Кафедра статистического моделирования
Дипломная работа
студентки 522-й группы Казаковой Виктории Николаевны

Оптимальные планы для коррелированных наблюдений

Научный руководитель:
к.ф.-м.н., доцент А.Н. Пепелышев
Рецензент:
д.ф.-м.н., профессор В.Б. Мелас

Основные понятия

- Пусть результаты эксперимента $y_1, \dots, y_n \in \mathbf{R}$ описываются уравнением

$$y_j = \eta(x_j, \beta) + \varepsilon(x_j),\ j = 1, \dots, n,$$

где β — вектор оцениваемых параметров,

$$E\varepsilon(x_j) = 0, \quad E\varepsilon(x_j)\varepsilon(x_i) = \rho(x_i, x_j).$$

- Точным планом эксперимента называется $\xi = (x_1, \dots, x_n)$.
- Информационной матрицей плана эксперимента ξ называется матрица

$$M(\xi) = F^T G^{-1} F,$$

где $F = \Big(f(x_1)\vdots\dots\vdots f(x_n)\Big)$, $G = \Big(\rho(x_i, x_j)\Big)_{i,j=1}^{n}$, $f(x) = \dfrac{\partial\eta(x,\beta)}{\partial\beta}$.

- Оценка МНК имеет ковариационную матрицу $M^{-1}(\xi)$.
- План ξ^* называется Д-оптимальным, если $\det M(\xi) \le \det M(\xi^*)$.

Полиномиальная модель с одинаковыми корреляциями

$$\eta(x,\beta) = \beta^T f(x), \quad f(x) = (1, x, \ldots, x^p)^T, \quad x \in [c,d]$$

- **Теорема.** *Пусть для полиномиальной модели при коррелированных наблюдениях в форме* $(E\varepsilon_i\varepsilon_j)_{i,j} = \sigma^2 G_n$, *где* σ^2 *— дисперсия ошибок, корреляционная матрица имеет вид*

$$G_n = \begin{pmatrix} 1 & r & \ldots & r \\ r & 1 & \ldots & r \\ \vdots & \vdots & \ddots & r \\ r & \ldots & r & 1 \end{pmatrix} \in \mathbf{R}^{n\times n}$$

и $\frac{-1}{n-1} < r < 1$.
Тогда точный Д-оптимальный план не зависит от r *и совпадает с точным Д-оптимальным планом* ξ_n^* *для некоррелированных наблюдений.*

Планы ξ_n^* получены в (Gaffke, Krafft, 1982).

Экспоненциальная модель

$$\eta(x,a,b) = ae^{-bx}, \qquad b>0,\ x\in[0,\infty).$$

- Корреляционная матрица ошибок измерений имеет вид

$$\mathbf{G} = \left(e^{-\lambda|x_i-x_j|}\right)_{i,j}, \quad \lambda>0.$$

- Матрицу $\mathbf{G^{-1}}$ можно представить в виде $\mathbf{G^{-1}} = \mathbf{V^T V}$, где

$$\mathbf{V} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ b_2 & a_2 & 0 & \dots & 0 \\ 0 & b_3 & a_3 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & b_n & a_n \end{pmatrix},$$

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

Локально Д-оптимальные планы

$$\det \mathbf{M} = a^2 \sum_{1\le i<j\le n} \Big[a_i a_j e^{-b(x_i+x_j)} (1-e^{(\lambda-b)(x_{i-1}-x_i)})(x_{j-1}e^{(\lambda-b)(x_{i-1}-x_i)}-x_j)-$$

$$-(1-e^{(\lambda-b)(x_{j-1}-x_j)})(x_{i-1}e^{(\lambda-b)(x_{i-1}-x_i)}-x_i)\Big]^2.$$

- **Теорема.** *Локально Д-оптимальный план* $\xi^* = \arg\max_\xi \det \mathbf{M}(\xi, a, b, \lambda)$

  1) не зависит от параметра a*,*

  2) содержит точку 0 *в своем носителе,*

  3) удовлетворяет условию

$$x_i^*(\gamma b, \gamma\lambda) = \frac{1}{\gamma} x_i^*(b, \lambda),$$

где $\gamma > 0$, $x_i^*(b,\lambda)$ *— точки локально Д-оптимального плана.*

Следствие. Достаточно изучить оптимальные планы при $b = 1$.

