Численное построение дискриминационных планов эксперимента

Зуйкова Алиса Эдуардовна, гр. 14.Б02-мм

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

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

Санкт-Петербург
2018г.

Введение

Задача — выбор лучшей модели из нескольких конкурирующих.

Используется T-критерий оптимальности [Atkinson, Fedorov, 1975] для таких дискриминирующего эксперимента.

T-оптимальные планы исследовались в работах:

- The non-uniqueness of some designs for discriminating between two polynomial models in one variable [Atkinson, 2010];
- T-optimal designs for discrimination between two polynomial models [Dette, Melas, Shpilev, 2012];
- T-optimal discriminating designs for fourier regression models [Dette, Melas, Shpilev, 2017] (рассмотрены не все случаи).

Постановка задачи

Пусть результаты эксперимента описываются стандартным уравнением регрессии:

$$y = \eta(x, \theta) + \varepsilon, \ \theta \in \Theta, \ x \in X.$$

План эксперимента — дискретная вероятностная мера:

$$\xi = \begin{pmatrix} x_1, \ldots, x_n \\ \omega_1, \ldots, \omega_n \end{pmatrix},$$

где

- n — число точек плана;
- x_i — точка, в которой проводится эксперимент;
- ω_i — доля проводимых экспериментов в точке.

Тригонометрические модели

Имеются конкурирующие модели $\eta_1(x,\theta_1)$ и $\eta_2(x,\theta_2)$, представленные тригонометрическими многочленами:

$$\eta_1(x,\theta_1) = \bar{q}_0 + \sum_{i=1}^{k_1} \bar{q}_{2i-1}\sin(ix) + \sum_{i=1}^{k_2} \bar{q}_{2i}\cos(ix),$$

$$\eta_2(x,\theta_2) = \tilde{q}_0 + \sum_{i=1}^{k_1} \tilde{q}_{2i-1}\sin(ix) + \sum_{i=1}^{k_2} \tilde{q}_{2i}\cos(ix) + \\ + \sum_{i=k_1+1}^{m} b_{2(i-k_1)-1}\sin(ix) + \sum_{i=k_2+1}^{m} b_{2(i-k_1)}\cos(ix),$$

где векторы параметров:

$$\theta_1 = (\bar{q}_0, \bar{q}_2, \ldots, \bar{q}_{2k_2}, \bar{q}_1, \ldots, \bar{q}_{2k_1-1})^{\mathrm{T}},$$

$$\theta_2 = (\tilde{q}_0, \tilde{q}_2, \ldots, \tilde{q}_{2k_2}, \tilde{q}_1, \ldots, \tilde{q}_{2k_1-1}, b_2, \ldots, b_{2m}, b_1, \ldots, b_{2m-1})^{\mathrm{T}}.$$

Определение T-оптимальности

Обозначим разность $\eta_2(x,\theta_2) - \eta_1(x,\theta_1)$ за $\bar{\eta}$. Тогда

$$\bar{\eta}(x,q,\bar{b}) = q_0 + \sum_{i=1}^{k_1} q_{2i-1}\sin(ix) + \sum_{i=1}^{k_2} q_{2i}\cos(ix) +$$

$$\sum_{i=k_1+1}^{m} b_{2(i-k_1)-1}\sin(ix) + \sum_{i=k_2+1}^{m} b_{2(i-k_1)}\cos(ix),$$

где $q_i = \tilde{q}_i - \bar{q}_i$, $\bar{b} = (b_1, b_2, \dots, b_{2(m-k_1)-1}, \dots, b_{2(m-k_2)})^T$.

Определение

T-критерий оптимальности:

$$T(\xi,\bar{b}) = \min_q \int\limits_X \bar{\eta}(x,q,\bar{b})^2 \xi(dx).$$

T-оптимальный план:

$$\xi^* = \operatorname*{argmax}_{\xi} T(\xi,\bar{b}).$$

Цели и задачи

Цель данной работы — исследование T-оптимальных планов для случаев, не рассмотренных в статье [Dette, Melas, Shpilev, 2017]:

1. $k_1 = m-1,\ k_2 = m-2$ при $b_1 \neq 0,\ b_2 = 0$ при четном m,
2. $k_1 = m-2,\ k_2 = m-1$, при $b_1 = 0,\ b_2 \neq 0$, при нечетном m,
3. $k_1 = m-2,\ k_2 = m-1$, при $b_1 \neq 0,\ b_2 = 0$ при $m \geq 2$.

