

И.Н. Аверина (Врест)

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ МЕТОДА ГЛАВНЫХ КОМПОНЕНТ

Известно, что истинная величина изучаемого явления содержит по крайней мере два компонента: истинную характеристику оцениваемого явления и ошибку измерения. Если измерения проводятся в экономике, то добавляется третья составляющая, зависящая от вариабельности изучаемого объекта. Таким образом зарегистрированное значение может быть представлено в виде суммы  $X_{ji} = X_{mi} + X_{Si} + X_{\theta i}$ , где  $X_{ji}$  - зарегистрированное значение измеряемого признака;  $X_{mi}$  - истинное значение (математическое ожидание) измеряемого признака;  $X_{Si}$  - вариативное значение измеряемого признака;  $X_{\theta i}$  - ошибка измерения при определении  $j$ -го признака у  $i$ -го объекта исследования.

В основу метода главных компонент положена линейная модель. Если  $N$  - число исследуемых объектов,  $n$  - число признаков (измеряемых характеристик объекта), то математическая модель принимает вид:

$$y_j' = \sum_{r=1}^n a_{jr} f_r, \quad (1)$$

где  $r, j = 1, 2, \dots, n$ ;  $f_r$  -  $r$ -я главная компонента;  $a_{jr}$  - вес  $r$ -й компоненты в  $j$ -й переменной;  $y_j'$  - нормированное значение  $j$ -го признака, полученное из модели;  $y_j$  - нормированное значение  $j$ -го признака, полученное из эксперимента, на основе наблюдений.

Если изучают зависимость между  $n$  признаками, то (1) можно переписать в виде  $n$  линейных уравнений, которые свяжут признаки с главными компонентами. Эти  $n$  линейных уравнений составят математическую модель при решении задачи методом главных компонент:

$$\begin{cases} y_1 = a_{11}f_1 + a_{12}f_2 + \dots + a_{1n}f_n; \\ y_2 = a_{21}f_1 + a_{22}f_2 + \dots + a_{2n}f_n; \\ \dots \\ y_n = a_{n1}f_1 + a_{n2}f_2 + \dots + a_{nn}f_n. \end{cases} \quad (2)$$

Итак, при проведении эксперимента получают результаты в виде матрицы  $X$  наблюдаемых величин. Элементы данной матрицы центрируются и нормируются, формируя матрицу  $Y$  порядка  $n \times N$ . В матричной форме это запишется  $Y = AF$  или  $y = Af$ , где матрица  $Y$  - представляет собой совокупность всех  $N$  наблюдаемых значений всех  $n$  параметров, а вектор  $y$  - совокупность  $n$  рассматриваемых признаков. Аналогично, матрица  $F$  включает совокупность всех  $N$  полученных значений всех  $n$  главных компонент, а вектор  $f$  представляет собой совокупность всех  $n$  главных компонент. Математическое описание изучаемого процесса заключается в определении оператора  $A$  данного процесса.