Chapter 4: Multiple Random Variables[1]

Yunghsiang S. Han

Graduate Institute of Communication Engineering,
National Taipei University
Taiwan

E-mail: yshan@mail.ntpu.edu.tw

[1]Modified from the lecture notes by Prof. Mao-Ching Chiu

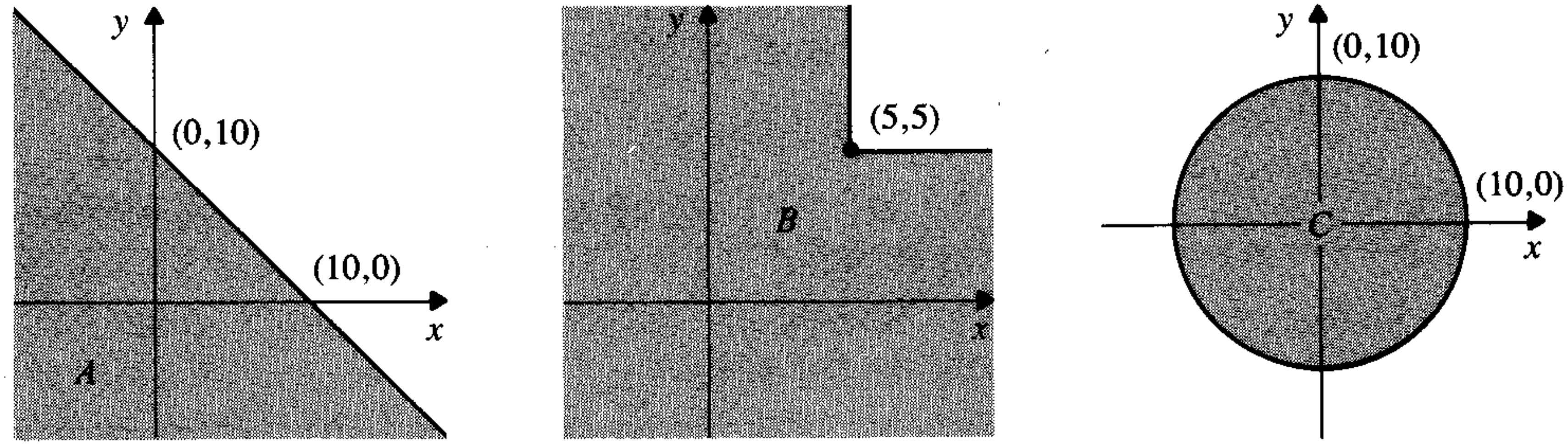
4.1 Vector Random Variables

Consider the two dimensional random variable $\boldsymbol{X} = (X, Y)$. Find the regions of the planes corresponding to the events

$A = \{X + Y \leq 10\}$,

$B = \{min(X, Y) \leq 5\}$ and

$C = \{X^2 + Y^2 \leq 100\}$.


y
(0,10)
(10,0)
x
A
y
(5,5)
B
x
y
(0,10)
(10,0)
C
x

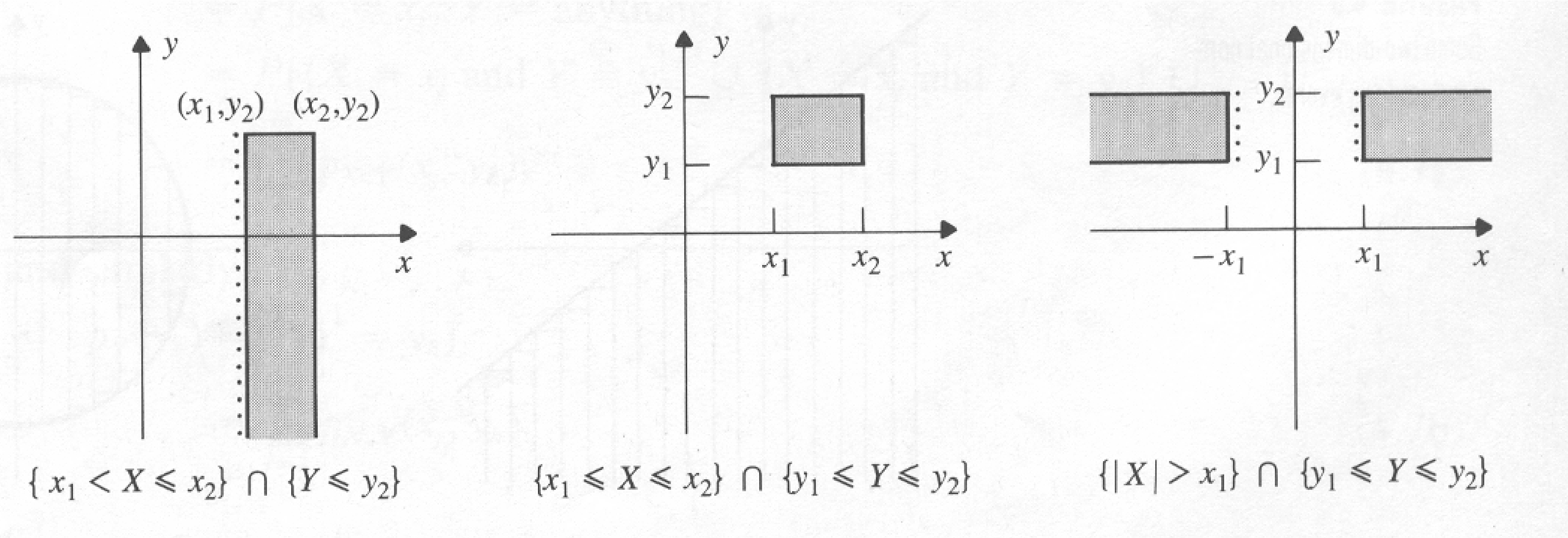
- Let the n-dimensional random variable $\boldsymbol{X}$ be $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ and A_k be a one dimensional event that involves X_k.
- Events with **product form** is defined as

$$A = \{X_1 \in A_1\} \cap \{X_2 \in A_2\} \cap \cdots \cap \{X_n \in A_n\}.$$

$$\begin{aligned} P[A] &= P[\{X_1 \in A_1\} \cap \{X_2 \in A_2\} \cap \cdots \cap \{X_n \in A_n\}] \\ &\triangleq P[X_1 \in A_1, \ldots, X_n \in A_n]. \end{aligned}$$

- Some events may not be of product form.

