определитель матрицы

$$F = (f_j(x_i))_{i,j=1}^{m}.$$

Умножим элементы i-ой строки матрицы F на $Q^2(x_i)$, $i = 1 \ldots m$. Тогда элементами i-ой строки новой матрицы будут многочлены степени $\le m - 1$ от x_i. Легко проверить, что линейным преобразованием столбцов определитель этой матрицы может быть приведен к определителю Вандермонда. Так как

$$\det M(\xi, \Theta) = \mu_1 \cdots \mu_m \det{}^2 F,$$

то отсюда вытекает формула (3.6). Формула (3.7) вытекает из того, что $\det F = 0$ при $\theta_{i+l+k} = \theta_{j+l+k}, i \ne j$. Аналогичным образом устанавливается формула (3.6) для функции регрессии (3.1).

При $\mathfrak{X} = [0, d]$ для функций регрессии (3.1) и (3.3) справедливо следующие утверждение.

Теорема 3.2.2. *Если выполнено условие (a) и одно из условий (b),(c), то локально D-оптимальный план существует, единственен и сосредоточен с равными весами в m точках, одна из которых равна 0. Кроме того, при $m > 2k$ одна из точек плана равна d.*

„®Є § вҐ«мбвў®. Существование локально D-оптимального плана вытекает из непрерывной дифференцируемости функций (3.1) и (3.3) и компактности отрезка $\mathfrak{X}$. Доказательство остальных утверждений проведем для случая функции регрессии (3.3).

Пусть $\xi^* = \{x_1^*, \dots, x_n^*, \mu_1^*, \dots, \mu_n^*\}$ — локально D-оптимальный план. Без ограничения общности положим

$$0 \le x_1^* < x_2^* < \ldots < x_n^* \le d.$$

Заметим, что $n \ge m$, так как иначе $\det M(\xi^*, \Theta) = 0$. Обозначим

$$g(x) = f^T(x) M^{-1}(\xi^*, \Theta) f(x),$$

где $f(x) = f(x, \Theta)$ определено выше и Θ фиксировано. Согласно теореме эквивалентности Кифера-Вольфовица (см. раздел 1.3)

$$g(x) \le m, x \in \mathfrak{X},\ g(x_i^*) = m,\ i = 1, \dots, m.$$

В силу дифференцируемости функции $g(x)$ отсюда вытекает, что

$$g'(x_i^*) = 0, i = 2, \dots, m-1,$$

причем, если $x_1^* \ne 0$, то $g'(x_1^*) = 0$, а если $x_n^* \ne d$, то $g'(x_n^*) = 0$. Функция $\tilde{g}(x) = g(x) - m$ имеет вид суммы многочленов и дробей со знаменателями вида $(x + \theta_{i+l+k})^{s_1}(x + \theta_{j+l+k})^{s_2}$, $s_1, s_2 = 1, 2$, $i, j = 1, \dots, k$. Приводя эти дроби к общему знаменателю, получим, что функция $\tilde{g}(x)$ имеет вид

$$\tilde{g}(x) = \tilde{P}(x)/Q^4(x),$$

где $Q(x) = \prod_{i=1}^k (x + \theta_{i+l+k})$, $\tilde{P}(x)$ — многочлен степени $\le 2m-2$ при $l > 0$ и степени $2m$ при $l = 0$. Так как функция $\tilde{g}(x)$ имеет не менее $2n-2$ нулей с учетом их кратности, то при $l > 0$ имеем $n = m, x_1^* = 0$, $x_m^* = d$. Следовательно, при $l > 0$ функция $\tilde{g}(x)$ имеет вид

$$const\, x(x-d) \prod_{i=2}^{m-1} (x - x_i^*)^2 / Q^4(x).$$

Пусть теперь $l = 0$. Так как $\det M(\xi, \Theta)$ убывает при увеличении всех точек плана на одну и тоже величину в силу формулы (3.6), то $x_1^* = 0$ и x_1^* — корень нечетной кратности функции $\tilde{g}(x)$. Кроме того, при $x \to \infty$ и при $x \to -\infty$ $\tilde{P}(x) \sim -(mx^m)^2$. Следовательно, функция $\tilde{g}(x)$ имеет корень $c_1 < 0$, корни кратности не меньше двух в точках $x_i^*, i = 2, \dots, n-1$, а также либо корень $x_n^* < d$ кратности ≥ 2, либо корень $c_2 = d$ и корень $c_3 > d$. Так как $\tilde{P}(x)$ — многочлен степени $2m$, то $n = m$ и функция $\tilde{g}(x)$ имеет вид

$$const\, x(x - {}_1)(x - {}_2)(x - {}_3) \prod_{i=2}^{m-1} (x - x_i^*)^2 / Q^4(x),$$

где $c_1 < 0$ и либо $x_m^* = c_2 = d, c_3 > d$, либо $c_2 = c_3 = x_m^* < d$. Так как число точек плана равно числу параметров, то веса точек локально D-оптимального плана одинаковы. Предположим, что существуют два различных оптимальных плана ξ_1 и ξ_2. Тогда план $(\xi_1 + \xi_2)/2$ является оптимальным. Однако этот план содержит по крайней мере $m+1$ различных точек, что невозможно.

