М. И. ЮДИЦКИЙ, ст. инж.
НИИАСС Госстроя УССР

О ПЛАНИРОВАНИИ ЭКСПЕРИМЕНТОВ ПРИ ОЦЕНКЕ ПАРАМЕТРОВ В РЕГРЕССИОННЫХ МОДЕЛЯХ С КОРРЕЛИРОВАННЫМИ ОШИБКАМИ НАБЛЮДЕНИЙ

Рассмотрим линейную модель множественной регрессии

$$y_i = \vec{\theta}\,'\vec{x}_i + \varepsilon_i, \quad i = \overline{1, n}, \tag{1}$$

где y_i — значение выходной переменной в i-м наблюдении; $\vec{x}_i = (x_{1i}, x_{2i}, \ldots, x_{ki})'$ — вектор значений контролируемых переменных; $\vec{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)'$—неизвестный параметр; ε_i — ошибка наблюдения.

Пусть ε_i — стационарная последовательность с $\mathbf{M}\varepsilon_i = 0$, $\mathbf{M}\,\varepsilon_i\varepsilon_j = r_{ij} = r(|i - j|)$ (положим $r(0) = 1$), $R_n = \| r_{ij} \|_{i,j=1}^{n}$. Запишем (1) в виде

$$\vec{y}_n = X_n\vec{\theta} + \vec{\varepsilon}_n, \tag{2}$$

где

$$\vec{y}_n = (y_1, y_2, \ldots, y_n)', \quad \vec{\varepsilon}_n = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)', \quad X_n = (\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n)'.$$

В качестве оценки вектора $\vec{\theta}$, построенной по результатам n наблюдений, возьмем оценку $\hat{\vec{\theta}}^{(n)}$, минимизирующую величину

$$(\vec{y}_n - X_n\vec{\theta})'\,(\vec{y}_n - X_n\vec{\theta}).$$

Тогда

$$\hat{\vec{\theta}}^{(n)} = (X_n'X_n)^{-1}X_n'\vec{y}_n \tag{3}$$

и обобщенная дисперсия оценки $\hat{\vec{\theta}}^{(n)}$ (детерминант ковариационной матрицы)

$$\mathbf{D}\,\hat{\vec{\theta}}^{(n)} = \frac{|X_n'R_nX_n|}{|X_n'X_n|^2}, \tag{4}$$

где $|A|$ — детерминант матрицы A.

Обобщенная дисперсия является в определенном смысле характеристикой точности оценки.

Рассмотрим задачу о выборе значений $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ из некоторой допустимой ограниченной замкнутой области $\Omega \subset R^k$ (R^k — k — мерное евклидово пространство) таким образом, чтобы $\mathbf{D}\hat{\vec{\theta}}^{(n)}$ была минимальной. Набор $\vec{x}_2, \vec{x}_1, \ldots, \vec{x}_n$, при котором достигается минимум $\mathbf{D}\hat{\vec{\theta}}^{(n)}$, будем называть оптимальным экспериментом.

Для модели с независимыми наблюдениями план эксперимента определяется набором точек $\vec{x}_2, \vec{x}_1, \ldots, \vec{x}_l$ ($\vec{x}_v \in \Omega$, $v = \overline{1, l}$, $l \leqslant n$, в которых проводятся наблюдения, и числами p_v ($p_v = \frac{n_v}{n}$, $v = \overline{1, l}$, n_v — целые, $n_v \geqslant 0$, $n_1 + n_2 + \ldots + n_l = n$), характеризующими «долю» измерений в каждой из точек, т. е. единичной дискретной мерой μ_n. Если n достаточно велико, то от требования кратности чисел p_v величине $\frac{1}{n}$ можно отказаться (в этом случае полученное решение будет приближенным решением исходной задачи) и рассматривать «обобщенный» план, задаваемый мерой $\mu(\vec{x})$ на Ω. При этом планирование эксперимента состоит в отыскании меры $\mu(\vec{x})$, минимизирующей $D\hat{\vec{\Theta}}$. Полученные результаты имеют следующий вид: указывается число точек, в которых достаточно проводить наблюдения, и определенные свойства, которыми должны обладать эти точки.

