Построение и исследование локально-оптимальных планов на основе функционального подхода

Ким Эрик Евгеньевич, гр.18.Б04-мм

Уровень образования: бакалавриат
Направление 01.03.02 «Прикладная математика и информатика»
Основная образовательная программа СВ.5004.2018 «Прикладная математика и информатика»
Профессиональная траектория «Вычислительная стохастика и статистические модели»

Научный руководитель: д. ф.-м. н., профессор Мелас В.Б.,
Рецензент: доктор технических наук, профессор Григорьев Ю. Д.

Санкт-Петербург
2022г.

Введение

- **Актуальная задача:** задача нахождения локально D-оптимального плана
- **В некоторых случаях** возможно построить явные формулы для планов во всей допустимой области
- **Иной подход** — функциональный подход, подход к исследованию и построению локально-оптимальных планов, основанный на представлении точек плана как неявно заданных функций параметров.
- **Эмпирическая оценка радиуса сходимости** определяет область действия формул, полученных на основе этого подхода, с заданной точностью.

Постановка задачи

Задача:

- Применить функциональный подход к задаче нахождения локально D-оптимального плана для экспоненциальной и дробно-рациональной моделей с 4 параметрами и построить соответствующие формулы для планов
- Нахождение эмпирической оценки радиуса сходимости ряда Тейлора неявной функции параметров для полученных формул

Задача нахождения локально D-оптимального плана

Рассмотрим нелинейную по параметрам регрессионную модель

$$\eta(x,\theta) = \sum_{i=1}^{k} \theta_i \eta_i(x, \theta_{i+k}),$$

где $\theta = (\theta_1, \ldots, \theta_{2k})$ — вектор параметров.

Определение

Приближенным планом эксперимента называется дискретная вероятностная мера

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

где n — некоторое натуральное число, $\omega_i > 0, x_i \in X, i = 1, \ldots, n$ и $\sum_{i=1}^{n} \omega_i = 1, x_1 < \ldots < x_n, X \subset \mathbb{R}$ — множество планирования.

Определение

Пусть $f(x,\theta) = \left[\frac{\partial\eta(x,\theta)}{\partial\theta_1}, \dots, \frac{\partial\eta(x,\theta)}{\partial\theta_{2k}}\right]^{\mathbb{T}}$. Тогда информационной матрицей плана называется

$$M(\xi,\theta) = \int_X f(x,\theta) f^{\mathbb{T}}(x,\theta) d\xi(x).$$

Задача нахождения локально D-оптимального плана

$$(\det M(\xi,\theta))^{\frac{1}{2k}} \to \max_{\xi\in\Xi,\theta=\theta_0},$$

где θ_0 — некий заданный вектор параметров, $M(\xi,\theta)$ — информационная матрица приближенного плана ξ для данной функции регрессии,
Ξ — некоторое множество приближенных планов эксперимента.

Подход к решению данной задачи — функциональный подход.

Функциональный подход (Melas 2006)

Пусть задана функция $g(\tau, z) : \mathbb{R}^{s+k} \to \mathbb{R}^s$
непрерывно-дифференциируемая в окрестности $U \subset \mathbb{R}^s \times \mathbb{R}^k$.
Необходимо решить относительно τ при различных z:

$$g(\tau, z) = 0$$

По теореме о неявной функции при выполнении следующих условий:

- $g(\tau_0, z_0) = 0$
- $\det J \neq 0$, где $J(\tau, z) = \left(\frac{\partial}{\partial \tau_i} g_j(\tau, z) \right)_{i,j=1}^{k} \Big|_{\tau=\tau_0, z=z_0}$
- g — вещественно-аналитическая

в некоторой окрестности $W \subset U$ задана неявная функция, то есть существует единственная непрерывная функция $\tau = \tau(z)$ такая, что $g(\tau, z) = 0$ при $\tau = \tau(z)$, более того — она вещественно-аналитическая.

