ME 233 Advance Control II

Lecture 3
Introduction to Probability Theory

Random Vectors and Conditional Expectation

(ME233 Class Notes pp. PR4-PR6)

Outline

- Multiple random variables
- Random vectors
  – Correlation and covariance
- Gaussian random variables
- PDFs of Gaussian random vectors
- Conditional expectation of Gaussian random vectors

Multiple Random Variables

Let X and Y be continuous random variables.

- Their joint cumulative distribution function (CDF) is given by

$$F_{XY}(x,y) = \underbrace{P(X \leq x,\, Y \leq y)}_{P(X \leq x \text{ and } Y \leq y)}$$

Multiple Random Variables

Let X and Y be continuous random variables with a differentiable joint CDF

$$F_{XY}(x,y) = P(X \leq x,\, Y \leq y)$$

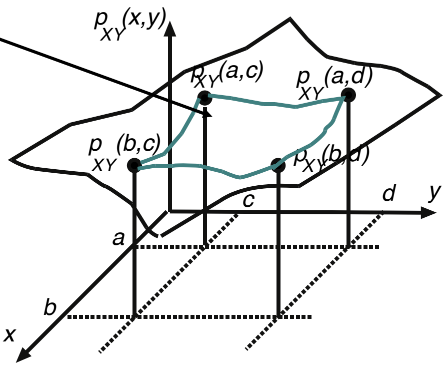
Their joint probability density function (PDF) is

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

5

Multiple Random Variables

$$P(a \le X \le b,\, c \le Y \le d) = \int_a^b \int_c^d p_{XY}(x,y)\, dy dx$$



$p_{XY}(x,y)$ has the usual meaning of density

6

Multiple Random Variables

Let X and Y be ***independent***

- Then:

$$F_{XY}(x,y) = F_X(x)\, F_Y(y)$$

Marginal CDF of X

Marginal CDF of Y

7

Multiple Random Variables

Let X and Y be ***independent***

- Then:

$$p_{XY}(x,y) = p_X(x)\, p_Y(y)$$

Marginal PDF of X

Marginal PDF of Y

8

Correlation and Covariance

Let X and Y be continuous random variables with joint PDF

$$p_{XY}(x,y)$$

- ***Correlation***:

$$R_{XY} = E\{XY\}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, p_{XY}(x,y)\, dy dx$$

Mean

9

Let X and Y be continuous random variables with joint PDF $p_{XY}(x,y)$

- ***Mean***:

$$m_X = E\{X\}$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, p_{XY}(x,y)\, dydx$$
$$= \int_{-\infty}^{\infty} x\, p_X(x)\, dx$$

where $$p_X(x) = \int_{-\infty}^{\infty} p_{XY}(x,y)\, dy$$

Correlation and Covariance

10

Let X and Y be continuous random variables with joint PDF $p_{XY}(x,y)$

- ***Covariance***:

$$\Lambda_{XY} = E\{(X - m_X)(Y - m_Y)\}$$

means

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - m_X)(y - m_Y)\, p_{XY}(x,y)\, dydx$$

Correlation and Covariance

11

Let X and Y be continuous random variables with joint PDF $p_{XY}(x,y)$

- X and Y ***are uncorrelated*** if :

$$\Lambda_{XY} = 0$$ ***their covariance is zero***

- X and Y ***are orthogonal*** if :

$$R_{XY} = 0$$ ***their correlation is zero***

Multiple Random Variables

12

- X and Y are uncorrelated if and only if

$$R_{XY} = E\{XY\} = E\{X\}\,E\{Y\} = m_X\, m_Y$$

Proof:

$$\Lambda_{XY} = E\{(X - m_X)(Y - m_Y)\}$$
$$= E\{XY\} - m_X \underbrace{E\{Y\}}_{m_Y} - \underbrace{E\{X\}}_{m_X} m_Y + m_X m_Y$$
$$= E\{XY\} - m_X m_Y$$

therefore $\Lambda_{XY} = 0 \quad \Leftrightarrow \quad E\{XY\} = m_X m_Y$ ∎

13

Variance

The ***variance*** of random variable X is:

$$\sigma_X^2 = E[(X - m_X)^2]$$

$$= E\{(X - m_X)(X - m_X)\}$$

$$= \Lambda_{XX}$$

14

Marginal PDF

Let X and Y have a joint PDF $p_{XY}(x,y)$

- ***Marginal or unconditional*** PDFs:

$$p_X(x) = \int_{-\infty}^{\infty} p_{XY}(x,y)\,dy$$

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

15

Marginal PDF

Let X and Y have a joint PDF $p_{XY}(x,y)$

- Expected value of X

$$m_X = E\{X\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\,p_{XY}(x,y)\,dydx$$

$$= \int_{-\infty}^{\infty} x\,p_X(x)\,dx$$

16

Conditional PDF

Let X and Y have a joint PDF $p_{XY}(x,y)$

- The ***Conditional*** PDF of X given an outcome of $Y = y_1$:

$$\underbrace{p_{X|Y=y_1}(x)}_{p_{X|y_1}(x)} = \frac{p_{XY}(x,y_1)}{p_Y(y_1)}$$

17

Conditional PDF

Let X and Y have a joint PDF $p_{XY}(x,y)$

- The ***Conditional*** PDF of Y given an outcome of $X = x_1$:

$$p_{Y|x_1}(y) = \frac{p_{XY}(x_1,y)}{p_X(x_1)}$$

18

Conditional PDF

Let X and Y have a joint PDF $p_{XY}(x,y)$

- ***Bayes' rule***:

$$\boxed{\begin{aligned} p_{X|y}(x)\,p_Y(y) &= p_{Y|x}(y)\,p_X(x) \\ &= p_{XY}(x,y) \end{aligned}}$$