Some two-dimensional product form events



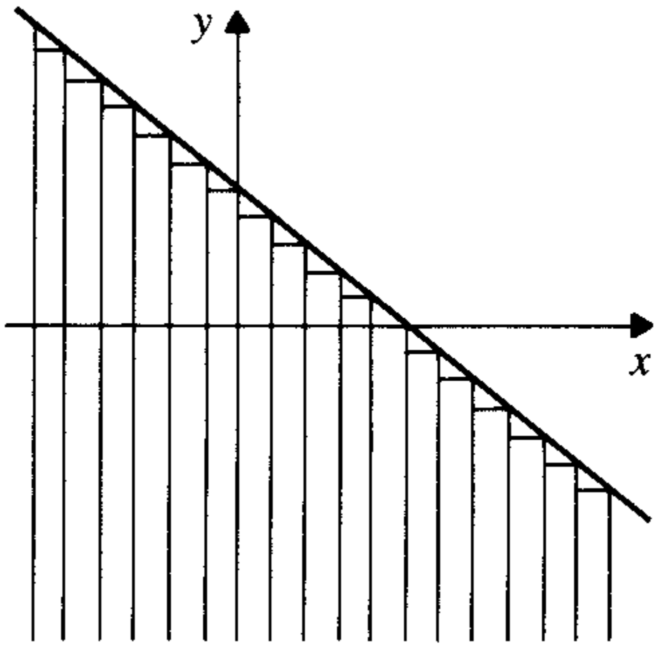
Probability of non-product-form event

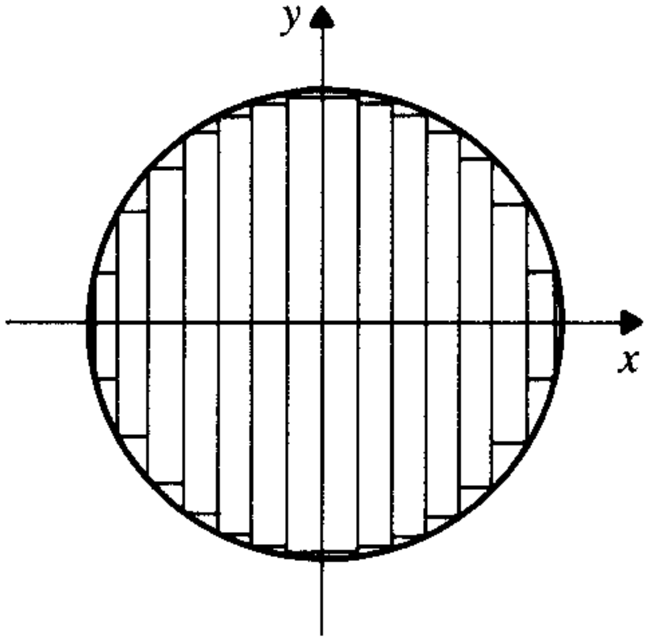
- B is partitioned into disjoint product-form events such as $B_1, B_2, \ldots, B_n$, and

$$P[B] \approx P\left[\bigcup_k B_k\right] = \sum_k P[B_k].$$

- Approximation becomes exact as B_k's become arbitrary fine.

Non-product-form events





Independence

- Two random variables X and Y are independent if

$$P[X \in A_1, Y \in A_2] = P[X \in A_1]P[Y \in A_2].$$

- Random variables $X_1, X_2, \ldots, X_n$ are independent if

$$P[X_1 \in A_1, \ldots, X_n \in A_n] = P[X_1 \in A_1] \cdots P[X_n \in A_n].$$

4.2 Pairs of Random Variables

Pairs of Discrete Random Variables

- Random variable $\boldsymbol{X} = (X, Y)$
- Sample space $S = \{(x_j, y_k) : j = 1, 2, \ldots, k = 1, 2, \ldots\}$ is countable.
- **Joint probability mass function (pmf)** of $\boldsymbol{X}$ is

$$\begin{aligned} & p_{X,Y}(x_j, y_k) \\ = \;& P[\{X = x_j\} \cap \{Y = y_k\}] \\ \triangleq \;& P[X = x_j, Y = y_k] \qquad j = 1, 2, \ldots \qquad k = 1, 2, \ldots \end{aligned}$$

- Probability of event A is

$$P[\boldsymbol{X} \in A] = \sum_{(x_j, y_k) \in A} p_{X,Y}(x_j, y_k).$$

- **Marginal probability mass function** is

$$\begin{aligned} p_X(x_j) &= P[X = x_j] \\ &= P[X = x_j, Y = \text{anything}] \\ &= P[\{X = x_j \text{ and } Y = y_1\} \cup \\ &\quad \{X = x_j \text{ and } Y = y_2\} \cup \cdots] \\ &= \sum_{k=1}^{\infty} p_{X,Y}(x_j, y_k). \end{aligned}$$

