Матричная модель дисперсионного анализа

$$x_{ij} = \mu + \alpha_i + e_{ij}, \quad i = 1, \ldots, r, \; j = 1, \ldots, n_i, \quad \text{где} \; \sum_{i=1}^{r} \alpha_i = 0,$$

можно записать в матричном виде:

$$Y = X\beta + \epsilon,$$

где $Y = (x_{11}, \ldots, x_{1n_1}, x_{21}, \ldots, x_{2n_2}, \ldots, x_{r1}, \ldots, x_{rn_r})'$ –вектор наблюдений, $\sum_{i=1}^{r} n_i = n$, $\beta = (\mu, \alpha_1, \ldots, \alpha_{r-1})'$ – вектор параметров

Матрица плана размерности n на r

$$
\mathrm{X} = \left[\begin{array}{ccccc}
1 & 1 & 0 & \dots & 0 \\
\dots & \dots & \dots & \dots & \dots \\
1 & 1 & 0 & \dots & 0 \\
\hline
1 & 0 & 1 & \dots & 0 \\
\dots & \dots & \dots & \dots & \dots \\
1 & 0 & 1 & \dots & 0 \\
\hline
\dots & \dots & \dots & \dots & \dots \\
\hline
1 & -1 & -1 & \dots & -1 \\
\dots & \dots & \dots & \dots & \dots \\
1 & -1 & -1 & \dots & -1
\end{array}\right], \qquad \beta = \left[\begin{array}{c}
\mu \\ \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_{\mathrm{r}-1}
\end{array}\right]
$$

Оценки параметров по методу наименьших квадратов имеют вид:

$$
\hat{\beta} = (\mathrm{X}'\mathrm{X})^{-}\mathrm{X}'\mathrm{Y}\,.
$$

Проверка гипотез

Для проверки гипотезы $\mathrm{H}_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_\mathrm{r} = 0$ используют модель с ограничением на параметры при помощи матрицы H размерности r (число параметров) на $\mathrm{s} = \mathrm{r} - 1$

$$\mathrm{H}'\beta = \theta_0, \quad \text{где} \quad \mathrm{H}' = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix} \quad \text{и} \quad \theta_0 = \begin{bmatrix} 0 \\ 0 \\ \ldots \\ 0 \end{bmatrix}$$

Можно показать, что $\mathrm{Z} = \mathrm{H}'\hat{\beta} \sim \mathcal{N}_\mathrm{s}(\mathrm{H}'\beta, \sigma^2\mathrm{D})$, где $\mathrm{D} = \mathrm{H}'(\mathrm{X}'\mathrm{X})^-\mathrm{H}$ и $\mathrm{R}_0^2 = (\mathrm{Y} - \mathrm{X}\hat{\beta})'(\mathrm{Y} - \mathrm{X}\hat{\beta}) \sim \sigma^2\chi^2(\mathrm{n} - \mathrm{r})$ распределены независимо. Кроме того $(\mathrm{Z} - \theta_0)'\mathrm{D}^{-1}(\mathrm{Z} - \theta_0) = \mathrm{R}_1^2 - \mathrm{R}_0^2 \sim \sigma^2\chi^2(\mathrm{s})$, где $\mathrm{R}_1^2 = (\mathrm{Y} - \mathrm{X}\beta^*)'(\mathrm{Y} - \mathrm{X}\beta^*)$, β^* оценка параметров усеченной модели. Тогда при справедливости нулевой гипотезы

$$\mathrm{F} = \frac{\mathrm{n} - \mathrm{r}}{\mathrm{s}} \frac{\mathrm{R}_1^2 - \mathrm{R}_0^2}{\mathrm{R}_0^2} \sim \mathrm{F}(\mathrm{s}, \mathrm{n} - \mathrm{r})\,.$$

Дифференцирование по вектору параметров

$$A\beta = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \dots \\ \beta_m \end{bmatrix} = \begin{bmatrix} \sum\limits_{i=1}^{m} a_{1i}\beta_i \\ \dots \\ \sum\limits_{i=1}^{m} a_{ni}\beta_i \end{bmatrix}$$

$$\frac{\partial A\beta}{\partial \beta} = \begin{bmatrix} \frac{\partial A\beta}{\partial \beta_1} & \dots & \frac{\partial A\beta}{\partial \beta_m} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nm} \end{bmatrix} = A$$