19

Conditional Expectation

Let X and Y have a joint PDF $p_{XY}(x,y)$

- Conditional Expectation of X given an outcome of $Y = y_1$:

$$\underbrace{m_{X|Y=y_1}}_{m_{X|y_1}} = E\{X|Y=y_1\}$$

$$= \int_{-\infty}^{\infty} x\, p_{X|y_1}(x)dx$$

20

Conditional Variance

Let X and Y have a joint PDF $p_{XY}(x,y)$

- Conditional variance of X given an outcome of $Y = y_1$:

$$\sigma^2_{X|y_1} = \Lambda_{X|y_1 X|y_1}$$

$$= E\{(X - m_{X|y_1})^2|Y = y_1\}$$

$$= \int_{-\infty}^{\infty} (x - m_{X|y_1})^2\, p_{X|y_1}(x)dx$$

21

Independent Variables

Let X and Y be independent. Then:

$$p_{XY}(x,y) = p_X(x)\, p_Y(y)$$

$$p_{X|y}(x) = p_X(x)$$

$$p_{Y|x}(y) = p_Y(y)$$

22

Independent Variables

If X and Y are independent random variables, then X and Y are uncorrelated

Proof:

$$\Lambda_{XY} = E\{(X - m_X)(Y - m_Y)\}$$

$$= E\{X - m_X\}E\{Y - m_Y\} \qquad \textbf{\textit{(independence)}}$$

$$= 0 \qquad \blacksquare$$

The converse statement is NOT true in general

23

Bilateral Laplace and Fourier Transforms

Given $f : \mathcal{R} \to \mathcal{R}$

- Laplace transform: $F(s) = \mathcal{L}\{f(\cdot)\}$

$$F(s) = \int_{-\infty}^{\infty} e^{-st} f(t)dt \qquad s \in \mathcal{C}$$

- Inverse Laplace transform:

$$f(t) = \frac{1}{2\pi j}\int_{\gamma - j\infty}^{\gamma + j\infty} e^{st} F(s)ds$$

for some real γ so that contour path of integration is in the region of convergence

24

Bilateral Laplace and Fourier Transforms

Given $f : \mathcal{R} \to \mathcal{R}$

- Fourier transform: $F(j\omega) = \mathcal{F}\{f(\cdot)\}$

$$F(j\omega) = \int_{-\infty}^{\infty} e^{-j\omega t} f(t)dt \qquad \omega \in \mathcal{R}$$

- Inverse Fourier transform:

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{j\omega t} F(\omega)d\omega$$

25

Moment Generating Function

The Fourier transform of the PDF of a random variable X is also called the *moment generating function* or characteristic function

Notice that, given the PDF $p_X(x)$

$$P_X(j\omega) \;=\; \mathcal{F}\{p_X(\cdot)\} = \int_{-\infty}^{\infty} e^{-j\omega x}\, p_X(x)\, dx$$

$$= E\left[e^{-j\omega X}\right]$$

it can be shown that $\quad E\left[X^n\right] = j^n P_X^{[n]}(j\omega)|_{\omega=0}$

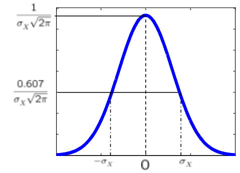
where [n] indicates the nth derivative w/r ω (see Poolla's notes)

26

Properties of Normal distributions

The *moment generating function* of a zero-mean normal distribution is also normal.

$$p_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}}\,\text{exp}\left(-\frac{x^2}{2\sigma_X^2}\right)$$



$$P_X(j\omega) \;=\; E\left[e^{-j\omega X}\right] = \int_{-\infty}^{\infty} e^{-j\omega x}\, p_X(x)\, dx$$

$$= \text{exp}\left(\frac{-\sigma_X^2\omega^2}{2}\right)$$

27

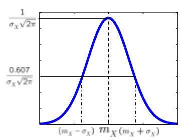
Moment generating functions of Normal PDFs

Let,

$$X \sim N(m_X, \sigma_X^2)$$

i.e.,

$$p_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}}\,\text{exp}\left(-\frac{(x-m_X)^2}{2\sigma_X^2}\right)$$



The moment generating functions of X is:

$$P_X(j\omega) = E\left\{e^{-j\omega X}\right\} = \text{exp}(-j\omega m_X)\,\text{exp}\left(\frac{-\sigma_X^2\omega^2}{2}\right)$$

