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.

Уравнение регрессии

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

- 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 $\mathbb{E}\varepsilon_i^2 = \sigma^2$ для всех i

План эксперимента и информационная матрица

Определение

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

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

C-оптимальный план эксперимента

Определение

C-оптимальным планом эксперимента для данного вектора c называется план минимизирующий функцию Φ

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

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

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

Постановка задачи

Определение

Если $c = f'(z) = (f_1'(z), \dots, f_m'(z))^\top$ *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 Элвинга [Dette et al., 2010]

Теорема

Допустимый план $\xi^\star$ *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:*

$$\left|p^\top f(x_i)\right| = 1 \qquad i = 1..m \leqslant n, \tag{1a}$$

$$\left|p^\top f(x)\right| \leqslant 1 \qquad x \in \mathcal{X}, \tag{1b}$$

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

Кроме того,

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

Веса у плана для оценки производной [Dette et al., 2019]

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

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

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:

$$\omega_i = \frac{|L_i'(z)|}{\sum_{j=1}^m \left|L_j'(z)\right|}.$$

Носитель плана

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 m. Таким образом носитель плана находится в точках

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

Корни производных базисных многочленов

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

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

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

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:*

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

Пример $m = 3$

С помощью этой теоремы найдем планы для оценки производной в случае $f(x) = (x, x^2, x^3)^\top$, $\mathcal{X} = [0, 1]$ и покажем их оптимальность. Опорные точки плана будут совпадать с экстремальными точками многочлена $S_3(x)$:

$$S_3(x) = 4\left(\frac{\sqrt{3}}{2}+1\right)^3 x^3 - 6\sqrt{3}\left(\frac{\sqrt{3}}{2}+1\right)^2 x^2 + 6\left(\frac{\sqrt{3}}{2}+1\right)x,$$

Которыми являются

$$x_1^* = 3\sqrt{3} - 5, \quad x_2^* = \sqrt{3} - 1, \quad x_3^* = 1.$$

Производные базисных многочленов $(m = 3)$

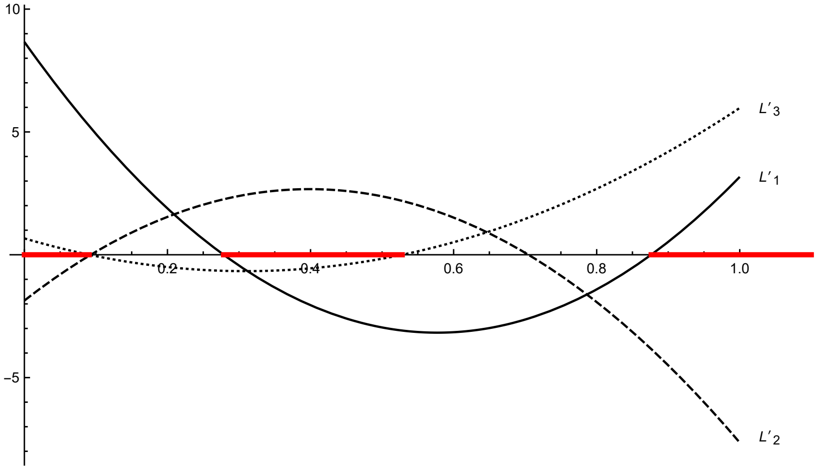

Рис.: Производные базисных многочленов. Красным выделены промежутки, на которых выполняется теорема.

Численное нахождение оптимальных планов

- С-оптимальность плана определяется как решение задачи оптимизации
- Теорема Элвинга дает критерий С-оптимальности
- Поэтому численный алгоритм может быть устроен следующим образом:
  1. Численно оптимизируем функцию $\Phi(\xi)$ из определения со случайным начальным планом
  2. Проверяем выполнение условий теоремы Элвинга
     - Если они выполнены — у нас есть результат
     - Если нет — возвращаемся к п. 1

Оптимизация $\Phi(\xi)$

- У любой матрицы $A^{-}(x)$ при условии постоянного ранга существует непрерывная вторая производная выражающаяся через производные $A(x)$
- Так как $\Phi(\xi) = c^{\top} M^{-}(\xi) c$, а ранг $M(\xi)$ зависит только от количества опорных точек, то существуют вторые производные у Φ по опорным точкам и весам
- Поэтому был использован квазиньютоновский алгоритм оптимизации с ограничениями L-BFGS-B

Проверка условий теоремы Элвинга

Нужно найти такой вектор p, что $\left|p^\top f(x)\right| \leqslant 1$, при этом равенство достигается в опорных точках. Такие p будут решением уравнения

$$\begin{pmatrix} f_1(x_1) & \dots & f_n(x_1) \\ f_1(x_2) & \dots & f_n(x_2) \\ \vdots & \ddots & \vdots \\ f_1(x_n) & \dots & f_n(x_n) \\ f_1'(x_{i_1}) & \dots & f_n'(x_{i_1}) \\ \vdots & \ddots & \vdots \\ f_1'(x_{i_m}) & \dots & f_n'(x_{i_m}) \end{pmatrix} p = \begin{pmatrix} 1 \\ s_2 \\ \vdots \\ s_n \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

где x_{i_j} — опорные точки не на границе области планирования, а $s_i = \pm 1$.

Проверка условий теоремы Элвинга

Для всех p с нулевой невязкой для предыдущего уравнения и данного плана ξ, проверяем третье условие теоремы Элвинга:

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

- Для того, чтобы не вычислять h достаточно проверять коллинеарность векторов
- Все сравнения с точностью до машинного нуля
- Если равенство выполняется, то ξ — c-оптимальный план

Результаты

- Описаны оптимальные планы размера n для нахождения производной в полиномиальной модели без свободного члена с $\mathcal{X} = [0, 1]$
- Приведен пример применения этого результата
- Разработан алгоритм численного нахождения c-оптимальных планов в общем случае