Задачи:

- Численное построение планов;
- Исследование зависимостей планов от параметров;
- Исследование зависимости числа точек плана от параметров;
- Поиск явного вида планов.

Теорема эквивалентности

Теорема [Dette, Titoff, 2009]

Для фиксированного $\bar{b}$ *эквивалентны следующие условия:*

$$\xi^* = \begin{pmatrix} x_1^*, \ldots, x_n^* \\ \omega_1, \ldots, \omega_n \end{pmatrix}, \; x_i^* \in [0, 2\pi], \; i = 1, \ldots, n.$$

(1) ξ^* *– T-оптимальный план для моделей* η_1 *и* η_2*.*
(2) Существует вектор q^* *и константа* $h > 0$ *такие, что функция* $\psi^*(x) = \bar{\eta}(x, q^*, \bar{b})$ *удовлетворяет условиям:*

- $|\psi^*(x)| \le h, \; x \in [0, 2\pi]$;
- $|\psi^*(x_i^*)| = h, \; i = 1, \ldots, n$;
- $\sum_{i=1}^n \psi^*(x_i^*) \cos(j x_i^*) \omega_i = 0, \; j = 0, \ldots, k_2$;
- $\sum_{i=1}^n \psi^*(x_i^*) \sin(j x_i^*) \omega_i = 0, \; j = 1, \ldots, k_1$.

Необходимое условие

Из теоремы эквивалентности следует система необходимых условий Т-оптимальности плана:

$$|\psi(x_i^*)| = |\psi(x_{i+1}^*)|, \; i = 1, \dots, n-1;$$
$$\psi(x_i^*)' = 0, \; i = 1, \dots, n;$$
$$\sum_{i=1}^{n} \omega_i = 1;$$
$$\sum_{i=1}^{n} \psi(x_i^*) \cos(j x_i^*) \omega_i = 0, \; j = 0, \dots, k_2;$$
$$\sum_{i=1}^{n} \psi(x_i^*) \sin(j x_i^*) \omega_i = 0, \; j = 1, \dots, k_1;$$

где x_i^* - точки плана, ω_i - веса плана.

$k_1 = m-1,\ k_2 = m-2,\ b_1 \neq 0,\ b_2 = 0,\ m$ четное

Будут рассмотрены случаи $m = 2$ и $m = 4$.

В первом случае вид модели:
$\bar{\eta}(x, q, \bar{b}) = q_0 + q_1 \sin(x) + \cos(x) + b \sin(2x)$.

$n = 2$ [Dette, Melas, Shpilev, 2016]. Алгоритм поиска $x_i,\ \omega_i,\ q_i$:

1. Решение системы необходимых условий с помощью fsolve;
2. Проверка на оптимальность по теореме эквивалентности по графику.

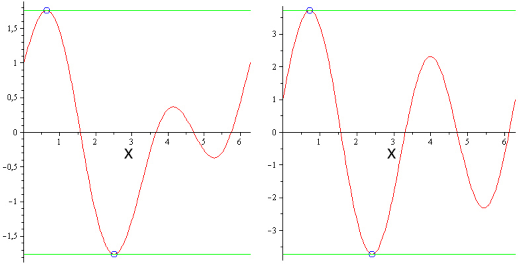
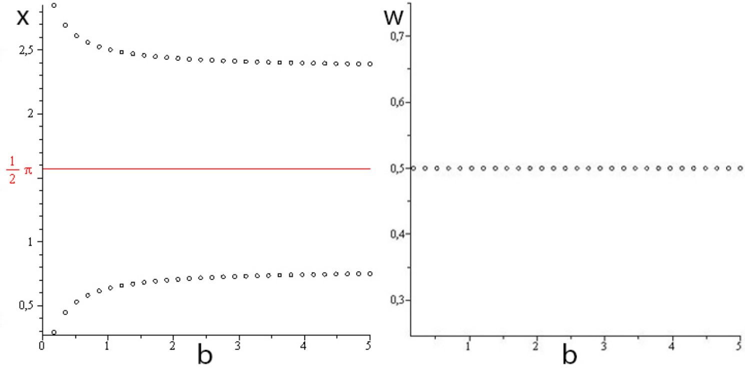


Рис.: Функция ψ^* при $b = 1$ (слева) и $b = 3$ (справа).

$k_1 = m-1,\ k_2 = m-2,\ b_1 \neq 0,\ b_2 = 0,\ m = 2$

Была численно построена зависимость точек (слева) и весов (справа) T-оптимального плана от параметра b.



