Построение оптимальных и локально оптимальных планов для регрессионных моделей

Корчажников Федор Васильевич, 422 гр.

Санкт-Петербургский государственный университет
Прикладная математика и информатика
Вычислительная стохастика и статистические модели

Научный руководитель: д.ф.-м.н., профессор Мелас В.Б.
Рецензент: к.ф.-м.н., доцент Шпилев П.В.

Санкт-Петербург
2016г.

Модель эксперимента с параметром $\theta \in \mathbb{R}^m$:

$$y_i = \eta(t_i, \theta) + \epsilon_i, \;\; i = 1, \ldots, N,$$

- y_i — результат эксперимента
- $\eta(t_i, \theta)$ — функция регрессии,
- $t_1, \ldots, t_N$ — условия проведения эксперимента,
- $\epsilon_1, \ldots, \epsilon_N$ — ошибки наблюдений, $\mathbb{E}\epsilon_i = 0, \mathbb{E}\epsilon_i\epsilon_j = 0$, при $i \neq j$.

План эксперимента — вероятностная мера:

$$\xi = \begin{pmatrix} t_1 & \ldots & t_n \\ \omega_1 & \ldots & \omega_n \end{pmatrix}, \;\; t_i \in \mathfrak{X}, \;\; i = 1, 2, \ldots, n,$$

где $\mathfrak{X}$ — множество планирования эксперимента,
$\omega_i > 0, \;\; \sum_{i=1}^{n} \omega = 1.$

Целью эксперимента, как правило, является:

- нахождение оценок параметров $\theta_1, \ldots, \theta_m$
- оценок функций от этих параметров
- проверка гипотез относительно их значений

Различные **критерии оптимальности** соответствуют получению наилучших (в некотором смысле) оценок параметра.

Задачи:

- Построить L-оптимальный план для полиномиальной модели (порядка 3 и 4)
- Обобщить алгебраический подход (Мелас, 1999) к построению локально D-оптимальных планов для дробно-рациональных моделей на случай малых промежутков

Информационная матрица плана:

$$M(\xi) = \int_{\mathfrak{X}} f(t) f^T(t) \xi(dt),$$

где $f = (f_1, \ldots, f_m)$, $f_i(t) = \frac{\partial \eta(t,\theta)}{\partial \theta_i}$.

Дисперсионная матрица:

$$D(\xi) = M^{-1}(\xi).$$

Критерий **L-оптимальности**:

$$trLD(\xi) \to \inf_{\xi \in \Xi_H},$$

- L — фиксированная неотрицательно определенная матрица
- Ξ_H — множество невырожденных планов

Рассмотрим **полиномиальную модель**:

$$\eta(x, \theta) = \theta_1 x^{m-1} + \dots + \theta_{m-1} x + \theta_m,$$

На отрезке $\mathfrak{X} = [-1, 1]$ находим L-оптимальные планы:

- $m = 3$, квадратичная модель:

$$\xi_2^* = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}$$

- $m = 4$, кубическая модель:

$$\xi_3^* = \begin{pmatrix} -1 & -0.44 & 0.44 & 1 \\ 0.146 & 0.354 & 0.354 & 0.146 \end{pmatrix}$$

Схема поиска оптимального плана:

- определяем класс возможных оптимальных планов
- получаем выражение для $trLD(\xi)$ (рассматриваем случай $L = I$)
- находим неизвестные x и ω из необходимых условий экстремума
- проверяем результат с помощью теоремы эквивалентности для L-критерия (Ермаков, Жиглявский, 1987)

Критерий **D-оптимальности**:

$$\log \det M(\xi) \to \sup_{\xi \in \Xi_H}.$$

Дробно-рациональная модель с четырьмя параметрами:

$$\eta(x, \theta) = \frac{\theta_1}{x + \theta_2} + \frac{\theta_3}{x + \theta_4}, \quad \theta_2, \theta_4 > 0.$$

Ищем локально D-оптимальный план на промежутке $[0, \infty)$.

Утверждение

Для дробно-рациональной модели локально D-оптимальный план сосредоточен с равными весами в четырех точках и имеет вид:

$$\xi = \begin{pmatrix} 0 & x_2^* & x_3^* & x_4^* \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Критерий **D-оптимальности**:

$$\log \det M(\xi) \to \sup_{\xi \in \Xi_H}.$$

Дробно-рациональная модель с четырьмя параметрами:

$$\eta(x,\theta) = \frac{\theta_1}{x+\theta_2} + \frac{\theta_3}{x+\theta_4}, \quad \theta_2, \theta_4 > 0.$$

Ищем локально D-оптимальный план на промежутке $[0, \infty)$.

