Dissecting Low-Rank Correlation Matrix

Navid Shokouhi

In most existing studies on model order selection, the fundamental assumption is that a low rank covariance matrix is buried in white Gaussian noise. Such a covariance is obtained from the following linear model for the observed data:

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{x}$ is an n-dimensional observation vector, $\mathbf{A}$ is an $n \times p$ linear mixing matrix, $\mathbf{s}$ is the p-dimensional signal (to be modeled), and $\mathbf{n}$ is white Gaussian noise, $\mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_n)$. The assumption used in calculating the log-likelihood function of the observation given a model order (more precisely, given the model) is that the full-rank covariance matrix $\mathbf{R}_x$ of the observations can be modeled as:

$$\mathbf{R}_x = \mathbf{U}_p(\mathbf{\Lambda} - \sigma^2\mathbf{I}_p)\mathbf{U}_p^T + \sigma^2\mathbf{I}_n. \tag{2}$$

where $\mathbf{U}_p$ contains the top p eigenvectors of $\mathbf{R}_x$, and $\mathbf{\Lambda}$ contains its top p eigenvalues. The objective of this document is to derive Eq. (2) starting from the assumption that the eigenvalue decomposition of $\mathbf{R}_x$ comprises of p large values corresponding to the signal and $n - p$ small values corresponding to the noise.

$$\mathbf{R}_x = \mathbf{U}\begin{pmatrix}\mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \sigma^2\mathbf{I}_{n-p}\end{pmatrix}\mathbf{U}^T, \tag{3}$$

where $\mathbf{\Lambda}$ is a diagonal $p \times p$ matrix containing the signal eigenvalues. $\mathbf{U}$ contains all of the eigenvectors of $\mathbf{R}_x$. We can split $\mathbf{U}$ by its columns, where the first p correspond to $\mathbf{\Lambda}$.

$$\mathbf{R}_x = \begin{pmatrix}\mathbf{U}_p \mathbf{U}_{n-p}\end{pmatrix}\begin{pmatrix}\mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \sigma^2\mathbf{I}_{n-p}\end{pmatrix}\begin{pmatrix}\mathbf{U}_p^T \\ \mathbf{U}_{n-p}^T\end{pmatrix}.$$

The matrices $\mathbf{U}_p$ and $\mathbf{U}_{n-p}$ are $n \times p$ and $n \times (n-p)$, respectively. Calculating the matrix multiplications:

$$\begin{aligned}\mathbf{R}_x &= \begin{pmatrix}\mathbf{U}_p\mathbf{\Lambda} & \sigma^2\mathbf{U}_{n-p}\end{pmatrix}\begin{pmatrix}\mathbf{U}_p^T \\ \mathbf{U}_{n-p}^T\end{pmatrix} \\ &= \mathbf{U}_p\mathbf{\Lambda}\mathbf{U}_p^T + \sigma^2\mathbf{U}_{n-p}\mathbf{U}_{n-p}^T \end{aligned} \tag{4}$$

We know that the eigenvector matrix $\mathbf{U}$ is orthonormal, $\mathbf{U}\mathbf{U}^T = \mathbf{I}_n$, therefore:

$$\mathbf{I}_n = \mathbf{U}_n\mathbf{U}_p^T + \mathbf{U}_{n-p}\mathbf{U}_{n-p}^T. \tag{5}$$

Replacing $\mathbf{U}_{n-p}\mathbf{U}_{n-p}^T$ with $\mathbf{I}_n - \mathbf{U}_p\mathbf{U}_p^T$:

$$\mathbf{R}_x = \mathbf{U}_p\mathbf{\Lambda}\mathbf{U}_p^T + \sigma^2\mathbf{I}_n - \sigma^2\mathbf{U}_p\mathbf{U}_p^T \tag{6}$$

$$= \mathbf{U}_p(\mathbf{\Lambda} - \sigma^2\mathbf{I}_p)\mathbf{U}_p^T + \sigma^2\mathbf{I}_n. \tag{7}$$

References