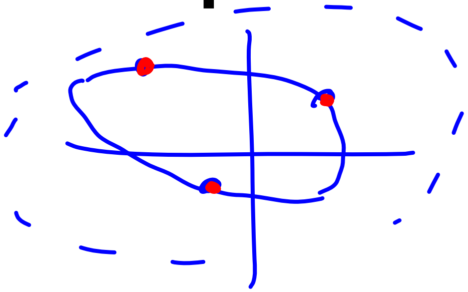
Ex: minimum volume ellipsoid



- Given points $x_1, x_2, \ldots, x_k$
- $\min_{A, x_C}$ vol(A) s.t.

$$(x_i - x_C)^T A (x_i - x_C) \le 1 \quad i = 1, \ldots, k$$

$$A \in S^{n*n}$$

$$A \succcurlyeq 0$$

$$\mathrm{vol}(A) \propto \frac{1}{\sqrt{|A|}}$$

$$\ln \mathrm{vol}(A) = -\frac{1}{2} \ln |A| + \text{const}$$

$$\min \; -\ln|A| \; \text{s.t}$$

$$\{\text{constraints}\}$$

Schur complement

- Symmetric block matrix M = $\begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$
- Schur complement is S = $C - B^T A^{-1} B$
- M $\overset{\succ}{\cancel{\succeq}}$ 0 iff $A \succ 0$ and $S \succ 0$

Back to min-volume ellipsoid

- $\max_{A, x_C} \log |A|$ s.t.
  $(x_i - x_C)^T A (x_i - x_C) \le 1 \quad i = 1, \ldots, k$
  $A = A^T,\ A \succcurlyeq 0$

$y_i = \begin{pmatrix} x_i \\ 1 \end{pmatrix}$

- $\max_{A,B,u,z} \log |A|$ s.t.

$B = \begin{pmatrix} A & -u \\ -u^T & z \end{pmatrix}$

$B = B^T$

$B \succcurlyeq 0$

$y_i^T B y_i \le 1$ for all i

$\boxed{x_i^T A x_i - 2u^T x_i + z \le 1}$

$A^* \succcurlyeq 0$

$A^* \quad B^* \quad u^* \quad z^*$ opt

$z^* - u^{*T} A^{*-1} u^* \succcurlyeq 0$

Def. $x_c^* = A^{*-1} u^*$

$z^* \ge u^{*T} A^{*-1} u^*$

$z^* = u^{*T} A^{*-1} u^*$

$\Rightarrow x_i^T A^* x_i - 2 (A^* x_c^*)^T x_i + x_c^{*T} A^* \cancel{A^{*-1}} \cancel{A^*} x_c^* \le 1$

$(x_i - x_c^*)^T A^* (x_i - x_c^*) \le 1$

Ex: manifold learning

$\leftarrow$ big

- Given points $x_1, \ldots, x_m$ $\in \mathbb{R}^n$

$d \ll n$

- Find points $y_1, \ldots, y_m$ $\in \mathbb{R}^d$
- Preserving **local geometry**
  - neighbor edges N $(i,j) \in N \iff$ preserve geometry btwn x_i, x_j
  - distances $\|x_i - x_j\| = \|y_i - y_j\|$ $(i,j) \in N$
- If we preserve distances among x_i, x_j, x_ℓ

  we also preserve angles $(x_i - x_j) \cdot (x_i - x_\ell)$

$x_i \in \mathbb{R}^n \rightarrow z_i \in \mathbb{R}^n \rightarrow y_i \in \mathbb{R}^d$

Step 1: “embed” R^n into R^n

- While preserving local distances, move points to make manifold as flat as possible
- max $\sum_{i=1}^{m} \sum_{j=1}^{m} \|z_i - z_j\|^2 = 2\sum_i \sum_j (z_i^T z_i) - 2\cancel{\underbrace{\sum_i \sum_j z_i^T z_j}}$

  $= 2m \sum_i z_i^T z_i$

  $\sum_i \sum_j z_i^T z_j = \sum_i z_i^T \underbrace{\left(\sum_j z_j\right)}_{=0}$

  s.t. $\|z_i - z_j\| = \|x_i - x_j\| \quad \forall (i,j) \in N$

  $\sum_{i=1}^{m} z_i = \vec{0}$

Step 2: reduce to R^d

Z

m exs

n dim

z_i^T

- Now that manifold is flat, just use PCA:

orthonormal

Define U, S: S diagonal, $\sum_i z_i z_i^T \approx \underbrace{U S U^T}$

$\mathbb{R}^{n\times d}$ $\quad$ $\mathbb{R}^{d\times d}$

$ZZ^T = K = Q \Sigma Q^T$

Take $y_i = U^T z_i = US$

$Z = Q\sqrt{\Sigma}$

Maximizing variance

- max $\sum_i z_i^T z_i$ s.t. $\sum_i z_i = \vec{0}$ $\|z_i - z_j\|^2 = \|x_i - x_j\|^2 \; (i,j) \in N$

K = "kernel matrix" $K_{ij} = z_i \cdot z_j$ $\quad K = ZZ^T \Rightarrow K \succeq 0$

$\|z_i - z_j\|^2 = z_i \cdot z_i - 2 z_i \cdot z_j + z_j \cdot z_j = K_{ii} - 2K_{ij} + K_{jj} = \|x_i - x_j\|^2$
$\forall (i,j) \in N$

$\rightarrow \sum_{i=1}^{m} \sum_{j=1}^{m} K_{ij} = 0 \iff \sum_i \sum_j z_i^T z_j = 0$

$\iff \left(\sum_i z_i\right)^T \left(\sum_j z_j\right) = 0 \iff \sum_i z_i = \vec{0}$

- max $\operatorname{tr}(K)$ s.t. $\sum_{ij} K_{ij} = 0$ $\quad K \succeq 0 \quad K = K^T$

$K_{ii} - 2K_{ij} + K_{jj} = \|x_i - x_j\|^2 \quad \forall (i,j) \in N$

Summary

- Solve SDP to "embed" R^n into R^n
- Use PCA to embed R^n into R^d
- Called "semidefinite embedding" or "maximum variance unfolding"

- Problems?

Problem: solving SDP

- Kernel matrix K $\approx X X^T$
- Idea: suppose we know a subspace that preserves geometry

A

column orthonormal

$$\tilde{K} \approx A \underbrace{A^T X X^T A}_{J} A^T \approx A J A^T = \tilde{K}$$

smaller

$$\tilde{K} \succeq 0 \iff J \succeq 0$$

Side note: non-Euclidean

$K_{ii} - 2K_{ij} + K_{jj} = d(x_i, x_j)$ might not be feasible

- If original distances are not Euclidean, might not be able to duplicate them exactly in Euclidean R^n
- Would need to soften constraints: **approximately** preserve local distances

$K_{ii} - 2K_{ij} + K_{jj} \leq d(x_i, x_j) \quad \forall (i,j) \in N$