Локально Д-оптимальные планы, случай $n = 2$

- Минимально возможное число точек в плане для экспоненциальной модели равно $n = 2$.

$$\det \mathbf{M}(x_2) = \frac{x_2^2 e^{-2bx_2}}{1 - e^{-2\lambda x_2}}$$

- Локально Д-оптимальный двух-точечный план имеет вид

$$\{0, x^*(b, \lambda)\},$$

где

$$x^*(b, \lambda) = \frac{-t^*}{2\lambda},$$

t^* есть решение уравнения

$$\frac{1}{1 - e^{-t}} = \frac{b}{\lambda} + \frac{2}{t}.$$

- $x_2^*(b, \lambda) \to 1/b$ при $\lambda \to \infty$.
- $x_2^*(b, \lambda) \to 1/(2b)$ при $\lambda \to 0$.

Локально Д-оптимальные планы, случай $n = 3$

- Локально Д-оптимальный план имеет вид $\xi^* = \{0, x_2^*, x_3^*\}$.

- Существуют такие значения параметра λ, при которых функция $\det \mathbf{M}(x_2, x_3, \lambda)$ имеет два локальных максимума по (x_2, x_3).

  Пусть λ^* есть то значение параметра λ, при котором происходит смена доминирования одного локального максимума над другим.

- Точки локально Д-оптимального трех-точечного плана имеют скачок с

$$x_2^* = 0.57029, \quad x_3^* = 3.23859$$

на

$$x_2^* = 0.34007, \quad x_3^* = 0.88700,$$

при переходе λ через значение $\lambda^* = 0.22367$.

Функциональный подход (Мелас, 1981)

- Введем систему уравнений

$$g(x_2, x_3, \lambda) = 0,$$

где

$$g(x_2, x_3, \lambda) = \begin{pmatrix} \frac{\partial}{\partial x_2} \det M(x_2, x_3, \lambda) \\ \frac{\partial}{\partial x_3} \det M(x_2, x_3, \lambda) \end{pmatrix}.$$

- Точки оптимального плана как функции от λ удовлетворяют

$$g(x_2(\lambda), x_3(\lambda), \lambda) \equiv 0$$

-

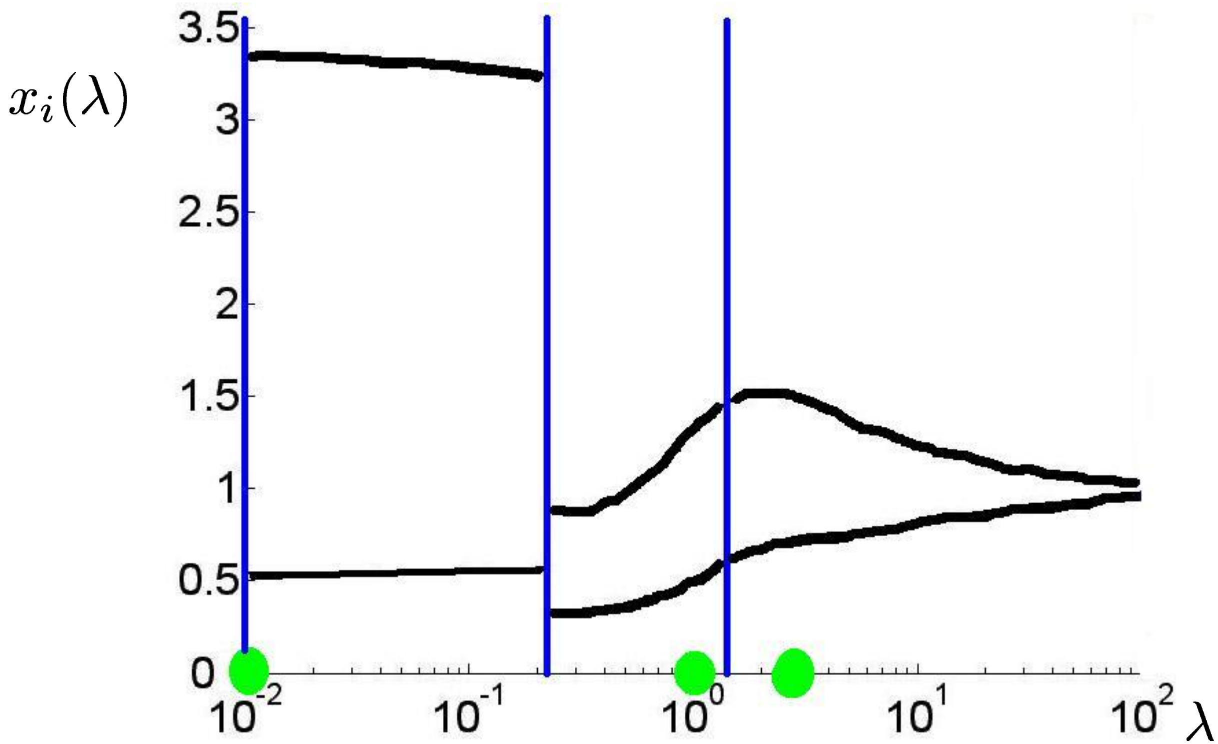
$$x(\lambda) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \sum_{k=0}^{\infty} x_{(k)} (\lambda - \lambda_0)^k$$

