

## КАНОНИЧЕСКАЯ МОДЕЛЬ ФАКТОРНОГО АНАЛИЗА

$$\text{Пусть: } \vec{\xi} = \vec{\alpha}^{(1)} f^{(1)} + \vec{\alpha}^{(2)} f^{(2)} + \dots + \vec{\alpha}^{(m)} f^{(m)} + \vec{\varepsilon} = \alpha \vec{f} + \vec{\varepsilon}, \quad (1)$$

где:

$\vec{\xi} = (\xi^{(1)}, \xi^{(2)}, \dots, \xi^{(k)})$  - вектор центрированных исходных признаков;

$\vec{f} = (f^{(1)}, f^{(2)}, \dots, f^{(m)})$  - вектор центрированных и нормированных обобщенных факторов, некоррелированных между собой;

$\alpha = (\vec{\alpha}^{(1)}, \vec{\alpha}^{(2)}, \dots, \vec{\alpha}^{(m)})$  - матрица факторных нагрузок;

$\vec{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k)$  - вектор центрированных характерных факторов, некоррелированных как между собой, так и с обобщенными факторами.

Пусть  $A = M(\vec{\xi}\vec{\xi}^T)$  - ковариационная матрица исходных признаков, а  $\Sigma = M(\vec{\varepsilon}\vec{\varepsilon}^T)$  - диагональная матрица ковариаций характерных факторов, тогда, согласно канонической модели факторного анализа:

$$A = \alpha \alpha^T + \Sigma \quad (2)$$

или:

$$\begin{cases} \text{cov}(\xi_i, \xi_j) = \sum_{s=1}^m \alpha_i^{(s)} \alpha_j^{(s)}, & i = \overline{1, k}, \quad j = \overline{1, k}, \quad i \neq j \\ D(\xi_i) = \sum_{s=1}^m (\alpha_i^{(s)})^2 + D(\varepsilon_i), & i = \overline{1, k} \end{cases} \quad (3)$$

Таким образом, ковариации исходных признаков полностью воспроизводятся матрицей нагрузок, а для воспроизведения их дисперсий нужны также дисперсии характерных факторов. Заметим, что ,

$$\text{cov}(\xi_j, f_s) = \alpha_j^{(s)}, \quad j = \overline{1, k}, \quad s = \overline{1, m}.$$

Модель (1) при условиях (2,3) носит название канонической модели факторного анализа. Заметим, что если решение системы (3) существует, то оно определяется с точностью до ортогонального преобразования матрицы факторных нагрузок.

## ОЦЕНИВАНИЕ ПАРАМЕТРОВ МОДЕЛИ ФАКТОРНОГО АНАЛИЗА МЕТОДОМ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Будем полагать, что обобщенные и характерные факторы имеют нормальное распределение (это соответственно означает, что исходные признаки имеют многомерное нормальное распределение). Для того, чтобы матрица факторных нагрузок определялась

однозначно потребуем, чтобы она удовлетворяла условию:  $\alpha^T \Sigma^{-1} \alpha = J$ , где  $J$  - диагональная матрица с упорядоченными по убыванию диагональными элементами. Заметим, что количество условий наложенных на параметры равно  $\frac{k(k+1)}{2} + \frac{m(m-1)}{2}$ , а число неизвестных параметров матрицы факторных нагрузок и матрицы дисперсий характерных факторов равно  $k(m+1)$ . Чтобы задача не являлась неопределенной необходимо, чтобы  $k(k+1)/2 + m(m-1)/2 \geq k(m+1)$  или  $(k-m)^2 \geq (k+m)$ . Поскольку, при  $(k-m)^2 = (k+m)$  решение теряет статистическую значимость, то максимальное число параметров для данной модели должно удовлетворять условию  $(k-m)^2 > (k+m)$ . Например, при  $k=5, k=6$  получаем, что  $m \leq 2$ .

