MDS降维算法推导以及代码实现
假设原始样本个数为m，样本维度为d，每一行代表一个样本

$$X=\begin{bmatrix} x_1^{(1)} & x_1^{(2)} & \cdots & x_1^{(d)} \\ x_2^{(1)} & x_2^{(2)} & \cdots & x_2^{(d)} \\ & & \vdots & \\ x_m^{(1)} & x_m^{(2)} & \cdots & x_m^{(d)} \end{bmatrix} \qquad (1)$$

$$\uparrow$$

$$R^{d\times m}$$

样本的距离矩阵 D：

$$D=\begin{bmatrix} 0 & dist_{12} & dist_{13} \cdots dist_{1m} \\ dist_{21} & 0 & dist_{23} \cdots dist_{2m} \\ & & \vdots \\ dist_{m1} & dist_{m2} & dist_{m3} \cdots 0 \end{bmatrix} \qquad (2)$$

$dist_{ij}$ 为第 i 个样本与第 j 个样本之间的距离

降维后（降低 d 的大小）：
假设 Z 是降维后的样本矩阵

$$Z=\begin{bmatrix} z_1^{(1)} & z_1^{(2)} & \cdots & z_1^{(d')} \\ z_2^{(1)} & z_2^{(2)} & \cdots & z_2^{(d')} \\ & & \vdots & \\ z_m^{(1)} & z_m^{(2)} & \cdots & z_m^{(d')} \end{bmatrix} \qquad (3)$$

$$\uparrow$$

$$R^{d'\times m}$$

欧式距离：
$$dist_{ij}=\| z_i - z_j \| \qquad (4)$$
$$\| x \|=(| x_1 |^2 + | x_2 |^2 + ... + | x_n |^2)^{1/2} \qquad (5)$$

我们要使 $dist_{ij}{}^2=\| z_i \|^2 + \| z_j \|^2 - 2z_i{}^T z_j \qquad (6)$

保证降维后样本间距离不变，定义内积矩阵 B，令

$$B=Z^T Z \qquad (7)$$

$$B = Z^T Z = \begin{bmatrix} z_1^{(1)} & z_1^{(2)} & z_1^{(3)} \\ z_2^{(1)} & z_2^{(2)} & z_2^{(3)} \\ z_3^{(1)} & z_3^{(2)} & z_3^{(3)} \end{bmatrix} \begin{bmatrix} z_1^{(1)} & z_2^{(1)} & z_3^{(1)} \\ z_1^{(2)} & z_2^{(2)} & z_2^{(3)} \\ z_1^{(3)} & z_2^{(2)} & z_3^{(3)} \end{bmatrix} =$$

$$\begin{bmatrix} (z_1^{(1)})^2 + (z_2^{(1)})^2 + (z_3^{(1)})^2 & * & * \\ * & (z_1^{(2)})^2 + (z_2^{(2)})^2 + (z_3^{(2)})^2 & * \\ * & * & (z_1^{(3)})^2 + (z_2^{(3)})^2 + (z_3^{(3)})^2 \end{bmatrix}$$

使得 $dist_{ij}{}^2 = \| z_i \|^2 + \| z_j \|^2 - 2 z_i{}^T z_j = b_{ii} + b_{jj} - 2b_{ij}$ (8)

为了方便，假设 $\sum_{i=1}^{m} z_i = 0$，降维后样本被标准化，使得 $z_1 + z_2 + z_3 + ... + z_m = 0$

即 B 的行与列之和为 0

接着，求 D 第 i 行的平方和：

$$\sum_{j=1}^{m} D_{ij}^2 = \sum_{j=1}^{m} (b_{ii} + b_{jj} - 2b_{ij})$$

$$= \sum_{j=1}^{m} b_{ii} + \sum_{j=1}^{m} b_{jj} - 2\sum_{j=1}^{m} b_{ij} \qquad (9)$$

