Санкт-Петербургский государственный университет

Прикладная математика и информатика

Статистическое моделирование

Кухтина Дарина Александровна

# Сравнение различных подходов к дискриминации моделей

Выпускная квалификационная работа

Научный руководитель:

д. ф.-м. н., профессор В. Б. Мелас

Рецензент:

д. ф.-т. н., профессор Ю. Д. Григорьев

Санкт-Петербург

2017

Saint Petersburg State University

Applied Mathematics and Computer Science

Statistical Modelling

Kukhtina Darina Alexandrovna

# Comparison of different approaches to model discrimination

Graduation Project

Научный руководитель:

Professor V. B. Melas

Рецензент:

Professor U. D. Grigorev

Saint Petersburg

2017

# Оглавление

# Введение

Под дискриминацией моделей подразумевается выбор одной из двух или нескольких альтернативных моделей. Проблема дискриминации регрессионных моделей представляет значительный теоретический и практический интерес. Соответствующая проблема построения планов эксперимента исследуется в научной литературе по теории планирования регрессионных экспериментов уже более пятидесяти лет. В отличие от (более изученной) проблемы по планированию эксперимента для оценивания параметров, модель не предполагается фиксированной. В пионерской работе ([1]) была изучена задача построения оптимального плана для выбора одной из двух полиномиальных моделей, одна из которых имеет известную степень, а степень другой на единицу больше. В качестве критерия оптимальности плана Стиглер предложил рассматривать величину дисперсии старшего коэффициента второй модели. Аткинсон и Федоров ([2]) предложили более общий критерий, который получил название критерия T-оптимальности. Этот критерий предполагает аппроксимацию сложной модели более простой, причем параметры более сложной модели считаются известными. В терминах теории проверки гипотез это означает, что при нулевой гипотезе модель полностью известна. Обзор работ в этом направлении можно найти в статье ([3]). Слабой стороной этого подхода является зависимость плана от информации о параметрах основной модели. В ряде работ ([4], [5]) было показано, что робастные версии T-критерия в случае вложенных моделей приводят к задаче оценивания старших коэффициентов основной модели, что перекликается с подходом Стиглера. Кроме того, Стиглер в той же работе поставил задачу построения планов, которые позволяют не только дискриминировать модели, но и (без дополнительных экспериментов) оценивать параметры модели, выбранной в результате процедуры дискриминации. Настоящая работа посвящена исследованию и развитию этого подхода для случая вложенных полиномиальных моделей.

Был предложен и численно исследован подход, обобщающий подход Стиглера к построению планов для дискриминации полиномиальных регрессионных моделей. Была выбрана более удобная, чем у Стиглера, параметризация. Рассмотрен и исследован случай моделей до четвертой степени, включительно. Установлено, что развиваемый подход позволяет эффективно проверить гипотезу о том, что верна модель $n$-ой степени при альтернативе в виде модели $n-1$-ой степени, причем при любом решении удается также достаточно эффективно оценить параметры выбранной модели.

# Глава 1

# Постановка задачи

Пусть наши данные описываются следующим уравнением:

$$y_j = \eta(x_j, \theta) + \varepsilon_j$$

,

где $j = 1, \ldots, N$, $x_j \in \chi$, $\theta \in \Omega \subset \mathrm{R}^m$. Здесь $\chi$ множество планирования, наделенное структурой компактного топологического пространства, $x_j$ — точки, в которых проводятся измерения, $\theta$ — вектор параметров, а $\varepsilon_j$ — ошибки измерений — случайные величины, которые некоррелированны, с нулевым математическим ожиданием и одинаковой дисперсией, а $\Omega$ — некоторое открытое множество . Функция $\eta(x_j, \theta)$ – регрессионная модель.

Введем формальное определение плана, и некоторые вспомогательные определения.

**Определение 1.** *План эксперимента — это дискретная вероятностная мера на множестве планирования $\chi$ с конечным носителем:*

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

*причем $\sum_{i=1}^{n} \omega_i = 1$ и $\omega_i \geq 0$ — весовые коэффициенты, $n$ — число различных точек плана, $x_i \in \chi$, где $\chi$ множество планирования, наделенное структурой компактного топологического пространства.*

В теории планирования используется предположение о линейности параметризации $\eta(x, \theta) = \theta^T f(x)$, где $f(x) = (f_1(x), \ldots, f_m(x))^T$, где $f_i(x)$ — заданные базисные функции, $i = 0, \ldots, n$.

**Определение 2.** *Информационная матрица плана — $M(\xi) = \int\limits_{\chi} f(x) f^T(x) \xi dx$.*

Отметим, что $y_i$ — результаты эксперимента, а $x_i$ (точки плана эксперимента) можно выбирать, исходя из задачи. В теории планирования эксперимента существуют следующие задачи:

1. Задача выбора одной из двух или нескольких конкурирующих моделей

2. Задача оценивания параметров выбранной модели

Для решения этих задач подходят различные критерии оптимальности. В частности, для решения второй задачи хорошо подходит критерий $D$-оптимальности:

**Определение 3.** *План называется $D$-оптимальным, если он максимизирует величину определителя информационной матрицы* $\det M(\xi) \to \max_{\xi \in \Xi}$.

В моей бакалаврской работе было выяснено, что для решения и первой, и второй задачи для полиномиальной модели хорошо подходят $D_s$-оптимальные планы. Введем их определение.

Разобьем множество параметров $\theta$ на 2 непересекающихся множества, где $\dim \theta_{(1)} = s$:

$$\theta^{\mathrm{T}} f(x) = \theta_{(1)}^{\mathrm{T}} f_{(1)}(x) + \theta_{(2)}^{\mathrm{T}} f_{(2)}(x).$$

**Определение 4.** *План, максимизирующий величину* $\det M_{(s)}(\xi)$ *называется $D_s$-оптимальным планом, где $M_{(s)}(\xi) = M_{11}(\xi) - X^{\mathrm{T}} M_{22}(\xi) X$. Причем, если матрица $X$ невырождена, то $X = M_{22}^{-1} M_{12}$; если матрица $X$ вырождена, то $X$ — произвольное решение системы $M_{22} X = M_{12}$, причем $M_{(s)}$ не зависит от выбора решения.*

Также и для первой, и для второй задачи рассматривается подход Стиглера, который будет описан позже.

## 1.1. Вспомогательные результаты, используемые в работе

Также введем понятие эффективности. Эффективность плана $\xi_1$, относительно $\xi_2$ — это величина, которая показывает во сколько раз нужно увеличить число экспериментов, если оценивать параметры, оптимально оцениваемые планом $\xi_1$, с помощью плана $\xi_2$. Если было проведено $n$ экспериментов, то информационная матрица имеет вид: $M = n \cdot M(\xi)$, где $n$ — количество экспериментов. Из линейной алгебры известно, что, если матрица $A$ размерности $n \times n$, то $\det mA = m^n \det A$, где $m = const$.

**Определение 5.** *Эффективность $D$-оптимального плана, относительно $D_s$-оптимального плана:*

$$\frac{\sqrt[s]{\det M_{(s)}(\xi_2)}}{\sqrt[s]{\det M_{(s)}(\xi_1)}},$$

*где $s$ — количество оцениваемых параметров модели, $\xi_2$ — $D$-оптимальный план и $\xi_1$ — $D_s$-оптимальный план.*

*Эффективность $D_s$-оптимального плана, относительно $D$-оптимального:*

$$\frac{\sqrt[n]{\det M(\xi_1)}}{\sqrt[n]{\det M(\xi_2)}},$$

*где $n$ — количество параметров модели, $\xi_2$ — $D$-оптимальный план и $\xi_1$ — $D_s$-оптимальный план.*

Для нахождения $D$-оптимальных и $D_s$-оптимальных планов для полиномиальной модели порядка $n$: $\eta(x,\theta) = \sum_{i=0}^{n} \theta_i x^i$ сформулируем теорему Стаддена:

**Теорема 1.** *([6], 1979г.)*

*Точки $D_s$-оптимального плана $\xi$: $-1, x_1, \ldots, x_{n-1}, 1$, где $x_1, \ldots, x_{n-1}$ корни уравнения $P_s'(x)U_{n-s}(x) - P_{s-1}'U_{n-s-1}(x) = 0$, где $s = 1, \ldots, n-1$, $P_i$ — многочлены Лежандра, $U_i$ — многочлены Чебышева второго рода.*

**Теорема 2.** *( [6], 1979г.)*

*Вес точек $D_s$-оптимального плана, $x_0, x_1, \ldots, x_n$ выражается формулой: $\omega_i = \frac{2}{2n+1+U_{2s}(x_i)}$, $i = 0, 1, \ldots, n$, где $U_{2s}$ — многочлены Чебышева второго рода.*

Также есть теорема, упрощающая численный поиск точек $D$-оптимального плана, что очень полезно в случаях, когда для поиска $D$-оптимального плана модели с помощью теоремы эквивалентности, например, приходится решать уравнения высоких степеней и сложно найти точное решение.

