ИЗВЕСТИЯ
ТОМСКОГО ОРДЕНА ОКТЯБРЬСКОЙ РЕВОЛЮЦИИ
И ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА ИМ. С. М. КИРОВА

Том 277 1977

ЛОКАЛЬНО-ОПТИМАЛЬНЫЕ ПОСЛЕДОВАТЕЛЬНЫЕ ПЛАНЫ В ДУАЛЬНОЙ ПОСТАНОВКЕ

С. М. БЕРСЕНЕВ, Н. И. САБЛИН

(Представлена научным семинаром НИИ ядерной физики, электроники и автоматики)

В настоящей работе вводится функционал, отражающий дуальность эксперимента, приводятся необходимые и достаточные условия локальной оптимальности, предлагается итерационная процедура построения локально-оптимального плана. Подобная задача решалась, например, в [1] с помощью информационного метода, однако без вывода условий оптимальности.

Пусть $\eta_1(x, \Theta_1)$ $\eta_2(x, \Theta_2)$, ..., $\eta_m(x, \Theta_m)$ непрерывные, линейные по параметрам функции, заданные на замкнутом множестве X, X — область возможных измерений, $X \subseteq \overline{X}$; $\xi(x)$ — вероятностная мера, заданная на области X, $y(x)$ — результаты наблюдения, причем

$$E(y/x)=\begin{cases}\eta_1(x,\theta_1),\\ \eta_m(x,\theta_m),\end{cases} \quad (1)$$

E — оператор усреднения по результатам наблюдений.

Будем предполагать, что результаты наблюдений распределены по нормальному закону с функцией плотности вероятности

$$\varphi_i(y_n)=(2\pi(\lambda(x)^{-1}+d_i(x,\varepsilon)))^{-1/2}\exp\left\{-\frac{1}{2}\cdot\frac{(y_n-\hat{\eta}_i(x,\hat{\theta}_i\varepsilon))^2}{\lambda(x)^{-1}+d_i(x,\varepsilon)}\right\};$$

здесь $\hat{\eta}_i(x, \hat{\Theta}_i, \varepsilon)$ — оценка i-й гипотезы по методу наименьших квадратов после проведения эксперимента по плану ε, который понимается как

$$\varepsilon=\begin{Bmatrix}\xi(x_1)\cdots\xi(x_n)\\ x_1\cdots\quad x_n\end{Bmatrix},\ x_i\in X,\ \int_X d\,\xi(x)=1;$$

$d_i(x,\varepsilon)=f_i^{\mathrm{T}}(x)M_i^{-1}(\varepsilon)f_i(x)$ — коридор ошибок i-й гипотезы;

$M_i(\varepsilon)=\int_X \lambda(x)f_i^{\mathrm{T}}(x)\,d\xi(x)$ — информационная матрица;

$\lambda(x)^{-1}$ — дисперсия наблюдений в точке x.

Введем вектор априорных вероятностей, связанных с соответствующими гипотезами

$$P^{\mathrm{T}}=\| p_{1,n},\dots, p_{m,n}\|,$$

элементы которого после каждого измерения в точке x могут быть рассчитаны по формуле Байеса:

$$p_{i,n}=\frac{p_{i,n-1}\varphi_i(y_n)}{\sum\limits_{j=1}^{m} p_{j,n-1}\varphi_j(y_n)}. \tag{2}$$

Поставим задачу о поиске плана эксперимента, осуществляя который мы проведем наилучший эксперимент не только для дискриминации, но и для уточнения параметров истинной модели в смысле D-оптимальности. Введем функционал:

$$\Delta(\varepsilon)=P^{\mathrm{T}} A(\varepsilon) P; \tag{3}$$

$$A(\varepsilon)=\omega_1 \int\limits_X D(\varepsilon,\ x)\, d\xi(x)+\omega_2 L(\varepsilon);$$

здесь ω_1 и ω_2 — весовые множители;

$D(\varepsilon,\ x)$ $L(\varepsilon)$ — матрицы, характеризующие дискриминацию и уточнение соответственно. Определим их элементы следующим образом:

$$d_{ij}=\begin{cases}0, & i=j\\ \lambda(x)\,(\hat{\eta}_i(\varepsilon)-\hat{\eta}_j(\varepsilon)), & i\neq j;\end{cases}$$

$$\hat{\eta}_i(\varepsilon)=\hat{\eta}_i(x,\hat{\theta}_i,\ \varepsilon),$$

$$l_{ij}=\begin{cases}\ln|M_j(\varepsilon)|, & i=j\\ 0, & i\neq j,\end{cases}$$

Кроме того, введем матрицы

$$\overline{A}(\varepsilon,x)=\omega_1 D(\varepsilon,x)+\omega_2\overline{L}(\varepsilon,x),$$

$$\overline{l}_{ij}=\begin{cases}0, & i\neq j\\ \lambda(x) d_j(x,\varepsilon), & i=j;\end{cases}$$

$$\widetilde{A}(\varepsilon)=\omega_1 \int\limits_X D(\varepsilon,x)\, d\,\xi(x)+\omega_2 K,$$

$$k_{ij}=\begin{cases}0, & i\neq j\\ k_j, & i=j,\end{cases}$$

k_j — число параметров в j-й гипотезе.

