Применение Функционального подхода для построения локально-оптимальных планов эксперимента

Ярославцева Ирина, гр. 20.Б04-мм

Санкт-Петербургский государственный университет
Математико-механический факультет
Вычислительная стохастика и статистические модели

Научный руководитель: д.ф.-м.н., профессор Мелас В. Б.
Рецензент: к.ф.-м.н., лектор Пепелышев А. Н.

Санкт-Петербург
2024г.

Введение

Пусть результаты эксперимента y_i описываются моделью:

$$y_i = \eta(x_i, \theta) + \varepsilon_i, \quad i = 1, \ldots, n,$$

- $\eta(x, \theta)$ — функция регрессии,
- $\theta = (\theta_1, \ldots, \theta_m)$ — вектор неизвестных параметров,
- $x_1, \ldots, x_n$ — условия проведения эксперимента,
- $\varepsilon_1, \ldots, \varepsilon_n$ — случайные ошибки наблюдений.

Линейные и нелинейные модели

- Параметр (скалярный) входит в модель линейным образом, если функция регрессии дифференцируема по этому параметру, и производная от него не зависит
- Иначе говорим, что параметр входит в модель нелинейно
- Модель называется линейной (по параметрам), если все параметры входят в неё линейным образом
- Иначе модель называется нелинейной (по параметрам)

- Необходимо оценивать значения неизвестных параметров $\theta_1, \ldots, \theta_m$, так как от них зависят результаты эксперимента
- Если эксперимент требует больших ресурсов, возникает необходимость разрабатывать оптимальные планы проведения эксперимента

Планом эксперимента называется дискретная вероятностная мера:

$$\xi = \begin{pmatrix} x_1, & \ldots, & x_n \\ \omega_1 & \ldots, & \omega_n \end{pmatrix}, \quad x_i \in \mathcal{X}, \quad i = 1, \ldots, n,$$

причем $x_i \neq x_j$ при $i \neq j$ — опорные точки, ω_i — веса, $\omega_i > 0$ и $\sum_{i=1}^{n} \omega_i = 1$, $\mathcal{X}$ — множество планирования.

Информационной матрицей плана называется

$$M(\xi, \theta) = \int_{\mathcal{X}} f(x, \theta) f^T(x, \theta) d\xi(x),$$

где $f(x, \theta) = \left(\frac{\partial \eta(x,\theta)}{\partial \theta_1}, \ldots, \frac{\partial \eta(x,\theta)}{\partial \theta_m} \right)^T$ — вектор частных производных по параметрам.

Локально D-оптимальным планом называется план ξ^*, который максимизирует величину:

$$\det M(\xi, \theta)$$

при некотором заданном векторе параметров θ_0.

Проблема оптимального планирования

- Оптимальное планирование для нелинейных регрессионных моделей остается недостаточно изученным
- В основном исследуются простые модели, для которых возможно построить оптимальные планы в явном виде
- Для остальных моделей приходится строить оптимальные планы численно
- Функциональный подход является методом аппроксимации точек локально-оптимальных планов эксперимента

Цель работы

- Применить функциональный подход для полиномиальной регрессионной модели линейной по параметру
- С помощью функционального подхода построить оптимальные планы для дробно-рациональной модели
- Исследовать эффективность функционального подхода путем сравнения планов, полученных с его помощью, с планами полученными в явном виде и численно

Теорема о неявной функции [Бибиков Ю.Н., 1991]

Пусть $g(x,z) : U \to \mathbb{R}^k$, где $U \subset \mathbb{R}^s \times \mathbb{R}^k$, причем в точке $(x_0, z_0) \in U$ выполнено:

- $g \in C^1(U)$ и $g(x_0, z_0) = 0$
- $\det J \neq 0$, где $J(x,z) = \left(\frac{\partial}{\partial x_i} g_j(x,z) \right)_{i,j=1}^{k} \Big|_{x=x_0, z=z_0}$

тогда в окрестности $W \subset U$ задана неявная функция, т.е. существует единственная непрерывная функция $x = x(z)$ такая, что $g(x(z), z) = 0$.

Более того, если $g(x,z)$ — вещественно-аналитическая, то и $x(z)$ вещественно-аналитическая функция.

Разложение в ряд Тейлора и рекуррентные формулы

Аналитичность функции означает, что она разложима в сходящийся ряд Тейлора. Пусть $x = (x_1, \ldots, x_s) \in \mathbb{R}^s$, $z \in \mathbb{R}$, тогда разложение принимает вид

$$x(z) = x_0 + \sum_{i=1}^{\infty} x_{(i)} (z - z_0)^i$$

Утверждение [Мелас В.Б., 2006]

Для коэффициентов разложения верны следующие рекуррентные формулы

$$x_{(n)} = -J_{(0)}^{-1} \frac{\partial^n}{\partial z^n} g\left(x_{<n-1>}(z), z\right)\Big|_{z=z_0}, \quad n = 1, 2, \ldots,$$

где $J_{(0)} = J(x_0, z_0)$, $x_{<n-1>}(z) = x_0 + \sum_{i=1}^{n-1} x_{(i)} (z - z_0)^i$.

