7-1 General Concepts

Definitions

- A *random vector* is the vector

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix}$$

  whose components are $\mathbf{x}_i$ are random variables.

- The *distribution* of $\mathbf{X}$ is the *joint distribution* of the elements of the vector. For

$$X = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$$

  The distribution of $\mathbf{X}$ is

$$F_{\mathbf{X}}(X) = P(\mathbf{x}_1 \le x_1, \ldots, \mathbf{x}_n \le x_n)$$

- If the random variables in $\mathbf{X}$ are jointly continuous then the joint density is

$$f_{\mathbf{X}}(X) = \frac{\partial^n F_{\mathbf{X}}(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n}$$

- If the random variables in $\mathbf{X}$ are jointly discrete then the probability mass function is

$$P(\mathbf{X} = X) = P(\mathbf{x}_1 = x_1, \ldots, \mathbf{x}_n = x_n)$$

- The random variables $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are called (mutually) *independent* if the events

$$\{\zeta \in \mathcal{S} \colon \mathbf{x}_1 \le x_1\}, \ldots, \{\zeta \in \mathcal{S} \colon \mathbf{x}_n \le x_n\}$$

  are independent.

Notes

- $F_{\mathbf{X}}(X) \colon \mathbb{R}^n \to [0,1] \in \mathbb{R}$
- $f_{\mathbf{X}}(X) \colon \mathbb{R}^n \to \mathbb{R}$.

Definitions

- A *random vector* is the vector

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix}$$

  whose components are $\mathbf{x}_i$ are random variables.

- The *distribution* of $\mathbf{X}$ is the *joint distribution* of the elements of the vector. For

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

  The distribution of $\mathbf{X}$ is

$$F_{\mathbf{X}}(X) = P(\mathbf{x}_1 \le x_1, \ldots, \mathbf{x}_n \le x_n)$$

- If the random variables in $\mathbf{X}$ are jointly continuous then the joint density is

$$f_{\mathbf{X}}(X) = \frac{\partial^n F_{\mathbf{X}}(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n}$$

- If the random variables in $\mathbf{X}$ are jointly discrete then the probability mass function is

$$P(\mathbf{X} = X) = P(\mathbf{x}_1 = x_1, \ldots, \mathbf{x}_n = x_n)$$

- The random variables $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are called (mutually) *independent* if the events

$$\{\zeta \in \mathcal{S} \colon \mathbf{x}_1 \le x_1\}, \ldots, \{\zeta \in \mathcal{S} \colon \mathbf{x}_n \le x_n\}$$

  are independent.

Notes

- $F_{\mathbf{X}}(X) \colon \mathbb{R}^n \to [0,1] \in \mathbb{R}$
- $f_{\mathbf{X}}(X) \colon \mathbb{R}^n \to \mathbb{R}$.

Properties

1. $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are independent:

$$F_{\mathbf{X}}(X) = F_{\mathbf{x}_1}(x_1) \cdots F_{\mathbf{x}_n}(x_n)$$
$$f_{\mathbf{X}}(X) = f_{\mathbf{x}_1}(x_1) \cdots f_{\mathbf{x}_n}(x_n)$$

2. $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are independent and identically distributed (IID):

$$F_{\mathbf{X}}(X) = F_{\mathbf{x}}(x_1) \cdots F_{\mathbf{x}}(x_n)$$
$$f_{\mathbf{X}}(X) = f_{\mathbf{x}}(x_1) \cdots f_{\mathbf{x}}(x_n)$$

   where $F_{\mathbf{x}}(\cdot)$ is the common CDF and $f_{\mathbf{x}}(\cdot)$ is the common PDF.