28

Sum of independent random variables

Let X and Y be two ***independent*** random variables with PDFs $p_X(x)$ $p_Y(y)$

Define

$$Z = X + Y$$

then

$$p_Z(z) \;=\; \int_{-\infty}^{\infty} p_X(x)p_Y(z-x)dx$$

$$= p_X(\cdot) * p_Y(\cdot) \qquad \textit{(convolution)}$$

29

Proof

Assume X and Y are two ***independent*** random variables and define

$$Z = X + Y$$

Let us now calculate the moment generating function of Z:

$$\begin{aligned} P_Z(j\omega) &= E\{e^{-j\omega Z}\} \\ &= E\{e^{-j\omega(X+Y)}\} = E\{e^{-j\omega X}\, e^{-j\omega Y}\} \\ &= E\{e^{-j\omega X}\}\, E\{e^{-j\omega Y}\} \quad \textbf{\textit{(independence)}} \\ &= P_X(j\omega)\, P_Y(j\omega) \end{aligned}$$

30

Proof

Since

$$P_Z(j\omega) = P_X(j\omega)\, P_Y(j\omega)$$

Applying the inverse Fourier transform,

$$\begin{aligned} p_Z(z) &= \int_{-\infty}^{\infty} p_X(x) p_Y(z-x)dx \\ &= p_X(\cdot) * p_Y(\cdot) \end{aligned}$$

■

31

Random Vectors

Let X_1 and X_2 be continuous random variables. Recall that:

- Their joint CDF is given by

$$F_{X_1X_2}(x_1, x_2) = P(X_1 \leq x_1,\, X_2 \leq x_2)$$

- Their joint PDF is

$$p_{X_1X_2}(x_1, x_2) = \frac{\partial^2 F_{X_1X_2}(x_1, x_2)}{\partial x_1 \partial x_2}$$

32

Random Vector

Define the random vector $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \in \mathcal{R}^2$

(and the dummy vector) $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathcal{R}^2$

with CDF

$$F_X(x) = P(X_1 \leq x_1,\, X_2 \leq x_2)$$

$$F_X : \mathcal{R}^2 \to \mathcal{R}_+$$

33

Random Vector

Define the random vector $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \in \mathcal{R}^2$

(and the dummy vector) $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathcal{R}^2$

with PDF

$$p_X(x) = \frac{\partial^2 F_X(x)}{\partial x_1 \, \partial x_2}$$

$$p_X : \mathcal{R}^2 \to \mathcal{R}_+$$

34

Random Vector

Define the random vector $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \in \mathcal{R}^2$

Mean:

$$\begin{aligned} m_X &= E\{X\} = \begin{bmatrix} m_{X_1} \\ m_{X_2} \end{bmatrix} \\ &= \int_{\mathcal{R}^2} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} p_X(x) dx_1 dx_2 \end{aligned}$$

35

Random Vector

Define the random vector $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \in \mathcal{R}^2$

Mean:

$$m_X = \begin{bmatrix} m_{X_1} \\ m_{X_2} \end{bmatrix} = \begin{bmatrix} \int_{-\infty}^{\infty} x p_{X_1}(x) dx \\ \int_{-\infty}^{\infty} y p_{X_2}(y) dy \end{bmatrix}$$

$$p_{X_1}(x) = \int_{-\infty}^{\infty} p_X(x, y) \, dy$$

Marginal PDFs

$$p_{X_2}(y) = \int_{-\infty}^{\infty} p_X(x, y) \, dx$$

36

Correlation

$$\begin{aligned} R_{XX} &= E\{XX^T\} \in \mathcal{R}^{2\times 2} \\ &= E\left\{ \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} \right\} \\ &= \begin{bmatrix} R_{X_1X_1} & R_{X_1X_2} \\ R_{X_2X_1} & R_{X_2X_2} \end{bmatrix} \end{aligned}$$

37

Covariance

$$\Lambda_{XX} = E\{(X-m_X)(X-m_X)^T\} \in \mathcal{R}^{2\times 2}$$

$$= E\left\{\begin{bmatrix} X_1 - m_{X_1} \\ X_2 - m_{X_2} \end{bmatrix} \begin{bmatrix} X_1 - m_{X_1} & X_2 - m_{X_2} \end{bmatrix}\right\}$$

$$= \begin{bmatrix} \Lambda_{X_1X_1} & \Lambda_{X_1X_2} \\ \Lambda_{X_2X_1} & \Lambda_{X_2X_2} \end{bmatrix}$$

38

Covariance

$$\Lambda_{XX} = \Lambda_{XX}^T \succeq 0$$

Proof:

- Define any deterministic vector $v \in \mathcal{R}^2$ $\|v\| \neq 0$
- $Q = (X-m_X)^T v$ is a scalar random variable.

