

Санкт-Петербургский государственный университет

А. Н. Пепелышев

**Некоторые современные методы
исследования оптимальных
планов эксперимента**

У ч е б н о е п о с о б и е

Санкт-Петербург
2006

УДК 519.24
ББК 22.172
П25

Рецензенты: проф. В. Б. Мелас (С.-Петербург. гос. ун-т),
проф. М. К. Чирков (С.-Петербург. гос. ун-т)

*Печатается по постановлению
Редакционно-издательского совета
С.-Петербургского государственного университета*

Пепельышев А.Н.

П25 Некоторые современные методы исследования оптимальных планов эксперимента: Учеб. пособие. — СПб., 2006. — 68 с.

В учебном пособии изложен подход к исследованию оптимальных планов эксперимента на основе асимптотического и функционального подходов.

Предназначено для студентов и аспирантов математико-механического факультета.

ББК 22.172

© А. Н. Пепельышев, 2006
© С.-Петербургский
государственный
университет, 2006

Предисловие

Учебное пособие посвящено ряду подходов к изучению оптимальных планов регрессионных экспериментов. В последние десятилетия исследования в этой области существенно продвинулись, однако результаты не достаточно полно нашли свое отражение в учебной литературе.

В пособии для двух регрессионных моделей применяется асимптотический анализ в сочетании с функциональным и алгебраическим подходами. Асимптотический анализ заключается в исследовании задачи в окрестностях точек специального вида. Функциональный подход состоит в изучении точек и весов оптимальных планов как функций некоторых параметров. Алгебраический подход заключается в составлении уравнения для полинома, корнями которого являются искомые величины. Для удобной компьютерной реализации функционального подхода выводятся эффективные алгоритмы.

Пособие состоит из трех разделов. В пункте 1.1 приводится ряд основных понятий и результатов необходимых для строгой постановки задачи оптимального планирования. В пункте 1.2 описывается функциональный подход к изучению оптимальных планов. В пункте 1.3 строятся рекуррентные формулы для вычисления коэффициентов разложения в ряды Тейлора неявно заданных функций. В разделе 2 эти формулы применяются к построению оптимальных планов для тригонометрической регрессии на произвольном отрезке. Для этой модели и для полиномиальной модели с коррелированными наблюдениями, изучаемой в разделе 3, применяется асимптотический анализ.

Отметим, что раздел 1 носит обзорный характер. Более полное развитие ряда основных понятий можно найти в учебнике Ермакова и Жиглявского (1987). Оригинальные теоретические результаты в разделах 2 и 3 получены автором совместно с проф. Н. Dette и проф. В. Б. Меласом. Они опубликованы в ряде совместных работ на английском языке, указанных в списке литературы. Все алгоритмы и соответствующие численные результаты получены автором.

1. Основные понятия теории оптимального планирования

1.1. Постановка задачи и обзор основных результатов теории оптимального планирования

На практике мы часто сталкиваемся с ситуациями, когда нам требуется проанализировать функциональную зависимость между одной переменной y (наблюдаемым откликом, измеряемой величиной) и другой переменной x (экспериментальными условиями), принадлежащей некоторой области \mathcal{X} планирования эксперимента. Эта функция зависит от неизвестных значений параметров. Мы измеряем значения функции в некоторых точках с тем, чтобы оценить неизвестные параметры. Во многих экспериментах эта функция измеряется только лишь с точностью до аддитивной случайной ошибки. В таких случаях эксперимент называется статистическим (Ермаков, 1983). Переходя от реальных экспериментов к математическому описанию, мы неизбежно приходим к математическим моделям эксперимента. Наука, занимающаяся проблемами, описанными выше, называется теорией оптимального планирования эксперимента.

1.1.1. Постановка задачи

Пусть результаты эксперимента $y_1, \dots, y_N \in \mathbb{R}$ описываются уравнением

$$y_j = \eta(x_j, \beta) + \varepsilon_j, \quad j = 1, \dots, N, \quad (1)$$

где $x_1, \dots, x_N \in \mathcal{X}$ — условия проведения эксперимента — точки некоторого множества планирования \mathcal{X} , $\eta(x, \beta)$ — известная вещественная функция, $\beta = (\beta_1, \dots, \beta_m)^T \in \Omega$ — вектор неизвестных параметров, N — общее число измерений, $\varepsilon_1, \dots, \varepsilon_N$ — ошибки измерений — некоррелированные случайные величины такие, что $E\varepsilon_j = 0$, $E\varepsilon_j^2 = \sigma^2$, $j = 1, \dots, N$. Большая часть результатов в теории планирования эксперимента получена при выполнении этих стандартных предположений относительно ошибок измерений. Постановка задачи в случае неравноточных или коррелированных измерений будет дана в параграфе 1.1.5.

Функцию $\eta(x, \beta)$ часто называют регрессионной моделью (Ермаков, 1983). Модель (1) называется линейной по параметрам, ес-

ли функцию регрессии $\eta(x, \beta)$ можно представить в виде $\beta^T h(x)$, где $h(x) = (h_1(x), \dots, h_m(x))^T$ — вектор функций, не зависящих от β . Обычно предполагают, что функции $h_1(x), \dots, h_m(x)$ являются непрерывными и линейно независимыми функциями. Модели, для которых такое представление невозможно, называются нелинейными по параметрам.

Таблица

$$\xi_N = \begin{pmatrix} x_1 & \dots & x_N \\ \frac{1}{N} & \dots & \frac{1}{N} \end{pmatrix} \quad (2)$$

называется дискретным (нормированным) планом эксперимента (точки плана могут совпадать). Дж. Кифером это понятие было обобщено. Под непрерывным (приближенным) планом эксперимента понимается дискретная вероятностная мера

$$\xi_N = \begin{pmatrix} x_1 & \dots & x_n \\ w_1 & \dots & w_n \end{pmatrix}, \quad (3)$$

где $x_i \neq x_j$, $\sum_{j=1}^n w_j = 1$, $w_j \geq 0$, n — число различных точек в плане, x_j — условия проведения эксперимента, w_j — весовые коэффициенты. При практическом использовании непрерывных планов следует проводить приблизительно $w_j N$ измерений в точке x_j , $j = 1, \dots, n$. Методы округления чисел $w_j N$ до целых N_j (таких что $\sum N_j = N$) можно найти в работах (Федоров, 1971a; Pukelsheim, 1993).

Целью эксперимента является оценка истинного значения β^* вектора параметров β . Параметр β оценивается с помощью некоторой статистики $\hat{\beta} = \hat{\beta}(x_1, \dots, x_N, y_1, \dots, y_N)$. Выбор статистики $\hat{\beta}$ составляет предмет отдельной теории. Пусть оценивание производится по методу наименьших квадратов. По определению МНК оценка $\hat{\beta}$ является решением экстремальной задачи

$$\sum_{j=1}^N (\eta(x_j, \beta) - y_j)^2 \rightarrow \min_{\beta \in \Omega}.$$

Эта оценка обеспечивает наиболее точное в определенном смысле оценивание вектора параметров при любом фиксированном плане эксперимента и обладает замечательным асимптотическим свойством.

Теорема 1.1. (теорема Гаусса–Маркова)

Пусть \mathcal{X}, Ω – замкнутые ограниченные множества, функции $h_1(x), \dots, h_m(x)$ являются непрерывными и линейно независимыми на \mathcal{X} , $\det X^T X \neq 0$, где $X = (h(x_1) \dots h(x_N))^T$, $Y = (y_1, \dots, y_N)^T$.

Тогда для линейной по параметрам модели $\eta(x, \beta) = \beta^T h(x)$ оценки метода наименьших квадратов определены однозначно, имеют вид

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (4)$$

и являются наилучшими линейными несмешенными оценками, причем

$$\mathbf{D}\hat{\beta} = \sigma^2 (X^T X)^{-1}.$$

Доказательство этой теоремы можно найти в (Rao, 1968; Pukelsheim, 1993). Справедливо следующее обобщение теоремы 1.1 на случай нелинейных по параметрам моделей.

Теорема 1.2. Пусть выполнены предположения:

1. Множества $\mathcal{X} \in \mathbb{R}^k$, $\Omega \in \mathbb{R}^m$ являются замкнутыми и ограниченными.
2. Функция $\eta(x, \beta)$ непрерывна на $\mathcal{X} \times \Omega$.
3. Последовательность планов ξ_N слабо сходится к плану ξ при $N \rightarrow \infty$, где ξ_N – N -точечный дискретный нормированный план.
4. Величина $\int (\eta(x, \beta) - \eta(x, \beta^o))^2 \xi(dx)$ равна нулю только при $\beta = \beta^o$.
5. Производные $\frac{\partial \eta}{\partial \beta_i}$, $\frac{\partial^2 \eta}{\partial \beta_i \partial \beta_j}$, $i, j = 1, \dots, m$, существуют и непрерывны.
6. Истинное значение β^* вектора параметров является внутренней точкой Ω .
7. Матрица

$$M(\xi) = M(\xi, \beta) = \int_{\mathcal{X}} f(x, \beta) f^T(x, \beta) \xi(dx), \quad (5)$$

где

$$f(x) = f(x, \beta) = \left(\frac{\partial \eta}{\partial \beta_1}(x, \beta), \dots, \frac{\partial \eta}{\partial \beta_m}(x, \beta) \right)^T, \quad (6)$$

является невырожденной при $\beta = \beta^*$.

Пусть $\hat{\beta}_N$ — МНК-оценка параметров согласно плану ξ_N .

Тогда последовательность оценок $\hat{\beta}_N$ сильно состоятельна.

Кроме того, последовательность распределений случайных векторов $\sqrt{N}(\hat{\beta}_N - \beta^*)$ сходится к нормальному распределению с нулевым вектором средних и дисперсионной матрицей $\sigma^2 M^{-1}(\xi, \beta^*)$.

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

Матрица $M(\xi)$ называется информационной матрицей, а вектор $f(x)$ — вектором регрессии. Вектор $f(x)$ совпадает с вектором $h(x)$ для линейной по параметрам модели вида $\beta^T h(x)$. Отметим, что некоторые условия в теореме 1.2 могут быть ослаблены.

Как видно из теорем 1.1 и 1.2, одним из способов повышения точности оценок является путь оптимального выбора условий проведения эксперимента. Это и составляет предмет теории оптимального планирования.

Сформулируем некоторые результаты, касающиеся исследования информационных матриц.

Пусть Ξ_n — множество непрерывных планов, сосредоточенных в n точках с ненулевыми весовыми коэффициентами, $\Xi = \cup_{n=1}^{\infty} \Xi_n$ — множество всех непрерывных планов,

$$\mathcal{M} = \{M : M = M(\xi), \xi \in \Xi\}$$

есть множество всех информационных матриц. Основные свойства информационных матриц даны в теореме 1.3.

Теорема 1.3.

1. Любая информационная матрица является неотрицательно определенной.

2. Если $n < m$, то $\det M(\xi) = 0$ для всех $\xi \in \Xi_n$.

3. Множество \mathcal{M} является выпуклым.

4. При выполнении условий 1, 2 и 5 теоремы 1.2 множество \mathcal{M} , рассматриваемое как множество векторов, состоящих из диагональных и наддиагональных элементов, является ограниченным и замкнутым множеством \mathbb{R}^s , $s = m(m+1)/2$.

5. Для любого плана $\xi \in \Xi$ найдется план $\tilde{\xi} \in \Xi_n$ с числом точек $n \leq m(m+1)/2 + 1$ такой, что $M(\tilde{\xi}) = M(\xi)$.

Отметим, что для нелинейных по параметрам моделей вектор $f(x) = f(x, \beta)$ зависит от β и, следовательно, информационная матрица $M(\xi) = M(\xi, \beta)$ также зависит от β . Поэтому в теореме 1.3 мы подразумеваем, что вектор β фиксирован.

Доказательство теоремы 1.3 можно найти в книге (Карлин, Стадден, 1976).

В силу пункта 5 теоремы 1.3 мы можем ограничить свое внимание непрерывными планами с конечным носителем. Полезным оказывается следующее представление информационной матрицы плана $\xi \in \Xi_n$:

$$M(\xi) = \tilde{F}W\tilde{F}^T, W = \text{diag}\{w_1, \dots, w_n\}, \tilde{F} = (f(x_1) \vdots \dots \vdots f(x_n)).$$

В силу формулы Бине—Коши мы имеем

$$\det M(\xi) = \sum_{1 \leq i_1 < \dots < i_m \leq n} w_{i_1} \cdots w_{i_m} (\det F(x_{i_1}, \dots, x_{i_m}))^2,$$

$$F(x_{i_1}, \dots, x_{i_m}) = (f(x_{i_1}) \vdots \dots \vdots f(x_{i_m})).$$

Как правило, не существует такого плана $\bar{\xi}$, для которого

$$M^{-1}(\bar{\xi}) \leq M^{-1}(\xi)$$

для любого плана $\xi \in \Xi$. Поэтому в качестве критериев оптимальности используются некоторые вогнутые функции, заданные на множестве информационных матриц, или выпуклые функции, заданные на множестве дисперсионных матриц и имеющие строгий статистический смысл.

1.1.2. Критерии оптимальности

Мы рассмотрим ряд критериев, которые являются наиболее важными в теории планирования эксперимента.

D-критерий. Критерий *D*-оптимальности имеет вид

$$(\det M(\xi))^{1/m} \rightarrow \sup_{\xi \in \Xi} .$$

Этот критерий соответствует требованию минимизации объема доверительного эллипсоида

$$\{\beta : (\beta - \hat{\beta})^T M^{-1}(\beta - \hat{\beta}) \leq \gamma\},$$

где γ — некоторая постоянная (зависящая только от доверительного уровня).

c-*критерий*. Критерий *c*-оптимальности имеет вид

$$c^T M^-(\xi) c \rightarrow \inf_{\xi \in \Xi_c},$$

где $\Xi_c = \{\xi \in \Xi : \exists u \in \mathbb{R}^m, c = M(\xi)u\}$, M^- — обобщенно обратная матрица к M . Этот критерий означает требование минимизации дисперсии линейной комбинации оценок параметров $c^T \hat{\beta}$.

e_i-*критерий*. Критерий *e_i*-оптимальности имеет вид

$$e_i^T M^-(\xi) e_i \rightarrow \inf_{\xi \in \Xi_{e_i}},$$

где $e_i = (0, \dots, 0, 1, 0, \dots, 0)^T$ — *i*-й орт евклидова пространства \mathbb{R}^m , и является частным случаем *c*-критерия. Этот критерий соответствует требованию минимизации дисперсии $(M^-(\xi))_{ii}$ оценки параметра β_i (*i*-й координаты вектора β).

E-*критерий*. Критерий *E*-оптимальности имеет вид

$$\lambda_{\max}(M^{-1}(\xi)) \rightarrow \inf_{\xi}.$$

Этот критерий означает требование минимизации максимальной оси доверительного эллипсоида, а также минимизацию наихудшей дисперсии линейной комбинации оценок параметров.

В ряде экспериментов нас может интересовать только часть всех параметров, т.е. экспериментатор хочет с большей точностью оценить только часть параметров, например, только β_1, \dots, β_k , $1 \leq k < m$. В этом случае критерием оптимальности может служить любой из выше рассмотренных критериев, заданных только на блоке матрицы $M^-(\xi)$, а именно на блоке, составленном из первых k строк и столбцов матрицы $M^-(\xi)$ и являющимся асимптотической дисперсионной матрицей оценок вектора параметров $(\hat{\beta}_1, \dots, \hat{\beta}_k)^T$. В общем случае критерий оптимальности может быть записан в виде

$$\Psi(M^-(\xi)) \rightarrow \inf_{\xi}.$$

План эксперимента ξ^* называется оптимальным, если на нем достигается минимум $\Psi(M^-(\xi))$.

Для нелинейных по параметрам моделей информационная матрица зависит от параметра β , и, следовательно, минимум величины $\Psi(M^-(\xi))$ также зависит от параметра β . Поэтому план

$\xi^*(\beta) = \arg \inf \Psi(M^-(\xi, \beta))$ называется локально оптимальным, поскольку он является оптимальным при одном фиксированном значении параметров β . Такие планы впервые рассматривались в работе (Chernoff, 1953). Такой подход к планированию эксперимента, заключающийся в изучении зависимости $\xi^*(\beta)$ от истинных значений параметров, называется локально оптимальным. Трудность этого подхода состоит в том, что при практическом применении мы не знаем истинного значения β^* параметров, а знаем лишь его некоторую оценку $\hat{\beta}$ (априорно или из предварительных экспериментов). Если план $\xi^*(\hat{\beta})$ является эффективным (в определенном смысле) по отношению к плану $\xi^*(\beta^*)$, то план $\xi^*(\hat{\beta})$ является пригодным для практического использования.

Другой возможностью преодоления указанной трудности является использование последовательного, минимаксного или байесовского подходов.

Последовательный подход заключается в том, что всю совокупность экспериментов разбивают на серии. Первая серия должна быть спланирована таким образом, чтобы на основе ее результатов можно было оценить (хотя бы грубо) все параметры регрессии. Для этого достаточно, чтобы $\inf_{\beta \in \Omega} \det M(\xi, \beta) > 0$. Обозначим оценку, полученную на основе обработки результатов первой серии, через $\hat{\beta}_{(1)}$. При планировании второй серии план выбирается на основе одного из известных критериев оптимальности, причем в информационную матрицу подставляется значение $\hat{\beta}_{(1)}$. Аналогично планируются следующие серии. Впервые такой подход без строгого обоснования был разработан в работе (Box, Hunter, 1965) для критерия D -оптимальности. Асимптотическая оптимальность последовательной процедуры, аналогичной процедуре Бокса–Хантера, была в общем случае доказана в работе (Федоров, 1971).

Последовательное планирование может оказаться чересчур громоздким или даже неосуществимым в некоторых практических ситуациях из-за необходимости чередовать эксперимент и вычисления. При наличии априорной информации о параметрах возможно находить оптимальные планы априори, на основе минимаксного или байесовского принципов. К положительным качествам такого априорного планирования следует отнести возможность составления таблиц оптимальных планов и возможность определения необходимого количества экспериментов для достижения заданной точности оценок.

При байесовском подходе план называется оптимальным, если он является решением задачи

$$\int_{\Omega} \Psi(M^-(\xi, \beta)) \pi(d\beta) \rightarrow \inf_{\xi},$$

где $\pi(d\beta)$ — априорная мера, заданная на множестве Ω значений параметров.

При минимаксном подходе план называется оптимальным, если он является решением задачи

$$\min_{\beta \in U} \frac{\min_{\zeta} \Psi(M^-(\zeta, \beta))}{\Psi(M^-(\xi, \beta))} \rightarrow \max_{\xi},$$

где U — априорное подмножество множества Ω .

Отметим, что задача построения локально оптимального плана является подзадачей во всех этих трех подходах.

1.1.3. Критерии эффективности

Для сравнения планов эксперимента нужно иметь меру превосходства одного плана над другим. Для этого в теории планирования есть понятие эффективности.

Эффективностью плана ξ_1 по отношению к плану ξ_2 называется величина

$$\text{eff}_{\Psi} = \frac{\Psi(M^-(\xi_2))}{\Psi(M^-(\xi_1))}.$$

Эта величина говорит нам, во сколько раз мы потеряем в точности оценок при применении плана ξ_1 по сравнению с применением плана ξ_2 , если общее число экспериментов одинаково, или во сколько раз больше потребуется экспериментов для достижения той же точности оценок при применении плана ξ_1 по сравнению с использованием плана ξ_2 .