План ε считаем локально-оптимальным, если он доставляет максимум (3) при заданном P.

Рассмотрим план, являющийся линейной комбинацией двух планов

$$\varepsilon=(1-\alpha)\varepsilon_1+\alpha\varepsilon_2,\ \alpha\in[0,\ 1];$$

ε_1 — невырожденный, т. е. $|M_j(\varepsilon_1)|=0$, $j=1, \ldots, m$.

Лемма.

$$\left.\frac{\partial \Delta(\varepsilon)}{\partial \alpha}\right|_{\alpha=0}=P^{\mathrm{T}}\int\limits_X \overline{A}(\varepsilon_1,x)\, d\,\xi_2(x) P-P^{\mathrm{T}}\widetilde{A}(\varepsilon_1)P, \tag{4}$$

$$\frac{\partial^2 \Delta(\varepsilon)}{\partial \alpha^2}<0;\ \alpha\in[0,\ 1]. \tag{5}$$

Доказательство. Вычислим указанные производные для элементов $L(\varepsilon)$, $\int\limits_X D(\varepsilon, x)\, d\xi(x)$.

$$\frac{\partial}{\partial \alpha}\ln|M_j(\varepsilon)|\Big|_{\alpha=0}= \frac{\partial}{\partial \alpha}\ln|(1-\alpha)M_j(\varepsilon_1)+\alpha M_j(\varepsilon_2)|\Big|_{\alpha=0}=$$

$$=\operatorname{sp} M_j^{-1}(\varepsilon)[M_j(\varepsilon_2)-M_j(\varepsilon_1)]\,|_{\alpha=0}=$$

$$=\operatorname{sp} M_j^{-1}(\varepsilon_1)M_j(\varepsilon_2)-\operatorname{sp} M_j^{-1}(\varepsilon_1)\cdot M_j(\varepsilon_1)=$$

$$\int\limits_X \lambda(x) f_j^T(x) M_j^{-1}(\varepsilon_1) f_j(x)\, d\xi(x)-k_j,$$

$$\frac{\partial^2}{\partial \alpha^2}\ln|M_j(\varepsilon)|=\operatorname{sp}\{-M_j^{-1}(\varepsilon)[M_j(\varepsilon_2)-M_j(\varepsilon_1)]M_j^{-1}(\varepsilon)[M_j(\varepsilon_2)-M_j(\varepsilon_1)]\}=$$

$$=-\operatorname{sp}\{M_j^{-1}(\varepsilon)[M_j(\varepsilon_2)-M_j(\varepsilon_1)]\}^2<0.$$

Преобразуем элемент матрицы $\int\limits_X D(\varepsilon, x)\, d\xi(x)$:

$$\int\limits_X \lambda(x)(\hat{\eta}_i(\varepsilon)-\hat{\eta}_j(\varepsilon))^2\, d\xi(x)=\int\limits_X \lambda(x)\,(\hat{\eta}_i(\varepsilon)-\eta_u+\eta_u-\hat{\eta}_j(\varepsilon))^2 d\xi(x),$$

где η_u — истинная гипотеза, равная в точке x результату наблюдения,

$\hat{\eta}_i(\varepsilon)$ — величина, построенная по типу наилучшей линейной оценки. Предполагая, что система гипотез независима, получим

$$\int\limits_X \lambda(x)(\hat{\eta}_i(\varepsilon)-\hat{\eta}_j(\varepsilon))^2 d\xi(x)=\int\limits_X \lambda(x)(\eta_u-\hat{\eta}_i(\varepsilon))^2 d\xi(x)+$$

$$+\int \lambda(x)(\eta_u-\hat{\eta}_j(\varepsilon))^2 d\xi(x).$$

При вычислении производных воспользуемся результатом [2]:

$$\frac{\partial}{\partial \alpha}\int\limits_X \lambda(x)(\eta_u-\hat{\eta}_j(\varepsilon))^2 d\xi(x)\Big|_{\alpha=0}=\int\limits_X \lambda(x)(\eta_u-\hat{\eta}_j(\varepsilon_1))^2 d\xi_2(x)-$$

$$-\int\limits_X \lambda(x)(\eta_u-\hat{\eta}_j(\varepsilon_1))^2 d\xi_1(x).$$

$$\frac{\partial^2}{\partial \alpha^2}\int\limits_X \lambda(x)(\eta_u-\hat{\eta}_j(\varepsilon))^2 d\xi(x)<0.$$

Тогда

$$\frac{\partial}{\partial \alpha}\int\limits_X \lambda(x)(\hat{\eta}_i(\varepsilon)-\hat{\eta}_j(\varepsilon))^2 d\xi(x)\Big|_{\alpha=0}\int\limits_X \lambda(x)(\hat{\eta}_i(\varepsilon_1)-\hat{\eta}_j(\varepsilon_1))^2 d\xi_2(x)-$$

$$-\int\limits_X \lambda(x)(\hat{\eta}_i(\varepsilon_1)-\hat{\eta}_j(\varepsilon_1))^2 d\xi_1(x).$$

Теперь, используя введенные обозначения, нетрудно получить (4) и (5).