Утверждение

Для дробно-рациональной модели локально D-оптимальный план сосредоточен с равными весами в четырех точках и имеет вид:

$$\xi = \begin{pmatrix} 0 & x_2^* & x_3^* & x_4^* \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Максимизируем $\log\det M(\xi)$. Следуя известному подходу (Мелас, 1999), получаем **дифференциальное уравнение**:

$$\psi''(x)xQ(x) + 2\psi'(x)[Q(x) - xQ'(x)] = \lambda(x)\psi(x),$$

$$\psi(x) = (x - x_2^*)(x - x_3^*)(x - x_4^*),\ Q(x) = (x + \theta_2)(x + \theta_4),\ \lambda(x) = \lambda_0 x + \lambda_1,$$

которое может быть представлено как **алгебраическое**:

$$\varphi^{\mathrm{T}}(x)A\psi = \varphi^{\mathrm{T}}(x)C_\lambda\psi,$$

где $\varphi(x) = (x^4, x^3, x^2, x, 1)$, $\psi = (1, \psi_1, \psi_2)$ — коэффициенты разложения $\psi(x)$, A и C_λ — матрицы порядка 5×4, Максимизируем $\log\det M(\xi)$. Следуя известному подходу (Мелас, 1999), получаем **дифференциальное уравнение**:

$$\psi''(x)xQ(x) + 2\psi'(x)[Q(x) - xQ'(x)] = \lambda(x)\psi(x),$$

$$\psi(x) = (x - x_2^*)(x - x_3^*)(x - x_4^*),\; Q(x) = (x + \theta_2)(x + \theta_4),\; \lambda(x) = \lambda_0 x + \lambda_1,$$

которое может быть представлено как **алгебраическое**:

$$\varphi^{\mathrm{T}}(x)A\psi = \varphi^{\mathrm{T}}(x)C_\lambda\psi,$$

где $\varphi(x) = (x^4, x^3, x^2, x, 1)$, $\psi = (1, \psi_1, \psi_2)$ — коэффициенты разложения $\psi(x)$, A и C_λ — матрицы порядка 5×4, Для $\mathfrak{X} = [0, \infty)$ получена явная формула (Мелас, 1999):

$$x^*_{2,4} = -\frac{1}{2}\left(1 + \frac{\lambda_1}{2} \pm \sqrt{(1 + \frac{\lambda_1}{2})^2 - 4}\right), \quad x^*_3 = 1,$$

$$\lambda_1 = -(a+3) - \sqrt{(a+3)^2 + 24}, \quad a = \theta_2 + \theta_4.$$

Промежуток планирования $\mathfrak{X} = [0, d]$:

- $d \geqslant x^*_4$ — **большие** промежутки,
- $d < x^*_4$ — **малые.**

В случае малых промежутков найдем численное решение, обобщив предложенный подход.

Случай **малых** промежутков:

$$\xi = \begin{pmatrix} 0 & x_2 & x_3 & d \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Дифференциальное уравнение:

$$x(d-x)Q(x) + \psi\prime(x)[Q(x)(d-2x) - 2x(d-x)Q\prime(x)] = \lambda(x)\psi(x),$$

- $Q(x) = (x+\theta_2)(x+\theta_4)$,
- $\psi(x) = (x - x_2^*)(x - x_3^*)$,
- $\lambda(x) = \sum_{i=0}^{2} \lambda_i x^i$.

Алгебраическое уравнение:

$$(A - \lambda_0 E_0 - \lambda_1 E_1 - \lambda_2 E_2)\psi = 0, \quad \lambda_0 = 3,$$

- A, E_i — матрицы 5×3,
- E_i — единицы по диагонали, начиная с i строки.

Введем матрицу $B = A - \sum_{i=0}^{2} \lambda_i E_i$.

$B_{(i)}$ — матрица, составленная из $(i+1), (i+2), (i+3)$-ей строк матрицы B.

$T_i = T_i(\lambda) = \det B_{(i)}(\lambda)$.