**Теорема 3.** *([7])*

*Для полиномиальной регрессии порядка $n$ на отрезке $[-1, 1]$ непрерывный $D$-оптимальный план существует, единственен и сосредоточен с равными весами в $n+1$ точках, которые являются корнями $(x^2-1)P_n'$, где $P_n(x)$ — полином Лежандра порядка $n$.*

# Глава 2

# Некоторые полиномиальные модели

## 2.1. Полиномиальная модель 5-го порядка

Полиномиальные модели представляют особый интерес, так как ими аппроксимируются многие более сложные модели.
Для полиномиальных моделей ниже 5-ой степени сравнение эффективностей $D$ и $D_s$ оптимальных планов было проведено в бакалаврской работе. Однако эффективности их также приведены (см. таблицу 2.1).

Рассмотрим полиномиальную модель 5 степени, у нее вектор регрессионных функций: $f(x) = (1, x, x^2, x^3, x^4, x^5)$. Тогда по теореме 3 $D$-оптимальный план сосредоточен в точках, которые являются корнями уравнения:
$(x^2 - 1)P_5' = 0$.
Применяя теорему 3, получаем, что $x_1 = -1$, $x_6 = 1$, и $x_2, x_3, x_4, x_5$ находятся из уравнения: $\frac{1}{8}(63x^5 - 70x^3 + 15x)' = 0$.
Отсюда получаем $D$-оптимальный план:

$$\begin{pmatrix} -1 & -0.76 & -0.29 & 0.29 & 0.76 & 1 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}.$$

Определитель информационной матрицы $M(\xi)$: $8.85385 \cdot 10^{-8}$.
$\sqrt[6]{M(\xi)} = 0.66 \cdot 10^{-5}$.
Определители усеченной матрицы:

1. $s = 5$; $4.811692 \cdot 10^{-7}$;
   $\sqrt[5]{M(\xi)} = 0.054$;

2. $s = 4$; $2.84585 \cdot 10^{-5}$;
   $\sqrt[4]{M(\xi)} = 0.072$;

3. $s = 3$; $0.001436888$;
   $\sqrt[3]{M(\xi)} = 0.11$;

4. $s = 2$; $0.1314229$;
   $\sqrt[2]{M(\xi)} = 0.34$;

5. $s = 1$; 0.35.

Теперь рассмотрим усеченные оптимальные планы, для различных $s$:

1. $s = 5$.

   $P_5'(x) \cdot t_1'(x) = 0$.

$$\begin{pmatrix} -1 & -0.76 & -0.29 & 0.29 & 0.76 & 1 \\ 0.15 & 0.16 & 0.16 & 0.16 & 0.16 & 0.15 \end{pmatrix}.$$

2. $s = 4$.

   $P_4'(x) \cdot U_1(x) - P_3'(x) \cdot U_0(x) = 0$. $\frac{1}{8}(35x^4 - 30x^2 + 3)' \cdot 2x - \frac{1}{2}(5x^3 - 3x)' = 0$.

   Решая, получаем:

$$\begin{pmatrix} -1 & -0.75 & -0.28 & 0.28 & 0.75 & 1 \\ 0.1 & 0.176 & 0.197 & 0.197 & 0.176 & 0.1 \end{pmatrix}.$$

   Определитель информационной матрицы: $8.780788 \cdot 10^{-8}$.

   $\sqrt[6]{M(\xi)} = 9.75 \cdot 10^{-2}$.

   Усеченный определитель: $2.820907 \cdot 10^{-5}$.

   $\sqrt[4]{D_4} = 0.72 \cdot 10^{-1}$.

3. $s = 3$.

   $P_3'(x) \cdot U_2(x) - P_2'(x) \cdot U_1(x) = 0$. Решая, получаем:

$$\begin{pmatrix} -1 & -0.75 & -0.3 & 0.3 & 0.75 & 1 \\ 0.111 & 0.208 & 0.173 & 0.173 & 0.208 & 0.111 \end{pmatrix}.$$

   Определитель информационной матрицы: $8.696061 \cdot 10^{-8}$.

   $\sqrt[6]{M(\xi)} = 6.56 \cdot 10^{-2}$.

   Усеченный определитель: 0.001413886. $\sqrt[3]{D_3} = 0.11$.

4. $s = 2$.

   $P_2'(x) \cdot U_3(x) - P_1'(x) \cdot U_2(x) = 0$. Решая, получаем:

$$\begin{pmatrix} -1 & -0.77 & -0.26 & 0.26 & 0.77 & 1 \\ 0.125 & 0.19 & 0.177 & 0.177 & 0.19 & 0.125 \end{pmatrix}.$$

   Определитель информационной матрицы: $8.690479 \cdot 10^{-8}$.

   $\sqrt[6]{M(\xi)} = 6.65 \cdot 10^{-2}$.

Усеченный определитель: 0.1339909.

$\sqrt{0.366}$

5. $s = 1$.

   $P_1'(x) \cdot U_4(x) - P_0'(x) \cdot U_3(x) = 0$. Решая, получаем:

$$\begin{pmatrix} -1 & -0.8 & -0.3 & 0.3 & 0.8 & 1 \\ 0.143 & 0.159 & 0.193 & 0.193 & 0.159 & 0.143 \end{pmatrix}.$$

   Определитель информационной матрицы: $8.327627 \cdot 10^{-8}$.

   $\sqrt[6]{M(\xi)} = 6.65 \cdot 10^{-2}$.

   Усеченный определитель: 0.3691267.

Нашли все необходимые величины, тогда можем, согласно определению 5, посчитать эффективности $D$-оптимальных планов и усеченных оптимальных планов.

Эффективности $D$-оптимального плана, относительно усеченных:

1. Относительно $D_4$ — 0.63;

2. Относительно $D_3$ — 0.47;

3. Относительно $D_2$ — 0.32;

4. Относительно $D_1$ — 0.27.

Эффективности усеченных планов, относительно $D$-оптимального:

1. $D_4$, относительно $D$ — 0.99;

2. $D_3$, относительно $D$ — 0.98;

3. $D_2$, относительно $D$ — 0.98;

4. $D_1$, относительно $D$ — 0.94.

Эти результаты говорят о том, что если использовать $D$-оптимальные планы для оценки части параметров, то количество экспериментов нужно будет увеличить в полтора раза и более (все эффективности меньше 0.63). Однако, если использовать планы, оптимально оценивающие часть параметров модели для оценки всех ее параметров, то число экспериментов увеличится незначительно (все эффективности более 0.94).

## 2.2. Полиномиальная модель 6-го порядка

Рассмотрим полиномиальную модель 6 степени, у нее вектор регрессионных функций: $f(x) = (1, x, x^2, x^3, x^4, x^5, x^6)$. Тогда по теореме 3 $D$-оптимальный план сосредоточен в точках, которые являются корнями уравнения:

$(x^2 - 1)P_6' = 0.$

Применяя теорему, получаем, что $x_1 = -1$, $x_6 = 1$, и $x_2, x_3, x_4, x_5, x_6$ находятся из уравнения: $\frac{1}{16}(231 \cdot 4x^6 - 315 \cdot x^4 + 105 \cdot x^2 - 5)' = \frac{1}{16}(231 \cdot 6 \cdot x^5 - 315 \cdot 4 \cdot x^3 + 105 \cdot 2 \cdot x) = 0.$

Отсюда получаем $D$-оптимальный план:

$$\begin{pmatrix} -1 & -\sqrt{\frac{5-2\sqrt{\frac{5}{3}}}{11}} & -\sqrt{\frac{5+2\sqrt{\frac{5}{3}}}{11}} & 0 & \sqrt{\frac{5-2\sqrt{\frac{5}{3}}}{11}} & \sqrt{\frac{5+2\sqrt{\frac{5}{3}}}{11}} & 1 \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} \end{pmatrix}.$$

Определитель информационной матрицы: $\det M = 4.521281 \cdot 10^{-11}$.

Рассмотрим усеченные определители:

1. $s = 6$; $3.164897 \cdot 10^{-10}$;
2. $s = 5$; $2.120134 \cdot 10^{-8}$;
   $\sqrt[5]{M(\xi)} = 0.1162 \cdot 10^{-4}$.
3. $s = 4$; $5.898116 \cdot 10^{-6}$;
   $\sqrt[4]{M(\xi)} = 0.1558 \cdot 10^{-3}$;
4. $s = 3$; $0.0006238777$; $\sqrt[3]{M(\xi)} = 0.0854$;
5. $s = 2$; $0.1044252$; $\sqrt{M(\xi)} = 0.0323$;
6. $s = 1$; $0.3162799$ .