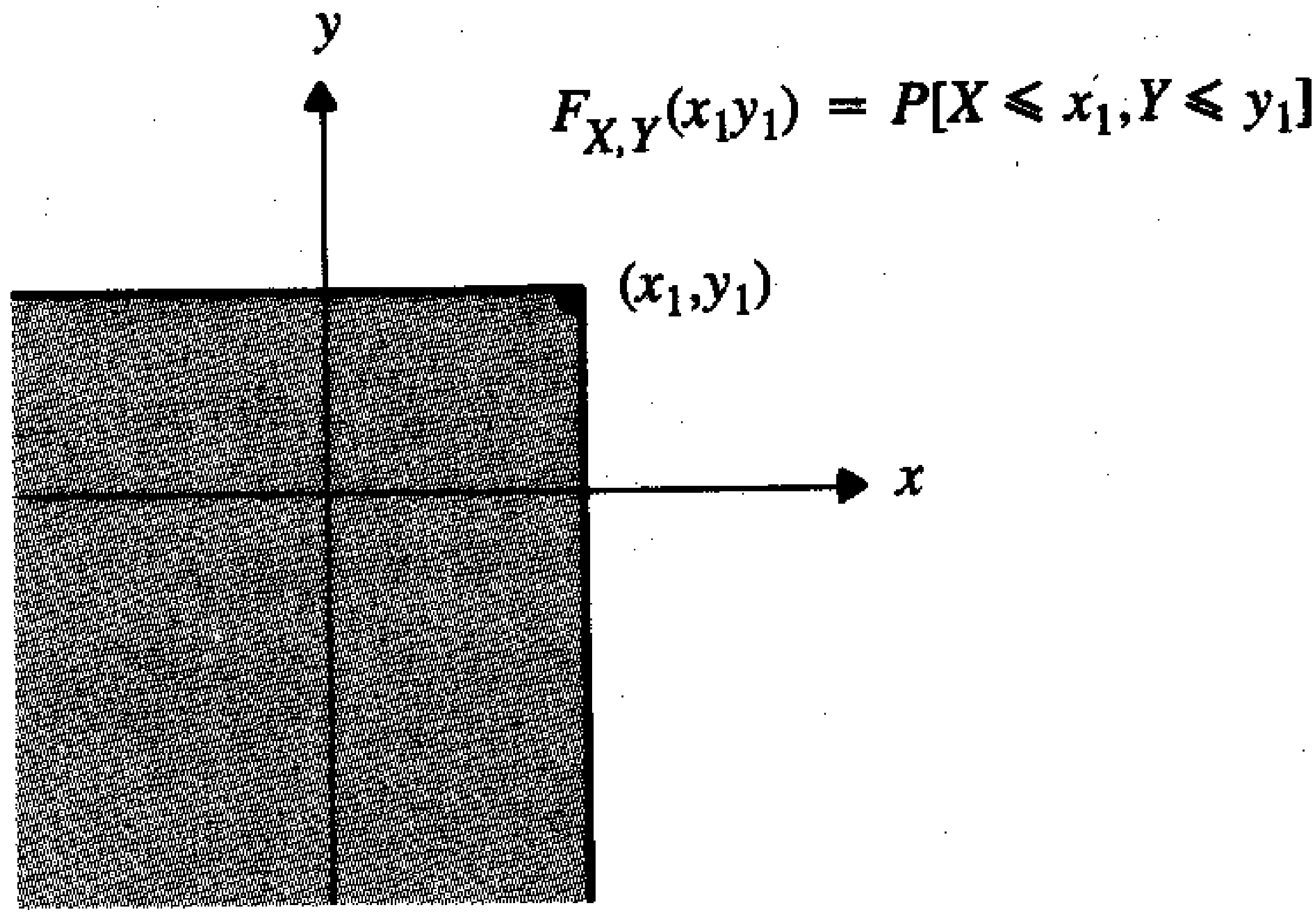
- Similarly

$$p_Y(y_k) = \sum_{j=1}^{\infty} p_{X,Y}(x_j, y_k).$$

Joint cdf of X and Y

- Joint cumulative distribution function of X and Y is given as

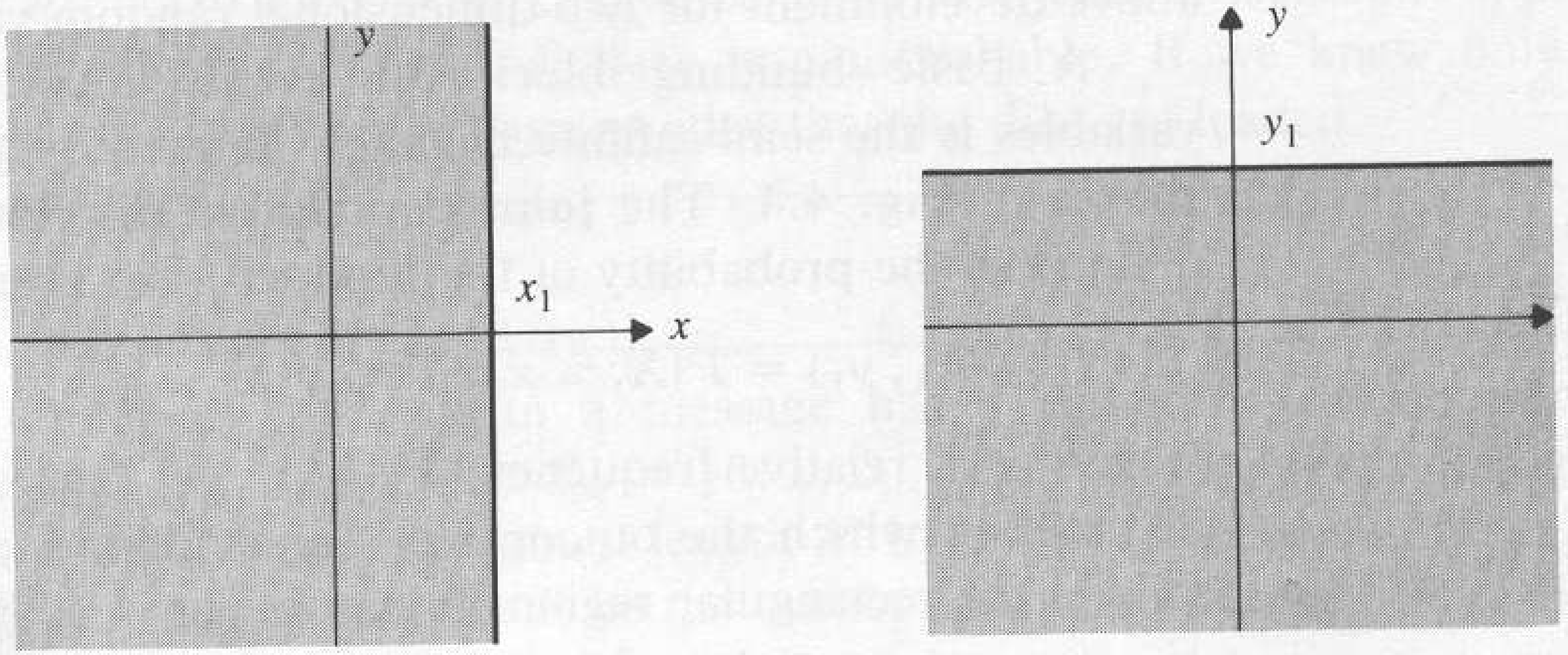
$$F_{X,Y}(x_1, y_1) = P[X \leq x_1, Y \leq y_1]$$


y
$F_{X,Y}(x_1 y_1) = P[X \leqslant x_1, Y \leqslant y_1]$
(x_1, y_1)
x

Properties

1. $F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_2)$, if $x_1 \leq x_2$ and $y_1 \leq y_2$.
2. $F_{X,Y}(-\infty, y_1) = F_{X,Y}(x_1, -\infty) = 0$.
3. $F_{X,Y}(\infty, \infty) = 1$.
4. $F_X(x) = F_{X,Y}(x, \infty) = P[X \leq x, Y < \infty] = P[X \leq x]$;
   $F_Y(y) = F_{X,Y}(\infty, y) = P[Y \leq y]$.
5. Continuous from the right

$$\begin{aligned} \lim_{x \to a^+} F_{X,Y}(x, y) &= F_{X,Y}(a, y) \\ \lim_{y \to b^+} F_{X,Y}(x, y) &= F_{X,Y}(x, b) \end{aligned}$$



$F_X(x_1) = P[X \leqslant x_1, Y < \infty]$

$F_Y(y_1) = P[X < \infty, Y \leqslant y_1]$

Example: Joint cdf of $\boldsymbol{X} = (X, Y)$ is given as

$$F_{X,Y}(x,y) = \begin{cases} (1-e^{-\alpha x})(1-e^{-\beta y}) & x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases}.$$

Find the marginal cdf's.

Sol:

$$F_X(x) = \lim_{y\to\infty} F_{X,Y}(x,y) = 1 - e^{-\alpha x} \qquad x \geq 0.$$

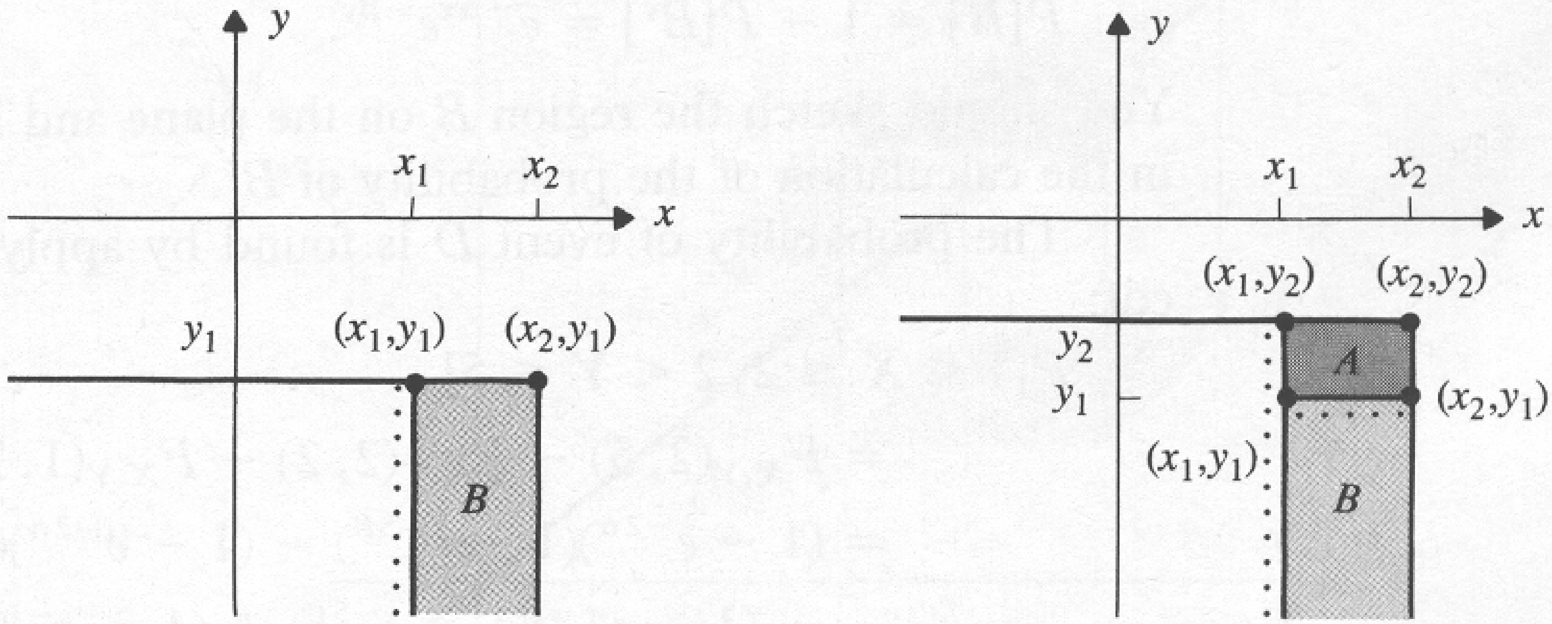
$$F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x,y) = 1 - e^{-\beta y} \qquad y \geq 0.$$