- Рекуррентные формулы для вычисления коэффициентов разложения в ряд Тейлора (Мелас, Пепелышев, 1999)

$$x_{(s)} = -J_0^{-1} \left(g \left(\sum_{k=0}^{s-1} x_{(k)} (\lambda - \lambda_0)^k, \lambda \right) \right)_{(s)}, \quad s = 1, 2, 3, \ldots, \quad J_0 = \frac{\partial}{\partial x} g(x_{(0)}, \lambda_0)$$

Локально Д-оптимальные планы, случай $n = 3$

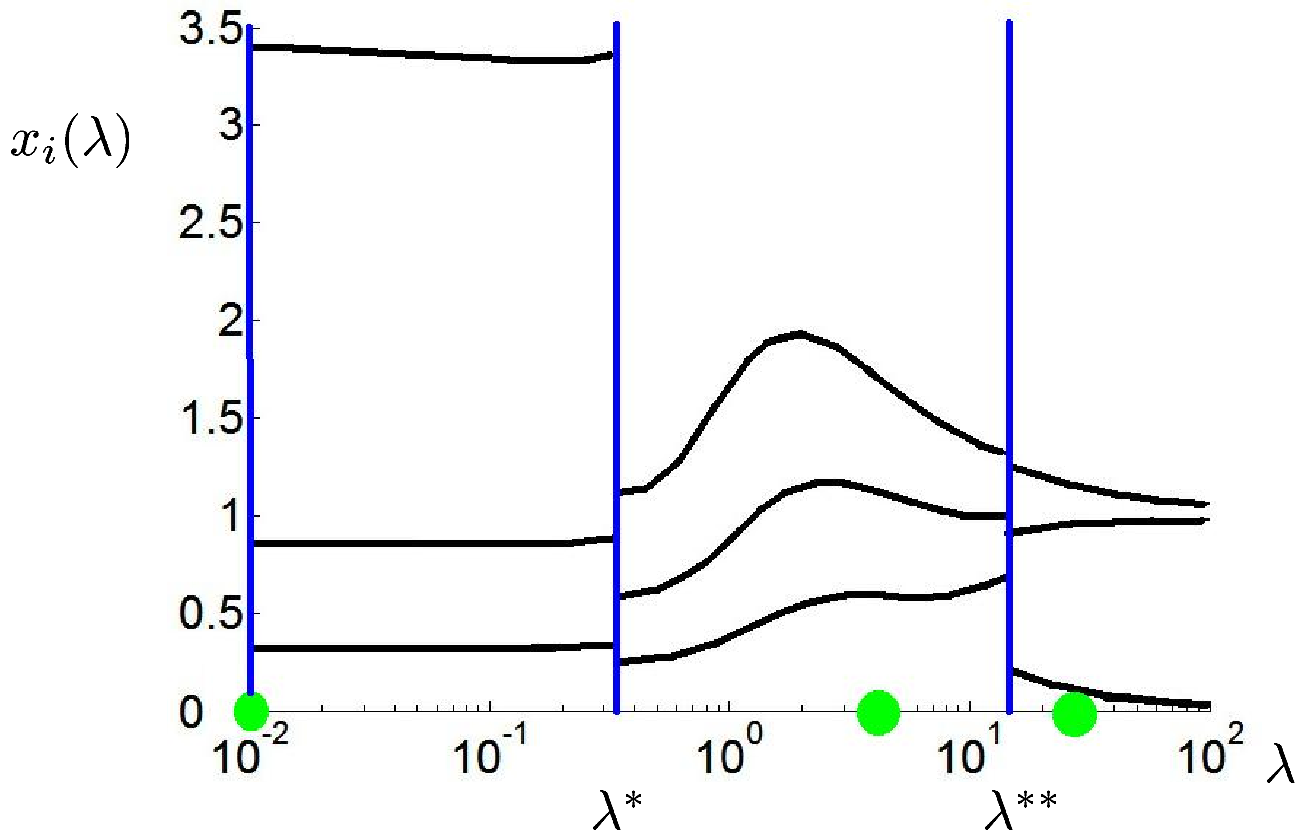
- $x_2(\lambda) = 0.5395 + 0.1096\lambda + 0.1156\lambda^2 + \dots$
  $x_3(\lambda) = 3.3560 - 0.6662\lambda + 1.8098\lambda^2 + \dots$
- $x_2(\lambda) = 0.5087 + 0.2687(\lambda - 1) - 0.0541(\lambda - 1)^2 + \dots$
  $x_3(\lambda) = 1.3056 - 0.4326(\lambda - 1) + 0.4930(\lambda - 1)^2 + \dots$
- $x_2(\lambda) = 0.6911 - 0.3836(\nu - 1/2) + 1.1401(\nu - 1/2)^2 + \dots$
  $x_3(\lambda) = 1.5177 - 0.0556(\nu - 1/2) - 1.4224(\nu - 1/2)^2 + \dots,\ \nu = 1/\lambda$