1. Оценим коэффициент перед $x, x^2, x^3, x^4, x^5, x^6$:

   $s = 6$;

   $P_6'(x)t_1'(x) = 0.$

   Решая, получим план:

   $$\begin{pmatrix} -1 & -\sqrt{\frac{5-2\sqrt{\frac{5}{3}}}{11}} & -\sqrt{\frac{5+2\sqrt{\frac{5}{3}}}{11}} & 0 & \sqrt{\frac{5-2\sqrt{\frac{5}{3}}}{11}} & \sqrt{\frac{5+2\sqrt{\frac{5}{3}}}{11}} & 1 \\ 0.8 & 0.145 & 0.152 & 0.154 & 0.152 & 0.145 & 0.8 \end{pmatrix}.$$

   Определитель информационной матрицы: $\det M = 1.796402 \cdot 10^{-09}$.

2. Оценим коэффициент перед $x^2, x^3, x^4, x^5, x^6$:

   $s = 5$;

   $P_5(x)' \cdot U_1(x) - P_4(x)' U_0(x) = 0$; $\frac{1}{8} \cdot (63 \cdot x^5 - 70 \cdot x^3 + 15 \cdot x)' \cdot 2x - \frac{1}{8} \cdot (35 \cdot x^4 - 30 \cdot x^2 + 3)' = 0$.

   Решая, получаем план:

$$\begin{pmatrix} -1 & -0.83 & -0.46 & 0 & 0.46 & 0.83 & 1 \\ 0.083 & 0.149 & 0.161 & 0.166 & 0.161 & 0.149 & 0.083 \end{pmatrix}.$$

   Определитель информационной матрицы: $\det M = 2.490946 \cdot 10^{-11}$.

   $\sqrt[5]{M(\xi)} = 0.757 \cdot 10^{-5}$.

3. Оценим коэффициент перед $x^3, x^4, x^5, x^6$:

   $s = 4$;

   $P_4(x)' \cdot U_2(x) - P_3(x)' U_1(x) = 0$; $\frac{1}{8} \cdot (35 \cdot x^4 - 30 \cdot x^2 + 3)' \cdot (4x^2 - 1) - \frac{1}{2} \cdot (5x^3 - 3x) \cdot 2x = 0$.

   Решая, получаем план:

$$\begin{pmatrix} -1 & -0.82 & -0.47 & 0 & 0.47 & 0.82 & 1 \\ 0.09 & 0.17 & 0.158 & 0.143 & 0.158 & 0.17 & 0.09 \end{pmatrix}.$$

   Определитель информационной матрицы: $\det M = 3.160438 \cdot 10^{-11}$.

   $\sqrt[4]{M(\xi)} = 1.258 \cdot 10^{-5}$.

4. Оценим коэффициент перед $x^4, x^5, x^6$:

   $s = 3$;

   $P_3(x)' \cdot U_3(x) - P_2(x)' U_2(x) = 0$; $\frac{1}{2} \cdot (5x^3 - 3x) \cdot (8x^3 - 4x) - \frac{1}{2}(3x^2 - 1)'(4x^2 - 1) = 0$.

   Получаем план:

$$\begin{pmatrix} -1 & -0.82 & -0.47 & 0 & 0.47 & 0.82 & 1 \\ 0.1 & 0.175 & 0.142 & 0.167 & 0.142 & 0.175 & 0.1 \end{pmatrix}.$$

   Определитель информационной матрицы: $\det M = 3.807664 \cdot 10^{-11}$.

   $\sqrt[3]{M(\xi)} = 1.258 \cdot 10^{-5}$.

5. Оценим коэффициент перед $x^5, x^6$:

   $s = 2$;

   $P_2(x)' \cdot U_4(x) - P_1(x)' U_3(x) = 0$; $\frac{1}{2}(3x^2 - 1)' \cdot (16x^4 - 12x^2 + 1) - (8x^3 - 4x) = 0$.

   Получаем план:

$$\begin{pmatrix} -1 & -0.84 & -0.45 & 0 & 0.45 & 0.84 & 1 \\ 0.11 & 0.148 & 0.164 & 0.143 & 0.164 & 0.148 & 0.11 \end{pmatrix}.$$

Определитель информационной матрицы: $\det M(\xi) = 3.767664 \cdot 10^{-11}$.

6. Оценим коэффициент перед $x^6$:

   $s = 1$;

   $P_1(x)' \cdot U_5(x) - P_0(x)' U_4(x) = 0$; $32x^5 - 32x^3 + 6x = 0$.

$$\begin{pmatrix} -1 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} & 1 \\ 0.125 & 0.133 & 0.154 & 0.167 & 0.154 & 0.133 & 0.125 \end{pmatrix}.$$

   Определитель информационной матрицы: $\det M(\xi) = 3.5767664 \cdot 10^{-11}$.

Эффективности $D$-оптимального плана, относительно усеченных (согласно определению 5):

1. Относительно $D_1$ — 0.52;
2. Относительно $D_2$ — 0.56;
3. Относительно $D_3$ — 0.63;
4. Относительно $D_4$ — 0.76;
5. Относительно $D_5$ — 0.87;

Эффективности усеченных планов, относительно $D$-оптимального:

1. $D_1$, относительно $D$ — 0.88;
2. $D_2$, относительно $D$ — 0.94;
3. $D_3$, относительно $D$ — 0.98;
4. $D_4$, относительно $D$ — 0.99;
5. $D_5$, относительно $D$ — 0.99;

Исходя из полученных значений эффективности, также можно сказать, что если использовать $D$-оптимальные планы для оценки части параметров, то количество экспериментов увеличится сильнее, чем если использовать усеченные оптимальные планы для оценки всех параметров модели.

Из рисунка 2.1 видно, что при увеличении количества оцениваемых параметров усеченные и $D$-оптимальные планы примерно одинаково эффективны.

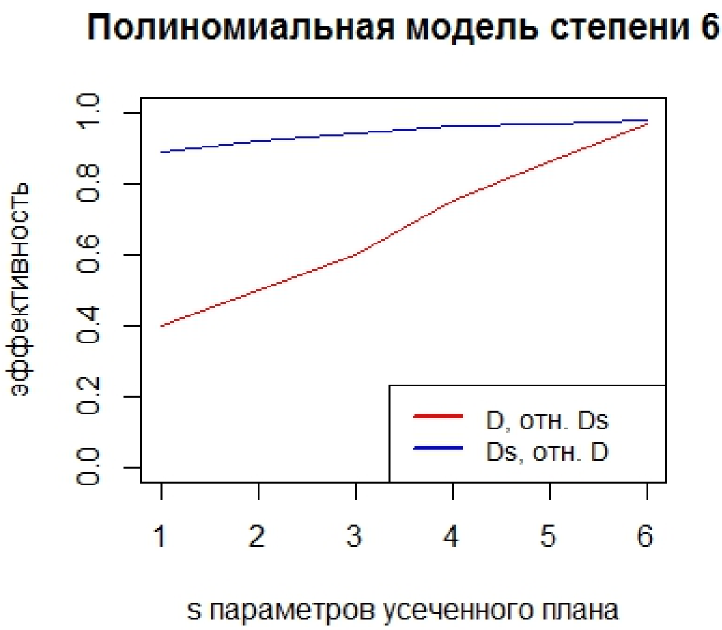


Рис. 2.1. Сравнение эффективности

## 2.3. Полиномиальная модель 7-го порядка

Рассмотрим полиномиальную модель 7 степени, у которой вектор регрессионных функций: $f(x) = (1, x, x^2, x^3, x^4, x^5, x^6, x^7)$. Тогда по теореме 1 $D$-оптимальный план сосредоточен в точках, которые являются корнями уравнения:
$(x^2 - 1)P_7' = 0$.
Применяя теорему, получаем, что $x_1 = -1$, $x_7 = 1$, и $x_2, x_3, x_4, x_5, x_6$ находятся из уравнения $\frac{1}{16}(429 \cdot x^7 - 693 \cdot x^5 + 315 \cdot x^3 - 35x)' = 0$.
Отсюда получаем $D$-оптимальный план:

$$\begin{pmatrix} -1 & -0.87 & -0.59 & -0.2 & 0.2 & 0.59 & 0.87 & 1 \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}.$$

Определитель информационной матрицы: $5.727722 \cdot 10^{-15}$.
Рассмотрим усеченные планы:

1. $s = 7$; $3.942459 \cdot 10^{-14}$;

2. $s = 6$; $6.148483 \cdot 10^{-12}$;

3. $s = 6$; $2.002705 \cdot 10^{-9}$;

4. $s = 4$; $1.751117 \cdot 10^{-6}$;

5. $s = 3$; $0.0003178107$;

6. $s = 2$; $0.08571896$;

7. $s = 1$; $0.2869012$.

1. Оценим коэффициент перед $x, x^2, x^3, x^4, x^5, x^6, x^7$:

   $s = 7$;

   $P_7'(x)t_1'(x) = 0$;

   $\frac{1}{16}(429 \cdot x^7 - 693 \cdot x^5 + 315 \cdot x^3 - 35x)' = 0$.