Для получения оценок МНК дифференцируем квадратичную форму.

$$\mathcal{L}_1 = (\mathrm{A}\beta)'(\mathrm{A}\beta) = \left(\sum_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{a}_{1\mathrm{i}}\beta_{\mathrm{i}}\right)^2 + \ldots + \left(\sum_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{a}_{\mathrm{ni}}\beta_{\mathrm{i}}\right)^2 .$$

Система нормальных уравнений имеет вид

$$\begin{cases} \frac{\partial \mathcal{L}_1}{\partial \beta_1} = 2\left(\mathrm{a}_{11}\left(\sum\limits_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{a}_{1\mathrm{i}}\beta_{\mathrm{i}}\right) + \ldots + \mathrm{a}_{\mathrm{n}1}\left(\sum\limits_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{a}_{\mathrm{ni}}\beta_{\mathrm{i}}\right)\right) = 0, \\ \ldots \\ \frac{\partial \mathcal{L}_1}{\partial \beta_{\mathrm{m}}} = 2\left(\mathrm{a}_{1\mathrm{m}}\left(\sum\limits_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{a}_{1\mathrm{i}}\beta_{\mathrm{i}}\right) + \ldots + \mathrm{a}_{\mathrm{nm}}\left(\sum\limits_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{a}_{\mathrm{ni}}\beta_{\mathrm{i}}\right)\right) = 0, \end{cases}$$

$$\frac{\partial (\mathrm{A}\beta)'(\mathrm{A}\beta)}{\partial \beta} = 0 \Longleftrightarrow \quad 2\frac{\partial (\mathrm{A}'\beta)}{\partial \beta}\mathrm{A}\beta = 2\mathrm{A}'\mathrm{A}\beta = 0\,.$$

ОМК с ограничением на параметры

$$\mathcal{L}_2 = \lambda'(1,\mathrm{s})\mathrm{H}'(\mathrm{s},\mathrm{m})\beta(\mathrm{m},1) =$$

$$= (\lambda_1, \dots, \lambda_\mathrm{s}) \begin{bmatrix} \mathrm{h}_{11} & \dots & \mathrm{h}_{\mathrm{m}1} \\ \dots & \dots & \dots \\ \mathrm{h}_{1\mathrm{s}} & \dots & \mathrm{h}_{\mathrm{ms}} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \dots \\ \beta_\mathrm{m} \end{bmatrix} =$$

$$= (\lambda_1, \dots, \lambda_\mathrm{s}) \begin{bmatrix} \sum\limits_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{h}_{\mathrm{i}1}\beta_\mathrm{i} \\ \dots \\ \sum\limits_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{h}_{\mathrm{is}}\beta_\mathrm{i} \end{bmatrix} = \sum_{\mathrm{j}=1}^{\mathrm{s}} \lambda_\mathrm{j} \left(\sum_{\mathrm{i}=1}^{\mathrm{m}} \mathrm{h}_{\mathrm{ij}}\beta_\mathrm{i} \right) .$$

Система нормальных уравнений

$$\begin{cases} \frac{\partial \mathcal{L}_2}{\partial \beta_1} = \sum\limits_{\mathrm{j}=1}^{\mathrm{s}} \lambda_\mathrm{j}\mathrm{h}_{1\mathrm{j}} = 0, \\ \dots \\ \frac{\partial \mathcal{L}_2}{\partial \beta_\mathrm{m}} = \sum\limits_{\mathrm{j}=1}^{\mathrm{s}} \lambda_\mathrm{j}\mathrm{h}_{\mathrm{mj}} = 0, \end{cases} \quad \leftrightarrow \quad \mathrm{H}\lambda = 0 .$$

Псевдообратная матрица

Определение. A^- - обобщенная обратная матрица, такая, что для любого вектора Y, при котором система $AX = Y$ совместна, $X = A^- Y$ является ее решением.

A^- существует $\iff AA^-A = A \iff A^-AA^- = A^-$.

Обобщенную обратную матрицу, иногда ее называют псевдообратной, можно вычислить через сингулярное разложение матрицы $A = U\Sigma V^T$, где Σ диагональная матрица с сингулярными числами на главной диагонали, матрицы U, V состоят из соответственно левых и правых собственных векторов (с.в. матриц AA^T, A^TA). Сингулярные числа вычисляются как квадратные корни из собственных чисел матриц AA^T, A^TA. Для псевдообратной матрицы справедливо $A^- = V\Sigma^{-1}U^T$, где у матрицы Σ^{-1} на диагонали обратные к ненулевым сингулярным числам.

