Random Vectors

Random Vectors

- We are now moving from a univariate random variable to multivariate random variables, also called as random vectors.
- An n-dimensional random vector is a column vector $\mathbf{X} = (X_1, \ldots X_n)^T$ whose components X_i are scalar valued random variables defined on the same space $(\Omega, \mathcal{F}, P)$.
- Since the components are on the same space, they may be correlated with each other.
- Example:$\mathbf{X} = (X_1, X_2)^T$ where $X_1 = Z_1$ ans $X_2 = Z_1 + Z_2$ where Z_1 and Z_2 are independent standard normal.
- What is the pdf, cdf, marginals, mean, variance/covariance of X?

Random Vectors - Notation

- The CDF and pdf of the random vector $\mathbf{X}$ is denoted as follows :

$$F_{\mathbf{X}}(\mathbf{x}) = F_{X_1,\ldots X_n}(x_1, \ldots x_n)$$

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1,\ldots X_n}(x_1, \ldots x_n)$$

- The joint CDF/pdf captures the correlation between components.

- The expected value vector $E[\mathbf{X}] = (E[X_1], \ldots, E[X_n])^T$

- Linearity of expectation hold here and so for any deterministic matrix $\mathbf{A}$ and vector $\mathbf{b}$ and $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ we have

$$E[\mathbf{Y}] = \mathbf{A}E[\mathbf{X}] + \mathbf{b}.$$

Covariance matrix

- The covariance matrix $C_{\mathbf{X}}$ captures the covariance between components and is defined by

$$
\begin{aligned}
C_{\mathbf{X}} &= E[(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])^T] \\
&= \begin{bmatrix}
Var(X_1) & Cov(X_1, X_2) & \dots & Cov(X_1, X_n) \\
Cov(X_2, X_1) & Var(X_2) & \dots & Cov(X_2, X_n) \\
\vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots \\
Cov(X_n, X_1) & Cov(X_n, X_2) & \dots & Var(X_n)
\end{bmatrix}
\end{aligned}
$$

Covariance matrix: Properties

- The covariance matrix $C_{\mathbf{X}}$ is always positive semi-definite, i.e., for any vector $a \neq 0$ we have $a^T C_{\mathbf{X}} a \geq 0$. Why ?

  Let $u = a^T(\mathbf{X} - E[\mathbf{X}])$, then $a^T C_{\mathbf{X}} a = E[uu^T] = E[u^2] \geq 0$

- If $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$, show that $C_{\mathbf{Y}} = AC_{\mathbf{X}}A^T$. (HW)

- Now recall how we obtained the pdf of Y from pdf of X when $Y = g(X)$

  Consider $Y = g(X)$ where g is monotone, continuous, differentiable. Then $f_Y(y) = f_X(h(y))|\frac{dh}{dy}(y)|$ where h is the inverse function of g.

- How does this generalize to $\mathbf{Y} = G(\mathbf{X})$? How do we get $f_{\mathbf{Y}}$ from $f_{\mathbf{X}}$?

Functions of random vectors

- Let $\mathbf{Y} = G(\mathbf{X})$ where $G : \mathbb{R}^n \to \mathbb{R}^n$, continuous invertible with continuous partial derivatives.
- Then one can write $\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} G_1(X_1, \dots X_n) \\ G_2(X_1, \dots X_n) \\ \vdots \\ G_n(X_1, \dots, X_n) \end{bmatrix}$
- For example if $\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} 2X_1 \\ X_1 + X_2 \end{bmatrix}$ then $G_1(X_1, X_2) = 2X_1$ and $G_2(X_1, X_2) = X_1 + X_2$.
- What does continuity of G mean? Continuity of components?

Functions of random vectors

- Let H denote inverse of G. We similarly have
$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \dots \\ X_n \end{bmatrix} = \begin{bmatrix} H_1(Y_1, \dots Y_n) \\ H_2(Y_1, \dots Y_n) \\ \vdots \\ H_n(Y_1, \dots, Y_n) \end{bmatrix}$$
- For the example we have $X_1 = H_1(Y_1, Y_2) =$ and $X_2 = H_2(Y_1, Y_2) = Y_2 - \frac{Y_1}{2}$.

Functions of random vectors

Let $\mathbf{Y} = G(\mathbf{X})$ where $G : \mathbb{R}^n \to \mathbb{R}^n$, continuous invertible with continuous partial derivatives. Let H denote its inverse. Then

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(H(\mathbf{y}))|J|$$

where J is the determinant of the Jacobian matrix given by

$$\begin{bmatrix} \frac{\partial H_1}{\partial y_1} & \frac{\partial H_1}{\partial y_2} & \cdots & \frac{\partial H_1}{\partial y_n} \\ \frac{\partial H_2}{\partial y_1} & \frac{\partial H_2}{\partial y_2} & \cdots & \frac{\partial H_2}{\partial y_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial H_n}{\partial y_1} & \frac{\partial H_n}{\partial y_2} & \cdots & \frac{\partial H_n}{\partial y_n} \end{bmatrix}$$

Jacobian determinant

- From Vector Calculus: The Jacobian gives the ratio of the incremental areas $dx_1 dx_2..dx_n$ and $dy_1, \ldots dy_n$.
- https://en.wikipedia.org/wiki/Jacobian_matrix_and_determinant
- HW1: For the running example, find $f_{\mathbf{Y}}(y)$.
- HW2: When $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$, how that

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{|det(A)|} f_{\mathbf{X}}(A^{-1}(\mathbf{y} - \mathbf{b}))$$

Standard Normal Vector

- An n length random vector **Z** is called as a standard normal vector if its components Z_i are independent and standard normal.
- What is $E[\mathbf{Z}]$ and $C_{\mathbf{Z}}$?
- Show that the pdf is given by

$$f_{\mathcal{Z}}(\mathbf{z}) = \frac{1}{(2\pi)^{n/2}} e^{\{-\frac{1}{2}\mathbf{z}^T\mathbf{z}\}}$$

- Now suppose $\mathbf{X} = A\mathbf{Z} + \mu$. What is $E[\mathbf{X}]$ and $C_{\mathbf{X}}$?
- $E[\mathbf{X}] = \mu$ and $C_{\mathbf{X}} = AA^T$.
- Note that A can have dimension $n \times l$ in which case **Z** is an l length standard normal.

Gaussian Random Vectors

- Consider $\mathbf{X} = A\mathbf{Z} + \mu$. and $E[\mathbf{X}] = \mu$ and $C_{\mathbf{X}} = AA^T$.
- What is $f_{\mathbf{X}}(\mathbf{x})$?

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &= \frac{1}{|det(A)|} f_{\mathbf{Z}}(A^{-1}(\mathbf{x} - \mu)) \\ &= \frac{1}{(2\pi)^{n/2}\sqrt{det(C_{\mathbf{X}})}} e^{\{-\frac{1}{2}(\mathbf{x}-\mu)^T C_X^{-1}(\mathbf{x}-\mu)\}} \end{aligned}$$

> A random vector is Gaussian iff for some A and μ, it can be written as $\mathbf{X} = A\mathbf{Z} + \mu$

For equivalent definitions see `https://en.wikipedia.org/wiki/Multivariate_normal_distribution`