Из теорем 3.2.1 и 3.2.2, очевидно, вытекает справедливость леммы 3.1.1, так как векторы, образованные точками локально D-оптимальных планов для функций регрессии (3.1) и (3.3), оказываются экстремальными точками функций, совпадающих с точностью до постоянного множителя.

Лемма 3.2.1. *В условиях теоремы 3.2.2 при $m = 2k$ и достаточно большом d имеет место неравенство $x_m^* < d$.*

„®Є § вҐ«мбвў®. Пусть $x_1 = 0 < x_2 < \ldots < x_m$, $\xi = \{x_1, \ldots, x_m, 1/m, \ldots, 1/m\}$, $\delta = \min_{1 \le i \le k} \theta_{i+k}$. Тогда в силу формулы (3.6) при $\overline{x}_i = x_i + \delta$

$$\det M(\xi, \Theta) \le const \frac{1}{\delta^m} \frac{1}{m^m x_m \overline{x}_{m-1}^2 \cdots \overline{x}_2^{m-1}} = \frac{1}{x_m} w(x_2, \ldots, x_{m-1}),$$

где $w(x_2, \ldots, x_{m-1})$ — ограниченная функция, то есть $w(x_2, \ldots, x_{m-1}) \le C_1$, где C_1 — некоторая константа. Отсюда $\det M(\xi, \Theta)$ мал при больших x_m. Следовательно, при достаточно большом d $x_m^* < d$.

3.3 Основное уравнение

Исследуем зависимость точек локально D-оптимального плана от параметров $\theta_{l+k+1}, \ldots, \theta_{l+2k}$, нелинейно входящих в рассматриваемую функцию регрессии, заданную формулой (3.3).

Обозначим $z = (z_1, \ldots, z_k)^T = (\theta_{1+l+k}, \ldots, \theta_{k+l+k})^T$,

$$\begin{aligned} \overline{Z} &= \{z : z_i > \varepsilon, \ i = 1, \ldots, k\}, \varepsilon > 0, \\ Z &= \overline{Z} \cap \{z : z_i \ne z_j, i \ne j, i, j = 1, \ldots, k\}. \end{aligned} \tag{3.8}$$

Пусть сначала $m > 2k$, то есть $l > 0$, $\mathfrak{X} = [0, d]$, где d — произвольное положительное число. Пусть

$$\begin{aligned} b &= (b_1, \ldots, b_{m-2})^T = (x_2, \ldots, x_{m-1})^T, \\ \xi_b &= \{0, x_2, \ldots, x_{m-1}, d; 1/m, \ldots, 1/m\}, \\ \overline{V} &= \{b : 0 \le b_1 \le \ldots \le b_{m-2} \le d\}, \\ V &= \overline{V} \cap \{b : b_i \ne b_j, \ i \ne j, \ i, j = 1, \ldots, m-2, b_1 \ne 0, b_{m-2} \ne d\}. \end{aligned}$$

Введем функцию

$$\varphi(b,z) = \left(\prod_{1\le i<j\le m} (x_j - x_i) / \prod_{i=1}^{m} Q^2(x_i,z) \right)^{2/m}, \tag{3.9}$$

где $Q(x,z) = \prod\limits_{j=1}^{k}(x+z_j)$, $x_1 = 0$, $x_m = d$, $x_j = b_{j-1}$, $j = 2,\ldots,m-1$. В силу теоремы 3.2.1 при любом фиксированном $z \in Z$

$$\varphi(b,z) = const\,[\det M(\xi_b,\Theta)]^{1/m}. \tag{3.10}$$

Определим при любом фиксированном z вектор $\tilde b = \tilde b(z)$

$$\tilde b(z) = \arg\max_{b\in\overline{V}} \varphi(b,z). \tag{3.11}$$

Так как $\varphi(b,z) = 0$ при $b \in \overline{V}\backslash V$ и $\varphi(b,z) > 0$ при $b \in V$, то максимум функции $\varphi(b,z)$ достигается при $b \in V = int\overline{V}$ и, значит, при $b = \tilde b(z)$ выполняется

$$\frac{\partial}{\partial b}\varphi(b,z) = 0. \tag{3.12}$$

Лемма 3.3.1. *При любом фиксированном $z \in Z$ уравнение (3.12) имеет единственное решение на множестве V.*

„®Є § вҐ«мбвў®. Пусть $z \in Z$ фиксировано, $M(\xi) = M(\xi,\Theta)$, $\Theta_2 = z$, $\tilde b = \tilde b(z)$ — любое решение уравнения (3.12),

$$g(x) = f^T(x) M^{-1}(\xi_{\tilde b}) f(x),$$

где $f(x)$ определено формулой (3.4). Так как веса выбраны оптимальным способом, то

$$\frac{\partial}{\partial\alpha} \ln\det M((1-\alpha)\xi_{\tilde b} + \alpha\xi_{x_i}) \Bigg|_{\alpha=+0} \le 0,$$

