Properties of the Covariance Matrix

The covariance matrix of a random vector $\mathbf{X} \in \mathbf{R}^n$ with mean vector $\mathbf{m}_x$ is defined via:

$$\mathbf{C}_x = E[(\mathbf{X}-\mathbf{m})(\mathbf{X}-\mathbf{m})^T].$$

The $(i,j)^{\text{th}}$ element of this covariance matrix $\mathbf{C}_x$ is given by

$$C_{ij} = E[(X_i - m_i)(X_j - m_j)] = \sigma_{ij}.$$

The diagonal entries of this covariance matrix $\mathbf{C}_x$ are the variances of the components of the random vector $\mathbf{X}$, i.e.,

$$C_{ii} = E[(X_i - m_i)^2] = \sigma_i^2.$$

Since the diagonal entries are all positive the trace of this covariance matrix is positive, i.e.,

$$\text{Trace}(\mathbf{C}_x) = \sum_{i=1}^{n} C_{ii} > 0.$$

This covariance matrix $\mathbf{C}_x$ is symmetric, i.e., $\mathbf{C}_x = \mathbf{C}_x^T$ because :

$$C_{ij} = \sigma_{ij} = \sigma_{ji} = C_{ji}.$$

The covariance matrix $\mathbf{C}_x$ is positive semidefinite, i.e., for $\mathbf{a} \in \mathbf{R}^n$:

$$\begin{aligned} E\{[(\mathbf{X}-\mathbf{m})^T\mathbf{a}]^2\} &= E\{[(\mathbf{X}-\mathbf{m})^T\mathbf{a}]^T[(\mathbf{X}-\mathbf{m})^T\mathbf{a}]\} \geq 0 \\ E[\mathbf{a}^T(\mathbf{X}-\mathbf{m})(\mathbf{X}-\mathbf{m})^T\mathbf{a}] &\geq 0, \quad \mathbf{a} \in \mathbf{R}^n \\ \mathbf{a}^T\mathbf{C}_x\mathbf{a} &\geq 0, \quad \mathbf{a} \in \mathbf{R}^n. \end{aligned}$$

Since the covariance matrix $\mathbf{C}_x$ is symmetric, i.e., self-adjoint with the usual inner product its eigenvalues are all real and positive and the eigenvectors that belong to distinct eigenvalues are orthogonal, i.e.,

$$\mathbf{C}_x = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T = \sum_{i=1}^{n} \lambda_i \vec{v}_i \vec{v}_i^T.$$

As a consequence, the determinant of the covariance matrix is positive, i.e.,

$$\text{Det}(\mathbf{C}_X) = \prod_{i=1}^{n} \lambda_i \geq 0.$$

The eigenvectors of the covariance matrix transform the random vector into statistically uncorrelated random variables, i.e., into a random vector with a diagonal covariance matrix. The Rayleigh coefficient of the covariance matrix is bounded above and below by the maximum and minimum eigenvalue :

$$\lambda_{\min} \;\leq\; \frac{\mathbf{a}^T\mathbf{C}_x\mathbf{a}}{\mathbf{a}^T\mathbf{a}}, \quad \mathbf{a} \in \mathbf{R} \;\leq\; \lambda_{\max}.$$