

## 宿題 1

適当な類似度行列に対して局所性保存射影を実装する。

目的関数は,

$$J = \frac{1}{2} \sum_{i,i'=1}^n W_{i,i'} \|\mathbf{T}\mathbf{x}_i - \mathbf{T}\mathbf{x}_{i'}\|^2 = \text{tr}(\mathbf{T}\mathbf{X}\mathbf{L}\mathbf{X}^T\mathbf{T}^T) \quad (1)$$

であり, 拘束条件は

$$\mathbf{T}\mathbf{X}\mathbf{D}\mathbf{X}^T\mathbf{T}^T = \mathbf{I} \quad (2)$$

である。 $J$  を最小にするような  $\mathbf{T}$  を求めることを考える。これはまず, 次の一般化固有値問題を解けばよい。

$$\mathbf{X}\mathbf{L}\mathbf{X}^T\boldsymbol{\xi} = \lambda\mathbf{X}\mathbf{D}\mathbf{X}^T\boldsymbol{\xi} \quad (3)$$

$$\lambda_1 \geq \dots \geq \lambda_d \quad (4)$$

$$\boldsymbol{\xi}_j^T\mathbf{X}\mathbf{D}\mathbf{X}^T\boldsymbol{\xi}_j = 1 \quad (5)$$

ここから, 下位  $m$  個の一般化固有ベクトルを並べ,

$$\mathbf{T} = \mathbf{T}_{\text{LPP}} = (\boldsymbol{\xi}_d, \dots, \boldsymbol{\xi}_{d-m+1})^T \quad (6)$$

とすればよい。

また, 類似度行列にはガウスカーネル

$$K(\mathbf{x}, \mathbf{c}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{c}\|^2}{2h^2}\right) \quad (7)$$

$$\mathbf{W} = \mathbf{K} = \begin{bmatrix} K(\mathbf{x}_1, \mathbf{x}_1) & \dots & K(\mathbf{x}_n, \mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ K(\mathbf{x}_n, \mathbf{x}_1) & \dots & K(\mathbf{x}_n, \mathbf{x}_n) \end{bmatrix} \quad (8)$$

を用いることとする。

$h = 1.0$  としてこれを実装したものが, ??ページの Listing ??である。結果は図 1, 2 に示した通りである。これより, 元のデータのクラスタ構造がうまく保存されたまま次元が削減されていることがわかる。

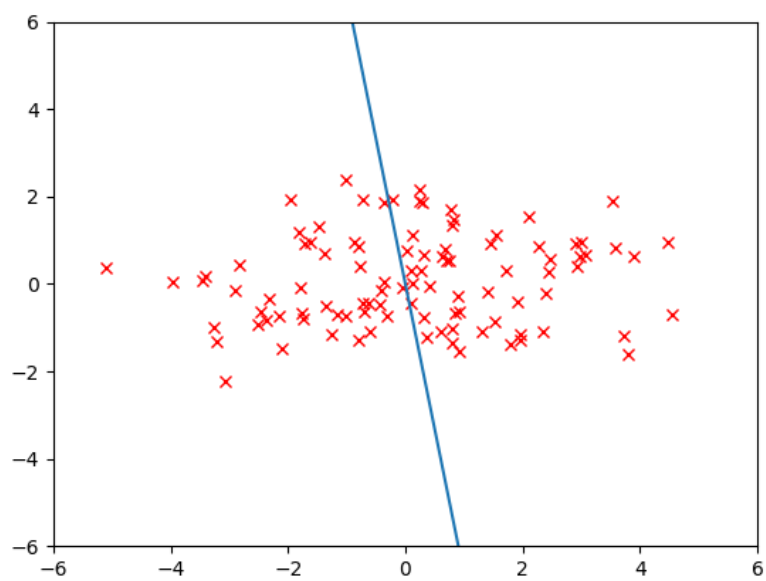


図 1: データ 1 に対する結果

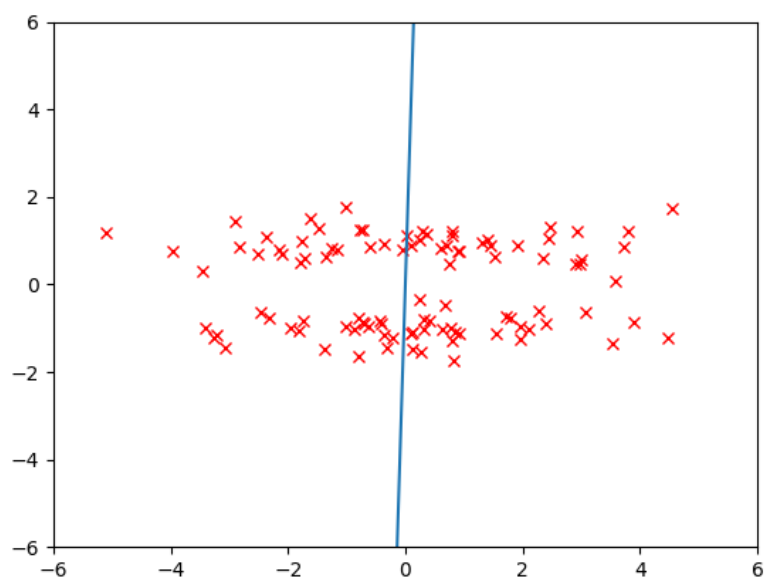


図 2: データ 2 に対する結果