Полиномиальная модель

Во многих случаях точки оптимальных планов оказываются корнями некоторых многочленов. Применим функциональный подход к задаче нахождения корней:

$$\eta(x,\theta) = x^3 + x^2 - 2x + \theta$$

В окрестности $\theta_0 = 0$, $x_0 = 1$ выполнены условия теоремы о неявной функции. Построим разложение

$$x_{<7>}(\theta) = 1 - \frac{1}{3}\theta - \frac{4}{27}\theta^2 - \frac{29}{243}\theta^3 - \frac{260}{2187}\theta^4 - \\ - \frac{2603}{19683}\theta^5 - \frac{9296}{59046}\theta^6 - \frac{1026217300}{5229910881}\theta^7$$

Оценка точности результатов

Ключевая задача: сколько членов разложения нужно взять, чтобы корни, вычисленные с помощью разложения и по формулам Кардано, отличались не более чем на 0.001?

По формулам Кардано

- для параметра $\theta = 0.5$ корень $x = 0.762750$:

$$x_{<1>}(\theta) \approx 0.833333 \Rightarrow |0.762750 - 0.833333| = 0.0705 > 0.001$$
$$x_{<7>}(\theta) \approx 0.76175 \Rightarrow |0.762750 - 0.76175| = 0.001$$

- для параметра $\theta = 0.2$ корень $x = 0.926231$:

$$x_{<1>}(\theta) \approx 0.933333 \Rightarrow |0.926231 - 0.933333| = 0.007 > 0.001$$
$$x_{<3>}(\theta) \approx 0.926452 \Rightarrow |0.926231 - 0.926452| = 0.0002 < 0.001$$

Дробно-рациональная модель

Дробно-рациональная регрессионная модель:

$$\eta(x,\theta) = \frac{\theta_1}{x+\theta_2} + \frac{\theta_3}{x+\theta_4}, \quad \theta_2, \theta_4 > 0$$

где $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ — вектор параметров, $x \in \mathcal{X} = [0, d)$.

Теорема о виде плана [Мелас В.Б., 1999]

Локально D-оптимальный план для дробно-рациональной модели с 4-мя параметрами имеет вид:

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

Случаи большого и малого промежутков

Множество планирования $\mathcal{X} = [0, d)$.

Если d велико, а именно $d > x_4^*$, где x_4^* определено ниже, можно получить оптимальные планы в явном виде по формулам [Мелас В.Б., 1999]:

$$x_1^* = 0, \quad x_3^* = \sqrt{\theta_2 \theta_4}$$

$$x_{2,4}^* = \frac{\sqrt{\theta_2 \theta_4}}{2} \left(-\frac{\lambda}{2} - 1 \pm \sqrt{(\frac{\lambda}{2} + 1)^2 - 4} \right)$$

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

Если d мало, а именно $d \leq x_4^*$, получить оптимальные планы в явном виде не удается. Применим функциональный подход.

Построение плана для случая малого промежутка

Оптимальный план для случая малого промежутка имеет вид:

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

Определитель информационной матрицы принимает вид:

$$\det M = C \times \frac{x_2 \cdot x_3 \cdot (x_3 - x_2)(d - x_2)(d - x_3)}{(x_2^2 + a x_2 + 1)^2 \cdot (x_3^2 + a x_3 + 1)^2}, \quad a = \theta_2 + \theta_4$$

В результате численной оптимизации определителя с помощью алгоритма BFGS на промежутке $[0, 2)$ при $a = 2$ была получена точка $x^* = (x_2^*, x_3^*) = (0.175430, 0.747127)$.

Функциональный подход для случая малого промежутка

Задача максимизации определителя информационной матрицы сводится к системе уравнений [Мелас В.Б., 1999]:

$$g_i = \frac{1}{x_i} + \sum_{i \neq j} \frac{1}{x_i - x_j} - 2\frac{2x_i + a}{x_i^2 + ax_i + 1} = 0, \quad i = 2, 3$$

Для численно найденной точки $x^* = (0.175430, 0.747127)$ и $a^* = 2$ выполнены условия теоремы о неявной функции:

- $g_i(x^*, a^*) = 0$, при $i = 2, 3$
- $\det J(x^*, a^*) \neq 0$

Можем применить функциональный подход.
С помощью программы на языке Python были подсчитаны первые 10 коэффициентов разложения.