Теорема. Для того, чтобы план ε был локально-оптимальным, необходимо и достаточно

$$P^T \tilde{A}(\varepsilon) P \geqslant P^T \bar{A}(\varepsilon, x) P, \ \mathrm{A}\, x \in X. \qquad (6)$$

Доказательство.

1. Необходимость. Пусть $\overset{*}{\varepsilon}$ — локально-оптимальный план и в точке $\overset{\circ}{X}$ не выполняется (6). Тогда рассмотрев план

$$\varepsilon = (1-\alpha)\overset{*}{\varepsilon} + \alpha\varepsilon(\overset{\circ}{x}),$$

$\varepsilon(\overset{\circ}{x})$ — план, сосредоточенный в одной точке, и, выбрав достаточно малое α, мы, используя (4), получим

$$\Delta(\varepsilon) > \Delta(\overset{*}{\varepsilon}),$$

что противоречит определению $\overset{*}{\varepsilon}$.

2. Достаточность. Предположим (6) выполняется для ε, который не является локально-оптимальным, т. е. существует $\tilde{\varepsilon}$, для которого

$$\Delta'(\tilde{\varepsilon}) > \Delta(\overset{*}{\varepsilon}).$$

Тогда, взяв линейную комбинацию

$$\varepsilon = (1-\alpha)\overset{*}{\varepsilon} + \alpha\tilde{\varepsilon}.$$

и в силу вогнутости (условие (5)), получим

$$\frac{\partial \Delta(\varepsilon)}{\partial \alpha}\Big|_{\alpha=0} = P^T \int_X \bar{A}(\overset{*}{\varepsilon}, x) d\tilde{\xi}(x) P - P^T \tilde{A}(\overset{*}{\varepsilon}) P > 0,$$

т. е.

$$\max_{x \in X} P^T \bar{A}(\overset{*}{\varepsilon}, x) P > P^T \tilde{A}(\overset{*}{\varepsilon}) P,$$

что противоречит (6). Теорема доказана.

Используя доказанную теорему, можно предложить следующую итерационную процедуру построения стратегии локально-оптимального плана эксперимента.

А. Проводится затравочный эксперимент с планом $\overset{s}{\varepsilon}$, невырожденным для всех гипотез, делаются оценки $\hat{\eta}_i(\overset{s}{\varepsilon})$, $i=1, \ldots, m$.

Б. Задается вектор априорных вероятностей

$$P_s^T = \| p_{1,s}, \ldots, p_{m,s} \|.$$

В. Находится точка $\overset{s}{x}$, удовлетворяющая

$$\max_{x \in X} P_s^T \bar{A}(\overset{s}{\varepsilon}, x) P_s.$$

Г. В этой точке проводится дополнительное измерение, пересчитывается вектор вероятностей, строятся новые оценки $\hat{\eta}_i(\overset{s+1}{\varepsilon})$, $i=1, \ldots, m$ затем возвращаемся к В.

Процедура (В, Г) проводится до тех пор, пока не выполнится (6). На каждом шаге итерационной процедуры мы имеем неравенство

$$\Delta(\overset{s+1}{\varepsilon}) > \overset{s}{\Delta}(\varepsilon),$$

(s — номер итерации), так как при малом α_s справедливы следующие соотношения:

$$\overset{s+1}{\varepsilon} = (1-\overset{s}{\alpha})\overset{s}{\varepsilon} + \overset{s}{\alpha}\,\overset{s}{\varepsilon}(x),$$

$$\frac{\Delta(\overset{s+1}{\varepsilon}) - \Delta(\overset{s}{\varepsilon})}{\overset{s}{\alpha}} \cong P_s^T \overline{A}(\overset{s}{\varepsilon}, \overset{s}{x}) P_s - P_s^T \widetilde{A}(\overset{s}{\varepsilon}) P_s.$$

Условие (6) при равенстве одной из $p_{j,s}$ единице переходит в условие D — оптимальности для j-й гипотезы.

ЛИТЕРАТУРА

1. Г. С. Федорова. «Заводская лаборатория», 1970, № 5.

2. В. В. Федоров. Локально-оптимальные последовательные планы дискриминирующих экспериментов, М., изд-во МГУ, 1970.