$x_1 = 0, x_m = d, x_i = \tilde b_{i-1}$, $i = 2,\ldots,m-1$, $\xi_{x_i} = \{x_i, 1\}$. Вычисляя производную, получаем

$$g(x_i) - m \le 0,\ i = 1,\ldots,m.$$

Так как

$$\sum_{i=1}^{m} g(x_i)/m = tr\ M^{-1}(\xi_{\tilde b}, z) M(\xi_{\tilde b}, z) = m,$$

то $g(x_i) = m, i = 1, \ldots, m$. Также имеем

$$\varphi'_{b_i}(\tilde{b}, z) = g'(\tilde{b}_i),\ i = 1, \ldots, m-2.$$

В силу того, что равенство (3.12) выполняется при $b = \tilde{b}$, имеем

$$g'(\tilde{b}_i) = 0,\ i = 1, \ldots, m-2.$$

При доказательстве теоремы 3.2.2 мы установили, что функция $\tilde{g}(x) = g(x) - m$ имеет вид

$$P_{2m-2}(x)/Q^4(x), \tag{3.13}$$

где $P_{2m-2}(x)$ — некоторой многочлен степени $2m-2$. Следовательно,

$$P_{2m-2}(x_i) = 0, i = 1, \ldots, m,\ \ P'_{2m-2}(x_i) = 0, i = 2, \ldots, m-1. \tag{3.14}$$

Отсюда $P_{2m-2}(x) = const\, x(x-d)\prod_{i=2}^{m-1}(x-x_i)^2$. При $x \in [0, d]$ этот многочлен не превосходит 0, а значит и $\tilde{g}(x) \le 0$ при $x \in [0, d]$, то есть $\max_{x \in \mathfrak{X}} g(x) = m$. По теореме эквивалентности Кифера-Вольфовица отсюда следует, что $\xi_{\tilde{b}}$ — локально D-оптимальный план. В силу теоремы 3.2.2 такой план является единственным.

Пусть

$$\begin{aligned} J(b,z) &= \left(\frac{\partial^2}{\partial b_i \partial b_j}\varphi(b,z)\right)_{i,j=1}^{m-2}, \\ J &= J(z) = J(\tilde{b}(z), z). \end{aligned}$$

Лемма 3.3.2. *При $z \in Z$ матрица J является невырожденной.*

„®Є § вҐ«мбвў®. Используем формулу

$$J = E - B^T \mathcal{D}^{-1} B,$$

где $E = diag\{g''(\tilde{b}_1), \ldots, g''(\tilde{b}_{m-2})\}$, $D > 0$, D, B определены формулой (2.27), $g(x) = f^T(x) M^{-1}(\xi_{\tilde{b}}, z) f(x)$. Достаточно проверить, что все диагональные элементы матрицы E отрицательны. Выше мы показали, что $\tilde{g}(x)$ имеет вид (3.13)-(3.14), где $x_1 = 0, x_m = d, x_i = \tilde{b}_{i-1}, i = 2, \ldots, m-1$. Непосредственное дифференцирование показывает, что $g''(\tilde{b}_i) < 0, i = 1, \ldots, m-2$.

Рассмотрим теперь случай $l = 0, m = 2k$. Пусть d достаточно велико, так что $x_m^* < d$ при любом $z \in \overline{Z}$. Пусть

$$\begin{aligned} b &= (b_1, \ldots, b_{m-1})^T = (x_2, \ldots, x_m)^T, \\ \xi_b &= \{0, x_2, \ldots, x_m, 1/m, \ldots, 1/m\}, \\ \overline{V} &= \{b :\ 0 \le b_1 \le \ldots \le b_{m-1} \le d\}, \\ V &= \overline{V} \cap \{b :\ b_i \ne b_j,\ i \ne j,\ i,j = 1, \ldots, m-1, b_1 > 0\}. \end{aligned}$$

Рассмотрим функцию $\varphi(b, z)$, заданную формулой (3.9) при $x_1 = 0, x_j = b_{j-1}, j = 2, \ldots, m$. Пусть $\tilde{b} = \tilde{b}(z)$ определено формулой (3.11) при любом фиксированном z. Тогда, очевидно, что

$$\frac{\partial}{\partial b}\varphi(b, z) = 0$$

при $b = \tilde{b}(z)$. Пусть

$$\begin{aligned} J(b, z) &= \left(\frac{\partial^2}{\partial b_i \partial b_j}\varphi(b, z)\right)_{i,j=1}^{m-1}, \\ J &= J(z) = J(\tilde{b}(z), z). \end{aligned}$$

Заметим, что леммы 3.3.1 и 3.3.2 сохраняют силу, а их доказательство аналогично приведенному выше.

Из теоремы 2.3.1 главы 2 вытекает следующий результат. Пусть при любом $z \in Z$ вектор $\tilde{b}(z)$ вычисляется по формуле (3.11).