Результаты для малого промежутка

Таблица: Коэффициенты ряда Тейлора в окрестности $a^* = 2$, $d = 2$

$x_{(0)}$	$x_{(1)}$	$x_{(2)}$	$x_{(3)}$	$x_{(4)}$
0.175430 0.747127	-0.035319 -0.048566	0.007203 0.005446	-0.001464 -0.000713	0.000295 9.9047e-5
$x_{(5)}$	$x_{(6)}$	$x_{(7)}$	$x_{(8)}$	$x_{(9)}$
-5.9200e-5 -1.3802e-5	1.1788e-5 1.8303e-6	-2.3373e-6 -2.0891e-7	4.6240e-7 1.3413e-8	-9.1513e-8 2.7071e-9

Исследование точности полученных результатов

Пусть x_h — результат численного нахождения опорных точек плана, $x_{<n>}$ — разложение в окрестности $a^* = 2$ до n-го члена.

Какое число членов разложения необходимо, чтобы планы отличались не более, чем на $\varepsilon = 0.0001,\ 0.01$?

- при $a = 4$

$$x_{<7>}(a) = \begin{pmatrix} 0.1251 \\ 0.6673 \end{pmatrix}, \quad x_h = \begin{pmatrix} 0.1251 \\ 0.6673 \end{pmatrix}$$

  заданная точность $\varepsilon = 0.0001$ достигается при $n = 7$.

- при $a = 4$

$$x_{<2>}(a) = \begin{pmatrix} 0.1336 \\ 0.6717 \end{pmatrix}, \quad x_h = \begin{pmatrix} 0.1251 \\ 0.6673 \end{pmatrix}$$

  заданная точность $\varepsilon = 0.01$ достигается при $n = 2$.

Исследование точности полученных результатов

- при $a = 3$

$$x_{<4>}(a) = \begin{pmatrix} 0.1461 \\ 0.7033 \end{pmatrix}, \quad x_h = \begin{pmatrix} 0.1461 \\ 0.7033 \end{pmatrix}$$

заданная точность $\varepsilon = 0.0001$ достигается при $n = 4$.

- при $a = 3$

$$x_{<1>}(a) = \begin{pmatrix} 0.1401 \\ 0.6985 \end{pmatrix}, \quad x_h = \begin{pmatrix} 0.1461 \\ 0.7033 \end{pmatrix}$$

заданная точность $\varepsilon = 0.01$ достигается при $n = 1$.

Результаты для большого промежутка

Таблица: Коэффициенты ряда Тейлора в окрестности $a^* = 5$

$x_{(0)}$	$x_{(1)}$	$x_{(2)}$	$x_{(3)}$	$x_{(4)}$
7.5581	-0.016506	0.00196	-0.000226	2.5373e-5
1.0	6.2450e-16	-6.9524e-18	-9.2846e-18	-2.0004e-18
0.13231	0.942907	0.005301	-0.000434	2.9063e-5
$x_{(5)}$	$x_{(6)}$	$x_{(7)}$	$x_{(8)}$	$x_{(9)}$
-2.8107e-6	3.1135e-7	-3.4932e-8	4.0114e-9	-4.7424e-10
-1.6020e-19	-2.6142e-20	9.0955e-22	6.0651e-21	5.3120e-21
-1.1141e-6	-8.5441e-8	2.8144e-8	-4.6133e-9	6.1684e-10

Исследование точности полученных результатов

Пусть $\tilde{x}$ — опорные точки плана, полученные в явном виде, $x_{<n>}$ — разложение в окрестности $a^* = 5$ до n-го члена.

- при $a = 2$

$$x_{<7>}(a) = \begin{pmatrix} 0.2086 \\ 1.0 \\ 4.7913 \end{pmatrix}, \quad \tilde{x} = \begin{pmatrix} 0.2086 \\ 1.0 \\ 4.7913 \end{pmatrix}$$

достигнута заданная точность $\varepsilon = 0.0001$ при $n = 7$.

- при $a = 3$

$$x_{<3>}(a) = \begin{pmatrix} 0.1510 \\ 1.0 \\ 6.6209 \end{pmatrix}, \quad \tilde{x} = \begin{pmatrix} 0.1510 \\ 1.0 \\ 6.6209 \end{pmatrix}$$

достигнута заданная точность $\varepsilon = 0.0001$ при $n = 3$.

Результаты

- Построены локально-оптимальные планы с помощью функционального подхода в случае, когда они не могут быть получены в явном виде
- Локально D-оптимальные планы были получены явно (для больших промежутков) и с помощью численного метода оптимизации (для малых промежутков)
- Продемонстрирована эффективность функционального подхода путем сравнения его результатов с результатами явного и численного построения планов
- Разработана программа на языке Python, реализующая явное и численное построение планов и алгоритм рекуррентного вычисления коэффициентов ряда Тейлора

Заключение

- Разложение опорных точек в ряд Тейлора позволяет получать оптимальные планы с высокой заданной точностью
- Полученные результаты показывают возможность использования приближений, полученных с помощью функционального подхода, как явных формул для построения локально-оптимальных планов

Благодарю за внимание!