Рекуррентные формулы для коэффициентов ряда Тейлора

Пусть $l = (l_1, \dots, l_k), l_i \geq 0$ — заданный вектор целых чисел. Предположим, что функция g имеет вид

$$g(\tau, z) = (z_1 - z_{1(0)})^{l_1} \dots (z_k - z_{k(0)})^{l_k} \tilde{g}(\tau, s),$$

причём $\det \tilde{J}(\tau_{(0)}, z_{(0)}) \neq 0$ и $\det J(\tau_{(0)}, z_{(0)}) = 0$, где $\tilde{J}(\tau_{(0)}, z_{(0)}) = \frac{\partial}{\partial \tau} \tilde{g}(\tau, z)$.
Тогда (Melas 2006) при $z_i \neq z_j (i \neq j), z_i \neq z_{i(0)}$, $i, j = 1, \dots, k$ неявно задана вещественно-аналитическая функция $\tau^*(z)$, удовлетворяющая исходному уравнению.

$$\tau^*(z) = \tau_{(0)} + \sum_{t=1}^{\infty} \sum_{s \in S_t} \tau^*_{(s)} (z_1 - z_{1(0)})^{s_1} \dots (z_k - z_{k(0)})^{s_k}.$$

Рекуррентные формулы для коэффициентов ряда Тейлора

Пусть I — произвольное множество индексов вида $s = (s_1, \ldots, s_k)$, $s_i \geq 0$, $i = 1, \ldots, k$. Обозначения:

- $\tau^*_{<I>}(z) = \sum_{s \in I} \tau^*_{(s)} (z_1 - z_{1(0)})^{s_1} \ldots (z_k - z_{k(0)})^{s_k}$,
- $f_{(s)} = (f(z))_{(s)} = \frac{1}{s_1! \ldots s_k!} \frac{\partial^{s_1}}{\partial z_1^{s_1}} \ldots \frac{\partial^{s_k}}{\partial z_k^{s_k}} f(z) \Big|_{z=z_0}$,
- $J_{(l)} = (J(\tau_{(0)}, z))_{(l)}$,
- $S_t = \{s = (s_1, \ldots, s_k); s_i \geq 0, \sum_{i=1}^{k} s_i = t\}, t = 0, 1, \ldots$

В работе В.Б.Меласа (Melas 2006) были введены рекуррентные формулы для коэффициентов ряда Тейлора:

Рекуррентные формулы для коэффициентов ряда Тейлора

$$\tau^*_{(s)} = -J^{-1}_{(l)} g(\tau^*_{<I>}(z), z)_{(s+l)}, \text{ где } s \in S_t, I = \cup_{j=0}^{t-1} S_j, t = 1, 2, \ldots$$

Эмпирическая оценка радиуса сходимости

- Ряды Тейлора неявной функции — явные формулы в окрестности теоретического радиуса сходимости ряда.
- Отрезки ряда длины n являются ими с заданной точностью δ в окрестности точки z_0, характеризующейся **эмпирической оценкой радиуса сходимости** — $r(\delta, n, z_0)$.

Обозначим для некоторой вектор-функции $f(z)$ в качестве i-ой компоненты — $f^{(i)}(z)$.

$$r(\delta, n, z_0) =$$
$$= \max\{||z - z_0|| : \forall z^* \in U, ||z^* - z_0|| \leq ||z - z_0|| : |\tau^{(i)}_{<n>}(z^*) - \tau^{*(i)}(z^*)| < \delta, \quad i = 1, \ldots, s\},$$

где $\tau_{<n>}(z)$ — отрезок из первых $n+1$ членов ряда Тейлора около точки z_0, $\tau^*(z)$ — неявная функция параметра.

Ключевая задача: найти эмпирическую оценку радиуса сходимости

Рассмотрим функцию регрессии вида:

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

где $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ — вектор параметров, $x \in [0; d), d$ — велико.

По теореме из работы (Мелас 1999) вид локально D-оптимального плана:

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

где при $\lambda = -(\theta_2+\theta_4+3) - \sqrt{(\theta_2+\theta_4+3)^2+24}$:

$$x_1 = 0, x_2 = \frac{\sqrt{\theta_2\theta_4}}{2}\left(-\frac{\lambda}{2} - 1 + \sqrt{(\lambda/2+1)^2-4}\right),$$

$$x_3 = \sqrt{\theta_2\theta_4}, x_4 = \frac{\sqrt{\theta_2\theta_4}}{2}\left(-\frac{\lambda}{2} - 1 - \sqrt{(\lambda/2+1)^2-4}\right).$$

Известный результат:
Задача нахождения локально D-оптимального плана для данной модели эквивалентна (Мелас 1999):

$$g_i(x,\theta) = \frac{1}{x_i} + \sum_{i \neq j} \frac{1}{x_i - x_j} - 2\frac{Q'(x_i)}{Q(x_i)} = 0, i = 2,3,4,$$