Задача 1.

Пусть A — симметричная матрица n на n ранга r, B — матрица m на n, BY — m линейных функций от $Y = (y_1, \ldots, y_n)^T$, где $y_i \sim \mathcal{N}(\mu, \sigma)$ независимые величины. Тогда если $BA = 0_{m,n}$, где $0_{m,n}$ матрица m на n из нулей, то BY и Y^TAY независимы.

Пусть $BA = 0$. Используем спектральное разложение симметричной матрицы A ранга r:

$$A = \lambda_1 P_1 P_1^T + \ldots + \lambda_r P_r P_r^T, \quad \lambda_i \neq 0,\ i = 1, \ldots, r.$$

$$BAP_1 = \lambda_1 BP_1 P_1^T P_1 + \ldots + \lambda_r BP_r P_r^T P_1 = 0_{m,1}, \Longrightarrow BP_i = 0_{m,1}$$

Аналогично $BP_i = 0_{m,1}$, т.е. л.ф. BY не зависят от функций $P_i^T Y$ (если $a^T b = 0$, то для Y с независимыми компонентами $\mathbb{E}(a^T Y)(Y^T b) = a^T I b = a^T b = 0$) и не зависят от $Y^T AY$, так как $Y^T AY = \sum \lambda_i (P_i^T Y)^2$.

Задача 2.

а) $X = X(X'X)^{-}X'X$,

б) $I - X(X'X)^{-}X'$ идемпотентна.

$$\text{а)}\quad G = X(I - (X'X)^{-}X'X),$$

$$\Longrightarrow \quad G'G = (I - (X'X)^{-}X'X)'X'X(I - (X'X)^{-}X'X) =$$

$$= (I - (X'X)^{-}X'X)'(X'X - X'X(X'X)^{-}X'X) =$$

$$= (I - (X'X)^{-}X'X)'(X'X - X'X) = 0, \quad \Longrightarrow \quad G = 0.$$

$$\text{б)}\quad (I - X(X'X)^{-}X')(I - X(X'X)^{-}X') =$$

$$= I - X(X'X)^{-}X' - X(X'X)^{-}X' + \underbrace{X(X'X)^{-}X'X}_{X}(X'X)^{-}X' =$$

$$= I - X(X'X)^{-}X'.$$

Theorem

Квадратичная форма ранга n $Y^TAY \sim chi^2 \iff A = A^2$, причем $df = rank(A) = TrA$.

Достаточность. Пусть $A^2 = A$. Применим теорему Фишера-Кочрена к

$$Y^TY = Y^TAY + Y^T(I - A)Y$$

и воспользуемся неравенством Сильвестра

$$rank(A) + rank(C) - rank(AC) \leq n$$

для идемпотентной матрицы $A = A^2$ и $C = I - A$, то

$$rank(A) + rank(I - A) - \underbrace{rank(A(I - A))}_{=0} \leq n$$

$$n = rank(I) = rank(A + I - A) \leq rank(A) + rank(I - A)\,,$$

Таким образом, $n = rank(A) + rank(I - A)$, и по теореме Фишера-Кочрена $Y^TAY \ \sim\ \chi^2(rank(A))$.

Задача 3.

Пусть $\beta = (\beta_1, \ldots, \beta_m)'$ вектор параметров, X - матрица плана ранга r, $(X'X)^-$ обобщенная обратная, $\hat{\beta} = (X'X)^- X'Y$, $\hat{\beta}$ решение уравнения $(X'X)\beta = X'Y$. Матрица H' размерности s на m ранга s. Показать, что распределены независимо $Z = H'\hat{\beta}$ и $R_0^2 = (Y - X\hat{\beta})'(Y - X\hat{\beta})$.

Введем матрицу C такую, что $H = X'C$. Тогда

$$H'\hat{\beta} = \overbrace{C'X(X'X)^- X'}^{B} Y = BY.$$

$$R_0^2 = Y'Y - \hat{\beta}'X'Y - Y'X\hat{\beta} + \hat{\beta}' \underbrace{X'X\hat{\beta}}_{X'Y} = Y'Y - Y'X\hat{\beta}, \qquad (1)$$

$$R_0^2 = Y'Y - Y'X(X'X)^- X'Y = Y'(\underbrace{I - X(X'X)^- X'}_{A})Y = Y'AY\,.$$

Таким образом $BA = C'X(X'X)^- X'(I - X(X'X)^- X') =$

$$= C'(X(X'X)^- X' - \underbrace{X(X'X)^- X'X}_{X}(X'X)^- X') = 0\,.$$

Задача 4.