Локально Д-оптимальные планы, случай $n = 4$

- $\lambda^* = 0.3348$, $\lambda^{**} = 14.3777$

λ	$\lambda^* - 0$	$\lambda^* + 0$	$\lambda^{**} - 0$	$\lambda^{**} + 0$
$x_2(\lambda)$	0.3462	0.2491	0.6966	0.2030
$x_3(\lambda)$	0.8919	0.5841	1.0074	0.9250
$x_4(\lambda)$	3.3611	1.1180	1.3133	1.2490



Максиминно-эффективные планы

- **Определение.** *План будем называть максиминно эффективным Д-оптимальным планом, если он максимизирует величину*

$$\min_{\beta\in\Omega}\left[\frac{\det M(\xi,\beta)}{\det M(\xi^*_{loc}(\beta),\beta)}\right]^{1/2},$$

где $\xi^*_{loc}(\beta)$ — локально Д-оптимальный план.

- **Теорема.** *Максиминный Д-оптимальный план*
  1) *не зависит от параметра a,*
  2) *содержит точку 0 в своем носителе,*
  3) *удовлетворяет условию*

$$x^*_i(\gamma\Omega)=\frac{1}{\gamma}x^*_i(\Omega),$$

где $\gamma>0$, $x^*_i(\Omega)$ — точки максиминного Д-оптимального плана.

