1 Principal Component Analysis

In the previous lecture we have dealt with the problems of signal denoising and deconvolution and proposed to solve them using the Wiener filter which is an optimal linear filter for denoising and an optimal LSI filter for deconvolution given the signal second order statistics. In this lecture we will focus on the problem of lossy compression.

1.1 Compression using \( k \)-term Approximation with Orthonormal Bases

In the lectures about sampling we have seen that we can represent any signal \( \varphi \in \mathbb{R}^N \) using an orthonormal basis \( \{ \psi_i \}_{i=1}^N \) with \( N \) coefficients as

\[
\varphi = \sum_{i=1}^{N} \varphi_i^\psi \psi_i.
\]

(1)

For compressing the signal one may suggest that we should use only \( k < N \) non-zero coefficients getting a compressed version of the signal

\[
\varphi^k = \sum_{i=1}^{k} \varphi_i^\psi \psi_i.
\]

(2)

Besides the fact that the \( k \)-largest coefficients are those that should be selected, it would be hard to evaluate whether such a compression is good or not and to select the "correct" basis for it without a prior knowledge on the signal. In this lecture we assume that the signal is a random process with known second order statistics. Given an autocorrelation matrix of a given signal or set of signals we would like to find the orthogonal basis that represents this signal with \( k \) coefficients with minimal error.

1.2 The \( k \)-term Approximation Error

Let \( R_\varphi \) be the autocorrelation matrix of a signal \( \varphi \). The MSE of the \( k \)-term approximation in (2) is

\[
E \left\| \varphi - \varphi^k \right\|_2^2 = E \left( \varphi - \varphi^k \right)^* \left( \varphi - \varphi^k \right).
\]

(3)

We want to optimize the basis used in (2) such that the compression error will decay fast as \( k \) increases. Opening the brackets in (2) leads to

\[
E \left\| \varphi - \varphi^k \right\|_2^2 = E \left[ \varphi^* \varphi \right] - E \left[ \varphi^* \varphi^k \right] - E \left[ \left( \varphi^k \right)^* \varphi \right] + E \left[ \left( \varphi^k \right)^* \varphi^k \right].
\]

(4)
First notice that $E[\varphi^* \varphi] = \sum_{i=1}^N E \left[ (\varphi_i^\psi)^2 \right] = \text{trace}(R_\varphi)$. Then note that $E[\varphi^* \varphi^k] = E[(\varphi^k)^* \varphi]$. By using (1) and (2) we have

$$E[\varphi^* \varphi^k] = E \left[ \left( \sum_{i=1}^N \varphi_i^\psi \psi_i \right)^* \sum_{i=1}^k \varphi_i^\psi \psi_i \right] = E \left[ \sum_{i=1}^k \sum_{i=1}^N \varphi_i^\psi \varphi_i^\psi^* \psi_i \psi_i \right] = \sum_{i=1}^k E \left[ (\varphi_i^\psi)^2 \right],$$

where in the last step we have used the linearity of the expectation and the orthogonality of the basis vectors $(\varphi_i^\psi, \psi_i) = \delta_{i,i}$. Notice that due to the orthogonality of the basis vectors we get in a very similar way that also

$$E[(\varphi^k)^* \varphi^k] = \sum_{i=1}^k E \left[ (\varphi_i^\psi)^2 \right] = E[\varphi^* \varphi^k].$$

Therefore we have

$$E \left\| \varphi - \varphi^k \right\|_2^2 = \text{trace}(R_\varphi) - \sum_{i=1}^k E \left[ (\varphi_i^\psi)^2 \right].$$

Since $\varphi_i^\psi$ is an optimal representation coefficient we have $\varphi_i^\psi = \langle \varphi, \psi_i \rangle = \psi_i^* \varphi = \varphi^* \psi_i$. Thus

$$\sum_{i=1}^k E \left[ (\varphi_i^\psi)^2 \right] = \sum_{i=1}^k E \left[ \psi_i^* \varphi \varphi^* \psi_i \right] = \sum_{i=1}^k \psi_i^* E[\varphi \varphi^*] \psi_i = \sum_{i=1}^k \psi_i^* R_\varphi \psi_i.$$

Plugging (8) into (7) leads to

$$E \left\| \varphi - \varphi^k \right\|_2^2 = \text{trace}(R_\varphi) - \sum_{i=1}^k \psi_i^* R_\varphi \psi_i.$$

### 1.3 The Principal Component Analysis

Having closed form of the error in terms of the $k$ orthonormal basis vectors used for the compression we may ask ourselves what is the best selection of $\{\psi_1, \psi_2, \ldots, \psi_k\}$ that minimizes (9). Notice that the first term in (9) is independent of the basis vectors. Therefore, finding the basis vectors that minimize (9) is equivalent to solving the maximization problem:

$$\max_{\psi_1, \psi_2, \ldots, \psi_k} \sum_{i=1}^k \psi_i^* R_\varphi \psi_i.$$

Recall that it is possible to write $R_\varphi = U \Lambda U^*$, where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$ is a diagonal matrix containing the eigenvalues of $R_\varphi$ and $U = [u_1, u_2, \ldots, u_N]$ is a unitary matrix containing the eigenvectors of $R_\varphi$. Assume w.l.o.g. (with out loss of generality) that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$. Considering first the case of $k = 1$ we have

$$\psi_1^* R_\varphi \psi_1 = \psi_1^* U \Lambda U^* \psi_1 = \sum_{i=1}^N \lambda_i \langle u_i, \psi_1 \rangle^2.$$

It is clear that the basis vector $\psi_1$ that maximizes the above term is $\psi_1 = u_1$.

