

Homework #5 solution key

Prayag

Neural networks and learning systems-I

April 23, 2019

Problem 8.5.

Solution. For the matched filter considered in Example 2, it is given that the random vector $\mathbf{X} = \mathbf{s} + \mathbf{V}$, where the signal component is represented by \mathbf{s} has a fixed Euclidean norm of one, i.e., $\sqrt{\mathbf{s}^T \mathbf{s}} = 1 \implies \mathbf{s}^T \mathbf{s} = 1$.

The random vector \mathbf{V} , represents the additive noise component and has zero mean and covariance matrix $\sigma^2 \mathbf{I}$. The correlation matrix of \mathbf{X} is given by $\mathbf{R} = \mathbf{s} \mathbf{s}^T + \sigma^2 \mathbf{I}$.

The largest eigenvalue of the correlation matrix \mathbf{R} and the corresponding eigenvector are $\lambda_1 = 1 + \sigma^2$ and $\mathbf{q}_1 = \mathbf{s}$ respectively. Let us post multiply the correlation matrix by the eigenvector \mathbf{q}_1 .

$$\begin{aligned} \mathbf{R} \mathbf{q}_1 &= (\mathbf{s} \mathbf{s}^T + \sigma^2 \mathbf{I}) \mathbf{q}_1 \\ &= (\mathbf{s} \mathbf{s}^T + \sigma^2 \mathbf{I}) \mathbf{s} \\ &= \mathbf{s} \mathbf{s}^T \mathbf{s} + \sigma^2 \mathbf{s} \\ &= \mathbf{s} + \sigma^2 \mathbf{s} \quad (\text{since } \mathbf{s}^T \mathbf{s} = 1) \\ &= (1 + \sigma^2) \mathbf{s} \\ &= \lambda_1 \mathbf{s} \\ &= \lambda_1 \mathbf{q}_1 \end{aligned}$$

Therefore the given parameters satisfy the basic relation of $\mathbf{R} \mathbf{q}_1 = \lambda_1 \mathbf{q}_1$. ■

Problem 8.15.

Solution. Let the total number of inputs be N . The data \mathbf{x}_i is projected using a kernel to get the projected data set $\phi(\mathbf{x}_i)$. Let $\tilde{\phi}(\mathbf{x}_i)$ be the projected data points after centralizing the data as given below

$$\tilde{\phi}(\mathbf{x}_i) = \phi(\mathbf{x}_i) - \frac{1}{N} \sum_{l=1}^N \phi(\mathbf{x}_l).$$

We can get the corresponding elements of the gram matrix $\tilde{\mathbf{K}}$ as below:

$$\begin{aligned}
\tilde{k}_{ij} &= \tilde{\phi}^T(\mathbf{x}_i)\tilde{\phi}(\mathbf{x}_j) \\
&= \left(\phi(\mathbf{x}_i) - \frac{1}{N} \sum_{m=1}^N \phi(\mathbf{x}_m) \right)^T \left(\phi(\mathbf{x}_j) - \frac{1}{N} \sum_{n=1}^N \phi(\mathbf{x}_n) \right) \\
&= \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) - \frac{1}{N} \sum_{m=1}^N \phi^T(\mathbf{x}_m) \phi(\mathbf{x}_j) - \frac{1}{N} \sum_{n=1}^N \phi^T(\mathbf{x}_i) \phi(\mathbf{x}_n) + \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N \phi^T(\mathbf{x}_m) \phi(\mathbf{x}_n) \\
&= k_{ij} - \frac{1}{N} \sum_{m=1}^N k_{mj} - \frac{1}{N} \sum_{n=1}^N k_{in} + \frac{1}{N^2} \sum_{m=1}^N \sum_{n=1}^N k_{mn}
\end{aligned}$$

A compact representation of the above in the matrix form is given by

$$\tilde{\mathbf{K}} = \mathbf{K} - \mathbf{N}\mathbf{K} - \mathbf{K}\mathbf{N} + \mathbf{N}\mathbf{K}\mathbf{N}$$

where \mathbf{N} is an $N \times N$ matrix with all entries as $\frac{1}{N}$. ■

Problem 8.16.

Solution. It is given that the eigenvector \tilde{q} of the correlation matrix $\tilde{\mathbf{R}}$ is normalized to unit length, that is,

$$\tilde{q}_k^T \tilde{q}_k = 1 \quad \text{for } k = 1, 2, \dots, l \tag{1}$$

where it is assumed that the eigenvalues of \mathbf{K} are arranged in descending order with λ_l being the smallest nonzero eigenvalue of the Gram matrix \mathbf{K} . We know that

$$\tilde{q} = \sum_{j=1}^N \alpha_j \phi(x_j) \tag{2}$$

$$\mathbf{K}\alpha = \lambda\alpha \tag{3}$$

Using (2) in (1) and upon simplifying, we get

$$\begin{aligned}
1 &= \tilde{q}_k^T \tilde{q}_k \\
&= \sum_{i=1}^N \sum_{j=1}^N \alpha_{ki}^T \alpha_{kj} \phi(x_i)^T \phi(x_j) \\
&= \alpha_k^T \mathbf{K} \alpha_k \\
&= \alpha_k^T \lambda_k \alpha_k \quad (\text{Using (3)}) \\
&= \lambda_k \alpha_k^T \alpha_k \\
\frac{1}{\lambda_k} &= \alpha_k^T \alpha_k \quad \text{for } k = 1, 2, \dots, l.
\end{aligned}$$

Therefore, we see that normalization of the eigenvector α of the Gram matrix \mathbf{K} is equivalent to the requirement of Eq(8.109) to be satisfied. ■