На основе графиков выдвинуто предположение: $x_1 = \pi - x_2$, $\omega_1 = \omega_2 = \frac{1}{2}$.

$k_1 = m - 1,\ k_2 = m - 2,\ b_1 \neq 0,\ b_2 = 0,\ m = 2$

Утверждение о явном виде плана

План

$$\xi^* = \begin{pmatrix} x^* & \pi - x^* \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},\ x^* = \arcsin\left(\frac{-1 + \sqrt{32b^2 + 1}}{8b}\right).$$

является T-оптимальным.

Доказательство проводится по теореме эквивалентности.

$$\bar{\eta}(x, q, \bar{b}) = q_0 + q_1 \sin(x) + q_2 \cos(x) + q_3 \sin(2x) + \\ + q_4 \cos(2x) + q_5 \sin(3x) + \cos(3x) + b \cos(4x).$$

Идея — построение T-оптимального плана путем поиска глобального минимума функции

$$(\psi(x_1) + \psi(x_2))^2 + \cdots + (\psi(x_{n-1}) + \psi(x_n))^2 + \sum_{i=1}^{n} (\psi'(x_i))^2 + \\ + (\sum_{i=1}^{n} \omega_i - 1)^2 + \sum_{j=0}^{m-2} (\sum_{i=1}^{n} \psi(x_i) \cos(j x_i) \omega_i)^2 + \sum_{j=1}^{m-1} (\sum_{i=1}^{n} \psi(x_i) \sin(j x_i) \omega_i)^2,$$

где функция $\psi(x) = \bar{\eta}(x, q, \bar{b})$, с помощью DirectSearch.

Возникает проблема поиска начальной точки приближения. Построен алгоритм для ее нахождения.

$k_1 = m - 1,\ k_2 = m - 2,\ b_1 \neq 0,\ b_2 = 0,\ m = 4$

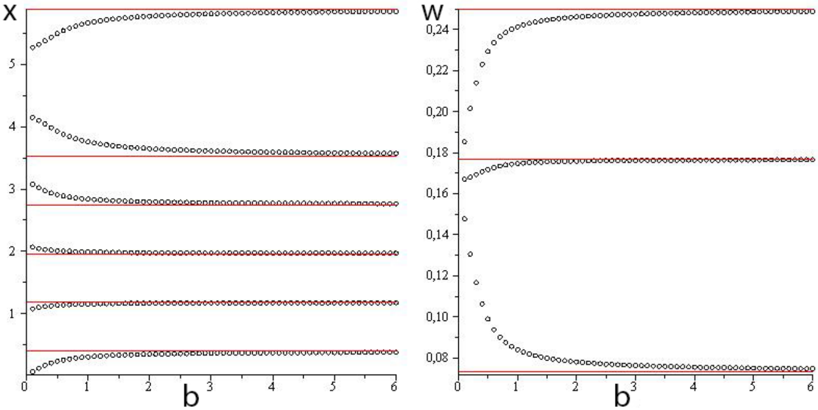


Рис.: Графики зависимости точек (слева) и весов (справа) плана от параметра b. Красными линиями показаны значения точек и весов предельного плана при $b \to \infty$.

$k_1 = m - 1,\ k_2 = m - 2,\ b_1 \neq 0,\ b_2 = 0,\ m = 4$

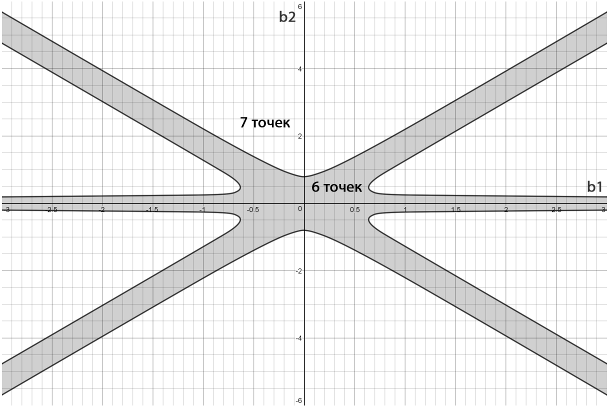


Рис.: График зависимости количества точек плана от параметров b_1 и b_2.

$$\bar{\eta}(x, q, \bar{b}) = q_0 + q_1 \sin(x) + q_2 \cos(x) + q_3 \cos(2x) + \sin(2x) + b\cos(3x).$$