$$v^T \Lambda_{XX} v = E\{\underbrace{v^T (X-m_X)}_{Q}\underbrace{(X-m_X)^T v}_{Q}\}$$

$$= E\{Q^2\} \geq 0$$

■

39

Random Vectors

X be a random n vector

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \in \mathcal{R}^n$$

with PDF

$$p_X(x) = \frac{\partial^n F_X(x)}{\partial x_1 \cdots \partial x_n}$$

$$p_X : \mathcal{R}^n \to \mathcal{R}_+$$

Y be a random m vector

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_m \end{bmatrix} \in \mathcal{R}^m$$

with PDF

$$p_Y(x) = \frac{\partial^m F_Y(x)}{\partial x_1 \cdots \partial x_m}$$

$$p_Y : \mathcal{R}^m \to \mathcal{R}_+$$

40

Cross-covariance

X be a random n vector $\quad$ Y be a random m vector

$$\Lambda_{XY} = E\{(X-m_X)(Y-m_Y)^T\} \in R^{n\times m}$$

$$= E\left\{\begin{bmatrix} X_1 - m_{X_1} \\ \vdots \\ X_n - m_{X_n} \end{bmatrix} \begin{bmatrix} Y_1 - m_{Y_1} & \cdots & Y_m - m_{Y_m} \end{bmatrix}\right\}$$

$$= \begin{bmatrix} \Lambda_{X_1Y_1} & \cdots & \Lambda_{X_1Y_m} \\ \vdots & & \vdots \\ \Lambda_{X_nY_1} & \cdots & \Lambda_{X_nY_m} \end{bmatrix} = \Lambda_{YX}^T$$

41

Cauchy-Schwarz inequality

For any scalar random variables X and Y

$$\Lambda_{XY}^2 \leq \Lambda_{XX}\Lambda_{YY}$$

42

Proof

Define the random vector $Z = \begin{bmatrix} X \\ Y \end{bmatrix} \in \mathcal{R}^2$

$$\Lambda_{ZZ} = \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix} \succeq 0$$

Thus,

$$\mathrm{Det}[\Lambda_{ZZ}] = \Lambda_{XX}\Lambda_{YY} - \Lambda_{XY}^2 \geq 0 \quad \blacksquare$$

43

Gaussian Random Variables (Review)

Let X be Gaussian with PDF

$$p_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}$$

Frequently-used notation

$$\boxed{X \sim N(m_X, \sigma_X^2)}$$

X is normally distributed with

mean m_X

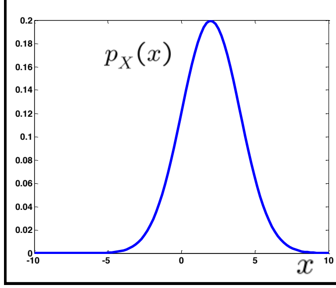
and variance $\sigma_X^2 = \Lambda_{XX}$

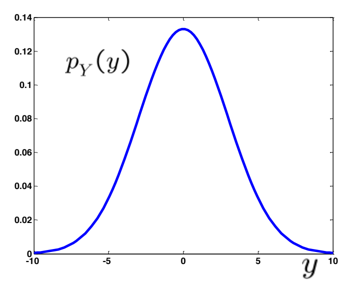
44

Two independent Gaussians

$X \sim N(m_X, \sigma_X^2)$ $\qquad$ $Y \sim N(m_Y, \sigma_Y^2)$

$\sigma_X = 2 \quad m_X = 2$ $\qquad$ $\sigma_Y = 3 \quad m_Y = 0$



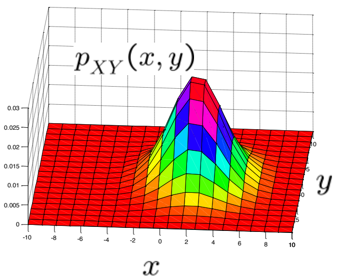


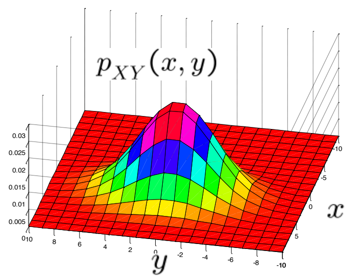
46

Two independent Gaussians

$$p_{XY}(x,y) = p_X(x)p_Y(y)$$

$\sigma_X = 2 \quad m_X = 2 \qquad \sigma_Y = 3 \quad m_Y = 0$





52

2-dimensional Gaussian random vector

$$Z = \begin{bmatrix} X \\ Y \end{bmatrix} \qquad m_Z = \begin{bmatrix} m_X \\ m_Y \end{bmatrix}$$

X and Y independent

$$\Lambda_{ZZ} = \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \end{bmatrix}$$

$$p_Z(z) = p_{XY}(x,y) = p_X(x)p_Y(y)$$

$$= \cdots = \frac{1}{2\pi\,|\Lambda_{ZZ}|^{\frac{1}{2}}}\, e^{-\frac{1}{2}(z-m_Z)^T \Lambda_{ZZ}^{-1}(z-m_Z)}$$