Рассмотрим следующие частные случаи эффективности, являющиеся эффективностями по отношению к локально оптимальным планам.

D -эффективностью называется величина

$$\text{eff}_D(\xi) = \frac{(\det M(\xi))^{1/m}}{\sup_{\zeta} (\det M(\zeta))^{1/m}}.$$

E -эффективностью называется величина

$$\text{eff}_E(\xi) = \frac{\lambda_{\min}(M(\xi))}{\sup_{\zeta} \lambda_{\min}(M(\zeta))}.$$

e_i -эффективностью называется величина

$$\text{eff}_i(\xi) = \frac{\inf_{\zeta} e_i^T M^-(\zeta) e_i}{e_i^T M^-(\xi) e_i}.$$

1.1.4. Теоремы эквивалентности

Центральными элементами в теории оптимального планирования являются теоремы эквивалентности и двойственные задачи. Теоремы эквивалентности устанавливают необходимые и достаточные условия оптимальности плана. Одновременный анализ пары двойственных задач предоставляет дополнительные методологические возможности.

Исторически первой появилась теорема эквивалентности для D -критерия в работе (Kiefer, Wolfowitz, 1960).

Теорема 1.4. План ξ^* является D -оптимальным тогда и только тогда, когда функция $d(x, \xi^*) = f^T(x) M^{-1}(\xi^*) f(x)$ удовлетворяет неравенству

$$\max d(x, \xi^*) \leq m.$$

Кроме того, $d(x_i^*, \xi^*) = m$, где x_i^* — точка плана ξ^* .

Для E -критерия справедлива следующая теорема эквивалентности (Мелас, 1997).

Теорема 1.5. План ξ^* является E -оптимальным тогда и только тогда, когда существует неотрицательно определенная матрица A^* , такая что $\text{tr} A^* = 1$ и

$$\max_{x \in \mathcal{X}} f^T(x) A^* f(x) \leq \lambda_{\min}(M(\xi^*)).$$

Кроме того, $f^T(x_i^*) A^* f(x_i^*) = \lambda_{\min}(M(\xi^*))$, $i = 1, \dots, n$ и матрица A^* может быть представлена в виде

$$A^* = \sum_{i=1}^s \alpha_i p_{(i)} p_{(i)}^T,$$

где s — кратность минимального собственного числа, $\alpha_i \geq 0$, $\sum \alpha_i = 1$, $\{p_{(i)}\}$ — ортонормированные собственные векторы, отвечающие минимальному собственному числу.

Сформулируем теперь теорему эквивалентности для e_i -критерия (Dette, et al., 2004).

Теорема 1.6. План ξ^* является e_i -оптимальным тогда и только тогда, когда существует положительное число r и вектор $q \in \mathbb{R}^{m-1}$ такие, что функция $\psi(x) = q_1 f_1(x) + \dots + q_{i-1} f_{i-1}(x) + f_i(x) + q_i f_{i+1}(x) + \dots + q_{m-1} f_m(x)$ удовлетворяет условиям:

$$\begin{aligned} r\psi^2(x) &\leq 1 \text{ при } x \in \mathcal{X}, \\ \text{supp } \xi^* &\subset \{x \in \mathcal{X} : r\psi^2(x) = 1\}, \\ \int \psi(x) f_j(x) \xi^*(dx) &= 0, \quad j = 1, \dots, m, \quad j \neq i. \end{aligned}$$

Кроме того, $r = e_i^T M^-(\xi^*) e_i$.

Еще раз подчеркнем, что основным инструментом для нахождения оптимальных планов являются теоремы эквивалентности. Однако в своей обычной формулировке такие теоремы являются скорее средством проверки на оптимальность некоторого плана, чем методом поиска оптимальных планов. Здесь мы ввели формулировки теорем эквивалентности, в известной мере, свободной от этого недостатка.

Для максиминных и байесовских критериев также есть теоремы эквивалентности. В общем случае она получена в статье (Dette, Haines, Imhof, 2003).

Теорема 1.7. План ξ^* является максиминно (эффективным) D -оптимальным тогда и только тогда, когда существует распределение π^* на множестве

$$N(\xi^*) = \left\{ \tilde{\beta} \in \Omega \mid \text{eff}_D(\xi^*, \tilde{\beta}) = \min_{\beta \in \Omega} \text{eff}_D(\xi^*, \beta) \right\}$$

такое, что неравенство

$$\int_{N(\xi^*)} f^T(x, \beta) M^{-1}(\xi^*, \beta) f(x, \beta) d\pi^*(\beta) \leq m \tag{7}$$

выполняется для всех $x \in \mathcal{X}$. Кроме того, равенство в (7) достигается во всех точках максиминного D -оптимального плана.

Множество $N(\xi)$ это множество тех значений параметров, на которых достигает минимум эффективности для плана ξ .

Идея получения неравенств в теоремах эквивалентности простая. Пусть $\Psi(M^-(\xi))$ — некоторый критерий оптимальности и пусть ξ^* — оптимальный план. Построим план $\xi_\alpha = (1 - \alpha)\xi^* + \alpha\xi_x$, где $\xi_x = \{x; 1\}$ — план сосредоточенный в точке x . По условию $\Psi(M^-(\xi_\alpha)) \geq \Psi(M^-(\xi^*))$. Тогда требуемое неравенство получается следующим образом:

$$\frac{\partial}{\partial \alpha} \Psi(M^-(\xi_\alpha)) \Big|_{\alpha=+0} = \lim_{\alpha \rightarrow +0} \frac{\Psi(M^-(\xi_\alpha)) - \Psi(M^-(\xi^*))}{\alpha} \geq 0,$$

т.е. надо вычислить производную по направлению.

1.1.5. Неравноточные и коррелированные измерения

Сначала рассмотрим случай неравноточных измерений. Пусть результаты эксперимента $y_1, \dots, y_N \in \mathbb{R}$ описываются уравнением

$$y_j = \beta^T h(x_j) + \varepsilon_{x_j}, \quad j = 1, \dots, N, \quad (8)$$

где $\varepsilon_{x_1}, \dots, \varepsilon_{x_N}$ — ошибки измерений — независимые случайные величины такие, что $\mathbf{E}\varepsilon_{x_j} = 0$, $\mathbf{E}\varepsilon_{x_j}^2 = \sigma^2 \lambda^{-1}(x_j)$, $j = 1, \dots, N$.

Пусть функция $\lambda(x)$ известна. Тогда, делая подстановку

$$y\sqrt{\lambda(x)} \rightarrow y, \quad h(x)\sqrt{\lambda(x)} \rightarrow h(x),$$

мы переходим к модели с равноточными измерениями, для которой результаты сформулированы в пункте 1.1.1. Таким образом, оценку параметров для модели (8) следует производить по методу взвешенных наименьших квадратов

$$\sum_{j=1}^N \lambda(x_j)(\beta^T h(x_j) - y_j)^2 \rightarrow \min_{\beta \in \Omega}$$

и информационная матрица принимает вид

$$M(\xi) = \int \lambda(x) h(x) h^T(x) \xi(dx).$$

Следуя этим формулам, функцию $\lambda(x)$ называют весовой. Для нелинейных по параметрам моделей рассуждения проводятся аналогично. В общем случае функция $\lambda(x) = \lambda(x, \beta, \theta)$ может зависеть

как от параметров модели, так и от дополнительных параметров θ . Например, весовая может иметь вид $\lambda(x, \beta, \theta) = (\eta(x, \beta))^{\theta}$. В этом случае, обобщенный вариант теоремы Гаусса—Маркова справедлив при дополнительных ограничениях (Pao, 1968), причем для ковариационной матрицы оценок метода взвешенных наименьших квадратов асимптотически выполняется

$$N\text{Cov}\hat{\beta} \approx \sigma^2 M^{-1}(\xi), \quad M^{-1}(\xi) = \int \lambda(x, \beta^*, \theta^*) f^T(x, \beta^*) f(x, \beta^*) \xi(dx).$$

Рассмотрим теперь вопрос планирования эксперимента для случая коррелированных наблюдений. Пусть результаты эксперимента $y_1, \dots, y_N \in \mathbb{R}$ описываются уравнением

$$y_j = \beta^T f(x_j) + \varepsilon_{x_j}, \quad j = 1, \dots, N, \quad (9)$$

где ошибки измерений имеют нулевое среднее и корреляционную структуру в форме

$$\text{Cov}(y_i, y_j) = \sigma^2 (\gamma \rho(x_i, x_j) + (1 - \gamma) \delta_{i,j}),$$

где $\gamma \in [0, 1]$, $\delta_{i,j}$ — символ Кронекера и $\rho(x_i, x_j)$ — известная функция. Обычно предполагают (Hoel, 1958; Bickel, Herzberg, 1979), что ковариация зависит от расстояния между наблюдениями, что приводит к условию $\rho(x_i, x_j) = \tilde{\rho}(x_i - x_j)$. При $\gamma = 0$ получаем случай некоррелированных наблюдений. При $\gamma < 1$ диагональные элементы ковариационной матрицы

$$\Sigma = (\text{Cov}(y_i, y_j))_{i,j=1}^N$$

больше вне диагональных, что означает, что повторные измерения в одной и той же точке дают дополнительную информацию. При $\gamma = 1$ говорят, что ошибки имеет авторегрессионную структуру.

Для получения оценок модель эксперимента перепишем в векторной форме

$$Y = F^T \beta, \quad Y = (y_1, \dots, y_N)^T, \quad F = (h(x_1) : \dots : h(x_N)).$$

Оценки параметров по методу наименьших квадратов определяются по следующей формуле:

$$\hat{\beta}_o = (F^T F)^{-1} F^T Y,$$

и ковариационная матрица для $\hat{\beta}_o$ имеет вид

$$\text{Cov}\hat{\beta}_o = (F^T F)^{-1} F^T \Sigma F (F^T F)^{-1}.$$

Оценки параметров по методу взвешенных наименьших квадратов вычисляются по формуле

$$\hat{\beta} = (F^T \Sigma^{-1} F)^{-1} F^T \Sigma^{-1} Y$$

и ковариационная матрица для $\hat{\beta}$ имеет вид

$$\text{Cov}\hat{\beta} = (F^T \Sigma^{-1} F)^{-1}.$$

В книге (Fedorov, Hackl, 1997) получены результаты при других условиях на корреляционную функцию.

1.2. Идея и методология функционального подхода

Функциональный подход к исследованию оптимальных планов был разработан В. Б. Меласом. Этот подход был впервые предложен в работе (Мелас, 1978). Дальнейшее развитие он получил с появлением персональных компьютеров (Мелас, 1997; Мелас, 1999а, 1999б; Мелас, Пепельышев, 1999).

Функциональный подход состоит в исследовании точек и весов оптимального плана как функций некоторых вспомогательных величин $u = (u_1, \dots, u_k)^T$. Эти величины могут быть истинными значениями параметров и/или метрическими характеристиками множества планирования. Например, если областью планирования является интервал, то в качестве таких величин могут быть использованы его длина и местоположение его середины.

Функциональный подход опирается на теорему о неявном отображении, позволяющую установить аналитичность. Как известно, аналитическая функция может быть разложена в сходящийся ряд Тейлора и начальный отрезок разложения можно рассматривать как хорошее приближение к самой функции. Коэффициенты разложения в ряд Тейлора могут быть вычислены, например, с помощью пакета символьной обработки Maple.

Опишем более подробно методологию функционального подхода.

Во-первых, необходимо составить вектор Θ , состоящий из нетривиальных точек, весов оптимального плана и, возможно, еще некоторых вспомогательных функций. Например, вспомогательные функции могут быть элементами двойственной задачи.

Вообще говоря, число точек в оптимальном плане может быть различным при разных значениях параметров. Поэтому множество U допустимых значений u следует разбить на подобласти, внутри которых число точек оптимального плана не меняется. Если требуется, то множество U следует размельчить еще больше, так, чтобы в каждой подобласти не нарушалась аналитичность оптимального плана и его характеристик. Например, это следует сделать в случае E -критерия, если меняется кратность минимального собственного числа информационной матрицы локально E -оптимального плана (поскольку в точке смены кратности будет нарушаться аналитичность оптимального плана). В результате мы должны получить, что вектор $\Theta = \Theta^*(u)$ имеет фиксированную размерность.

Следующий шаг состоит в составлении уравнения $g(\Theta, u) = 0$ (где g — вектор-функция, размерность которой равна размерности Θ), решением которой является вектор $\Theta^*(u)$, соответствующий оптимальному плану. Это уравнение можно рассматривать как уравнение задающее искомую вектор-функцию $\Theta^*(u)$ неявным образом. Это уравнение может отражать необходимые условия экстремума прямой задачи (критерия оптимальности). Оно также может включать уравнения, связывающие элементы прямой и двойственной задачи.

Третий шаг состоит в изучении невырожденности матрицы Якоби введенного уравнения $g(\Theta, u) = 0$ в каждой подобласти. Это условие является основным в теореме о неявном отображении. Доказав невырожденность якобиана в некоторой подобласти возможных значений переменной u , мы сразу получаем аналитичность вектор-функции $\Theta^*(u)$ в этой подобласти.

Для рекуррентного вычисления коэффициентов разложения функции в ряд Тейлора в некоторой точке необходимо знать значение функции в этой точке. Поэтому на четвертом шаге мы выбираем точку разложения, в которой якобиан является невырожденным, и строим в ней оптимальный план. В ряде случаев удается построить оптимальный план в некоторой точке в явном аналитическом виде.

На пятом шаге мы строим разложение в ряд функции $\Theta^*(u)$.

Далее нам необходимо провести анализ построенного разложения. Теоретически найти радиус сходимости не удается. Поэтому исследование радиуса сходимости остается проводить численным образом.

1.3. Рекуррентные формулы для вычисления коэффициентов рядов Тейлора неявных функций

Настоящий пункт посвящен построению рекуррентных формул для вычисления коэффициентов разложения точек и весов оптимальных планов, рассматриваемых как неявно заданные функции некоторых вспомогательных величин. Формулы для вычисления коэффициентов рядов Тейлора неявно заданных функций были известны уже Эйлеру. Однако отсутствие компьютеров в те времена наложило на их вид свой отпечаток. Описываемые далее рекуррентные формулы, удобные для вычисления в математических пакетах, являются оригинальными результатами автора. Практические расчеты в пакете Maple показывают, что построенные алгоритмы являются эффективными. Кроме того, разработаны рекуррентные формулы в случае, когда якобиан системы уравнений имеет вырожденность некоторого специального типа, которая имеет место в рассматриваемых задачах теории оптимального планирования.

1.3.1. Постановка задачи

Пусть m, k – произвольные натуральные числа, $x = (x_1, \dots, x_m)$, $u = (u_1, \dots, u_k)$, $g(x, u) = (g_1(x, u), \dots, g_m(x, u))^T$ – некоторая вещественная аналитическая вектор-функция, заданная на некотором открытом множестве. Переменные u_1, \dots, u_k будем называть параметрами. Пусть $x_{[0]}$ является решением системы уравнений $g(x, u) = 0$ при $u = u_{(0)}$ и пусть якобиан в этой точке является невырожденным, т.е.

$$\det \frac{\partial}{\partial x} g(x_{[0]}, u_{(0)}) \neq 0.$$

Будем говорить, что при выполнении этого предположения имеет место невырожденный случай. Не уменьшая общности (за счет линейной замены параметров), можно считать, что $u_{(0)} = 0$.

По теореме о неявной функции (Ганнинг, Росси, 1969) существует единственная вещественная аналитическая функция $x(u)$ такая, что $x(0) = x_{[0]}$ и $g(x(u), u) \equiv 0$ в некоторой окрестности нуля. Как известно, аналитическая функция $x(u)$ разлагается в сходя-

щийся ряд Тейлора

$$x(u) = \sum_j x_{[j]} u^j = x_{[0]} + \sum_{t=1}^{\infty} \sum_{|j|=t} x_{[j]} u^j,$$

где при $k > 1$ символ $j = (j_1, \dots, j_k)$ есть мультииндекс и $|j| = j_1 + \dots + j_k$. В общем случае найти функцию $x(u)$ как решение уравнения в явном виде не удается. Поэтому возможным путем решения задачи является построение разложения функции $x(u)$ в ряд Тейлора. Далее в пункте 1.3 выводится эффективный алгоритм вычисления коэффициентов разложения в ряд Тейлора.

Для любой (скалярной, векторной или матричной) функции $\phi(v)$, $v \in \mathbb{R}^s$ коэффициент разложения этой функции в ряд Тейлора в нуле при степени vv^j обозначим через $\phi_{[j]}$:

$$\phi_{[j]} = \phi_{[j_1, \dots, j_s]} = (\phi(v))_{[j]} = \frac{1}{j!} \frac{\partial^j \phi}{\partial v^j}(0) = \frac{1}{j_1! \dots j_s!} \frac{\partial^{j_1+\dots+j_s} \phi}{\partial v_1^{j_1} \dots \partial v_s^{j_s}}(0).$$

1.3.2. Метод Эйлера

При $m = 1, k = 1$ процесс нахождения коэффициентов $x_{[j]}$, носящий рекуррентный характер, был предложен еще Эйлером (Матвеев, 1995). Он заключается в следующем. Рассмотрим уравнение

$$x'(u) = \psi(x(u), u). \quad (10)$$

Это уравнение получается дифференцированием (полным образом) по u тождества $g(x(u), u) \equiv 0$, где $\psi(x, u) = -\frac{\partial g}{\partial u}(x, u)/\frac{\partial g}{\partial x}(x, u)$. Метод Эйлера — это метод последовательного дифференцирования уравнения (10) с учетом условия $x(0) = x_{[0]}$ и того, что коэффициент $x_{[k]} = x^{(k)}(0)/k!$. Полагая в (10) $u = 0$, $x(0) = x_{[0]}$, получаем $x'(0) = \psi(x_{[0]}, 0)$. Для нахождения $x''(0)$ продифференцируем сначала (10) (полным образом) по u :

$$x'' = \psi'_x x' + \psi'_u.$$

Полагая здесь $u = 0$, $x(0) = x_{[0]}$ и учитывая уже найденное выше значение $x'(0)$, получаем $x''(0) = \psi'_x(x_{[0]}, 0)x'(0) + \psi'_u(x_{[0]}, 0)$. Продолжая этот процесс последовательного дифференцирования, находим все значения $x^{(k)}(0)$. Отметим, что в этом алгоритме мы вычисляем смешанные производные функции $\psi(x, u)$ по переменным x, u .

1.3.3. Рекуррентный алгоритм

В этом пункте мы построим рекуррентный алгоритм вычисления коэффициентов разложения функции, неявно заданной уравнением $g(x, u) = 0$, который является более эффективным, чем приведенный выше.