Наличие корреляции между ошибками наблюдений значительно усложняет задачу, так как, в частности, изменение последовательности точек, в которых проводятся измерения, влечет изменение дисперсии оценки, и поэтому одна лишь мера $\mu(\vec{x})$ не задает полностью план эксперимента.

Для решения поставленной задачи предлагается следующий подход. Будем считать, что число наблюдений достаточно велико, и будем минимизировать асимптотическую обобщенную дисперсию. Кроме того, предположим, что переменные $\vec{x}_v$—независимые одинаково распределенные случайные величины с $P\{\vec{x}_v \in \Omega_1\} = \mu(\Omega_1)$ для любого $\Omega_1 \in \sigma(\Omega)$ ($\sigma(\Omega)$ — σ- алгебра борелевских множеств). Тогда (4) дает выражение условной обобщенной дисперсии оценки $\hat{\vec{\theta}}^{(n)}$ при гипотезе, что первые n наблюдений проведены в точках $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$.

Что касается существования оценок (3), то отметим следующее. Если $\vec{x}_v$ имеет непрерывное распределение вероятностей, то матрица

$X_n'X_n$ с вероятностью 1 невырожденная и оценка существует почти наверное (п. н.). В общем случае может оказаться, что $\mathbf{P}\{|X_n'X_n| = 0\} \neq 0$. Оценим эту вероятность. Будем в дальнейшем рассматривать лишь те распределения вероятностей, для которых матрица вторых моментов $Z = \| z_{ij} \|_{i,j=1}^k$ невырожденная. Так как $\vec{x}_v$ $(v = \overline{1,n})$ имеют одинаковые распределения и сосредоточены в ограниченной области Ω, то $\mathbf{M}|x_{iv}x_{jv}|^r < \infty$ для всех $i, j = \overline{1, k}$ и для любого $r = 1, 2, \ldots$. Следовательно, согласно теореме 28 ([1], стр. 349) для любых $\varepsilon > 0$, $r = 1, 2, \ldots$ и всех $i, j = \overline{1,k}$ $\mathbf{P}\{|\rho_{ij}^{(n)}| \geqslant \varepsilon\} = o(n^{-r+1})$, где $\| \rho_{ij}^{(n)} \|_{i,j=1}^k = \frac{1}{n} X_n'X_n - Z$.

Так как $|Z| > 0$, то, очевидно, существует такое $\varepsilon_0 > 0$, что при $|\rho_{ij}^{(n)}| < \varepsilon_0$ $(i, j = \overline{1,k})$ $\frac{1}{n} |X_n'X_n| > 0$. Поэтому

$$\mathbf{P}\left\{\frac{1}{n}|X_n'X_n| = 0\right\} \leqslant \sum_{i,j=1}^{k} \mathbf{P}\{ |\rho_{ij}^{(n)}| \geqslant \varepsilon_0\} = o(n^{-r+1})$$

и вероятность обращения в нуль определителя при больших n мала.

Следовательно, можно положить

$$\hat{\vec{\theta}}^{(n)} = (X_n'X_n)^{-1}X_n'\vec{y}_n \text{ при } |X_n'X_n| \neq 0. \qquad (3')$$

Отметим также, что поскольку изменение конечного числа начальных наблюдений не влияет на значение асимптотической обобщенной дисперсии, то можно неслучайно разместить первые k наблюдений так, чтобы обеспечить невырожденность матрицы $X_n' X_n$ при $n > k$. Таким образом, всегда можно добиться, чтобы $|X_n'X_n| \neq 0$; именно этот случай будем далее рассматривать.

Пусть Ω — некоторый выпуклый многогранник в R^k. В заметке показано, что асимптотически оптимальному эксперименту соответствует распределение, сосредоточенное не более чем в $C_k^2 + 2k$ вершинах Ω.

Перейдем к доказательству этого факта. Найдем предельное выражение для $D\hat{\vec{\Theta}}^{(n)}$.