53

n-dimensional Gaussian random vector

Joint PDF of a Gaussian vector

$$Z = \begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix}$$

$$Z \sim N(m_Z, \Lambda_{ZZ})$$

$$p_Z(z) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Lambda_{ZZ}|^{\frac{1}{2}}}\, e^{-\frac{1}{2}(z-m_Z)^T \Lambda_{ZZ}^{-1}(z-m_Z)}$$

n: dimension of Z

54

Linear combination of Gaussians

If X is Gaussian and

$$Z = AX + b$$

where

- A is a deterministic matrix
- b is a deterministic vector

then Z is also Gaussian

Conditional PDF (Review)

55

Let X and Y have a joint PDF $p_{XY}(x,y)$

- The ***Conditional*** PDF of X given an outcome of $Y = y_1$:

$$p_{X|y_1}(x) = \frac{p_{XY}(x,y_1)}{p_Y(y_1)}$$

Conditional Expectation (Review)

56

Let X and Y have a joint PDF $p_{XY}(x,y)$

- Conditional Expectation of X given an outcome of $Y = y_1$:

$$m_{X|y_1} = E\{X|y_1\}$$

$$= \int_{-\infty}^{\infty} x\, p_{X|y_1}(x)dx$$

Motivation for Gaussians

57

When X and Y are Gaussians

The conditional probabilities $p_{X|y}(x)$

and conditional expectations $m_{X|y}$
(for any outcome y)

can be calculated very easily!

Random Vectors

58

Define the Gaussian random $n + m$ vector

$$Z = \begin{bmatrix} X \\ Y \end{bmatrix} \sim N(m_Z, \Lambda_{ZZ})$$

X is Gaussian n vector Y is a Gaussian m vector

$$m_Z = \begin{bmatrix} m_X \\ m_Y \end{bmatrix} \qquad \Lambda_{ZZ} = \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix}$$

59

Random Vectors

X is Gaussian n vector $\qquad$ Y is a Gaussian m vector

$$m_X = E\{X\} \qquad m_Y = E\{Y\}$$

$$\Lambda_{XX} = E\{(X - m_X)(X - m_X)^T\} \qquad (n \times n)$$

$$\Lambda_{YY} = E\{(Y - m_Y)(Y - m_Y)^T\} \qquad (m \times m)$$

$$\Lambda_{XY} = E\{(X - m_X)(Y - m_Y)^T\} \qquad (n \times m)$$

60

Conditional PDF for Gaussians

- The conditional PDF of X given $Y = y$

$$p_{X|y}(x) = \frac{p_{XY}(x,y)}{p_Y(y)}$$

$$\boxed{p_{X|y}(x) = \frac{1}{(2\pi)^{n/2}\sqrt{|\Lambda_{X|yX|y}|}} e^{-\frac{1}{2}(x - m_{X|y})^T \Lambda_{X|yX|y}^{-1}(x - m_{X|y})}}$$

also a Gaussian PDF

61

Conditional PDF for Gaussians

The conditional random vector X given and *outcome* $Y = y$

$$\boxed{X|y \sim N(m_{X|y}, \Lambda_{X|yX|y})}$$

is also normally distributed
(also a Gaussian random vector)

62

Conditional PDF for Gaussians

$$p_{X|y}(x) = \frac{1}{(2\pi)^{n/2}\sqrt{|\Lambda_{X|yX|y}|}} e^{-\frac{1}{2}(x - m_{X|y})^T \Lambda_{X|yX|y}^{-1}(x - m_{X|y})}$$

$$X|y \sim N(m_{X|y}, \Lambda_{X|yX|y})$$

$$m_{X|y} = m_X + \Lambda_{XY}\Lambda_{YY}^{-1}(y - m_Y)$$

conditional expectation of X given $Y = y$

affine function of the outcome *y*

Conditional PDF for Gaussians

$$p_{X|y}(x) = \frac{1}{(2\pi)^{n/2}\sqrt{|\Lambda_{X|yX|y}|}} e^{-\frac{1}{2}(x-m_{X|y})^T \Lambda_{X|yX|y}^{-1}(x-m_{X|y})}$$

$$X|y \sim N(m_{X|y}, \Lambda_{X|yX|y})$$

$$\Lambda_{X|yX|y} = \Lambda_{XX} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YX}$$

The conditional covariance of X given $Y = y$

independent of the outcome y !!

Conditional covariance of X given $Y = y$

$$\Lambda_{X|yX|y} = E\{(x - m_{X|y})(x - m_{X|y})^T |_{Y=y}\}$$

$$= \Lambda_{XX} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YX}$$

$$E\{(X - m_X)(X - m_X)^T\}$$

$$\lambda_{max}\left[\Lambda_{X|yX|y}\right] \leq \lambda_{max}\left[\Lambda_{XX}\right] - \lambda_{min}\left[\Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YX}\right]$$

max eigenvalues *min eigenvalue*

Independent Gaussians

Let X and Y be jointly Gaussian random vectors.