Введем множества индексов

$$\prec\gamma\succ = \{j : j_r \leq \gamma_r, r = 1, \dots, k\}, \\ \prec\gamma\succ^- = \prec\gamma\succ \setminus \{\gamma\} = \{j : j_r \leq \gamma_r, r = 1, \dots, k, |j| < |\gamma|\}.$$

Рассмотрим сначала невырожденный случай. Итак, мы имеем тождество

$$g(x(u), u) \equiv 0. \quad (11)$$

Разложим функции в ряды Тейлора

$$g(x, u) = \sum_{i,j} g_{[i,j]} (x - x_{[0]})^i u^j = \sum_{i,j_1, \dots, j_k} g_{[i,j_1, \dots, j_k]} (x - x_0)^i u_1^{j_1} \cdots u_k^{j_k}, \\ x(u) = \sum_{\alpha} x_{[\alpha]} u^{\alpha} = x_{[0]} + \sum_{|\alpha| > 0} x_{[\alpha]} u^{\alpha}, \\ g(x(u), u) = \sum_j G_j u^j = \sum_{j_1, \dots, j_k} G_{j_1, \dots, j_k} u_1^{j_1} \cdots u_k^{j_k}.$$

В силу (11) для любого j выполняется равенство $G_j = 0$ и $J_{[0]} = g_{[1,0]}$, где $J_{[0]} = \frac{\partial}{\partial x} g(x_{[0]}, 0)$. Выпишем выражения, связывающие коэффициенты G_j с коэффициентами $g_{[i,j]}$ и $x_{[\alpha]}$. При $k = 1$ имеем

$$G_1 = g_{[1,0]} x_{[1]} + g_{[0,1]}, \\ G_2 = g_{[1,0]} x_{[2]} + g_{[2,0]} x_{[1]}^2 + g_{[1,1]} x_{[1]} + g_{[0,2]}, \\ G_3 = g_{[1,0]} x_{[3]} + g_{[3,0]} x_{[1]}^3 + g_{[2,1]} x_{[1]}^2 + 2g_{[2,0]} x_{[1]} x_{[2]} + g_{[1,2]} x_1 + g_{[1,1]} x_2 + g_{[0,3]}.$$

При $k = 2$ имеем

$$G_{1,0} = g_{[1,0,0]} x_{[1,0]} + g_{[0,1,0]}, \\ G_{0,1} = g_{[1,0,0]} x_{[0,1]} + g_{[0,0,1]}, \\ G_{2,0} = g_{[1,0,0]} x_{[2,0]} + g_{[1,1,0]} x_{[1,0]} + g_{[2,0,0]} x_{[1,0]}^2 + g_{[0,2,0]}, \\ G_{1,1} = g_{[1,0,0]} x_{[1,1]} + g_{[1,1,0]} x_{[0,1]} + g_{[1,0,1]} x_{[1,0]} + 2g_{[2,0,0]} x_{[0,1]} x_{[1,0]} + g_{[0,1,1]}, \\ G_{0,2} = g_{[1,0,0]} x_{[0,2]} + g_{[1,0,1]} x_{[0,1]} + g_{[2,0,0]} x_{[0,1]}^2 + g_{[0,0,2]}, \\ G_{3,0} = g_{[1,0,0]} x_{[3,0]} + 2g_{[2,0,0]} x_{[1,0]} x_{[2,0]} + g_{[3,0,0]} x_{[1,0]}^3 + g_{[2,1,0]} x_{[1,0]}^2 + \\ + g_{[1,1,0]} x_{[2,0]} + g_{[1,2,0]} x_{[1,0]} + g_{[0,3,0]}.$$

При произвольном k имеем

$$\begin{aligned} G_\beta &= \sum_{j \leq \beta} \sum_{i=0}^{\infty} g_{[i,j]} \left(\left(\sum_{|\alpha|>0} x_{[\alpha]} u^\alpha \right)^i \right)_{\beta-j} \\ &= \sum_{j \leq \beta} \sum_i g_{[i,j]} \left(\left(\sum_{\alpha \leq \beta-j, |\alpha|>0} x_{[\alpha]} u^\alpha \right)^i \right)_{\beta-j}. \end{aligned} \quad (12)$$

Из формулы (12) вытекает справедливость следующего результата.

Лемма 1.1. В невырожденном случае для любого β ($|\beta| > 0$) выполняется:

- 1) G_β зависит только от $x_{[\alpha]}$ таких, что $\alpha \leq \beta$,
- 2) $G_\beta - J_{[0]} x_{[\beta]}$ не зависит от $x_{[\beta]}$.

Отсюда вытекает рекуррентная формула для нахождения коэффициентов

$$x_{[\beta]} = -J_{[0]}^{-1} h_{[\beta]}, \text{ где } h_{[\beta]} = (g(x_{<I>}(u), u))_{[\beta]}, I = \prec\beta\succ^- . \quad (13)$$

Рекуррентная формула остается верной, если $I \supset \prec\beta\succ^-$, так как по лемме 1.1 $h_{[\beta]}$ не зависит от коэффициентов с индексами из $I \setminus \prec\beta\succ$. Заметим, что при вычислениях на компьютере разумно находить несколько коэффициентов на каждом шаге. В связи с этим на основе (13) получаем следующий алгоритм. Введем множества

$$S_t = \{j : |j| = j_1 + \dots + j_k = t\}, I_t = \cup_{l=0}^t S_l.$$

Алгоритм 1.1. На шаге t ($t = 1, 2, \dots$) вычисляем все коэффициенты с индексами из S_t по формуле

$$x_{[\beta]} = -J_{[0]}^{-1} h_{[\beta]}, \text{ где } h_{[\beta]} = (g(x_{<I_{t-1}>}(u), u))_{[\beta]} .$$

Рассмотрим теперь вырожденный случай некоторого специального типа. Пусть мы имеем уравнение

$$\bar{g}(x, u) = 0$$

и пусть функция $\bar{g}(x, u)$ имеет следующий вид: $\bar{g}(x, u) = u^l p(u) g(x, u)$, где $p(u) = \sum_{|j|=d} a_j u^j$ — однородный многочлен такой, что $a_{d,0,\dots,0} \neq 0$,

и уравнение $g(x, u) = 0$ задает невырожденный случай. Будем считать, что порядок d полинома $p(u)$ минимальный. При $k = 1$ имеем, что $p(u) \equiv \text{const}$ и $d = 0$. В вырожденном случае для вычисления коэффициентов справедлив следующий алгоритм.

Положим $c = (p, 0, \dots, 0)$ и введем множества

$$\overline{S}_t = \{j : j_1 + \sum_{r=2}^k 2j_r = t\}, \quad \overline{I}_t = \cup_{l=0}^t \overline{S}_l.$$

Алгоритм 1.2. На шаге t ($t = 1, 2, \dots$) вычисляем все коэффициенты с индексами из S_t^* по формуле

$$x_{[\beta]} = -\overline{J}_{[l+c]}^{-1} \overline{h}_{[\beta+l+c]}, \quad \text{где } \overline{h}_{[\gamma]} = \left(\overline{g}(x_{< I_{t-1}^* >} (u), u) \right)_{[\gamma]},$$

где $S_i^* = S_i$, $I_i^* = I_i$ при $d = 0$ и $S_i^* = \overline{S}_i$, $I_i^* = \overline{I}_i$ при $d > 0$.

Перейдем к обоснованию алгоритма 1.2. Мы имеем

$$\sum_j \overline{G}_j u^j = \overline{g}(x, u) = u^l p(u) g(x, u) = \sum_s G_s u^s.$$

Следовательно,

$$\overline{G}_{j+l+c} = \sum_{|\alpha|=p} a_\alpha G_{j+c-\alpha}.$$

Здесь и далее, если в выражении встречается мультииндекс β такой, что существует i_0 и $\beta_{i_0} < 0$, то это означает, что элемент с этим мультииндексом β в выражении отсутствует.

Пусть $d = 0$. В этом случае $\overline{G}_{j+l} = a_{d,0,\dots,0} G_j$, $\overline{J}_{[l]} = a_{d,0,\dots,0} J_{[0]}$. Отсюда получаем справедливость рекуррентной формулы при $d = 0$ и также при $k = 1$.

Пусть теперь $k = 2$ и $d > 0$. Ради простоты обозначений пусть $l = 0$ (случай $l \neq 0$ рассматривается аналогично). По условию мы имеем, что

$$\overline{G}_{s_1+d, s_2} = a_{d,0} G_{s_1, s_2} + a_{d-1,1} G_{s_1+1, s_2-1} + \dots + a_{0,d} G_{s_1+p, s_2-d}.$$

Введем множества индексов

$$\begin{aligned} \prec s_1, s_2 \succ_d &= \prec s_1, s_2 \succ \cup \prec s_1 + 1, s_2 - 1 \succ \cup \dots \cup \prec s_1 + d, s_2 - d \succ, \\ \prec s_1, s_2 \succ_d^- &= \prec s_1, s_2 \succ_d \setminus \{(s_1, s_2)\}. \end{aligned}$$

Из леммы 1.1 следует, что \bar{G}_{s_1, s_2} зависит от коэффициентов $x_{[\alpha]}$ с индексами из $\prec s_1, s_2 \succ_d$. Отсюда и в силу того, что коэффициенты $(g(x_{\prec k_1, k_2 \succ}(u), u))_{[\beta_1, \beta_2]} = 0$ при $(\beta_1, \beta_2) \in \prec k_1, k_2 \succ$, получаем формулу для вычисления коэффициента $x_{[s_1, s_2]}$

$$x_{[s_1, s_2]} = -\bar{J}_{[l+c]}^{-1} \bar{h}_{[s_1 + l_1 + p, s_2 + l_2]}, \quad \bar{h}_{[j_1, j_2]} = (\bar{g}(x_{\prec I \succ}(u), u))_{[j_1, j_2]}, \quad (14)$$

где $I = \prec s_1, s_2 \succ_d^-$ (или $I \supset \prec s_1, s_2 \succ_d^-$). Таким образом, на основе (14) получаем алгоритм 1.2.

При $k > 2$ рассуждения проводятся аналогично (со следующим определением для множества $\prec s \succ_d = \cup_{|j|=d} \prec s + c - j \succ$).

Обсудим более подробно последовательность вычисления коэффициентов. Для наглядности рассмотрим случай двух скалярных параметров u и v . Исследуем последовательность вычисления коэффициентов в невырожденном случае ($g(x, u, v) = 0$) и в вырожденном случае ($\bar{g}(x, u, v) = p(u, v)g(x, u, v)$). Для наглядности коэффициенты разложения функции $x = x(u, v) = \sum_{i,j} x_{[i,j]} u^i v^j$ будем изображать в виде квадратной таблицы, бесконечной вправо и вниз. На пересечении i строки и j столбца таблицы находится коэффициент $x_{[i,j]}$.

$x_{[0,0]}$	$x_{[0,1]}$	$x_{[0,2]}$	
$x_{[1,0]}$	$x_{[1,1]}$	$x_{[1,2]}$	

Пусть задано $x_{[0,0]}$, которое необходимо для начала рекуррентного вычисления коэффициентов. При нахождении коэффициента $x_{[k,l]}$ должны быть известны некоторые «предыдущие» коэффициенты $x_{[i,j]}$. Изобразим эту зависимость на некотором шаге. На схематическом рисунке используются следующие обозначения. Если в клетке таблицы стоит «?», то этот коэффициент вычисляется на данном шаге, если «*», то этот коэффициент должен быть известен (т.е. он обязательно должен присутствовать в правой части рекуррентной формулы), если клетка пустая, то значение этого коэффициента не влияет на нахождение коэффициента, определяемого на данном шаге.

В (15) в левой таблице представлена зависимость в невырожденном случае, в правой таблице — в вырожденном случае, когда порядок d полинома $p(u, v) = \sum_{i=0}^d a_{i,p-i} u^i v^{d-i}$ равен 2.

...
...	*	*	*		
...	*	*	*		
...	*	*	*		
...	*	*	?		

...
...	*	*	*	*	*	*
...	*	*	*	*	*	?
...	*	*	*	*	*	
...	*	*	*			

(15)

Если требуется вычислить коэффициентов как можно больше за меньшее число шагов и за меньшее время, то лучшей последовательностью вычисления будет следующая. На каждом шаге вычислять не один, а сразу несколько коэффициентов, при этом на k -м шаге должны быть известны все коэффициенты, которые находятся с 1-го по $(k - 1)$ -й шаги включительно. Изобразим лучшую последовательность вычисления. В клетке таблицы стоит число — номер шага, на котором должен быть найден коэффициент, отвечающий этой клетке.

	1	2	3	
1	2	3	4	
2	3	4	5	
3	4	5	6	
4	5	6	7	

	2	4	6	
1	3	5	7	
2	4	6	8	
3	5	7	9	
4	6	8	10	

(16)

В (16) в левой таблице — в невырожденном случае, в правой таблице — в вырожденном случае.

2. *D*-оптимальные планы для тригонометрической модели на произвольном отрезке

2.1. Постановка задачи

Рассмотрим тригонометрическую модель

$$\eta(t) = \beta_0 + \sum_{j=1}^k (\beta_{2j-1} \sin(jt) + \beta_{2j} \cos(jt)), \quad (17)$$

где $t \in \mathcal{X} = [b, c]$, $f(t) = (1, \sin(t), \dots, \cos(kt))^T$ — вектор регрессии, $\beta = (\beta_0, \dots, \beta_{2k})^T$ — вектор неизвестных параметров.

В настоящем разделе мы изучим *D*-оптимальные планы на произвольном отрезке. Так как вектор регрессии является функцией с периодом 2π , то, не уменьшая общности, можно считать, что $c - b \leq 2\pi$.

Задача построения оптимальных планов на произвольном отрезке имеет практический интерес, поскольку встречаются эксперименты, в которых невозможно проводить наблюдения на полном круге (Kitsos et al., 1988). Решение этой задачи для стандартного отрезка $[-\pi, \pi]$ является классическим (Карлин, Стадден, 1976). Для случая произвольных отрезков эта задача до сих пор не была решена. Ниже выводится полное решение этой задачи, которое получено в совместной работе (Dette, Melas, Pepelyshev, 2002).

Отметим, что *E*-оптимальные планы на $[-a, a]$ были исследованы в работе (Dette, Melas, 2001). В ней показано, что кратность минимального собственного числа информационной матрицы равна единице при малых a и может быть любой при больших a . На каждом интервале значений a , на котором кратность не меняется, *E*-оптимальный план может быть построен с помощью функционального подхода. В работе (Dette, Melas, 2003) получены некоторые результаты, касающиеся построения e_i -оптимальных планов для тригонометрической регрессии.

2.2. Предварительные результаты

В первой лемме мы покажем, что достаточно рассмотреть симметричные относительно нуля отрезки. На произвольном отрезке

оптимальный план получается сдвигом точек оптимального плана на симметричном отрезке той же длины на $(c - b)/2$.

Лемма 2.1. *Пусть $\xi_\omega = \{t_1 + \omega, \dots, t_n + \omega; w_1, \dots, w_n\}$ — произвольный план эксперимента. Тогда для любого ω выполняется*

$$\det M(\xi_\omega) = \det M(\xi_0).$$

Доказательство. Если $n < 2k + 1$, то в этом случае информационные матрицы являются вырожденными и их определители равны нулю. Пусть теперь $n = 2k + 1$. Из равенства

$$M(\xi_\omega) = \sum_{j=1}^{2k+1} f(t_j + \omega) f^T(t_j + \omega) w_j = F(\xi_\omega) \text{diag}(w_1, \dots, w_{2k+1}) F^T(\xi_\omega)$$

следует, что

$$\det M(\xi_\omega) = (\det F(\xi_\omega))^2 \prod_{j=1}^{2k+1} w_j,$$

где матрица $F(\xi_\omega)$ имеет вид

$$F(\xi_\omega) = \left(f(t_1 + \omega) : \dots : f(t_{2k+1} + \omega) \right).$$

Легко видеть, что

$$f(t + \omega) = Pf(t),$$

где P — блочно-диагональная матрица:

$$P = \begin{pmatrix} 1 & & & \\ & Q(\omega) & & \\ & & \ddots & \\ & & & Q(k\omega) \end{pmatrix}, \quad Q(\gamma) = \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ -\sin(\gamma) & \cos(\gamma) \end{pmatrix}.$$

Очевидно, что выполняется следующее равенство:

$$\det F(\xi_\omega) = \det P \det F(\xi_0) = \det F(\xi_0),$$

которое доказывает лемму при $n = 2k + 1$. При $n > 2k + 1$ утверждение леммы следует из формулы Бине—Коши и рассуждений, приведенных для случая $n = 2k + 1$. ■

Из леммы следует, что если ξ_0^* — D -оптимальный план на отрезке $[-a, a]$, то план ξ_ω^* будет оптимальным планом на отрезке $[-a + \omega, a + \omega]$. Таким образом, достаточно рассмотреть случай симметричных отрезков $[b, c] = [-a, a]$ ($0 < a \leq \pi$).

Пусть теперь ξ_a^* обозначает оптимальный план на $[-a, a]$. Заметим, что в общем случае D -оптимальный план не является единственным (Федоров, 1971а). Однако известно (теорема 1.3), что оптимальная информационная матрица является единственной и невырожденной, т.е. информационные матрицы всех оптимальных планов совпадают. Кроме того, по теореме эквивалентности план ξ_a^* удовлетворяет неравенству $d(t, \xi_a^*) \leq 2k + 1$ для всех $t \in [-a, a]$ с равенством в точках плана, где $d(t, \xi) = f^T(t)M^{-1}(\xi)f(t)$. Введем множество планов $\Xi_a^{(1)}$, состоящее из всех планов вида

$$\xi = \xi_a = \begin{pmatrix} -t_k & \dots & -t_1 & t_0 & t_1 & \dots & t_k \\ \frac{1}{2k+1} & \dots & \frac{1}{2k+1} & \frac{1}{2k+1} & \frac{1}{2k+1} & \dots & \frac{1}{2k+1} \end{pmatrix}, \quad (18)$$

где $0 = t_0 < t_1 < \dots < t_k = a$, и множество

$$\Xi_a^{(2)} = \{\xi : \text{supp } \xi \subset [-a, a], d(t, \xi) \equiv 2k + 1\}.$$

Справедлив следующий результат.

