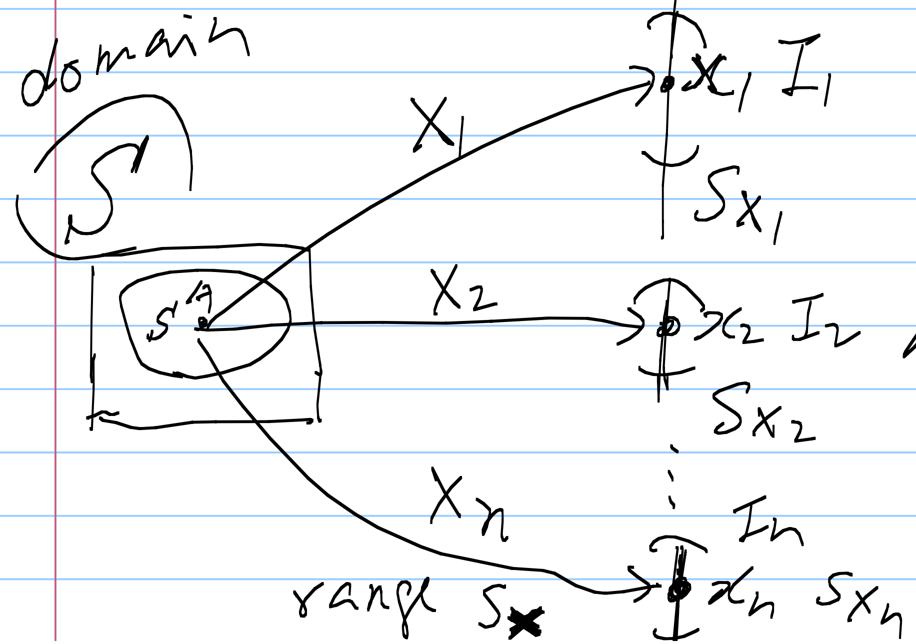
便箋標題 2008/5/2

CHAP 5 Random Vectors

Defn: $\mathbf{X} \equiv [X_1, X_2, \cdots, X_n]'$ where $'$ denotes transpose and $X_1, X_2, \cdots, X_n$ are random variables, is called an n-dimensional random vector.

Note: When $n=1$, $\mathbf{X}$ becomes a random variable.



We can denote any event A in S by

$$A = \{s \mid X_1(s) \in I_1, X_2(s) \in I_2, \cdots, X_n(s) \in I_n\}$$

$$\equiv \{X_1 \in I_1, X_2 \in I_2, \cdots, X_n \in I_n\}$$

where $\mathbb{I} \equiv \{ \underbrace{[x_1, x_2, \dots, x_n]'}_{\equiv \mathbf{X}'} \mid x_1 \in I_1, x_2 \in I_2, \dots, x_n \in I_n \}$ $\textcircled{=} \{ \mathbf{X} \in \mathbb{I} \}$

Then, $P[A] = P[\mathbf{X} \in \mathbb{I}]$

Defn 5.3: Multivariate Joint PDF

The joint PDF of a continuous $X_1, X_2, \cdots, X_n$ is defined by

$$f_{X_1, \cdots, X_n}(x_1, \cdots, x_n) = \frac{\partial^n F_{X_1, \cdots, X_n}(x_1, \cdots, x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n}$$

provided that $F_{X_1, \cdots, X_n}(x_1, \cdots, x_n)$ is differentiable.

$$\mathbf{W} \triangleq [\mathbf{X}', \mathbf{Y}']' \Rightarrow f_{\mathbf{W}}(\mathbf{w}) = f_{\mathbf{X}, \mathbf{Y}}(\mathbf{x}, \mathbf{y})$$

$$f_{\underline{V},\underline{W}}(\underline{v},\underline{w}) = \begin{cases} 4, & 0 \le v_1 \le v_2 \le 1,\ 0 \le w_1 \le w_2 \le 1 \\ 0, & \text{otherwise} \end{cases}$$

$$f_{\underline{V}}(\underline{v}) = \int_0^1 \int_{w_1}^1 4 \, dw_2 \, dw_1 = \int_0^1 4(1-w_1)\, dw_1 = \left(4w_1 - 2w_1^2\right)\Big|_0^1$$

$$\Longrightarrow f_{\underline{V}}(\underline{v}) = \begin{cases} 2, & 0 \le v_1 \le v_2 \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Similarly,

$$f_{\underline{W}}(\underline{w}) = \int_0^1 \int_{v_1}^1 4 \, dv_2 \, dv_1 \qquad 0 \le w_1 \le w_2 \le 1$$

$$\Rightarrow f_{\underline{W}}(\underline{w}) = \begin{cases} 2, & 0 \le w_1 \le w_2 \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Thus, $f_{\underline{V},\underline{W}}(\underline{v},\underline{w}) = f_{\underline{V}}(\underline{v})\, f_{\underline{W}}(\underline{w})$. Thus, $\underline{V}$ and $\underline{W}$ are independent.

