

# Оптимальные планы для оценивания производных в полиномиальной регрессионной модели без свободного члена

Барсуков Егор Вячеславович, гр. 16-Б.04мм

Санкт-Петербургский государственный университет  
Математико-Механический факультет  
Кафедра статистического моделирования

Научный руководитель: д.ф.-м.н., профессор В. Б. Мелас

Санкт-Петербург  
2020 г.

$$y_j = \theta^\top f_j(x_j) + \varepsilon_j, \quad j = 1 \dots N, x_j \in \mathcal{X}$$

- $N$  — количество экспериментов;
- $f(x)$  — вектор регрессионных функций;
- $\theta = (\theta_1, \dots, \theta_m)^\top$  — неизвестные параметры;
- $x_1, \dots, x_N$  — условия проведения эксперимента;
- $\mathcal{X}$  — множество планирования;
- $\varepsilon_1, \dots, \varepsilon_N$  — случайные величины, характеризующие ошибки наблюдений.
  - Несмещенные  $\mathbb{E}\varepsilon_i = 0$
  - Некоррелированные  $\mathbb{E}\varepsilon_i\varepsilon_j = 0$  для  $i \neq j$
  - Равноточные  $\mathbb{E}\varepsilon_i^2 = \sigma$  для всех  $i$

## Определение

*Точным планом эксперимента называется следующая дискретная вероятностная мера*

$$\xi = \begin{pmatrix} x_1 & \dots & x_m \\ \omega_1 & \dots & \omega_m \end{pmatrix}, \quad x_i \in \mathcal{X},$$

## Определение

*Для плана эксперимента определим его информационную матрицу*

$$M(\xi) = \int_{\mathcal{X}} f(x) f^{\top}(x) \xi(dt).$$

## Определение

*C-оптимальным планом эксперимента для данного вектора  $c$  называется план минимизирующий функцию  $\Phi$*

$$\Phi(\xi) = \begin{cases} c^T M(\xi)^- c, & \text{если } \exists v, \text{ такой, что } c = M(\xi)v \\ +\infty, & \text{иначе} \end{cases},$$

*где  $M(\xi)^-$  — псевдообратная матрица к информационной матрице плана*

- C-оптимальный план минимизирует дисперсию МНК-оценки  $\theta^T c$ ;
- В общем виде задача нахождения таких планов не решена

## Определение

*Если  $c = f'(z) = (f'_1(z), \dots, f'_m(z))^T$  то соответствующий план называется планом для оценки производной в точке  $z$ .*

- Целью работы является описание оптимальных планов для оценки производной в модели  $f(x) = (x, \dots, x^m)$  при носителе  $\mathcal{X} = [0, d]$ .
- Это имеет практический смысл, если существует априорное знание о значении функции в нулевой точке и эксперимент проводится на положительном отрезке.
- Решение задачи отличается от того, которое получается при носителе, границе которого не принадлежит ноль, которое описано в [Dette et al., 2019].
- Разработать алгоритм численно решающий задачу нахождения  $c$  оптимальных планов для произвольного  $c$ .

## Теорема

Допустимый план  $\xi^*$  с носителем  $x_1, \dots, x_m \in \mathcal{X}$  и весами  $\omega_1, \dots, \omega_m$  является  $c$ -оптимальным тогда и только тогда, когда существует  $p \in \mathbb{R}^k$  и константа  $h$  такие, что выполняются следующие условия:

$$\left| p^\top f(x_i) \right| = 1 \quad i = 1..m \leq n \quad (1a)$$

$$\left| p^\top f(x) \right| \leq 1 \quad x \in \mathcal{X} \quad (1b)$$

$$c = h \sum_{i=1}^m \omega_i f(x_i) p^\top f(x_i). \quad (1c)$$

Кроме того

$$h^2 = c^\top M^-(\xi^*) c$$

Для набора точек  $x_1^*, \dots, x_k^*$  определим множество базисных многочленов

$$L_i(z) = \frac{z \prod_{l \neq i} (z - x_l^*)}{x_i^* \prod_{l \neq i} (x_i^* - x_l^*)}, \quad i = 1, \dots, k$$

## Теорема

*Оптимальный план для оценивания производной полиномиальной модели без свободного члена с опорными точками  $x_1^*, \dots, x_m^*$ , где  $m = n$  или  $m = n - 1$  имеет веса вычисленные по следующей формуле:*

$$\omega_i = \frac{|L'_i(z)|}{\sum_{j=1}^m |L'_j(z)|},$$

С помощью теоремы Элвинга было доказано, что если оптимальный план для оценивания производной с носителем  $[0, 1]$  состоит из  $n$  точек, то носитель состоит из экстремальных точек следующего многочлена

$$S_n(x) = T_n \left( x \left( 1 + \cos \frac{\pi}{2n} \right) - \cos \frac{\pi}{2n} \right),$$

где  $T_n$  — многочлен Чебышёва первого рода степени  $n$ . Таким образом носитель плана находится в точках

$$x_i^* = \frac{\cos \frac{(n-i)\pi}{n} + \cos \frac{\pi}{2n}}{1 + \cos \frac{\pi}{2n}}, \quad i = 1, \dots, n$$



- $\{L_i(z)\}_{i=1}^n$  — базисные многочлены построенные по точкам  $x_1^*, \dots, x_n^*$ ;
- $u_1^i, \dots, u_{n-1}^i$  — корни производной  $i$ -го базисного многочлена, упорядоченные по возрастанию

Было доказано, что тогда корни производных базисных многочленов упорядочены следующим образом

$$u_1^n < u_1^{n-1} < \dots < u_1^1 < u_2^n < u_2^{n-1} < \dots < u_{n-1}^1.$$

# Оптимальный план на отрезке с началом в нуле для оценки производной

## Теорема

*План с носителем  $\{x_i^*\}_{i=1}^n$  является оптимальным планом для оценивания производной полиномиальной модели без свободного члена в точке  $z$  при  $\mathcal{X} = [0, 1]$  тогда и только тогда, когда выполняется одно из следующих условий:*

- $z \in (-\infty, u_1^n)$
- $z \in (u_i^1, u_{i+1}^n), i = 1, \dots, n - 2$
- $z \in (u_{n-1}^1, +\infty)$

Были описаны оптимальные планы размера  $n$  для нахождения производной в полиномиальной модели без свободного члена с областью моделирования на положительном отрезке и показано, в каких точках  $z$  их не существует.

- H. Dette, V. Melas, and Pepelyshev A. Optimal designs for estimating the slope of a regression. *Statistics*, 44(6):617–628, 2010.
- H. Dette, V. Melas, and P. Shpilev. Some explicit solutions ofc-optimal design problems for polynomial regression with no intercept. *Annals of the Institute of Statistical Mathematics*, 2019.