- Probability of region $B = \{x_1 < X < x_2, Y \leq y_1\}$
  $F_{X,Y}(x_2, y_1) = F_{X,Y}(x_1, y_1) + P[x_1 < X < x_2, Y \leq y_1]$
  $\rightarrow\ P[x_1 < X < x_2, Y \leq y_1] = F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_1)$

- Probability of region $A = \{x_1 < X \leq x_2, y_1 < Y \leq y_2\}$

$$
\begin{aligned}
F_{X,Y}(x_2, y_2) &= P[x_1 < X \leq x_2, y_1 < Y \leq y_2] \\
&+ F_{X,Y}(x_2, y_1) + F_{X,Y}(x_1, y_2) - F_{X,Y}(x_1, y_1)
\end{aligned}
$$

$$
\begin{aligned}
&P[x_1 < X \leq x_2, y_1 < Y \leq y_2] \\
&= F_{X,Y}(x_2, y_2) - F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_2) + F_{X,Y}(x_1, y_1)
\end{aligned}
$$

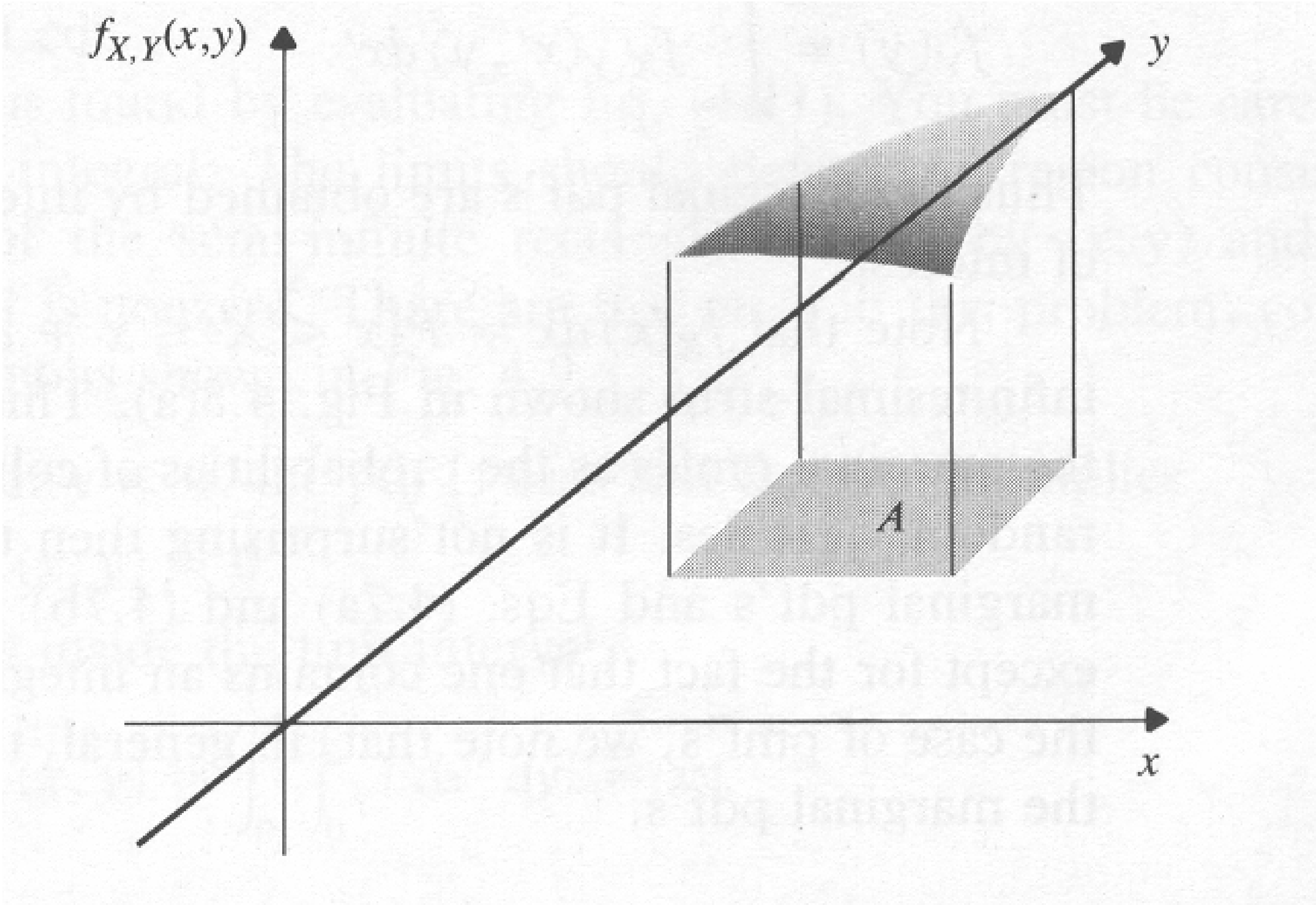


(a) (b)

Joint pdf of Two Jointly Continuous Random Variables

- Random variable $\boldsymbol{X} = (X, Y)$
- Joint probability density function $f_{X,Y}(x, y)$ is defined such that for every event A

$$P[\boldsymbol{X} \in A] = \int\int_A f_{X,Y}(x', y')dx'dy'.$$


$f_{X,Y}(x,y)$
y
A
x

Properties

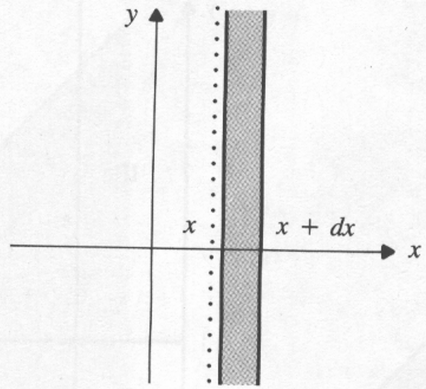
1. $1 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x', y') dx' dy'$.
2. $F_{X,Y}(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(x', y') dx' dy'$.
3. $f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y}$.
4. $P[a_1 < X \leq b_1, a_2 < Y \leq b_2] = \int_{a_2}^{b_2} \int_{a_1}^{b_1} f_{X,Y}(x', y') dx' dy'$.
5. Marginal pdf's

   $$f_X(x) = \frac{d}{dx} F_X(x) = \frac{d}{dx} F_{X,Y}(x, \infty)$$

$$= \frac{d}{dx}\int_{-\infty}^{x}\left\{\int_{-\infty}^{+\infty} f_{X,Y}(x', y')dy'\right\}dx'$$

