Планирование эксперимента для дискриминации моделей

Гученко Роман Александрович

Кафедра статистического моделирования
Санкт-Петербургский государственный университет
Математико-механический факультет

Научный руководитель: д.ф.-м.н., профессор В.Б. Мелас
Рецензент: к.ф.-м.н., доцент П.В. Шпилев

Санкт-Петербург
2014г.

План эксперимента — дискретная вероятностная мера:

$$\xi = \begin{bmatrix} x_1 & \dots & x_n \\ \omega_1 & \dots & \omega_n \end{bmatrix}, \qquad x_i \in \mathcal{X},\ \omega_i \geq 0,\ \sum_{i=1}^{n} \omega_i = 1.$$

Общее уравнение регрессии:

$$y_{i,j} = \eta(x_i, \theta) + \varepsilon_{i,j}, \qquad i = 1, \dots, n,\ j = 1, \dots, r_i,\ r_i = N\omega_i,$$

где

- Точки x_i и веса ω_i задаются планом.
- Значения $y_{i,j}$ — это результаты наблюдений.
- Функция $\eta(x, \theta)$ называется регрессионной моделью.
- Вектор θ отвечает за неизвестные параметры этой модели.
- $\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma^2)$ — независимые случайные ошибки.
- Общее число доступных измерений равно N.

В работе [Atkinson, Fedorov, 1975] было доказано, что мощность F-теста для проверки гипотезы

$$H_0 : \eta(x, \theta) = \eta_1(x, \theta_1)$$

против альтернативы

$$H_1 : \eta(x, \theta) = \eta_2(x, \theta_2)$$

является монотонно возрастающей функцией от величины

$$\Delta(\xi) = \sum_{i=1}^{n} \omega_i \left[\eta_1(x_i, \theta_1) - \eta_2(x_i, \widehat{\theta}_2)\right]^2;$$

$$\widehat{\theta}_2 = \arg\inf_{\theta_2 \in \Theta_2} \sum_{i=1}^{n} \omega_i \left[\eta_1(x_i, \theta_1) - \eta_2(x_i, \theta_2)\right]^2.$$

Пусть $\overline{\theta}_1$ — это некоторая априорная оценка параметров θ_1.

Определение (Atkinson, Fedorov, 1975)

План ξ^* называется локальным T-оптимальным планом для дискриминации моделей η_1 и η_2, если он максимизирует

$$T(\xi) = \int_{\mathcal{X}} \left[\eta_1(x, \overline{\theta}_1) - \eta_2(x, \widehat{\theta}_2)\right]^2 \xi(dx), \text{ где}$$

$$\widehat{\theta}_2 = \arg\inf_{\theta_2 \in \Theta_2} \int_{\mathcal{X}} \left[\eta_1(x, \overline{\theta}_1) - \eta_2(x, \theta_2)\right]^2 \xi(dx).$$

Определение (Braess, Dette, 2013)

План ξ^* называется локальным T_p-оптимальным планом для дискриминации моделей $\eta_1, \ldots, \eta_\nu$, если он максимизирует

$$T_p(\xi) = \sum_{i,j=1}^{\nu} p_{i,j} \inf_{\theta_{i,j} \in \Theta_j} \int_{\mathcal{X}} \left[\eta_i(x, \overline{\theta}_i) - \eta_j(x, \widehat{\theta}_{i,j})\right]^2 \xi(dx), \text{ где}$$

$$\widehat{\theta}_{i,j} = \arg\inf_{\theta_{i,j} \in \Theta_j} \int_{\mathcal{X}} \left[\eta_i(x, \overline{\theta}_i) - \eta_j(x, \theta_{i,j})\right]^2 \xi(dx),$$

$$p_{i,i} = 0,\ p_{i,j} \geq 0,\ i, j = 1, \ldots, \nu.$$

Теорема (Braess, Dette, 2013)

Введем обозначение:

$$\Psi(x,\xi) = \sum_{i,j=1}^{\nu} p_{i,j}\left[\eta_i(x,\overline{\theta}_i) - \eta_j(x,\widehat{\theta}_{i,j})\right]^2.$$

[R1] *План* ξ^* *—* T_p*-оптимальный* $\Leftrightarrow \forall x \in \mathcal{X} : \Psi(x,\xi^*) \le T_p(\xi^*)$, *причем в опорных точках* ξ^* *достигается равенство.*

[R2] *План* ξ *— не* T_p*-оптимальный* $\Rightarrow \exists \dot{x} \in \mathcal{X} : \Psi(\dot{x},\xi) > T_p(\xi^*)$.

