04/02/2019

$$f_Y(y) = P(Y=-1)\ \delta(y+1) + P(Y=1)\ \delta(y-1)$$
$$= \frac{1}{4}\delta(y+1) + \frac{3}{4}\delta(y-1)$$

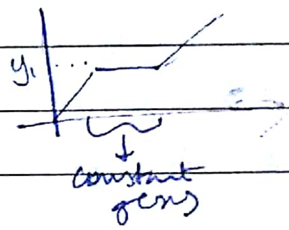
⇒ Transformation of a Random Variable $Y = g(X)$

given the density function of X, density function y as –

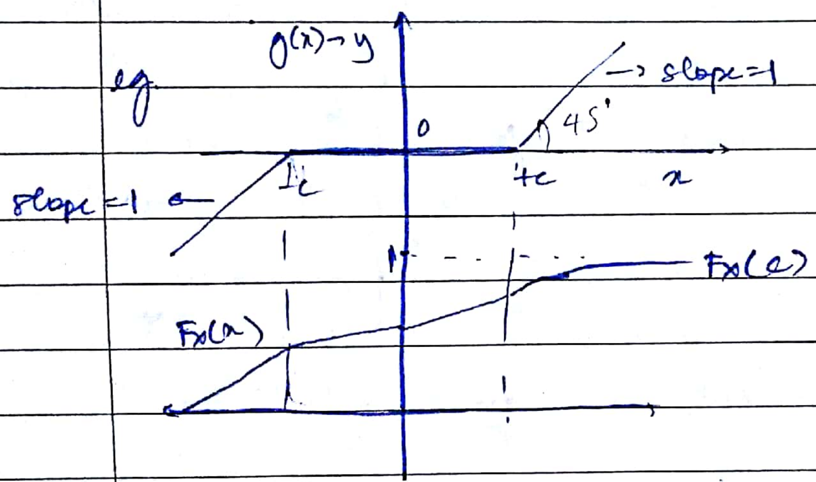
$$f_Y(y) = \frac{f_X(x_1)}{|g'(x_1)|} + \frac{f_X(x_2)}{|g'(x_2)|} + \cdots$$

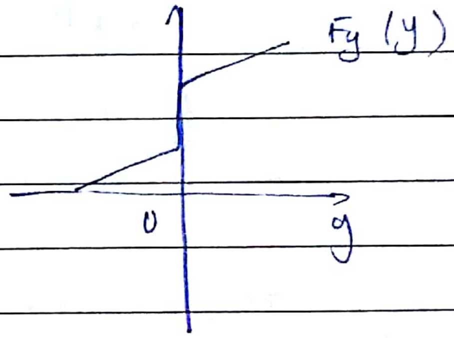
$$g'(x) = \frac{d}{dx}g(x) = \frac{dy}{dx}$$

If $g(x) = y_1$ in $(x_0, x_1)$ interval




eg




→ $Z = X + Y = g(x, y)$
$W = X - Y = h(x, y)$ } functions of two Random Variable

* Joint Density functions & Conditional Density functions.

$$F_Y(y|M) = \frac{P(Y \le y, M)}{P(M)} \quad \text{if } M\{X \le x\}$$

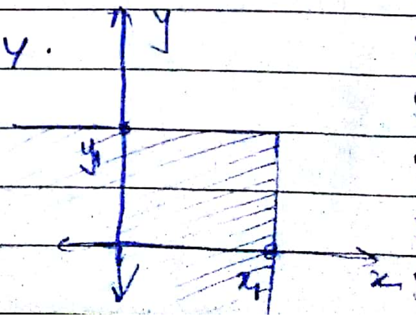
then $F_Y(y|M) = \dfrac{P(Y \le y, X \le x)}{P(M)} = \dfrac{F_{XY}(x, y)}{P(M)}$

$F_{XY}(x, y)$ → Joint CDF of R.V ~~x,y~~ X and Y.

Events — Space of $\{X(s), Y(s)\}$

A → $X \le x$ — Joint occurrences

B → $Y \le y$



__/__/__

$F_{XY}(x,y) \geq 0$  $-\infty < x < \infty$

$-\infty < y < \infty$

$F_{XY}(-\infty, y) = P(X < -\infty, Y \le y) = 0 = F_{XY}(x, -\infty)$

$F_{XY}(\infty, \infty) = 1$

$F_{XY}(\infty, y) = F_Y(y) \rightarrow$ CDF of $Y$

$F_{XY}(x, \infty) = F_X(x) \longrightarrow$ CDF of $X$

$$f_{XY}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y) \Rightarrow$$

$$F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(\alpha, \beta) \cancel{dy\, dx}\, d\alpha\, d\beta$$

$$\rightarrow F_{XY}(x,\infty) = F_X(x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f_{XY}(\alpha,\beta)\, d\beta\, d\alpha$$

We know, $f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx}\left[\int_{-\infty}^{x} f_X(\alpha)\, d\alpha\right]$

$$= \frac{d}{dx}\left[\int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} f_{XY}(\alpha,\beta)\, d\beta\right) d\alpha\right]$$

$$= \int_{-\infty}^{\infty} f_{XY}(x,\beta)\, d\beta = \int_{-\infty}^{\infty} f_{XY}(x,y)\, dy$$

$\rightarrow F_{Y|X}(y|x) = P(Y \le y \mid X = x)$ X taking a value $x$

$$= \lim_{\Delta x \to 0} P(Y \le y \mid x < X \le x + \Delta x)$$

$$= \lim_{\Delta x \to 0} F_Y(y \mid x < X \le x+\Delta x) \propto \frac{P(Y \le y,\ x < X \le \Delta x + x)}{P(x < X \le \Delta x + x)}$$

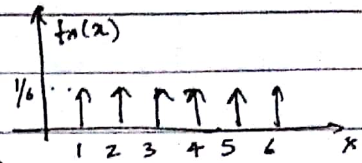
$$= \lim_{\Delta x \to 0} \frac{F_{XY}(x+\Delta x, y) - F_{XY}(x,y)}{F_X(x+\Delta x) - F_X(x)}$$

$$= \frac{\frac{\partial}{\partial x} F_{XY}(x,y)}{f_X(x)} = \frac{f_{XY}(x,y)}{f_X(x)}$$

04/02/2019

# Tutorial - 3

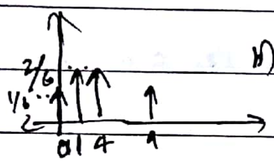
**Q.1.** $Y = X^2$ a) $f_X(x) = \frac{1}{6} \sum_{i=1}^{6} \delta(x-i)$

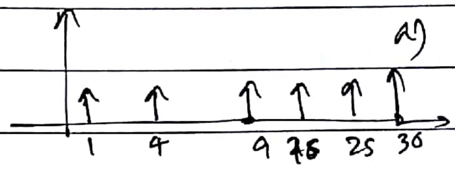



$$f_Y(y) = \frac{1}{6}\Big[\delta(x-1) + \delta(x-4) + \delta(x-9) + \delta(x-16) + \delta(x-25) + \delta(x-36)\Big]$$

b) $f_X(x) = \frac{1}{6} \sum_{i=-2}^{3} \delta(x-i)$

$$f_Y(y) = \frac{2}{6}\delta(x-1) + \frac{2}{6}\delta(x-4) + \frac{1}{6}(\delta(x)) + \frac{1}{6}\delta(x-9)$$




**Q.2.** $Y = 2X + 3$

$F_Y(y) = P(Y \le y)$

$= P\left(X \le \frac{y-3}{2}\right)$

$= F_X\left(\frac{y-3}{2}\right)$

$Y = 2X + 3$

$X = \frac{Y-3}{2}$

given $f_X(x) = \begin{cases} 3x^2 & 0 < x < 1 \\ 0 & \text{o/w} \end{cases}$

$f_Y(y) = f_X\left(\frac{y-3}{2}\right)\frac{dx}{dy}$

$\frac{dx}{dy} = \frac{d}{dy}\left(\frac{y-3}{2}\right)$

$= 3\left(\frac{y-3}{2}\right)^2 \times \frac{1}{2}$

$= \frac{3}{8}(y-3)^2 \rightarrow 3 < y < 5$

$0$ otherwise

**Q.3.** $Y = X^2$ If $f_X(x)$ is PDF of $X$, what is $f_Y(y)$

$F_Y(y) = P(Y \le y)$

since $Y = X^2$

$X = \pm\sqrt{y}$

$= P(-\sqrt{y} \le X \le \sqrt{y})$

$= F_X(\sqrt{y}) - F_X(-\sqrt{y})$

$f_Y(y) = f_X(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} - f_X(-\sqrt{y})\left(\frac{-1}{2\sqrt{y}}\right)$

