

パターン認識特論

カーネル主成分分析

和田俊和

主成分分析とは？

- 通常のPCAは、正準化された M 個の N 次元データ $\mathbf{x}_i \in R^N$, $\sum_{i=1}^M \mathbf{x}_i = \mathbf{0}$ から計算される共分散行列 $C = \frac{1}{M} \sum_{i=1}^M \mathbf{x}_i \mathbf{x}_i^T$ の固有値問題を解くことと等価である.
- これは C の固有値を λ , 固有ベクトルを \mathbf{v} と表すと, $C\mathbf{v} = \lambda\mathbf{v}$ を満足する λ と \mathbf{v} を求める問題である.

固有ベクトルはデータの線形結合

- この C と固有ベクトル \boldsymbol{v} を掛け合わせると,

$$C\boldsymbol{v} = \frac{1}{M} \sum_{i=1}^M (\boldsymbol{x}_i \cdot \boldsymbol{v}) \boldsymbol{x}_i$$

となる. これと, $C\boldsymbol{v} = \lambda\boldsymbol{v}$ を考慮すると, \boldsymbol{v} が

$\boldsymbol{x}_i (i=1, \dots, M)$ の線形結合で表わされることが分かる. つまり,

$$\boldsymbol{v} = \frac{1}{M\lambda} \sum_{i=1}^M (\boldsymbol{x}_i \cdot \boldsymbol{v}) \boldsymbol{x}_i \quad \text{である.}$$

固有方程式は \mathbf{x}_i に射影しても成立

- $C\mathbf{v} = \lambda\mathbf{v}$ は各 \mathbf{x}_i に射影した後にも成立する.

$$\forall (i = 1 \cdots M), \quad \mathbf{x}_i \square C\mathbf{v} = \lambda(\mathbf{x}_i \square \mathbf{v})$$

- この性質は後のカーネル化で用いる.

高次元化とカーネル化

- データを適当な非線形高次元写像

$$\Phi : R^N \mapsto F$$

によって高次元特徴空間 F に移して, そこで処理を行う.

- 実際に, この写像を行うと, 次元数が高く計算不可能となるため, $k(x, y) = \Phi(x) \cdot \Phi(y)$ となるカーネル関数を用いて内積計算のみを行い, 実際に $\Phi(x)$ を計算しないようにする.

主成分分析の場合1

- 主成分分析の場合には

$$C = \frac{1}{M} \sum_{i=1}^M \Phi(\mathbf{x}_i) \Phi(\mathbf{x}_i)^T$$

と計算することになるが、これは F^2 の空間に属するので、実際には計算できない。

- ここで、前の議論と同じく次式が成り立つ。

$$\forall (i = 1 \cdots M), \quad \Phi(\mathbf{x}_i) \cdot CV = \lambda (\Phi(\mathbf{x}_i) \cdot V)$$

主成分分析の場合2

- また、固有ベクトルはデータの線形結合で表現できることから、

$$V = \sum_{i=1}^M \alpha_i \Phi(x_i)$$

が成り立つ。

- これら2つの式から、次式が成り立つ。

$$\Phi(x_k) \cdot C \sum_{i=1}^M \alpha_i \Phi(x_i) = \lambda \Phi(x_k) \cdot \sum_{i=1}^M \alpha_i \Phi(x_i)$$

主成分分析の場合3

• この式に, $C = \frac{1}{M} \sum_{i=1}^M \Phi(\mathbf{x}_i) \Phi(\mathbf{x}_i)^T$ を代入し,

$$\Phi(\mathbf{x}_k) \cdot \frac{1}{M} \sum_{j=1}^M \Phi(\mathbf{x}_j) \Phi(\mathbf{x}_j)^T \sum_{i=1}^M \alpha_i \Phi(\mathbf{x}_i) = \lambda \Phi(\mathbf{x}_k) \cdot \sum_{i=1}^M \alpha_i \Phi(\mathbf{x}_i)$$

が得られる. これを整理して,

$$\lambda \sum_{i=1}^M \alpha_i \Phi(\mathbf{x}_k) \cdot \Phi(\mathbf{x}_i) = \frac{1}{M} \sum_{i=1}^M \alpha_i (\Phi(\mathbf{x}_k) \cdot \sum_{j=1}^M \Phi(\mathbf{x}_j)) (\Phi(\mathbf{x}_j) \cdot \Phi(\mathbf{x}_j))$$

$$M\lambda K\boldsymbol{\alpha} = K^2\boldsymbol{\alpha} \quad \text{但し} \quad M\lambda\boldsymbol{\alpha} = K\boldsymbol{\alpha}$$

$$K_{ij} \square \Phi(\mathbf{x}_i) \square \Phi(\mathbf{x}_j)$$

注意点

- $M\lambda\alpha = K\alpha$ を解くと, $V = \sum_{i=1}^M \alpha_i \Phi(x_i)$ の係数が分かる.
- V の正規化をしなければならない.

$$\|V\|^2 = V \cdot V = \sum_{i=1}^M \alpha_i \Phi(x_i) \cdot \sum_{j=1}^M \alpha_j \Phi(x_j)$$

$$= \sum_{i,j=1}^M \alpha_i \alpha_j K_{ij} = \alpha \cdot K \alpha = \lambda M \alpha \cdot \alpha \quad \|\alpha\| = \frac{1}{\sqrt{\lambda M}}$$

結果の使い方

- 固有ベクトルへの射影成分は

$$V \cdot \Phi(\mathbf{x}_j) = \sum_{i=1}^M \alpha_i \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j) = \sum_{i=1}^M \alpha_i K_{ij}$$

と計算できる.

