Сравнение различных подходов для дискриминации моделей

Кухтина Дарина Александровна, гр. 622

Санкт-Петербургский государственный университет
Кафедра статистического моделирования

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

Рецензент: д.т.н. Григорьев Ю.Д.

Санкт-Петербург
2017г.

$y_j = \eta(x_j, \theta) + \varepsilon_j$, где $j = 1, \ldots, N$, $x_j \in \chi$, $\theta \in \Omega \subset \mathrm{R}^m$
— общее уравнение регрессии.

Задачи:

1. Оценить параметры модели
2. Выбрать одну из конкурирующих моделей

Теория планирования эксперимента: $y_1, \ldots, y_N$ — результаты эксперимента, $\eta(x, \theta)$– функция регрессии, $\theta = (\theta_1, \cdots, \theta_m)$ — параметры, $\varepsilon_1, \ldots, \varepsilon_N$ — ошибки наблюдений, $x = (x_1, \ldots, x_N)$ — условия проведения эксперимента, $x_i \in \chi$. $\mathbb{E}\varepsilon_j = 0$, $\mathbb{E}\varepsilon_j^2 = \sigma^2$, $j = 1, \ldots, m$ $\mathbb{E}\varepsilon_j\varepsilon_i = 0$ при $i \neq j$.

D-оптимальность

План эксперимента — это вероятностная мера на множестве планирования χ:

$$\begin{pmatrix} x_1 & \dots & x_m \\ \omega_1 & \dots & \omega_m \end{pmatrix},$$

причем $\sum_{i=1}^{m} \omega_i = 1$ и $\omega_i \geq 0$.

$f_i(t)$ — непрерывные, линейно независимые функции.
Предполагается, что : $\eta(t, \theta) = \theta^{\mathrm{T}} f(t)$.
Информационная матрица плана: $M(\xi) = \int\limits_{\chi} f(x) f^T(x) \xi dx$.

План ξ^* называется D-**оптимальным**, если
$\det M(\xi^*) \to \max_{\xi \in \Xi}$.

D_s-оптимальность

Хотим найти оптимальный план для оценки только s параметров:

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

$$M(\xi) = \begin{pmatrix} M_{11}(\xi) & M_{12}(\xi)^{\mathrm{T}} \\ M_{12}(\xi) & M_{22}(\xi) \end{pmatrix},$$

Оцениваем s параметров, тогда план, максимизирующий величину $\det M_s(\xi)$ называется усеченным **D_s-оптимальным планом.**

D_s-оптимальность

Информационная матрица представима в виде:

$$M(\xi) = \begin{pmatrix} M_{11}(\xi) & M_{12}(\xi)^{\mathrm{T}} \\ M_{12}(\xi) & M_{22}(\xi) \end{pmatrix}.$$

Матрица $M_{(s)}(\xi)$ определяется следующим образом:

$M_{(s)}(\xi) = M_{11}(\xi) - X^{\mathrm{T}} M_{22}(\xi) X$. Если матрица $M_{22}(\xi)$ невырождена, то $X = M_{22}^{-1} M_{12}$; если матрица $M_{22}(\xi)$ вырождена, то X — произвольное решение системы $M_{22} X = M_{12}$, причем $M_{(s)}(\xi)$ не зависит от выбора решения.

Подход Стиглера

Модель: $P_m = \beta_m x^m + \beta_{m-1} x^{m-1} + \ldots + \beta$

Задачи:

1. Выводы о значении старшего коэффициента β_m;
2. Выводы про модель P_m.

Определение

Будем называть план ξ_0 C-ограниченным D-оптимальным планом для модели P_m, если ξ_0 максимизирует $\det M_{(m)}$ среди всех планов ξ, удовлетворяющих условию:

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

Постановка задачи: Цели

- Исследовать задачу о дискриминации для полиномиальной модели;
- Оптимально оценивать параметры для предполагаемой модели, и для альтернативной;
- Оптимально оценивать старший коэффициент полиномиальной модели;
- Сравнить различные подходы для решения этих задач.

Теорема Стаддена

Теорема(Stadden, 1979 г.)

Для полиномиальной модели на отрезке $[-1, 1]$ непрерывный D-оптимальный план существует, единственен и сосредоточен с равными весами в m точках, которые являются корнями $(x^2 - 1)P'_{m-1}$, где $P_{m-1}(x)$ — полином Лежандра порядка $m - 1$.

Точки усеченного D_s-оптимального плана ξ, состоят из -1 и 1 и $m - 1$ корней $P'_s(x)U_{m-s}(x) - P'_{s-1}U_{m-s-1}(x) = 0$, где $s = 0, 1, \ldots, m - 1$, P_i — многочлены Лежандра, U_i — многочлены Чебышева второго рода.

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

Понятие эффективности

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

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

s — количество оцениваемых параметров модели, ξ_2 — D-оптимальный план и ξ_1 — D_s-оптимальный план.
Эффективность D_s-оптимального плана, относительно D-оптимального:

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

m — количество параметров модели.