$= \frac{1}{2\sqrt{y}}\left[f_X(\sqrt{y}) + f_X(-\sqrt{y})\right]$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad f_Y(y) = \frac{1}{\sqrt{2\pi y}} e^{-y/2}$$

Q.4. $Y = \ln X$ $\qquad f_X(x) = \frac{\theta}{x^{\theta+1}}, \quad x>1, \ \theta>0$ — Pareto distribution

$$F_Y(y) = P(Y \le y)$$
$$= P(X \le e^y)$$
$$= F_X(e^y)$$

$Y = \ln X$

$X = e^Y$

$\frac{dx}{dy} = e^y$

$$f_Y(y) = f_X(x) \cdot \frac{dx}{dy}$$
$$= \frac{\theta}{e^{y(\theta+1)}} \cdot e^y = \theta e^{-\theta y}$$

~~x>1~~ $y > 0, \ \theta > 0$

Q.5. $Y = \frac{X-a}{b-a}$ $\qquad f_X(x) = \begin{cases} \frac{1}{b-a} & a<x<b \\ 0 & \text{otherwise} \end{cases}$

$$F_Y(y) = P(Y \le y)$$
$$= P(X \le Y(b-a)+a)$$
$$= F_X(Y(b-a)+a)$$

$Y = \frac{X-a}{b-a}$

$X = Y(b-a)+a$

$\frac{dx}{dy} = (b-a)$

$$f_Y(y) = f_X(Y(b-a)+a)\,(b-a)$$
$$= \frac{1}{b-a} \times (b-a) = 1$$

for $Y \in (0,1)$

$= 0$ otherwise

$X = a$, $Y = 0$ | $X = b$, $Y = 1$

HI PAN KI!!

Scanned by CamScanner

05/02/2019

$$\frac{\partial}{\partial x} F_{XY}(x,y) = \int_{-\infty}^{y} f_{XY}(x,\beta)\, d\beta$$

$$\therefore \quad F_{Y|X}(y|x) = \frac{\int_{-\infty}^{y} f_{XY}(x,\beta)\, d\beta}{f_X(x)}$$

Differentiating w.r.t $y$

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}$$

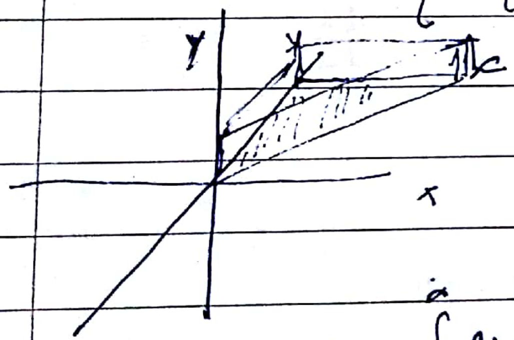
$$\therefore \quad f_{XY}(x,y) = f_X(x) \cdot f_{Y|X}(y|x)$$
$$= f_Y(y) \cdot f_{X|Y}(x|y)$$

* Two R.V. $X$ & $Y$ are said to be statistically independent when the condition - $f_{XY}(x,y) = \underbrace{f_X(x)}_{\text{Marginal PDF}} f_Y(y)$ is true.

and then -

$$\left.\begin{aligned} f_{Y|X}(y|x) &= f_Y(y) \\ f_{X|Y}(x|y) &= f_X(x) \end{aligned}\right\} \text{Conditional Density} = \text{Marginal density}$$

eg. $f_{XY}(x,y) = \begin{cases} \cancel{1/2} & 0 \le x \le y \quad 0 \le y \le 2 \\ 0 & \text{otherwise} \end{cases}$




find (a) $f_{Y|X}(y|1)$

(b) $f_{Y|X}(y|1.5)$

(c) Are $X$ & $Y$ statistically dependent?

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\, dy = 1 - \frac{x}{2}$$

$$f_{Y|X} = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{1}{\cancel{2}} \quad \in [1, 2]$$

$$f_{X|Y} = \frac{f_{XY}(x,y)}{f_X(x)} = \begin{cases} 2 & 1.5 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

$\rightarrow$ $Z = X + Y = g(X+Y)$

$E[Z] = E_X[B] + E_X$ [random]

$x$ $\quad$ $y$

$= E[X+Y] = E(g(X,Y))$

$= \cancel{\iint_{-\infty}^{\infty} g(x,y)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\, f_X(x,y)\, dx\, dy$

$g(x,y)$

$= \iint_{-\infty}^{\infty} x\, f_{XY}(x,y)\, dx\, dy + \iint_{-\infty}^{\infty} y\, f_{XY}(x,y)\, dx\, dy.$

If $g(x,y) = xy$

$\cancel{E[XY] = \int f_{XY}}$

$E[XY] = \iint xy\, f_{XY}(x,y)\, dx\, dy.$

If $X$ & $Y$ are independent $E[XY] = E[X]E[Y]$

$E[X^m Y^m] = E(X)^m E(Y)^m$

08/02/2019

eg $f_{X,Y}(x,y) = \begin{cases} 1/2 & 0 \le x \le y,\ 0 < y \le 2 \\ 0 & \text{otherwise} \end{cases}$

$f_{Y|X}(y|x) = \dfrac{f_{X,Y}(x,y)}{f_X(x)}$ $\qquad f_{X|Y}$ ?

→ Functions of two random variables

$Z = g(X,Y)$ $\qquad$ given $f_{XY}(x,y)$, Find $f_{ZW}(z,w)$

$W = h(X,Y)$ $\qquad\qquad f_Z(z),\ f_W(w)$

$\{Z \le z,\ W \le w\} = \{(X,Y) \in \underbrace{D_{zw}}_{\text{Region}}\}$

$f_{ZW}\,dz\,dw = P\{z < Z \le z+dz,\ w < W \le w+dw\}$

We have $g(x,y) = z$, $\quad h(x,y) = w$

For a given $z, w$ denote the $n$ real roots

$$f_{Z,W}(z,w) = \frac{f_{XY}(x_1,y_1)}{|J(x_1,y_1)|} + \frac{f_{X,Y}(x_2,y_2)}{|J(x_2,y_2)|} + \dots + \frac{f_{XY}(x_n,y_n)}{|J(x_n,y_n)|}$$

where $|J(x,y)| = \left|J\left(\frac{z,w}{x,y}\right)\right| = \left|\det\begin{bmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{bmatrix}\right|$

Jacobian

it is the determinant of the appropriate partial derivatives.

eg. Let $Z = X+Y$, $W = X-Y$

find $f_{ZW}(z,w)$ given $\quad f_X(x) = 1/2,\ |x| < 1$

$\qquad\qquad\qquad\qquad\qquad f_Y(y) = 1/2,\ |y| < 1$

* $f_{XY}(x,y) = f_X(x)\, f_{Y|X}(y|x) = f_Y(y)\, f_{X|Y}(x|y)$

If $X$ & $Y$ are independent then $f_{Y|X}(y|x) = f_Y(y)$

$\qquad\qquad\qquad\qquad$ & $f_{X|Y}(x|y) = f_X(x)$

In this problem, $X$ & $Y$ are statistically independent.

