Сингулярное разложение (SVD)

Опр. 1 SVD — декомпозиция вещественной матрицы с целью её приведение к каноническому виду

Теорема 1. $A \in \mathbb{R}^{m \times n}$. $\exists$ такие ОНБ $e^k_{k=1}{}^n \subset \mathbb{C}^m$ и положит. числа $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_r$, $0 \leqslant r \leqslant \min(m, n)$, что

$$Ae^k = \begin{cases} \sigma_k q^k, & k \leqslant r \\ 0, & k > r \end{cases}$$

$\sigma_1, \ldots, \sigma_r$ — синг. числа матрицы A

$\{e^k\}_{k=1}^n$, $\{q^k\}_{k=1}^m$ — базисы (сингулярные) A

$\operatorname{Im}(A)$ - размерность r, которая ранг матрицы A

Д-во: Матрица $A*A$ самосопр. и неотр. $((A^*A)^* = A^*A$; $(A^*Ax, x) = (Ax, Ax) \geqslant 0)$

$\Longrightarrow$ $\exists\, e^k_{k=1}{}^n$ - ОНБ собств. векторов матрицы A^*A $\Longrightarrow$

$\Longrightarrow$ $A^*Ae^k = \sigma_k^2 e^k$, $\sigma_k^2 \geqslant 0$, $k = 1, 2, \ldots, n$ $\Longrightarrow$

$\Longrightarrow$ $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_r$, $0 \leqslant r \leqslant 0$, $\sigma_{r+1} = \ldots = \sigma_n = 0$

Положим $z^k = Ae^k$ для $l = \overline{1, r}$ и

$$(z^p, z^q) = (Ae^p, Ae^q) = (A^*Ae^p, e^q) = \sigma_p^2 (e^p, e^q)$$

$$\Longrightarrow (z^p, z^q) = \begin{cases} \sigma_p^2, & p = q \\ 0, & p \neq q \end{cases} \Longrightarrow$$ векторы $q^k = \sigma_k^{-1} Ae^k$, $k = \overline{1, r}$,

образуют ортонормированную систему в пространстве $\mathbb{C}^m$. Если $r < m$, дополним её произвольно векторами q^k $k = \overline{r+1, m}$ до ОНБ пр-ва $\mathbb{C}^m$. Из этого следует утв. теоремы

Опр. 1.2 $A = V\Sigma W^*$ – сингулярное выражение матрицы A, где $\Sigma \in \mathbb{R}^{m\times n}$, у которой эл-ты, лежт. на главной диагонали – это сингулярные числа (остальные нулевые)

$$V = \{q^k\}_{k=1}^{m} \in \mathbb{R}^{m\times m}, \quad W = \{e^k\}_{k=1}^{n} \in \mathbb{R}^{n\times n}$$

Метод главных компонент

Опр 2.1 Метод главных компонент – один из основных способов уменьшить размерность данных, потеряв наименьшее кол-во инф-и

Теорема 2.1 Для данной матрицы A $\exists A_k$ – ее апроксимация: $rang(A_k) = k \leq rang(A)$ и $\forall B$, $B \in \mathbb{R}^{m\times n}$, $rang(B) = k$: $\|A - A_k\|_F \leq \|A - B\|_F$, где $\|.\|_F$ – норма Фробениуса

Формальная постановка задачи

Пусть имеется n числовых признаков $f_j(x)$ $j = \overline{1,n}$.
Объекты обуч. выборки отождествляем с их призн. опис-ями: $x_i(f_1(x_i), \dots, f_n(x_i))$, $i = \overline{1,\ell}$
Рассмотрим матрицу F, строки которых соотв. призн. опис. обуч. объектов:

$$F_{\ell\times n} = \begin{bmatrix} f_1(x_1) & \dots & f_n(x_1) \\ \dots & \dots & \dots \\ f_1(x_n) & \dots & f_n(x_n) \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_\ell \end{bmatrix}$$

Обозначим через $z_i = (g_1(x_i), \dots, g_m(x_i))$ признаковые описания тех же объектов в новом пр-ве $Z = \mathbb{R}^m$, $m < n$:

$$G_{\ell \times n} = \begin{bmatrix} g_1(x_1) & \cdots & g_n(x_1) \\ & \text{''} & \\ g_1(x_n) & \cdots & g_n(x_n) \end{bmatrix} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$$

Потребуем, чтобы исх. призн. опис. можно было восстановить по новым описаниям с помощью некоторого лин. преобр., опр. матрицей $U = (u_{js})_{n \times m}$: $\hat{f}_j(x) = \sum_{s=1}^{m} g_s(x) u_{js}$, $j = \overline{1,n}$, $x \in X$ или в векторной записи $\hat{x} = zU^T$. Восстановленное опис-е $\hat{x}$ не обязано в точности совпадать с исх. описанием x, но их отличия на объектах обучающей выборки должно быть как можно меньше при выбранной размерности m. Будем искать одновременно и матрицу новых признаковых описаний G, и матрицу лин. преобр U, при которых суммарная невязка $\Delta^2(G,U)$ восст. описаний минимальна: $\Delta^2(G,U) = \sum_{i=1}^{\ell} \| \hat{x}_i U^T - x_i \|^2 = \| GU^T - F \|^2 \to \min_{G,U}$

G, U невырождены: $rg(G) = rg(U) = m$. Иначе бы $\exists\ \hat{G}\hat{U}^T = GU^T$ с числом столбцов в матрице $\hat{G}$, меньшим m. Поэтому интересны только случаи $m \leq rg(F)$

Опр. 2.2 $u_1, \dots, u_m$ — главные компоненты

Теорема 2.2 Если $m \leq rg(F)$, то минимум $\Delta^2(G,U)$ достигается когда столбцы матрицы U есть собств. векторы F^TF, соотв. m максимальным с.зн. При этом $G = FU$, матрицы U и G ортогональны.