3. Marginal Distributions

$$F_{\mathbf{x}_1}(x_1) = F_{\mathbf{X}}(x_, \infty, \ldots, \infty)$$

4. Marginal PDFs for jointly continuous RVs

$$f_{\mathbf{x}_1}(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, x_2, \ldots, x_n)\, dx_2 \cdots dx_n$$

5. Marginal PMFs for jointly discrete RVs

$$P(\mathbf{x}_1 = x_1) = \sum_{i_2} \cdots \sum_{i_n} P(\mathbf{x}_1 = x_1, \mathbf{x}_2 = x_{i_2}, \ldots, \mathbf{x}_n = x_{i_n})$$

6. Expectation (continuous RVs)

$$\begin{aligned} E\{g(\mathbf{X})\} &= E\{g(\mathbf{x}_1, \ldots, \mathbf{x}_n)\} \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n) f_{\mathbf{X}}(x_1, \ldots, x_n)\, dx_1 \cdots dx_n \end{aligned}$$

7. Expectation (discrete RVs)

$$\begin{aligned} E\{g(\mathbf{X})\} &= E\{g(\mathbf{x}_1, \ldots, \mathbf{x}_n)\} \\ &= \sum_{i_1} \cdots \sum_{i_n} g(x_{i_1}, \ldots, x_{i_n}) P(\mathbf{x}_1 = x_{i_1}, \ldots, \mathbf{x}_n = x_{i_n}) \end{aligned}$$

Vector/Matrix Definitions

Vector Definitions

- Vector, conjugate, transpose, Hermitian (conjugate transpose)

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

- Conjugate

$$X^* = \begin{bmatrix} x_1^* \\ \vdots \\ x_n^* \end{bmatrix}$$

- Transpoose

$$X^t = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$$

- Hermitian (conjugate-transpose)

$$X^H = \begin{bmatrix} x_1^* & \cdots & x_n^* \end{bmatrix}$$

Matrix Definitions

- Matrix

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{bmatrix}$$

- Conjugate

$$X^* = \begin{bmatrix} x_{11}^* & \cdots & x_{1n}^* \\ \vdots & & \vdots \\ x_{n1}^* & \cdots & x_{nn}^* \end{bmatrix}$$

- Transpose

$$X^t = \begin{bmatrix} x_{11} & \cdots & x_{n1} \\ \vdots & & \vdots \\ x_{1n} & \cdots & x_{nn} \end{bmatrix}$$

- Hermitian (conjugate-transpose)

$$X^H = \begin{bmatrix} x_{11}^* & \cdots & x_{n1}^* \\ \vdots & & \vdots \\ x_{1n}^* & \cdots & x_{nn}^* \end{bmatrix}$$

Vector Operations for $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$

- Inner product

$$X^H Y = \begin{bmatrix} x_1^* & \cdots & x_n^* \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} x_i^* y_i$$

  inner product: $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$

- Outer product

$$XY^H = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1^* & \cdots & y_n^* \end{bmatrix} = \begin{bmatrix} x_1 y_1^* & \cdots & x_1 y_n^* \\ \vdots & & \vdots \\ x_n y_1^* & \cdots & x_n y_n^* \end{bmatrix}$$

  outer product $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}^{n \times n}$

Matrix Operations

- Determinant: $\det(X)\colon \mathbb{C}^{n\times n} \to \mathbb{C}$
- Inverse: X^{-1} is the *inverse* of the square matrix X means

$$X^{-1}X = XX^{-1} = I$$

- Symmetric
  - the real matrix X is *symmetric* means $X = X^t$.
  - the complex-valued matrix X is *conjugate symmetric* or *Hermitian* means $X = X^H$.
- Unitary: the complex-valued matrix X is *unitary* means

$$XX^H = X^H X = I$$

- Eigen-decomposition $X = Q\Lambda Q^{-1}$ where

the n eigenvalues of X

$$Q = \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & & \vdots \\ v_{n1} & \cdots & v_{nn} \end{bmatrix} \quad \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

the eigenvector corresponding to λ_n

the eigenvector corresponding to λ_1

Statistical Vectors and Matrices for $\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix}$

