My Viewpoints on Subspace Theory

Huiping Huang (Homepage: https://huipinghuang.work)

November 2018

Consider a data model as $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)$, where t denotes the time index, $\mathbf{A}$ is the steering matrix, and $\mathbf{s}(t)$ and $\mathbf{n}(t)$ stand for signal and noise components, respectively. $\mathbf{x}(t) \in \mathbb{C}^{M\times 1}$, $\mathbf{s}(t) \in \mathbb{C}^{K\times 1}$, $\mathbf{n}(t) \in \mathbb{C}^{M\times 1}$, $\mathbf{A} \in \mathbb{C}^{M\times K}$, and $M > K$. The covariance matrices are defined as follows

$$\begin{aligned} \mathbf{R}_x &\triangleq E\{\mathbf{x}(t)\mathbf{x}^H(t)\} \\ \mathbf{R}_s &\triangleq E\{\mathbf{s}(t)\mathbf{s}^H(t)\} \\ \mathbf{R}_n &\triangleq E\{\mathbf{n}(t)\mathbf{n}^H(t)\} \end{aligned} \tag{1}$$

where $E\{\cdot\}$ and $(\cdot)^H$ are mathematical expectation and Hermitian transport operators, respectively. It is easy to find that

$$\mathbf{R}_x = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \mathbf{R}_n. \tag{2}$$

Defining $\bar{\mathbf{R}}_s \triangleq \mathbf{A}\mathbf{R}_s\mathbf{A}^H$, (2) can be rewritten as

$$\mathbf{R}_x = \bar{\mathbf{R}}_s + \mathbf{R}_n. \tag{3}$$

According to the definition of $\bar{\mathbf{R}}_s$ and $M > K$, it can be found that $\bar{\mathbf{R}}_s$ is a rank-reduction matrix with rank $Rank(\bar{\mathbf{R}}_s) = K$, where $Rank(\cdot)$ is the rank of a matrix. Therefore, the eigenvalue decomposition of $\bar{\mathbf{R}}_s$ can be depicted as follows

$$\bar{\mathbf{R}}_s = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H \tag{4}$$

where $\mathbf{U}$ and $\mathbf{\Sigma}$ contain all the eigenvectors and eigenvalues of $\bar{\mathbf{R}}_s$, respectively, and $\mathbf{\Sigma}$ is a diagonal matrix which can be formulated as $\mathbf{\Sigma} = diag\{\sigma_1, \sigma_2, \cdots, \sigma_K, 0, \cdots, 0\}$, where $diag\{\cdot\}$ denotes a diagonal matrix.

As for $\mathbf{R}_n$, if the noise component is independent and identically distributed Gaussian process, it satisfies $\mathbf{R}_n = \sigma_0\mathbf{I}$, where σ_0 is the noise power and $\mathbf{I}$ stands for the identity matrix. Because the matrix $\mathbf{U}$ in (4) is a unitary matrix, $\mathbf{R}_n$ can be further reformulated as follows

$$\mathbf{R}_n = \mathbf{U}\mathbf{\Sigma}_0\mathbf{U}^H \tag{5}$$

where $\mathbf{\Sigma}_0 = diag\{\sigma_0, \sigma_0, \cdots, \sigma_0\}$.

Substituting both (4) and (5) back into (3), one can obtain

$$\mathbf{R}_x = \mathbf{U}\mathbf{\Sigma}_x\mathbf{U}^H \tag{6}$$

where $\mathbf{\Sigma}_x = \mathbf{\Sigma} + \mathbf{\Sigma}_0 = diag\{\sigma_1 + \sigma_0, \sigma_2 + \sigma_0, \cdots, \sigma_K + \sigma_0, \sigma_0, \cdots, \sigma_0\}$. The equation (6) can be viewed as the eigenvalue decomposition of $\mathbf{R}_x$.

Comparing (4) and (6), it can be seen that $\bar{\mathbf{R}}_s$ and $\mathbf{R}_x$ share the same eigenvectors (i.e., $\mathbf{U}$). (**Conclusion 1**)

The signal subspace, I think, should be defined as $Span\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_K\}$, where $\mathbf{a}_k$ is the k-th column of $\mathbf{A}$, i.e., $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_K]$. Because of $\bar{\mathbf{R}}_s = \mathbf{A}\mathbf{R}_s\mathbf{A}^H$, it can be obtained that

$$Col(\bar{\mathbf{R}}_s) = Span\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_K\} \tag{7}$$

where $Col(\cdot)$ stands for the column subspace. Denoting $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_M]$, and based on (4), we have

$$Col(\bar{\mathbf{R}}_s) = Span\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_K\}. \tag{8}$$

Comparing (7) and (8), we can have $Span\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_K\} = Span\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_K\}$, i.e., the signal subspace is exactly equal to the subspace spanned by the K eigenvectors corresponding to K biggest eigenvalues of $\mathbf{R}_x$. (**Conclusion 2**)

The noise subspace should be defined as the subspace which is orthogonal with the signal subspace. Because of the orthogonality between every two eigenvectors of $\mathbf{R}_x$, we have $Span\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_K\}$ is orthogonal with $Span\{\mathbf{u}_{K+1}, \cdots, \mathbf{u}_M\}$. Therefore, $Span\{\mathbf{u}_{K+1}, \cdots, \mathbf{u}_M\}$ can be viewed as the noise subspace, i.e., the noise subspace is equal to the subspace spanned by the $M-K$ eigenvectors corresponding to $M-K$ smallest eigenvalues of $\mathbf{R}_x$. (**Conclusion 3**)