Теорема 3.3.1. *Вектор-функция $\tilde{b}(z)$ определена единственным образом и является вещественной аналитической вектор-функцией при $z \in Z$, где Z определено формулой (3.8), а ε — любое положительное число. Координаты этой вектор-функции при $z \in \overline{Z}$ являются точками локально D-оптимального плана для функции регрессии (3.3) при $\mathfrak{X} = [0, d]$, где $d > 0$ произвольно при $l > 0$ и достаточно велико при $l = 0$.*

3.4 Алгебраический подход и предельные планы

В силу теорем 3.2.1 и 3.2.2 задача нахождения локально D-оптимального плана для функции регрессии (3.3) при $\mathfrak{X} = [0, d]$ сводится к нахождению максимума следующей функции

$$T(x_1, \ldots, x_m) = \prod_{i=1}^{m} \omega(x_i) \prod_{1 \le i < j \le m} (x_j - x_i)^2,$$

при $0 \le x_1 < \ldots < x_m \le d$,

$$\omega(x) = \prod_{i=1}^{k} (x + z_i)^{-4},\ 0 < z_1 < \ldots < z_k,\ k \ge 1.$$

Эта задача для некоторых других функций $\omega(x)$, в частности для $\omega(x) = x(d - x)$, была поставлена и решена Стильтьесом (Сеге, 1959, с. 159). Далее мы разработаем метод нахождения локально D-оптимальных планов, основанный на развитии идеи Стильтьеса.

В этом разделе мы будем использовать два параметра r и n определяемых по следующему правилу: если $l = 0$ и d достаточно велико, то положим $r = k-1, n = m-1$, в остальных случаях положим $r = k, n = m-2$.

Пример 3.4.1. Пусть $k = 1, l = 0$, т.е. функция регрессии имеет вид

$$\eta(x) = \frac{\theta_1}{x+\theta_2},\ \theta_2 > 0, \theta_1 \neq 0.$$

Прямое вычисление показывает, что локально D-оптимальный план имеет вид

$$\begin{array}{cc} \{0, \theta_2; 1/2, 1/2\} & d > \theta_2 \\ \{0, d; 1/2, 1/2\} & d \leq \theta_2. \end{array}$$

Пусть $r \geq 1, \psi(x)$ — некоторый многочлен степени n, $\psi(x) = \sum_{i=0}^{n} \psi_i x^{n-i}$, $\psi = (\psi_0, \ldots, \psi_n)^T$ — вектор его коэффициентов. В этом разделе мы будем использовать обозначение

$$f(x) = (x^{r+n}, x^{r+n-1}, \ldots, x, 1)^T.$$

Рассмотрим случай $r \geq 1, l = 0, r = k-1$. Пусть матрица $A = (a_{ij})$ размером $(r+n+1) \times (n+1)$ определена равенством

$$\begin{aligned} f^T(x)A\psi &= \psi''(x)x\prod_{i=1}^{k}(x+z_i) + \\ &+ 2\psi'(x)\left(\prod_{i=1}^{k}(x+z_i) - 2x\sum_{j=1}^{k}\prod_{i\neq j}(x+z_i)\right). \end{aligned} \tag{3.15}$$

Легко проверить, что матрица A определена однозначно и $a_{ij} = 0, j > i, i = 1, \ldots, n-1$. Введем следующие матрицы размера $(r+n+1) \times (n+1)$

$$\begin{aligned} E_0 &= (E, \mathbf{0}_{(n+1)\times r})^T, \\ E_1 &= (\mathbf{0}_{(n+1)\times 1}, E, \mathbf{0}_{(n+1)\times(r-1)})^T, \ldots, E_r = (\mathbf{0}_{(n+1)\times r}, E)^T, \end{aligned}$$

где E — единичная матрица размера $(n+1) \times (n+1)$, $\mathbf{0}_{s_1 \times s_2}$ — нулевая матрица размера $s_1 \times s_2$,

$$B = A - \sum_{i=0}^{r} \lambda_i E_i,\ \lambda_0 = n(n-1) + 2n(1-2k).$$

Элементы матрицы B обозначим через b_{ij}. Пусть $B_{(i)}$ матрица, составленная из $(i+1),\dots,(i+n+1)$-ой строк матрицы B (т.е. $B_{(i)} = E_i^T B$), $B_{(i)} = B_{(i)}(\lambda), \lambda = (\lambda_1,\dots,\lambda_r)$. Положим $Q_i = Q_i(\lambda) = \det B_{(i)}(\lambda)$. Заметим, что при заданных $\lambda_1,\dots,\lambda_r$ элементы вектора ψ удовлетворяют рекуррентным соотношениям

$$\psi_0 = \lambda_0,\ \psi_{s+1} = -\sum_{j=1}^{s} b_{s+1,j}\psi_j/b_{s+1,s+1}, s = 0,\dots,n-1. \tag{3.16}$$

Если $l > 0$ или $l = 0$, но $x_m^* > d$, то заменим равенство (3.15) на равенство

$$\begin{aligned} f^T(x)A\psi &= \psi''(x)x(x-d)\prod_{i=1}^{k}(x+z_i) + \\ &+ 2\psi'(x)\left((2x-d)\prod_{i=1}^{k}(x+z_i) - 2x(x-d)\sum_{j=1}^{k}\prod_{i\neq j}(x+z_i)\right). \end{aligned}$$

Справедлива следующая теорема.