08/02/2019

$$f_{ZW}(z,w) = \frac{f_{XY}(x,y)}{\left|J\left(\frac{z,w}{x,y}\right)\right|}$$

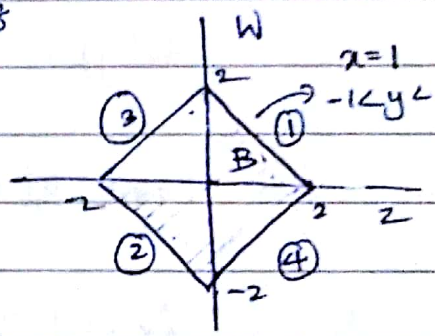
$z = X + Y$ $\qquad$ $x = (z+w)/2$

$w = X - Y$ $\qquad$ $y = (z-w)/2$

$$\left|J\left(\frac{z,w}{x,y}\right)\right| = \left|\det\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right| = |-2| = \underline{\underline{2}}$$

$$f_{ZW}(z,w) = \frac{f_X(x)\, f_Y(y)}{2} = \frac{\cancel{1/2 \times 1/2}}{\cancel{2}} = \frac{1}{8}$$

$$= \frac{f_X((z+w)/2)\, f_Y((z-w)/2)}{2} = \frac{1}{8}$$




for Range : $|x| < 1$, $|y| < 1$

i.e. $-1 < x < 1$, $-1 < y < 1$

$z \in (-2, 2)$ $\qquad$ $w \in (-2, 2)$

$z, w \in \underline{\underline{B}}$

fix $x = 1$ $\qquad$ $-1 < y < 1$

$x = \frac{z+w}{2}$ $\qquad$ $\underline{\underline{w = 2 - z}}$ ①

fix $x = -1$ $\rightarrow$ $w = -z - 2$ ② $\rightarrow$ $x = -1$ $\quad$ $\underline{-1 < y < 1}$

fix $y = -1$

$y = (z-w)/2$ $\rightarrow$ $w = z + 2$ ③ $\qquad$ $y = -1$

$-1 < x < 1$

fix $y = 1$ $\rightarrow$ $w = z - 2$ ④ $\rightarrow$ $y = 1$

$-1 < x < 1$

$\rightarrow$ find $f_Z(z)$ & $f_W(w)$

$$f_Z(z) = \int_{z-2}^{2-z} f_{ZW}(z,w)\, dw \rightarrow z \geq 0$$

$$= \int_{-z-2}^{z+2} f_{ZW}(z,w)\, dw \rightarrow z < 0$$

11/02/2019

Tutorial - 4

Q.1. $f_{X,Y}(x,y) = \begin{cases} 4xy, & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & \text{elsewhere.} \end{cases}$

$$f_X(x) = \int_0^1 4xy\, dy = 4x \int_0^1 y\, dy = 4x \left[\frac{y^2}{2}\right]_0^1 = \underline{\underline{2x}}.$$

$$f_Y(y) = \int_0^1 4xy\, dx = 4y \int_0^1 x\, dx = 2y.$$

$$f_X(x) f_Y(y) = 4xy = f_{XY}(x,y)$$

Q.2. $f_{X,Y}(x,y) = \begin{cases} 1/2, & 0 \le x \le y,\ 0 \le y \le 2 \\ 0, & \text{elsewhere.} \end{cases}$

$$f_X(x) = \int_0^2 \frac{1}{2}\, dy = 1$$

$$f_Y(y) = \int_0^y \frac{1}{2}\, dx = \frac{y}{2}$$

Not independent

$$f_{Y|X}(y|x) = 1/2 \qquad f_{X|Y}(x|y) = 1/y$$

Q.3. $f_{X,Y}(x,y) = \begin{cases} c e^{-x} e^{-y}, & 0 \le x \le y \le \infty \\ 0, & \text{elsewhere.} \end{cases}$

$$\cancel{f_{X,Y}(x,y) = \int_x^\infty \int_0^\infty c e^{-x} e^{-y}\, dx\, dy}$$

$$\cancel{= \int_x^\infty c e^{-y} \left[-e^{-x}\right]_0^\infty dy = c \int_x^\infty e^{-y}\, dy}$$

$$\cancel{= c\left[-e^{-y}\right]_x^\infty = c e^{-x} = 1 \quad \therefore c = e^{-x}}$$

$$\cancel{f_X(x) = \int_0^\infty e^{-y}\, dy = 1}$$

$$\cancel{f_Y(y) = \int_x^\infty e^{-y}\, dx = -x e^{-y}}$$

$$f_{X,Y}(x,y) = \int_0^\infty \left[\int_x^\infty c e^{-x} e^{-y}\, dy\right] dx = \int_0^\infty c e^{-x} \left[-e^{-y}\right]_x^\infty dx$$

$c = 2$     __/__/____

$$= c\int_0^\infty e^{-x} e^{-x}\, dx = c\int_0^\infty e^{-2x}\, dx = -\frac{c}{2}\left[e^{-2x}\right]_0^\infty = \frac{+c}{2}$$

$$f_X(x) = \int_x^\infty 2e^{-x}e^{-y}\, dy = 2e^{-x}\left[-e^{-y}\right]_x^\infty = \cancel{2e^{-x}}\ 2e^{-2x}$$

$$f_Y(y) = \int_0^\infty 2e^{-x}e^{-y}\, dx = 2e^{-y}\left[-e^{-x}\right]_0^\infty = 2e^{-y}$$

Q. 5.    $Y = W$     $X = Z - W$     $Z = X + Y$    $W = Y$

$$\frac{dw}{dy} = 1 \qquad \frac{dw}{dx} = 0 \qquad \frac{dz}{dx} = 1 \qquad \frac{dz}{dy} = 1$$

$$|J(x,y)| = \left|\det\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right| = \underline{\underline{1}}.$$

$$f_{Z,W}(z,w) = f_{X,Y}(x,y)$$

Scanned by CamScanner

___/___/___

Q 6. $Y_1 = X_1^2 - X_2^2$

$Y_2 = X_1^2 + X_2^2$

$Y_3 = X_3$

$X_1 = \pm\sqrt{\frac{Y_1 + Y_2}{2}}$

$X_2 = \pm\sqrt{\frac{Y_2 - Y_1}{2}}$

$X_3 = Y_3$

$y_2 > y_1$

$y_2 - y_1 > 0$

$y_1 > 0$

$$J(x_1, x_2, x_3) = \left| \det \begin{bmatrix} \frac{dy_3}{dx_1} & \frac{dy_3}{dx_2} & \frac{dy_3}{dx_3} \\ \frac{dy_2}{dx_1} & \frac{dy_2}{dx_2} & \frac{dy_2}{dx_3} \\ \frac{dy_1}{dx_1} & \frac{dy_1}{dx_2} & \frac{dy_1}{dx_3} \end{bmatrix} \right| = \begin{bmatrix} 0 & 0 & 1 \\ 2x_1 & +2x_2 & 0 \\ 2x_1 & -2x_2 & 0 \end{bmatrix}$$

$$= \left| -4x_1x_2 - 4x_1x_2 \right| = 8x_1x_2$$

$x_1^{(1)} = \sqrt{y_1+y_2/2}$

$x_2^{(1)} = \sqrt{y_2-y_1/2}$

$x_3^{(1)} = y_3$

$x_1^{(2)} = \sqrt{y_1+y_2/2}$

$x_2^{(2)} = -\sqrt{y_2-y_1/2}$

$x_3^{(2)} = y_3$

$x_1^{(3)} = -\sqrt{y_1+y_2/2}$

$x_2^{(3)} = \sqrt{y_2-y_1/2}$

$x_3^{(3)} = y_3$

$x_1^{(4)} = -\sqrt{y_1+y_2/2}$

$x_2^{(4)} = -\sqrt{y_2-y_1/2}$

$x_3^{(4)} = y_3$

$$J_1 = \left| 8\sqrt{\frac{y_1+y_2}{2}}\sqrt{\frac{y_2-y_1}{2}} \right| = 4\sqrt{y_2 - y_1}$$

$$= J_2 = J_3 = J_4$$

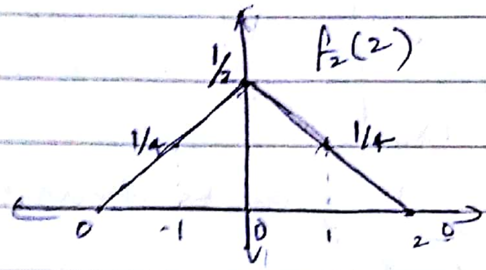
11/02/19

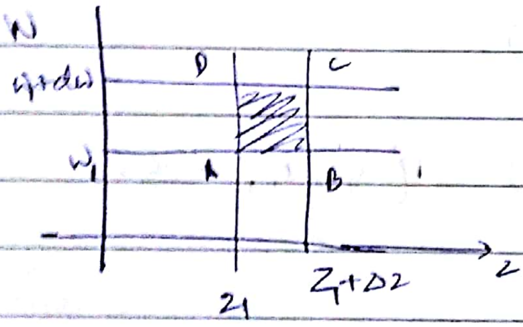
cont.