   Решая, получаем план:

$$\begin{pmatrix} -1 & -0.87 & -0.59 & -0.2 & 0.2 & 0.59 & 0.87 & 1 \\ 0.07 & 0.13 & 0.14 & 0.12 & 0.12 & 0.14 & 0.13 & 0.07 \end{pmatrix}.$$

   Определитель информационной матрицы: $1.39743 \cdot 10^{-15}$.

   Определитель усеченной матрицы: $9.988046 \cdot 10^{-14}$.

2. Оценим коэффициент перед $x^2, x^3, x^4, x^5, x^6, x^7$:

   $s = 6$;

   $P_6'(x)U_1(x) - P_5'(x)U_0(x) = 0$;

   $\frac{1}{16}(231 \cdot x^6 - 315 \cdot x^4 + 105 \cdot x^2 - 5)' \cdot 2x - \frac{1}{8}(63 \cdot x^5 - 70 \cdot x^3 + 15x) = 0$.

   Решая, получаем план:

$$\begin{pmatrix} -1 & -0.87 & -0.58 & -0.2 & 0.2 & 0.58 & 0.87 & 1 \\ 0.07 & 0.11 & 0.13 & 0.15 & 0.15 & 0.13 & 0.11 & 0.07 \end{pmatrix}.$$

   Определитель информационной матрицы: $2.889897 \cdot 10^{-15}$.

   Определитель усеченной матрицы: $2.472375 \cdot 10^{-12}$.

3. Оценим коэффициент перед $x^3, x^4, x^5, x^6, x^7$:

   $s = 5$;

   $P_5'(x)U_2(x) - P_4'(x)U_1(x) = 0$;

   $\frac{1}{8}(63 \cdot x^5 - 70 \cdot x^3 + 15x)'(4x^2 - 1) - \frac{1}{8}(35 \cdot x^4 - 30 \cdot x^2 + 3)' \cdot 2x = 0$.

   Решая, получаем план:

$$\begin{pmatrix} -1 & -0.86 & -0.59 & -0.22 & 0.22 & 0.59 & 0.86 & 1 \\ 0.08 & 0.12 & 0.14 & 0.14 & 0.14 & 0.14 & 0.12 & 0.08 \end{pmatrix}.$$

Определитель информационной матрицы: $3.578557 \cdot 10^{-15}$.

Определитель усеченной матрицы: $1.138478 \cdot 10^{-9}$.

4. Оценим коэффициент перед $x^4, x^5, x^6, x^7$:

   $s = 4$;

   $P_4'(x)U_3(x) - P_3'(x)U_2(x) = 0$;

   $\frac{1}{8}(35 \cdot x^4 - 30 \cdot x^2 + 3)' \cdot (8x^3 - 4x) - \frac{1}{2}(5x^3 - 3x)'(4x^2 - 1) = 0.$

   Решая, получаем план:

$$\begin{pmatrix} -1 & -0.86 & -0.59 & -0.22 & 0.22 & 0.59 & 0.86 & 1 \\ 0.08 & 0.15 & 0.13 & 0.12 & 0.12 & 0.13 & 0.15 & 0.08 \end{pmatrix}.$$

   Определитель информационной матрицы: $4.305941 \cdot 10^{-15}$.

   Определитель усеченной матрицы: $1.21643 \cdot 10^{-6}$.

5. Оценим коэффициент перед $x^5, x^6, x^7$:

   $s = 3$;

   $P_3'(x)U_4(x) - P_2'(x)U_3(x) = 0$;

   $\frac{1}{2}(5x^2 - 3x) \cdot (16x^4 - 12x^2 + 1) - \frac{1}{2}(3x^2 - 1)' \cdot (8x^3 - 4x) = 0;$

   Решая, получаем план:

$$\begin{pmatrix} -1 & -0.87 & -0.58 & -0.22 & 0.22 & 0.58 & 0.87 & 1 \\ 0.9 & 0.14 & 0.13 & 0.12 & 0.12 & 0.13 & 0.14 & 0.9 \end{pmatrix}.$$

   Определитель информационной матрицы: $4.779823 \cdot 10^{-15}$.

   Определитель усеченной матрицы: 0.0002046849.

6. Оценим коэффициент перед $x^6, x^7$:

   $s = 2$;

   $P_2'(x)U_5(x) - P_1'(x)U_4(x) = 0$;

   $\frac{1}{2}(3x^2 - 1)'(32x^5 - 32x^3 + 6x) - (16x^4 - 12x^2 + 1) = 0.$

   Решая, получаем план:

$$\begin{pmatrix} -1 & -0.89 & -0.58 & -0.19 & 0.19 & 0.58 & 0.89 & 1 \\ 0.1 & 0.12 & 0.15 & 0.13 & 0.13 & 0.15 & 0.12 & 0.1 \end{pmatrix}.$$

   Определитель информационной матрицы: $4.577615 \cdot 10^{-15}$.

   Определитель усеченной матрицы: 0.06440268.

7. Оценим коэффициент перед $x^7$:

   $s = 1$;

   $P_1'(x)U_6(x) = 0$;

   $64x^6 - 80x^4 + 24x^2 - 1 = 0$.

   Решая, получаем план:

$$\begin{pmatrix} -1 & -0.9 & -0.62 & -0.2 & 0.2 & 0.62 & 0.9 & 1 \\ 0.11 & 0.12 & 0.13 & 0.14 & 0.14 & 0.13 & 0.12 & 0.11 \end{pmatrix}.$$

Определитель информационной матрицы: $4.788364 \cdot 10^{-15}$.

Определитель усеченной матрицы: 0.2755611.

Эффективности $D$-оптимального плана, относительно усеченных:

1. Относительно $D_1$ — 0.4;
2. Относительно $D_2$ — 0.5;
3. Относительно $D_3$ — 0.6;
4. Относительно $D_4$ — 0.75;
5. Относительно $D_5$ — 0.86;
6. Относительно $D_6$ — 0.97.

Эффективности усеченных планов, относительно $D$-оптимального:

1. $D_1$, относительно $D$ — 0.89;
2. $D_2$, относительно $D$ — 0.92;
3. $D_3$, относительно $D$ — 0.94;
4. $D_4$, относительно $D$ — 0.96;
5. $D_5$, относительно $D$ — 0.97;
6. $D_6$, относительно $D$ — 0.98.

Здесь видим, что эффективность усеченных планов также больше, чем эффективность $D$-оптимальных планов, хоть и не так сильно, как в предыдущих случаях. Вероятно, это связано с тем, что на рассматриваемом промежутке $[0, 1]$, при увеличении количества точек плана, сами точки становятся очень близки.

# Глава 3

# Подход Стиглера

В статье ([1]) Стиглер описывает критерий, который он предложил, обладающий свойствами:

1. По своему построению план должен предусматривать проверку того, является ли модель адекватной подгонкой к истинной регрессионной функции.

2. Если было решено, что модель адекватна, то с помощью плана можно было сделать какие-либо выводы, относительно модели.

3. Оптимальный план не должен зависеть от неизвестных параметров модели.

Положим, что $D(\xi) = M^{-1}(\xi)$ — дисперсионная матрица.

**Определение 6.** *Критерий $G$-оптимальности имеет вид:*

$$\max_{t \in \chi} d(t, \xi) \to \sup_{\xi \in \Xi_H},$$

*где $\Xi_H$ — такое множество планов, где* $\det M(\xi) \neq 0$, *а* $d(t, \xi) = f(t)^T D(\xi) f(t)$.

Критерий, который придумал Стиглер, является компромиссом между $D$-оптимальным критерием и $D_s$-оптимальным критерием. Рассмотрим класс полиномиальных моделей. Пусть $P_m$ – модель вида $a + a_1 x + a_2 x^2 + \ldots a_m x^m$.

**Определение 7.** *Будем называть план $\xi_0$ $C$-ограниченным $D$-оптимальным планом для модели $P_m$, если $\xi_0$ максимизирует* $\det M_{(m)}$ *среди всех планов $\xi$, удовлетворяющих условию:*

$$\det M_{(m)} \leq C \det M_{(m+1)}.$$

**Определение 8.** *Будем называть план $\xi_0$ $C$-ограниченным $G$-оптимальным планом для модели $P_m$, если $\xi_0$ минимизирует:*

$$\max_{-1 \leq x \leq 1} d_m(x, \xi)$$

*среди всех планов $\xi$, удовлетворяющих условию:*

$$\det M_{(m)} \leq C \det M_{(m+1)},$$

*где* $d_m(x, \xi) = g(x)' \det M_{(m)}^{-1} g(x)$, *а* $g(x)' = (1, x, \ldots, x^m)$.

Оптимальный план, удовлетворяющий и первому и второму определению, должен также (автоматически) удовлетворять условию:

$$\mathrm{var}(\hat{\beta_{m+1}}) = \sigma^2 \frac{C}{n}.$$

Возникает естественный вопрос: как выбрать $C$? Выбирать эту константу нужно таким образом, чтобы был достигнут баланс между двумя целями:

1. Точные выводы о значении старшего коэффициента $\beta_m$;

2. Точные выводы про модель $P_m$.