Лемма 2.2. *Либо D -оптимальный план ξ_a^* на $[-a, a]$ может быть симметричным относительно нуля планом, сосредоточенным в $2k + 1$ точках, включая концы интервала планирования с равными весами, либо функция $d(t, \xi_a^*)$ для этого плана тождественно равна $2k + 1$, т.е.*

$$\xi_a^* \in \Xi_a^{(1)} \cup \Xi_a^{(2)}.$$

Доказательство. По теореме эквивалентности любой план $\xi \in \Xi_a^{(2)}$ является D -оптимальным планом. Пусть $\xi = \{u_1, \dots, u_n; w_1, \dots, w_n\}$ — некоторый план, являющийся D -оптимальным планом на $[-a, a]$ ($-a \leq u_1 < \dots < u_n \leq a$). Если $\xi \notin \Xi_a^{(2)}$, то $d(t, \xi) \neq 2k + 1$ и по теореме эквивалентности мы имеем

$$\begin{aligned} d(t, \xi) - 2k - 1 &\leq 0, \quad \forall u \in [-a, a], \\ d(u_i, \xi) - 2k - 1 &= 0, \quad i = 1, \dots, n, \\ \frac{d}{dt}d(t, \xi)|_{t=u_i} &= 0, \quad i = 2, \dots, n - 1. \end{aligned}$$

То есть функция $d(t, \xi) - 2k - 1$ имеет не менее $2n - 2$ корней с учетом их кратности. Пусть $\tilde{\xi}$ обозначает отражение плана ξ

относительно нуля, $\tilde{\xi} = \{-u_1, \dots, -u_n; w_1, \dots, w_n\}$. Легко видеть, что $\det M(\tilde{\xi}) = \det M(\xi)$ и, следовательно, план $\tilde{\xi}$ является D -оптимальным. В силу выпуклости D -критерия симметричный план $\xi^* = (\xi + \tilde{\xi})/2$ является также D -оптимальным планом на интервале $[-a, a]$. Заметим, что существует перестановочная матрица P такая, что

$$PM(\xi)P^T = \begin{pmatrix} M_1(\xi) & M_2(\xi) \\ M_2^T(\xi) & M_3(\xi) \end{pmatrix},$$

где

$$\begin{aligned} M_1(\xi) &= \int_{-a}^a f_c(t) f_c^T(t) d\xi(t), \\ M_2(\xi) &= \int_{-a}^a f_c(t) f_s^T(t) d\xi(t), \\ M_3(\xi) &= \int_{-a}^a f_s(t) f_s^T(t) d\xi(t), \\ f_c(t) &= (1, \cos(t), \dots, \cos(kt))^T, \\ f_s(t) &= (\sin(t), \dots, \sin(kt))^T. \end{aligned} \tag{19}$$

В силу единственности информационной матрицы D -оптимального плана мы получаем, что

$$M_2(\xi) = M_2(\tilde{\xi}) = M_2(\xi^*) = 0.$$

Следовательно, мы имеем

$$g(t) = d(t, \xi) = f_c^T(t) M_1^{-1}(\xi) f_c(t) + f_s^T(t) M_3^{-1}(\xi) f_s(t) - 2k - 1 = \sum_{i=0}^{2k} \gamma_i \cos(it),$$

где $\gamma_0, \dots, \gamma_{2k}$ — некоторые константы. Так как $\xi \notin \Xi_a^{(2)}$, то функция $g(t)$ не тождественно равна константе. Заметим, что система функций $\{1, \cos(t), \dots, \cos(2kt)\}$ является чебышевской на $[0, a]$ и на $[-a, 0]$ и, следовательно, функция $g(t)$ может иметь не более $2k$ корней на $[0, a]$ и $2k$ корней на $[-a, 0]$, или не более $4k$ корней на $[-a, a]$ (учитывая кратности). Таким образом, любой D -оптимальный план сосредоточен ровно в $2k+1$ точках, включая концы отрезка $[-a, a]$. D -оптимальный план с числом точек, равным числу неизвестных параметров, имеет равные веса. Если D -оптимальный план $\xi \notin \Xi_a^{(1)} \cup \Xi_a^{(2)}$, то $\xi \neq \tilde{\xi}$ и, следовательно, D -оптимальный план $\xi^* = (\xi + \tilde{\xi})/2$ сосредоточен более чем в $2k+1$ точках, что невозможно. Таким образом, $\xi \in \Xi_a^{(1)}$. ■

2.3. Аналитичность точек D -оптимального плана

Лемма 2.2 показывает, что достаточно рассматривать планы вида (18). В следующей лемме мы найдем явное представление для определителя информационной матрицы плана такого вида.

Лемма 2.3. *Пусть ξ – произвольный симметричный $(2k+1)$ -точечный план с равными весами (т.е. ξ имеет вид (18)) и пусть $x_i = \cos(t_i)$, $i = 0, 1, \dots, k$, тогда*

$$\det M(\xi) = \frac{2^{2k^2}}{(2k+1)^{2k+1}} \prod_{l=1}^k (1 - x_l^2) (1 - x_l)^2 \prod_{1 \leq i < j \leq k} (x_i - x_j)^4.$$

Доказательство. Для любого плана ξ вида (18) мы имеем

$$\det M(\xi) = \det M_1(\xi) \det M_3(\xi),$$

где матрицы $M_1(\xi), M_3(\xi)$ определены в (19). Введем план

$$\eta = \eta_\xi = \begin{pmatrix} x_0 & x_1 & \dots & x_k \\ \frac{1}{2k+1} & \frac{2}{2k+1} & \dots & \frac{2}{2k+1} \end{pmatrix},$$

где $x_0 = \cos(t_0) = \cos(0) = 1$. Тогда мы получаем, что

$$M_1(\xi) = \left(\int_{-1}^1 T_i(x) T_j(x) d\eta_\xi(x) \right)_{i,j=0}^k,$$

$$M_2(\xi) = \left(\int_{-1}^1 (1 - x^2) U_i(x) U_j(x) d\eta_\xi(x) \right)_{i,j=0}^{k-1},$$

где $T_i(x), U_i(x)$ – многочлены Чебышева первого и второго рода. Так как $T_i(x)$ есть многочлен степени i со старшим коэффициентом равным 2^{i-1} , то

$$M_1(\xi) = L \int_{-1}^1 f_p(x) f_p^T(x) d\eta_\xi(x) L^T,$$

где $f_p(x) = (1, x, \dots, x^k)^T$, элементы нижнетреугольной матрицы L определяются равенством $(T_0(x), \dots, T_k(x))^T = L f_p(x)$, и, учитывая формулу для определителя Вандермонда, получаем

$$\det M_1(\xi) = 2^{k(k-1)} \frac{2^k}{(2k+1)^{k+1}} \prod_{l=1}^k (1 - x_l)^2 \prod_{1 \leq i < j \leq k} (x_i - x_j)^2.$$

Аналогично рассуждая, получаем

$$\det M_3(\xi) = 2^{k(k-1)} \frac{2^k}{(2k+1)^k} \prod_{l=1}^k (1-x_l^2) \prod_{1 \leq i < j \leq k} (x_i - x_j)^2.$$

Комбинируя последние две формулы, получаем утверждение леммы. ■

Изучим функцию

$$\phi(x) = \phi(x, a) = \prod_{l=1}^k (1-x_l^2)(1-x_l)^2 \prod_{1 \leq i < j \leq k} (x_i - x_j)^4$$

на множестве $\cos(a) = x_k < x_{k-1} \dots < x_1 < x_0 = 1$ как функцию от $x = (x_1, \dots, x_{k-1})$. Заметим, что функция $\ln \phi(x)$ является вогнутой и, следовательно, функция $\phi(x)$ имеет единственный максимум, который мы обозначим через $x^* = x^*(a)$. Функция $x^*(a)$ может быть получена как единственное решение уравнения

$$\frac{\partial}{\partial x} \phi(x, a) = 0.$$

Кроме того, матрица Якоби

$$J(a) = \left(\frac{\partial^2}{\partial x_i \partial x_j} \phi(x, a) \right)_{i,j=1}^{k-1} \Big|_{x=x^*(a)}$$

положительно определена, и, следовательно, функция $x^*(a)$ является вещественной аналитической функцией.

Изучим поведение точек D -оптимального плана при $a \rightarrow 0$. С этой целью введем вектор $\tau = (\tau_1, \dots, \tau_{k-1})^T$, связанный с точками плана следующим образом: $t_i = a\tau_i$, ($0 < \tau_1 < \dots < \tau_{k-1} < 1$). Отметим, что любому плану $\xi \in \Xi_a^{(1)}$ соответствует единственный вектор τ и вектор x . Следовательно, мы можем определить функцию $\phi(\tau, a) = \phi(x_\tau, a)$ ($x_\tau = (\cos(a\tau_1), \dots, \cos(a\tau_{k-1}))^T$) и функцию

$$\tau^* = \tau^*(a) = \left(\frac{\arccos x_1^*(a)}{a}, \dots, \frac{\arccos x_{k-1}^*(a)}{a} \right)^T, \quad (20)$$

которая также будет вещественной аналитической функцией.

Лемма 2.4. Функция $\tau^*(a)$ имеет предел при $a \rightarrow 0$

$$\tau^*(0) = \lim_{a \rightarrow 0} \tau^*(a) = (\tau_1^*(0), \dots, \tau_{k-1}^*(0))^T,$$

где $\tau_1^*(0) < \dots < \bar{\tau}_{k-1}^*(0)$ есть положительные корни многочлена

$$P_{k-1}^{(1,1/2)}(2x^2 - 1) = \frac{1}{2x} P_{2k-1}^{(1,1)}(x),$$

где $P_i^{(\alpha,\beta)}(x)$ – многочлен Якоби степени i относительно меры $(1-x)^\alpha(1+x)^\beta dx$.

Доказательство. Используя разложения $\sin(t) = t + o(t)$, $\cos(t) = 1 - t^2/2 + o(t^2)$, мы получаем

$$\phi(\tau, a) = \phi(x_\tau, a) = \frac{a^{2k(2k+1)}}{2^{2k^2}} \prod_{l=1}^k \tau_l^6 \prod_{1 \leq i < j \leq k} (\tau_i^2 - \tau_j^2)^4 (1 + o(a)),$$

$\tau_k = 1$. Следовательно, $\tau^*(0)$ существует и может быть получен максимированием функции

$$\bar{\phi}(\tau, a) = \prod_{l=1}^k \tau_l^3 (1 - \tau_l^2)^2 \prod_{1 \leq i < j \leq k-1} (\tau_i^2 - \tau_j^2)^2.$$

Заметим, что эта функция обладает одним экстремумом и, следовательно, вектор $\tau^*(0)$, заданный как точка максимума, определен корректно. Вычисляя частные производные от логарифма функции $\bar{\phi}(\tau)$, получаем систему

$$\frac{3}{\tau_i} + \frac{4\tau_i}{\tau_i^2 - 1} + \sum_{j=1, \dots, k-1, j \neq i} \frac{4\tau_i}{\tau_i^2 - \tau_j^2} = 0, \quad i = 1, \dots, k-1.$$

Делая подстановку $y_i = \tau_i^2$, получаем

$$\frac{3}{y_i} + \frac{4}{y_i - 1} + \sum_{j=1, \dots, k-1, j \neq i} \frac{4}{y_i - y_j} = 0, \quad i = 1, \dots, k-1.$$

Рассмотрим многочлен $\psi(y) = \prod_{i=1}^{k-1} (y - y_i) = y^{k-1} + \sum_{i=0}^{k-2} \psi_i y^i$. Заметим, что

$$\frac{1}{2} \frac{\psi''(y_i)}{\psi'(y_i)} = \sum_{j=1, \dots, k-1, j \neq i} \frac{1}{y_i - y_j}. \tag{21}$$

Тогда

$$\frac{3}{y_i} + \frac{4}{y_i - 1} + 2 \frac{\psi''(y_i)}{\psi'(y_i)} = 0, i = 1, \dots, k-1.$$

Приводя к общему знаменателю, мы получаем

$$2y(y-1)\psi''(y) + (7y-3)\psi'(y) = 0 \text{ при } y = y_i, i = 1, \dots, k-1.$$

Так как левая часть равенства есть многочлен степени не более k и обращается в нуль в точках $y_i, i = 1, \dots, k-1$, то

$$2y(y-1)\psi''(y) + (7y-3)\psi'(y) = (k-1)(2(k-1)+5)\psi(y). \quad (22)$$

Собирая коэффициенты при одинаковых степенях y и приравнивая их к нулю, мы получаем рекуррентную формулу для вычисления коэффициентов ψ_i ($i = 0, \dots, k-2$)

$$\psi_{k-2} = -\frac{(2k-1)(k-1)}{4k-1}, \quad \psi_{\nu-1} = \frac{\nu(2\nu+1)}{(2\nu+1+2k)(\nu-k)}\psi_\nu,$$

$\nu = 1, \dots, k-2$. Известно (Сеге, 1962), что дифференциальное уравнение (22) имеет единственное решение в виде полинома Якоби $P_{k-1}^{(1/2,1)}(1-2y)$ и, используя равенство $P_n^{(1/2,1)}(-y) = (-1)^n P_n^{(1,1/2)}(y)$, получаем требуемое. ■

В табл. 1 представлен многочлен $\psi(y)$ и вектор $\tau^*(0)$ для низших значений $k = 2, 3, 4, 5$.

Теорема 2.1. 1) Если $a \geq a^* = \pi(1 - \frac{1}{2k+1})$, то план ζ^* , средоточенный с равными весами в $2k+1$ точках

$$t_i^* = 2\pi \frac{i-1-k}{2k+1}, \quad i = 1, \dots, 2k+1$$

является D -оптимальным планом на $[-a, a]$;

2) Если $a < a^*$, то D -оптимальный план ξ_a^* является единственным и имеет вид

$$\xi_a^* = \begin{pmatrix} -a & \dots & -a\tau_1^*(a) & 0 & a\tau_1^*(a) & \dots & a\tau_{k-1}^*(a) & a \\ \frac{1}{2k+1} & \dots & \frac{1}{2k+1} & \frac{1}{2k+1} & \frac{1}{2k+1} & \dots & \frac{1}{2k+1} & \frac{1}{2k+1} \end{pmatrix}, \quad (23)$$

где $\tau^*(a)$ — аналитическая вектор-функция, определенная в (20) и лемме 2.4.

Таблица 1. Многочлен $\psi(y)$ и координаты вектора $\tau^*(0)$

k	$\psi(y)$ и $\tau_j(0)$
2	$\psi(y) = y - 3/7$ $\tau_1^*(0) = \sqrt{3/7} = 0.6546$
3	$\psi(y) = y^2 - 10/11y + 5/33$ $\tau_1^*(0) = 0.4688, \tau_2^*(0) = 0.8302$
4	$\psi(y) = y^3 - 7/5y^2 + 7/13y - 7/143$ $\tau_1^*(0) = 0.3631, \tau_2^*(0) = 0.6772, \tau_3^*(0) = 0.8998$
5	$\psi(y) = y^4 - 36/19y^3 + 378/323y^2 - 84/323y + 63/4199$ $\tau_1^*(0) = 0.2958, \tau_2^*(0) = 0.5652, \tau_3^*(0) = 0.7845, \tau_4^*(0) = 0.9340$

Доказательство. Предположим, что план $\xi^* \in \Xi_a^{(2)}$ является D -оптимальным на $[-a, a]$. По определению множества $\Xi_a^{(2)}$ функция $d(t, \xi^*)$ тождественно равна $2k + 1$ на $[-a, a]$. Следовательно, $d(t, \xi^*) \equiv 0$ на полном круге $[-\pi, \pi]$. Поэтому план ξ^* является также D -оптимальным на $[-\pi, \pi]$. По единственности информационной матрицы оптимального плана мы имеем

$$M(\xi^*) = \text{diag}(1, 1/2, \dots, 1/2), \det M(\xi^*) = 2^{-2k}.$$

С другой стороны, мы имеем, что

$$\lim_{a \rightarrow 0} \max_{\xi} \det M(\xi) = 0.$$

Таким образом, для достаточно малых a D -оптимальный план не может быть элементом множества $\Xi_a^{(2)}$. По лемме 2.2 следует, что D -оптимальный план должен принадлежать множеству $\Xi_a^{(1)}$. Из рассуждений, приведенных после леммы 2.3, следует, что для достаточно малых a D -оптимальный план $\xi^*(a)$ является единственным и имеет вид (23). Пусть

$$\hat{a} = \sup\{a \in (0, \pi] \mid \det M(\xi_a^*) < 2^{-2k}\}.$$

Покажем, что $\hat{a} = a^*$. Известно (Федоров, 1971а), что равномерное распределение ζ^* является D -оптимальным планом для тригоно-

метрической модели на $[-\pi, \pi]$ и также на любом отрезке $[-a, a]$ при $a \in [a^*, \pi]$. Заметим, что $\xi_{a^*}^* = \zeta$. Отсюда получаем, что

$$\xi_{a^*}^* \in \Xi_{a^*}^{(1)} \cap \Xi_{a^*}^{(2)}.$$

Из этого включения следует, что $\hat{a} \leq a^*$. Напомним, определитель информационной матрицы является строго вогнутой функцией на множестве $\Xi_a^{(1)}$. Для плана $\xi_{a^*}^*$

$$\det M(\xi_{a^*}^*) = 2^{-2k}$$

и для любого плана $\xi \in \Xi_a^{(1)}$

$$\det M(\xi) < 2^{-2k}.$$

Это следует из того, что в противном случае выпуклая комбинация планов $\xi_{a^*}^*$ и ξ дала бы информационную матрицу, определитель которой был бы больше, чем 2^{-2k} , что невозможно. Отсюда следует, что $\hat{a} = a^*$. Таким образом, если $a < a^*$, то D -оптимальный план принадлежит множеству $\Xi_a^{(1)}$, и из определения $\tau^*(a)$ и леммы 2.4 следует, что D -оптимальный план имеет вид (23). ■

Отметим, что теорема 2.1 дает полное решение задачи построения D -оптимальных планов для тригонометрической регрессии. При $a > a^*$ равномерное распределение в точках $t_i^* = 2\pi \frac{i-1-k}{2k+1}$, $i = 1, \dots, 2k+1$ является D -оптимальным планом, но такой оптимальный план не является единственным. При $a \leq a^*$ D -оптимальный план является единственным и определен в (23), где вектор $\tau^*(a) = (\tau_1^*(a), \dots, \tau_{k-1}^*(a))^T$ может быть получен посредством разложения в ряд Тейлора в точке $a = 0$

$$\tau^*(a) = \sum_{i=0}^{\infty} \tau_{[i]} a^i$$

с начальным вектором $\tau_{[0]} = \tau^*(0)$, найденным в лемме 2.4, с помощью рекуррентной процедуры, разработанной в пункте 1.3.

Однако прямое вычисление коэффициентов разложения в ряд Тейлора приводит к значительным затратам компьютерного времени. Поэтому в следующем пункте мы введем иные уравнения для вычисления коэффициентов разложения точек оптимального плана на основе алгебраического подхода, предложенного в работе (Мелас, 1999а).

2.4. Функционально-алгебраический подход

В пункте 2.3 мы получили, что $x^*(a) = (x_1^*(a), \dots, x_{k-1}^*(a))^T$ ($x_i^*(a) = \cos t_i^*(a)$) является точкой максимума функции

$$\phi(x, a) = \prod_{j=1}^k \left((1 - x_j^2)(1 - x_j)^2 \right) \prod_{1 \leq i < j \leq k} \left(x_j - x_i \right)^4,$$

где $\cos(a) = x_k < \dots < x_1 < 1$. Обозначим $\alpha = 1 - \cos(a)$. Целью этого пункта является построение разложения вектор-функции $x^*(a)|_{a=\arccos(1-\alpha)}$ по степеням α в нуле ($\alpha \in [0, 2]$).