Алгоритм (Аткинсона–Федорова)

Пусть на шаге s *имеется план* ξ_s. *Тогда*

(1.) *Выбираем точку* $x_{s+1} = \arg\max_x \Psi(x,\xi_s)$.

(2.) *Берем новый план* $\xi_{s+1} = [1-\alpha_s]\,\xi_s + \alpha_s \xi(x_{s+1})$, *где* $\alpha_s \to 0$, $\sum_{s=0}^{\infty}\alpha_s = \infty$, $\sum_{s=0}^{\infty}\alpha_s^2 < \infty$.

Алгоритм сходится в том смысле, что $\lim_{s\to\infty} T_p(\xi_s) = \max_\xi T_p(\xi)$.

Утверждение

Если план ξ^* — T_p-оптимальный, тогда

$$\int_{\mathcal{X}} \left[\eta_i(x,\overline{\theta}_i) - \eta_j(x,\widehat{\theta}_{i,j}(\xi^*))\right] \frac{\partial \eta_j(x,\theta_{i,j})}{\partial \theta_{i.j(q)}}\bigg|_{\theta_{i,j}=\widehat{\theta}_{i,j}(\xi^*)} \xi^*(dx) = 0,$$

$$i,j : p_{i,j} \neq 0,\ q = 1,\ldots,\dim(\theta_{i,j}).$$

Лишние точки предлагается удалять решая задачу ЛП:

$$\sum_{i,j=1}^{\nu} p_{i,j} \sum_{k=1}^{n} \omega_k \left[\eta_i(x_k,\overline{\theta}_i) - \eta_j(x_k,\widehat{\theta}_{i,j}(\xi))\right]^2 \to \max_{\omega};$$

$$\sum_{k=1}^{n} \omega_k \left[\eta_r(x_k,\overline{\theta}_r) - \eta_v(x_k,\widehat{\theta}_{r,v}(\xi))\right] \frac{\partial \eta_v(x_k,\theta_{r,v})}{\partial \theta_{r,v}}\bigg|_{\theta_{r,v}=\widehat{\theta}_{r,v}(\xi)} = 0;$$

$$\sum_{k=1}^{n} \omega_k = 1;\ \omega_k \geq 0,\ k = 1,\ldots,n,\ n = \#\mathrm{supp}(\xi);$$

где $\theta_{r,v} = \arg\max_{i,j:p_{i,j}\neq 0} \theta_{i,j}$.

Алгоритм (Идея)

Пусть на шаге s имеется план ξ_s.

(1.) *Добавляем в носитель плана локальные максимумы $\Psi(x, \xi_s)$.*

(2.) *Находим ω, максимизирующие $T_p(\rho, \omega)$ при фиксированных ρ. Здесь ρ — это опорные точки плана, а ω — веса.*

T_p-эффективность:

$$\mathrm{Eff}_{T_p}(\xi, \theta_{\mathrm{fix}}) = \frac{T_p(\xi, \theta_{\mathrm{fix}})}{\sup_\eta T_p(\eta, \theta_{\mathrm{fix}})} \in [0, 1].$$

Возможное условие остановки:

$$\underline{\mathrm{Eff}}_{T_p}(\xi, \theta_{\mathrm{fix}}) = \frac{T_p(\xi, \theta_{\mathrm{fix}})}{\max_x \Psi(x, \xi)} > 1 - \delta.$$

Далее во все численных примерах будем полагать, что $\delta = 10^{-3}$.

Модель Вейбулла и экспоненциальная модель:

$$\begin{cases} \eta_1(x, \theta_1) = a_1 - b_1 e^{-\lambda_1 x^{h_1}}; \\ \eta_2(x, \theta_2) = a_2 - b_2 e^{-\lambda_2 x}; \end{cases}$$

Рассмотрим случай $p_{1,1} = p_{2,1} = p_{2,2} = 0,\ p_{1,2} = 1$. Оптимальный план не зависит от параметров a_1 и b_1 первой модели. Априорные значения для оставшихся параметров: $\overline{\lambda}_1 = 0.1$ и $\overline{h}_1 = 1.5$.