Для поиска $C$-ограниченных $G$- и $D$-оптимальных планов для полиномиальной модели порядка 1 удобно пользоваться следующей теоремой:

**Теорема 4.** *([1]) Для $C \geq 4$, $C$-ограниченные $G$- и $D$-оптимальные планы сосредоточены в точках: $\pm 1$ и $0$ с весами:*

$$\xi_0(-1) = \xi_0(1) = \frac{1}{4} + \frac{1}{2}\sqrt{\frac{1}{4} - \frac{1}{C}}, \xi_0(0) = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{C}}.$$

## 3.1. Сравнение результатов

Сравним эффективности $C$-усеченных $D$-оптимальных планов, предложенных Стиглером, относительно $D$-оптимальных планов, эффективных для оценки всех параметров модели, и относительно усеченных $D$-оптимальных планов, эффективных для оценки части параметров.

Для начала вспомним, как выглядят усеченные $D$-оптимальные планы для оценки старшего параметра, для полиномиальных моделей некоторых порядков, которые были найдены ранее:

1. Для полиномиальной модели второго порядка:

$$\xi = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}.$$

2. Для кубической полиномиальной модели:

$$\xi = \begin{pmatrix} -1 & -\frac{1}{2} & \frac{1}{2} & 1 \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}.$$

Приведем $D$-оптимальный план для линейной полиномиальной модели:

$$\xi = \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Далее, приведем $C$-ограниченные $D$-и $G$-оптимальные планы, в зависимости от разных $C$ для линейной модели.

Таблица 3.1. Планы, в зависимости от константы $C$

| Значение $C$ | Вес точки -1 | Вес точки 0 | Вес точки 1 |
|---|---|---|---|
| 4 | 0.25 | 0.5 | 0.25 |
| 4.5 | 0.33 | 0.33 | 0.33 |
| 5 | 0.362 | 0.276 | 0.362 |
| 6 | 0.394 | 0.211 | 0.394 |
| 7 | 0.414 | 0.173 | 0.414 |
| 25 | 0.479 | 0.042 | 0.479 |
| 1000 | 0 | 1 | 0 |

Отсюда видно, что при $C = 4$ план совпадает с $D_s$-оптимальным планом для оценки старшего коэффициента квадратичной модели. При $C = 4.5$ план совпадает с $D$-оптимальным планом в случае линейной модели. При $C \to \infty$ план становится $D$-оптимальным планом для линейной модели.

Для наглядности нарисуем график зависимости от $C$ следующих эффективностей:

- Эффективность плана Стиглера, относительно $D$-оптимального плана для линейной модели;
- Эффективность плана Стиглера, относительно $D$-оптимального плана для квадратичной модели;
- Эффективность плана Стиглера, относительно $D_s$-оптимального плана для квадратичной модели.

Рассмотрим эффективности этих планов, в зависимости от $C$:

Таблица 3.2. Эффективность, в зависимости от константы $C$

| Значение $C$ | для линейной | для квадратичной | для $D_1$ |
|---|---|---|---|
| 4 | 0.70 | 0.97 | 1 |
| 4.5 | 0.82 | 1 | 0.97 |
| 6 | 0.85 | 0.96 | 0.96 |
| 9 | 0.93 | 0.86 | 0.87 |
| 20 | 0.97 | 0.68 | 0.76 |
| 30 | 0.98 | 0.85 | 0.70 |
| 100 | 1 | 0.45 | 0.63 |

На рисунке 3.1 красным показана эффективность $C$-ограниченных $D$-оптимальных планов, по отношению к $D$-оптимальным для квадратичной полиномиальной модели, синим — эффективность $C$-ограниченных
$D$-оптимальных планов, по отношению к $D$-оптимальным для линейной модели, а зеленым — по отношению к $D_s$-оптимальным.

Как мы видим, с ростом $C$ повышается эффективность планов, рассматриваемых в статье Стиглера, относительно $D$-оптимальных для линейной модели, что является логичным, ведь при $C \to \infty$ $C$-ограниченные планы для рассмотренной модели становятся $D$-оптимальными планами для линейной модели. Эффективность $C$-ограниченных $D$-оптимальных планов, относительно $D$-оптимальных планов, относительно $D$-оптимальных планов для квадратичной модели повышается до значения $C = 4.5$, затем начинает резко падать. Для $D_s$-оптимального плана ситуация аналогичная после значения $C = 4$ эффективность начинает падать.

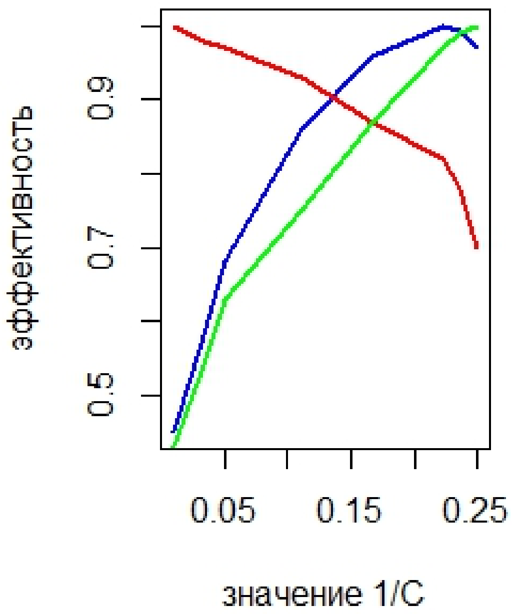


Рис. 3.1. Эффективность $C$-ограниченных $D$-оптимальных планов, по отношению к $D$-оптимальным и $D_s$-оптимальным.

# Глава 4

# Переход к более широкому классу планов

Вспомним, что мы хотим решить задачу о дискриминации моделей. В то же время оптимально оценить параметры модели и в случае принятия нулевой гипотезы, и в случае принятия альтернативной. Возможно ли найти планы, удовлетворяющие этим условиям? Рассмотрим новый класс планов на примере квадратичной и кубической модели.

Рассмотрим функцию

$$\Psi_\gamma(\xi) = \gamma \ln det(M_1(\xi)) + (1-\gamma) \ln det(M_2(\xi)),$$

где $M_1(\xi)$ и $M_2(\xi)$ - информационные матрицы для полиномиальных моделей порядка $n-1$ и $n$. Данная функция является строго вогнутой по леммам 2.4,2.5,2.6 (см. [7]). Введем следующие обозначения:

$$d_1(\xi, x) = f_{(n-1)}^T(x) M_1^{-1}(\xi) f_{(n-1)}(x), \quad d_2(\xi, x) = f_{(n)}^T(x) M_2(\xi)^{-1} f_{(n)}(x),$$

$f_{(n-1)}(x), f_{(n)}(x)$ — вектора регрессионных функций полиномиальных моделей порядка $n-1$ и $n$ соответственно.

**Определение 9.** *Будем говорить, что план $\xi$ $G_\gamma$-оптимальный, если*

$$\xi = \arg\min_\xi \max_x d_\gamma(x, \xi), \text{ где } d_\gamma(x,\xi) = \gamma d_1(x,\xi) + (1-\gamma) d_2(x,\xi).$$

**Определение 10.** *Будем говорить, что план $\xi$ $\Psi_\gamma$-оптимальный, если*

$$\xi = \arg\max_x \Psi_\gamma(\xi).$$

Имеет место следующая лемма:

**Лемма 1.** *Пусть $\xi_\alpha = \alpha\xi_1 + (1-\alpha)\xi_0$, $\xi_1$ — произвольный план, $\xi_0$ — невырожденный. Тогда*

$$\frac{\partial \Psi_\gamma(\xi_\alpha)}{\partial \alpha}\Big|_{\alpha=0+} = \gamma \int d_1(x,\xi_0)\xi_1(dx) + (1-\gamma) \int d_2(x,\xi_0)\xi_1(dx) - (n+1-\gamma).$$

**Доказательство Леммы 1**

Для доказательства мы воспользуемся вспомогательным результатом (см. Лемму 2.7 в [7])

**Лемма 2.** *Для произвольной матрицы $A > 0$ имеет место следующая формула:*

$$\frac{\partial}{\partial \alpha} \ln \det A(\alpha) = \mathrm{tr} A^{-1}(\alpha) \frac{\partial}{\partial \alpha} A(\alpha).$$

По Лемме 2 имеем:

$$\frac{\partial \Psi_\gamma(\xi_\alpha)}{\partial \alpha} = \gamma \mathrm{tr} M_1^{-1}(\xi_\alpha) \frac{\partial M_1(\xi_\alpha)}{\partial \alpha}\Big|_{\alpha=0+} + (1-\gamma) \mathrm{tr} M_2^{-1}(\xi_\alpha) \frac{\partial M_2(\xi_\alpha)}{\partial \alpha}\Big|_{\alpha=0+} =$$

$$= \gamma \mathrm{tr} M_1^{-1}(\xi_0)(M_1(\xi_1) - M_1(\xi_0)) + (1-\gamma) \mathrm{tr} M_2^{-1}(\xi_0)(M_2(\xi_1) - M_2(\xi_0)) =$$

$$\gamma \int f^T_{(n-1)}(x) M_1^{-1}(\xi_0) f_{(n-1)}(x) \xi_1(dx) - \gamma n + (1-\gamma) \int d_2(x, \xi_0) \xi_1(dx) - (1-\gamma)(n+1) =$$

$$\gamma \int d_1(x, \xi_0) \xi_1(dx) + (1-\gamma) \int d_2(x, \xi_0) \xi_1(dx) - (n + 1 - \gamma).$$

Лемма доказана.