If we continue with $k = 2, 3, \ldots$ we will get that the optimal selection for the basis vectors is $\psi_i = u_i$. This technique for selecting the optimal basis is called principal component analysis (PCA). The selected basis is also known as the discrete karhunen-Loève transform (KLT) or the Hotelling transform.
To summarize, given a signal $\varphi$ with a known autocorrelation matrix $R_\varphi$, its optimal $k$-term approximation with an orthonormal basis is done by taking the first $k$ columns in $U$, $U_{[1,k]}$. The optimal coefficients are simply $U^*_{[1,k]} \varphi$.

Since $\text{trace} R_{\varphi} = \text{trace} A$ it is easy to show that the optimal error is

$$E \left\| \varphi - \varphi^k \right\|_2^2 = \sum_{i=k+1}^N \lambda_i. \tag{12}$$

Thus, the optimal error decay with respect to $k$ depends on the decay rate of the eigenvalues of $R_{\varphi}$. In the worst case, all eigenvalues will be equal to each other and we will have a linear decay in the error. Our hope is that the decay will be faster. Note that the more structure the signal have the more it is likely to happen since structure create dominant directions in the space.

### 1.4 Optimal Basis for Natural Signals and Images

In many cases we do not have the autocorrelation matrix for a given signal. One option is to take many signals similar to this signal and learn their statistics. Another option is to use an autocorrelation matrix that characterize properties known for such signals.

Such an example we have with natural signals and images. An autocorrelation matrix commonly used with these signals is

$$R_\rho = \begin{bmatrix} 1 & \rho & \rho^2 & \ldots & \rho^{N-1} \\ \rho & 1 & \rho & \ldots & \rho^{N-2} \\ \rho^2 & \rho & 1 & \ldots & \rho^{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{N-1} & \rho^{N-2} & \rho^{N-3} & \ldots & 1 \end{bmatrix}, \tag{13}$$

with $\rho$ close to 1. This matrix characterizes the fact that in natural signals the correlation between close elements is high and it decays fast as the distance between the elements increases.

As we will see hereafter, the KLT of this matrix is approximately the rows of the discrete cosine transform (DCT). Notice that the above autocorrelation matrix characterizes correlations only in one dimension. Therefore, for images we will assume that each column and row separately obey this statistics. This result with having the KLT for images the two dimensional DCT (2D-DCT) which is a generalization of the DCT to two dimensions in the same way that 2D-DFT generalizes the DFT to 2D. Applying the 2D-DCT on an image is equivalent to applying the DCT twice: Once vertically on the columns and then horizontally on the rows of the outcome.

The DCT $D \in \mathbb{R}^{N \times N}$ is defined as

$$D_{k,n} = \begin{cases} \frac{1}{\sqrt{N}} & k = 0 \\ \frac{2}{\sqrt{N}} \cos \left( \frac{\pi(2n+1)k}{2N} \right) & 1 \leq k \leq N - 1 \end{cases}. \tag{14}$$

It has the following properties

1. The DCT is a real unitary matrix. This implies that

$$D^{-1} = D^T = D^* \Rightarrow D^T D = I. \tag{15}$$

2. The discrete cosine basis vectors are the rows of $D$.

3. The DCT is a fast transform. There exists an algorithm that can apply it on a vector with $O(N \log(N))$ operations.
4. The basis vectors of the DCT (its rows) are eigenvectors of the following symmetric tridiagonal matrix

\[ Q_c = \begin{bmatrix}
1 - \alpha & -\alpha & 0 & \cdots & 0 \\
-\alpha & 1 & -\alpha & \ddots & \vdots \\
0 & \ddots & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & 1 -\alpha \\
0 & \cdots & 0 & -\alpha & 1 -\alpha \\
\end{bmatrix} \]  

(16)

5. The reason that the discrete cosine basis vectors are approximately the KLT of \( R_\Omega \) is that

\[ \beta^2 R_p^{-1} = \begin{bmatrix}
1 - \rho \alpha & -\alpha & 0 & \cdots & 0 \\
-\alpha & 1 & -\alpha & \ddots & \vdots \\
0 & \ddots & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & 1 -\alpha \\
0 & \cdots & 0 & -\alpha & 1 -\rho \alpha \\
\end{bmatrix} \]  

(17)

for \( \beta^2 = (1 - \rho^2)(1 + \rho^2) \) and \( \alpha = \frac{\rho}{1+\rho} \). Due to the facts that \( \beta^2 R_p^{-1} \approx Q_c \) for \( \rho \) close to 1 and that \( R_p^{-1} \) and \( R_\Omega \) has the same eigenvectors we have that the rows in DCT are the eigenvectors of \( R_\Omega \) implying that it is the KLT of natural signals and images.

The facts that DCT is a fast transform and that it is the KLT of \( R_\rho \) have made it very popular in signal and image processing. For example, the JPEG image compression uses the DCT.