Запишем для точки максимума необходимые условия экстремума. Вычисляя частные производные первого порядка функции $\ln \phi(x, \alpha)$, получим

$$\frac{1}{1 + x_i^*} - \frac{3}{1 - x_i^*} + \frac{4}{x_i^* - 1 + \alpha} + \sum_{j=1, \dots, k-1, j \neq i} \frac{4}{x_i^* - x_j^*} = 0,$$

$i = 1, \dots, k-1$. Введем многочлен $\psi(z) = \prod_{i=1}^{k-1} (z - x_i^*) = z^{k-1} + \sum_{i=0}^{k-2} \psi_i^* z^i$. В силу (21) после приведения к общему знаменателю, имеем

$$(1 - z^2)(z - 1 + \alpha)\psi''(z) + (-4z^2 + (1 - 2\alpha)z + 3 - \alpha)\psi'(z) = 0$$

при $z = x_i^*$, $i = 1, \dots, k-1$. Так как левая часть равенства есть многочлен степени не более $k+1$ и обращается в нуль в точках x_i^* , $i = 1, \dots, k-1$, то

$$P(z) := (1 - z^2)(z - 1 + \alpha)\psi''(z) + (-4z^2 + (1 - 2\alpha)z + 3 - \alpha)\psi'(z) - (\lambda_0 z + \lambda_1)\psi(z) \quad (24)$$

тождественно равен нулю, где $\lambda_0 = -(m-1)(m+2)$, λ_1 — неизвестная константа. Выражение для λ_0 мы нашли после приравнивания старшего коэффициента многочлена $P(z)$ к нулю. Запишем $P(z)$ в иной форме.

$$P(z) = (z^k, \dots, z, 1)A(\lambda_1, \alpha)\psi, \quad \psi = (\psi_{k-1}, \dots, \psi_0)^T,$$

где $A = A(\lambda_1, \alpha) = (k+1) \times k$ матрица, $\psi_{k-1} = 1$. Заметим, что первая строка матрицы A состоит из нулей. Пусть B — матрица, полученная из A удалением первой строки. Таким образом,

$\psi_1^*, \dots, \psi_{k-1}^*, \lambda_1^*$ удовлетворяют уравнению $B\psi = 0$ относительно переменных $\psi_1, \dots, \psi_{k-1}, \lambda_1$. Поскольку $\psi^* \neq 0$ и $B(\lambda_1^*, \alpha)\psi^* = 0$, то матрица $B(\lambda_1^*, \alpha)$ — вырожденная и имеет вид ($B = (b_{i,j})$)

$$b_{i,j} = \begin{cases} -(k-j)(k+1-j) + 4(k-j) - \lambda_0, & j-i=1, \\ (1-\alpha)(k-j)(k+1-j) + (1-2\alpha)(k-j) - \lambda_1, & j-i=0, \\ (k-j)(k+1-j) + (3-\alpha)(k-j), & j-i=-1, \\ (\alpha-1)(k-j)(k+1-j), & j-i=-2, \\ 0, & \text{иначе.} \end{cases}$$

Отметим, что матрицу B можно представить следующим образом: $B = B(\lambda_1, \alpha) = \tilde{B}(\alpha) - \lambda_1 I$, где I — единичная матрица. Таким образом, λ_1^* является собственным числом матрицы $\tilde{B}(\alpha)$, т.е. λ_1^* — корень уравнения $\det B(\lambda, \alpha) = 0$. При известном λ_1^* значения $\psi_1^*, \dots, \psi_{k-1}^*$ могут быть вычислены рекуррентным образом по формулам

$$\psi_{k-1}^* = 1, \quad \psi_\nu^* = - \sum_{j=\nu+1}^{k-1} b_{k-1-\nu, k-j} \psi_j / b_{k-1-\nu, k-\nu}, \quad (25)$$

$$\nu = k-2, \dots, 1, 0.$$

Теперь, когда мы ввели ряд алгебраических структур, разбив исходную задачу на последовательность простых вложенных задач, применим функциональный подход к их решению.

Рассмотрим функцию

$$g(\lambda, \alpha) = \det(\tilde{B}(\alpha) - \lambda I).$$

Искомая функция $\lambda^* = \lambda^*(\alpha)$ задана решением уравнения

$$g(\lambda, \alpha) = 0.$$

Так как λ является простым собственным числом матрицы $\tilde{B}(\alpha)$ (рекуррентная формула (25) показывает, что собственный вектор является единственным), то выполняется неравенство

$$\left. \frac{\partial}{\partial \lambda} g(\lambda, \alpha) \right|_{\lambda=\lambda^*(\alpha)} \neq 0.$$

По теореме о неявной функции мы получаем, что функция $\lambda^*(\alpha)$ является вещественной аналитической. Следовательно, она может

быть разложена в ряд Тейлора. Найдем предел функции $\lambda^*(\alpha)$ при $\alpha \rightarrow 0$ с помощью предельного перехода в (24). Поскольку точки D -оптимального плана стремятся к нулю, то $x_i^* \rightarrow 1$ и полином $\psi(z) \rightarrow (z - 1)^{k-1}$. Отсюда получаем, что

$$\lim_{\alpha \rightarrow 0} \lambda^*(\alpha) = 1 - k^2.$$

Первые несколько коэффициентов разложения функции

$$\lambda^*(\alpha) = \sum_{i=0}^{\infty} \lambda_{[i]} \alpha^i, \quad \lambda_{[0]} = 1 - k^2,$$

представлены в табл. 2. Для удобства перейдем к шкалированным коэффициентам $\bar{\lambda}_{[i]} = 2^i \lambda_{[i]}$ (так как $\alpha \in [0, 2]$). Разложения функций

$$\psi_j^*(\alpha) = \sum_{i=0}^{\infty} \bar{\psi}_{j[i]} \left(\frac{\alpha}{2}\right)^i, \quad \bar{\psi}_{j[0]} = \frac{(m-1)!}{j!(m-1-j)!}, \quad j = 0, \dots, k-2,$$

могут быть вычислены или через решение уравнения $B\psi = 0$, или по формулам (25). Первые несколько коэффициентов разложения этих функций приведены в табл. 2.

Таким образом, с помощью рядов Тейлора мы нашли коэффициенты многочлена $\psi(z)$, корнями которого являются искомые функции $x_i^*(\alpha)$, $i = 1, \dots, k-1$, которые в свою очередь также могут быть разложены в ряды. А именно для вычисления отрезка ряда $x_{i<s>}^*(\alpha)$ сначала найдем отрезки рядов $\psi_{j<s>}^*(\alpha)$ и построим многочлен $\psi_{<s>}(z) = z^{k-1} + \sum_{j=0}^{k-2} \psi_{j<s>}^*(\alpha)z^j$. Введем полином $h(u) = h(u, \alpha) = \alpha^{1-k} \psi_{<s>}(1 - \alpha u)$ (поскольку $x_i^*(\alpha) \rightarrow 1$). По лемме 2.4 мы имеем, что $h(u, 0) = \text{const} P_{k-1}^{(1, 1/2)}(2u - 1)$. Для нахождения корней $u_i^*(\alpha)$ многочлена $h(u, \alpha)$ мы можем применить метод из пункта 1.3. Таким образом, $x_{i<s>}^*(\alpha) = 1 - \alpha u_{i<s-1>}^*(\alpha)$. Окончательно находя разложения точек D -оптимального плана, вычисляем $t_{i<s>}^*(a) = a \tau_{i<s-1>}^*(a)$, раскладывая в ряд функцию $\text{arccos } x_{i<s>}^*(1 - \cos a)$. Шкалированные коэффициенты $\bar{\tau}_{i[j]}^* = \pi^j \tau_{i[j]}^*$ функций $\tau_j^*(a)$, $j = 1, \dots, k-1$ представлены в табл. 3 (нечетные коэффициенты равны нулю).

На рис. 1 изображена зависимость точек D -оптимального плана для $a \in (0, \pi]$ при некоторых значениях k .

Таблица 2. Шкалированные коэффициенты разложения
функций $\lambda^*(\alpha)$, $\psi_j^*(\alpha)$, $j = 0, \dots, k - 2$

	i	0	1	2	3	4	5
$k = 2$	$\psi_{0[i]}$	-1	.85714	-.06997	-.02856	-.00886	-.00133
	$\bar{\lambda}_{[i]}$	-3	-.57143	-.27988	-.11424	-.03544	-.00533
$k = 3$	$\psi_{0[i]}$	1	-1.81818	.67618	-.02666	-.00204	.00190
	$\bar{\psi}_{1[i]}$	-2	1.81818	-.07012	-.03345	-.01963	-.01248
	$\bar{\lambda}_{[i]}$	-8	-1.09091	-.42074	-.20068	-.11779	-.07489
$k = 4$	$\psi_{0[i]}$	-1	2.80000	-2.22277	.48317	-.00808	.00222
	$\bar{\psi}_{1[i]}$	3	-5.60000	2.29169	-.05781	-.01309	-.00261
	$\bar{\psi}_{2[i]}$	-3	2.80000	-.06892	-.03375	-.02060	-.01400
	$\bar{\lambda}_{[i]}$	-15	-1.60000	-.55138	-.27001	-.16480	-.11199
$k = 5$	$\psi_{0[i]}$	1	-3.78947	4.74904	-2.23711	.32001	-.00380
	$\bar{\psi}_{1[i]}$	-4	11.36842	-9.56600	2.36016	-.03433	-.00019
	$\bar{\psi}_{2[i]}$	6	-11.36842	4.88488	-.08948	-.02490	-.00831
	$\bar{\psi}_{3[i]}$	-4	3.78947	-.06792	-.03357	-.02072	-.01430
	$\bar{\lambda}_{[i]}$	-24	-2.10526	-.67923	-.33566	-.20720	-.14296

Таблица 3. Шкалированные коэффициенты разложения
функций $\tau_j^*(a)$

	i	0	2	4	6	8	10
$k = 2$	$\bar{\tau}_{1[i]}$.65465	-.21977	-.07747	.04852	.06118	-.02116
	$\bar{\tau}_{2[i]}$.46885	-.19145	-.00875	.02584	-.00184	-.00283
$k = 3$	$\bar{\tau}_{1[i]}$.83022	-.13502	-.10286	-.05465	-.00161	.03946
	$\bar{\tau}_{2[i]}$.36312	-.15556	.00820	.01117	-.00368	-.00011
	$\bar{\tau}_{3[i]}$.67719	-.18093	-.07349	.00094	.02393	.01100
$k = 4$	$\bar{\tau}_{1[i]}$.89976	-.08456	-.07603	-.06025	-.03806	-.01256
	$\bar{\tau}_{2[i]}$.29576	-.12851	.01204	.00501	-.00238	.00036
	$\bar{\tau}_{3[i]}$.56524	-.18316	-.03971	.01585	.01178	-.00245
	$\bar{\tau}_{4[i]}$.78448	-.14366	-.08805	-.03360	.00483	.01980
$k = 5$	$\bar{\tau}_{1[i]}$.93400	-.05677	-.05431	-.04874	-.03965	-.02762

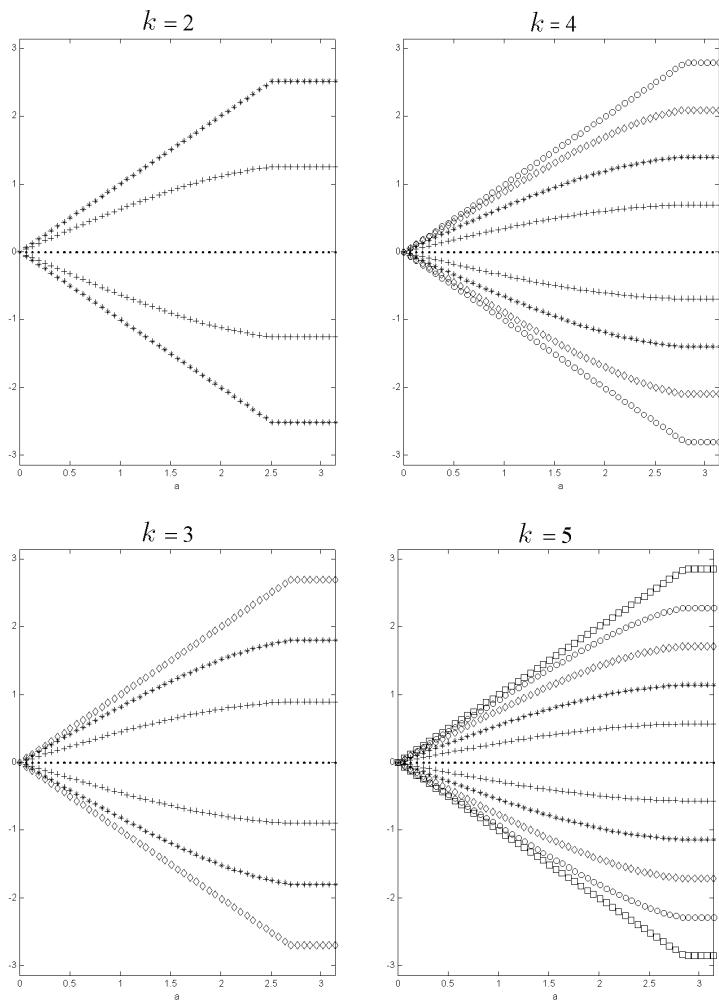


Рис. 1. Точки и веса D -оптимального плана.

2.5. Сравнение D - и E -оптимальных планов

В теореме 2.1 мы установили, что D -оптимальный план имеет вид

$$\xi_D^*(a) = \left(-a, -a\tau_{k-1D}^*(a), \dots, -a\tau_{1D}^*(a), 0, a\tau_{1D}^*(a), \dots, a\tau_{k-1D}^*(a), a \right) \begin{matrix} \\ \frac{1}{2k+1}, \dots, \frac{1}{2k+1}, \frac{1}{2k+1}, \dots, \frac{1}{2k+1} \end{matrix}.$$

В работе (Dette, Melas, 2001) получено, что E -оптимальный план имеет вид

$$\xi_E^*(a) = \left(-a, -a\tau_{k-1E}^*(a), \dots, -a\tau_{1E}^*(a), 0, a\tau_{1E}^*(a), \dots, a\tau_{k-1E}^*(a), a \right) \begin{matrix} \\ w_{kE}^*(a), \dots, w_{1E}^*(a), w_{0E}^*(a), w_{1E}^*(a), \dots, w_{kE}^*(a) \end{matrix},$$

и при $a \rightarrow 0$

$$w_{kE}^*(a) \rightarrow \frac{1}{4k}, \quad w_{iE}^*(a) \rightarrow \frac{1}{2k}, \quad \tau_{iE}^*(a) \rightarrow \cos\left(\frac{k-i}{2k}\pi\right), \quad i = 0, \dots, k-1,$$

Отметим, что при $a > \pi(1 - \frac{1}{2k+1})$ D - и E -оптимальные планы совпадают. В табл. 4 и 5 приведены оптимальные планы при некоторых значениях $\bar{a} = a/\pi$ при $k = 2$ и $k = 3$ соответственно.

Таблица 4. D - и E -оптимальные планы при $k = 2$

\bar{a}	0.1	0.2	0.3	0.4	0.5	0.6	0.7
τ_{1D}^*	0.652	0.646	0.634	0.618	0.596	0.569	0.536
τ_{1E}^*	0.706	0.701	0.694	0.682	0.667	0.646	0.620
w_{0E}^*	0.247	0.237	0.220	0.195	0.162	0.122	0.095
w_{1E}^*	0.250	0.250	0.249	0.245	0.238	0.226	0.212
w_{2E}^*	0.127	0.132	0.142	0.157	0.181	0.213	0.241

В табл. 6, 7 проведено сравнение оптимальных планов по величинам $C_D(\xi_E^*) = (\text{eff}_D(\xi_E^*(a), a))^{-1}$ и $\text{eff}_E(\xi_D^*) = \text{eff}_E(\xi_D^*(a), a)$. Мы видим, что эффективности слабо зависят от a и то, что D - и E -оптимальные планы различаются. E -оптимальный план имеет примерно 95% эффективность по D -критерию. D -оптимальный план имеет примерно 85% эффективность по E -критерию.

Таблица 5. D - и E -оптимальные планы при $k = 3$

\bar{a}	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
τ_{1D}^*	0.466	0.461	0.451	0.438	0.421	0.400	0.376	0.349
τ_{2D}^*	0.828	0.825	0.818	0.806	0.789	0.768	0.735	0.694
τ_{1E}^*	0.498	0.494	0.486	0.475	0.460	0.442	0.420	0.353
τ_{2E}^*	0.869	0.863	0.857	0.850	0.839	0.823	0.802	0.752
w_{0E}^*	0.163	0.158	0.148	0.133	0.116	0.098	0.086	0.121
w_{1E}^*	0.163	0.162	0.156	0.146	0.131	0.110	0.081	0.081
w_{2E}^*	0.169	0.171	0.176	0.183	0.190	0.194	0.189	0.155
w_{3E}^*	0.087	0.088	0.094	0.105	0.121	0.147	0.188	0.203

Таблица 6. Сравнение D - и E -оптимальных планов при $k = 2$

\bar{a}	0.1	0.2	0.3	0.4	0.5	0.6	0.7
$C_D(\xi_E^*)$	1.07	1.06	1.05	1.04	1.04	1.06	1.09
$\text{eff}_E(\xi_D^*)$	0.84	0.85	0.86	0.88	0.90	0.91	0.89

Таблица 7. Сравнение D - и E -оптимальных планов при $k = 3$

\bar{a}	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$C_D(\xi_E^*)$	1.06	1.06	1.05	1.05	1.05	1.07	1.13	1.09
$\text{eff}_E(\xi_D^*)$	0.82	0.83	0.83	0.84	0.85	0.86	0.84	0.82

3. Точные оптимальные планы при коррелированных наблюдениях

В этом разделе мы изучим точные D -оптимальные планы для оценок метода взвешенных наименьших квадратов (МВНК) для линейной и квадратичной регрессионных моделей с авторегрессионной структурой ошибок. Мы покажем, что в случае высоко коррелированных наблюдений оптимальный план близок к плану в равномерно отстоящих точках и для несильно коррелированных наблюдений этот план имеет большую эффективность, чем оптимальный план, полученный в условиях некоррелированности. Мы также рассмотрим проблему построения оптимальных планов для оценивания параметра наклона в линейной регрессии и сравним оптимальные планы для оценок МНК и МВНК.

Задача построения точных оптимальных планов в линейной модели при коррелированных наблюдениях, в частности, возникает в токсикологии, где при факторном планировании несколько компонентов при различных количествах сравниваются их способностью сдерживать рост бактерий. На каждом этапе факторного планирования количество бактерий измеряется в трех временных точках. Пусть влияние только одного компонента изучается. Предполагается, что измерения на различных этапах независимы, а измерения в различных временных точках на одном этапе коррелированы с одной и той же ковариационной матрицей. Если задана параметрическая форма для кривой роста бактерий, то параметры могут быть оценены по МВНК. Такой анализ является обобщенным дисперсионным анализом (GMANOVA) (Rao, 1968). Таким образом, задача планирования эксперимента состоит в выборе временных точек, в которых будет измеряться количество бактерий.

В случае некоррелированных наблюдений данная проблема хорошо отражена в литературе (Hoel, 1958; Hohmann, Jung, 1975; Gaffke and Krafft, 1982; Imhof, Krafft and Schaefer, 2000). В этих статьях рассматривались D -, G -, A - и D_1 -оптимальные планы для линейной или квадратичной модели. Нахождение оптимальных планов при коррелированных наблюдениях является значительно более сложной задачей и эта проблема менее исследована. Первой работой по оптимальному планированию для линейной регрессионной модели при коррелированных наблюдениях является работа Хоела (Hoel, 1958), в которой рассматривался МВНК и только равномер-

ные планы. В работах (Bickel, Herzberg, 1979) и (Bickel, Herzberg, Schilling, 1981) рассматривался МНК и определялись асимптотические (при увеличении числа наблюдений) оптимальные планы для регрессии в виде константы, прямой линии через начало координат и оценивания параметра наклона в линейной регрессии. Оптимальные планы также изучались в работах (Abt, Liski, Mandal, Sinha, 1997, 1998) для линейной и квадратичной модели при коррелированных наблюдениях. Следуя Хоелу (Hoel, 1958), эти авторы искали оптимальные планы в классе планов с равномерно отстоящими точками. Далее мы ослабим эти условия и построим планы, когда число наблюдений невелико.

Мы также исследуем эффективность планов, полученных в предположении независимости наблюдений (Hohmann, Jung, 1975; Gaffke, Krafft, 1982).

Теоретические результаты этого пункта получены в совместной работе (Dette, Kunert, Pepelyshev, 2006).