Найдем корни алгебраического уравнения, применив теорему:

Теорема (Мелас, 1999)

Существует и притом единственное решение задачи

$$\sum_{i=1}^{2} T_i^2(\lambda) \to \min_{\lambda \in \mathbb{R}^2}$$

такое, что точки локально D-оптимального плана, не совпадающие с 0 и d, являются корнями многочлена

$$\psi(x) = \psi_0 x^2 + \psi_1 x + \psi_2,$$

где коэффициенты ψ_0, ψ_1, ψ_2 вычисляются по рекуррентным формулам

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

Алгоритм поиска:

- находим $\lambda^* \in \mathbb{R}^2$, на котором достигается минимум
- при известном λ^* находим ψ_0, ψ_1, ψ_2
- находим корни квадратного трехчлена $\psi(x)$ — опорные точки оптимального плана, не совпадающие с 0 и d

Минимизация по $\lambda = (\lambda_1, \lambda_2)$:

- ограничения на λ_1 и λ_2,
- начальное приближение λ^0.

Алгоритм поиска:

- находим $\lambda^* \in \mathbb{R}^2$, на котором достигается минимум
- при известном λ^* находим ψ_0, ψ_1, ψ_2
- находим корни квадратного трехчлена $\psi(x)$ — опорные точки оптимального плана, не совпадающие с 0 и d

Минимизация по $\lambda = (\lambda_1, \lambda_2)$:

- ограничения на λ_1 и λ_2,
- начальное приближение λ^0.

Ограничения на $\lambda = (\lambda_1, \lambda_2)$:

$$-a - 5d < \lambda_1 < -a - 3d,$$
$$-10ad - 38d^2 - 10 < \lambda_2 < 2ad + 32d^2 - 10,$$

где $a = \theta_2 + \theta_4$.

Начальное приближение λ^0:

- находим решение для большого промежутка $(d^* > d)$
- по опорным точкам находим $\psi_0^*, \psi_1^*, \psi_2^*$
- по ψ_i^* восстанавливаем начальное приближение λ^0

Пример работы алгоритма:

- $\theta_1 = \theta_3 = 1, \quad \theta_2 = \frac{1}{5}, \quad \theta_4 = 5, \quad \mathfrak{X} = [0, 7],$

  локально D-оптимальный план

$$\xi_1^* = \begin{pmatrix} 0 & 0.128 & 0.979 & 7 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix},$$

- $\theta_1 = \theta_3 = 1, \quad \theta_2 = \frac{1}{10}, \quad \theta_4 = 10, \quad \mathfrak{X} = [0, 10],$

  локально D-оптимальный план

$$\xi_2^* = \begin{pmatrix} 0 & 0.079 & 0.956 & 10 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Теорема эквивалентности (Кифер, Вольфовиц, 1960):

- $d(x, \xi^*) = f^T(x)D(\xi^*)f(x) <= n$ — число параметров
- равенство достигается в опорных точках плана

Рис.: Пример 1, $\mathfrak{X} = [0, 7]$

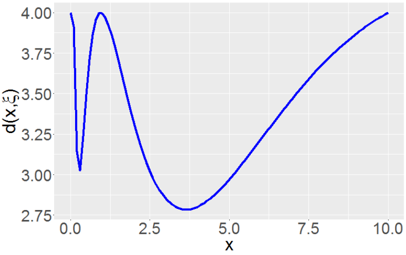

Рис.: Пример 2, $\mathfrak{X} = [0, 10]$

- Построен L-оптимальный план для полиномиальной модели (порядка 3 и 4).
- Восстановлены детали решения (Мелас, 1999) задачи поиска локально D-оптимального плана для дробно-рациональных моделей с четырьмя параметрами и случая больших промежутков.
- Алгебраический подход обобщен на случай малых промежутков, реализован алгоритм построения численного решения на языке R.
- Показан пример работы алгоритма для двух конкретных моделей, найдены локально D-оптимальные планы.

Спасибо за внимание!

- Построен L-оптимальный план для полиномиальной модели (порядка 3 и 4).
- Восстановлены детали решения (Мелас, 1999) задачи поиска локально D-оптимального плана для дробно-рациональных моделей с четырьмя параметрами и случая больших промежутков.
- Алгебраический подход обобщен на случай малых промежутков, реализован алгоритм построения численного решения на языке R.
- Показан пример работы алгоритма для двух конкретных моделей, найдены локально D-оптимальные планы.

Спасибо за внимание!