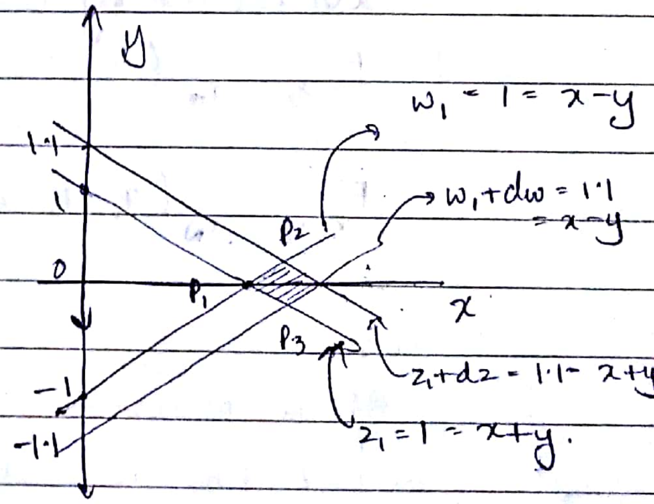
$$f_z(z) = \int_{z-2}^{2-z} f_{z,w}(z, w)\, dw \qquad z \geq 0$$

$$= \int_{z-2}^{2-z} \frac{1}{8}\, dw = \frac{1}{8}\left[2-z-(z-2)\right] = \left(\frac{1}{2} - \frac{1}{4}z\right)$$

$$\& \int_{-2-z}^{2+z} f_{z,w}(z, w)\, dw = \frac{1}{8}\int_{-2-z}^{2+z} dw = \frac{1}{8}\left[2+z-(-2-z)\right] = \frac{1}{2} + \frac{z}{4} \qquad z < 0$$




$z = x+y$
$w = x-y.$

take $(z_1, w_1) = (1, 1)$

$\left.\begin{array}{l} x+y = 1 \\ x-y = 1 \end{array}\right\} \rightarrow$ intersect at $(1, 0)$

$(z_1+dz,\ w_1+dw) = (1.1, 1.1)$

$$P_1 = \begin{pmatrix} x \\ y \end{pmatrix} \qquad P_2 = \begin{pmatrix} x + \frac{dx}{dz}.dz \\ y + \frac{dy}{dz}.dz \end{pmatrix} \qquad P_3 = \begin{pmatrix} x + \frac{dx}{dw}\, dw \\ y + \frac{dy}{dw}.dw \end{pmatrix}$$

Difference Vectors are -

$$V_1 = \begin{pmatrix} \frac{dx}{dz}\, dz \\ \frac{dy}{dz}\, dz \end{pmatrix} \qquad V_2 = \begin{pmatrix} \frac{dx}{dw}\, dw \\ \frac{dy}{dw}\, dw \end{pmatrix}$$

12/02/2019

Area of required $\|^{gm}$ $= |V_1 \times V_2| = \begin{bmatrix} i & j & k \\ \frac{dx}{dz}dz & \frac{dy}{dz}.dz & 0 \\ \frac{dx}{dw}dw & \frac{dy}{dw}.dw & 0 \end{bmatrix}$

$A = \left| \frac{dx}{dz}(dz) . \frac{dy}{dw}(dw) - \frac{dx}{dw}(dw) \frac{dy}{dz}(dz) \right|$

$f_{ZW}(z, w) \cancel{dz\, dw} = f_{XY}(x, y) \left| \frac{dx}{dz} . \frac{dy}{dw} - \frac{dx}{dw} \frac{dy}{dz} \right|$

$\Rightarrow$ Random Vector $X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$

If we have two R.V.s $X_1$ & $X_2$, they can be characterised by using joint PDF $f_{X_1 X_2}(x_1, x_2)$

Here we have N number of R.V.s, they can be characterised by using Multivariate PDF

$f_{X_1 X_2, \dots X_N}(x_1, x_2, \dots x_N)$

$F_{X_1 X_2 \dots X_N}(x_1, x_2, x_3, \dots x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots X_N \leq x_n)$

$Y = X + n$

- Covariance

If the outcome of one R.V. is a given value, what can we say about the outcome of another R.V.?

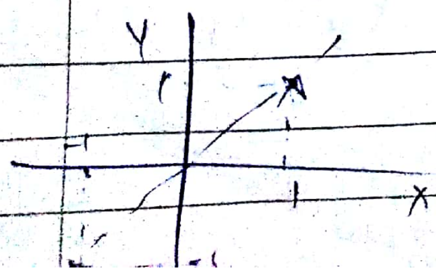
$Var(X) = E(X^2) - (E(X))^2 = E((X - m_X)^2)$

Covariance b/n X & Y is given by,

$Cov(X, Y) = E[(X - m_x)(Y - m_y)]$

eg. $P(X=1, Y=1) = P(X=-1, Y=-1) = 1/2$

$X = Y$ $\qquad E(XY) = ? = 1 \times 1 \times \frac{1}{2} + (-1)(-1) \times \frac{1}{2} = 1$

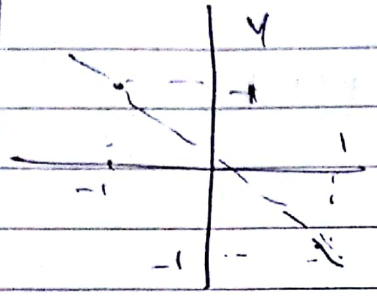


Since $E(XY) = \sum_i x_i y_i P(X = x_i, Y = y_i)$

$E(X) = \sum_i x_i P(X = x_i)$ $\qquad E(X) = 0$

→ $P(X=1, Y=-1) = P(X=-1, Y=1) = 1/2$

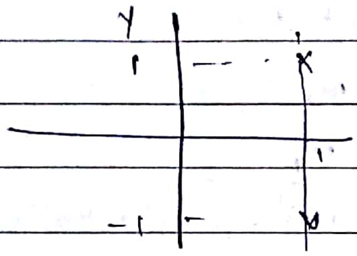
$X = -Y$



$E(XY) = -1$

$E(X) = 0$

→ $P(X=1, Y=+1) = P(X=1, Y=-1) = 1/2$



$E(XY) = 0$

→ $P(X=0, Y=0) = P(X=2, Y=2) = 1/2$

$E(X) = 1$    Still $X = Y$    $E(XY) = 2$

$E(Y) = 1$

Subtract $m_x$ from $X$ and $E((X-m_x)(Y-m_y)) = 1$

same as case 1, which is a better estimate of cov.

* Cov $\in (-\infty, \infty)$

→ Another measure which takes values between $-1$ & $1$ can also be used to measure how two R.V.'s $X$ & $Y$ co-vary

That measure is called correlation coefficient $\rho_{XY}$

$$\rho_{XY} = \frac{E[(X-m_x)(Y-m_y)]}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

$$-1 \le \rho_{XY} \le +1$$

Proof: $U = X - m_X$    consider $X$ & $Y$

$E(X) = m_x$

$E(U) = E(X - m_x) = E(X) - m_x = m_x - m_x = 0$

$U$ is the zero-mean R.V.