Теорема 1. Пусть существуют четвертые моменты меры μ, матрица Z вторых моментов меры невырожденна и спектральная плотность $f(\lambda)$ стационарной последовательности ε_i является кусочно-гладкой непрерывной в нуле функцией. Пусть $\vec{m} = \mathbf{M}\vec{x}_v = (m_1, m_2, \ldots, m_k)'$

$$\delta = 2\pi f(0),\ S = \mathbf{M}(\vec{x}_v - \vec{m})(\vec{x}_v - \vec{m})' = \| \sigma_{ij} \|_{i,j=1}^k \text{ (здесь } \sigma_{ii} = \sigma_i^2,\ S + \vec{m}\vec{m}' = Z),\ F(\mu) = |S + \delta\vec{m}\vec{m}'|\,|S + \vec{m}\vec{m}'|^{-2}. \qquad (5)$$

Тогда при $n \to \infty$ имеет место сходимость по вероятности

$$n^k \mathbf{D}\hat{\vec{\theta}}^{(n)} \xrightarrow{p} F(\mu).$$

Доказательство. Рассмотрим формулу (4). По усиленному закону больших чисел

$$\frac{1}{n} X_n' X_n \xrightarrow{\text{п.н}} \| m_i m_j + \sigma_{ij} \|_{i,j=1}^k = S + \vec{m}\vec{m}'.$$

Кроме того,

$$\frac{1}{n} X_n' R_n X_n = \frac{1}{n} \left\| \sum_{v,w=1}^{n} r_{vw} x_{iv} x_{jw} \right\|_{i,j=1}^{k} \xrightarrow{p} \| \delta m_i m_j + \sigma_{ij} \|_{i,j=1}^k = S + \delta \vec{m}\vec{m}'.$$

Действительно, рассмотрим отдельно три слагаемых:

$$\frac{1}{n} \sum_{v,w=1}^{n} r_{vw} x_{iv} x_{jw} = \alpha_n + \beta_n + \gamma_n = \frac{1}{n} m_i m_j \sum_{v,w=1}^{n} r_{vw} +$$

$$+ \frac{1}{n} \sum_{v,w=1}^{n} r_{vw} [m_j (x_{iv} - m_i) + m_i (x_{jw} - m_j)] +$$

$$+ \frac{1}{n} \sum_{v,w=1}^{n} r_{vw} (x_{iv} - m_i)(x_{jw} - m_j).$$

Так как при $n \to \infty$ $\frac{1}{n} \sum_{v,w=1}^{n} r_{vw} \to 2\pi f(0)$, то $\alpha_n \to \delta m_i m_j$. Функция $f(\lambda)$ — кусочно-гладкая, поэтому ряд из ее коэффициентов Фурье $r(u)$ сходится абсолютно. Учитывая, что

$$\frac{1}{n} \sum_{v,w=1}^{n} r_{vw}^2 \leqslant \frac{1}{n} \sum_{v,w=1}^{n} |r(|v-w|)| \leqslant \sum_{u=0}^{n-1} |r(u)| < \infty; \ \mathbf{M}\beta_n = 0,\ \mathbf{M}\beta_n^2 \to 0,$$

получим, что $\beta_n \xrightarrow{p} 0$. Представим γ_n в виде

$$\gamma_n = \frac{1}{n} \sum_{v,w=1}^{n} r_{vw} \xi_v^{(i)} \xi_w^{(j)},\ \mathbf{M}\xi_v^{(i)} = \mathbf{M}\xi_w^{(j)} = 0,$$

$$\mathbf{M}\xi_v^{(i)} \xi_w^{(j)} = \begin{cases} \sigma_{ij}, & v = w, \\ 0, & v \neq w, \end{cases} \quad v, w = \overline{1, n},\ \ i, j = \overline{1, k}.$$

Очевидно, что $\mathbf{M}\gamma_n = \sigma_{ij}$ при всех n. Легко показать, что $\mathbf{D}\gamma_n \to 0$ при $n \to \infty$ и, следовательно, $\gamma_n \xrightarrow{p} \sigma_{ij}$. Теорема доказана.

Замечание. Если при всех $u = 1, 2, \ldots$ $r(u) \geqslant 0$, то в условии теоремы для функции $f(\lambda)$ достаточно предположить лишь ее непрерывность в нуле.

Найдем теперь такую меру $\mu(\vec{x})$ на Ω, чтобы значение функционала $F(\mu)$ было минимальным.