$\square$

**Лемма 3.** *Пусть план $\xi^*$ оптимален в смысле $\Psi_\gamma$-критерия. Обозначим за $\xi_\alpha = \alpha\xi + (1-\alpha)\xi^*$, $\xi$ — план, сосредоточенный в точке $x$, $0 < \alpha < 1$. Тогда*

$$d_\gamma(x, \xi^*) \leq (n + 1 - \gamma)$$

**Доказательство Леммы 3**

В силу $\Psi_\gamma$-оптимальности плана $\xi^*$ имеет место следующее неравенство:

$$\Psi_\gamma(\xi_\alpha) \leq \Psi_\gamma(\xi^*) \Leftrightarrow \frac{\Psi_\gamma(\xi_\alpha) - \Psi_\gamma(\xi^*)}{\alpha} \leq 0.$$

Переходя к пределу при $\alpha \to 0+$, с учетом Леммы 1, получаем

$$\frac{\partial}{\partial \alpha} \Psi_\gamma(\xi_\alpha)\Big|_{\alpha=0+} = \gamma d_1(x, \xi^*) + (1-\gamma) d_2(x, \xi^*) - (n + 1 - \gamma) \leq 0.$$

Лемма доказана.

$\square$

**Лемма 4.** *Для любого невырожденного плана $\xi$ имеет место неравенство* $\max_x d_\gamma(x,\xi) \geq (n+1-\gamma)$.

**Доказательство.** Прямое вычисление дает

$$\max_x d_\gamma(x,\xi) \geq \gamma \int d_1(x,\xi)\xi(dx) + (1-\gamma)\int d_2(x,\xi)\xi(dx) =$$
$$= \gamma \int f^T_{(n-1)}(x) M_1^{-1}(\xi) f_{(n-1)}(x) d\xi + (1-\gamma)\int f^T_{(n)}(x) M_2^{-1}(\xi) f_{(n)}(x) d\xi =$$
$$= \gamma \mathrm{tr} M_1^{-1}(\xi) M_2(\xi) + (1-\gamma)\mathrm{tr} M_2^{-1}(\xi) M_2(\xi) = \gamma n + (1-\gamma)(n+1) = n+1-\gamma.$$

Лемма доказана.

□

Сформулируем и докажем аналог теоремы эквивалентности для $\Psi_\gamma$-критерия

**Теорема 5** (Эквивалентности для $\Psi_\gamma$-критерия)**.** *Рассмотрим полиномиальные модели порядка $n-1$ и $n$. Если множество информационных матриц $M_1(\xi)$ и $M_2(\xi)$ компактно, то следующие условия эквивалентны:*

*(a) план $\xi^*$ — $\Psi_\gamma$-оптимальный;*

*(b) план $\xi^*$ — $G_\gamma$-оптимальный;*

*(c)* $\max_{x\in\chi} d_\gamma(x,\xi^*) = (n+1-\gamma)$.

*Причем план $\xi^*$ сосредоточен в конечном числе точек и последнее равенство достигается в точках $x_i^*$ оптимального плана $\xi^*$. Кроме того, информационные матрицы всех оптимальных планов совпадают.*

**Доказательство теоремы 5.**

Схема доказательства будет следующей: $(a) \Rightarrow (c) \Rightarrow (b) \Rightarrow (a)$.

**1**. (**a**) $\Rightarrow$ (**c**). Пусть план $\xi^*$ оптимален в смысле $\Psi_\gamma$-критерия.

Обозначим за $\xi_\alpha = \alpha\xi + (1-\alpha)\xi^*$, $\xi$ — план, сосредоточенный в точке $x$, $0 < \alpha < 1$. По Лемме 3 $d_\gamma(x,\xi^*) \leq (n+1-\gamma)$ для $\forall x$, и следовательно, $\max_x d_\gamma(x,\xi^*) = (n+1-\gamma)$.

**2**. (**c**) $\Rightarrow$ (**b**). Пусть для плана $\xi^*$ выполнено $\max_x d_\gamma(x,\xi^*) = (n+1-\gamma)$, тогда по лемме 4 план $\xi^*$ — $G_\gamma$-оптимальный.

**3**. (**b**) $\Rightarrow$ (**a**). Доказательство будем проводить от противного. Предположим, что существует $G_\gamma$-оптимальный план, который не является $D_\gamma$-оптимальным. Обозначим этот план через $\xi^*$. Пусть план $\xi_{\Psi_\gamma}$ — некоторый $\Psi_\gamma$-оптимальный план. Для плана $\xi_\alpha = (1-\alpha)\xi^* + \alpha\xi_{\Psi_\gamma}$ в силу строгой вогнутости функции $\Psi_\gamma(\xi)$ имеем

$$\Psi_\gamma(\xi_\alpha) \geq (1-\alpha)\Psi_\gamma(\xi^*) + \alpha\Psi_\gamma(\xi_{\Psi_\gamma}).$$

Вычитая из обеих частей $\Psi_\gamma(\xi^*)$, получаем

$$\Psi_\gamma(\xi_\alpha) - \Psi_\gamma(\xi^*) \geq \alpha\left(\Psi_\gamma(\xi_{\Psi_\gamma}) - \Psi_\gamma(\xi^*)\right) > 0,$$

так как план $\xi^*$ не является $\Psi_\gamma$-оптимальным. Из полученного неравенства следует, что

$$\frac{\Psi_\gamma(\xi_\alpha) - \Psi_\gamma(\xi^*)}{\alpha} > 0,$$

и значит,

$$\frac{\partial}{\partial\alpha}\Psi_\gamma(\xi_\alpha)\Big|_{\alpha=0+} = \gamma\int d_1(x,\xi^*)\xi_{\Psi_\gamma}(dx) + (1-\gamma)\int d_2(x,\xi^*)\xi_{\Psi_\gamma}(dx) - (n+1-\gamma) > 0,$$

т.е. мы получили, что $\int d_\gamma(x,\xi^*)\xi_{\Psi_\gamma}(dx) > (n+1-\gamma)$. С другой стороны, так как план $\xi^*$ — $G_\gamma$-оптимальный, имеет место неравенство

$$d_\gamma(x,\xi^*) \leq (n+1-\gamma) \ \ \forall x.$$

Получили противоречие. Следовательно, $G_\gamma$-оптимальный план $\xi^*$ также является $\Psi_\gamma$-оптимальным. Таким образом, мы доказали эквивалентность трех условий $(a), (b), (c)$. То, что максимум функции $d_\gamma(x,\xi)$ достигается в точках оптимального плана (если план сосредоточен в конечном числе точек), следует из Леммы 4 (для плана $\xi = \xi^*$ неравенство в Лемме превращается в равенство). То, что информационные матрицы всех оптимальных планов совпадают, следует из строгой вогнутости функции $\Psi_\gamma$ и выпуклости множества непрерывных планов.
Теорема доказана.

$\square$

Применим теорему эквивалетности для введенных планов для численного поиска $\Psi_\gamma$-оптимальных планов для полиномиальных моделей некоторых степеней.

## 4.1. Квадратичная модель

Основная модель имеет вид: $\theta_1 + \theta_2 x + \theta_3 x^2$, где $\theta_i$ — оцениваемые параметры, а $x \in [-1, 1]$. Задача: найти максимум величины $\alpha \ln\det M_1(\xi) + (1-\alpha)\ln\det M_2(\xi)$ по планам $\xi$ вида:

$$\begin{pmatrix} -1 & 0 & 1 \\ \frac{1-\nu}{2} & \nu & \frac{1-\nu}{2} \end{pmatrix},$$

где $0 < \nu < 1$, $M_2$ — информационная матрица для квадратичной модели, $M_1$ — информационная матрица для линейной модели.

Заметим, что при некотором значении $\alpha$ эти планы могут совпасть как с $M_2$ - оптимальным, так и с $M_1$- оптимальным.