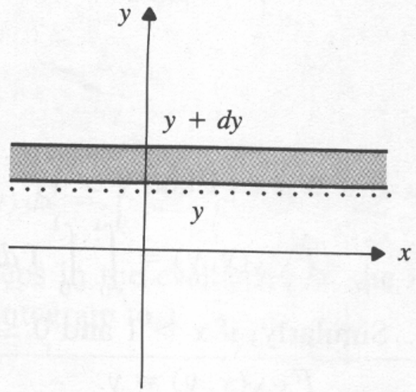
$$= \int_{-\infty}^{+\infty} f_{X,Y}(x, y')dy'.$$

6. $$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x', y)dx'.$$



$f_X(x)dx \cong P[x < X \leqslant x + dx, Y < \infty]$

(a)



$f_Y(y)dy \cong P[X < \infty, y < Y \leqslant y + dy]$

(b)

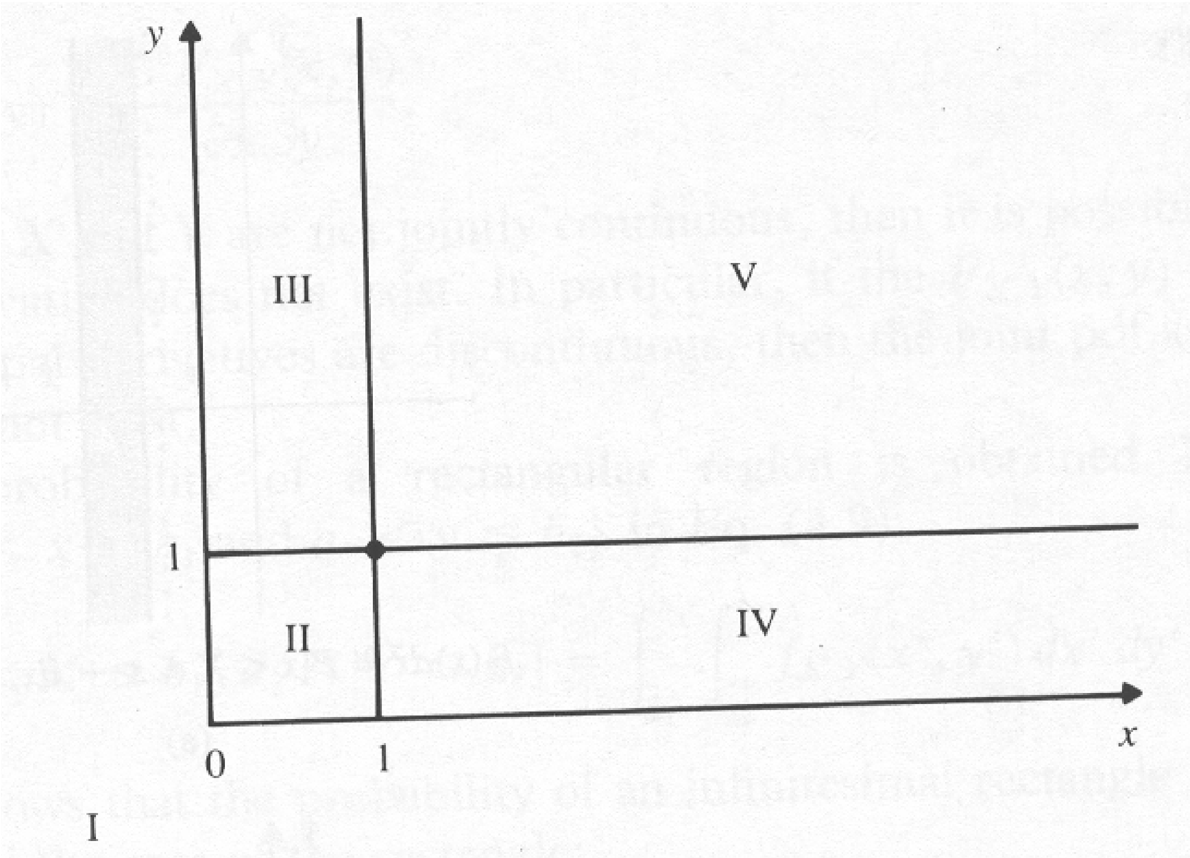
Example: Let the pdf of $\boldsymbol{X} = (X, Y)$ be

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

Find the joint cdf.

Sol: Consider five cases:

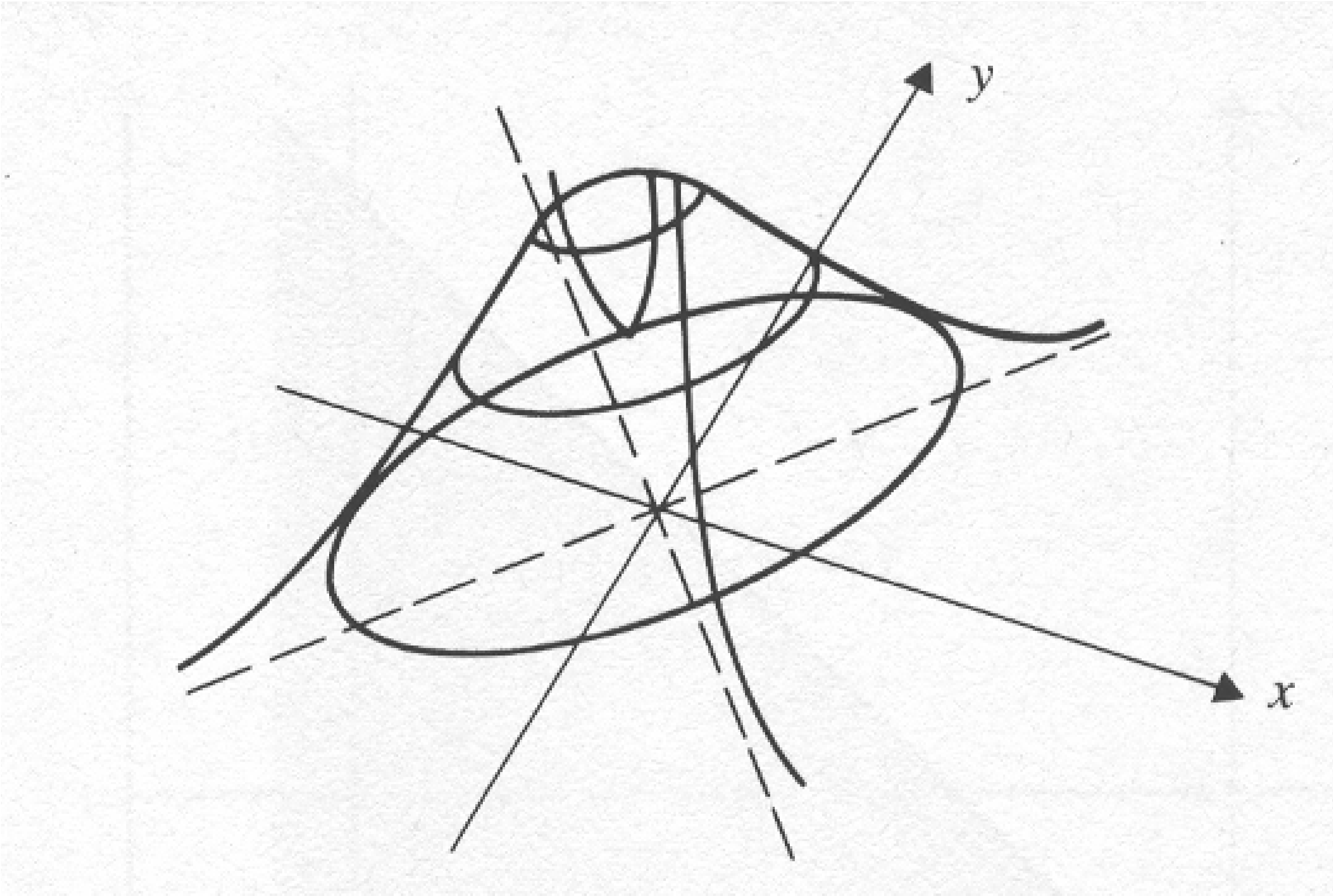
1. $x < 0$ or $y < 0$, $F_{X,Y}(x, y) = 0$;
2. $(x, y) \in$ unit interval, $F_{X,Y}(x, y) = \int_0^y \int_0^x 1 dx' dy' = xy$;
3. $0 \le x \le 1$ and $y > 1$, $F_{X,Y}(x, y) = \int_0^1 \int_0^x 1 dx' dy' = x$;
4. $x > 1$ and $0 \le y \le 1$, $F_{X,Y}(x, y) = y$;
5. $x > 1$ and $y > 1$, $F_{X,Y}(x, y) = \int_0^1 \int_0^1 1 dx' dy' = 1$.