а) $Z = H'\hat{\beta} \sim \mathcal{N}_s(H'\beta, \sigma^2 D)$, где $D = H'(X'X)^- H$ и
б) $R_0^2 \sim \sigma^2\chi^2(n-r)$, где r – ранг матрицы плана X.

а) Действительно, $E(\hat{\beta} - E\hat{\beta})(\hat{\beta} - E\hat{\beta})' =$

$$= E((X'X)^- X'(Y - EY))((X'X)^- X'(Y - EY))' =$$

$$= (X'X)^- X'\sigma^2 I((X'X)^- X')' =$$

$$= \sigma^2 (X'X)^- X'X(X'X)^- = \sigma^2 (X'X)^- .$$ Отсюда $D = H'(X'X)^- H$.

б) $R_0^2 = Y'(I - X(X'X)^- X')Y$,

матрица $I - X(X'X)^- X'$ идемпотентна,

$$\Longrightarrow R_0^2 \sim \sigma^2\chi^2(n-r),$$

где $\mathrm{Tr}I - \mathrm{Tr}X(X'X)^- X' =$

$$= n - \mathrm{Tr}(X'X)^- X'X =$$

$$= n - \mathrm{rang}(X'X) = n - \mathrm{rang}(X) = n - r .$$

Задача 5.

При дифференцировании $(Y - X\beta)'(Y - X\beta)$ по β при условии $H'\beta = \theta_0$ получена система уравнений (разд. ??):

$$\begin{cases} X'X\beta + H\lambda = X'Y, \\ H'\beta = \theta_0 \end{cases} \tag{2}$$

λ^* и β^* – ее решение. Показать, что

$$R_1^2 = (Y - X\beta^*)'(Y - X\beta^*) = (Z - \theta_0)'D^{-1}(Z - \theta_0) + R_0^2\,.$$

Заметим, что

$$R_1^2 = (Y - X\beta^* + X\hat{\beta} - X\hat{\beta})'(Y - X\beta^* + X\hat{\beta} - X\hat{\beta}) =$$
$$= (Y - X\hat{\beta})'(Y - X\hat{\beta}) + (\hat{\beta} - \beta^*)'X'X(\hat{\beta} - \beta^*)',$$
так как $(\hat{\beta} - \beta^*)'X'(Y - X\hat{\beta}) = 0$ по определению $\hat{\beta}$.

Достаточно показать, что

$$(\hat{\beta} - \beta^*)'X'X(\hat{\beta} - \beta^*)' = (Z - \theta_0)'D^{-1}(Z - \theta_0).$$

Пусть $H = X'XC$, тогда
$D = H'(X'X)^{-}H = C'X'X(X'X)^{-}X'XC = C'X'XC$. Вычитаем из $X'X\hat{\beta} = X'Y$ уравнение (2) $X'X\beta^* + H\lambda^* = X'Y$.

$$X'X(\hat{\beta} - \beta^*) = H\lambda^* = X'XC\lambda^*, \quad \hat{\beta} - \beta^* = (X'X)^{-}X'XC\lambda^*,$$
$$(\hat{\beta} - \beta^*)'X'X(\hat{\beta} - \beta^*) = \underbrace{\lambda^{*'}C'X'X(X'X)^{-}}_{(\hat{\beta}-\beta^*)'}X'XC\lambda^* =$$
$$= \lambda^{*'}C'X'XC\lambda^* = \lambda^{*'}D\lambda^*$$
$$Z - \theta_0 = H'\hat{\beta} - H'\beta^* = H'(\hat{\beta} - \beta^*) = C'X'X(X'X)^{-}X'XC\lambda^* =$$
$$= C'X'XC\lambda^* = D\lambda^*$$
$$\Longrightarrow \lambda^* = D^{-1}(Z - \theta_0),$$
$$\Longrightarrow \lambda^{*'}D\lambda^* = (Z - \theta_0)'D^{-1}DD^{-1}(Z - \theta_0) = (Z - \theta_0)'D^{-1}(Z - \theta_0).$$