Используя для нахождения оценок матрицы факторных нагрузок и дисперсионной матрицы характерных факторов метод максимального правдоподобия, получим следующие уравнения для оценок  $\hat{A}, \hat{\alpha}, \hat{\Sigma}, \hat{J}$ :

$$\text{diag}(\hat{A} - \bar{A}) = 0 \quad \text{или} \quad \bar{a}_{ii} = \sum_{s=1}^m (\hat{\alpha}_i^{(s)})^2 + \hat{v}_i, \quad i = \overline{1, k}, v_i = D(\varepsilon_i) \quad (1)$$

$$\hat{\alpha} \hat{J} = (\bar{A} - \hat{\Sigma}) \hat{\Sigma}^{-1} \hat{\alpha} \quad (2)$$

Для упрощения записи в дальнейшем знаки оценок для  $A, \alpha, \Sigma, J$  опущены.

Поскольку  $J$  диагональная матрица, то из уравнения (2) вытекает, что оценки векторов факторных нагрузок  $\bar{\alpha}^{(i)}$  должны являться собственными векторами матрицы  $(\bar{A} - \Sigma) \Sigma^{-1}$ , соответствующими собственным числам  $\lambda_i$  и удовлетворяющими условию:

$$\bar{\alpha}^{(i)T} \Sigma^{-1} \bar{\alpha}^{(i)} = \lambda_i. \quad (3)$$

При заданном числе факторов  $m$  итерационная процедура решения системы

$$\begin{cases} \alpha J = (\bar{A} - \Sigma) \Sigma^{-1} \alpha \\ \bar{A} = \alpha \alpha^T + \Sigma \end{cases}$$

может быть построена следующим образом.

Задаемся начальным приближением  $\Sigma_0$  для  $\Sigma$ . Можно взять, например, в качестве диагональных элементов  $\Sigma_0$  величины  $(1 - |r(i)|) \bar{D}_i$ , где  $r(i)$  наибольшая (по абсолютной величине) выборочная корреляция для  $i$ -го исходного признака с другими признаками. Далее находим  $m$  собственных чисел и собственных векторов матрицы  $(\bar{A} - \Sigma_0) \Sigma_0^{-1}$ , упорядоченных по убыванию собственных чисел, нормируем их, используя условие (3) и, соответственно, получаем первое приближение для матрицы факторных нагрузок  $\alpha_1$ .

Далее находим следующее приближение для  $\Sigma$ :  $\text{diag}\Sigma_1 = \text{diag}(\bar{A} - \alpha_1\alpha_1^T)$  и процесс повторяется. Итерационный процесс заканчивают, когда очередное приближение матрицы  $\Sigma$  мало отличается от предыдущего, то есть, если  $\|\Sigma_i - \Sigma_{i-1}\| < \varepsilon$ , где  $\varepsilon$  - заранее заданное число. В качестве нормы можно использовать, например, абсолютную величину разности следов матриц.

Гипотезу о возможности представления исходного вектора признаков с помощью модели факторного анализа с числом факторов равным  $m$  можно проверить, используя следующую статистику:

$$\eta = -\left(n - 1 - \frac{1}{6}(2k + 5) - \frac{2}{3}m\right) \left( \ln \frac{|\bar{A}|}{|\alpha\alpha^T + \Sigma|} - \text{Sp}(A(\alpha\alpha^T + \Sigma)^{-1}) + k \right).$$

При истинности  $H_0$ : допустимо представление исходных признаков в виде  $m$ -факторной модели (матрицы  $\hat{A} = \hat{\alpha}\hat{\alpha}^T + \hat{\Sigma}$  и  $\bar{A}$  статистически не различаются), статистика приближенно имеет распределение  $\chi^2$  с числом степеней свободы  $\nu = \frac{1}{2}((k - m)^2 - (k + m))$

Для оценивания значений факторов канонической модели факторного анализа используют либо метод Бартлетта, либо метод Томпсона.

