Session 24

Statistical Independence of Multiple Random Variables

Defn: The random variables $X_1, \ldots, X_n$ are statistically independent iff all events of the form

$\{X_1 \in A_1\}, \{X_2 \in A_2\}, \ldots, \{X_n \in A_n\}$

are statistically independent, where

$A_1 \in B(\mathbb{R}), A_2 \in B(\mathbb{R}), \ldots, A_n \in B(\mathbb{R}),$

for all Borel sets $A_1, \ldots, A_n$.

Equivalently,

Defn': The jointly distributed random variables $X_1, X_2, \ldots, X_n$ are <u>statistically independent</u> iff

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = f_{X_1}(x_1)\, f_{X_2}(x_2) \cdots f_{X_n}(x_n).$$

Conditional Densities, Characteristic Functions and Normality (n-dim Gaussian)

Conditional densities involving $X_1, \ldots, X_n$ generalize easily from the 2-dim case:

$$f(x_1,\ldots,x_k \mid x_{k+1},\ldots,x_n) = \frac{f(x_1,\ldots,x_n)}{f(x_{k+1},\ldots,x_n)}$$

(<u>n.b.</u> $f(x_1, x_3 \mid x_2, x_4) = \dfrac{f(x_1,x_2,x_3,x_4)}{f(x_2,x_4)}$)

24.4

Recall that the correlation between two RVs $\mathbb{X}_j$ and $\mathbb{X}_k$ is

$$R_{jk} = E[\mathbb{X}_j \mathbb{X}_k]$$

and their covariance is given by

$$C_{jk} = E\left[(\mathbb{X}_{.} - \overline{\mathbb{X}}_j)(\mathbb{X}_k - \overline{\mathbb{X}}_k) \right]$$

We are interested in tabulating these values for the RVec

$$\underline{\mathbb{X}} = (\mathbb{X}_1, \dots, \mathbb{X}_n)$$

24.5

<u>Defn</u>: Given a RVec $\underline{\mathbb{X}} = (\mathbb{X}_1, \dots, \mathbb{X}_n)$ we define the <u>correlation matrix</u>

$$R_{\underline{X}} = \begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1n} \\ R_{21} & 22 & \cdots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \cdots & R_{nn} \end{pmatrix}$$

and the <u>covariance matrix</u>

$$C_{\underline{X}} = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{2} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{pmatrix}$$

24.6

If $\mathbb{X}_j$ and $\mathbb{X}_k$ are complex RVs, then

$$R_{jk} \triangleq E[\mathbb{X}_j \mathbb{X}_k^*]$$

$$C_{jk} \triangleq E[(\mathbb{X}_j - \overline{\mathbb{X}}_j)(\mathbb{X}_k - \overline{\mathbb{X}}_k)^*]$$

n.b. A complex RV $\mathbb{Z}$ has the form

$$\mathbb{Z} = \mathbb{X} + i\,\mathbb{Y},$$

where $\mathbb{X}$ and $\mathbb{Y}$ are j-dist real RVs. $\mathbb{Z}^* = \mathbb{X} - i\,\mathbb{Y}$

23-20

24.7

n.b. $\underline{\mathbb{X}} = (\mathbb{X}_1, \dots, \mathbb{X}_n) \sim$ row vector

$$E[\underline{\mathbb{X}}] \triangleq (E[\mathbb{X}_1], \dots, E[\mathbb{X}_n]) = \overline{\underline{\mathbb{X}}}.$$

We can also write

$$R_{\underline{\mathbb{X}}} = E[\underline{\mathbb{X}}^T \underline{\mathbb{X}}^*]$$

and

$$C_{\underline{\mathbb{X}}} = E[(\underline{\mathbb{X}} - \overline{\underline{\mathbb{X}}})^T (\underline{\mathbb{X}} - \overline{\underline{\mathbb{X}}})^*]$$

n.b. For a row vector $\underline{x} = (x_1, \dots, x_n)$

$$\underline{x}^T \underline{x} = \begin{pmatrix} x_1x_1 & x_1x_2 & \cdots & x_1x_n \\ x_2x_1 & x_2x_2 & \cdots & x_2x_n \\ \vdots & \vdots & & \vdots \\ x_nx_1 & x_nx_2 & \cdots & x_nx_n \end{pmatrix} \quad \text{(outer product)}$$

$$\underline{x}\,\underline{x}^T = x_1x_1 + x_2x_2 + \cdots + x_nx_n \quad \text{(inner product)}$$

Opposite of what we are used to for column vectors.