Теорема 3.4.1. *При $r \geq 1$ существует и притом единственное решение задачи*

$$\sum_{i=1}^{r} Q_i^2(\lambda) \to \min_{\lambda\in\mathrm{R}^r} \tag{3.17}$$

такое, что точки локально D-оптимального плана для функции регрессии (3.3), не совпадающие с 0 и d, являются корнями многочлена

$$\psi(x) = \psi_0 x^n + \psi_1 x^{n-1} + \dots + \psi_n,$$

причем коэффициенты $\psi_0,\dots,\psi_n$ могут быть вычислены при λ, на котором достигается минимум, по рекуррентным формулам (3.16).

„®Є § вҐ«мбвў®. Рассмотрим случай $l = 0, d > x_m^*$, где x_m^* — наибольшая точка локально D-оптимального плана для функции регрессии (3.3) на $\mathfrak{X} = [0, d]$. Раннее было показано, что $x_1^* = 0$, а функция $T(0, x_2,\dots,x_m)$ имеет единственную стационарную точку и

$$\frac{\partial}{\partial x_i}T(0, x_2,\dots,x_m) = 0,\ i = 2,\dots,m, \tag{3.18}$$

при $x_j = x_j^*, j = 2,\dots,m$. Рассмотрим функцию $\psi^*(x) = \lambda_0\prod_{i=2}^{m}(x-x_i^*)$. Нетрудно проверить, что для $\psi(x) = \psi^*(x)$ имеет место равенство

$$\frac{1}{2}\frac{\psi''(x_i)}{\psi'(x_i)} = \sum_{j\neq i}\frac{1}{x_i - x_j}$$

при $x_l = x_l^*, l = 2, \ldots, m$. Заметим, что равенство (3.18) имеет вид

$$\sum_{j \neq i} \frac{1}{x_i^* - x_j^*} + \frac{1}{x_i^*} - \sum_{l=1}^{k} \frac{2}{x_i^* + z_l} = 0,\ i = 2, \ldots, m.$$

Следовательно,

$$\frac{1}{2}\frac{\psi''(x)}{\psi'(x)} + \frac{1}{x} - \sum_{l=1}^{k} \frac{2}{x + z_l} = 0$$

при $x = x_i^*, i = 2, \ldots, m$. Приводя левую часть этого равенства к общему знаменателю, получим

$$\psi''(x)x\prod_{l=1}^{k}(x+z_l) + 2\psi'(x)\left(\prod_{l=1}^{k}(x+z_l) - 2x\sum_{i=1}^{k}\prod_{l\neq i}(x+z_l)\right) = 0 \tag{3.19}$$

при $x = x_i^*, i = 2, \ldots, m$. Многочлен в левой части этого равенства имеет степень $\leq n + r$ и делится на $\psi(x)$. Следовательно, он имеет вид

$$f^T(x)A\psi = q(x)\psi(x)/\psi_0, \tag{3.20}$$

где $q(x) = \sum_{i=0}^{r} \lambda_i x^{r-i}$, A — некоторая $(n + r + 1) \times (n + 1)$ матрица. Заметим, что равенство (3.20) можно переписать в виде

$$f^T(x)B\psi = 0,$$

откуда $\det B_{(i)} = 0$, $i = 1, \ldots, r$. Так как стационарная точка функции $T(0, x_2, \ldots, x_m)$ при $0 < x_2 < \ldots < x_m$ существует и единственна в силу леммы 3.3.1, то многочлен $\psi(x)$, имеющий $m - 1$ вещественных корней $x_2, \ldots, x_m$ на $(0, \infty)$ и удовлетворяющий равенству (3.19) при $x = x_i, i = 2, \ldots, m$, существует и определен единственным образом. Следовательно, существует точка $\lambda = \lambda^* \in \mathbb{R}^r$, являющаяся решением задачи (3.17) и такая, что коэффициенты многочлена $\psi(x)$ определяются по рекуррентным формулам (3.16). Так как матрица A определена однозначно и $f^T(x)A\psi = q(x)\psi(x)/\psi_0$, то λ^* определено однозначно.

В случае $d < x_m^*$ а также в случае $l > 0$ рассуждения проводятся аналогично.

Проиллюстрируем применение теоремы 3.4.1 на следующем примере.

Пример 3.4.2. Пусть $k = 2, l = 0$, т.е. функция регрессии имеет вид

$$\eta(x, \Theta) = \frac{\theta_1}{x + \theta_3} + \frac{\theta_2}{x + \theta_4},$$

$\theta_1, \theta_2 \neq 0,\ \theta_3, \theta_4 > 0,\ x \in \mathfrak{X} = [0, d]$, d достаточно велико. В этом случае $n = m - 1 = 3, r = k - 1 = 1$. Обозначим $z_1 = \theta_3, z_2 = \theta_4, \Delta = z_1 + z_2$. Пусть сначала $z_1 z_2 = 1$. Матрица $B_{(1)}$ имеет вид

$$\begin{pmatrix} -\lambda & 2 & 0 & 0 \\ 12 & -2\Delta - \lambda & 6 & 0 \\ 0 & 6 & -2\Delta - \lambda & 12 \\ 0 & 0 & 2 & -\lambda \end{pmatrix},$$