Таблица: Стартовый план: равномерный на $x_{\text{init}} = (0, 1, \ldots, 9, 10)$.

Итоговый план				Время			
x_1	x_2	x_3	x_4	0	1	2a	2b
ω_1	ω_2	ω_3	ω_4				
0.000	1.466	5.896	10.000	21.28	8.58	0.05	0.09
0.213	0.380	0.287	0.120				

Четыре dose-response модели ($p_{i,j} = 1$ при $i > j$):

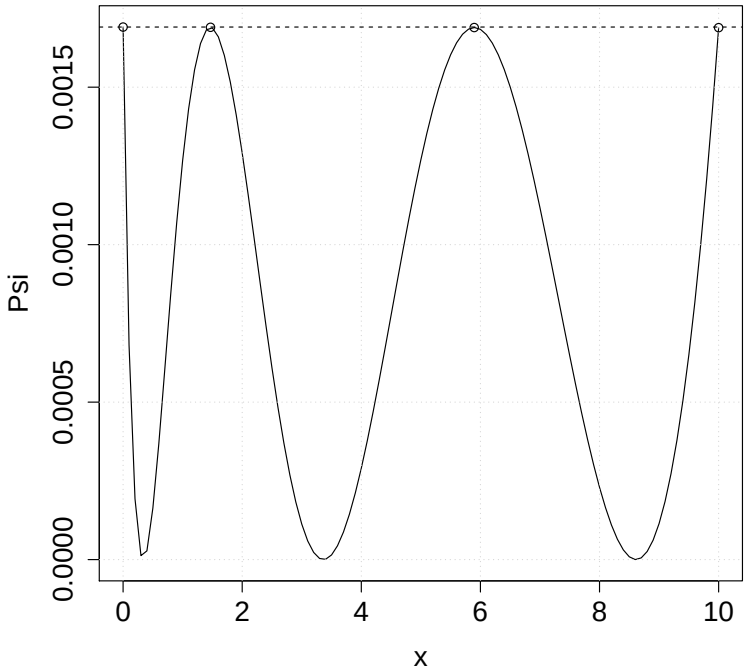
$$\begin{cases} \eta_1(x, \theta_1) = \theta_{1,1} + \theta_{1,2} x; \\ \eta_2(x, \theta_2) = \theta_{2,1} + \theta_{2,2} x(\theta_{2,3} - x); \\ \eta_3(x, \theta_3) = \theta_{3,1} + \theta_{3,2} x/(\theta_{3,3} + x); \\ \eta_4(x, \theta_4) = \theta_{4,1} + \theta_{4,2}/(1 + \exp(\theta_{4,3} - x)/\theta_{4,4}). \end{cases}$$

Априорные значения параметров:

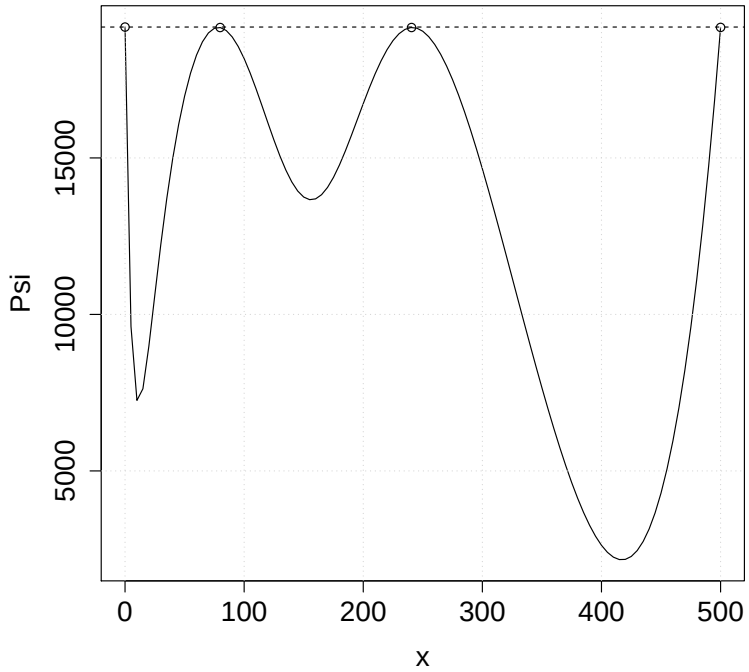
$$\overline{\theta}_1 = (60, 0.56); \qquad \overline{\theta}_2 = (60, 7/2250, 600);$$
$$\overline{\theta}_3 = (60, 294, 25); \qquad \overline{\theta}_4 = (49.62, 290.51, 150, 45.51).$$

Таблица: Стартовый план: равномерный на $x_{\text{init}} = (0, 50, \ldots, 450, 500)$.