Согласно методу Бартлетта модель для каждого наблюдения  $X^{(i)}, i = \overline{1, n}$  рассматривается как регрессия величины  $X^{(i)}$  на факторы  $\alpha^{(j)}, j = \overline{1, m}$  с неизвестными коэффициентами  $f_i^{(j)}$ :

$$X^{(i)} = f_i^{(1)}\alpha^{(1)} + f_i^{(2)}\alpha^{(2)} + \dots + f_i^{(m)}\alpha^{(m)} + \varepsilon.$$

Соответственно для значения  $X_j^{(i)}$  признака  $\xi_j$  в  $i$ -ом наблюдении справедливо:

$$X_j^{(i)} = f_i^{(1)}\alpha_j^{(1)} + f_i^{(2)}\alpha_j^{(2)} + \dots + f_i^{(m)}\alpha_j^{(m)} + \varepsilon_j, \quad j = \overline{1, k}.$$

Заметим, что наблюдения неравноточные, характерные факторы в общем случае имеют различные дисперсии  $D(\varepsilon_j) = \nu_j$ . Следовательно, для оценки факторов следует применять обобщенный метод наименьших квадратов, согласно которому:

$$\hat{f}_i = (\hat{f}_i^{(1)}, \hat{f}_i^{(2)}, \dots, \hat{f}_i^{(m)})^T = (\alpha^T \Sigma^{-1} \alpha)^{-1} \alpha^T \Sigma^{-1} X^{(i)T}.$$

Последнее выражение удобнее переписать в виде:

$$\hat{f}^{(i)T} = X^{(i)} \Sigma^{-1} \alpha (\alpha^T \Sigma^{-1} \alpha)^{-1},$$

тогда матрица оценок значений факторов:

$$\hat{F} = X \Sigma^{-1} \alpha (\alpha^T \Sigma^{-1} \alpha)^{-1}.$$

В методе Томпсона, исходя из того, что оценки обобщенных факторов есть линейная комбинация исходных признаков, коэффициенты которой определяются из условия минимума среднеквадратичного отклонения оценки от фактора, то есть:

$$\hat{f}^{(j)} = \beta^{(j)T} \xi = \beta_1^{(j)} \xi_1 + \beta_2^{(j)} \xi_2 + \dots + \beta_k^{(j)} \xi_k, \quad j = \overline{1, m},$$

где вектор коэффициентов удовлетворяет условию:

$$M(f^{(j)} - \hat{f}^{(j)})^2 \rightarrow \min.$$

Таким образом, коэффициенты  $\beta_i^{(j)}, i = \overline{1, k}$  являются коэффициентами линейной регрессии и могут быть выражены через ковариации величин  $\xi_1, \xi_2, \dots, \xi_k$  и  $f^{(j)}$ :

$$\bar{\beta}^{(j)} = A^{-1}(\bar{\xi}) M(f^{(j)} \cdot \bar{\xi}),$$

здесь:  $A$  - матрица ковариаций величин  $\xi_1, \xi_2, \dots, \xi_k$ ,  $M(f^{(j)} \cdot \xi)$  - вектор ковариаций величины  $f^{(j)}$  и величин  $\xi_1, \xi_2, \dots, \xi_k$ .

Соответственно

$$\hat{f}^{(j)} = \xi^T A^{-1}(\xi) M(f^{(j)} \cdot \xi).$$

Заменяя  $A(\xi)$  и  $M(f^{(j)} \cdot \xi)$  их оценками  $\hat{A}(\xi) = \alpha \alpha^T + \Sigma$  и  $\hat{M}(f^{(j)} \cdot \xi) = \alpha^{(j)}$  получим оценку для  $\hat{f}^{(j)}$ :

$$\hat{f}^{(j)} = \xi^T (\alpha \alpha^T + \Sigma)^{-1} \alpha^{(j)}, \quad j = \overline{1, m}.$$

Соответственно оценка для вектора обобщенных факторов:

$$\hat{f}^T = \xi^T (\alpha \alpha^T + \Sigma)^{-1} \alpha.$$

Пусть  $X$  матрица значений исходных признаков размеров  $n \times k$ ,  $F$  - матрица оценок значений обобщенных факторов размеров  $n \times m$ , тогда:

$$\hat{F} = X (\alpha \alpha^T + \Sigma)^{-1} \alpha.$$