y
III
V
1
II
IV
0
1
x
I

Example: Random variables X and Y are jointly Gaussian

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-(x^2-2\rho xy+y^2)/2(1-\rho^2)} \quad -\infty < x, y < \infty.$$

Find the marginal pdf's.


y
x

- Marginal pdf of X

$$f_X(x) = \frac{e^{-x^2/2(1-\rho^2)}}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} e^{-(y^2-2\rho xy)/2(1-\rho^2)} dy$$

- Add and subtract $\rho^2 x^2$ in the exponent, i.e., $y^2 - 2\rho xy + \rho^2 x^2 - \rho^2 x^2 = (y - \rho x)^2 - \rho^2 x^2$.

$$\begin{aligned} f_X(x) &= \frac{e^{-x^2/2(1-\rho^2)}}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} e^{-[(y-\rho x)^2 - \rho^2 x^2]/2(1-\rho^2)} dy \\ &= \frac{e^{-x^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \underbrace{\frac{e^{-(y-\rho x)^2/2(1-\rho^2)}}{\sqrt{2\pi(1-\rho^2)}}}_{N(\rho x; 1-\rho^2)} dy \end{aligned}$$

$$= \frac{e^{-x^2/2}}{\sqrt{2\pi}} \quad \rightarrow \quad N(0,1).$$

Example: Let X be the input to a communication channel and Y the output. The input to the channel is $+1$ volt or -1 volt with equal probability. The output of the channel is the input plus a noise voltage N that is uniformly distributed in the interval $[-2, +2]$ volts. Find $P[X = +1, Y \leq 0]$.

Sol:

$$P[X = +1, Y \leq y] = P[Y \leq y | X = +1] P[X = +1],$$

where $P[X = +1] = 1/2$. When the input $X = 1$, the output Y is uniformly distributed in the interval $[-1, 3]$. Therefore,

$$P[Y \leq y | X = +1] = \frac{y+1}{4} \quad \text{for } -1 \leq y \leq 3.$$

Thus

$$
\begin{aligned}
P[X = +1, Y \leq 0] &= P[Y \leq 0 | X = +1] P[X = +1] \\
&= (1/4)(1/2) = 1/8.
\end{aligned}
$$

4.3 Independence of Two Random Variables

- X and Y are independent random variables if for every events A_1 and A_2

$$P[X \in A_1, Y \in A_2] = P[X \in A_1]P[Y \in A_2]$$

- Suppose X and Y are discrete random variables. We are interesting in the probability of event $A = A_1 \cap A_2$. Let $A_1 = \{X = x_j\}$ and $A_2 = \{Y = y_k\}$, then the independence of X and Y implies

$$\begin{aligned} p_{X,Y}(x_j, y_k) &= P[X = x_j, Y = y_k] \\ &= P[X = x_j]P[Y = y_k] \end{aligned}$$

$$= \quad p_X(x_j)p_Y(y_k)$$

$\rightarrow$ joint pmf is equal to the product of the marginal pmf's.

- Let X and Y be random variables with $p_{X,Y}(x_j, y_k) = p_X(x_j)p_Y(y_k)$. Let $A = A_1 \cap A_2$.

$$\begin{aligned} P[A] &= \sum_{x_j \in A_1} \sum_{y_k \in A_2} p_{X,Y}(x_j, y_k) \\ &= \sum_{x_j \in A_1} \sum_{y_k \in A_2} p_X(x_j)p_Y(y_k) \\ &= \sum_{x_j \in A_1} p_X(x_j) \sum_{y_k \in A_2} p_Y(y_k) \\ &= P[A_1]P[A_2] \end{aligned}$$

- **Discrete random variables X and Y are independent if and only if the joint pmf is equal to the product of the marginal pmf's for all**

x_j, y_k.

- Random variables X and Y are independent if and only if

$$F_{X,Y}(x,y) = F_X(x)F_Y(y) \qquad \text{for all } x \text{ and } y.$$

- If X and Y are jointly continuous, then X and Y are independent if and only if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) \qquad \text{for all } x \text{ and } y.$$

- If X and Y are independent random variables, then $g(X)$ and $h(Y)$ are also independent.

  Proof: Let A and B are any two events involve $g(X)$ and $h(Y)$, respectively. Define

$$A' = \{x : g(x) \in A\} \quad \text{and} \quad B' = \{y : h(y) \in B\}.$$

  Then

$$\begin{aligned} P[g(X) \in A, h(Y) \in B] &= P[X \in A', Y \in B'] \\ &= P[X \in A']P[Y \in B'] \\ &= P[g(X) \in A]P[h(Y) \in B]. \end{aligned}$$

4.4 Conditional Probability & Conditional Expectation

Conditional Probability

- Probability of $Y \in A$ given that the exact value of X is known as

$$P[Y \in A | X = x] = \frac{P[Y \in A, X = x]}{P[X = x]}.$$

- Conditional cdf of Y given $X = x_k$ is

$$F_Y(y|x_k) = \frac{P[Y \leq y, X = x_k]}{P[X = x_k]}, \quad \text{for } P[X = x_k] > 0.$$

- Conditional pdf of Y given $X = x_k$ is

$$f_Y(y|x_k) = \frac{d}{dy} F_Y(y|x_k).$$

- Probability of event A given $X = x_k$ is

$$P[Y \in A | X = x_k] = \int_{y \in A} f_Y(y|x_k) dy.$$

- If X and Y are independent, then $F_Y(y|x) = F_Y(y)$ and $f_Y(y|x) = f_Y(y)$.

