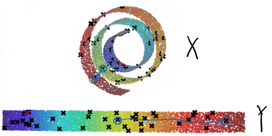
Semi-definite embedding (SDE)

Again, given data $X \in \mathbb{R}^{d \times n}$, we want to find $Y \in \mathbb{R}^{p \times n}$, $p < d$.




The idea of SDE is to unfold the manifold so

(1) Euclidean distances of neighbor points are preserved
(local structures will be retained in low-dimensional space)

(2) maximise the pairwise distance of every pair of points
(pulling points a far away from each other as possible so this manifold can unfolded)

For (1):
$$\boxed{|\underset{\sim}{x}_i - \underset{\sim}{x}_j|^2 = |\underset{\sim}{y}_i - \underset{\sim}{y}_j|^2 \quad \text{if } \eta_{ij} = 1}$$

$$X = [\underset{\sim}{x}_1, \underset{\sim}{x}_2 \cdots \underset{\sim}{x}_n] \qquad Y = [\underset{\sim}{y}_1, \cdots \underset{\sim}{y}_n]$$

$$\eta_{ij} = \begin{cases} 1 & \text{if } i \text{ is a neighbor of } j \\ 0 & \text{otherwise} \end{cases}$$

If we expand the constraint

$$\boxed{\begin{aligned} |\underset{\sim}{x}_i - \underset{\sim}{x}_j|^2 &= \underset{\sim}{y}_i^2 + \underset{\sim}{y}_j^2 - 2\underset{\sim}{y}_i^T\underset{\sim}{y}_j \\ &= \underset{\sim}{y}_i^T\underset{\sim}{y}_i + \underset{\sim}{y}_j^T\underset{\sim}{y}_j - 2\underset{\sim}{y}_i^T\underset{\sim}{y}_j \\ &= K_{ii} + K_{jj} - 2K_{ij} \end{aligned}}$$

K is the kernel, which should satisfy

— K should be a symmetric positive semi-definite matrix, denote by $K \succeq 0$

— K should be centralised, meaning

$$\sum_{ij} K_{ij} = 0 = \sum_{ij} \underset{\sim}{y}_i^T \underset{\sim}{y}_j$$

$$K = \begin{pmatrix} K_{11}, & \cdots, & K_{1n} \\ \vdots & & \\ K_{n1}, & \cdots, & K_{nn} \end{pmatrix} = \begin{pmatrix} \underset{\sim}{y}_1^T\underset{\sim}{y}_1, & \cdots, & \underset{\sim}{y}_1^T\underset{\sim}{y}_n \\ \vdots & \ddots & \vdots \\ \underset{\sim}{y}_n^T\underset{\sim}{y}_1, & \cdots, & \underset{\sim}{y}_n^T\underset{\sim}{y}_n \end{pmatrix}$$

For (2): we want to maximise the pairwise distance in low-dimensional space

$$\begin{aligned} \max_Y \ & \frac{1}{2}\sum_{ij} |\underset{\sim}{y}_i - \underset{\sim}{y}_j|^2 \\ &= \frac{1}{2}\sum_{ij}(\underset{\sim}{y}_i^T\underset{\sim}{y}_i + \underset{\sim}{y}_j^T\underset{\sim}{y}_j - 2\underset{\sim}{y}_i^T\underset{\sim}{y}_j) \\ &= \frac{1}{2}\Big(\sum_i \underset{\sim}{y}_i^T\underset{\sim}{y}_i + \sum_j \underset{\sim}{y}_j^T\underset{\sim}{y}_j - \cancel{2\underbrace{\sum_{ij}\underset{\sim}{y}_i^T\underset{\sim}{y}_j}_{0}}\Big) \\ &= \underbrace{\sum_i \underset{\sim}{y}_i^T\underset{\sim}{y}_i}_{\text{variance}} \\ &= \mathrm{Trace}(K) \end{aligned}$$

SDE is also known as maximum variance unfolding

In summary, we put all together and have the optimisation

$$\boxed{\begin{aligned} &\max_K \ \mathrm{Trace}(K) \qquad \text{Semi-definite programming} \\ &\text{s.t.} \quad K \succeq 0, \ \sum_{ij} K_{ij} = 0 \\ &\qquad |\underset{\sim}{x}_i - \underset{\sim}{x}_j|^2 = K_{ii} + K_{jj} - 2K_{ij} \quad \text{if } \eta_{ij} = 1 \end{aligned}}$$

Once we have K, we can compute embedding Y from

$$Y = \Sigma V^T$$

where V consists of eigenvectors and $\Sigma$ is a diagonal matrix of $\sqrt{\text{eigenvalues}}$ of K