$\det B_{(1)} = (\lambda^2 - 2\Delta\lambda - 24)^2 - 36\lambda^2, \lambda_0 = -12$. Отсюда $\psi_1 = \lambda/2, \psi_2 = ((\lambda+2\Delta)\lambda/2 - 12)/6,\ \psi_3 = (-(\lambda + 2\Delta)\psi_2 - 6\lambda/2)/12, \psi_3 = 2\psi_2/\lambda$. Из этих равенств находим, что $\psi_2 = \pm\lambda/2, \psi_3 = \pm 1$. Так как $\psi(x)$ имеет только положительные корни, то $\psi_1 < 0, \psi_2 > 0$, откуда $\lambda < 0, \psi_2 = -\lambda/2$ и $\lambda^2 + (2\Delta + 6)\lambda - 24 = 0$. Следовательно, единственное решение уравнения $\det B_{(1)} = 0$, для которого соответствующий многочлен $\psi(x)$ имеет только положительные корни, равно

$$\lambda^* = -\Delta - 3 - \sqrt{(\Delta + 3)^2 + 24},$$

а $\psi(x)$ имеет вид

$$x^3 - \frac{\lambda^*}{2}x^2 + \frac{\lambda^*}{2}x - 1 = (x - 1)(x^2 + (1 - \frac{\lambda^*}{2})x + 1).$$

Следовательно,

$$x_3^* = 1,\ x_{2,4}^* = \frac{1}{2}\left(\frac{\lambda^*}{2} - 1 \mp \sqrt{(\frac{\lambda^*}{2} - 1)^2 + 4}\right).$$

Для произвольных z_1, z_2 линейным преобразованием модели получаем

$$x_3^* = \sqrt{z_1 z_2},\ x_{2,4}^* = \frac{\sqrt{z_1 z_2}}{2}\left(\frac{\lambda^*}{2} - 1 \mp \sqrt{(\frac{\lambda^*}{2} - 1)^2 + 4}\right).$$

Локально D-оптимальный план имеет вид

$$\{0, x_2^*, x_3^*, x_4^*; 1/4, 1/4, 1/4, 1/4\}.$$

При $r \geq 2$ получение явных аналитических выражений для точек локально D-оптимального плана с помощью изложенного выше алгебраического метода не представляется возможным. Однако функциональный подход позволяет в этом случае построить разложение точек плана в ряд Тейлора по степеням

$z_1, \ldots, z_k$. Для этого необходимо найти выражение для точек плана в некоторой специально выбранной точке $z_{(0)}$.

Заметим, что при $z \to z_\alpha = (\alpha, \ldots, \alpha)^T$ величина $\det M(\xi, z) \to 0$. Однако вектор $\tilde{b}$, состоящий из точек плана, не совпадающих с концами промежутка $[0, d]$, может быть найден для случая $z_\alpha = (\alpha, \ldots, \alpha)^T$ и имеет своими координатами пределы точек локально D-оптимального плана при $z \to z_\alpha$. Положим $\alpha = 1$. Рассмотрим сначала случай, когда $l = 0$ и d достаточно велико. В этом случае для функции $\psi(x)$, определенной выше, имеем уравнение

$$\begin{aligned}
&\psi''(x)x(x+1) + 2\psi'(x)(x(1-2k)+1) = \lambda_0\psi(x),\\
&\lambda_0 = (m-1)(m-2) + 2(m-1)(1-2k).
\end{aligned}$$

Приравнивая коэффициенты при одинаковых степенях x в левой и правой частях, получим

$$\begin{aligned}
\psi_1 &= \frac{m(m-1)}{2(m-2k-1)},\\
\psi_{s+1} &= \psi_s \frac{(m-s)(m-s-1)}{m(2s-1) - 2(s+1)(1-2k)}, s = 1, 2, \ldots, m-1.
\end{aligned} \tag{3.21}$$

Таким образом справедлива следующая теорема.

Теорема 3.4.2. *При $m = 2k$, $\mathfrak{X} = [0, d]$, где d достаточно велико, ненулевые точки локально D-оптимального плана для функции регрессии (3.3) сходятся при $z \to (1, \ldots, 1)^T$ к нулям многочлена $\psi(x)$, коэффициенты которого при любом $k \geq 1$ могут быть найдены по формулам (3.21).*

Проиллюстрируем результат теоремы 3.4.2 на следующем примере.

Пример 3.4.3. Пусть $k = 3, m = 6$, т.е. функция регрессии имеет вид

$$\eta(x, \Theta) = \frac{\theta_1}{x+\theta_4} + \frac{\theta_2}{x+\theta_5} + \frac{\theta_3}{x+\theta_6},$$

$\theta_1, \theta_2, \theta_3 \neq 0$, $\theta_4, \theta_5, \theta_6 > 0$, $x \in \mathfrak{X} = [0, d]$, d достаточно велико. Пусть $z_i = \theta_{i+3} \to 1, i = 1, 2, 3$. Используя теорему 3.4.2, получим

$$\begin{aligned}
\psi(x) &= x^5 - 15x^4 + 50x^3 - 50x^2 + 15x - 1 =\\
&= (x-1)(x^4 - 14x^3 + 36x^2 - 14x + 1) =\\
&= (x-1)(x^2 + (-7+\sqrt{15})x + 1)(x^2 + (-7-\sqrt{15})x + 1).
\end{aligned}$$

