А. В. Бернштейн (Москва, ИСА РАН). **Обобщенная задача анализа главных компонент.**

Пусть $X_1, X_2, \ldots, X_n$ — множество p-мерных векторов. Метод главных компонент решает следующую оптимизационную задачу: найти аффинное подпространство $L(q)$ размерности $q < p$, минимизирующее величину

$$\varepsilon_N = \frac{1}{N} \sum_{i=1}^{N} \|X_i - P_{L(q)} X_i\|^2, \tag{1}$$

в которой $P_{L(q)} X$ есть проекция вектора $X \in \mathbf{R}^p$ на аффинное подпространство $L(q)$. Если минимальное значение величины ε_N (1) мало, то вместо исходных p-мерных векторов $\{X_i\}$ можно рассматривать их проекции $\{P_{L(q)} X_i\}$ на построенное аффинное подпространство, а каждая проекция определяется q коэффициентами разложения по базису, выбранному в $L(q)$. Поэтому метод главных компонент используется в задаче снижения размерности, и он является наилучшим в линейной постановке этой задачи.

В прикладных задачах анализа данных [1] от процедуры снижения размерности обычно требуется, чтобы обеспечивалась близость не только между векторами $\{X_i\}$ и $\{P_{L(q)} X_i\}$, но и значениями заданной векторной функции

$$F(X) = (F_1(X), F_2(X), \ldots, F_m(X))^T. \tag{2}$$

Например [1], в аэродинамических приложениях важной является задача снижения размерности описания профилей крыла, которые рассматриваются как замкнутые кривые в двумерном пространстве и описываются набором координат конечного числа точек на контуре профиля. Методы снижения размерности позволяют снизить размерность описания в десятки раз, обеспечивая малую «геометрическую» ошибку между описанием X исходного профиля и описанием X^* редуцированного профиля, построенного по его малоразмерному описанию. Однако при этом может не обеспечиваться близость векторов аэродинамических характеристик $F(X)$ и $F(X^*)$.

Поэтому рассматривается функциональная задача снижения размерности, качество решения которой описывается функционалом вида

$$\varepsilon_N(F) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \biggl\{ \theta_0^2 \, \|X_i - X_i^*\|^2 + \sum_{j=1}^{m} \theta_j^2 |F_j(X_i) - F_j(X_i^*)|^2 \biggr\}}, \tag{3}$$

где $\theta_0, \theta_1, \ldots, \theta_m$ — заданные неотрицательные веса, удовлетворяющие условиям $\sum_{j=0}^{m} \theta_j^2 = 1$, $\theta_0 > 0$, и эта задача является нелинейной.

Большинство современных методов нелинейного снижения размерности основаны на предварительном изучении локального поведения выборки в окрестностях ее точек, в которых для каждой точки по подвыборке ее «близких соседей» решается локальная задача снижения размерности с использованием метода главных компонент.

В функциональной задаче снижения размерности также используется локальный подход, и если вектор-функция F (2) гладкая, то в окрестности произвольной точки X_0 можно рассмотреть линеаризованный функционал $\varepsilon_N(F)$ (3) заменой разностей $F(X) - F(X^*)$ линейными приращениями $(\nabla F(X_0))^T (X - X^*)$, где $\nabla F(X_0)$ — матрица градиентов порядка $m \times p$, j-й столбец которой есть градиент функции F_j в точке X_0, $j = 1, 2, \ldots, m$.

Поэтому необходимо решать обобщенную задачу анализа главных компонент, которая формулируется следующим образом.

Пусть $X_1, X_2, \ldots, X_N$ — заданное множество p-мерных векторов, и A — заданная матрица размера $n \times p$, $n \geqslant p$, ранга p. Ищутся аффинное подпространство $L(q)$

размерности $q < p$ в $\mathbf{R}^p$ и аффинное линейное преобразование $B_{L(q)}$ из $\mathbf{R}^p$ в $L(q)$, минимизирующие величину $S(L(q), B_{L(q)}) = N^{-1}\sum_{i=1}^{N}\|AX_i - AB_{L(q)}X_i\|^2$.

Теорема. *Пусть e_0 и Σ_X — построенные по данным выборочные вектор средних и ковариационная матрица. Пусть $\mathbf{E}_{0,L(q)}$ — матрица, столбцы которой состоят из первых q ортонормированных собственных векторов $\{e_1, e_2, \ldots, e_q\} \in \mathbf{R}^p$ матрицы $(\Sigma_X \times A^T A)$, записанных в порядке убывания соответствующих собственных чисел. Обозначим $L_0(q) = \{X \in \mathbf{R}^p\colon\ X = e_0 + \sum_{j=1}^{q} y_j e_j,\ y_i \in \mathbf{R}^1,\ i = 1, 2, \ldots, q\}$, — q-мерное аффинное подпространство, и $B_{0,L(q)}\colon\ X \to e_0 + \mathbf{E}_{0,L(q)} \times (\mathbf{E}_{0,L(q)})^T \times \Sigma_A \times (X - e_0)$, — аффинное линейное отображение из $\mathbf{R}^p$ в $L_0(q)$.*

Тогда $\{L_0(q), B_{0,L(q)}\}$ минимизируют по $L(q)$ и $B_{L(q)}$ величину $S(L(q), B_{L(q)})$.

СПИСОК ЛИТЕРАТУРЫ

1. *Бернштейн А.В., Кулешов А.П.* Когнитивные технологии в проблеме снижения размерности описания геометрических объектов. — Информационные технологии и вычислительные системы, 2008, № 2, с. 6–19.