Сравнение эффективности

Таблица: Эффективность D-оптимального плана, относительно D_s-оптимальных для полиномиальных моделей

Планы	Степень 2	Степень 3	Степень 4
D_1	0.54	0.10	0.03
D_2	1	0.17	0.13
D_3	—	1	0.25
D_4	—	—	1

Сравнение эффективности

Таблица: Эффективность D_s-оптимальных планов, относительно D-оптимального плана

Планы	Степень 2	Степень 3	Степень 4
D_1	0.95	0.94	0.93
D_2	1	0.95	0.96
D_3	—	1	0.98
D_4	—	—	1

Обобщение критерия Стиглера

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 m и $m+1$.
$d_1(\xi, x) = f^T_{(m)}(x) M_1^{-1}(\xi) f_{(m)}(x)$,
$d_2(\xi, x) = f^T_{(m+1)}(x) M_2(\xi)^{-1} f_{(m+1)}(x)$.

Определение

План ξ G_γ-оптимальный, если $\xi = \arg\min_\xi \max_x d_\gamma(x, \xi)$, где $d_\gamma(x,\xi) = \gamma d_1(x,\xi) + (1-\gamma) d_2(x,\xi)$.

Определение

План ξ Ψ_γ-оптимальный, если $\xi = \arg\max_x \Psi_\gamma(\xi)$.

Теорема эквивалентности для Ψ_γ-критерия

Теорема (Эквивалентности для Ψ_γ-критерия)

Рассмотрим полиномиальные модели порядка m и $m+1$ на стандартном отрезке [-1, 1]. Следующие условия эквивалентны:

(a) *план ξ^* — Ψ_γ-оптимальный;*

(b) *план ξ^* — G_γ-оптимальный;*

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

Существует единственный Ψ_γ-оптимальный план . Этот план сосредоточен в m+1 точке, причем концы отрезка являются опорными точками плана и все опорные точки плана и их веса симметричны относительно начала.

Обобщенные планы для квадратичной модели

Модель:

$\theta_1 + \theta_2 x + \theta_3 x^2$, где θ_i – оцениваемые параметры,а $x \in [-1, 1]$.

Задача: $\gamma \ln \det M_1(\xi) + (1-\gamma) \ln \det M_2(\xi) \to max$ по планам ξ 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.

Обобщенные планы для кубической модели

Модель:

$\theta_1 + \theta_2 x + \theta_3 x^2 + \theta_4 x^3$, где θ_i — оцениваемые параметры,а $x \in [-1, 1]$.

Задача: $\gamma \ln\det M_2(\xi) + (1-\gamma) \ln\det M_3(\xi) \to max$ по планам ξ 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.

Обобщенные планы для полиномиальной модели 4-ой степени

Модель:

$\theta_1 + \theta_2 x + \theta_3 x^2 + \theta_4 x^3 + \theta_5 x^4$, где θ_i – оцениваемые параметры,а $x \in [-1, 1]$.

Задача: $\gamma \ln\det M_3(\xi) + (1-\gamma)\ln\det M_4(\xi) \to max$ по планам ξ вида:

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

где M_4 — информационная матрица для модели 4-ой степени, M_3 — информационная матрица для кубической модели.

Сравнение эффективности для квадратичной модели

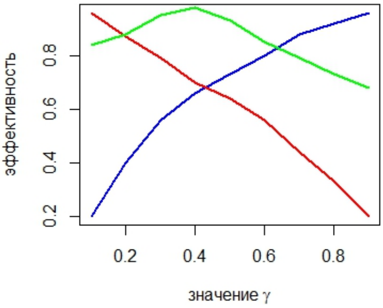

Рис.: Эффективность обобщенно оптимальных планов, по отношению к D-оптимальным и D_s-оптимальным.

Таблица: Эффективность, в зависимости от γ

γ	x^2	x	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

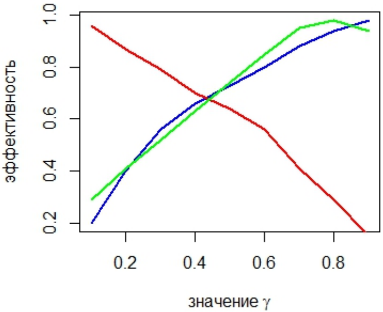

Рис.: Эффективность обобщенно оптимальных планов, по отношению к D-оптимальным и D_s-оптимальным.

Таблица: Эффективность, в зависимости от γ

γ	x^3	x^2	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

Сравнение эффективности полиномиальной модели 4-ой степени

Таблица: Эффективность обощенно оптимальных планов, в зависимости от γ

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

Результаты

- Был предложен и численно исследован подход, обобщающий подход Стиглера к построению планов для дискриминации полиномиальных регрессионных моделей.
- Было проведено сравнение D-оптимальных и D_s оптимальных планов для полиномиальных моделей.
- Численно построены обобщенно оптимальные планы для моделей, до 4-ой степени включительно.
- Установлено, что данный подход позволяет эффективно проверить гипотезу о том, что верна модель $m+1$-ой степени при альтернативе в виде модели m-ой степени, причем при любом решении удается также достаточно эффективно оценить параметры выбранной модели.

СПАСИБО ЗА ВНИМАНИЕ!