Отсюда получаем

$$\begin{array}{rcl} x_2^* &=& (7+\sqrt{15}-\sqrt{60+14\sqrt{15}})/2 \approx 0.0927, \\ x_3^* &=& (7-\sqrt{15}-\sqrt{60-14\sqrt{15}})/2 \approx 0.3616, \\ x_4^* &=& 1, \\ x_5^* &=& (7-\sqrt{15}+\sqrt{60-14\sqrt{15}})/2 \approx 2.765, \\ x_6^* &=& (7+\sqrt{15}+\sqrt{60+14\sqrt{15}})/2 \approx 10.78. \end{array}$$

С помощью этого предельного плана в следующем разделе мы найдем разложение точек плана по $\Delta_1 = z_2 - 1, \Delta_2 = z_3 - 1$ при $z_1 = 1$. При произвольном z_1 план получается умножением всех точек на z_1.

Рассмотрим теперь случай, когда либо $d > 0$ произвольно и $l > 0$, либо $l = 0$ и d достаточно мало. В обоих случаях согласно теореме 3.2.2 точки x_1^* и x_m^* имеют вид $x_1^* = 0, x_m^* = d$. Пусть $\psi(x) = \prod_{i=2}^{m-1}(x - x_i^*) = \sum_{i=0}^{m-2}\psi_i x^{m-2-i}$, $\psi_0 = 1$. Заметим, что при некотором $\lambda \in \mathbb{R}$ эта функция удовлетворяет уравнению

$$\begin{array}{rcl} \psi''(x)x(x-d)(x+1) &+& 2\psi'(x)((2x-d)-x(x-d)2k)) = \\ &=& (\lambda_0 x + \lambda)\psi(x), \\ \lambda_0 &=& (m-2)(m-3)+2(m-2)(2-2k). \end{array} \qquad (3.22)$$

Пусть A — $m \times (m-1)$ матрица, такая что левая часть уравнения равна

$$f^T(x)A\psi, \;\; f(x) = (x^{m-1}, \ldots, x, 1)^T.$$

Пусть

$$B = B(\lambda) = A - \lambda_0 E_0 - \lambda E_1, \;\; B_{(1)} = B_-,$$

где '$_-$' означает вычеркивание первой строки матрицы, матрицы E_0 и E_1 введены выше. Аналогично доказательству теоремы 3.4.1 можно установить следующий результат.

Теорема 3.4.3. *При $m = 2k$ и достаточно малом $d > 0$ и при $m > 2k$ и произвольном $d > 0$ существует единственное решение уравнения*

$$\det B_{(1)}(\lambda) = 0$$

при $\lambda = \lambda^* \in \mathbb{R}$ такое, что точки локально D-оптимального плана, отличные от 0 и d, сходятся при $z \to (1, \ldots, 1)^T$ к нулям многочлена $\psi(x)$, являющегося решением уравнения (3.23) при $\lambda = \lambda^*$. Коэффициенты этого многочлена могут быть вычислены по рекуррентным формулам (3.16).

Подставляя $z_1 = z_2 = 1$ в выражение для точек плана, полученное в примере 3.4.2, имеем

$$x_4^* \to (5+\sqrt{21})/2 \text{ при } z_1, z_2 \to 1.$$

Пример 3.4.4. Пусть $l = 0, k = 2$, d достаточно мало, а точнее $d < (5+\sqrt{21})/2$. В этом случае

$$\psi(x) = (x - x_2^*)(x - x_3^*) = x^2 + \psi_1 x + \psi_2.$$

Уравнение (3.23) принимает вид

$$\begin{aligned} &2x(x+1)(x-d) + 2(2x+\psi_1)(-2x^2 + (1+4d)x - d) = \\ &= (\lambda_0 x + \lambda)(x^2 + \psi_1 x + \psi_2),\ \lambda_0 = -6. \end{aligned}$$

Матрица $B_{(1)}$ имеет вид

$$\begin{pmatrix} 14d + 6 - \lambda & 2d & 0 \\ -6 & 2(1+4d) - \lambda & 6 \\ 0 & -2d & -\lambda \end{pmatrix},$$

$$\psi_1 = (\lambda - 6 - 14d)/2d,\ \ \psi_2 = (6 + 14d - \lambda)/\lambda. \tag{3.23}$$

Заметим, что при $d = 1$

$$\det B_{(1)} = (\lambda - 10)(\lambda^2 - 20\lambda + 24)$$

и $\lambda^* = 10 + 2\sqrt{19}$ есть единственное решение уравнения $\det B_{(1)}(\lambda) = 0$, для которого соответствующий многочлен $\psi(x)$ имеет два корня внутри интервала $(0, d)$. Отсюда из соображений непрерывности вытекает, что при произвольном d λ^* есть максимальный положительный корень уравнения $\det B_{(1)}(\lambda) = 0$, т.е. уравнения

$$\lambda^3 - (22d + 8)\lambda^2 + (112d^2 + 100d + 12)\lambda - 12d(14d + 16) = 0.$$

Точки оптимального плана имеют вид $x_1^* = 0, x_4^* = d$,

$$x_{2,3}^* = -\frac{\psi_1}{2} \mp \sqrt{\frac{\psi_1^2}{4} - \psi_2},$$

где ψ_1 и ψ_2 определяются по формуле (3.23) при $\lambda = \lambda^*$.