$\underline{X}$, $\underline{W} = \underline{g}(\underline{X})$

(a) $Y = \max\{X_1, X_2, \cdots, X_n\}$

$$F_Y(y) \overset{\triangle}{=} P[Y \le y] = P[\max\{X_1, X_2, \cdots, X_n\} \le y]$$

$$= P[X_1 \le y, X_2 \le y, \cdots, X_n \le y]$$

X_i's iid

$$\overset{?}{=} \prod_{i=1}^{n} P[X_i \le y] \quad (\because \text{independent})$$

$$= \prod_{i=1}^{n} F_X(y) \quad (\because \text{identically distributed})$$

$$= (F_X(y))^n$$

$$\Rightarrow f_Y(y) = \frac{dF_Y(y)}{dy} = n(F_X(y))^{n-1} f_X(y) \quad \text{Q.E.D.}$$

(b) $W = \min\{X_1, X_2, \cdots, X_n\}$

$$P[W > w] = P[\min\{X_1, X_2, \cdots, X_n\} > w]$$
$$= P[X_1 > w, X_2 > w, \cdots, X_n > w]$$
$$= \prod_{i=1}^{n} \underbrace{P[X_i > w]}_{} \quad (\because \text{independent})$$
$$= \prod_{i=1}^{n} (1 - F_X(w)) \quad (\because \text{identically distributed})$$

$$\Rightarrow F_W(w) = 1 - P[W > w] = 1 - (1 - F_X(w))^n$$

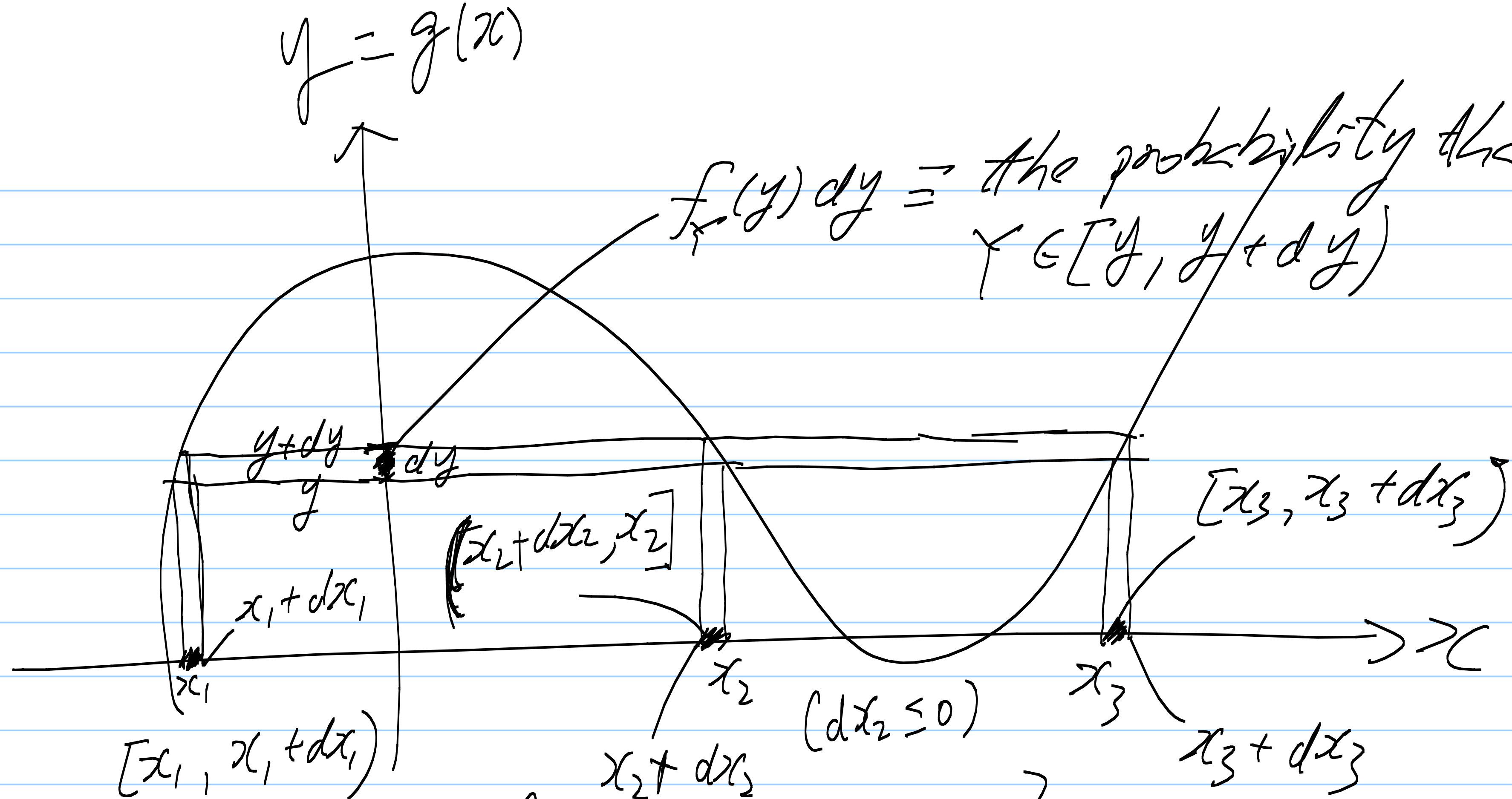
$$f_W(w) = \frac{dF_W(w)}{dw} = n(1 - F_X(w))^{n-1} f_X(w) \quad \text{Q.E.D.}$$

