Signal and Image Processing (236327)

Tutorial 13 – Part A

Principle Component Analysis (PCA)
Principle Component Analysis (PCA)

Consider a class of signals that is described by a random vector \( \overline{\varphi}_\omega \) with zero mean \( E_\omega\{\overline{\varphi}_\omega\} = \overline{0} \) and an autocorrelation matrix \( R_{\overline{\varphi}} = E_\omega\{\overline{\varphi}_\omega \overline{\varphi}_\omega^*\} \).

The diagonalization of the autocorrelation matrix \( R_{\overline{\varphi}} \) is obtained using the unitary matrix \( U \) via

\[
R_{\overline{\varphi}} = U \Lambda U^*
\]

where \( \Lambda \) is a diagonal matrix, composed of \( R_{\overline{\varphi}} \)'s eigenvalues:

\[
\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N \geq 0
\]
Principle Component Analysis (PCA)

We consider the matrix $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N]$ by its columns, and the direct transform to be applied via

$$\tilde{\Phi}_\omega^{(\mathbf{U})} = \mathbf{U}^* \Phi_\omega$$

where the $k^{th}$ component of $\tilde{\Phi}_\omega^{(\mathbf{U})}$ is the scalar $\langle \tilde{\Phi}_\omega, \bar{u}_k \rangle = \bar{u}_k^* \Phi_\omega$.

The corresponding $K$-term approximation, for any $K \in \{1, \ldots, N\}$, is

$$\tilde{\Phi}_\omega^{(K)} = \sum_{k=1}^{K} \langle \tilde{\Phi}_\omega, \bar{u}_k \rangle \bar{u}_k$$
Principle Component Analysis (PCA)

The above matrix $\mathbf{U}$ is the PCA transformation, providing the optimal $K$-term approximation in the sense of minimal expected squared-error

$$E\{\mathcal{E}^2_\omega(K)\} = E\left\{\left(\overline{\phi}_\omega - \hat{\phi}^{(K)}_\omega\right)^* \left(\overline{\phi}_\omega - \hat{\phi}^{(K)}_\omega\right)\right\}$$

for any $K \in \{1, \ldots, N\}$.

Specifically, the **minimal** expected error is

$$E\{\mathcal{E}^2_\omega(K)\} = \sum_{k=K+1}^{N} \lambda_k$$
PCA: An Example for A Simple Class

A class of one-dimensional discrete signals is defined as follows.

A signal is the column vector of $N$ samples of the form:

$$\overline{\varphi} = [M, \ldots, M, M+L, M, \ldots, M]^T$$

i.e., all the vector components have the value $M$ except for the $K^{th}$ component that has the value $M + L$.

The vector components are indexed starting at 1, i.e., the vector can be generally formulated as $\overline{\varphi} = [\varphi_1, \ldots, \varphi_N]^T$
PCA: An Example for A Simple Class

A class of one-dimensional discrete signals is defined as follows.

A signal is the column vector of \( N \) samples of the form:

\[
\overline{\varphi} = [M, \ldots, M, M + L, M, \ldots, M]^T
\]

\( K \)th component

\( M, L, \) and \( K \) are **independent random variables**:

\( K \) is a uniform random variable over the integers \( \{1, \ldots, N\} \).

\( M \) obeys \( E\{M\} = 0 \) and \( E\{M^2\} = c \)

\( L \) obeys \( E\{L\} = 0 \) and \( E\{L^2\} = N(1 - c) \)

where \( 0 < c < 1 \) is a (deterministic) constant.
PCA: An Example for A Simple Class

a. Show that the random vector $\bar{\phi}$ has a zero mean.

**Solution:**

The expected value of the $i^{th}$ component is

$$E\{\phi_i\} = E\{\phi_i|K = i\}P(K = i) + E\{\phi_i|K \neq i\}P(K \neq i)$$

$$= E\{M + L\}P(K = i) + E\{M\}P(K \neq i) = 0$$

Hence, the mean signal (vector) is $E\{\bar{\phi}\} = 0$.
PCA: An Example for A Simple Class

b. Calculate the autocorrelation matrix of $\bar{\varphi}$, denoted as $R_{\bar{\varphi}}$, and show it is circulant.