- Mean Vector

$$\mu_{\mathbf{X}} = E\{\mathbf{X}\} = \begin{bmatrix} E\{\mathbf{x}_1\} \\ \vdots \\ E\{\mathbf{x}_n\} \end{bmatrix} = \begin{bmatrix} \mu_{\mathbf{x}_1} \\ \vdots \\ \mu_{\mathbf{x}_n} \end{bmatrix}$$

- Covariance Matrix

$$C_{\mathbf{XX}} = E\{(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{X} - \mu_{\mathbf{X}})^t\}$$

$$= E\left\{ \begin{bmatrix} \mathbf{x}_1 - \mu_{\mathbf{x}_1} \\ \vdots \\ \mathbf{x}_n - \mu_{\mathbf{x}_n} \end{bmatrix} \begin{bmatrix} (\mathbf{x}_1 - \mu_{\mathbf{x}_1}) & \cdots & (\mathbf{x}_n - \mu_{\mathbf{x}_n}) \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} (\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{x}_1 - \mu_{\mathbf{x}_1}) & \cdots & (\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{x}_n - \mu_{\mathbf{x}_n}) \\ \vdots & & \vdots \\ (\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{x}_1 - \mu_{\mathbf{x}_1}) & \cdots & (\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{x}_{\mathbf{x}_n} - \mu_{\mathbf{x}_n}) \end{bmatrix} \right\}$$

$$= \begin{bmatrix} E\{(\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{x}_1 - \mu_{\mathbf{x}_1})\} & \cdots & E\{(\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{x}_n - \mu_{\mathbf{x}_n})\} \\ \vdots & & \vdots \\ E\{(\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{x}_1 - \mu_{\mathbf{x}_1})\} & \cdots & E\{(\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{x}_n - \mu_{\mathbf{x}_n})\} \end{bmatrix}$$

$$= \begin{bmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & & \vdots \\ C_{n1} & \cdots & C_{nn} \end{bmatrix}$$

- Correlation Matrix

$$R_{\mathbf{XX}} = E\{\mathbf{X}\mathbf{X}^t\} = E\left\{ \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} \mathbf{x}_1\mathbf{x}_1 & \cdots & \mathbf{x}_1\mathbf{x}_n \\ \vdots & & \vdots \\ \mathbf{x}_n\mathbf{x}_1 & \cdots & \mathbf{x}_n\mathbf{x}_n \end{bmatrix} \right\} = \begin{bmatrix} E\{\mathbf{x}_1\mathbf{x}_1\} & \cdots & E\{\mathbf{x}_1\mathbf{x}_1\} \\ \vdots & & \vdots \\ E\{\mathbf{x}_n\mathbf{x}_1\} & \cdots & E\{\mathbf{x}_n\mathbf{x}_n\} \end{bmatrix}$$

$$= \begin{bmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & & \vdots \\ R_{n1} & \cdots & R_{nn} \end{bmatrix}$$

Statistical Matrices for

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_n \end{bmatrix}$$

- Cross-Covariance Matrix

$$C_{\mathbf{XY}} = E\{(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{Y} - \mu_{\mathbf{Y}})^t\}$$

$$= E\left\{ \begin{bmatrix} \mathbf{x}_1 - \mu_{\mathbf{x}_1} \\ \vdots \\ \mathbf{x}_n - \mu_{\mathbf{x}_n} \end{bmatrix} \begin{bmatrix} (\mathbf{y}_1 - \mu_{\mathbf{y}_1}) & \cdots & (\mathbf{y}_n - \mu_{\mathbf{y}_n}) \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} (\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{y}_1 - \mu_{\mathbf{y}_1}) & \cdots & (\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{y}_n - \mu_{\mathbf{y}_n}) \\ \vdots & & \vdots \\ (\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{y}_1 - \mu_{\mathbf{y}_1}) & \cdots & (\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{y}_n - \mu_{\mathbf{y}_n}) \end{bmatrix} \right\}$$