где $Q(y) = (y+\theta_2)(y+\theta_4), x = (x_2, x_3, x_4), \theta = (\theta_2, \theta_4)$.
Осуществим замену переменных $a = \theta_2 + \theta_4, b = \theta_2\theta_4$ и примем $b = 1$:

$$g_i(x,a) = \frac{1}{x_i} + \sum_{i \neq j} \frac{1}{x_i - x_j} - 2\frac{2x_i + a}{x_i^2 + ax_i + 1} = 0, i = 2,3,4.$$

Применение функционального подхода

Полученные результаты:
Используя явные формулы, получим, что для $a^* = 5$ точки плана равны $x^* = (7.5581, 1, 0.1323)$. Заметим, что для x^*, a^* справедливо:

1. $g_i(x^*, a^*) = 0, i = 2, 3, 4$ — по выбору x^*;
2. $\det J(x^*, a^*) = \det \frac{\partial}{\partial x} g(x, a)\Big|_{x=x^*, a=a^*} \neq 0$

То есть можно применить функциональный подход. Разложение неявной функции в ряд Тейлора:

$$\tilde{x}(a) = x_{(0)} + \sum_{i=1}^{\infty} x_{(i)} (a - a^*)^i.$$

По рекуррентным формулам были подсчитаны первые 10 коэффициентов ряда Тейлора функции $\tilde{x}(a)$ около точки a^*.

Исследование для дробно-рациональной модели

Наличие явной формулы для нахождения опорных точек плана дает возможность найти эмпирическую оценку радиуса сходимости.

$n = 10$, Точность $\delta = 0.0001$, обозначение: $f(a)$ — явное отображение, сопоставляющее параметру a ненулевые опорные точки локально D-оптимального плана, $a^* = 5$.

- $a = 9$

$$f(a) = \begin{pmatrix} 11.3930 \\ 1 \\ 0.0878 \end{pmatrix}, x_{<10>}(a) = \begin{pmatrix} 11.3930 \\ 1 \\ 0.0877 \end{pmatrix}.$$

  Получен результат с точностью 0.0001.

- $a = 8.9$

$$f(a) = \begin{pmatrix} 11.2960 \\ 1 \\ 0.0885 \end{pmatrix}, x_{<10>}(a) = \begin{pmatrix} 11.2960 \\ 1 \\ 0.0885 \end{pmatrix}.$$

- $a = 1.1$

$$f(a) = \begin{pmatrix} 3.9937 \\ 1 \\ 0.2504 \end{pmatrix}, x_{<10>}(a) = \begin{pmatrix} 3.9939 \\ 1 \\ 0.2503 \end{pmatrix}.$$

Получен результат с точностью 0.0002. Продолжаем сужать радиус.

- $a = 1.3$

$$f(a) = \begin{pmatrix} 4.1694 \\ 1 \\ 0.2398 \end{pmatrix}, x_{<10>}(a) = \begin{pmatrix} 4.1694 \\ 1 \\ 0.2398 \end{pmatrix}.$$

Получен результат с необходимой точностью.

В результате: $r(\delta, n, a^*) \approx 3.7$.

Экспоненциальная модель

Рассмотрим функцию регрессии вида:

$$\eta(x,\theta) = \theta_1 e^{-\theta_3 x} + \theta_2 e^{-\theta_4 x}, \text{ где } x \in X, X = [0; +\infty).$$

Пусть $\theta_1 \neq 0, \theta_2 \neq 0, \theta_2 > \theta_1, \theta_i > \epsilon, i = 1 \ldots k$, где ϵ — некоторое бесконечно малое число.

Вид локально D-оптимального плана известен из работы (Melas 2006):

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

Более того он не зависит от параметров, линейно входящих в рассматриваемую модель.