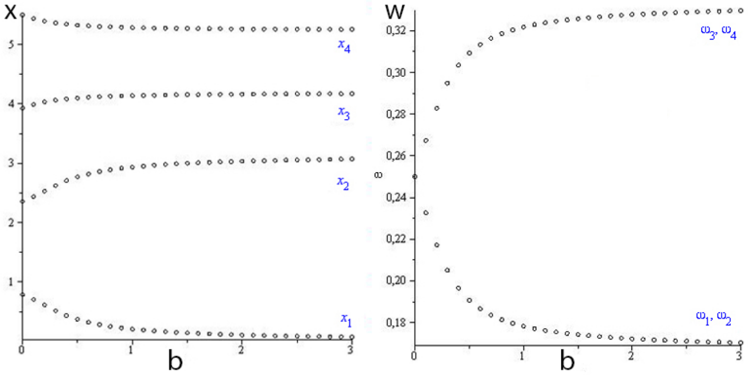


Рис.: Зависимость точек (слева) и весов (справа) T-оптимального плана от параметра b.

$k_1 = m - 2,\ k_2 = m - 1,\ b_1 = 0,\ b_2 \neq 0,\ m = 3$

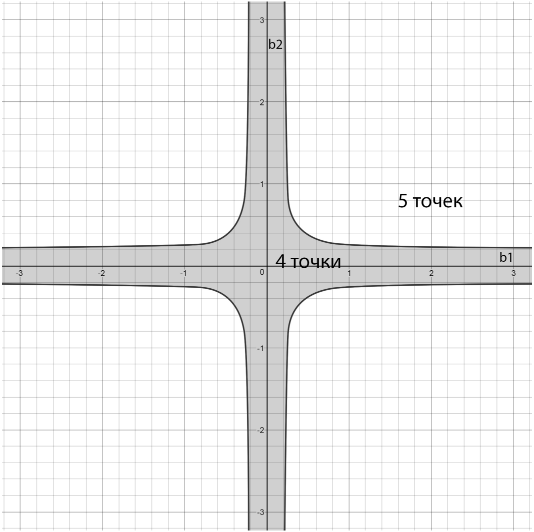


Рис.: График зависимости количества точек плана от параметров b_1 и b_2.

$$\bar{\eta}(x, q, \bar{b}) = q_0 + q_1 \sin(x) + q_2 \cos(x) + q_3 \sin(2x) + \\ + q_4 \cos(2x) + q_5 \cos(3x) + \sin(3x) + b \sin(4x).$$

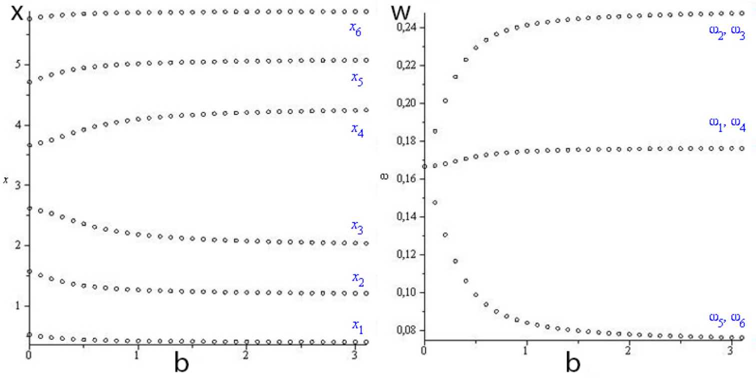


Рис.: Зависимость точек (слева) и весов (справа) T-оптимального плана от параметра b.

Итоги

- Найдены методы и построены алгоритмы численного построения T-оптимальных планов с помощью Maple.
- Рассмотренны модели:
  1. $m = 2,\ k_1 = 1,\ k_2 = 0,\ b_1 \neq 0,\ b_2 = 0;$
  2. $m = 4,\ k_1 = 3,\ k_2 = 2,\ b_1 \neq 0,\ b_2 = 0;$
  3. $m = 3,\ k_1 = 1,\ k_2 = 2,\ b_1 = 0,\ b_2 \neq 0;$
  4. $m = 4,\ k_1 = 2,\ k_2 = 3,\ b_1 \neq 0,\ b_2 = 0;$

  для каждой из которых:
  - численно построены T-оптимальные планы и изучена зависимость от параметра;
  - границы областей n-1 и n-точечных планов найдены численно.
- Для модели 1 T-оптимальный план найден в явном виде;
- Для модели 2 предельные значения точек и весов плана найдены в явном виде.