X and Y are independent if and only if they are uncorrelated

Proof:

$(\Rightarrow)$ We already showed this this is true even if X and Y are not jointly Gaussian

$(\Leftarrow)$ $X|y \sim N(m_{X|y},\ \Lambda_{X|yX|y})$

$$m_{X|y} = m_X + \cancel{\Lambda_{XY}}\Lambda_{YY}^{-1}(y - m_Y) \quad = m_X$$

$$\Lambda_{X|yX|y} = \Lambda_{XX} - \cancel{\Lambda_{XY}}\Lambda_{YY}^{-1}\Lambda_{YX} \quad = \Lambda_{XX}$$

$$\Rightarrow \quad X|y \sim N(m_X,\ \Lambda_{XX}) \quad \Rightarrow \quad p_{X|y}(x) = p_X(x) \quad \blacksquare$$

Proof of conditional PDF for Gaussians

Idea of proof

- Some details regarding Schur complements
- A lot of algebra…

67

Schur complement

- Given

$$M = \begin{bmatrix} A & D \\ C & B \end{bmatrix}$$

- Schur complement of B:

$$\Delta = A - DB^{-1}C$$

- Then

$$|M| = \det\left(\begin{bmatrix} A & D \\ C & B \end{bmatrix}\right) = |B|\,|\Delta|$$

68

Schur complement

- Given

$$M = \begin{bmatrix} A & D \\ C & B \end{bmatrix}$$

- If Schur complement of B

$$\Delta = A - DB^{-1}C$$

is nonsingular

- Then

$$M^{-1} = \begin{bmatrix} \Delta^{-1} & -\Delta^{-1}F \\ -E\Delta^{-1} & B^{-1} + E\Delta^{-1}F \end{bmatrix}$$

$$E = B^{-1}C \qquad F = DB^{-1}$$

69

Proof

- Given

$$M = \begin{bmatrix} A & D \\ C & B \end{bmatrix}$$

- Define

$$Q = \begin{bmatrix} I & 0 \\ \underbrace{-B^{-1}C}_{E} & B^{-1} \end{bmatrix}$$

- Then

$$MQ = \begin{bmatrix} \underbrace{A - DB^{-1}C}_{\Delta} & \underbrace{DB^{-1}}_{F} \\ 0 & I \end{bmatrix} = R$$

- Results follow by computing inverses and determinants of matrices Q and R

70

details

$$R = \begin{bmatrix} \Delta & F \\ 0 & I \end{bmatrix} \Rightarrow R^{-1} = \begin{bmatrix} \Delta^{-1} & -\Delta^{-1}F \\ 0 & I \end{bmatrix} \qquad Q = \begin{bmatrix} I & 0 \\ -E & B^{-1} \end{bmatrix}$$

$$M = RQ^{-1} \Rightarrow M^{-1} = Q\,R^{-1}$$

$$M^{-1} = \begin{bmatrix} I & 0 \\ -E & B^{-1} \end{bmatrix} \begin{bmatrix} \Delta^{-1} & -\Delta^{-1}F \\ 0 & I \end{bmatrix}$$

$$M^{-1} = \begin{bmatrix} \Delta^{-1} & -\Delta^{-1}F \\ -E\Delta^{-1} & B^{-1} + E\Delta^{-1}F \end{bmatrix} \qquad E = B^{-1}C \quad F = DB^{-1}$$

71

Conditional covariance $\Lambda_{X|yX|y}$

- Given

$$\Lambda_{ZZ} = \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix}$$

- The Schur complement of Λ_{YY}

$$\Delta = \Lambda_{XX} - \Lambda_{XY}\,\Lambda_{YY}^{-1}\Lambda_{YX}$$

$$= \Lambda_{X|yX|y}$$

72

Schur complement of Λ_{YY}

- Given

$$\Lambda_{ZZ} = \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix} \qquad \Delta = \Lambda_{XX} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YX}$$

- Then

$$|\Lambda_{ZZ}| = \det\left(\begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix}\right) = |\Lambda_{YY}|\,|\Delta|$$

$$\Delta = \Lambda_{X|yX|y}$$

73

Schur complement of Λ_{YY}

- Given

$$\Lambda_{ZZ} = \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix} \qquad \Delta = \Lambda_{XX} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YX}$$

- and

$$\Lambda_{ZZ}^{-1} = \begin{bmatrix} \Delta^{-1} & -\Delta^{-1}F \\ -F^T\Delta^{-1} & \Lambda_{YY}^{-1} + F^T\Delta^{-1}F \end{bmatrix}$$

$$\Delta = \Lambda_{X|yX|y} \qquad F = \Lambda_{XY}\Lambda_{YY}^{-1}$$

74

Theorem

Given $\begin{bmatrix} X \\ Y \end{bmatrix} \sim N(\begin{bmatrix} m_X \\ m_Y \end{bmatrix}, \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix})$

Then $X|y \sim N(m_{X|y}, \Lambda_{X|yX|y})$

with

$$m_{X|y} = m_X + \Lambda_{XY}\,\Lambda_{YY}^{-1}\,(y - m_Y)$$

$$\Lambda_{X|yX|y} = \Lambda_{XX} - \Lambda_{XY}\,\Lambda_{YY}^{-1}\Lambda_{YX}$$