Применение функционального подхода

Известные результаты:

- Задача нахождения локально D-оптимального плана эквивалентна системе ур-й (Melas 2006):

$$g_i(\tau, z) = \frac{\partial}{\partial \tau_i} \det \mathrm{F}(\zeta_\tau, z) = 0, i = 1, 2, 3$$

где при $\tau_0 = 0, (\tau_1, \tau_2, \tau_3) = (x_2, x_3, x_4), z = (z_1, z_2) = (\theta_3, \theta_4)$

$$\mathrm{F}(\zeta_\tau, z) = \left(e^{-z_1 \tau_s}, -\tau_s e^{-z_1 \tau_s} e^{-z_2 \tau_s} - \tau_s e^{-z_2 \tau_s}\right)_{s=0}^{3}$$

- Известно решение при $z_1 \to \alpha, z_2 \to \alpha, \alpha > 0$ (там же): $\tau_0 = (\frac{\gamma_1}{2\alpha}, \frac{\gamma_2}{2\alpha}, \frac{\gamma_3}{2\alpha})$, где $\gamma_1, \gamma_2, \gamma_3$ — корни многочлена Лаггера 3-его порядка с параметром 0

Полученные результаты:

Осуществим замену переменных $z^* = \frac{z_1 - z_2}{2}$, приняв $z_1 + z_2 = 2$. Тогда система уравнений примет вид:

$$g_i^*(\tau, z^*) = \frac{\partial}{\partial \tau_i} \det \mathrm{F}(\zeta_\tau, 1 + z^*, 1 - z^*) = 0, i = 1, 2, 3,$$

$$z^* = 0 \Longleftrightarrow z_1 = 1, z_2 = 1.$$

Примем $z_0^* = 0$.

1. Вектор-функция $g^*(\tau, z)$ представима в следующем виде:

$$g^*(\tau, z^*) = z^{*4} \tilde{g^*}(\tau, z^*).$$

2. $\det J(\tau, z_0^*) = \det \frac{\partial}{\partial \tau} g(\tau, z_0^*) = 0.$
3. $\det \tilde{J}(\tau_0, z_0^*) = \det \frac{\partial}{\partial \tau} \tilde{g}(\tau_0, z_0^*) \neq 0.$

Можно применить функциональный подход и рекуррентные формулы. По ним были найдены первые 7 ненулевых коэфф-тов ряда Тейлора.

Исследование для экспоненциальной модели

В предыдущих работах (Мелас 1981) были найдены численными методами локально D-оптимальные планы для исследуемой модели для различных значений параметра $z^* = \frac{z_1 - z_2}{2}$.

Пусть $n = 12$, точность $\delta = 0.02$, обозначим τ_{z^*} — известные значения ненулевых точек локально D-оптимального плана, соответствующего параметру z^*, $z_0^* = 0$. Значения точек плана даны с точностью в 2 знака после запятой.

- $z^* = 0.1$

$$\tau_{z^*} = \begin{pmatrix} 0.47 \\ 1.66 \\ 3.91 \end{pmatrix}, \tau^*_{<12>}(z^*) = \begin{pmatrix} 0.47 \\ 1.66 \\ 3.90 \end{pmatrix}.$$

Получен результат с необходимой точностью.

- $z^* = 0.2$
$$\tau_{z^*} = \begin{pmatrix} 0.47 \\ 1.66 \\ 3.97 \end{pmatrix}, \tau^*_{<12>}(z^*) = \begin{pmatrix} 0.47 \\ 1.67 \\ 3.96 \end{pmatrix}.$$
Получен результат с необходимой точностью.
- $z^* = 0.3$
$$\tau_{z^*} = \begin{pmatrix} 0.48 \\ 1.69 \\ 4.08 \end{pmatrix}, \tau^*_{<12>}(z^*) = \begin{pmatrix} 0.47 \\ 1.69 \\ 4.08 \end{pmatrix}.$$
Получен результат с необходимой точностью.
- $z^* = 0.4$
$$\tau_{z^*} = \begin{pmatrix} 0.48 \\ 1.70 \\ 4.23 \end{pmatrix}, \tau^*_{<12>}(z^*) = \begin{pmatrix} 0.47 \\ 1.72 \\ 4.26 \end{pmatrix}.$$
Результат с необходимой точностью не получен.

Таким образом, $r(\delta, n, z_0^*) \approx 0.3$.

Заключение

- Применен функциональный подход к задаче нахождения локально D-оптимального плана для двух моделей с 4 параметрами и построены формулы для опорных точек плана в виде отрезков ряда Тейлора неявных функций около некоторой точки
- Найдена эмпирическая оценка радиуса сходимости для данных формул с заданной точностью δ около некоторых заданных точек
- Полученные результаты показывают возможность использования приближений неявных функций, полученных в результате применения функционального подхода, как явных формул для построения локально D-оптимальных планов в нелинейных по параметрам регрессионных моделях