- 一般のベクトル \mathbf{t} の射影成分は,

$$V \cdot \Phi(\mathbf{t}) = \sum_{i=1}^M \alpha_i \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{t}) = \sum_{i=1}^M \alpha_i k(\mathbf{x}_i, \mathbf{t})$$

と書ける.

正準化1

• $\sum_{i=1}^M \bar{\Phi}(\mathbf{x}_i) = 0$ を保証する正準化処理は、カーネルの変形によってなされる.

$$\begin{aligned} \bar{K}_{ij} &= \bar{\Phi}(\mathbf{x}_i)^T \bar{\Phi}(\mathbf{x}_j) \\ &= \left(\Phi(\mathbf{x}_i) - \frac{1}{M} \sum_{n=1}^M \Phi(\mathbf{x}_n) \right)^T \left(\Phi(\mathbf{x}_j) - \frac{1}{M} \sum_{m=1}^M \Phi(\mathbf{x}_m) \right) \\ &= \Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}_j) - \frac{1}{M} \sum_{n=1}^M \Phi(\mathbf{x}_n)^T \Phi(\mathbf{x}_j) \\ &\quad - \frac{1}{M} \sum_{m=1}^M \Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}_m) + \frac{1}{M^2} \sum_{n,m=1}^M \Phi(\mathbf{x}_n)^T \Phi(\mathbf{x}_m) \\ &= K_{ij} - \frac{1}{M} \sum_{n=1}^M K_{nj} - \frac{1}{M} \sum_{m=1}^M K_{im} + \frac{1}{M^2} \sum_{n,m=1}^M K_{nm} \end{aligned}$$

• この変形により、元のデータが正規化された場合のカーネル関数が得られる.

正準化2

- このカーネル関数から固有値と固有ベクトルが求められる。
- また、固有ベクトルへの射影を計算する際にも平均値を引く必要があるが、

$$V \cdot \left(\Phi(t) - \frac{1}{M} \sum_{n=1}^M \Phi(x_n) \right) = \sum_{i=1}^M \alpha_i \Phi(x_i) \cdot \left(\Phi(t) - \frac{1}{M} \sum_{n=1}^M \Phi(x_n) \right)$$
$$= \sum_{i=1}^M \alpha_i \left(k(x_i, t) - \frac{1}{M} \sum_{n=1}^M K_{in} \right)$$

という形で計算をすることができる。

2次形式の計算方法1

- 共分散行列を以下のように表わす.

$$C = \frac{1}{M} \sum_{i=1}^M \Phi(\mathbf{x}_i) \Phi(\mathbf{x}_i)^T = \sum_{i=1}^M \lambda_i V_i V_i^T$$

- 逆行列は $C^{-1}C = I$ から,

$$C^{-1} = \sum_{i=1}^M \frac{1}{\lambda_i} V_i V_i^T$$

- $V = \sum_{i=1}^M \alpha_i \Phi(\mathbf{x}_i)$ を代入すると

$$C^{-1} = \sum_{k=1}^M \frac{1}{\lambda_k} \sum_{i,j=1}^M \alpha_i^k \alpha_j^k \Phi(\mathbf{x}_i) \Phi(\mathbf{x}_j)^T$$

2次形式の計算方法2

- 2次形式 $\Phi(\mathbf{t})^T C^{-1} \Phi(\mathbf{t})$ の計算を行うと,

$$\Phi(\mathbf{t})^T C^{-1} \Phi(\mathbf{t})$$

$$= \Phi(\mathbf{t})^T \left(\sum_{n=1}^M \frac{1}{\lambda_n} \sum_{i,j=1}^M \alpha_i^n \alpha_j^n \Phi(\mathbf{x}_i) \Phi(\mathbf{x}_j)^T \right) \Phi(\mathbf{t})$$

$$= \sum_{n=1}^M \frac{1}{\lambda_n} \sum_{i,j=1}^M \alpha_i^n \alpha_j^n \Phi(\mathbf{t})^T \Phi(\mathbf{x}_i) \Phi(\mathbf{x}_j)^T \Phi(\mathbf{t})$$

$$= \sum_{n=1}^M \frac{1}{\lambda_n} \sum_{i,j=1}^M \alpha_i^n \alpha_j^n k(\mathbf{t}, \mathbf{x}_i) k(\mathbf{x}_j, \mathbf{t})$$

となる.

課題

- Transpose Trickにおいて求めた元の空間での固有ベクトルに, 高次元のベクトルを射影する計算を示しなさい.