3.1. Постановка задачи

Рассмотрим линейную по параметрам регрессионную модель

$$Y_{t_i} = \beta_1 f_1(t_i) + \dots + \beta_p f_p(t_i) + \varepsilon_{t_i}, \quad i = 1, \dots, n, \quad (26)$$

где f_1, \dots, f_p ($p \in \mathbb{N}$) — некоторые заданные регрессионные функции, условия проведения измерения t_i выбираются из компактного интервала. Не уменьшая общности, можно считать, что $t_i \in [0, 1]$. Параметры β_1, \dots, β_p неизвестны. Предположим, что ошибки $\varepsilon_{t_1}, \dots, \varepsilon_{t_n}$ центрированы и удовлетворяют стационарному авторегрессионному процессу, причем корреляция между двумя измерениями зависит от времени t так, что $E[\varepsilon_t] = 0$ и

$$\sigma_{ts} := \text{Cov}(Y_t, Y_s) = \text{Cov}(\varepsilon_t, \varepsilon_s) = \sigma^2 \lambda^{|t-s|}. \quad (27)$$

Здесь $t, s \in [0, 1]$ и λ — известная константа такая, что $0 \leq \lambda < 1$. Без потери общности, предположим что $\sigma^2 = 1$. Точный план $\xi = \{t_1, \dots, t_n\}$ есть вектор длины n , содержащий условия измерения для модели (26). Пусть $0 \leq t_1 \leq \dots \leq t_n \leq 1$. Перепишем (29) в векторной форме

$$Y = X_\xi \beta + \varepsilon_\xi,$$

где $Y = (Y_{t_1}, \dots, Y_{t_n})^T$ обозначает вектор измерений, $\beta = (\beta_1, \dots, \beta_p)^T$,

$$X_\xi = \begin{bmatrix} f_1(t_1) & \dots & f_p(t_1) \\ \vdots & \vdots & \vdots \\ f_1(t_n) & \dots & f_p(t_n) \end{bmatrix}$$

есть матрица плана и вектор ошибок $\varepsilon_\xi = (\varepsilon_{t_1}, \dots, \varepsilon_{t_n})^T$ имеет ковариационную матрицу

$$\Sigma_\xi = \begin{bmatrix} 1 & \lambda^{(t_2-t_1)} & \lambda^{(t_3-t_1)} & \dots & \lambda^{(t_n-t_1)} \\ \lambda^{(t_2-t_1)} & 1 & \lambda^{(t_3-t_2)} & & \lambda^{(t_n-t_2)} \\ \lambda^{(t_3-t_1)} & \lambda^{(t_3-t_2)} & 1 & \dots & \lambda^{(t_n-t_3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda^{(t_n-t_1)} & \lambda^{(t_n-t_2)} & \lambda^{(t_n-t_3)} & \dots & 1 \end{bmatrix}.$$

Если $t_i = t_{i+1}$ при некотором $1 \leq i \leq n-1$, соответствующие измерения имеют корреляцию 1 и, следовательно, повторное измерение в точке t_{i+1} не дает увеличение информации. По этой причине мы будем в этом разделе всюду считать, что $t_1 < \dots < t_n$. В этом случае матрица Σ_ξ обратима и непосредственные вычисления показывают, что $\Sigma_\xi^{-1} = V_\xi^T V_\xi$, где матрица V_ξ имеет вид

$$V_\xi = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ \frac{-\lambda^{(t_2-t_1)}}{\sqrt{1-\lambda^{2(t_2-t_1)}}} & \frac{1}{\sqrt{1-\lambda^{2(t_2-t_1)}}} & & & \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \frac{-\lambda^{(t_n-t_{n-1})}}{\sqrt{1-\lambda^{2(t_n-t_{n-1})}}} & \frac{1}{\sqrt{1-\lambda^{2(t_n-t_{n-1})}}} \end{bmatrix}.$$

Оценка параметра по методу взвешенных наименьших квадратов имеет вид $\hat{\beta} = (X_\xi^T V_\xi^T V_\xi X_\xi)^{-1} X_\xi^T V_\xi^T V_\xi Y$ с ковариационной матрицей

$$\text{Cov}(\hat{\beta}) = (X_\xi^T V_\xi^T V_\xi X_\xi)^{-1}.$$

Точный D -оптимальный план ξ^* минимизирует $\det(\text{Cov}(\hat{\beta}))$ по множеству планов вида $\xi = \{t_1, \dots, t_n\}$. Эта задача эквивалентна задаче максимизации $\det M_\xi$, где матрица M_ξ задана выражением

$$M_\xi = X_\xi^T V_\xi^T V_\xi X_\xi. \quad (28)$$

Далее мы сосредоточимся на линейной и квадратичной модели.

3.2. Линейная модель

В этом пункте мы рассмотрим линейную модель

$$Y_{t_i} = \mu + \beta t_i + \varepsilon_{t_i}, \quad i = 1, \dots, n. \quad (29)$$

Сперва мы выведем более наглядное представление определителя матрицы M_ξ , определенной в (28). С этой целью мы введем обозначения $d_1 = 0$, $d_i = t_i - t_{i-1}$, $2 \leq i \leq n$, $a_1 = 1$, $b_1 = 0$,

$$a_j = \frac{1}{\sqrt{1 - \lambda^{2d_j}}}, \quad b_j = \frac{\lambda^{d_j}}{\sqrt{1 - \lambda^{2d_j}}}, \quad j = 2, \dots, n,$$

и найдем, что

$$V_\xi X_\xi = \begin{bmatrix} 1 & t_1 \\ a_2 - b_2 & t_1(a_2 - b_2) + d_2 a_2 \\ a_3 - b_3 & t_1(a_3 - b_3) + d_2(a_3 - b_3) + d_3 a_3 \\ a_n - b_n & t_1(a_n - b_n) + (d_2 + \dots + d_{n-1})(a_n - b_n) + d_n a_n \end{bmatrix}.$$

Используя формулу Бине—Коши (Карлин, Стадден, 1976), мы получаем

$$\begin{aligned} \det M_\xi &= \det X_\xi^T V_\xi^T V_\xi X_\xi \\ &= \sum_{1 \leq i < j \leq n} \det^2 \begin{pmatrix} a_i - b_i & t_{i-1}(a_i - b_i) + d_i a_i \\ a_j - b_j & t_{j-1}(a_j - b_j) + d_j a_j \end{pmatrix} \\ &= \sum_{1 \leq i < j \leq n} \det^2 \begin{pmatrix} a_i - b_i & (t_{i-1} - t_1)(a_i - b_i) + d_i a_i \\ a_j - b_j & (t_{j-1} - t_1)(a_j - b_j) + d_j a_j \end{pmatrix}. \quad (30) \end{aligned}$$

Следовательно, план $\tilde{\xi} = \{1 - t_n, \dots, 1 - t_1\}$ есть план, полученный из $\xi = \{t_1, \dots, t_n\}$ отражением точек t_i относительно $t = 1/2$. Тогда $\det M_{\tilde{\xi}} = \det M_\xi$, где матрица M_ξ определена в (28), при $p = 2$, $f_1(t) = 1$ и $f_2(t) = t$.

Доказательство. Отметим, что определители в представлении (30) могут быть представлены в следующем виде:

$$\det^2 \begin{pmatrix} a_i - b_i & a_i t_i - b_i t_{i-1} \\ a_j - b_j & a_j t_j - b_j t_{j-1} \end{pmatrix}.$$

Теперь, записывая это выражение для плана $\tilde{\xi}$, мы получаем утверждение леммы. ■

Теорема 3.1. Пусть ξ – произвольный план с точками $0 \leq t_1 < \dots < t_n \leq 1$. Пусть план ξ^* содержит точки, которые задаются следующими выражениями: $t_1^* = 0$, $t_2^* = t_2 - t_1 = d_2$, $t_3^* = t_3 - t_1 = d_2 + d_3$, …, $t_{n-1}^* = t_{n-1} - t_1 = d_2 + \dots + d_{n-1}$, и $t_n^* = 1$. Тогда $\det M_\xi \leq \det M_{\xi^*}$.

Доказательство. Мы уже видели, что план $\tilde{\xi}$, полученный сдвигом точек к 0, имеет тоже значение D -критерия, что и ξ . Отличие плана ξ^* от $\tilde{\xi}$ состоит только в крайней правой точке $t_n \in [0, 1]$. Мы сейчас покажем, что производная функции критерия $\det(X_{\tilde{\xi}}^T V_{\tilde{\xi}}^T V_{\tilde{\xi}} X_{\tilde{\xi}})$ по переменной d_n положительна, что докажет утверждение теоремы. Введем величины

$$f_i(d_j) = \det \begin{pmatrix} a_i - b_i & (d_1 + \dots + d_{i-1})(a_i - b_i) + d_i a_i \\ a_j - b_j & (d_1 + \dots + d_{j-1})(a_j - b_j) + d_j a_j \end{pmatrix}$$

для $1 \leq i < j \leq n$. Из (30) следует, что

$$\det M_{\tilde{\xi}} = \sum_{1 \leq i < j \leq n} (f_i(d_j))^2$$

и, следовательно,

$$\frac{\partial}{\partial d_n} \det M_{\tilde{\xi}} = \sum_{1 \leq i < n} 2 f_i(d_n) f'_i(d_n),$$

где $f'_i(d_n)$ есть производная $f_i(d_n)$ по d_n . Таким образом, достаточно показать, что $f_i(d_n) > 0$ и $f'_i(d_n) > 0$ для всех $1 \leq i < n$ и для любого $0 < d_n \leq 1$.

С этой целью мы заметим, что для $2 \leq j \leq n$ и $d_j > 0$

$$a_j = \frac{1}{1 - \lambda^{d_j}} (a_j - b_j).$$

Следовательно, для $2 \leq i < n$, мы можем записать

$$\begin{aligned} f_i(d_n) &= \det \begin{pmatrix} a_i - b_i & (d_1 + \dots + d_{i-1} + \frac{d_i}{1-\lambda^{d_i}})(a_i - b_i) \\ a_n - b_n & (d_1 + \dots + d_{n-1} + \frac{d_n}{1-\lambda^{d_n}})(a_n - b_n) \end{pmatrix} \\ &= (a_i - b_i)(a_n - b_n)[d_{i+1} + \dots + d_{n-1} + g(d_n) + \ell(d_i)], \end{aligned}$$

где функции $g(x)$ и $\ell(x)$ заданы выражениями $g(x) = \frac{x}{1-\lambda^x}$ и $\ell(x) = x - \frac{x}{1-\lambda^x}$. Отметим, что $a_j - b_j \geq 0$ для всех j и тогда

$$f_i(d_n) \geq (a_i - b_i)(a_n - b_n)[g(d_n) + \ell(d_i)].$$

При $x \rightarrow 0$ мы имеем $g(x) \rightarrow -1/\ln \lambda > 0$, и производная $g(x)$ равна

$$g'(x) = \frac{1}{(1-\lambda^x)^2}(1 - \lambda^x + x\lambda^x \ln \lambda).$$

Пусть $h(x)$ есть числитель g' . Тогда $h(0) = 0$ и для h' выполняется

$$h'(x) = -\lambda^x \ln(\lambda) + \lambda^x \ln(\lambda) + x\lambda^x (\ln \lambda)^2 = x\lambda^x (\ln \lambda)^2 > 0,$$

для всех $x > 0$. Следовательно, $h(x) > 0$ для всех $x > 0$ и значит, что $g'(x) > 0$. Тогда мы получим

$$g(x) > \lim_{x \rightarrow 0} g(x) = -\frac{1}{\ln \lambda}$$

для всех $x > 0$. С другой стороны,

$$\ell'(x) = 1 - \frac{1}{(1-\lambda^x)^2}(1 - \lambda^x + x\lambda^x \ln \lambda) = -\frac{\lambda^x}{(1-\lambda^x)^2}(1 - \lambda^x + x \ln \lambda).$$

Задав $q(x) = 1 - \lambda^x + x \ln \lambda$, мы найдем, что ее производная

$$q'(x) = -\lambda^x \ln \lambda + \ln \lambda < 0,$$

что влечет $q(x) < q(0) = 0$, для всех $x > 0$. Следовательно, $\ell'(x) > 0$, для всех $x > 0$ и

$$\ell(x) > \lim_{x \rightarrow 0} \ell(x) = \frac{1}{\ln \lambda}$$

для всех $x > 0$. В итоге мы имеем, что для любых $d_i \geq 0$ и для любого $d_n > 0$ выполняется $g(d_n) + \ell(d_i) > -1/\ln \lambda + 1/\ln \lambda = 0$. Отсюда вытекает, что

$$f_i(d_n) > 0,$$

для всех $2 \leq i < n$ и любого $d_n > 0$. Теперь рассмотрим $f'_i(d_n)$. Мы получим, что для $2 \leq i < n$

$$\begin{aligned} f'_i(d_n) &= (a_i - b_i)(a'_n - b'_n)(d_{i+1} + \dots + d_{n-1} + g(d_n) + \ell(d_i)) \\ &\quad + (a_i - b_i)(a_n - b_n)g'(d_n), \end{aligned}$$

где $(a'_n - b'_n)$ есть производная $(a_n - b_n)$ по d_n . Мы уже видели, что $a_i - b_i > 0$, $a_n - b_n > 0$, $g'(d_n) > 0$ и $d_{i+1} + \dots + d_{n-1} + g(d_n) + \ell(d_i) > 0$. Так как

$$a_n - b_n = \frac{1}{\sqrt{1 - \lambda^{2d_n}}}(1 - \lambda^{d_n}) = \sqrt{\frac{1 - \lambda^{d_n}}{1 + \lambda^{d_n}}},$$

мы получаем для $a'_n - b'_n$ неравенство

$$a'_n - b'_n = -\frac{\lambda^{d_n} \ln \lambda}{(1 + \lambda^{d_n})\sqrt{1 - \lambda^{2d_n}}} > 0$$

для любого $d_n > 0$. Следовательно, $f'_i(d_n) > 0$ для любого $d_n > 0$ ($i = 2, \dots, n-1$).

Остается рассмотреть случай $i = 1$. Ясно, что

$$\begin{aligned} f_1(d_n) &= \det \begin{pmatrix} 1 & 0 \\ a_n - b_n & (d_1 + \dots + d_{n-1} + \frac{d_n}{1-\lambda^{d_n}})(a_n - b_n) \end{pmatrix} \\ &= (a_n - b_n)(d_1 + \dots + d_{n-1} + g(d_n)) \end{aligned}$$

положительно. Аналогично,

$$f'_1(d_n) = (a'_n - b'_n)(d_1 + \dots + d_{n-1} + g(d_n)) + (a_n - b_n)g'(d_n)$$

также положительно.

Учитывая все аргументы, мы имеем, что

$$\frac{\partial}{\partial d_n} \det M_{\tilde{\xi}} = \sum_{1 \leq i < n} 2f_i(d_n)f'_i(d_n) > 0$$

для любого $d_n > 0$, что влечет утверждение теоремы. ■

Замечание 3.1. Заметим, что при $\lambda \rightarrow 0$ мы получаем линейную регрессионную модель с независимыми наблюдениями. В этом случае соответствующая информационная матрица $M_{\xi^*}(\lambda)$, определенная в (28), для точного D -оптимального плана не обязательно сходится к информационной матрице для D -оптимального плана при некоррелированных наблюдениях. Для случая некоррелированных наблюдения хорошо известно, что точный n -точечный D -оптимальный план имеет вид

$$\xi_{\text{lim}}^* = \{0, 0, \dots, 0, 1, \dots, 1\},$$

где $k = \lfloor \frac{n}{2} \rfloor$ наблюдений берется в каждой граничной точке интервала $[0, 1]$ и при нечетном n еще одна точка берется либо в 0, либо в 1 (Hohmann, Jung, 1975). Для этого плана мы имеем

$$\det M_{\xi_{\text{lim}}^*} = \frac{1}{1 - \lambda^2}$$

независимо от числа наблюдений n .

Теперь рассмотрим противоположный случай ($\lambda \rightarrow 1$), который соответствует высоко коррелированным наблюдениям. Следующий результат показывает, что в этом случае точный D -оптимальный план сходится к плану в равномерно отстоящих точках на интервале $[0, 1]$.

Теорема 3.2. При $\lambda \rightarrow 1$ точный n -точечный D -оптимальный план для линейной регрессионной модели с корреляционной структурой (27) сходится к (дискретному) равномерному плану $\xi_n = \{0, 1/(n-1), 2/(n-1), \dots, 1\}$.

Доказательство. Вспомним определение обозначений

$$a_i = \frac{1}{\sqrt{1 - \lambda^{2(t_i - t_{i-1})}}}, \quad b_i = \frac{\lambda^{t_i - t_{i-1}}}{\sqrt{1 - \lambda^{2(t_i - t_{i-1})}}}.$$

Раскладывая в ряд Тейлора в точке $\lambda = 1$, мы получаем

$$\begin{aligned} (a_i - b_i)^2 &= \frac{t_i - t_{i-1}}{2}(1 - \lambda) + \frac{t_i - t_{i-1}}{4}(1 - \lambda)^2 \\ &\quad - \frac{(t_i - t_{i-1})^3 - 4(t_i - t_{i-1})}{24}(1 - \lambda)^3 + o((1 - \lambda)^3), \end{aligned}$$

$$(a_i t_i - b_i t_{i-1})^2 = \frac{t_i - t_{i-1}}{2(1-\lambda)} - \frac{(t_i - t_{i-1})(2t_i + 2t_{i-1} - 1)}{4} + (4(t_i - t_{i-1})^3 - (t_i - t_{i-1})) \frac{2(1-\lambda) + (1-\lambda)^2}{48} + o((1-\lambda)^2).$$

$$(a_i - b_i)(a_i t_i - b_i t_{i-1}) = \frac{t_i - t_{i-1}}{2} + (t_i^2 - t_{i-1}^2) \frac{2(1-\lambda) + (1-\lambda)^2}{8} + o((1-\lambda)^2).$$

Подставляя их в выражение $M_\xi = (V_\xi X_\xi)^T (V_\xi X_\xi)$ учитывая, что $\sum_i t_i^p - t_{i-1}^p = 1$, $t_1 = 0$, $t_n = 1$ мы находим, что $\det M_\xi = AB - C^2$, где A , B и C имеет вид

$$A = 1 + \frac{(1-\lambda)}{2} + \frac{(1-\lambda)^2}{4} + [1 + S_{1,1}] \frac{(1-\lambda)^3}{8} + o((1-\lambda)^3), \quad (31)$$

$$\begin{aligned} B(1-\lambda)^2 &= \sum_i (a_i t_i - b_i t_{i-1})^2 (1-\lambda)^2 \\ &= \frac{(1-\lambda)}{2} + \frac{(1-\lambda)^2}{4} + \frac{(1-\lambda)^3}{8} + \frac{(1-\lambda)^4}{16} + o((1-\lambda)^3), \end{aligned} \quad (32)$$

$$\begin{aligned} C(1-\lambda) &= \sum_i (a_i - b_i)(a_i t_i - b_i t_{i-1})(1-\lambda) \\ &= \frac{(1-\lambda)}{2} + \frac{(1-\lambda)^2}{4} + \frac{(1-\lambda)^3}{8} + o((1-\lambda)^3), \end{aligned} \quad (33)$$

где $S_{1,1} = \sum t_i t_{i-1} (t_i - t_{i-1})$. Таким образом, мы имеем

$$\det M_\xi = \frac{1/2}{1-\lambda} + \frac{1}{4} + \frac{1-\lambda}{8} + (1 + S_{1,1}) \frac{(1-\lambda)^2}{16} + o((1-\lambda)^2)$$

и, следовательно, D -оптимальный план сходится (при $\lambda \rightarrow 1$) к плану, который максимизирует выражение

$$S_{1,1} = \sum_{i=2}^n t_i t_{i-1} (t_i - t_{i-1})$$

Учитывая $t_1 = 0$ и $t_n = 1$ и дифференцируя его по оставшимся переменным t_2, \dots, t_{n-1} , мы получаем систему уравнений

$$2t_j t_{j-1} - t_{j-1}^2 + t_{j+1}^2 - 2t_{j+1} t_j = 0,$$

$j = 2, \dots, n-1$, которая дает единственное решение $t_i = (i-1)/(n-1)$, $i = 2, 3, \dots, n-1$. Вычисление матрицы вторых производных в этой

точке показывает, что она является точкой максимума, что завершает доказательство теоремы. ■

Далее мы покажем, что равномерный план $\xi_n = \{0, 1/(n-1), 2/(n-1), \dots, 1\}$ имеет высокую D -эффективность для широко множества значений параметра λ . Но сейчас мы сформулируем соответствующие результаты для оценивания параметра наклона в линейной регрессионной модели. В этом случае оптимальный план максимизирует величину

$$D_1(\xi) = (e_2^T M_\xi^{-1} e_2)^{-1} = \frac{\det M_\xi}{\sum_{j=1}^n (a_j - b_j)^2}, \quad (34)$$

где $e_2 = (0, 1)^T$, $a_1 = 1$, $b_1 = 0$ и последнее равенство следует из (30). Далее оптимальный план, который максимизирует функцию (34), мы будем называть точным D_1 -оптимальным планом.

