Spectral methods

Principal component analysis (PCA)

$\vec{x}^{(i)} \quad i = 1, \dots, N \qquad \vec{x}^{(i)} \in \mathbb{R}^p$

Find a description of the data set that captures as much information as possible in reduced dimensions.

Original dimension: p $\longrightarrow$ Reduced dimension $M < p$

$$\vec{x}^{(i)} = \sum_{j=1}^{p} a_j^{(i)} \vec{\phi}_j$$

Search for an orthonormal basis! $\{\vec{\phi}_j\} \qquad \vec{\phi}_j \in \mathbb{R}^p$

$$\vec{\phi}_j^T \cdot \vec{\phi}_k = \delta_{jk} \begin{cases} 1, & \text{if } j = k \\ 0, & \text{otherwise} \end{cases}$$

$$\Rightarrow a_j^{(i)} = \vec{\phi}_j^T \cdot \vec{x}^{(i)} = \vec{x}^{(i)T} \cdot \vec{\phi}_j$$

Minimisation of $\mathcal{L}$ under constraints:

$$\mathcal{L} = \sum_{j=M+1}^{P} \vec{\phi}_j^T C_x \vec{\phi}_j + \sum_{j=M+1}^{P} \lambda_j \left(1 - \vec{\phi}_j^T \cdot \vec{\phi}_j\right)$$

$$\nabla_{\phi_j} \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \vec{\phi}_j} = 2\, C_x \vec{\phi}_j - 2\, \lambda_j \vec{\phi}_j$$

$$\left.\frac{\partial \mathcal{L}}{\partial \vec{\phi}_j}\right|_{\vec{\phi}_j^*} = 0 \qquad \underline{C_x \vec{\phi}_j^* = \lambda_j \vec{\phi}_j^*}$$

Solution: $\{ b_j^*, \phi_j^* \}$

$\uparrow$ eigenvectors of C_x

$$MSE = \sum_{j=M+1}^{P} \vec{\phi}_j^{*T} \underline{C_x \cdot \vec{\phi}_j^*} = \sum_{j=M+1}^{P} \vec{\phi}_j^{*T} \lambda_j \vec{\phi}_j^*$$

$$= \sum_{j=M+1}^{P} \lambda_j$$

Sum of the eigenvalues that have been discarded.

$$MSE = \sum_{j=M+1}^{p} \frac{1}{N} \sum_{i=1}^{N} \Big(\overbrace{\vec{x}^{(i)T} \cdot \vec{\phi}_j}^{a_j^{(i)}} - \underbrace{\frac{1}{N} \sum_{i=1}^{N} \vec{x}_i^{(i)T} \cdot \vec{\phi}_j}_{b_j^{*}} \Big)^2 =$$

$$= \sum_{j=M+1}^{p} \frac{1}{N} \sum_{i=1}^{N} \left[\left(\vec{x}^{(i)T} - \frac{1}{N} \sum_{i=1}^{N} \vec{x}^{T(i)} \right) \cdot \vec{\phi}_j \right]^2$$

$$= \sum_{j=M+1}^{p} \frac{1}{N} \left[\underset{1\times p}{\vec{\phi}_j^T} \cdot \underbrace{\underset{p\times 1}{\left(\vec{x}^{(i)} - \frac{1}{N} \sum_{i=1}^{N} \vec{x}^{(i)} \right)} \cdot \underset{1\times p}{\left(\vec{x}^{(i)} - \frac{1}{N} \sum_{i=1}^{N} \vec{x}^{(i)} \right)^T}}_{} \underset{p\times 1}{\vec{\phi}_j} \right]$$

$$\parallel$$

$$\mathbb{E}_i \left[\left(\vec{x} - \mathbb{E}_i(\vec{x}) \right) \cdot \left(\vec{x} - \mathbb{E}_i(\vec{x}) \right)^T \right]$$

$$= \phi_j$$

$$\parallel$$

covariance matrix

$$(C_x)_{p\times p}$$

$$MSE = \sum_{j=M+1}^{p} \underset{(1\times p)}{\vec{\phi}_j^T} \underset{(p\times p)}{\boxed{C_x}} \underset{(p\times 1)}{\vec{\phi}_j}$$

Data

Approximate $\vec{x}^{(i)}$ by:

$$\vec{x}_M^{(i)} = \sum_{j=1}^{M} a_j^{(i)} \vec{\phi}_j + \sum_{j=M+1}^{P} \underline{\underline{b_j}} \vec{\phi}_j$$

Find $\{b_j, \vec{\phi}_j\}$ that apply to all samples.

For each sample:

$$\Delta \vec{x}^{(i)} = \vec{x}^{(i)} - \vec{x}_M^{(i)} = \sum_{j=M+1}^{P} \left[a_j^{(i)} - b_j \right] \vec{\phi}_j$$

$i = 1, \dots, N$

Error on the dataset:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} \| \Delta \vec{x}^{(i)} \|^2 = \frac{1}{N} \sum_{i=1}^{N} \left[\sum_{j,k=M+1}^{P} (a_j^{(i)} - b_j)(a_k^{(i)} - b_k) \vec{\phi}_j^T \cdot \vec{\phi}_k \right]$$