$$= mb_{ii} + tr(B) - 0$$

求 D 第 j 列的平方和：

$$\sum_{i=1}^{m} D_{ij}^2 = \sum_{i=1}^{m} (b_{ii} + b_{jj} - 2b_{ij})$$

$$= \sum_{i=1}^{m} b_{ii} + \sum_{i=1}^{m} b_{jj} - 2\sum_{i=1}^{m} b_{ij} \qquad (10)$$

$$= mb_{jj} + tr(B) - 0$$

D 的所有元素之和为：

$$\sum_{j=1}^{m} \sum_{i=1}^{m} D_{ij}^2 = \sum_{j=1}^{m} \sum_{i=1}^{m} (b_{ii} + b_{jj} - 2b_{ij})$$

$$= 2m * tr(B) \qquad (11)$$

由公式（8）~（11）可以推导：

$$b_{ii} = \frac{\sum_{j=1}^{m} D_{ij}{}^2 - tr(B)}{m} = \frac{\sum_{j=1}^{m} D_{ij}{}^2}{m} - \frac{tr(B)}{m} \qquad (12)$$

$$b_{jj} = \frac{\sum_{i=1}^{m} D_{ij}{}^2 - tr(B)}{m} = \frac{\sum_{i=1}^{m} D_{ij}{}^2}{m} - \frac{tr(B)}{m} \qquad (13)$$

为了计算方便，我们做这样的定义：

$$dist_{i\cdot}^{\ 2} = \frac{\sum_{j=1}^{m} D_{ij}^{\ 2}}{m} \quad (14)$$

$$dist_{\cdot j}^{\ 2} = \frac{\sum_{i=1}^{m} D_{ij}^{\ 2}}{m} \quad (15)$$

因此，由公式（8）~（13）可以得到：

$$b_{ij} = \frac{b_{ii} + b_{jj} - dist_{ij}^{\ 2}}{2} = \frac{dist_{i\cdot}^{\ 2} + dist_{\cdot j}^{\ 2} - dist_{ij}^{\ 2} - 2\frac{tr(B)}{m}}{2} \qquad (16)$$

由上式可知，在已知 D 的情况下，可以求出降维后的内积矩阵 B.

接着我们要做的是根据 B 求出 Z

回顾一下相关定理：

设 A 为 n 阶对称矩阵，则必有正交矩阵 P，使得 $P^{-1}AP = P^{T}AP = \Lambda$（$\Lambda$ 以 A 的 n 个特征值为对角元的对角矩阵）。

根据上述定理，可得 $B = P\Lambda P^{T}$，P 是特征向量矩阵。

$$\Lambda = diag(\lambda_1, \lambda_2, \lambda_3, \ldots\lambda_d), \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_d$$

$$B = P\Lambda P^{T} = \left[p_1\ p_2 \ldots p_m\ \right] diag(\lambda_1, \lambda_2, \lambda_3, \ldots\lambda_d) \begin{bmatrix} p_1^{\ T} \\ p_2^{\ T} \\ \vdots \\ p_m^{\ T} \end{bmatrix} \qquad (17)$$

$$= \lambda_1 p_1 p_1^{\ T} + \lambda_2 p_2 p_2^{\ T} + \ldots + \lambda_m p_m p_m^{\ T}$$

假定有 d^{*} 个非零特征值，则

$$\Lambda_{*} = diag(\lambda_1, \lambda_2, \lambda_3, \ldots\lambda_{d^{*}}), \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{d^{*}}$$

根据 $B = P\Lambda P^{T}$ 和公式（7），可推出：

$$Z^{T}Z = P_{*}\Lambda P_{*}^{T} = (P_{*}\sqrt{\Lambda})(\sqrt{\Lambda}P_{*}^{T})$$

即 $Z = \Lambda_{*}^{1/2} P_{*}^{T} \in R^{d^{*} \times m}$ $\qquad (18)$