$\Rightarrow$ Order Statistic

Consider $Y = g(X)$ where $y = g(x)$ is a well-defined function. For a given y, there exist $x_1(y), x_2(y), \dots, x_n(y)$ satisfying $y = g(x_i)$, $i \in \{1, 2, \dots, n\}$. Then, $f_X(x)$ and $f_Y(y)$ are related by

$$f_Y(y) = \sum_{i=1}^{n} f_X(x_i(y)) \left\{ \left| \left. \frac{dg(x)}{dx} \right|_{x = x_i(y)} \right| \right\}^{-1}$$

This is called the Jacobian rule of obtaining $f_Y(y)$ from $f_X(x)$.



Thus, the event $\{Y \in [y, y+dy)\}$

$= \{X \in [x_1, x_1+dx_1)\} \cup \{X \in [x_2+dx_2, x_2]\}$

$$\cup \underbrace{\{X \in [x_3, x_3 + dx_3)\}}$$

$$\xRightarrow{dy \to 0} f_Y(y)\,dy = P[Y \in [y, y+dy)]$$

$$\sim = P[X \in [x_1, x_1 + dx_1)] + P[X \in (x_2 + dx_2, x_2]]$$

$$+ P[X \in [x_3, x_3 + dx_3)]$$

$$= f_X(x_1)\,dx_1 + f_X(x_2)\,|dx_2| + f_X(x_3)\,dx_3$$

$$\Rightarrow f_Y(y) = \sum_{i=1}^{3} f_X(x_i) \left| \frac{dx}{dy} \right|_{x=x_i}$$

Defn: $X_1, X_2, \cdots, X_n$ are called jointly Gaussian, or jointly normal, if and only if any of their linear combinations, i.e., $\sum_{i=1}^{n} a_i X_i$ for any a_i's with $a_i \neq 0$ for all i, is marginally Gaussian.

Notes: ① X_i is marginally Gaussian.

② Given $X_1, X_2, \cdots, X_n$ are marginally Gaussian, $X_1, X_2, \cdots, X_n$ may not be jointly Gaussian.