## 4.2. Планы для некоторых значений $\alpha$

Мы знаем какую величину мы хотим максимизировать и знаем вид планов. Следовательно, можем найти численно планы для некоторых значений $\alpha$:

1. $\alpha$=0.1
$$\begin{pmatrix} -1 & 0 & 1 \\ 0.34 & 0.32 & 0.34 \end{pmatrix}.$$
2. $\alpha$=0.5
$$\begin{pmatrix} -1 & 0 & 1 \\ 0.375 & 0.25 & 0.375 \end{pmatrix}.$$
3. $\alpha$=0.9
$$\begin{pmatrix} -1 & 0 & 1 \\ 0.46 & 0.08 & 0.46 \end{pmatrix}.$$

## 4.3. Сравнение эффективности

Как мы видим из постановки задачи, мы ищем планы, которые максимизируют: $\alpha \ln\det M_1(\xi) + (1-\alpha)\ln\det M_2(\xi)$, следовательно при $\alpha = 0$ мы получим просто оптимальный план, близкий к квадратичной модели, а при $\alpha = 1$ – оптимальный план близкий к линейной модели.

Рассмотрим эффективность найденных планов для квадратичной модели, относительно

$D$-оптимальных планов для квадратичной модели, $D$-оптимальных планов для линейной модели, и усеченных $D$-оптимальных планов для оценки старшего коэффициента квадратичной модели. Для начала вспомним вид этих планов.

$D$-оптимальный план для квадратичной модели:

$$\xi = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

$D$-оптимальный план для линейной модели:

$$\xi = \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Усеченный $D$-оптимальный план для оценки старшего коэффициента квадратичной модели:

$$\xi = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}.$$

Приведем таблицу зависимости эффективности от значения $\alpha$:(таблица 4.1)

Таблица 4.1. Эффективность, в зависимости от $\alpha$

| Значение $\alpha$ | для квадратичной | для линейной | для $D_s$ |
|---|---|---|---|
| 0.1 | 0.97 | 0.21 | 0.84 |
| 0.2 | 0.87 | 0.42 | 0.88 |
| 0.3 | 0.79 | 0.56 | 0.95 |
| 0.4 | 0.70 | 0.65 | 0.98 |
| 0.5 | 0.64 | 0.73 | 0.93 |
| 0.6 | 0.56 | 0.80 | 0.85 |
| 0.7 | 0.44 | 0.86 | 0.79 |
| 0.8 | 0.33 | 0.92 | 0.73 |
| 0.9 | 0.20 | 0.96 | 0.68 |

Для наглядности приведен график зависимости эффективности планов Стиглера, относительно остальных, в зависимости от значения $\alpha$. На рисунке 4.1 синим — эффек-

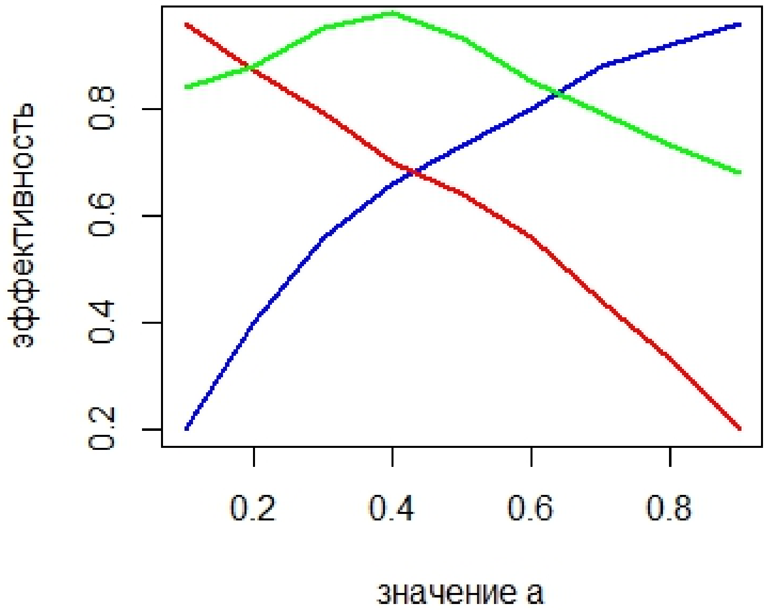


Рис. 4.1. Эффективность планов Стиглера, по отношению к $D$-оптимальным и $D_s$-оптимальным.

тивность найденного плана, относительно $D$-оптимального плана для линейной модели. Красным — эффективность, относительно $D$- оптимального плана для квадратичной модели, а зеленым — относительно усеченного $D$-оптимального плана для оценки старшего коэффициента квадратичной модели.

Как мы видим, с ростом $\alpha$ эффективность найденных планов, относительно $D$-оптимальных планов для линейной модели растет. Тогда как эффективность, относительно $D$-оптимальных планов для квадратичной модели уменьшается, а эффективность, относительно усеченного $D$-оптимального плана для оценки старшего коэффициента квадратичной модели растет до определенного значения ($\alpha \approx 0.4$), а затем начинает убывать.

## 4.4. Кубическая модель

Основная модель имеет вид: $\theta_1+\theta_2 x+\theta_3 x^2+\theta_4 x^3$, где $\theta_i$ – оцениваемые параметры,а $x \in [-1,1]$. Задача: найти максимум величины $\alpha \ln\det M_2(\xi) + (1-\alpha)\ln\det M_3(\xi)$ по планам $\xi$ вида:

$$\begin{pmatrix} -1 & -t & t & 1 \\ \nu & \mu & \mu & \nu \end{pmatrix},$$

где $0 < t < 1$, $\nu+\mu = \frac{1}{2}$, $M_2$ – информационная матрица для квадратичной модели, $M_3$ – информационная матрица для кубической модели.

Заметим, что при некотором значении $\alpha$ эти планы могут совпасть как с $M_3$ - оптимальным, так и с $M_2$- оптимальным.

## 4.5. Планы для некоторых значений $\alpha$

Мы знаем какую величину мы хотим максимизировать и знаем вид планов. Можем найти $\det M_3$ и $\det M_2$.

Для различных $\alpha$ получаем планы:

1. $\alpha = 0.1$

$$\begin{pmatrix} -1 & -0.44 & 0.44 & 1 \\ 0.26 & 0.24 & 0.24 & 0.26 \end{pmatrix}.$$

2. $\alpha = 0.5$

$$\begin{pmatrix} -1 & -0.4 & 0.4 & 1 \\ 0.27 & 0.22 & 0.22 & 0.27 \end{pmatrix}.$$

3. $\alpha = 0.9$

$$\begin{pmatrix} -1 & -0.23 & 0.23 & 1 \\ 0.32 & 0.18 & 0.18 & 0.32 \end{pmatrix}.$$

## 4.6. Сравнение эффективности

Как мы видим из постановки задачи, мы ищем планы, которые максимизируют: $\alpha \ln\det M_2(\xi) + (1-\alpha)\ln\det M_3(\xi)$, следовательно при $\alpha = 0$ мы получим план близкий к оптимальному плану для кубической модели, а при $\alpha = 1$ — план, близкий к плану для кубической модели.

Рассмотрим эффективность планов Стигера для кубической модели, относительно

$D$-оптимальных планов для квадратичной модели, $D$-оптимальных планов для кубической модели, и усеченных $D$-оптимальных планов для оценки старшего коэффициента кубической модели. Для начала вспомним вид этих планов.

$D$-оптимальный план для квадратичной модели:

$$\xi = \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

$D$-оптимальный план для кубической модели:

$$\xi = \begin{pmatrix} -1 & -\frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Усеченный $D$-оптимальный план для оценки старшего коэффициента кубической модели:

$$\xi = \begin{pmatrix} -1 & -\frac{1}{2} & \frac{1}{2} & 1 \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}.$$

Приведем таблицу зависимости эффективности от значения $\alpha$(таблица 4.2):

Таблица 4.2. Эффективность, в зависимости от $\alpha$

| Значение $\alpha$ | для кубической | для квадратичной | для $D_1$ |
|---|---|---|---|
| 0.1 | 0.96 | 0.20 | 0.29 |
| 0.2 | 0.87 | 0.40 | 0.41 |
| 0.3 | 0.79 | 0.56 | 0.52 |
| 0.4 | 0.70 | 0.66 | 0.63 |
| 0.5 | 0.64 | 0.73 | 0.74 |
| 0.6 | 0.56 | 0.80 | 0.85 |
| 0.7 | 0.41 | 0.88 | 0.95 |
| 0.8 | 0.29 | 0.94 | 0.98 |
| 0.9 | 0.15 | 0.98 | 0.94 |

Для наглядности приведен график зависимости эффективности планов Стиглера, относительно остальных, в зависимости от значения $\alpha$. На рисунке 4.2 синим — эффек-

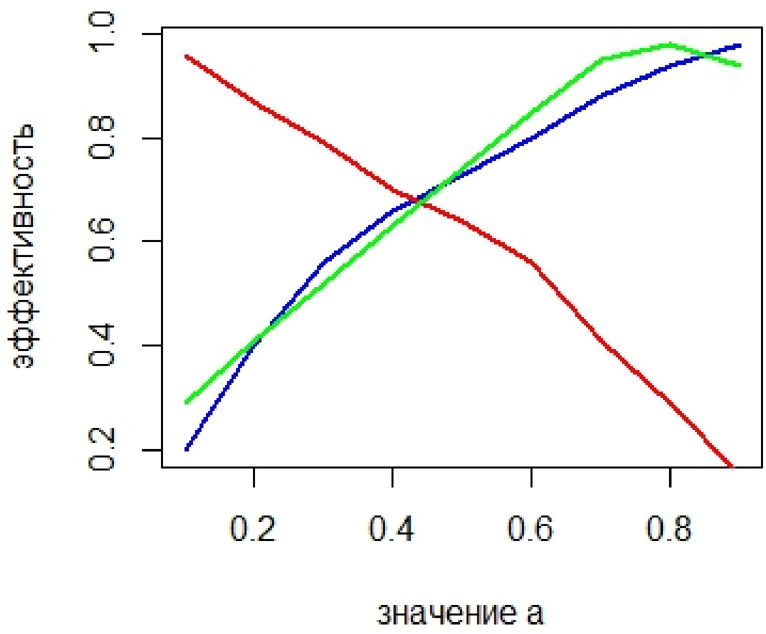


Рис. 4.2. Эффективность планов Стиглера, по отношению к $D$-оптимальным и $D_s$-оптимальным.