$$= \begin{bmatrix} E\{(\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{y}_1 - \mu_{\mathbf{y}_1})\} & \cdots & E\{(\mathbf{x}_1 - \mu_{\mathbf{x}_1})(\mathbf{y}_n - \mu_{\mathbf{y}_n})\} \\ \vdots & & \vdots \\ E\{(\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{y}_1 - \mu_{\mathbf{y}_1})\} & \cdots & E\{(\mathbf{x}_n - \mu_{\mathbf{x}_n})(\mathbf{y}_n - \mu_{\mathbf{y}_n})\} \end{bmatrix}$$

$$= \begin{bmatrix} C_{\mathbf{x}_1\mathbf{y}_1} & \cdots & C_{\mathbf{x}_1\mathbf{y}_n} \\ \vdots & & \vdots \\ C_{\mathbf{x}_n\mathbf{y}_1} & \cdots & C_{\mathbf{x}_n\mathbf{y}_n} \end{bmatrix}$$

- Cross-Correlation Matrix

$$R_{\mathbf{XY}} = E\{\mathbf{X}\mathbf{Y}^t\} = E\left\{ \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 & \cdots & \mathbf{y}_n \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} \mathbf{x}_1\mathbf{y}_1 & \cdots & \mathbf{x}_1\mathbf{y}_n \\ \vdots & & \vdots \\ \mathbf{x}_n\mathbf{y}_1 & \cdots & \mathbf{x}_n\mathbf{y}_n \end{bmatrix} \right\} = \begin{bmatrix} E\{\mathbf{x}_1\mathbf{y}_1\} & \cdots & E\{\mathbf{x}_1\mathbf{y}_1\} \\ \vdots & & \vdots \\ E\{\mathbf{x}_n\mathbf{y}_1\} & \cdots & E\{\mathbf{x}_n\mathbf{y}\mathbf{x}_n\} \end{bmatrix}$$

$$= \begin{bmatrix} R_{\mathbf{x}_1\mathbf{y}_1} & \cdots & R_{\mathbf{x}_1\mathbf{y}_n} \\ \vdots & & \vdots \\ R_{\mathbf{x}_n\mathbf{y}_1} & \cdots & R_{\mathbf{x}_n\mathbf{y}_n} \end{bmatrix}$$

Normal Random Vector: $\mathbf{X} \sim N(\mu_{\mathbf{X}}, C_{\mathbf{XX}})$ means

$$f_{\mathbf{X}}(X) = \frac{1}{\sqrt{(2\pi)^n \det(C_{\mathbf{XX}})}} \exp\left\{-\frac{1}{2}(X-\mu_{\mathbf{X}})^T C_{\mathbf{XX}}^{-1}(X-\mu_{\mathbf{X}})\right\}$$

where

$$\mu_{\mathbf{X}} = E\{\mathbf{X}\}$$

$$C_{\mathbf{XX}} = E\left\{(\mathbf{x}-\mu_{\mathbf{X}})(\mathbf{x}-\mu_{\mathbf{X}})^t\right\}$$

Because $X \in \mathbb{R}^n$, $f_{\mathbf{X}}(X)\colon \mathbb{R}^n \to \mathbb{R}$.

If $\mathbf{X} \sim N(\mu_{\mathbf{X}}, C_{\mathbf{XX}})$, then

$$\mathbf{y} = a_1\mathbf{x}_1 + \cdots + a_n\mathbf{x}_n = \underbrace{\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_n \end{bmatrix}}_{\mathbf{X}}$$

is also normal.

This generalizes: If $\mathbf{X} \sim N(\mu_{\mathbf{X}}, C_{\mathbf{XX}})$, then

$$\underbrace{\mathbf{Y}}_{k\times 1} = \underbrace{A}_{k\times n}\underbrace{\mathbf{X}}_{n\times 1}$$

is normal: $\mathbf{Y} \sim N(\mu_{\mathbf{Y}}, C_{\mathbf{YY}})$