$R_{\underline{X}}$ and $C_{\underline{X}}$ satisfy the following properties:

1. They are nonnegative definite, i.e., for any $\underline{\alpha}$, $\underline{\alpha}' R_{\underline{X}} \underline{\alpha} \geq 0$.

$$\left(\underline{\alpha}' R_{\underline{X}} \underline{\alpha} = \underline{\alpha}' E\left[\underline{X}\,\underline{X}'\right] \underline{\alpha} = E\left[(\underline{\alpha}'\underline{X})(\underline{\alpha}'\underline{X})'\right] = E\left[(\underline{\alpha}'\underline{X})^2\right] \geq 0 \right)$$

2. Their determinants are nonnegative.
3. Their eigenvalues are nonnegative.

Singular Value Decomposition
(Eigenvalue Decomposition)

Consider an $n \times n$ symmetric ~~square~~ matrix A.
Find its eigenvectors $\underline{e}_i$, $i = 1, 2, \cdots, n$, and eigenvalues λ_i, $i = 1, 2, \cdots, n$, i.e., $\boxed{A\underline{e}_i = \lambda_i \underline{e}_i}$, for $i = 1, 2, \cdots, n$. (λ_i is the eigenvalue of A corresponding to eigenvector $\underline{e}_i$).

$$\underset{n\times n}{A} \underbrace{\begin{bmatrix} \underline{e}_1 & \underline{e}_2 & \cdots & \underline{e}_n \end{bmatrix}}_{n\times n \;\equiv\; E \text{ eigenmatrix}} = \underbrace{\begin{bmatrix} \underline{e}_1 & \underline{e}_2 & \cdots & \underline{e}_n \end{bmatrix}}_{E} \underbrace{\begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_n \end{bmatrix}}_{\text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) \;\equiv\; D}$$

$\Rightarrow AE = ED$

That is, A can be decomposed into

$$A = EDE'$$

Since $E'E = I = EE'$ ($e_i' e_j = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$)

i.e., E contains orthonormal column vectors.

Since A is symmetric, $\lambda_i \geq 0$, $\forall i$. Thus,

$$A = ED^{1/2} D^{1/2} E' = (\underline{ED^{1/2}})(ED^{1/2})'$$

where $D^{1/2} = \mathrm{diag}[\lambda_1^{1/2}, \lambda_2^{1/2}, \cdots, \lambda_n^{1/2}]$. If we let $B = ED^{1/2}$, then $A = BB'$.

$$\underline{X} = \underbrace{U D^{1/2}}_{A} \underline{Z} + \underline{\mu}_X$$

where $C_X = U D U'$ and $\underline{Z}$ is standard normal. Thus,

$$\underline{X} = \left[\underline{u}_1 \,\middle|\, \underline{u}_2 \,\middle|\, \cdots \,\middle|\, \underline{u}_n\right] \begin{bmatrix} \sqrt{d_1} & & & \\ & \sqrt{d_2} & & \\ & & \ddots & \\ & & & \sqrt{d_n} \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix} + \begin{bmatrix} \mu_{x_1} \\ \mu_{x_2} \\ \vdots \\ \mu_{x_n} \end{bmatrix}$$

$$= \left[\sqrt{d_1}\,\underline{u}_1 \,\middle|\, \sqrt{d_2}\,\underline{u}_2 \,\middle|\, \cdots \,\middle|\, \sqrt{d_n}\,\underline{u}_n\right] \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix} + \begin{bmatrix} \mu_{x_1} \\ \mu_{x_2} \\ \vdots \\ \mu_{x_n} \end{bmatrix}$$

$$\Rightarrow \underline{X} = \sum_{i=1}^{n} \sqrt{d_i}\, \underline{u}_i\, Z_i + \underline{\mu}_X$$

Recall that we need $\underbrace{2^n - n - 1}$ equations to justify the independence of n events $A_1, A_2, \cdots, A_n$.

Here, we only need ONE equation to justify the independence of n random variables $X_1, X_2, \cdots, X_n$, i.e.,

$$f_{X_1, \cdots, X_n}(x_1, \cdots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$$

$\Longleftrightarrow$ For any $2 \le k \le n$, $X_{i_1}, \cdots, X_{i_k}$ are independent for $i_1, i_2, \cdots, i_k$ being k distinct integers in

$\{1, 2, \cdots, n\}$.

Note: $f_{X,Y}(x, y) = f_X(x) f_Y(y)$ does NOT guarantee that

$$f_{X_1, \cdots, X_n, Y_1, \cdots, Y_m}(x_1, \cdots, x_n, y_1, \cdots, y_n)$$

$$= \prod_{i=1}^{n} f_{X_i}(x_i) \prod_{k=1}^{m} f_{Y_k}(y_k)$$

& (i.e., $X_1, \cdots, X_n, Y_1, \cdots, Y_m$ are independent).