2.0 Random Vectors

2.1 Definitions of Correlation Matrices

Definition: Let $X_1, X_2, \ldots, X_n$ be n random variables defined on (U, F, P). Then,

$$\mathbf{X}(u) = \begin{bmatrix} X_1(u) \\ X_2(u) \\ \vdots \\ X_n(u) \end{bmatrix}$$

is a *random vector.*

To fully characterize $\mathbf{X}(u)$ we need an $n-$dimensional joint *pdf*. This would often result in very cumbersome calculations when utilized. Instead of doing this we usually just use 1st and 2nd order statistics in our study of random vectors. This is sufficient for most needs. We often write $\mathbf{X}$ for $\mathbf{X}(u)$.

Definition: $\mu_{\mathbf{X}} = E[\mathbf{X}]$ is called the *mean vector.* Note

$$\mu_{\mathbf{X}} = [E[X_1(u)], \ldots, E[X_n(u)]]^t \,.$$

Here t denotes the transpose.

Definition: $R_{\mathbf{X}} = E\left[\mathbf{X}\mathbf{X}^\dagger\right]$ is called the *correlation matrix.* Here $\dagger$ denotes the conjugate transpose, i.e.,

$$\mathbf{X}^\dagger(u) = (X_1^*(u), \ldots, X_n^*(u)) \,.$$

Thus

$$R_{\mathbf{X}} = \begin{bmatrix} E[X_1(u)X_1^*(u)] & \ldots & E[X_1(u)X_n^*(u)] \\ \vdots & \ldots & \vdots \\ E[X_n(u)X_1^*(u)] & \ldots & E[X_n(u)X_n^*(u)] \end{bmatrix}.$$

Definition: $K_{\mathbf{X}} = E\left[(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{X} - \mu_{\mathbf{X}})^\dagger\right]$ is called the *covariance matrix.*

Note: $K_{\mathbf{X}} = R_{\mathbf{X}} - \mu_{\mathbf{X}}\mu_{\mathbf{X}}^\dagger$.

Definition: If $\mathbf{X}(u)$ and $\mathbf{Y}(u)$ are two random vectors then $R_{\mathbf{XY}} = E\left[\mathbf{X}\mathbf{Y}^\dagger\right]$ is called the *cross-correlation matrix.*

Note: $R_{\mathbf{XY}} = R^{\dagger}_{\mathbf{YX}}$.

Definition: $K_{\mathbf{XY}} = E\left[(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{Y} - \mu_{\mathbf{Y}})^{\dagger}\right]$ is called the *cross-covariance matrix.*

Note: $K_{\mathbf{XY}} = R_{\mathbf{XY}} - \mu_{\mathbf{X}}\mu^{\dagger}_{\mathbf{Y}}$.

Let $\mathbf{Z} = [\mathbf{XY}]^t$. Then the correlation matrix for $\mathbf{Z}$ is

$$R_{\mathbf{Z}} = E\left[\mathbf{Z}\mathbf{Z}^{\dagger}\right] = \begin{bmatrix} R_{\mathbf{X}} & R_{\mathbf{XY}} \\ R_{\mathbf{YX}} & R_{\mathbf{Y}} \end{bmatrix}.$$

2.2 Properties of Correlation Matrices

Definition: A matrix $\mathbf{M}$ is said to be *Hermitian symmetric* if $\mathbf{M} = \mathbf{M}^{\dagger}$.

Note:

$$R^{\dagger}_{\mathbf{X}} = \left(E\left[\mathbf{X}\mathbf{X}^{\dagger}\right]\right)^{\dagger} = E\left[\left(\mathbf{X}\mathbf{X}^{\dagger}\right)^{\dagger}\right] = E\left[\left((\mathbf{X}^{\dagger})^{\dagger}\mathbf{X}^{\dagger}\right)\right] = E\left[\mathbf{X}\mathbf{X}^{\dagger}\right] = R_{\mathbf{X}}$$

so correlation matrices are Hermitian symmetric.

Definition: A Hermitian symmetric matrix $\mathbf{M}$ is said to be *non-negative definite* if for any complex vector $\mathbf{a}$

$$\mathbf{a}^{\dagger}\mathbf{M}\mathbf{a} \geq 0.$$

Claim: Correlation matrices are non-negative definite.