3.5 Разложение в ряд Тейлора

В лемме 3.3.2 мы доказали, что матрица $J = J(z)$ является невырожденной при любом $z \in Z$, а значит и при $z = z_{(0)} = (1, \ldots, 1)^T$. Кроме того, в предыдущем разделе указан метод решения уравнения

$$\frac{\partial}{\partial b}\varphi(b, z) = 0 \tag{3.24}$$

в точке $z = z_{(0)}$. Поэтому для разложения точек локально D-оптимального плана в ряды Тейлора можно использовать Алгоритм 1 из главы 2.

Рассмотрим случай $k = 3, m = 6, d > x^*_m$. Пусть $z_1 = 1, z_2 = 1 + \Delta_1, z_3 = 1 + \Delta_2$. Построим разложение точек локально D-оптимального плана в ряд Тейлора по степеням величин Δ_1 и Δ_2. Для произвольных $z_1 = \gamma, z_2, z_3$ план может быть получен умножением на γ точек плана, построенного при $z_1 = 1, \Delta_1 = z_2/\gamma - 1, \Delta_2 = z_3/\gamma - 1$. Запишем уравнение (3.24) в этом случае

$$\frac{1}{x_j} + \sum_{i \neq j} \frac{1}{x_j - x_i} - 2\left(\frac{1}{x_j + 1} + \frac{1}{x_j + 1 + \Delta_1} + \frac{1}{x_j + 1 + \Delta_2}\right) = 0,$$

$j = 2, \ldots, 6$. Обозначим $u = \Delta_1\Delta_2, v = \Delta_1 + \Delta_2$. Перепишем наше уравнение в виде

$$\frac{1}{x_j} + \sum_{i \neq j} \frac{1}{x_j - x_i} - 2\left(\frac{1}{x_j + 1} + \frac{2x_j + 2 + v}{x_j^2 + (2 + v)x_j + 1 + u + v}\right) = 0,$$

$j = 2, \ldots, 6$. Из этого уравнения следует, что точки оптимального плана являются функциями параметров u и v и могут быть разложены в ряды Тейлора по степеням u и v в окрестности точки $(0, 0)$

$$x^*_{i+1}(u, v) = \tilde{b}_i(u, v) = \sum_{s_1=0}^{\infty}\sum_{s_2=0}^{\infty} b_{i(s_1,s_2)} u^{s_1} v^{s_2},$$

$i = 1, \ldots, 5$. Применяя алгоритм, описанный в главе 2, построим таблицу коэффициентов $\{b_{i(s_1,s_2)}\}$ (табл. 3.1). В этой таблице в блоке с номером (s_1, s_2) представлен вектор коэффициентов $(b_{1(s_1,s_2)}, \ldots, b_{5(s_1,s_2)})^T$.

Таблица 3.1: Коэффициенты разложения.

$s_1 \backslash s_2$	0	1	2	3
0	0.093	0.031	-0.015	0.009
	0.362	0.121	-0.050	0.029
	1.000	0.333	-0.111	0.058
	2.765	0.922	-0.229	0.113
	10.780	3.593	-0.655	0.321
1	0.045	-0.033	0.027	
	0.151	-0.103	0.084	
	0.333	-0.207	0.164	
	0.686	-0.395	0.307	
	1.966	-1.118	0.866	
2	-0.015	0.026		
	-0.048	0.080		
	-0.095	0.155		
	-0.180	0.286		
	-0.509	0.804		
3	0.008			
	0.025			
	0.049			
	0.090			
	0.253			

Литература

1. *Бибиков Ю. Н.* Курс обыкновенных дифференциальных уравнений. М.: Высшая школа, 1991.

2. *Демьянов В.Ф., Малоземов В.Н.* (1972). Введение в минимакс. М.: Наука.

3. *Ермаков С. М., Жиглявский А. А.* (1987). Математическая теория оптимального эксперимента. М.: Наука.

4. *Карлин С., Стадден В.* (1976). Чебышевские системы и их применение в анализе и статистике. М.: Наука.

5. *Мелас В. Б.* (1981). Оптимальные планы для экспоненциальной регрессии // Математические методы планирования эксперимента / Под ред. В. В. Пененко. Новосибирск: Наука. С. 174–198.

6. *Мелас В. Б.* (1999). Общая теория функционального подхода к оптимальному планированию эксперимента. СПб.: Изд-во С.-Петерб. ун-та.

7. *Мелас В. Б., Пепелышев А. Н.* (1999). Степенные разложения неявных функций и локально оптимальные планы эксперимента // Статистические модели с приложениями в эконометрике. СПб.: Изд-во НИХИ СПбГУ. С. 108–117.

8. *Сеге Г.* (1962). Ортогональные многочлены. М.: Наука.

9. *Фихтенгольц Г.М.* (1966). Курс дифференциального и интегрального исчисления. Т. 1. М.: Наука.

10. *JennrichR. J.* (1969). Asymptotic properties of non-linear least squares estimators // Ann. Math. Stat., 40, 633–643.

Оглавление