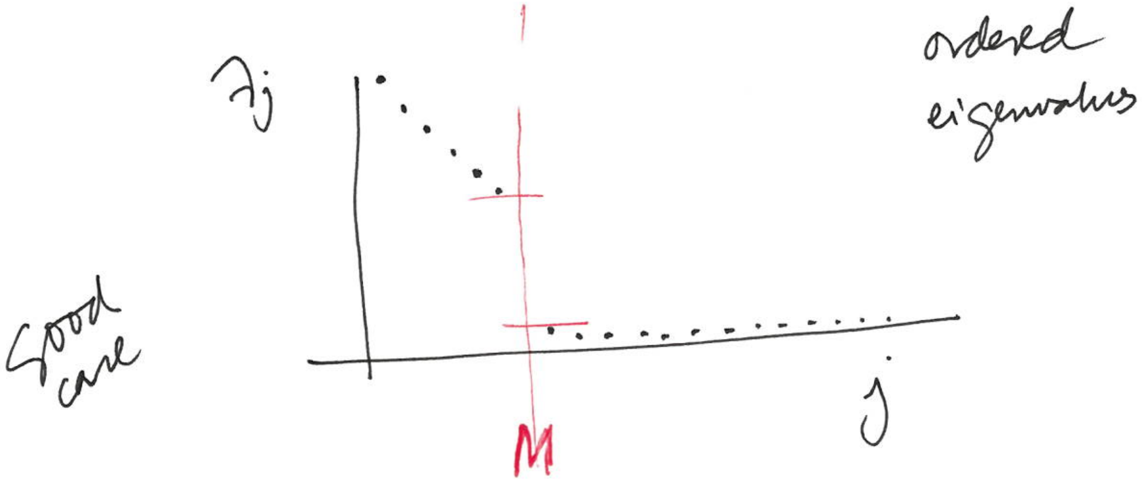
$$= \frac{1}{N} \sum_{i=1}^{N} \sum_{j=M+1}^{P} (a_j^{(i)} - b_j)^2$$ (by orthogonality)

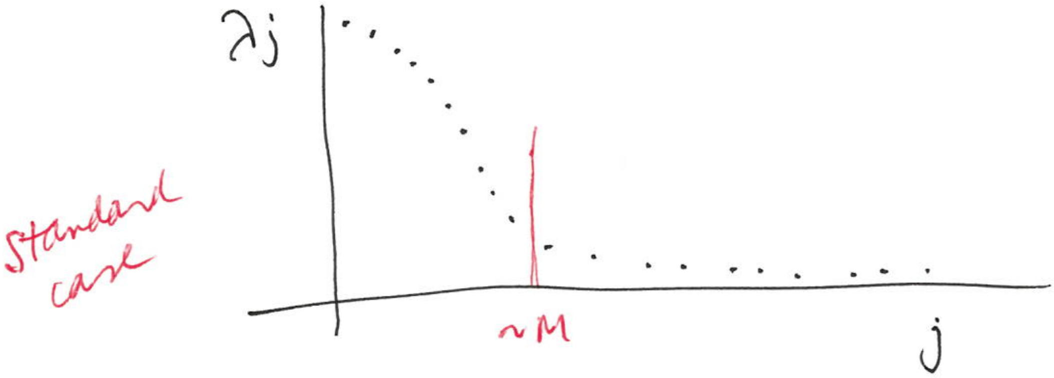
Minimise MSE:

(1) $$\frac{\partial MSE}{\partial b_j} = \frac{1}{N} \sum_{i=1}^{N} (-2)(a_j^{(i)} - b_j)$$

$$\left. \frac{\partial MSE}{\partial b_j} \right|_{b_j^*} = 0 \qquad \underline{b_j^* = \frac{1}{N} \sum_{i=1}^{N} a_j^{(i)}}$$

Obtain $C_{\mathbf{x}}$ from data: Find eigendecomposi





Tolerance : ε

$$M(\varepsilon) \Big/ \sum_{j=M+1}^{P} \lambda_i < \varepsilon$$

Principal components are the eigenvectors of C_x

$$\hat{\vec{x}}^{(i)} = \sum_{j=1}^{M} \left(\vec{x}^{(i)T} \cdot \vec{\phi}_j \right) \vec{\phi}_j$$

and the $\vec{\phi}_j$ are the eigenvectors of the covariance matrix

Mathematical basis for PCA:

Singular value decomposition (SVD).

Eckart & Young 39.

If B has rank M

then $\| A - B \| \geqslant \| A - A_M \|$

$$A_M = \sum_{j=1}^{M} \sigma_i \vec{u}_i \vec{v}_i^T$$

where $\vec{u}_i$ are the right eigenvectors of A

$\vec{v}_i$ are the left eigenvectors of A

σ_i are the singular values of A

These definitions follow from the SVD

$\boxed{\text{SVD}}$:

$A_{n \times n}$

Analogue to Diagonalization of square matrices

$$\left[\begin{array}{l} AV = V\Lambda \\ V = (\vec{v}_1 \cdots \vec{v}_n) \\ \Lambda = \mathrm{diag}(\lambda i) \\ A = V \Lambda V^{-1} \end{array}\right.$$

but for rectangular matrices.

Analogue for rectangular matrices:

Given $A_{m\times n} \in \mathbb{R}^{m\times n}$

$$\underset{m\times n}{A}\, V_{n\times n} = \underset{m\times m}{U}\, \Sigma_{m\times n}$$

$$\underset{n\times n}{V} = \left(\vec{v}_1 \;\cdots\; \vec{v}_n \right)$$

$$\underset{m\times m}{U} = \left(\vec{u}_1 \;\cdots\; \vec{u}_m \right)$$

$$\underset{m\times n}{\Sigma} = \left(\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right)$$

$$\min(m, n) \geq r$$