__/__/__

Consider two R.V.s $X$ & $Y$ $\rightarrow \geq 0$ $f(\alpha)$ is always positive.

$$E(X+\alpha Y)^2 = f(\alpha)$$

$$f(\alpha) = E[X^2 + \alpha^2 Y^2 + 2\alpha XY]$$

$$= E(Y^2)\alpha^2 + 2\alpha E(XY) + E(X^2)$$

$$= A\alpha^2 + b\alpha + C$$

Discriminant $\sqrt{b^2-4ac} < 0$

$$b^2 \leq 4ac$$

$$\cancel{4}(E(XY))^2 \leq \cancel{4} E(Y^2)E(X^2)$$

$$|E(XY)| \leq \sqrt{E(X^2)}\sqrt{E(Y^2)}$$

Replace $X \rightarrow U$ where $U = X - m_x$

and $Y \rightarrow W$ where $W = Y - m_y$

$$|E(UW)| \leq \sqrt{E((Y-m_y)^2)}\sqrt{E((X-m_x)^2)}$$

$$|E((X-m_x)(Y-m_y))| \leq \sigma_y \sigma_x$$

$$|\rho_{XY}| \leq 1$$

⇒ Chebyshev's Inequality

Is it possible to obtain the probability bound, based on mean and variance, without using PDF?

$$P(|X-m_x| > \gamma) \leq B \qquad B \in [0,1]$$

$f_X(x) \Rightarrow$ continuous r.v. of $X$ pdf

18/02/2019

Tutorial - 5

Q.1. Table

| $X \backslash Y$ | $y_1$ | $y_2$ | . | $f_X(x)$ ↓ |
|---|---|---|---|---|
| $x_1$ | $f_{X,Y}(x_1,y_1)$ | | | |
| $x_2$ | | | | |
| . | | | | |
| . | | | | |
| $f_Y(y) \rightarrow$ | | | | |

HHH
HHT
HTH
HTT
TTT
TTH
THT
THH.

| $X \backslash Y$ | 0 | 1 | 2 | $f_X(x)$ |
|---|---|---|---|---|
| 0 | 1/8 | 1/8 | 0 | 2/8 |
| 1 | 1/8 | 2/8 | 1/8 | 4/8 |
| 2 | 0 | 1/8 | 1/8 | 2/8 |
| $f_Y(y)$ | 2/8 | 4/8 | 2/8 | =1 |

$E(X) = 1$

$E(Y) = 1$

$$Cov(X,Y) = E[(X-m_X)(Y-m_Y)] = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

$$E[X^2] = \frac{4}{8} + \frac{8}{8} = 1.5$$

$$E[Y^2] = 1.5.$$

$$Var(X) = E(X^2) - (E(X))^2 = 1/2 = Var(Y)$$

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var Y}} = \frac{1/4}{\sqrt{1/2}\sqrt{1/2}} = \frac{1/4}{1/2} = 1/2$$

$$E(X,Y) = \sum x_i y\, P_{x,y}(x,y)$$

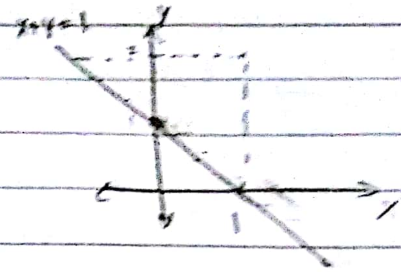
Q.2. $f_{X,Y}(x,y) = \begin{cases} cx^2 + xy/3 & 0 \le x \le 1,\ 0 \le y \le 2 \\ 0 & \text{otherwise.} \end{cases}$

(a) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y) = 1$ $\qquad \int_0^2\int_0^1 \left(cx^2 + \frac{xy}{3}\right) dx\, dy$

$$\Rightarrow \int_0^2 \left[c\frac{x^3}{3} + \frac{y}{3}\left(\frac{x^2}{2}\right)\right]_0^1 dy = \int_0^2 \left(\frac{c}{3} + \frac{y}{6}\right) dy = \left(\frac{cy}{3} + \frac{y^2}{12}\right)_0^2$$

$$\Rightarrow \frac{2c}{3} + \frac{\cancel{4}^{1}}{\cancel{12}_{3}} = 1 \qquad \frac{2c}{3} = \frac{2}{3} \qquad \boxed{c = 1}$$

(b) $P[X+Y \geq 1]$ $\qquad Y = 1-X$




$$\int_0^1 \int_{1-x}^{2} \left(x^2 + \frac{xy}{3}\right) dy\, dx$$

$$= \cancel{\int \left[x^2 + \frac{x(1+x)^2}{6}\right] dx}$$

$$= \int_0^1 \left[x^2 y + \frac{xy^2}{6}\right]_{1-x}^{2} dx = \int_0^1 \left(x^2(2-(1-x)) + \frac{x}{6}\left[2^2 - (1-x)^2\right]\right) dx$$

$$= \int_0^1 \left(x^2(1+x) + \frac{x}{6}(4-1-x^2+2x)\right] dx = \int_0^1 \left(\frac{5x^3}{6} + \frac{4}{3}x^2 + \frac{x}{2}\right) dx$$

$$= \left[\frac{5x^4}{6\times 4} + \frac{4x^3}{9} + \frac{x^2}{4}\right]_0^1 = \frac{5}{24} + \frac{4}{9} + \frac{1}{4} = \frac{65}{72}$$

(c) $$f_X(x) = \int_0^2 \left(x^2 + \frac{xy}{3}\right) dy = \left[x^2 y + \frac{x}{3}\left(\frac{y^2}{2}\right)\right]_0^2 = 2x^2 + \frac{2x}{3}$$

$0 < x < 1$

$$f_Y(y) = \int_0^1 \left(x^2 + \frac{xy}{3}\right) dx = \left[\frac{x^3}{3} + \frac{y}{3}\left(\frac{x^2}{2}\right)\right]_0^1 = \frac{1}{3} + \frac{y}{6}$$

$0 < y < 2$

(d) $$f_X(x)\, f_Y(y) = \left(2x^2 + \frac{2x}{3}\right)\left(\frac{1}{3} + \frac{y}{6}\right) = \frac{2x^2}{3} + \frac{2x^2 y}{6} + \frac{2x}{9} + \frac{2xy}{18}$$

$\neq f_{XY}(x,y)$ $\quad \therefore$ Not independent

(e) $Cov(X,Y)$

Q.3. $[A - \lambda I] = 0$ → for eigen vector

$AX = \lambda X$

→ find $\lambda$s

$|A - \lambda I| = 0$

→ $Var(X) = \sigma_x^2 = \int_{-\infty}^{\infty} (x - m_x)^2 f_x(x)\, dx$

$$= \int_{|(x-m_x)| > r} (x-m_x)^2 f_x(x)\, dx$$

$$+ \int_{|(x-m_x)| \le r} (x-m_x)^2 f_x(x)\, dx.$$

$$Var(X) \geq \int_{|(x-m_x)| > r} (x-m_x)^2 f_x(x)\, dx.$$

$$Var(X) \geq r^2 \int_{|x-m_x| > r} f_x(x)\, dx.$$

$$\geq r^2 P(|(X-m_x)| > r)$$

$$\therefore \boxed{P(|X-m_x| > r) \le \frac{Var(X)}{r^2}}$$

→ $r = k\sigma_x$

$$P(|(X-m_x)| > k\sigma_x) \le \frac{\cancel{\sigma_x^2}}{k^2 \cancel{\sigma_x^2}}$$

$$\le \frac{1}{k^2}$$

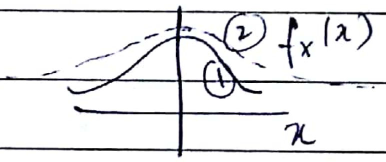
Consider $\sigma_x = 1$, $k = 3$

$$P(|X-m_x|) > 3) \le \frac{1}{9}$$

Used in DIP → gaussian filter

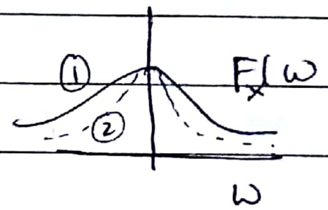
gaussian $\xrightarrow{\mathcal{F}}$ gaussian

Shape remains same




$[h(t)$ or $h[n]]$

$\downarrow \mathcal{F}$




$\{H(\omega)$ or $H(\Omega)\}$

curve ① will allow more no. of frequencies to pass than curve ②. ① → HPF, ② → LPF } used as impulse response

Shape changes with variance

Variance ↑ Freq ↓ LPF

→ when LPF applied to image it will get blurred.

___/___/______

Continue with covariance.

$$Z = X + Y$$

$Var(X+Y) = Var(Z)$

We know that $Var(X) = \sigma_x^2$

$= E[X^2] - (E[X])^2$ $\rightarrow (E[X] + E[X])^2 = (m_x + m_x)^2$

$\therefore$ $Var(X+X) = E[(X+X)^2] - (E[X+Y])^2$

$= E[X^2 + X^2 + 2XX - \{m_x^2 + m_y^2 + 2m_x m_y\}$

$= Var(X) + Var(X) + 2cov(X,Y)$

If the two r.v.'s X & Y are un-correlated

$Var(X+Y) = Var(X) + Var(Y)$

$Var(X-Y) = Var(X) + Var(Y) - 2cov(X,Y)$

$\rightarrow$ $Var(X+Y) = Var(X) + Var(Y) + cov(X,Y) + cov(Y,X)$

$Var(X_1+X_2) = Var(X_1) + Var(X_2) + cov(X_1,X_2) + cov(X_2,X_1)$

$Var(X_1+X_2+X_3) = Var(X_1) + Var(X_2) + Var(X_3) + cov(X_1,X_2)$
$+ cov(X_2,X_3) + cov(X_1,X_3) + cov(X_2,X_1) + cov(X_3,X_2)$
$+ cov(X_3,X_1)$

$\rightarrow$ Consider

$$Var(X_1+X_2) = \underset{1\times 2}{[1\ 1]} \underset{2\times 2}{\begin{bmatrix} Var(X_1) & Cov(X_1,X_2) \\ Cov(X_2,X_1) & Var(X_2) \end{bmatrix}} \underset{2\times 1}{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}$$

19/02/2019.