**Solution:**

The variance of the $i^{th}$ component is the $(i, i)$ entry in $R_{\bar{\varphi}}$:

$$r_{ii} = E\{\varphi_i^2\} = E\{\varphi_i^2 | K = i\}P(K = i) + E\{\varphi_i^2 | K \neq i\}P(K \neq i)$$

$$= (E\{L^2\} + E\{M^2\})P(K = i) + E\{M^2\}P(K \neq i)$$

$$= E\{L^2\}P(K = i) + E\{M^2\} \cdot (P(K = i) + P(K \neq i)) =$$

$$= N(1 - c) \frac{1}{N} + c = 1$$
b. Calculate the autocorrelation matrix of $\varphi$, denoted as $R_{\varphi}$, and show it is circulant.

**Solution:**
The correlation between the $i^{th}$ and the $j^{th}$ components of $\varphi$, which is the $(i, j)$ entry in $R_{\varphi}$:

$$r_{ij} = E\{\varphi_i \varphi_j\}$$

$$= E\{\varphi_i \varphi_j | K = i\}P(K = i) + E\{\varphi_i \varphi_j | K = j\}P(K = j) + E\{\varphi_i \varphi_j | K \neq i, j\}P(K \neq i, j)$$

$$= E\{(M + L)M | K = i\}P(K = i) + E\{M(M + L) | K = j\}P(K = j) + E\{M^2 | K \neq i, j\}P(K \neq i, j)$$

$$= E\{M^2\}P(K = i) + E\{M^2\}P(K = j) + E\{M^2\}P(K \neq i, j)$$

$$= E\{M^2\}$$

$$= c$$

where above we used the independence of $M$, $L$, and $K$ in the calculation of

$$E\{(M + L)M | K = i\} = E\{(M + L)M\} = E\{M^2 + LM\} = E\{M^2\} + E\{L\}E\{M\} = E\{M^2\}$$
b. Calculate the autocorrelation matrix of $\overline{\varphi}$, denoted as $R_{\overline{\varphi}}$.

**Solution:**

To conclude the structure of $R_{\overline{\varphi}}$ is

$$R_{\overline{\varphi}} = \begin{bmatrix}
1 & c & c & \cdots & c \\
c & 1 & c & \ddots & \vdots \\
c & c & \ddots & \ddots & c \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
c & \cdots & c & c & 1
\end{bmatrix}$$

exhibiting a **circulant** form.
PCA: An Example for A Simple Class

c. Is the random signal is cyclo-stationary?

**Solution:**

Since the autocorrelation matrix is circulant, all the \((i, j)\) entries that obey \(j - i \pmod{N} = a\) have the same value (for \(a = 0, \ldots, N - 1\)).

This means that \(E\{\phi_i \phi_{i+a(\text{mod } N)}\} = E\{\phi_l \phi_{l+a(\text{mod } N)}\}\) for any \(i, l = 1, \ldots, N\). Then, by definition, the cyclo-stationarity property is satisfied.
d. What is the PCA matrix corresponding to the autocorrelation matrix $\mathbf{R}_\phi$?

**Solution:**

Since the autocorrelation matrix is circulant and of size $N \times N$, it is diagonalized by the $N \times N$ $[DFT]^*$ matrix.

Hence, the PCA transform matrix is the $[DFT]^*$ matrix!
e. Explain how the eigenvalues of $\mathbf{R}_\varphi$ can be computed in a way that is simpler than an explicit eigendecomposition procedure.

**Solution:**

Since $\mathbf{R}_\varphi$ is a circulant matrix, it is diagonalized by the $[DFT]^*$ matrix, and the eigenvalues can be obtained by applying the the $[DFT]^*$ matrix on the column-ordering of the first row of $\mathbf{R}_\varphi$. 

236327, CS Department, Technion, Spring 2018
Principle Component Analysis (PCA) for A Dataset

We considered the PCA for a class with a probabilistic definition.

The PCA approach can be employed also for a given set of data vectors \( \{\overline{\phi}_1, \overline{\phi}_2, \ldots, \overline{\phi}_M\} \) that, for a large \( M \), are assumed to empirically represent the class.

Then, one can **empirically estimate** the class mean via

\[
\hat{\mu}_\overline{\phi} = \frac{1}{M} \sum_{m=1}^{M} \overline{\phi}_m
\]

and the corresponding autocorrelation matrix

\[
\hat{R}_\overline{\phi} = \frac{1}{M} \sum_{m=1}^{M} (\overline{\phi}_m - \hat{\mu}_\overline{\phi})(\overline{\phi}_m - \hat{\mu}_\overline{\phi})^*
\]

and the PCA can be computed based on \( \hat{R}_\overline{\phi} \).