Теорема 3.3.

- (a) Пусть $\xi = \{t_1, \dots, t_n\}$ есть некоторый план и план $\tilde{\xi} = \{1 - t_n, \dots, 1 - t_1\}$ состоит из точек, отраженных относительно точки $t = 1/2$. Тогда $D_1(\xi) = D_1(\tilde{\xi})$.
- (б) Если $\xi = \{t_1, \dots, t_n\}$ есть точный D_1 -оптимальный план для модели (29) с корреляционной структурой (27), тогда $t_1 = 0, t_n = 1$.
- (в) При $\lambda \rightarrow 1$ точный n -точечный D_1 -оптимальный план для модели (29) с корреляционной структурой (27) сходится к плану $\bar{\xi} = \{0, t_2, t_3, \dots, t_{n-1}, 1\}$, где точки t_2, \dots, t_{n-1} минимизируют величину

$$\frac{S_{1,2}}{6} - \frac{S_{1,1}}{8} - \frac{S_{2,1}}{18} - \frac{S_{1,3}}{18} \quad (35)$$

при $t_1 = 0, t_n = 1$, где

$$S_{p,q} = \sum_{i=2}^n t_i^p t_{i-1}^q (t_i^q - t_{i-1}^q). \quad (36)$$

Доказательство. Пункты (а) и (б) могут быть доказаны аналогично лемме 3.1 и теореме 3.1. Переходим к доказательству пункта (в). Здесь нам потребуются разложения более высоких порядков для $\det M_\xi = AB - C^2$ и A . Непосредственные вычисления показывают, что справедливы следующие выражения:

$$\begin{aligned}
A &= 1 + \frac{(1-\lambda)}{2} + \frac{(1-\lambda)^2}{4} + (1 + S_{1,1}) \frac{(1-\lambda)^3}{8} \\
&\quad + (1 + 3S_{1,1}) \frac{(1-\lambda)^4}{16} + o((1-\lambda)^4), \\
B(1-\lambda)^2 &= \frac{(1-\lambda)}{2} + \frac{(1-\lambda)^2}{4} + \frac{(1-\lambda)^3}{8} + \frac{(1-\lambda)^4}{16} \\
&\quad + \left(\frac{1}{32} + \frac{S_{2,1}+S_{1,3}}{72} \right) (1-\lambda)^5 + o((1-\lambda)^5), \\
C(1-\lambda) &= \frac{(1-\lambda)}{2} + \frac{(1-\lambda)^2}{4} + \frac{(1-\lambda)^3}{8} + (3+2S_{1,2}) \frac{(1-\lambda)^4}{48} + o((1-\lambda)^4).
\end{aligned}$$

Отсюда вытекает, что

$$\begin{aligned}
\det M_\xi &= (1-\lambda)^{-1}/2 + 1/4 + (1-\lambda)/8 + (1 + S_{1,1})(1-\lambda)^2/16 \\
&\quad + \left(\frac{1}{32} + \frac{S_{1,1}}{8} - \frac{S_{1,2}}{24} + \frac{S_{2,1}+S_{1,3}}{72} \right) (1-\lambda)^3 + o((1-\lambda)^3).
\end{aligned}$$

и, используя $1/D_1(\xi) = A/\det M_\xi$, получаем

$$\frac{1}{D_1(\xi)} = 2(1-\lambda) + \left(\frac{S_{1,2}}{6} - \frac{S_{1,1}}{8} - \frac{S_{2,1}}{18} - \frac{S_{1,3}}{18} \right) (1-\lambda)^2 + o((1-\lambda)^2).$$

Следовательно, точный D_1 -оптимальный план для линейной модели с корреляционной структурой (27) сходится к плану $\xi = \{0, t_2, t_3, \dots, t_{n-1}, 1\}$, где точки t_2, \dots, t_{n-1} минимизируют функцию (35). ■

3.3. Численные результаты

В этом пункте представлены некоторые численные результаты для точных D -оптимальных планов. Также будет исследована эффективность D -оптимального плана ξ_{lim}^* для линейной модели с некоррелированными наблюдениями и эффективность равномерного плана ξ_n , полученного в теореме 3.2.

Сначала рассмотрим случай $n = 3$. Из теоремы 3.1 следует, что достаточно искать оптимальный план в виде $t_1 = 0, t_2 = d$ и $t_3 = 1$. При $n = 3$ D -критерий принимает вид

$$\begin{aligned}
\det(X_\xi^T V_\xi^T V_\xi X_\xi) &= (1 + (a_2 - b_2)^2 + (a_3 - b_3)^2) (d^2 a_2^2 + (a_3 - db_3)^2) \\
&\quad - (da_2(a_2 - b_2) + (a_3 - db_3)(a_3 - b_3))^2 \\
&= \frac{2 \left((1 - (1-d)\lambda^d)(1 - d\lambda^{1-d}) - d(1-d) \right)}{(1 - \lambda^{2d})(1 - \lambda^{2(1-d)})} = \psi(d), \quad (37)
\end{aligned}$$

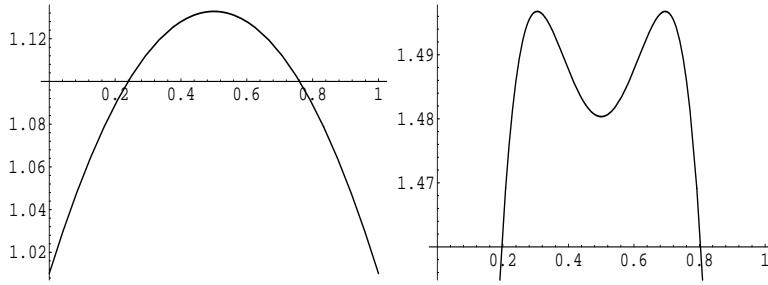


Рис. 2. Функция ψ , определенная в (37), при $\lambda = 0.1$ (слева) и при $\lambda = 0.0001$ (справа).

где последняя строка определяет функцию ψ . Следовательно, точный 3-точечный D -оптимальный план может быть получен посредством максимизирования функции ψ по $d \in (0, 1)$. Нетрудно видеть, что функция (37) симметрична относительно точки $d = 1/2$. Однако не представляется возможным найти максимум функции ψ аналитически.

Мы найдем решение этой задачи численно для некоторых значений параметра λ . Для некоторого множества значений λ величина $\det(X_\xi^T V_\xi^T V_\xi X_\xi)$ достигает максимума при $d = 1/2$. А именно, если параметр λ не очень мал, план $\xi = \{0, 1/2, 1\}$ является D -оптимальным для линейной модели (29). В левой части рис. 2 приведен пример функции ψ , соответствующий случаю $\lambda = 0.1$.

При λ стремящимся к 0 ситуация меняется и становится более сложной. В правой части рис. 2 показан график функции ψ при $\lambda = 0.0001$. В этом случае функция ψ имеет локальный минимум в точке $d = 1/2$ и имеет два глобальных максимума, соответствующих различным D -оптимальным планам. А именно существуют два D -оптимальных плана для линейной модели для $n = 3$ наблюдений, которые имеют вид $\{0, 0.2459, 1\}$ и $\{0, 0.7541, 1\}$. То есть при очень маленьких значениях параметра λ существуют два неравномерных точных D -оптимальных плана. Для продолжения исследования мы рассмотрим вторую производную функции ψ

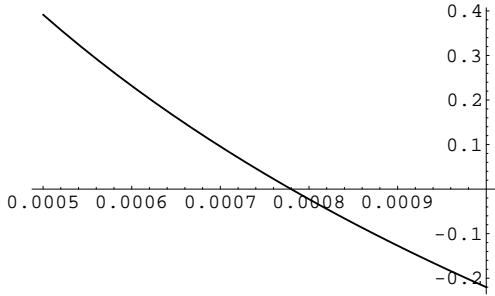


Рис. 3. Зависимость второй производной функции ψ , определенной в (37), в точке $d = 1/2$ от λ .

$$\psi''(1/2) = -4 \frac{\left(\lambda + 1/2 \sqrt{\lambda} (\ln(\lambda))^2 - 2 \sqrt{\lambda} \ln(\lambda) - 1 - \frac{(-4\sqrt{\lambda} + \lambda + 3)\lambda(\ln(\lambda))^2}{(1-\lambda)^2} \right)}{(1-\lambda)^2}, \quad (38)$$

в точке $d = 1/2$ как функцию от λ . Эта функция представлена на рис. 3. Видно, что она отрицательна при всех $\lambda > 0.0007798 = \lambda^*$ и положительна при $\lambda < \lambda^*$. Отсюда следует, что оптимальный план сосредоточен в равноотстоящих точках 0, 1/2, 1 при любых $\lambda > \lambda^*$. При $\lambda < \lambda^*$ оптимальный план является неравномерным и его средняя точка лежит ближе к границе отрезка планирования. Таким образом, численные исследования позволяют сформулировать следующий результат.

Для линейной регрессионной модели с коррелированными наблюдениями точный 3-точечный D-оптимальный план равен $\xi_3 = \{0, 1/2, 1\}$ при $\lambda \geq \lambda^$ и равен $\{0, d, 1\}$ или $\{0, 1-d, 1\}$ при $\lambda < \lambda^*$, где $d = d(\lambda) \in [0, 1/2]$ есть единственное решение уравнения*

$$\psi'(d) = 0,$$

где функция ψ определена в (37).

В табл. 8 представлены нетривиальные точки точного D-оптимального плана для оценки по МВНК для линейной модели (29) с корреляционной структурой (27) для $n = 3$ наблюдений. В таблице также содержатся D-эффективность точного D-оптимального

плана ξ_{lim}^* при некоррелированных наблюдениях и эффективности равномерного плана ξ_3 . Видно, что равномерный план имеет очень высокую эффективность при разных уровнях корреляций.

Таблица 8. Нетривиальная точка $d(\lambda)$ точного D -оптимального плана для оценки по МВИК для линейной модели (29) с корреляционной структурой (27) для $n = 3$ наблюдений для различных значений параметра λ , а также D -эффективность точного D -оптимального плана $\xi_{lim}^* = \{0, 0, 1\}$ для некоррелированных наблюдений и эффективность равномерного плана $\xi_3 = \{0, 0.5, 1\}$.

Точный D -оптимальный план равен $\xi^* = \{0, d(\lambda), 1\}$

λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d(\lambda)$.5	.5	.5	.5	.5	.211	.155
$eff(\xi_3)$	1.0	1.0	1.0	1.0	1.0	.972	.947
$eff(\xi_{lim}^*)$.999	.996	.944	.867	.831	.794	.773

Рассмотрим теперь случай $n = 4, 5, 6$. При $n \geq 4$ исследование D -оптимальных планов для линейной модели (29) с корреляционной структурой (27) проведем численными методами. При $n = 4$ из теоремы 3.1 следует что точный D -оптимальный план имеет вид $\xi^* = \{0, t_2, t_3, 1\}$. Однако численный анализ показывает, что точный 4-точечный D -оптимальный план имеет вид

$$\{0, d, 1 - d, 1\},$$

где $d = d(\lambda) \in (0, 0.5)$. В табл. 9 представлены D -оптимальные планы для линейной модели (29) с корреляционной структурой (27) для $n = 4$ наблюдений при различных значениях λ . В ней также даны D -эффективности плана $\xi_{lim}^* = \{0, 0, 1, 1\}$, оптимального для некоррелированных наблюдений и равномерного плана $\xi_4 = \{0, 1/3, 2/3, 1\}$. Интересно отметить, что равномерный план снова имеет высокую эффективность при разных значениях параметра λ .

Ситуация для $n = 5$ и $n = 6$ аналогична. Точные оптимальные планы для $n = 5$ и $n = 6$ наблюдений представлены в табл. 9. Численный анализ показывает, что для пяти наблюдений точный

D -оптимальный план имеет вид $\{0, d_1, d_2, 1 - d_1, 1\}$ (или его отражение относительно точки $t = 1/2$), где $d_1 = d_1(\lambda) \in (0, 0.5)$ и $d_2 = d_2(\lambda) \in (0, 0.5]$. Для $n = 6$ наблюдений точный D -оптимальный план имеет вид $\{0, d_1, d_2, 1 - d_2, 1 - d_1, 1\}$, где $d_1 = d_1(\lambda) \in (0, 0.5)$ и $d_2 = d_2(\lambda) \in (0, 0.5)$.

Таблица 9. Нетривиальные точки точных D -оптимальных планов для оценок МВНК для линейной регрессионной модели (29) с корреляционной структурой (27) при $n = 4$, $n = 5$ и $n = 6$ наблюдений. D -эффективность D -оптимального плана ξ_{lim}^* для некоррелированных наблюдений и эффективность равномерного плана ξ_n

$n = 4$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d(\lambda)$.332	.328	.317	.303	.281	.192	.146
$\text{eff}(\xi_4)$	1.0	1.0	1.0	.998	.993	.947	.900
$\text{eff}(\xi_{\text{lim}}^*)$	1.0	.996	.928	.806	.731	.642	.604
$n = 5$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d_1(\lambda)$.249	.243	.233	.224	.215	.181	.142
$d_2(\lambda)$.5	.5	.5	.5	.5	.424	.276
$\text{eff}(\xi_5)$	1.0	1.0	1.0	1.0	.996	.975	.931
$\text{eff}(\xi_{\text{lim}}^*)$	1.0	1.0	.922	.780	.685	.573	.530
$n = 6$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d_1(\lambda)$.199	.194	.184	.177	.171	.146	.115
$d_2(\lambda)$.399	.397	.391	.385	.380	.340	.261
$\text{eff}(\xi_6)$	1.0	1.0	1.0	.999	.997	.981	.942
$\text{eff}(\xi_{\text{lim}}^*)$	1.0	.995	.919	.767	.659	.519	.469

Для больших значений n структуру точного D -оптимального плана можно описать следующим образом: если $n = 2k$, точный $2k$ -точечный D -оптимальный план имеет вид

$$\{0, d_1, \dots, d_{k-1}, 1 - d_{k-1}, \dots, 1 - d_1, 1\}, \quad (39)$$

где $d_i = d_i(\lambda) \in (0, 0.5)$, а если $n = 2k + 1$, точный $(2k + 1)$ -точечный

D -оптимальный план имеет вид

$$\{0, d_1, \dots, d_{k-1}, d_k, 1 - d_{k-1}, \dots, 1 - d_1, 1\} \quad (40)$$

(или его отражение относительно точки $t = 1/2$, где $d_i = d_i(\lambda) \in (0, 0.5)$).

Замечание 3.2. Стоит отметить одну особенность критерия оптимальности. Пусть ξ произвольный план, содержащий концы отрезка планирования. Тогда при $\lambda \rightarrow 1$ эффективность такого плана стремится к 1, что видно из разложения в ряд Тейлора критерия оптимальности.

Оценивание параметра наклона. Сейчас мы кратко опишем результаты о точных оптимальных планах для оценки по МВНК для параметра наклона в линейной модели.

В табл. 10 даны точные оптимальные планы для $n = 3, 4, 5, 6$. В ней также приведены D_1 -эффективности $\text{eff}(\xi) = D_1(\xi)/D_1(\xi^*)$ равномерного плана и точного D_1 -оптимального плана, полученного в предположении некоррелированности наблюдений, который совпадает с D -оптимальным планом для некоррелированных наблюдений. Структура D_1 оптимального плана приведена в (39) и (40) для случаев четного и нечетного числа наблюдений соответственно.

3.4. Точные оптимальные планы для квадратичной модели

В этом пункте мы кратко обсудим задачу о точных D -оптимальных планах для оценок по МВНК для квадратичной регрессионной модели

$$Y_{t_i} = \beta_1 + \beta_2 t_i + \beta_3 t_i^2 + \varepsilon_{t_i} \quad i = 1, \dots, n, \quad (41)$$

с корреляционной структурой наблюдений (27). Построение оптимальных планов проведем численными методами. Однако методами, рассмотренными в предыдущем параграфе, можно получить некоторые простые свойства оптимальных планов для квадратичной модели.

Таблица 10. Нетривиальные точки точного D -оптимального плана для оценивания по МВНК параметра наклона в линейной модели (29) с корреляционной структурой (27) для $n = 3, n = 4, n = 5$ и $n = 6$ наблюдений. D_1 -эффективность точного D_1 -оптимального плана ξ_{lim}^* для некоррелированных наблюдений и эффективность равномерного плана

$n = 3$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d(\lambda)$.146	.147	.151	.151	.145	.118	.097
$eff(\xi_3)$	1.0	1.0	.996	.975	.947	.888	.857
$eff(\xi_{lim}^*)$	1.0	1.0	.996	.975	.947	.888	.857
$n = 4$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d(\lambda)$.180	.180	.178	.172	.163	.133	.109
$eff(\xi_4)$	1.0	1.0	.996	.973	.935	.826	.759
$eff(\xi_{lim}^*)$	1.0	1.0	.990	.941	.877	.747	.683
$n = 5$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d_1(\lambda)$.186	.186	.184	.177	.168	.138	.114
$d_2(\lambda)$.239	.239	.238	.235	.229	.200	.171
$eff(\xi_5)$	1.0	1.0	.998	.986	.962	.866	.789
$eff(\xi_{lim}^*)$	1.0	1.0	.989	.935	.863	.710	.634
$n = 6$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d_1(\lambda)$.112	.112	.111	.109	.106	.093	.081
$d_2(\lambda)$.252	.251	.250	.246	.239	.210	.181
$eff(\xi_6)$	1.0	1.0	.999	.989	.970	.882	.798
$eff(\xi_{lim}^*)$	1.0	1.0	.988	.928	.847	.669	.579

Теорема 3.4.