In general if there are $N$ r.v.s then the covariance matrix $C_X$ can be written as

$$C_X = \begin{bmatrix} Var(X_1) & Cov(X_1 X_2) & \cdots & Cov(X_1 X_N) \\ Var(X_2, X_1) & Cov(X_2) & \cdots & Cov(X_2, X_N) \\ \vdots & & & \vdots \\ Var(X_N, X_1) & \cdots & \cdots & Var(X_N) \end{bmatrix}_{N \times N}$$

(captures the $2^{nd}$ order statistics of the R.V.s).

→ Properties of $C_X$

1/. If all random variables are uncorrelated, $C_X$ is a diagonal matrix

2/. Since $Cov(X_1, X_2) = Cov(X_2, X_1)$ it is a real symmetric matrix

$C_X = C_X^T$

3/. Covariance matrix is positive semi-definite.

If a matrix ~~is~~ $A_{m \times n}$ is semi-definite, $x^T A x \geq 0$

(If it is positive definite, then $x^T A x > 0$)

$x^T A x$ is a scalar quantity. $x \in R^n$

Consider $U_1 = X_1 - m_{X_1}$ & $U_2 = X_2 - m_{X_2}$ (Mean subtracted R.V.s)

Then variance, $Var(a_1 U_1 + a_2 U_2)$

$$= E(a_1 U_1 + a_2 U_2)^2 - \underbrace{\{E(a_1 U_1 + a_2 U_2)\}^2}_{\to 0} \quad (\because U_1 \text{ \& } U_2 \text{ have 0 mean})$$

$$= E(a_1^2 U_1^2 + a_2^2 U_2^2 + 2 a_1 a_2 U_1 U_2) \geq 0$$

$\{\because Var(x) \geq 0$ cannot be negative$\}$

$$= a_1^2 E(X_1 - m_{X_1})^2 + a_2^2 E(X_2 - m_{X_2})^2 + 2 a_1 a_2 E[(X_1 - m_{X_1})(X_2 - m_{X_2})]$$

$$= \underbrace{(a_1 \;\; a_2)}_{x} \underbrace{\begin{bmatrix} E(X_1 - m_{X_1})^2 & E[(X_1 - m_{X_1})(X_2 - m_{X_2})] \\ E[(X_2 - m_{X_2})(X_1 - m_{X_1})] & E(X_2 - m_{X_2})^2 \end{bmatrix}}_{C_X} \underbrace{\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}}_{x^T} \geq 0$$

Hence, $C_X$ is positive semi-definite.

4/. From this we can say that the

$Det(C_X)$

19/02/2019.

$v^T C_x v \geq 0$

$v^T \lambda v \geq 0$

$\lambda \|v\|^2 \geq 0$

$\{Av = \lambda v\}$ → For a square matrix
↳ $C_x v = \lambda v$
$\{v^T v = \|v\|^2\}$

- $(A - \lambda I)x = 0 \Leftrightarrow$ homogenous equation → trivial solution is that $x = 0$. $\{|B| = 0$ → B is linearly dependent$\}$. We don't want $x \neq 0$, so we force $|A - \lambda I| = 0$

* Since $C_x$ is symmetric real matrix, what can we say about the eigen values & eigen vectors of $C_x$?

i) Eigen values are all real

ii) Eigen vectors of distinct eigen values are all orthogonal
↓
dot product = 0.

* Principle component analysis (PCA) or KL Transform.

Consider $Y = AX$.

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Find a transformation such that Y values are highly de-correlated - we pack all X values into some q. Y values.

A depends on X. PCA is a data-dependent transform.

$Ax = \lambda x$ — ① $x \in C^n$ $\lambda \in C$

$Ax^* = \lambda^* x^*$ — ②

$\lambda = \lambda^*$ indicates that $\lambda \in \mathbb{R}$

Pre-multiply ① with $x^{*T}$

$x^{*T} A x = x^{*T} \lambda x = \lambda x^{*T} x$ ($\because \lambda$ is a constant)

can be written as

$(Ax^*)^T x = \lambda x^{*T} x$ ($\because (Ax^*)^T = x^{*T} A^T = x^{*T} A$, $\because A^T = A$)

also $(\lambda^* x^*)^T x = \lambda x^{*T} x$ ($\lambda$ is a constant)

$\therefore \lambda^* x^{*T} x = \lambda x^{*T} x$

Scanned by CamScanner

21/02/2019

Consider $A x_1 = \lambda_1 x_1$ — ①

$A x_2 = \lambda_2 x_2$ — ②

Pre-multiplying ① with $x_2^T$

$$x_2^T A x_1 = x_2^T \lambda_1 x_1$$
$$(A x_2)^T x_1 = x_2^T \lambda_1 x_1$$
$$\lambda_2 x_2^T x_1 = \lambda_1 x_2^T x_1$$
$$(\lambda_2 - \lambda_1) x_2^T x_1 = 0$$

if $\lambda_2 \neq \lambda_1$, $x_2^T x_1 = 0$

$\therefore$ $x_2 \cdot x_1 = 0 \rightarrow$ orthogonal.

Q HW — If repeated eigen values are present, one can always find set of orthonormal (orthogonal + unit norm) eigen vectors.

Q. — If the vectors are orthogonal, they are indeed linearly independent. Is the converse true?

Consider $x_1, \ldots x_n$ vectors are all orthogonal

Constants $c_1, \ldots c_n$ such that,

$$c_1 x_1 + c_2 x_2 + c_3 x_3 + \cdots + c_n x_n = 0$$

If they are all independent then $c_1 \ldots c_n = 0$.

$$(c_1 x_1 + c_2 x_2 + \cdots c_n x_n) \cdot x_1 = 0$$

$$c_1 \|x_1\|^2 = 0$$ (rest all are zero for $x_i \cdot x_1 = 0$ orthogonal ($i \neq 1$))

$c_1 = 0$.

Similarly $c_2 \ldots c_n$ are all zeros.

- PCA is also a linear transform like DFT, DCT etc.

$Y = AX$

$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix}$ transformed random vector.

$X = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix}$ input random vector.

A is the transformation matrix.

Objective : get Y as uncorrelated

i.e. $C_Y$ has to be diagonal

21/02/2019

$C_Y$ is $N \times N$ matrix and choose data X such that $C_X$ is not diagonal

Aim: Transform correlated data into decorrelated data.

$$C_X = E \begin{bmatrix} (x_1 - m_{x_1}) \\ \vdots \\ (x_N - m_{x_N}) \end{bmatrix}_{N \times 1} \begin{bmatrix} (x_1 - m_{x_1}) & \cdots & (x_N - m_{x_N}) \end{bmatrix}_{1 \times N}$$

↓ Covariance matrix of $X_N$ $N \times N$.

$$\underset{N \times N}{C_Y} = E\left[ \underset{N \times 1}{(\underline{Y} - m_Y)} \underset{1 \times N}{(\underline{Y} - m_Y)^T} \right] = E\left[ (A\underline{X} - Am_{\underline{x}})(A\underline{X} - Am_x)^T \right]$$

$$m_X = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_N) \end{bmatrix} \qquad m_Y = \begin{bmatrix} E(Y_1) \\ E(Y_2) \\ \vdots \\ E(Y_N) \end{bmatrix} = E(Y) = E(AX) = A\,E(X)$$

$$\begin{aligned} C_Y &= E\left[ (A\underline{X} - Am_{\underline{x}})(A\underline{x} - Am_{\underline{x}})^T \right] \\ &= E\left[ A(\underline{X} - m_{\underline{x}})(\underline{x} - m_{\underline{x}})^T A^T \right] \\ &= \left[ A\, C_{\underline{X}}\, A^T \right] \end{aligned}$$

$C_X = U \Sigma U^{-1}$ $\quad \Sigma \rightarrow$ diagonal $\lambda$'s of A

Possible because eigen vectors matrix U is invertible because eigen vectors are orthogonal.