24.8

Defn: An $n \times n$ matrix $\mathbb{B}$ is said to be <u>non-negative definite</u> if

$$\sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j^* b_{ij} \geq 0$$

for all complex numbers $x_1, \ldots, x_n$, where

$$\mathbb{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}.$$

If the inequality is strict (>0) for all $x_1, \ldots, x_n$ not all zero, then we say that the matrix $\mathbb{B}$ is <u>positive definite</u>.

24.9

Theorem: The correlation matrix $R_{\mathbb{X}}$ of any RVec $\underline{\mathbb{X}}$ is non-negative definite.

Proof:

$$0 \leq E\left[\,|a_1 \mathbb{X}_1 + a_2 \mathbb{X}_2 + \cdots + a_n \mathbb{X}_n|^2\right]$$

$$E\left[\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j^* \mathbb{X}_i \mathbb{X}_j^*\right]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j^* E[\mathbb{X}_i \mathbb{X}_j^*]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j^* R_{ij}, \quad \text{for all } a_1, \ldots, a_n \in \mathbb{C}$$

$\Rightarrow R_{\underline{\mathbb{X}}}$ is non-negative definite ∎

Corollary: $C_{\mathbb{X}}$ is non-negative definite.

n.b. $C_{\underline{\mathbb{X}}} = R_{\underline{\tilde{\mathbb{X}}}}$, $\tilde{\mathbb{X}}_1 = \mathbb{X}_1 - \overline{\mathbb{X}_1}$, ..., $\tilde{\mathbb{X}}_n = \mathbb{X}_n - \overline{\mathbb{X}_n}$.

Char. Ftn. of a RVec

Let $\underline{X} = (X_1, \ldots, X_n)$ be a RVec.

Then the characteristic function of $\underline{X}$ is

$$\Phi_{\underline{X}}(\omega_1, \ldots, \omega_n) = \Phi_{\underline{X}}(\underline{\Omega})$$

$$\triangleq E\left[e^{i(\omega_1 X_1 + \omega_2 X_2 + \cdots + \omega_n X_n)}\right]$$

$$= E\left[e^{i \underline{\Omega}\, \underline{X}^T}\right]$$

where

$$\underline{\Omega} = (\omega_1, \ldots, \omega_n).$$

n.b. Let $Z = X_1 + X_2 + \cdots + X_n$

Then if $\Phi_{\underline{X}}(\omega_1, \ldots, \omega_n)$ is the char. fth. of $X_1, \ldots, X_n$, then

$$\Phi_Z(\omega) = E[e^{i\omega Z}] = E\left[e^{i\omega(X_1 + X_2 + \cdots + X_n)}\right]$$

$$= E\left[e^{i(\omega X_1 + \omega X_2 + \cdots + \omega X_n)}\right]$$

$$= \Phi_{\underline{X}}(\omega, \omega, \ldots, \omega).$$

Furthermore, if $X_1, \ldots, X_n$ are stat. indep., Then

$$\Phi_{X_1, \ldots, X_n}(\omega_1, \ldots, \omega_n) = \Phi_{X_1}(\omega_1) \cdot \Phi_{X_2}(\omega_2) \cdots \Phi_{X_n}(\omega_n)$$

and thus

$$\Phi_Z(\omega) = \Phi_{X_1}(\omega) \cdot \Phi_{X_2}(\omega) \cdots \Phi_{X_n}(\omega).$$

Gaussian RVecs and their Char. Ftns.

24.12

Defn: Let $X_1, \ldots, X_n$ be a RVec of dimension n. Then $\underline{X} = (X_1, \ldots, X_n)$ is a <u>Gaussian RVec</u>, and $X_1, \ldots, X_n$ are jointly Gaussian iff

$$Z = a_0 + \sum_{j=1}^{n} a_j X_j$$

is a Gaussian RV for all $a_0, a_1, \ldots, a_n \in \mathbb{R}$.