Итоговый план				Время			
x_1	x_2	x_3	x_4	0	1	2a	2b
ω_1	ω_2	ω_3	ω_4				
0.000	79.171	240.870	500.000	30.23	11.17	0.18	0.65
0.255	0.213	0.357	0.175				



(a) Первый пример.



(b) Второй пример.

Рис.: Иллюстрация к теореме эквивалентности. Непрерывная линия — это график функции $\Psi(x, \xi')$, пунктирная линия — это значение функционала $T(\xi')$, а кругами обозначены положения опорных точек итогового плана ξ'.

Рассмотрим следующие модели ($x \in [-1, 1]$):

$$\begin{cases} \eta_1 = \sum_{i=0}^{m} \theta_{1,i} x^i + \frac{\theta_{1,m+1}}{\theta_{1,m+2} x^2 - 1}; \\ \eta_2 = \sum_{i=0}^{m} \theta_{2,i} x^i; \end{cases} \qquad \begin{cases} \eta_1 = \sum_{i=0}^{m} \theta_{1,i} x^i + \frac{\theta_{1,m+1} x}{\theta_{1,m+2} x^2 - 1}; \\ \eta_2 = \sum_{i=0}^{m} \theta_{2,i} x^i; \end{cases}$$

Теорема

T-оптимальные планы для представленных моделей сосредоточены в $(m+2)$*-х точках. Если* m *— нечетное, то опорные точки из* $(-1, 1)$ *плана для левой пары моделей совпадают с корнями полиномов*

$$\begin{aligned} \Psi_1(x) &= U_m(x) - 2\alpha^2 U_{m-2}(x) + \alpha^4 U_{m-4}(x); \\ \Psi_2(x) &= 2x\left[\alpha^4 - 1\right] T_{m-1}(x) + \\ &\quad + \left[\alpha^4 + 2\alpha^2 + 1 - 2x^2\left\{\alpha^4 + 1\right\}\right] U_{m-2}(x) \end{aligned}$$

а если m *— четное, то это верно для правой пары. Точки* ± 1 *принадлежат носителю оптимального плана. Тут* $\alpha = a - \sqrt{a^2 - 1}$; $a = 1/\theta_{1,m+2}$; $T_k(x)$ *и* $U_k(x)$ *— полиномы Чебышева.*

- $m = 2$ (Правая пара)

$$x_1 = -1 \qquad x_2 = -\frac{1}{2}(\alpha^2 + 1) \qquad x_3 = \frac{1}{2}(\alpha^2 + 1) \qquad x_4 = 1$$

- $m = 3$ (Левая пара)

$$x_1 = -1 \quad x_2 = -\sqrt{\frac{\alpha^2 + 1}{2}} \quad x_3 = 0 \quad x_4 = \sqrt{\frac{\alpha^2 + 1}{2}} \quad t_5 = 1$$

- $m = 4$ (Правая пара)

$$x_1 = -1 \qquad x_2 = -\frac{\sqrt{4\alpha^2 + 5} + 1}{4} \qquad x_3 = -\frac{\sqrt{4\alpha^2 + 5} - 1}{4}$$

$$x_4 = \frac{\sqrt{4\alpha^2 + 5} - 1}{4} \qquad x_5 = \frac{\sqrt{4\alpha^2 + 5} + 1}{4} \qquad x_6 = 1$$