Обозначим через U произвольный фиксированный конечный набор точек, среди которых имеются по крайней мере k линейно независимых, $U = \{\vec{u}_v = (u_{1v}, u_{2v}, \ldots u_{kv})' \in \Omega, v = \overline{1, l}\}$; M_u — семейство вероятностных мер, сосредоточенных в точках множества U. Будем считать в дальнейшем, что $2\pi f(0) \geqslant \frac{1}{2}$; отметим, что из положительности всех $r(|v - w|)$ следует, что $2\pi f(0) \geqslant 1$.

Теорема 2. Пусть $\min\limits_{\mu \in M_u} F(\mu) = F(\mu_0)$. Тогда μ_0 сосредоточена в точках $\vec{u}_{i_1} \vec{u}_{i_2}, \ldots, \vec{u}_{i_t} \in U$, лежащих на некотором эллипсоиде, содержащем внутри себя все остальные точки множества U, причем $t \leqslant C_k^2 + 2k$ (меру μ_0 назовем оптимальной в классе M_u).

Доказательство. Представим $F(\mu)$ как функцию от $(p_1, p_2, \ldots, p_l) = p$ и обозначим ее $F(p)$. Очевидно, что существует минимум $F(p)$ по области $T = \{p : p_v \geqslant 0,\ p_1 + p_2 + \ldots + p_l = 1\}$. Используя метод, описанный в работе [2] (стр. 16—18), получим, что в точке минимума p существует число c такое, что

$$\frac{\partial F}{\partial p_v} = c, \text{ если } p_v > 0; \quad \frac{\partial F}{\partial p_v} \geqslant c, \text{ если } p_v = 0. \tag{6}$$

Рассмотрим $\frac{\partial F}{\partial p_v}$ $(v = \overline{1, l})$:

$$\frac{\partial F}{\partial p_v} = \left[\frac{\frac{\partial}{\partial p_v} | S + \delta \vec{m}\vec{m}' |}{| S + \delta \vec{m}\vec{m}' |} - 2 \frac{\frac{\partial}{\partial p_v} | S + \vec{m}\vec{m}' |}{| S + \vec{m}\vec{m}' |} \right] | S + \delta \vec{m}\vec{m}' | \, | S + \vec{m}\vec{m}' |^{-2}. \tag{7}$$

Так как

$$m_j = \sum_{v=1}^{l} p_v u_{jv}, \quad \sigma_{ij} = \sum_{v=1}^{l} p_v (u_{iv} - m_i)(u_{jv} - m_j), \quad i, j = \overline{1, k},$$

то выражения (7) являются соответственно функциями от векторов $\vec{u}_v (v = \overline{1, l})$ вида $(A\vec{u}_v, \vec{u}_v) + L\vec{u}_v$, где $A = A(U, p)$, $L = L(U, p)$, т. е. матрицы A и L сохраняют свой вид для одного и того же набора $p_1, p_2, \ldots, p_l$ при всех v $(v = \overline{1, l})$.

Покажем, что при $\delta \geqslant \frac{1}{2}$ форма $[-(A\vec{u}_v, \vec{u}_v)]$ неотрицательно определена (н. о. о.). Для этого достаточно показать, что матрица

$$-A=\left[\frac{1}{2}(S+\vec{m}\vec{m}')\right]^{-1}-[S+\delta\vec{m}\vec{m}']^{-1}=B^{-1}-G^{-1}\ \text{н. о.о-} \quad (8)$$

Для доказательства воспользуемся известным из теории матриц результатом (см., например, [3], стр. 71—72): если матрица B — положительно определена (п. о.) и матрица $G—B$ — н. о. о., то матрица $B^{-1}-C^{-1}$ — н. о. о. Из (8) вытекает, что B — п. о. и при $\delta\geqslant\frac{1}{2}$ $G-B=\frac{1}{2}S+\left(\delta-\frac{1}{2}\right)\vec{m}\vec{m}'$ — н. о. о., т. е. форма $[-(A\vec{u}_v,\vec{u}_v)]$ — н. о. о. и соотношения (7) являются уравнениями эллипсоида, причем на этом эллипсоиде лежат все точки, которым соответствуют значения $p_v>0$, т. е. точки, в которых сосредоточена мера, а точки с $p_v=0$ лежат внутри него (см. (6)).