- If X and Y are discrete, then conditional pdf will consist of delta functions with probability mass given by the conditional pmf of Y given $X = x_k$:

$$\begin{aligned} p_Y(y_j|x_k) &= P[Y = y_j|X = x_k] \\ &= \frac{P[X = x_k, Y = y_j]}{P[X = x_k]} \\ &= \frac{p_{X,Y}(x_k, y_j)}{p_X(x_k)}. \end{aligned}$$

- If X and Y are discrete, the probability of any event A given $X = x_k$ is

$$P[Y \in A|X = x_k] = \sum_{y_j \in A} p_Y(y_j|x_k).$$

Example: Let X be the input to a communication channel and let Y be the output. The input to the channel is $+1$ volt or -1 volt with equal probability. The output of the channel is the input plus a noise voltage N that is uniformly distributed in the interval $[-2, +2]$ volts. Find the probability that Y is negative given that X is $+1$.

Sol: If $X = +1$, then Y is uniformly distributed in the interval $[-1, 3]$ and

$$f_Y(y|1) = \begin{cases} \frac{1}{4} & -1 \le y \le 3 \\ 0 & \text{elsewhere} \end{cases} .$$

Thus

$$P[Y < 0|X = +1] = \int_{-1}^{0} \frac{dy}{4} = \frac{1}{4}.$$

Continuous Random Variables

- If X is a continuous random variable, then $P[X = x] = 0$.
- Conditional cdf of Y given $X = x$ is

$$F_Y(y|x) = \lim_{h \to 0} F_Y(y|x < X \leq x + h).$$

$$\begin{aligned} F_Y(y|x < X \leq x + h) &= \frac{P[Y \leq y, x < X \leq x + h]}{P[x < X \leq x + h]} \\ &= \frac{\int_{-\infty}^{y} \int_{x}^{x+h} f_{X,Y}(x', y') dx' \, dy'}{\int_{x}^{x+h} f_X(x') dx'} \end{aligned}$$

$$\approx \quad \frac{\int_{-\infty}^{y} f_{X,Y}(x,y')dy'h}{f_X(x)h}.$$

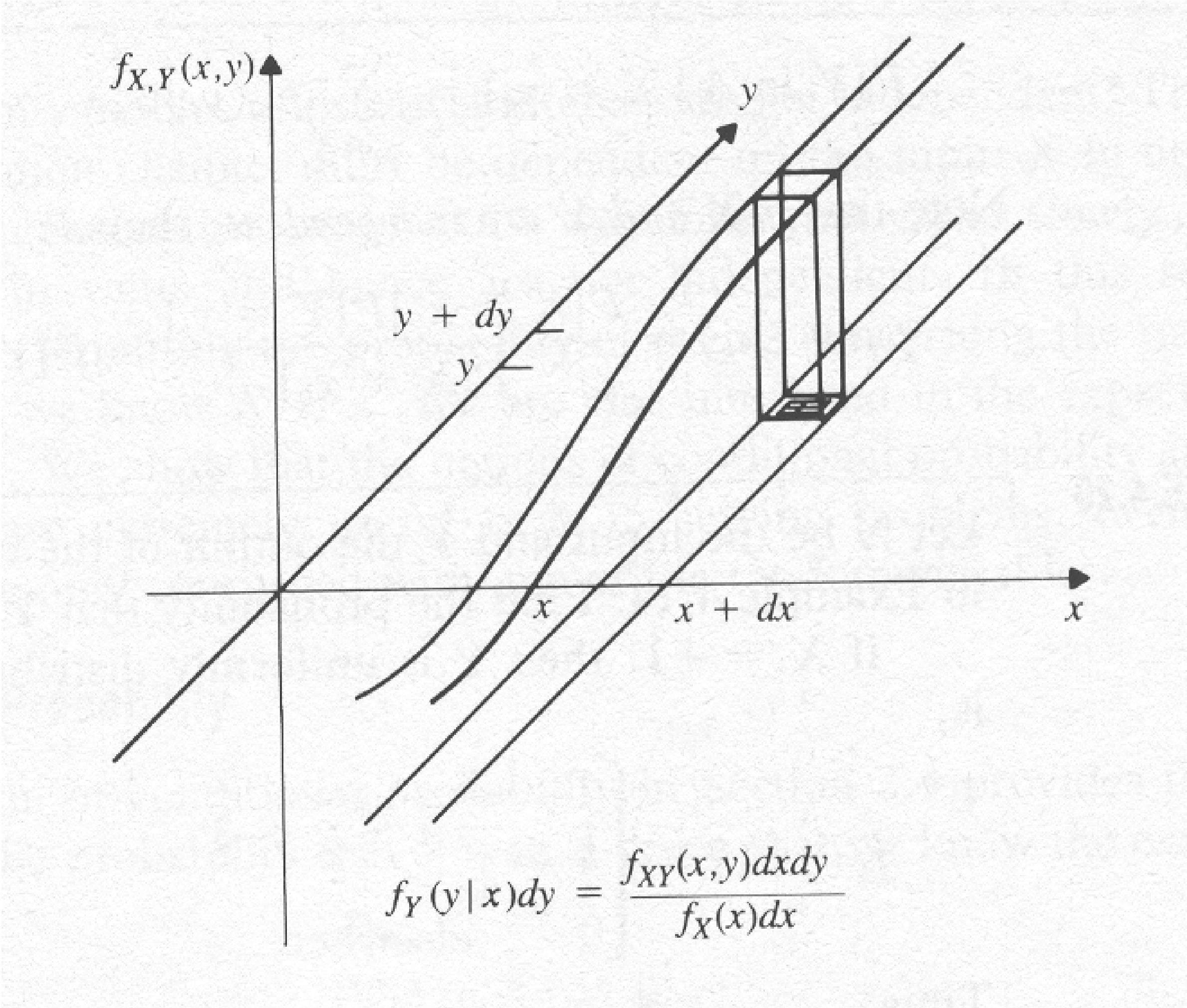
- As h approach zero

$$F_Y(y|x) = \frac{\int_{-\infty}^{y} f_{X,Y}(x,y')dy'}{f_X(x)}.$$

- Conditional pdf of Y given $X = x$ is

$$f_Y(y|x) = \frac{d}{dy}F_Y(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}.$$

- If X and Y are independent, then $f_{X,Y}(x,y) = f_X(x)f_Y(y)$, $f_Y(y|x) = f_Y(y)$, and $F_Y(y|x) = F_Y(y)$.