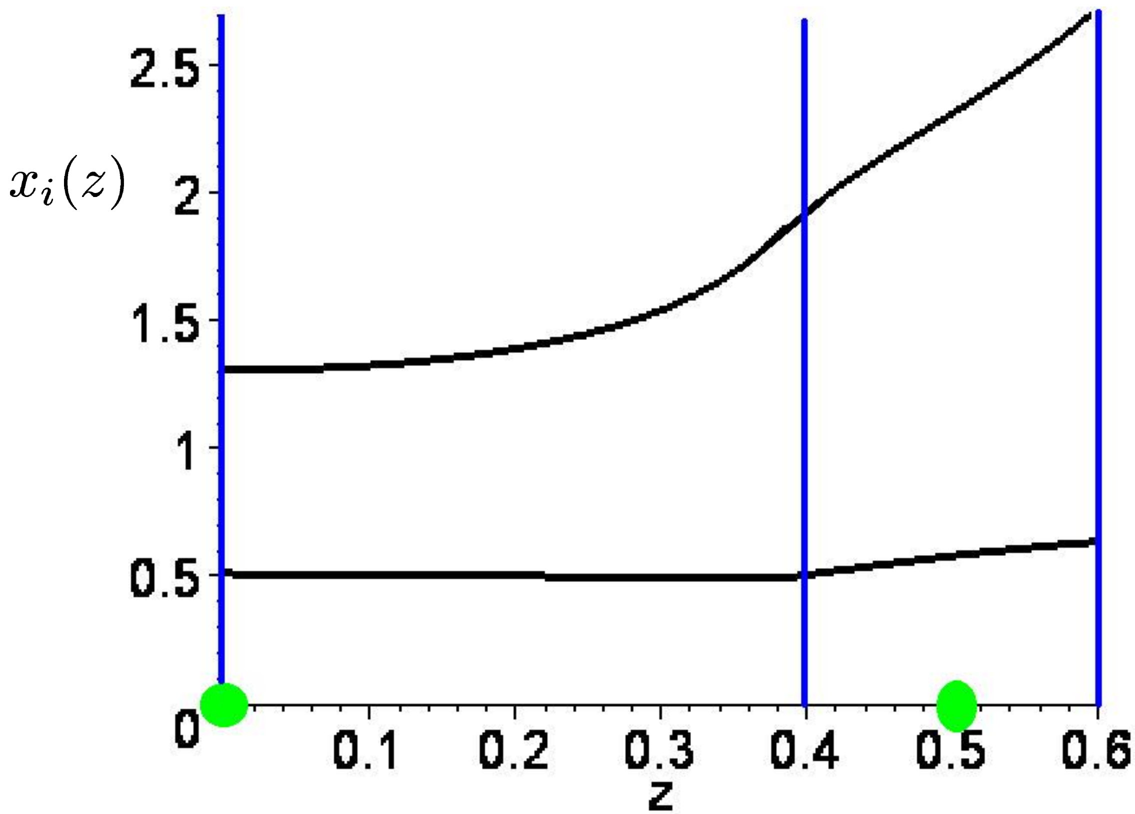
Изучим зависимость максиминного плана $\xi^*(z)$ от параметра z для

$$\Omega=\Omega(z)=\{\beta\ :\ (1-z)c_i\le\beta_i\le(1+z)c_i,\ i=1,\dots,m\}.$$

Функциональный подход для максиминно-эффективных планов

$$\Omega(z) = \{(b,\lambda)\ :\ (1-z) \le b \le (1+z),\ (1-z) \le \lambda \le (1+z)\}$$

- $x_2(z) = 0.5395 - 0.1887z^2 - 0.0263z^4 + \dots$
  $x_3(z) = 3.3560 + 1.8505z^2 + 5.7658z^4 + \dots$
- $x_2(z) = 0.5789 + 0.6649(z-1/2) - 0.9839(z-1/2)^2 + \dots$
  $x_3(z) = 2.3109 + 3.5827(z-1/2) + 2.3505(z-1/2)^2 + \dots$



Заключение

- Для полиномиальной модели с одинаковыми корреляциями доказана независимость от значения корреляции r
- Для экспоненциальной модели
  - изучены двух-, трех- и четырех-точечные локальные Д-оптимальные планы,
  - на основе функционального подхода построены разложения трех- и четырех-точечных локальных Д-оптимальных планов и максиминно-эффективных планов, с помощью этих разложений можно вычислять планы для любых значений параметров,
  - планы в равноотстоящих точках имеют умеренную эффективность и могут быть использованы на начальных стадиях эксперимента, на последующих стадиях следует использовать максиминно-эффективные планы.