Далее докажем, что при отыскании наборов $p_1, p_2, \ldots, p_l$ минимизирующих $F(p)$, достаточно рассматривать лишь те из них, в которых не более чем $t_0=C_k^2+2k$ чисел p_i отличны от нуля. Для этого введем аналогично [2] t_0-мерные векторы
$\vec{\alpha}_i'=\{u_{1i}, u_{2i}, \ldots, u_{ki}, u_{1i}^2+\sigma_1^2, u_{1i}u_{2i}, \ldots, u_{k-1i}u_{ki}, u_{ki}^2+\sigma_k^2\}$, $i=\overline{1, l}$. Тогда функционалу $F(p)$ соответствует выпуклая комбинация векторов $\vec{\alpha}_i$, а именно $\sum_{i=1}^{l}p_i\vec{\alpha}_i=\vec{F}$, причем вектор $\vec{F}$, отвечающий оптимальному набору $p_1, p_2, \ldots, p_l$, является граничной точкой наименьшего выпуклого многогранника Q, содержащего векторы $\vec{\alpha}_1, \vec{\alpha}_2, \ldots$ $\ldots, \vec{\alpha}_l$. Действительно, если точка $\vec{F}$ не является граничной, то точка $\lambda\vec{F}$ при некотором $\lambda>1$ принадлежит многограннику:

$$\lambda\vec{F}=\sum_{i=1}^{l}\tilde{p}_i\vec{\alpha}_i.$$

Покажем, что эксперимент, соответствующий $p_1, p_2, \ldots, p_l$, придает функционалу значение $\tilde{F}(p)<F(p)$. Так как $\sigma_{ij}=\sum_{v=1}^{l}p_v u_{iv}u_{jv}-m_i m_j$, то, обозначив $Z=\left\|\sum_{v=1}^{l}p_v u_{iv}u_{jv}\right\|_{i,j=1}^{k}$, перепишем правую часть (5) в виде

$$F(p)=|Z+(\delta-1)\vec{m}\vec{m}'|\,|Z|^{-2}.$$

Учитывая, что если матрица D — п. о., E=п.о ., то $|D|<|D+E|$, при $\lambda>1$ получим

$$F(\tilde{p})=|\lambda Z+\lambda^2(\delta-1)\vec{m}\vec{m}'|\,|\lambda Z|^{-2}=\left|\frac{1}{\lambda}Z+(\delta-1)\vec{m}\vec{m}'\right||Z|^{-2}<$$

$$< \left| \frac{1}{\lambda} Z + (\delta - 1) \vec{m}\vec{m}' + \left(1 - \frac{1}{\lambda}\right) Z \right| |Z|^{-2} = F(p).$$

Так как любая граничная точка многогранника $Q \subset R^{t_0}$ может быть представлена как выпуклая комбинация не более чем t_0 векторов, то теорема 2 полностью доказана.

Следствие. Пусть H — наименьший выпуклый многогранник, содержащий все l векторов множества U. Тогда оптимальная в классе M_u мера μ_0 сосредоточена не более чем в $C_k^2 + 2k$ вершинах H.

Теорема 3. Пусть Ω — выпуклый многогранник в R^k. Тогда минимум $F(\mu)$ по всем дискретным мерам μ на Ω, сосредоточенным в конечном числе точек, совпадает с минимумом $F(\mu)$ по дискретным мерам, сосредоточенным не более чем в $C_k^2 + 2k$ вершинах Ω.

Доказательство. Пусть μ' — некоторая дискретная мера, сосредоточенная на множестве точек U', среди которых есть точки, лежащие внутри Ω. Тогда, пополнив U' вершинами многогранника Ω, получим конечный набор точек U, причем $\mu' \in M_u$. Согласно теореме 2 и следствию, $\min\limits_{\mu \in M_u} F(\mu) = F(\mu_0)$, где μ_0 — некоторая мера, сосредоточенная не более чем в $C_k^2 + 2k$ вершинах Ω и $F(\mu_0) \leqslant F(\mu')$. Теорема 3 доказана.