Proof

- Random vector

$$Z = \begin{bmatrix} X \\ Y \end{bmatrix} \sim N(\underbrace{\begin{bmatrix} m_X \\ m_Y \end{bmatrix}}_{m_Z}, \underbrace{\begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix}}_{\Lambda_{ZZ}})$$

•dummy variables

$$\tilde{z} = z - m_Z \quad = \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} x - m_X \\ y - m_Y \end{bmatrix}$$

Proof: use Schur complement

- Now compute:

$$\tilde{z}^T \Lambda_{ZZ}^{-1} \tilde{z} \quad = \begin{bmatrix} \tilde{x}^T & \tilde{y}^T \end{bmatrix} \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix}$$

- Using:

$$\Lambda_{ZZ}^{-1} = \begin{bmatrix} \Delta^{-1} & -\Delta^{-1}F \\ -F^T\Delta^{-1} & \Lambda_{YY}^{-1} + F^T\Delta^{-1}F \end{bmatrix}$$

$$\Delta = \Lambda_{XX} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YX} \qquad F = \Lambda_{XY}\Lambda_{YY}^{-1}$$

Proof

- Now compute:

$$\tilde{z}^T \Lambda_{ZZ}^{-1} \tilde{z} \quad = \begin{bmatrix} \tilde{x}^T & \tilde{y}^T \end{bmatrix} \begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix}$$

$$= (\tilde{x} - F\tilde{y})^T \Delta^{-1} (\tilde{x} - F\tilde{y})$$

$$+ \tilde{y}^T \Lambda_{YY}^{-1} \tilde{y}$$

$$\Delta = \Lambda_{XX} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YX} \qquad F = \Lambda_{XY}\Lambda_{YY}^{-1}$$

Proof: compute the conditional PDF

$$p_{X|y}(x) = \frac{p_{XY}(x,y)}{p_Y(y)} \quad = \frac{p_Z(x,y)}{p_Y(y)}$$

where:

$$p_Y(y) \quad = \quad \frac{1}{(2\pi)^{\frac{m}{2}} |\Lambda_{YY}|^{\frac{1}{2}}} exp\left(-\frac{1}{2}\tilde{y}^T \Lambda_{YY}^{-1} \tilde{y}\right)$$

dimension of Y

$$\tilde{y} = y - m_Y$$

79

Proof: compute the conditional PDF

$$p_{X|y}(x) = \frac{p_{XY}(x,y)}{p_Y(y)} \quad = \frac{p_Z(x,y)}{p_Y(y)}$$

where:

$$p_Z(z) \;=\; \frac{1}{(2\pi)^{\frac{n+m}{2}}\, |\Lambda_{ZZ}|^{\frac{1}{2}}}\, exp\left(-\frac{1}{2}\, \tilde{z}^T\, \Lambda_{ZZ}^{-1}\, \tilde{z}\right)$$

dimension of X + dimension of Y

$$\tilde{z} = \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} x - m_X \\ y - m_Y \end{bmatrix}$$

80

Proof

$$p_{X|y}(x) = \frac{p_{XY}(x,y)}{p_Y(y)}$$

$$= \frac{(2\pi)^{\frac{m}{2}} |\Lambda_{YY}|^{\frac{1}{2}}}{(2\pi)^{\frac{n+m}{2}}\, |\Lambda_{ZZ}|^{\frac{1}{2}}}$$

$$exp\left(-\frac{1}{2}\, \underbrace{\tilde{z}^T\, \Lambda_{ZZ}^{-1}\, \tilde{z}}\, - \frac{1}{2}\, \tilde{y}^T\, \Lambda_{YY}^{-1}\, \tilde{y}\right)$$

$$\tilde{z}^T\, \Lambda_{ZZ}^{-1}\, \tilde{z}^T \quad = (\tilde{x} - F\tilde{y})^T\, \Delta^{-1} (\tilde{x} - F\tilde{y}) \quad + \tilde{y}^T\, \Lambda_{YY}^{-1}\, \tilde{y}$$

81

Proof

$$p_{X|y}(x) = \frac{p_{XY}(x,y)}{p_Y(y)}$$

$$= \frac{|\Lambda_{YY}|^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}\, |\Lambda_{ZZ}|^{\frac{1}{2}}}$$

$$exp\left[-\frac{1}{2}\, (\tilde{x} - F\tilde{y})^T\, \Delta^{-1} (\tilde{x} - F\tilde{y})\right]$$

$$\Delta = \Lambda_{XX} - \Lambda_{XY}\Lambda_{YY}^{-1}\Lambda_{YX} \qquad\qquad F = \Lambda_{XY}\Lambda_{YY}^{-1}$$