- (a) Пусть план $\xi^- = \{1 - t_n, \dots, 1 - t_1\}$ получен из плана $\xi = \{t_1, \dots, t_n\}$ отражением точек t_i относительно $t = 1/2$, тогда $\det M_\xi = \det M_{\tilde{\xi}}$, где M_ξ определена в (28) при $p = 3, f_1(t) = 1, f_2(t) = t, f_3(t) = t^2$.
- (б) Точный D -оптимальный план $\xi_n^* = \{t_1, \dots, t_n\}$ для квадра-

тичной модели (41) с корреляционной структурой (27) содержит концы интервала планирования, а именно $t_1 = 0$ и $t_n = 1$.

- (в) При $\lambda \rightarrow 1$ точный n -точечный D -оптимальный план для квадратичной модели (41) сходится к равномерному плану $\xi_n = \{0, 1/(n-1), 2/(n-1), \dots, 1\}$.

Доказательство. Утверждения (а) и (б) могут получены методами из предыдущего параграфа. При $\lambda \rightarrow 1$ для элементов матрицы

$$M_\xi = \begin{pmatrix} A & C & D \\ C & B & E \\ D & E & F \end{pmatrix}$$

справедливы следующие разложения (где $\delta = 1 - \lambda$)

$$\begin{aligned} A &= 1 + \frac{\delta}{2} + \frac{\delta^2}{4} + (1 + S_{1,1}) \frac{\delta^3}{8} + (1 + 3S_{1,1}) \frac{\delta^4}{16} + o(\delta^4), \\ B\delta^2 &= \frac{\delta}{2} + \frac{\delta^2}{4} + \frac{\delta^3}{8} + \frac{\delta^4}{16} + o(\delta^4), \\ C\delta &= \frac{\delta}{2} + \frac{\delta^2}{4} + \frac{\delta^3}{8} + (3 + 2S_{1,2}) \frac{\delta^4}{48} + o(\delta^4), \\ D\delta &= \frac{\delta}{2} + (1 - S_{1,1}) \frac{\delta^2}{4} + (1 - S_{1,1}) \frac{\delta^3}{8} + o(\delta^3), \\ E\delta^2 &= \frac{\delta}{2} + \frac{\delta^2}{4} + (3 - 2S_{1,2}) \frac{\delta^3}{24} + o(\delta^3), \\ F\delta^2 &= (1 + S_{1,1}) \frac{\delta}{2} + (1 - S_{1,1}) \frac{\delta^2}{4} + o(\delta^2), \end{aligned}$$

где $S_{1,1}$ и $S_{1,2}$ определены в (36). Непосредственное вычисление определителя матрицы M_ξ показывает, что

$$\det M_\xi = \frac{1}{16} S_{1,1} (1 - \lambda)^{-2} + o((1 - \lambda)^{-2}).$$

Теперь утверждение теоремы следует из аргументов, данных в доказательстве теоремы 3.2. ■

Численные расчеты показывают, что точный оптимальный план имеет вид (39) при $n = 2k$ и вид (40) при $n = 2k + 1$. В табл. 11 представлены точные D -оптимальные планы для квадратичной модели для $n = 4, n = 5$ и $n = 6$ для различных значений параметра

Таблица 11. Нетривиальные точки точного D -оптимального плана для оценивания по МВНК в квадратичной регрессионной модели (41) с корреляционной структурой (27) для $n = 4$, $n = 5$ и $n = 6$ наблюдений.

D -эффективность точного D -оптимального плана $\xi_{\text{lim}}^* = \{0, 1/2, 1/2, 1\}$

($n = 4$), $\xi_{\text{lim}}^* = \{0, 1/2, 1/2, 1/2, 1\}$ ($n = 5$), $\xi_{\text{lim}}^* = \{0, 0, 1/2, 1/2, 1, 1\}$

($n = 6$) для некоррелированных наблюдений и эффективность равномерного плана ξ_n

$n = 4$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d(\lambda)$.333	.335	.345	.355	.362	.386	.407
$\text{eff}(\xi_4)$	1.0	1.0	1.0	.998	.995	.984	.975
$\text{eff}(\xi_{\text{lim}}^*)$.945	.944	.929	.892	.860	.820	.809
$n = 5$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d_1(\lambda)$.250	.252	.265	.273	.274	.286	.315
$d_2(\lambda)$.500	.500	.500	.500	.500	.500	.500
$\text{eff}(\xi_5)$	1.0	1.0	1.0	.999	.998	.996	.992
$\text{eff}(\xi_{\text{lim}}^*)$.928	.926	.907	.854	.803	.730	.712
$n = 6$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d_1(\lambda)$.200	.202	.215	.220	.214	.194	.142
$d_2(\lambda)$.400	.401	.407	.410	.409	.409	.415
$\text{eff}(\xi_6)$	1.0	1.0	1.0	.999	.999	.999	.992
$\text{eff}(\xi_{\text{lim}}^*)$.921	.919	.897	.835	.772	.668	.634

коррелированности λ . При $n = 3$ равномерный план $\xi_3 = \{0, 1/2, 1\}$ является точным D -оптимальным планом для всех λ . В таблице также приведены D -эффективности равномерного плана ξ_n и точного D -оптимального плана, полученного в предположении некоррелированности наблюдений (Gaffke, Krafft, 1982). Видно, что равномерный план имеет высокую эффективность при разных значениях параметра λ . Например, при $n = 5$ D -эффективность плана ξ_5 как минимум 99.0% при условии, что параметр λ лежит в интервале $[10^{-10}, 1]$. Также видно, что точки точного D -оптимального плана слабо зависят от параметра λ .

Теперь мы кратко сравним точные оптимальные планы для линейной и квадратичной модели. Во-первых, точки оптимальных планов для линейной модели расположены ближе к границе. Во-вторых, из того факта, что в обоих случаях точные оптимальные планы стремятся к равномерному плану при $\lambda \rightarrow 1$, следует оптимальный план для квадратичной модели имеет высокую эффективность в линейной модели, и обратно. Например, при $n = 6$ и $\lambda = 0.01$ эффективность точного D -оптимального плана для квадратичной модели в линейной модели равна 99.7%, а эффективность точного D -оптимального плана для линейной модели в квадратичной модели равна 99.5%.

3.5. Оптимальные планы для оценивания по МНК

В этом пункте мы кратко опишем задачу построения точных D -оптимальных планов для оценивания по МНК. Для простоты ограничимся линейной моделью (29) с корреляционной структурой (27). Заметим, что ковариационная матрица для оценок параметров модели по МНК имеет вид

$$\tilde{M}_\xi^{-1} = (X_\xi^T X_\xi)^{-1} (X_\xi^T (V_\xi^T V_\xi)^{-1} X_\xi) (X_\xi^T X_\xi)^{-1}, \quad (42)$$

где матрицы X_ξ и V_ξ определены в пункте 3.1 при $p = 2$, $f_1(t) = 1$, $f_2(t) = t$. Точный D -оптимальный план для оценивания по МНК в линейной модели с корреляционной структурой (27) максимизирует $\det \tilde{M}_\xi$.

Теорема 3.5.

- (a) Пусть план $\tilde{\xi} = \{1 - t_n, \dots, 1 - t_1\}$ получен из плана $\xi = \{t_1, \dots, t_n\}$ отражением точек относительно $t = 1/2$. Тогда $\det \tilde{M}_{\tilde{\xi}} = \det \tilde{M}_\xi$, где матрица \tilde{M}_ξ определена в (42).
- (б) Точный D -оптимальный план $\xi = \{t_1, \dots, t_n\}$, максимизирующий $\det \tilde{M}_\xi$ удовлетворяет условиям $t_1 = 0$, $t_n = 1$.

В табл. 12 представлены точные D -оптимальные планы для оценивания по МНК в линейной модели с корреляционной структурой (27). Для квадратичной модели оптимальные планы приведены в таблице 13, причем для 3 наблюдений равномерный план $\xi_3 = \{0, 1/2, 1\}$ является D -оптимальным при всех λ . Отметим, что

Таблица 12. Нетривиальные точки точного D -оптимального плана для оценивания по МНК в линейной модели (29) с корреляционной структурой (27) для $n = 4$, $n = 5$ и $n = 6$ наблюдений.
 D -эффективность D -оптимального плана ξ_{lim}^* для некоррелированных наблюдений и эффективность равномерного плана

$n = 3$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d(\lambda)$.000	.000	.500	.500	.500	.212	.155
$\text{eff}(\xi_3)$.997	.994	1.0	1.0	1.0	.972	.947
$\text{eff}(\xi_{\text{lim}}^*)$	1.0	1.0	.950	.867	.833	.794	.773
$n = 4$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d(\lambda)$.000	.000	.334	.312	.288	.193	.147
$\text{eff}(\xi_4)$.986	.977	1.0	.999	.994	.948	.900
$\text{eff}(\xi_{\text{lim}}^*)$	1.0	1.0	.950	.813	.734	.642	.604
$n = 5$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d_1(\lambda)$.000	.000	.000	.216	.216	.183	.142
$d_2(\lambda)$.000	.000	.500	.500	.500	.433	.278
$\text{eff}(\xi_5)$.975	.961	.978	.997	.996	.976	.932
$\text{eff}(\xi_{\text{lim}}^*)$	1.0	1.0	.941	.795	.690	.573	.531
$n = 6$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d_1(\lambda)$.000	.000	.000	.135	.164	.149	.116
$d_2(\lambda)$.000	.402	.339	.387	.388	.346	.263
$\text{eff}(\xi_6)$.966	.946	.951	.992	.995	.982	.943
$\text{eff}(\xi_{\text{lim}}^*)$	1.0	.997	.929	.788	.668	.521	.470

для линейной модели оптимальные планы имеют сложное поведение. При $n = 3, 4$ и 5 существует критическое значение λ^* такое, что точный D -оптимальный план для некоррелированных наблюдений является также D -оптимальным для коррелированных наблюдений при $\lambda > \lambda^*$. При $\lambda < \lambda^*$ поведение плана меняется. Оптимальные планы приведены в табл. 12. Такой критической точки не существует для квадратичной модели. В обеих моделях равномерный план также оказывается очень эффективным, а эффективность точного

D -оптимального плана для некоррелированных наблюдений ощущимо меньше.

Таблица 13. Нетривиальные точки точного D -оптимального плана для оценивания по МНК в квадратичной модели (29) с корреляционной структурой (27) для $n = 3$, $n = 4$, $n = 5$ и $n = 6$ наблюдений.

D -эффективность D -оптимального плана ξ_{lim}^* для некоррелированных наблюдений и эффективность равномерного плана

$n = 4$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d(\lambda)$.352	.356	.359	.359	.363	.386	.407
$\text{eff}(\xi_4)$.999	.999	.998	.996	.994	.984	.975
$\text{eff}(\xi_{\text{lim}}^*)$.951	.950	.933	.894	.861	.820	.809
$n = 5$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d_1(\lambda)$.304	.310	.305	.288	.279	.286	.315
$d_2(\lambda)$.500	.500	.500	.500	.500	.500	.500
$\text{eff}(\xi_5)$.996	.995	.994	.996	.997	.996	.992
$\text{eff}(\xi_{\text{lim}}^*)$.944	.943	.920	.861	.805	.730	.712
$n = 6$							
λ	.9	.5	.1	.01	.001	10^{-6}	10^{-9}
$d_1(\lambda)$.290	.301	.289	.250	.228	.197	.144
$d_2(\lambda)$.413	.422	.425	.415	.410	.408	.415
$\text{eff}(\xi_6)$.991	.989	.990	.994	.998	.999	.993
$\text{eff}(\xi_{\text{lim}}^*)$.945	.945	.921	.850	.780	.668	.635

Литература

1. Ганнинг Р., Росси Х. Аналитические функции многих комплексных переменных. М.: Мир, 1969.
2. Математическая теория планирования эксперимента / Под ред. С.М. Ермакова. М.: Наука, 1983.
3. Ермаков С.М., Жигляевский А.А. Математическая теория оптимального эксперимента. М.: Наука, 1987.
4. Матвеев Н.М. Лекции по аналитической теории дифференциальных уравнений. СПб.: Изд-во С.-Петерб. ун-та, 1995.
5. Карлин С., Стадден В. Чебышевские системы и их применение в анализе и статистике. М.: Наука, 1976.
6. Мелас В.Б. E-оптимальные планы эксперимента. СПб.: Изд-во С.-Петерб. ун-та, 1997.
7. Мелас В.Б. Локально оптимальные планы эксперимента. СПб.: Изд-во СПбГТУ, 1999а.
8. Мелас В.Б. Общая теория функционального подхода к оптимальному планированию эксперимента. СПб.: Издательство С.-Петерб. ун-та, 1999б.
9. Мелас В.Б., Пепельшев А.Н. Степенные разложения неявных функций и локально оптимальные планы эксперимента // Статистические модели с приложениями в эконометрике и смежных областях. СПб.: Изд-во НИИХ СПбГУ, 1999. С. 108–118.
10. Rao C.R. Линейные статистические методы и их применение. М.: Наука, 1968.
11. Сеге Г. Ортогональные многочлены. М.: Наука, 1962.
12. Федоров В.В. Теория оптимального эксперимента. М.: Наука, 1971а.
13. Федоров В.В. Планирование при нелинейной зависимости поверхности отклика от оцениваемых параметров. Препринт 21. М.: Изд-во Моск. ун-та, 1971б.

14. *Abt M., Liski E.P., Mandal N.K., Sinha B.K.* Optimal designs in growth curve models: Part I Correlated model for linear growth: Optimal designs for slope parameter estimation and growth prediction // *Journal of Statistical Planning and Inference*. 1997. N 64. P. 141–150.
15. *Abt M., Liski E.P., Mandal N.K., Sinha B.K.* Optimal designs in growth curve models: Part II Correlated model for quadratic growth: Optimum designs for slope parameter estimation and growth prediction // *Journal of Statistical Planning and Inference*. 1998. N 67. P. 287–296.
16. *Bickel, P.J.; Herzberg, A.M.* Robustness of design against autocorrelation in time I: Asymptotic theory, optimality for location and linear regression // *Ann. Stat.* 1979. N 7. P. 77–95.
17. *Bickel P.J., Herzberg A.M., Schilling M.F.* Robustness of Design Against Autocorrelation in Time II: Optimality, Theoretical and Numerical Results for the First-Order Autoregressive Process // *Journal of the American Statistical Association*. 1981. N 76. P. 870–877.
18. *Box G.E.P., Hunter W.G.* Sequential design of experiments for nonlinear models // *Proc. IBM Scient. Comp. Symposium in Statistics*. 1965. P. 113–137.
19. *Chernoff H.* Locally optimal design for estimating parameters // *Ann. Math. Stat.* 1953. Vol. 24. P. 586–602.
20. *Dette H., Haines L.M., Imhof L.* Maximin and Bayesian optimal designs for regression models. Preprint. 2003. Available at: <http://ruhr-uni-bochum.de/mathematik3/preprint.htm>
21. *Dette H., Kunert J., Pepelyshev A.* Exact optimal designs for weighted least squares analysis with correlated errors. Preprint Bochum: Ruhr-Universität Bochum, 2006. P. 22.
22. *Dette H., Melas V.B.* *E*-optimal designs in Fourier regression model. Preprint Bochum: Ruhr-Universität Bochum. 2001. P. 39.
23. *Dette H., Melas V.B.* Optimal designs for estimating individual coefficients in Fourier regression models // *Ann. Stat.* 2003. Vol. 3. N 5. P. 1669–1692.

24. *Dette H., Melas V.B., Pepelyshev A.* Optimal designs for estimating individual coefficients in polynomial regression – a functional approach // *J. of Statis. Plan. and Inference*. 2004. N 118. P. 201–219.
25. *Dette H., Melas V.B., Pepelyshev A.N.* D-optimal designs for trigonometric regression models on a partial circle // *Ann. Inst. Stat. Math.* 2002. Vol. 54. N 4. P. 945–959.
26. *Fedorov, V.V., Hackl P.* Model-oriented design of Experiments. Springer. 1997.
27. *Gaffke N., Krafft O.* Exact D-optimum designs for quadratic regression // *J. R. Stat. Soc.* 1982. Ser. B. N 44. P. 394–397.
28. *Hoel P.G.* Efficiency Problems in Polynomial Estimation // *Annals of Mathematical Statistics* 1958. N 29. P. 1134–1145,
29. *Hohmann G., Jung W.* On sequential and nonsequential D-optimal experimental design // *Biometr. Z.* 17, 329–336. 1975.
30. *Imhof L., Krafft O., Schaefer M.* D-optimal exact designs for parameter estimation in a quadratic model // *Sankhya*. 2000. Ser. B. Vol. 62. N 2. P. 266–275.
31. *Jennrich R.J.* Asymptotic properties of non-linear least squares estimators // *Ann. Math. Stat.* 1969. N. 40. P. 633–643.
32. *Kiefer J., Wolfowitz J.* The equivalence of two extremum problems // *Canad. J. Math.* 1960. Vol. 14, P. 363–366.
33. *Kitsos C.P., Titterington D.M., Torsney B.* An optimal design problem in rhythmometry // *Biometrics*. 1998. N. 44. P. 657–671,
34. *Melas V.B.* Optimal designs for exponential regression // *Math. Operations forsh. Statist.* 1978. Ser. Statistics. N 9. P. 45–59.
35. *Pukelsheim F.* Optimal design of experiments. N.Y.: Wiley, 1993.

Содержание

Предисловие	3
1. Основные понятия теории оптимального планирования	4
1.1. Постановка задачи и обзор основных результатов теории оптимального планирования	4
1.1.1. Постановка задачи	4
1.1.2. Критерии оптимальности	8
1.1.3. Критерии эффективности	11
1.1.4. Теоремы эквивалентности	12
1.1.5. Неравноточные и коррелированные измерения	14
1.2. Идея и методология функционального подхода	16
1.3. Рекуррентные формулы для вычисления коэффици- ентов рядов Тейлора неявных функций	18
1.3.1. Постановка задачи	18
1.3.2. Метод Эйлера	19
1.3.3. Рекуррентный алгоритм	20
2. D-оптимальные планы для тригонометрической модели на произвольном отрезке	25
2.1. Постановка задачи	25
2.2. Предварительные результаты	25
2.3. Аналитичность точек D-оптимального плана	29
2.4. Функционально-алгебраический подход	35
2.5. Сравнение D- и E-оптимальных планов	40
3. Точные оптимальные планы при коррелированных наблюдениях	42
3.1. Постановка задачи	43
3.2. Линейная модель	45
3.3. Численные результаты	52
3.4. Точные оптимальные планы для квадратичной модели	57
3.5. Оптимальные планы для оценивания по МНК	61
Литература	64

Учебное издание

Андрей Николаевич Пепелышев

**Некоторые современные методы исследования
оптимальных планов эксперимента**

Учебное пособие

Зав. редакцией *Г.И. Чередниченко*
Редактор *Ф.С. Бастиан*
Обложка *А.В. Калининой*
Техн. редактор *Л.Н. Иванова*

Подписано в печать 21.08.2006.
Ф-т 60×84/16. Усл. печ. л. 3,95.
Уч.-изд. л. 3,71. Тираж 110 экз. Заказ N

Редакция оперативной подготовки изданий
С.-Петербургского государственного университета.
199034, С.-Петербург, Университетская наб., 7/9.

Типография Издательства СПбГУ.
199061, С.-Петербург, Средний пр., 41.