- $m = 5$ (Левая пара)

$$x_1 = -1 \qquad x_2 = -\frac{\sqrt{\alpha^2 + 3}}{2} \qquad x_3 = -\frac{\sqrt{\alpha^2 + 1}}{2}$$

$$x_4 = 0 \qquad x_5 = \frac{\sqrt{\alpha^2 + 1}}{2} \qquad x_6 = \frac{\sqrt{\alpha^2 + 3}}{2}$$

$$x_7 = 1$$

Модели $\eta_k(t, \theta_k)$ удовлетворяет дифференциальным уравнениям:

$$\eta'(t) = \mu_k(t)\eta(t),$$

где

$$\mu_1(t) = \theta_{1,1}\left[1 + \frac{\theta_{1,2}}{S(t)}\right]^{-1};$$

$$\mu_2(t) = \theta_{2,1}\left[1 + \frac{\theta_{2,2}}{S(t)}\left\{1 + \frac{I(t)}{\theta_{2,3}}\right\}\right]^{-1};$$

$$\mu_3(t) = \theta_{3,1}\left[1 + \frac{\theta_{3,2}}{S(t)}\right]^{-1}\left[1 + \frac{I(t)}{\theta_{3,3}}\right]^{-1};$$

$$\mu_4(t) = \theta_{4,1}\left[1 + \frac{\theta_{4,2}}{S(t)} + \frac{I(t)}{\theta_{4,3}}\right]^{-1};$$

Связь между $S(t)$, $I(t)$ и $\eta(t)$ задается соотношениями:

$$S(t) - S_0 = \frac{\eta_0 - \eta(t)}{\theta_S};\ I(t) - I_0 = \frac{\eta_0 - \eta(t)}{\theta_I};$$

при этом $\eta_0 = \eta(0) > 0$, $S_0 = S(0) > 0$, $I_0 = I(0) > 0$, $I(t) > 0$, и все параметры в θ_k положительны для $k = 1, \ldots, 4$.

Лемма

Введем обозначения: $a = S_0\theta_S + \eta_0;\ b = I_0\theta_I + \eta_0.$
Модели типа Моно неявно получаются из уравнений:

$$t = K_1\left[A_1 \log\frac{\eta(t)}{\eta_0} - B_1 \log\frac{a-\eta(t)}{a-\eta_0}\right];\ K_1 = 1/\theta_{1,1};$$

$$A_1 = 1 + \theta_{1,2}\theta_S/a;\ B_1 = \theta_{1,2}\theta_S/a.$$

$$t = K_2\left[A_2 \log\frac{\eta(t)}{\eta_0} - B_2 \log\frac{a-\eta(t)}{a-\eta_0}\right];\ K_2 = 1/\theta_{2,1}\theta_{2,3}\theta_I;$$

$$A_2 = R + \theta_{2,3}\theta_I;\ B_2 = R - \theta_{2,2}\theta_S;\ R = b\theta_{2,2}\theta_S + \theta_{2,2}\theta_{2,3}\theta_S\theta_I/a.$$

$$t = K_3\left[A_3 \log\frac{\eta(t)}{\eta_0} - B_3 \log\frac{a-\eta(t)}{a-\eta_0} + \eta_0 - \eta(t)\right];\ K_3 = 1/\theta_{3,1}\theta_{3,3}\theta_I;$$

$$A_3 = (b + \theta_{3,3}\theta_I)(a + \theta_{3,2}\theta_S)/a;\ B_3 = \theta_{3,2}\theta_S(b + \theta_{3,3}\theta_I - a)/a.$$

$$t = K_4\left[A_4 \log\frac{\eta(t)}{\eta_0} - B_4 \log\frac{a-\eta(t)}{a-\eta_0} + \eta_0 - \eta(t)\right];\ K_4 = 1/\theta_{4,1}\theta_{4,3}\theta_I;$$

$$A_4 = b + \theta_{4,3}\theta_I + \theta_{4,2}\theta_{4,3}\theta_S\theta_I/a;\ B_4 = \theta_{4,2}\theta_{4,3}\theta_S\theta_I/a.$$

В дипломной работе сделано следующее:

1. Предложены новые численные процедуры для построения T и T_p-оптимальных планов эксперимента для случая произвольного числа конкурирующих регрессионных моделей, которые оказались весьма эффективными.
2. Для нескольких пар дробно-рациональных моделей сформулированы и доказаны теоремы, позволяющие аналитически находить опорные точки для T-оптимальных планов.
3. Для моделей типа Моно установлено, что сама постановка задачи дискриминации возможна не для всех комбинаций моделей из класса. В возможных случаях планы дискриминации были найдены численно.