Теорема 4. Пусть M — семейство всех вероятностных мер на выпуклом многограннике Ω. Тогда

$$\min_{\mu \in M} F(\mu) = F(\mu_0),$$

где мера μ_0 сосредоточена не более чем в $C_k^2 + 2k$ вершинах Ω.

Доказательство. Для любой меры $\mu \in M$ и для любого $\varepsilon > 0$ существует такая дискретная мера μ_ε на Ω, сосредоточенная в конечном числе точек, что $|F(\mu) - F(\mu_\varepsilon)| < \varepsilon$. Пусть μ_0 оптимальная в классе дискретных мер на Ω, тогда $F(\mu_0) \leqslant F(\mu_\varepsilon)$. Выберем последовательность $\varepsilon_n \to 0$. Получим $\lim\limits_{n \to \infty} F(\mu_{\varepsilon_n}) = F(\mu)$ и $F(\mu_0) \leqslant F(\mu)$, а свойства μ_0 определяются теоремой 3.

Из теоремы 4 для модели с одномерным неизвестным параметром θ и допустимой областью значений контролируемых переменных $\Omega = [a, b]$ вытекает, что наблюдения следует проводить не более чем в двух точках Ω, а именно, в точках a и b. Докажем в этом случае более сильный результат.

Теорема 5. Минимум функционала

$$F_1(\mu) = \frac{2\pi f(0) m^2 + \sigma^2}{(m^2 + \sigma^2)^2} \tag{5'}$$

по всем вероятностным мерам, заданным на отрезке $[a, b]$, $0 < a < b$, достигается на мере μ_0, сосредоточенной в одной точке b.

Доказательство. Пусть $\delta = 2\pi f(0) > 1$. Тогда

$$F_1(\mu) = \frac{(\delta - 1) m^2}{(m^2 + \sigma^2)^2} + \frac{1}{m^2 + \sigma^2}. \tag{9}$$

Пусть наблюдения проводятся в точке b с вероятностью p и в точке a с вероятностью $1-p$. Тогда из (9) получим

$$\varphi(p)=\frac{(\delta-1)[pb+(1-p)a]^2}{[pb^2+(1-p)a^2]^2}+\frac{1}{pb^2+(1-p)a^2}. \tag{10}$$

Второе слагаемое в (10) минимально при $p=1$. Первое слагаемое можно представить в виде $(\delta-1)\psi^2(p)$, где $\psi(p)>0$, $\psi'(p)<0$. Поэтому $\min\limits_{0\leqslant p\leqslant 1}\varphi(p)$ достигается при $p=1$. Если $\delta\leqslant 1$, то

$$F_1(\mu)=\frac{\delta}{m^2+\sigma^2}+\frac{(1-\delta)\sigma^2}{(m^2+\sigma^2)^2}.$$

Очевидно, что $F_1(\mu)$ минимально, если $m^2+\sigma^2$ максимально и $\sigma^2=0$, т. е. если $m=b$. Таким образом, при $k=1$ оптимальные значения x_v неслучайны, $x_v\equiv b$ $(v=\overline{1,n})$. Причем в этом случае формула (4) дает значение дисперсии оценки $\hat{\theta}^{(n)}$ и $nD\hat{\theta}^{(n)}\to F_1(\mu)$ при $n\to\infty$.

Автор благодарит доцента КГУ А. Я. Дороговцева за постановку задачи и постоянное внимание к работе.

ЛИТЕРАТУРА

1. Петров В. В. Суммы независимых случайных величин. «Наука», М., 1972.

2. Дороговцев А. Я. Оптимальные методы планирования регрессионных экспериментов. Лекции по спецкурсу. КГУ, Киев, 1971.

3. Рао С. Р. Линейные статистические методы и их применения. М., «Наука», 1968.

M. I. Yuditsky

ABOUT OPTIMAL DESIGN IN ESTIMATION OF PARAMETERS OF REGRESSION MODELS WITH CORRELATED ERRORS OF OBSERVATIONS

Summary

The problem of optimal design under criteria of minimum asymptotic generalized variance for estimating unknown parameters in the linear regression models with correlated noise is studied.

Поступила в редколлегию 26.IV 1973.