тивность плана Стиглера для квадратичной модели, относительно $D$-оптимального плана для квадратичной модели. Красным — эффективность, относительно $D$- оптимального плана для кубической модели, а зеленым — относительно усеченного $D$-оптимального плана для оценки старшего коэффициента кубической модели.

Как мы видим, с ростом $\alpha$ эффективность планов Стиглера, относительно $D$-оптимальных планов для квадратичной модели растет. Тогда как эффективность, относительно $D$-оптимальных планов для кубической модели уменьшается, а эффективность, относительно усеченного $D$-оптимального плана для оценки старшего коэффициента кубической модели растет до определенного значения, а затем начинает убывать.

## 4.7. Полиномиальная модель четвертой степени

Основная модель имеет вид: $\theta_1 + \theta_2 x + \theta_3 x^2 + \theta_4 x^3 + \theta_5 x^4$, где $\theta_i$ – оцениваемые параметры,а $x \in [-1, 1]$.

Задача: найти максимум величины $\alpha \ln \det M_3(\xi) + (1-\alpha) \ln \det M_4(\xi)$ по планам $\xi$ вида:

$$\begin{pmatrix} -1 & -t & 0 & t & 1 \\ \omega_1 & \omega_2 & 1 - 2\omega_1 - 2\omega_2 & \omega_2 & \omega_1 \end{pmatrix}.$$

Заметим, что при некотором значении $\alpha$ эти планы могут совпасть как с $M_3$ - оптимальным, так и с $M_4$- оптимальным.

## 4.8. Планы для некоторых значений $\alpha$

Так как мы знаем вид планов, и величину, которую мы хотим максимизировать, то можем численно найти эти планы для некоторых значений $\alpha$.

1. $\alpha = 0.1$
$$\begin{pmatrix} -1 & -0.65 & 0 & 0.65 & 1 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \end{pmatrix}.$$
2. $\alpha = 0.5$
$$\begin{pmatrix} -1 & -0.61 & 0 & 0.61 & 1 \\ 0.22 & 0.19 & 0.17 & 0.19 & 0.22 \end{pmatrix}.$$
3. $\alpha = 0.6$
$$\begin{pmatrix} -1 & -0.59 & 0 & 0.59 & 1 \\ 0.225 & 0.19 & 0.16 & 0.19 & 0.225 \end{pmatrix}.$$
4. $\alpha = 0.7$
$$\begin{pmatrix} -1 & -0.57 & 0 & 0.57 & 1 \\ 0.23 & 0.19 & 0.14 & 0.19 & 0.23 \end{pmatrix}.$$
5. $\alpha = 0.8$
$$\begin{pmatrix} -1 & -0.55 & 0 & 0.55 & 1 \\ 0.23 & 0.2 & 0.12 & 0.2 & 0.23 \end{pmatrix}.$$
6. $\alpha = 0.9$
$$\begin{pmatrix} -1 & -0.51 & 0 & 0.51 & 1 \\ 0.24 & 0.22 & 0.08 & 0.22 & 0.24 \end{pmatrix}.$$

Заметим, что как и в предыдущих случаях, в случае полинома степени 4 при росте $\alpha$ план становится похож на $D$-оптимальный план для полиномиальной модели степени 3. В то же время, при маленьком значении $\alpha$, наоборот, на $D$-оптимальный план для полиномиальной модели степени 4.

## 4.9. Сравнение эффективности

Как мы видим из постановки задачи, мы ищем планы, которые максимизируют:$\alpha \ln \det M_3(\xi$ $(1-\alpha) \ln \det M_4(\xi)$, следовательно при $\alpha = 0$ мы получим просто оптимальный план для квадратичной модели, а при $\alpha = 1$ – оптимальный план для кубической модели. Рассмотрим эффективность планов Стигера для кубической модели, относительно $D$-оптимальных планов для полиномиальной модели 4-ой степени, $D$-оптимальных планов для кубической модели, и усеченных $D_s$-оптимальных планов для оценки старшего коэффициента полиномиальной модели 4-ой степени. Для начала вспомним вид этих планов.

$D$-оптимальный план для полиномиальной модели 4-ой степени:

$$\xi = \begin{pmatrix} -1 & -\sqrt{\frac{3}{7}} & 0 & \sqrt{\frac{3}{7}} & 1 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}.$$

$D$-оптимальный план для кубической модели:

$$\xi = \begin{pmatrix} -1 & -\frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Усеченный $D$-оптимальный план для оценки старшего коэффициента полиномиальной модели 4-ой степени:

$$\xi = \begin{pmatrix} -1 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 1 \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} \end{pmatrix}.$$

Приведем таблицу зависимости эффективности от значения $\alpha$(таблица 4.3):

Из таблицы 4.3 видно, что с ростом $\alpha$ эффективность планов Стиглера для кубической модели, относительно $D$-оптимальных планов для кубической модели возрастает. Эффективность планов Стиглера для кубической модели, относительно $D$-оптимальных

Таблица 4.3. Эффективность, в зависимости от $\alpha$

| Значение $\alpha$ | для кубической | для 4 степени | для $D_1$ |
|---|---|---|---|
| 0.1 | 0.13 | 0.96 | 0.31 |
| 0.2 | 0.29 | 0.94 | 0.43 |
| 0.3 | 0.40 | 0.87 | 0.56 |
| 0.4 | 0.55 | 0.80 | 0.68 |
| 0.5 | 0.63 | 0.73 | 0.81 |
| 0.6 | 0.71 | 0.65 | 0.90 |
| 0.7 | 0.79 | 0.56 | 0.98 |
| 0.8 | 0.88 | 0.43 | 0.92 |
| 0.9 | 0.96 | 0.15 | 0.86 |

планов для полиномиальной модели четвертой степени убывает. А эффективность, относительно усеченного оптимального плана для оценки старшего коэффициента возрастает до определенного значения $\alpha$, а зачем начинает убывать.

# Заключение

В работе был численно исследован подход Стиглера к дискриминации регрессионных моделей полиномиального вида. Под дискриминацией моделей подразумевается выбор одной из двух, или нескольких альтернативных моделей. Был предложен и численно исследован подход, обобщающий подход Стиглера к построению планов для дискриминации полиномиальных регрессионных моделей. Была выбрана более удобная, чем у Стиглера, параметризация. Рассматривался случай полиномиальной модели до 4 степени включительно. Установлено, что развиваемый подход позволяет эффективно проверить гипотезу о том, что верна модель $n$-ой степени при альтернативе в виде модели $n-1$-ой степени, причем при любом решении удается также достаточно эффективно оценить параметры выбранной модели.

# Список литературы


1. Stigler S. M. Optimal experimental design for polynomial regression // Journal of the American Statistical Association. — 1971. — Vol. 66, no. 334. — P. 311–318.
2. Atkinson A. C., Fedorov V. V. The design of experiments for discriminating between two rival models // Biometrika. — 1975. — no. 62. — P. 57 – 70.
3. Dette H., Titoff S. Optimal discrimination designs // Annals of Statistics. — 2009. — no. 37. — P. 2056 — 2082.
4. Dette H., Melas V. B., Shpilev P. T-optimal designs for discrimination between two polynomial models // Annals of Statistics. — 2009. — no. 40(1). — P. 188 – 205.
5. Melas V. B. On standardized maximin designs for discrimination between two polynomial models // Journal of Statistical Theory and Practice. — 2009. — Vol. 7, no. 4. — P. 674 – 686.
6. Studden W. J. $D_s$-Optimal designs for polynomial regression using continued fractional // Mimeograph series. — 1979. — Vol. 79, no. 2. — P. 1–17.
7. Мелас В., Шпилев П. Планирование и анализ для регрессионных моделей. — Санкт-Петербург : Санкт-Петербургский государственный университет, 2012. — 102 с.