MDS算法流程

- 降维后的数据维度为d’，则降维后的数据矩阵为d’×m，记为Z，每一列表示一个样本记为z_i。MDS算法的目的就是要使得降维后的距离矩阵仍为D，即$\|z_i - z_j\| = dist_{ij}$。
- 令$B = Z^T Z$，则$b_{ij} = z_i^T z_j$，有
- $dist_{ij}^2 = \boxed{\|z_i - z_j\|^2 = \|z_i\|^2 + \|z_j\|^2 - 2z_i^T z_j} = b_{ii} + b_{jj} - 2b_{ij}$

注释：假设有两个向量$(x_1, y_1, z_1)^T, (x_2, y_2, z_2)^T$

$$则\left\|\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} - \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}\right\|^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2$$

$$= \left\|\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right\|^2 + \left\|\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right\|^2 - 2(x_1, y_1, z_1)\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$$

MDS算法流程

令降维后的样本中心化，即每一个维度的均值均为0. 可以得到

$$\sum_{i=1}^{m} b_{i1} = \left(b_{11} + b_{21} + b_{31} + \cdots\right) = \sum_{i=1}^{m} \left(z_i^T z_1\right) = \left(\sum_{i=1}^{m} \left(z_i^T\right)\right) z_1 = 0$$

即$\sum_{i=1}^{m} b_{ij} = 0$

同理$\sum_{j=1}^{m} b_{ij} = 0$

MDS算法流程

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

$$= tr\left(B\right) + mb_{jj}\left(\begin{array}{l} j\text{是常数，}tr\left(B\right)\text{表示矩阵}B \\ \text{主对角线元素相加} \end{array}\right)$$

$$\sum_{j=1}^{m} dist_{ij}^{2} = tr\left(B\right) + mb_{ii}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{m} dist_{ij}^{2} = \sum_{i=1}^{m}\left(tr\left(B\right) + mb_{ii}\right) = \sum_{i=1}^{m} tr\left(B\right) + m\sum_{i=1}^{m} b_{ii} = 2m\ tr\left(B\right)$$

MDS算法流程

$$dist_{i\cdot}^2 = \frac{1}{m}\sum_{j=1}^{m} dist_{ij}^2 , \qquad (10.7)$$

$$dist_{\cdot j}^2 = \frac{1}{m}\sum_{i=1}^{m} dist_{ij}^2 , \qquad (10.8)$$

$$dist_{\cdot\cdot}^2 = \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m} dist_{ij}^2 , \qquad (10.9)$$

西瓜书上这几个公式是规定，即规定了左边这三个符号来表示右边的内容，通过以上公式可以得到

$$b_{ij} = -\frac{1}{2}(dist_{ij}^2 - dist_{i\cdot}^2 - dist_{\cdot j}^2 + dist_{\cdot\cdot}^2) ,$$

MDS算法流程

- 通过以上计算可以得到矩阵B

$B = Z^T Z$,所以B是实对称矩阵，可以进行特征值分解

$B = V\Lambda V^T = V\left(\Lambda^{1/2}\right)^T \Lambda^{1/2} V^T$

$Z = \Lambda^{1/2} V^T$,

如果有d^*个非零特征值，这些特征值构成对角矩阵

Λ_*,对应的特征向量矩阵为V_*, 则$Z = \Lambda_*^{1/2} V_*^T \in R^{d^* \times m}$

思考

在奇异值分解中$AA^T = P\Lambda_1 P^T, A^T A = Q\Lambda_2 Q^T$

$A^T A = Q\Sigma^T \Sigma Q^T = Q\Sigma^T P^T P\Sigma Q^T \Rightarrow A = P\Sigma Q^T$

为什么在MDS 算法中计算Z 矩阵时，令$Z = \Lambda^{1/2} V^T$ ？

在奇异值分解中，公式两边都是确定的，必须保证两边可以相等，但在MDS算法中，矩阵Z是不确定的，只要矩阵Z可以满足使得降维后的各样本点之间的距离与降维前的距离相等就可以了，所以其形式可以根据情况改变，按照以上方式确定的矩阵Z各个行之间是正交的，这也符合数据各个维度正交的直观印象(比如x轴和y轴正交)。