$U^{-1} = U^T$

$$\begin{aligned} C_Y &= A U \Sigma U^{-1} A^T \\ &= A U \Sigma U^T A^T \\ &= A U \Sigma (AU)^T \end{aligned}$$

$\therefore \boxed{A}$ ↓ eigen vectors of covariance matrix of x

Let $A = U^T$

$$\begin{aligned} \therefore C_Y &= U^T U \Sigma (U^T U)^T \\ &= \Sigma \qquad (\because U^T = U^{-1}) \end{aligned}$$

↳ Diagonal matrix ⟹ Decorrelated.

23/02/2019

Tutorial - 6

Q1. $\mu_x = 64.5$ $\quad \sigma^2 = 144$

$P(44 < X < 85) = ?$

$P(|X - m_x| > \gamma) = \frac{\sigma^2}{\gamma^2}$

$P(44 < X < 85) = P(44 - m_x < X - m_x < 85)$

$= P(44 - 64.5 < X - m_x < 85 - 64.5)$

$= P(-20.5 < X - m_x < 20.5) = P(|X - m_x| < 20.5)$

$= 1 - \frac{\sigma^2}{\gamma^2} = 1 - \frac{144}{(20.5)^2} = \underline{0.658}$

$P(36 < X < 93) = P(36 - 64.5 < X - m_x < 93 - 64.5)$

$= P(-28.5 < X - m_x < 28.5) = P(|X - m_x| < 28.5)$

$= 1 - \frac{\sigma^2}{\gamma^2} = 1 - \frac{144}{(28.5)^2} = \underline{0.8228}$

Q2. $X = [X1, X2]^T$ $\quad \sigma^2 = 1$ $\quad \underbrace{Cov(X1, X2) = 0}$

$Y = [Y1, Y2]^T$, $\quad Y = AX$

$C_x = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$C_y = A C_x A^T$

$C_y = A I A^T = A A^T$

$\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} = AA^T$

$C_y = U \Sigma U^T$

$U \Sigma^{1/2} \Sigma^{1/2} U^T = A A^T$

$\therefore A = U \Sigma^{1/2}$

$\downarrow$

$[A - \lambda I] = 0$

$\begin{bmatrix} 1-\lambda & 0.5 \\ 0.5 & 1-\lambda \end{bmatrix} = 0$

$(1-\lambda)^2 - 0.25 = 0$

$\lambda^2 - 2\lambda + 0.75 = 0$

$\lambda^2 - 0.75\lambda - 1.25\lambda + 0.75 = 0$

$\lambda(\lambda - 0.5) - 1.5(\lambda - 0.5) = 0$

$(\lambda - 0.5)(\lambda - 1.5) = 0$

$\therefore \lambda_1 = 0.5$

$\lambda_2 = 1.5$

$\lambda_1 = 1.5$

$$\begin{bmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -0.5 & 0.5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$x + y = 0$

$\therefore x = -1$

$y = 1$

(1) $(C_x - \lambda I) = 0 \rightarrow$ To find eigen values

(2) Find eigen vectors per each eigen value using $C_x \lambda = \lambda X$

$$\left[ C_x - \lambda I \mid X \right] = 0$$

Bring in diagonal form & get equation & find $x, y$...

(3) ~~the~~ Basis from eigen vectors.

$$A = U \Sigma^{1/2} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2.5 & 0 \\ 0 & 0.5 \end{bmatrix}^{1/2}$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1.22 & 0 \\ 0 & 0.707 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1.22 & -0.707 \\ 1.22 & 0.707 \end{bmatrix}$$

Q.3. $f_{X_1 X_2}(x_1, x_2) = 2 \quad 0 < x_1 < x_2 < 1$

$= 0$ otherwise.

$$Y_1 = \underbrace{\frac{X_1}{X_2}}_{\in (0,1)} \qquad Y_2 = \underbrace{X_2}_{\in (0,1)} \qquad f_{Y_1 Y_2}(y_1, y_2) = ?$$

$$f_{Y_1 Y_2}(y_1, y_2) = \frac{f_{X_1 X_2}(x_1, x_2)}{|J|} \qquad J = D_x = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{bmatrix}$$

$\frac{u'v - v'u}{v^2}$

__/__/____

$$D_x = \begin{bmatrix} 1/x_2 & -x_1/x_2 \\ 0 & 1 \end{bmatrix} = \frac{1}{x_2} = \frac{1}{y_2}$$

$\therefore f_{Y_1Y_2}(y_1 y_2) = 2y_2$

$x_1 = y_1 x_2 = y_1 y_2$

$x_2 = y_2$

Q.4: $f_{X_1X_2}(x_1, x_2) = \begin{cases} e^{-(x_1+x_2)} & , \; x_1 \geq 0 \; \& \; x_2 \geq 0 \\ 0 & , \; \text{elsewhere} \end{cases}$

$Y_1 = X_1 + X_2$

$Y_2 = \frac{X_1}{X_1 + X_2} = a.$

$f_{Y_1Y_2}(y_1, y_2) = \frac{f_{X_1X_2}(x_1, x_2)}{|J|}$

$Y_2 X_1 + Y_2 X_2 = X_1$

$X_1 =$

$X_1 = Y_1 - X_2$

$X_1 = -X_2 Y_2$

$$\frac{dy_1}{dx_1} = 1$$

$$\frac{dy_2}{dx_1} = \frac{(1)(X_1+X_2) - (1)(X_1)}{(X_1+X_2)^2} = \frac{X_2}{(X_1+X_2)^2}$$

$$\frac{dy_1}{dx_2} = 1$$

$$\frac{dy_2}{dx_2} = \frac{0 - (1)(X_1)}{(X_1+X_2)^2} = \frac{-X_1}{(X_1+X_2)^2}$$

$$J = \begin{vmatrix} 1 & \frac{X_2}{(X_1+X_2)^2} \\ 1 & \frac{-X_1}{(X_1+X_2)^2} \end{vmatrix} = \frac{-X_1}{(X_1+X_2)^2} - \frac{X_2}{(X_1+X_2)^2} = \frac{-1}{X_1+X_2} = \frac{-1}{Y_1}$$

$$f_{Y_1Y_2}(y_1, y_2) = \frac{e^{-(y_1)}}{(1/y_1)} = y_1 e^{-y_1}$$

Since $x_1 \geq 0$, $x_2 \geq 0$ $\therefore$ $y_1 \geq 0$

$y_2 = \frac{X_1}{X_1 + X_2}$ since denomi > nu $0 \leq y_2 \leq 1$

25/02/2019

eg: $U = \begin{bmatrix} U_{11} & U_{21} \\ U_{12} & U_{22} \end{bmatrix}$ → eigen vectors (each column) of covariance matrix of X (2x2)

choosing $A = U^T$, $C_Y$ becomes diagonal i.e. $(Y_1, \dots, Y_N)$ are uncorrelated which is useful in compression, feature extraction etc.

$Y = U^T X$ — decided by $X_1, \dots X_N$

$$\underset{N\times 1}{\begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix}} = \underset{N\times N}{\begin{bmatrix} U_1^T \dots \\ U_2^T \dots \\ \vdots \\ U_N^T \dots \end{bmatrix}} \underset{N\times 1}{\begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix}}$$

unlike DFT, DCT etc.

∴ Data Dependent — which leads to an optimum transform in MSE between the original input and the reconstructed output when M number of Y's are retained where $M < N$

$Y = U^T X \longrightarrow X = U^{T^{-1}} Y = UY$

MSE = 0 when all N Y's are used. $\hat{X} = X$ when $M = N$

suppose $N = 8$, $M = 3$ then $\hat{X} = UY$ where $Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ 0 \\ 0 \\ \vdots \end{bmatrix}_{8\times 1}$

$$\hat{X} = \underset{8\times 1}{\begin{bmatrix} \ \end{bmatrix}} = \underset{8\times 8}{\overset{U}{\begin{bmatrix} \ \end{bmatrix}}} \underset{8\times 1}{\overset{Y}{\begin{bmatrix} \ \end{bmatrix}}}$$

↓ last 5 columns not needed.