Proof: We just need to show $\mathbf{a}^\dagger R_{\mathbf{X}}\mathbf{a} \geq 0$.

$$\mathbf{a}^\dagger R_{\mathbf{X}}\mathbf{a} = (a_1^* \ldots a_n^*)\, E\left[\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} (\, X_1^* \cdots X_n^* \,)\right] \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

$$= E\left[\left(\sum_{i=1}^{n} a_i^* X_i\right)\left(\sum_{j=1}^{n} a_j X_j^*\right)\right] = E\left[\left(\sum_{i=1}^{n} a_i^* X_i\right)\left(\sum_{j=1}^{n} a_j^* X_j\right)^*\right]$$

$$= E\left[\left|\sum_{i=1}^{n} a_i^* X_i\right|^2\right] \geq 0.$$

Also, $\mathbf{a}^\dagger K_{\mathbf{X}}\mathbf{a} \geq 0$.

2.3 Linear Transformations of Random Vectors

$\mathbf{Y}(u)$ is formed by a linear transformation of $\mathbf{X}(u)$. Here $\mathbf{X}(u) \in \mathbf{R}^n$ and $\mathbf{Y}(u) \in \mathbf{R}^m$.

$$Y_i(u) = \sum_{j=1}^{n} h_{ij} X_j(u), \quad i = 1, 2, \ldots, m$$

or

$$\mathbf{Y}(u) = \mathbf{H}\mathbf{X}(u)$$

where

$$H = \begin{bmatrix} h_{11} & \ldots & h_{1n} \\ \vdots & \ldots & \vdots \\ h_{m1} & \ldots & h_{mn} \end{bmatrix}.$$

Let us now look at the first and second moments.

$$\mu_{\mathbf{Y}} = E[\mathbf{Y}(u)] = E[\mathbf{H}\mathbf{X}(u)] = \mathbf{H}E[\mathbf{X}(u)] = \mathbf{H}\mu_{\mathbf{X}}.$$

$$\mathbf{R}_{\mathbf{Y}} = E\left[\mathbf{Y}(u)\mathbf{Y}^\dagger(u)\right] = E\left[\mathbf{H}\mathbf{X}(u)\,(\mathbf{H}\mathbf{X}(u))^\dagger\right] = \mathbf{H}E\left[\mathbf{X}(u)\mathbf{X}(u)^\dagger\right]\mathbf{H}^\dagger$$

$$= \mathbf{H}\mathbf{R}_{\mathbf{X}}\mathbf{H}^\dagger.$$

Also, $\mathbf{K}_{\mathbf{Y}} = \mathbf{H}\mathbf{K}_{\mathbf{X}}\mathbf{H}^\dagger$.

Question: Given a vector $\mathbf{X}$ of n uncorrelated random variables with zero mean and unit variance how do we transform this vector into a vector $\mathbf{Y}$ with mean $\mathbf{c}$ and covariance $\mathbf{K_Y}$?

Let

$$\tilde{\mathbf{Y}}(u) = \mathbf{H}\mathbf{X}(u).$$

Then

$$\mu_{\tilde{\mathbf{Y}}} = \mathbf{H}\mu_{\mathbf{X}} = 0$$

and

$$\mathbf{K}_{\tilde{\mathbf{Y}}} = \mathbf{H}\mathbf{K_X}\mathbf{H}^{\dagger}.$$

Now $\mathbf{K_X}$ is an $n \times n$ identity matrix, $\mathbf{I}_n$. Thus

$$\mathbf{K}_{\tilde{\mathbf{Y}}} = \mathbf{H}\mathbf{I}_n\mathbf{H}^{\dagger} = \mathbf{H}\mathbf{H}^{\dagger}.$$

We now let

$$\mathbf{Y}(u) = \tilde{\mathbf{Y}}(u) + \mathbf{c}.$$

Then

$$\mathbf{Y}(u) = \mathbf{H}\mathbf{X}(u) + \mathbf{c}.$$

Hence,

$$\mu_{\mathbf{Y}} = \mathbf{c}$$

and

$$\mathbf{K_Y} = \mathbf{K}_{\tilde{\mathbf{Y}}} = \mathbf{H}\mathbf{H}^{\dagger}.$$

Problem: We need to find $\mathbf{H}$ given some $\mathbf{K_Y}$. This is a matrix factorization problem that we will deal with later in the course.