82

Proof

$$p_{X|y}(x) = \frac{|\Lambda_{YY}|^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}\, |\Lambda_{ZZ}|^{\frac{1}{2}}}$$

$$exp\left[-\frac{1}{2}\, (\tilde{x} - F\tilde{y})^T\, \Delta^{-1} (\tilde{x} - F\tilde{y})\right]$$

use Schur determinant result:

$$|\Lambda_{ZZ}| = \det\left(\begin{bmatrix} \Lambda_{XX} & \Lambda_{XY} \\ \Lambda_{YX} & \Lambda_{YY} \end{bmatrix}\right) = |\Lambda_{YY}|\, |\Delta|$$

Proof

83

$$p_{X|y}(x) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Delta|^{\frac{1}{2}}}$$

$$exp\left[-\frac{1}{2}(\tilde{x} - F\tilde{y})^T \Delta^{-1} (\tilde{x} - F\tilde{y})\right]$$

Now use:

$$\Lambda_{X|yX|y} = \Delta = \Lambda_{XX} - \Lambda_{XY} \Lambda_{YY}^{-1} \Lambda_{YX}$$

Proof

84

$$p_{X|y}(x) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Lambda_{X|yX|y}|^{\frac{1}{2}}}$$

$$exp\left[-\frac{1}{2}(\tilde{x} - F\tilde{y})^T \Lambda_{X|yX|y}^{-1} (\tilde{x} - F\tilde{y})\right]$$

Now use: $F = \Lambda_{XY} \Lambda_{YY}^{-1}$ $\quad \tilde{x} = x - m_X$

$$\tilde{x} - F\tilde{y} = x - \underbrace{m_X - \Lambda_{XY} \Lambda_{YY}^{-1} \tilde{y}}_{m_{X|y}} = x - m_{X|y}$$

Proof

85

$$p_{X|y}(x) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Lambda_{X|yX|y}|^{\frac{1}{2}}}$$

$$exp\left[-\frac{1}{2}(\tilde{x} - F\tilde{y})^T \Lambda_{X|yX|y}^{-1} (\tilde{x} - F\tilde{y})\right]$$

Therefore,

$$X|y \sim N(m_{X|y}, \Lambda_{X|yX|y})$$

Proof

86

Therefore,

$$X|y \sim N(m_{X|y}, \Lambda_{X|yX|y})$$

with

$$m_{X|y} = m_X + \Lambda_{XY} \Lambda_{YY}^{-1} (y - m_Y)$$

$$\Lambda_{X|yX|y} = \Lambda_{XX} - \Lambda_{XY} \Lambda_{YY}^{-1} \Lambda_{YX} \qquad \blacksquare$$

This result is important and constitutes the basis for the Kalman Filter!

Supplemental Material (You are not responsible for this…)

- Laplace and Fourier transform of Gaussian PDF
- Transformation of random variables

Laplace transform of normal PDF

$$p_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}}$$

$$P_X(s) = \int_{-\infty}^{\infty} e^{-sx} p_X(x)\, dx = \frac{1}{\sigma_X\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-sx} e^{-\frac{(x-m_X)^2}{2\sigma_X^2}} dx$$

$$= \frac{1}{\sigma_X\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-A(x)} dx$$

where, after "completing the squares",

$$A(x) = sx + \frac{x^2}{2\sigma_X^2} + \frac{m_X^2}{2\sigma_X^2} - \frac{2m_X x}{2\sigma_X^2}$$

$$= \frac{1}{2\sigma_X^2}\left\{\left[x + (s\sigma_X^2 - m_X)\right]^2 - s^2\sigma_X^4 + 2m_X s\sigma_X^2\right\}$$

Laplace transform of normal PDF

substituting,

$$P_X(s) = e^{(s^2\sigma_X^2/2)-sm_X} \underbrace{\int_{-\infty}^{\infty} \left\{\frac{1}{\sqrt{2\pi}\sigma_X} e^{-(x+s\sigma_X^2-m_X)^2/2\sigma_X^2}\right\} dx}_{=1 \text{ (area under a PDF = 1)}}$$

$$\boxed{P_X(s) = e^{(s^2\sigma_X^2/2)-sm_X}}$$

Fourier transform: $P_X(j\omega) = e^{\frac{-\omega^2\sigma_X^2}{2}} e^{-j\omega m_X}$

Transformation of random variables

Given a real valued function f of random variable X

$$Y = f(X)$$

Assume that Y is also a random variable.

Also assume that $g(\cdot) = f^{-1}(\cdot)$ exists. Then,

$$\boxed{p_Y(y_o) = p_X(g(y_o)) \left|\frac{dg(y_o)}{dy}\right|}$$

Transformation of random variables

Let $y_o = f(x_o)$ and $x_o = g(y_o)$

$$P(x_o \leq X \leq x_o + dx) = P(y_o \leq Y \leq y_o + dy)$$

$$\int_{x_o}^{x_o+dx} p_X(x)dx = \begin{cases} \int_{y_o}^{y_o+dy} p_Y(y)dy & dy > 0 \\ -\int_{y_o}^{y_o+dy} p_Y(y)dy & dy < 0 \end{cases}$$

$$p_Y(y_o) = p_X(x_o) \left|\frac{dx}{dy}\right|_{x=x_o} = p_X(g(y_o)) \left|\frac{dg(y_o)}{dy}\right|$$