8x3 can be used instead & Y
U 3x1

$$\underset{8\times 1}{\hat{X}} = \underset{8\times 3}{\overset{U}{\begin{bmatrix} \ \end{bmatrix}}} \underset{3\times 1}{\overset{Y}{\begin{bmatrix} \ \end{bmatrix}}}$$

Taking $E(X - \hat{X})^2$ will give the MSE — will be lesser than other transforms

→ $C_Y = \Sigma \Rightarrow \begin{bmatrix} \sigma_{Y_1}^2 & Cov(Y_1,Y_2) & \dots & Cov(Y_1,Y_N) \\ Cov(Y_2,Y_1) & \sigma_{Y_2}^2 & & \vdots \\ \vdots & & \ddots & \\ Cov(Y_N,Y_1) & \dots & & \sigma_{Y_N}^2 \end{bmatrix}_{N\times N} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{bmatrix}$

* Addition of all elements of $C_X$ gives Total Power content of X
$Var(X_1 + X_2 + \dots X_N)$

$2 C_x a_1 - 2\lambda a_1 = 0$

$a_1 C_x a_1 = \lambda a_1$

__/__/____

- Similarly Total Power Contents of Y only uses diagonal elements
  $Var(Y_1 + Y_2 + \cdots Y_N) = Var(Y_1) + Var(Y_2) + \cdots Var(Y_N)$

Total power in input is distributed in covariances but in output only concentrated in variances. For PCA, high power will be found in $Y_1$, lesser in $Y_2$, and so on. So retaining first few $Y_i$ can give for example 90% of power.

→ Max. $\lambda$ value corresponds to max Variance ($\lambda_1$).

$$C_Y = A\, C_X\, A^T$$

$$\cancel{\begin{bmatrix} \sigma_{Y_1}^2 & Cov(Y_1,Y_2) \\ Cov(Y_2,Y_1) & \sigma_{Y_2}^2 \end{bmatrix}_{2\times2} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}_{2\times2} \begin{bmatrix} \sigma_{X_1}^2 & Cov(X_1,X_2) \\ Cov(X_2,X_1) & \sigma_{X_2}^2 \end{bmatrix}_{2\times2} \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}_{2\times2}}$$

$$\sigma_{Y_1}^2 = a_1^T C_X a_1$$

$$\begin{bmatrix} \sigma_{Y_1}^2 & Cov(Y_1,Y_2) \\ Cov(Y_2,Y_1) & \sigma_{Y_2}^2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} \sigma_{X_1}^2 & Cov(X_1,X_2) \\ Cov(X_2,X_1) & \sigma_{X_2}^2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$a_1 = \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix} \quad a_1^T = \begin{bmatrix} a_{11} & a_{12} \end{bmatrix} \quad a_2 = \begin{bmatrix} a_{21} \\ a_{22} \end{bmatrix} \quad a_2^T = \begin{bmatrix} a_{21} & a_{22} \end{bmatrix}$$

$$\sigma_{Y_1}^2 = a_1^T C_X a_1$$

$a_1 \cdot a_2 = 0$ orthogonal + unit norm

$$L(a_1, \lambda) = a_1^T C_X a_1 - \lambda(\|a_1\|^2 - 1)$$

to obtain $a_1$ that maximises this constraint: $\|a_1\| = 1$

Diff. wrt $a_1$ & equate to 0 to maximise,

$$\frac{dL(a_1,\lambda)}{da_1} = 2 C_x a_1 - 2\lambda a_1 = 0$$

$$C_x a_1 = \lambda a_1$$

26/02/2018

$\sigma_{y_1}^2 = a_1^T C_x a_1 = a_1^T \lambda_1 a_1 = \lambda_1 a_1^T a_1 = \lambda_1$

III'ly $\sigma_{y_2}^2 = a_2^T C_x a_2 = a_2^T [\lambda_2 a_2] = \lambda_2$

PCA for image compression - Using hyperspectral images captured by satellites.

Cuprite data - 188 bands (AVIRIS)

250 x 190 size each band

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} \; \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

→ first band
→ second band

$X_1 = \begin{bmatrix} \; \end{bmatrix}_{N \times N} \xrightarrow{N^2 \text{ transform}} \begin{bmatrix} x_1 \\ \vdots \\ x_{N} \end{bmatrix}_{188 \times N^2}$

$A = U^T \rightarrow 188 \times 188$

$$\underbrace{\begin{bmatrix} X_1 \end{bmatrix}}_{N^2 \times 1} \quad \underbrace{\begin{bmatrix} X_2 \end{bmatrix}}_{N^2 \times 1} \quad \cdots \quad \underbrace{\begin{bmatrix} X_{188} \end{bmatrix}}_{N^2 \times 1}$$

↓ convert to $X_1 - m_{x_1}$

Mean subtracted R.V.

& now $E[X_1^2]$ ... gives variance

& $E(X_1 X_2)$ gives $\mathrm{cov}(x_1, x_2)$

$\hat{X} = UY$

Scanned by CamScanner

01/03/2019

PCA - Linear transform - optimum in the mean sense
$(E(X-\hat{X})^2)$ is minimum - decorrelate the data.

· How to generate RV's having given covariance matrix

→ Example on compressing hyperspectral images using PCA.

- Linear independency of RVs

~~Statistical independency of RVs $X_1$ & $X_2$~~

~~$f_{X_1X_2}(x_1,x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$~~ - This has to be true

~~$E(X_1X_2) = E(X_1) \cdot E(X_2)$~~

$X_1, X_2, \ldots X_N$ are dependent if

$a_1X_1 + a_2X_2 + \cdots + a_N X_N = 0$, $a_1, a_2, a_3$ - few of them atleast should be non-zero, if all are zero then they are independent

$$aX_1 + a_2 X_2 = 0$$

$$X_2 = -\frac{a_1}{a_2} X_1$$

· Covariance matrix - $a^T C_X a \geq 0$ → +ve semidefinite.

· One can show that $C_X$ matrix is not invertible if $X_1, X_2 \ldots X_N$ are linearly dependent.

· Consider $\mathrm{Var}(a_1U_1 + a_2U_2)^2$ where $U_1 = X_1 - m_{X_1}$, $U_2 = X_2 - m_{X_2}$

$$= E((a_1U_1 + a_2U_2)^2) - 0$$

$$= a^T C_X a \geq 0$$

$$[E(a_1U_1 + a_2U_2)]^2 = 0$$

Now consider $U_1$ & $U_2$ as dependent R.V.'s ∴ Non-zero $a_1, a_2$ such that $a_1U_1 + a_2U_2 = 0$

$$\Rightarrow E((a_1U_1 + a_2U_2)^2) = 0$$

$$a^T C_X a = 0 \qquad C_X X = \lambda X$$

$$a^T \lambda a = 0$$

$$\lambda \|a\|^2 = 0$$

Since $a_1, a_2$ are non-zero $\|a\|^2 \neq 0$

$$\underline{\lambda = 0}$$

Indicating $C_X$ is Non-invertible.

Entries of $C_X$ not in our control, depends on $X_1 X_2 \cdots X_N$.

01/03/2019

⇒ Minimum Mean Squared Estimation (MMSE) of a R.V.

X is accessible, how will you estimate Y?

$\hat{Y}$ has to be as close to Y as possible

$$\underset{\text{R.V.}}{Y} \longrightarrow \boxed{\text{channel}} \longrightarrow \underset{\text{R.V.}}{X}$$

- One of the ways in which this can be achieved is by using MMSE estimate of Y, i.e. $\min (Y-\hat{Y})^2 = \min (Y-g(X))^2$
- Let us say we want to estimate Y as a constant, $\hat{Y} = b$

Then we have to minimize $E((Y-b)^2)$ wrt b → differentiate

$$E((Y-b)^2) = \int (y-b)^2 f_Y(y)\, dy.$$

diff. wrt b & equating to 0

$$-2\int (y-b) f_Y(y)\, dy = 0$$

$$E((y-b)) = 0$$

$$E(y) = b \rightarrow \hat{Y}$$

$$\min E((Y-\hat{Y})^2) = E((Y-E(Y))^2) = Var(Y)$$