24.13

The joint characteristic function of such a Gaussian RVec has the following form:

$$\Phi_{\underline{X}}(\underline{\Omega}) = \Phi_{X_1,\ldots,X_n}(\omega_1, \ldots, \omega_n)$$

$$= e^{i \underline{\Omega}\, \underline{m}_{\underline{X}}^T} e^{-\frac{1}{2} \underline{\Omega}\, C_{\underline{X}}\, \underline{\Omega}^T}$$

$$= e^{i \sum_{j=1}^{n} \omega_j \mu_j} e^{-\frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \omega_j \omega_k c_{jk}} \qquad (*)$$

where

$$\underline{m}_{\underline{X}} = (E[X_1], \ldots, E[X_n]) = (\mu_1, \ldots, \mu_n)$$

and

$$C_{\underline{X}} = E\left[(\underline{X} - \underline{m}_{\underline{X}})^T (\underline{X} - \underline{m}_{\underline{X}}) \right]$$

$$= [c_{jk}] \quad n \times n \text{ covariance matrix of } \underline{X}$$

n.b. Suppose that $X_1, \dots, X_n$ are j-Gaussian with char. fn (*)

Let $Z = \sum_{k=1}^{n} a_k X_k$. What is distribution of Z?

Show that Z is Gaussian and find its mean and variance (exercise)

So we have shown the following important result:

Theorem: Let $X_1, X_2, \dots, X_n$ be jointly distributed, jointly Gaussian random variables. Then any linear combination

$$Z = \sum_{j=1}^{n} a_j X_j$$

is a Gaussian random variable, for any $a_1, a_2, \dots, a_n \in \mathbb{R}$.

Joint Gaussian Examples:

24.16

Recall:

- We have seen that

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

- We also know that

$$f_{Y}(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx$$

and

$$f_{XY}(x,y) = f_{Y}(y \mid \{X=x\})\, f_{X}(x)$$

$$\Rightarrow f_{X}(x \mid \{Y=y\}) = \frac{f_{Y}(y \mid \{X=x\})\, f_{X}(x)}{\int_{-\infty}^{\infty} f_{Y}(y \mid \{X=\alpha\}) f_{X}(\alpha)\,d\alpha}$$

(Bayes Theorem)

Also recall: We investigated two estimators for estimating the value of X given $\{Y=y\}$:

24.17

$$\hat{X}_{MMS}(y) = E[X \mid \{Y=y\}],$$

and

$$\hat{X}_{MAP}(y) = \arg\max_{x} \{ f_{X}(x \mid \{Y=y\}) \}.$$

Both of these estimators require that we find

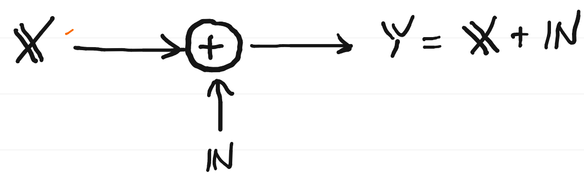
$$f_{X}(x \mid \{Y=y\}).$$

24.18

Example: Let X and N be two zero-mean, jointly distributed independent Gaussian RVs, with variances σ_X^2 and σ_N^2, respectively.

$X \perp\!\!\!\perp N$

Now consider a new RV Y:



Suppose I observe $Y = y$. What is the Minimum Mean-Square Error estimator of X given $\{Y = y\}$?

24.9

24.19

$X = X$

$Y = X + N$

$$\Phi_{X,Y}(\omega_1, \omega_2) = E\left[e^{i(\omega_1 X + \omega_2 Y)}\right]$$

$$= E\left[e^{i(\omega_1 X + \omega_2(X + N))}\right]$$

$$= E\left[e^{i((\omega_1 + \omega_2)X + \omega_2 N)}\right] = E\left[e^{i(\omega_1+\omega_2)X}\right] \cdot E\left[e^{i\omega_2 N}\right]$$

$$= \Phi_X(\omega_1 + \omega_2) \cdot \Phi_N(\omega_2)$$

24.20

$\hat{X}_{MMS}(y) = E[X \mid \{Y = y\}]$

So I need to find $f_X(x \mid \{Y=y\}) = \dfrac{f_{XY}(x,y)}{f_Y(y)}$

We need to find the j-pdf $f_{XY}(x,y)$ of X and Y:

- Easy to show $\sigma_Y^2 = \sigma_X^2 + \sigma_N^2$.
- Also, $r_{XY} = \dfrac{E[XY] - E[X] \cdot E[Y]}{\sigma_X \sigma_Y}$

(exercise)

$= \cdots = \sqrt{\dfrac{\sigma_X^2}{\sigma_X^2 + \sigma_N^2}}$

24.21

Aside 1:

$r_{XY} = \dfrac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$

$\mathrm{var}(Y) = \sigma_X^2 + \sigma_N^2, \quad \sigma_Y = \sqrt{\sigma_X^2 + \sigma_N^2}$

$E[XY] = E[X(X+N)] = E[X^2 + X \cdot N]$

$= E[X^2] + \overset{0}{E[X]} \cdot E[N] = E[X^2] = \sigma_X^2$

$\Rightarrow \mathrm{cov}(X,Y) = \sigma_X^2$

$r = \dfrac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \dfrac{\sigma_X^2}{\sigma_X \sigma_Y} = \dfrac{\sigma_X}{\sigma_Y}$

$= \sqrt{\dfrac{\sigma_X^2}{\sigma_X^2 + \sigma_N^2}}$

24.22

$$f_X(x \mid \{Y=y\}) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

$$= \frac{\frac{1}{2\pi\sigma_X\sigma_Y(1-r^2)} \exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{x^2}{\sigma_X^2} - \frac{2rxy}{\sigma_X\sigma_Y} + \frac{y^2}{\sigma_Y^2}\right]\right\}}{\frac{1}{\sqrt{2\pi}\,\sigma_Y} \exp\left\{\frac{-y^2}{2\sigma_Y^2}\right\}}$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_X\sqrt{1-r^2}} \exp\left\{\frac{-1}{2(1-r^2)}\left[\frac{x^2}{\sigma_X^2} - \frac{2rxy}{\sigma_X\sigma_Y} + \frac{y^2}{\sigma_Y^2} - (1-r^2)\frac{y^2}{\sigma_Y^2}\right]\right\}$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_X\sqrt{1-r^2}} \exp\left\{\frac{-1}{2(1-r^2)}\left[\frac{x^2}{\sigma_X^2} - \frac{2rxy}{\sigma_X\sigma_Y} + r^2\frac{y^2}{\sigma_Y^2}\right]\right\}$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_X\sqrt{1-r^2}} \exp\left\{\frac{-1}{2(1-r^2)}\left[\frac{x}{\sigma_X} - \frac{ry}{\sigma_Y}\right]^2\right\}$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_X\sqrt{1-r^2}} \exp\left\{\frac{-1}{2(1-r^2)\sigma_X^2}\left[x - r\frac{\sigma_X}{\sigma_Y}y\right]^2\right\}$$

24.23

So we see here that

$$f_X(x \mid Y=y) = \frac{1}{\sqrt{2\pi}\,\sigma_X\sqrt{1-r^2}} \exp\left\{\frac{-1}{2\sigma_X^2(1-r^2)}\left[x - r\frac{\sigma_X}{\sigma_Y}y\right]^2\right\},$$

which is a Gaussian pdf with mean $r\frac{\sigma_X}{\sigma_Y}y$ and variance $\sigma_X^2(1-r^2)$. So by inspection, we have

$$\hat{X}_{MMS}(y) = r\frac{\sigma_X}{\sigma_Y} = \sqrt{\frac{\sigma_X^2}{\sigma_X^2+\sigma_N^2}} \cdot \frac{\sigma_X}{\sqrt{\sigma_X^2+\sigma_N^2}} \cdot y = \boxed{\frac{\sigma_X^2}{\sigma_X^2+\sigma_N^2} \cdot y}$$

24.24

Also, it can be shown (exercise) that

$$f_{\mathbb{X}}(x \mid \{\mathbb{Y}=y\}) = \frac{f_{\mathbb{X}\mathbb{Y}}(x,y)}{f_{\mathbb{Y}}(y)}$$

(exercise)

$$= \cdots = \frac{1}{\sqrt{2\pi}\,\sigma_X\sqrt{1-r_{XY}^2}} \exp\left\{ \frac{-\left(x - r_{XY}\frac{\sigma_X}{\sigma_Y}y\right)^2}{2\sigma_X^2(1-r_{XY}^2)} \right\}$$

$$\therefore \hat{X}_{MMS}(y) = E[\mathbb{X} \mid \{\mathbb{Y}=y\}] = r_{XY}\frac{\sigma_X}{\sigma_Y}\,y$$

$$= \cdots = \left(\frac{\sigma_X^2}{\sigma_X^2+\sigma_N^2}\right)y$$

24.25

Exam 3 Cut-off

↑ ↑ ↑ ↑ ↑ ↑