$f_{X,Y}(x,y)$
y
$y + dy$
y
x
$x + dx$
x
$f_Y(y|x)dy = \frac{f_{XY}(x,y)dxdy}{f_X(x)dx}$

Example: Let X and Y be random variables with joint pdf

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-y} & 0 \le y \le x < \infty \\ 0 & \text{elsewhere} \end{cases}.$$

Find $f_X(x|y)$ and $f_Y(y|x)$.
Sol: The marginal pdfs of X and Y are

$$f_X(x) = \int_0^{\infty} f_{X,Y}(x,y)dy = \int_0^{x} 2e^{-x}e^{-y}dy = 2e^{-x}(1-e^{-x}) \quad 0 \le x < \infty$$

$$f_Y(y) = \int_0^{\infty} f_{X,Y}(x,y)dx = \int_y^{\infty} 2e^{-x}e^{-y}dx = 2e^{-2y} \quad 0 \le y < \infty$$

$$f_X(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{2e^{-x}e^{-y}}{2e^{-2y}} = e^{-(x-y)} \quad \text{for } x \ge y$$

$$f_Y(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{2e^{-x}e^{-y}}{2e^{-x}(1-e^{-x})} \quad \text{for } 0 < y < x$$

- The relation of joint probability and conditional probability for discrete random variables X and Y are

$$\begin{aligned} P[X = x_k, Y = y_j] &= P[Y = y_j | X = x_k] P[X = x_k] \\ p_{X,Y}(x, y) &= p_Y(y|x) p_X(x) \end{aligned}$$

- Suppose that we are interested in the probability of $Y \in A$. Then

$$\begin{aligned} P[Y \in A] &= \sum_{\text{all } x_k} \sum_{y_j \in A} p_{X,Y}(x_k, y_j) \\ &= \sum_{\text{all } x_k} \sum_{y_j \in A} p_Y(y_j | x_k) p_X(x_k) \end{aligned}$$

$$= \sum_{\text{all } x_k} p_X(x_k) \sum_{y_j \in A} p_Y(y_j | x_k).$$

Thus,

$$P[Y \in A] = \sum_{\text{all } x_k} P[Y \in A | X = x_k] p_X(x_k).$$

- If X and Y are continuous, then

$$f_{X,Y}(x,y) = f_Y(y|x)f_X(x).$$

- Probability of $Y \in A$ is

$$P[Y \in A] = \int_{-\infty}^{+\infty} P[Y \in A|X = x]f_X(x)dx.$$

Example: The random variable X is selected at random from the unit interval; the random variable Y is then selected at random from the interval $(0, X)$. Find the cdf of Y.

Sol: We have

$$F_Y(y) = P[Y \leq y] = \int_0^1 P[Y \leq y | X = x] f_X(x) dx.$$

When $X = x$, Y is uniformly distributed in $(0, x)$. Thus,

$$P[Y \leq y | X = x] = \begin{cases} \frac{y}{x} & 0 \leq y \leq x \\ 1 & x \leq y \end{cases}$$

and

$$F_Y(y) = \int_0^y 1 dx' + \int_y^1 \frac{y}{x'} dx' = y - y \ln y.$$

The pdf of Y is then

$$f_Y(y) = -\ln y \qquad 0 \leq y \leq 1.$$

Conditional Expectation

- Conditional expectation of Y given $X = x$ is

$$E[Y|x] = \int_{-\infty}^{+\infty} y f_Y(y|x) dy.$$

- For discrete random variables, we have

$$E[Y|x] = \sum_{y_j} y_j p_Y(y_j|x).$$

- Define a function $g(x) = E[Y|x]$.
- $g(X)$ is a random variable.
- Consider $E[g(X)] = E[E[Y|X]]$. Then, We have

$$E[Y] = E[E[Y|X]],$$

where

$$\begin{aligned} E[E[Y|X]] &= \int_{-\infty}^{+\infty} E[Y|x] f_X(x) dx \quad \text{when } X \text{ is continuous;} \\ E[E[Y|X]] &= \sum_{x_k} E[Y|x_k] p_X(x_k) \quad \text{when } X \text{ is discrete.} \end{aligned}$$

- For continuous random variables,

$$\begin{aligned} E[E[Y|X]] &= \int_{-\infty}^{+\infty} E[Y|x] f_X(x) dx \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f_Y(y|x) dy f_X(x) dx \\ &= \int_{-\infty}^{+\infty} y \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx dy \\ &= \int_{-\infty}^{+\infty} y f_Y(y) dy = E[Y]. \end{aligned}$$

- The expected value of a function $h(Y)$ of Y is

$$E[h(Y)] = E[E[h(Y)|X]].$$

4.5 Multiple Random Variables

- Let $X_1, X_2, \ldots, X_n$ be an n-dimensional vector random variable.
- Joint cdf of $X_1, X_2, \ldots, X_n$ is

$$F_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = P[X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n].$$

- Joint cdf of $X_1, X_2, \ldots, X_{n-1}$ is

$$F_{X_1,X_2,\ldots,X_{n-1}}(x_1, x_2, \ldots, x_{n-1}) = F_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_{n-1}, \infty).$$

Example: Let event A be defined as follows:

$$A = \{\max(X_1, X_2, X_3) \leq 5\}.$$

Find the probability of A.

Sol: $\max(X_1, X_2, X_3) \leq 5$ if and only if each of the three numbers is less than 5; therefore

$$\begin{aligned} P[A] &= P[\{X_1 \leq 5\} \cap \{X_2 \leq 5\} \cap \{X_3 \leq 5\}] \\ &= F_{X_1,X_2,X_3}(5, 5, 5). \end{aligned}$$

- Joint probability mass function of n discrete random variables is

$$p_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = P[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n].$$

- Probability of event A is

$$P[(X_1,\ldots,X_n) \in A] = \sum \cdots \sum_{\boldsymbol{x} \in A} p_{X_1,X_2,\ldots,X_n}(x_1,\ldots,x_n),$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$.

- Marginal pmf for X_j is

$$p_{X_j}(x_j) = \sum_{x_1} \cdots \sum_{x_{j-1}} \sum_{x_{j+1}} \cdots \sum_{x_n} p_{X_1,X_2,\ldots,X_n}(x_1,\ldots,x_n).$$

- Marginal pmf for $X_1, X_2, \ldots, X_{n-1}$ is

$$p_{X_1,X_2,\ldots,X_{n-1}}(x_1, x_2, \ldots, x_{n-1}) = \sum_{x_n} p_{X_1,X_2,\ldots,X_n}(x_1, \ldots, x_n).$$

- Conditional pmf is

$$p_{X_n}(x_n|x_1,\ldots,x_{n-1}) = \frac{p_{X_1,\ldots,X_n}(x_1,\ldots,x_n)}{p_{X_1,\ldots,X_{n-1}}(x_1,\ldots,x_{n-1})}.$$

-

$$\begin{aligned}
& p_{X_1,\ldots,X_n}(x_1,\ldots,x_n) \\
&= p_{X_n}(x_n|x_1,\ldots,x_{n-1}) \\
&\quad \times p_{X_{n-1}}(x_{n-1}|x_1,\ldots,x_{n-2}) \cdots p_{X_2}(x_2|x_1) p_{X_1}(x_1).
\end{aligned}$$

Example: A computer system receives message over three communications lines. Let X_j be the number of messages received on line j in one hour. Suppose that the joint pmf of X_1, X_2, and X_3 is given by

$$p_{X_1,X_2,X_3}(x_1,x_2,x_3) = (1-a_1)(1-a_2)(1-a_3)a_1^{x_1}a_2^{x_2}a_3^{x_3} \quad \text{for } x_1 \geq 0, x_2 \geq 0, x_3 \geq 0.$$

Find $p_{X_1,X_2}(x_1,x_2)$ and $p_{X_1}(x_1)$ given that $0 < a_i < 1$.

Sol: The marginal pmf of X_1 and X_2 is

$$\begin{aligned} p_{X_1,X_2}(x_1,x_2) &= (1-a_1)(1-a_2)(1-a_3)\sum_{x_3=0}^{\infty} a_1^{x_1}a_2^{x_2}a_3^{x_3} \\ &= (1-a_1)(1-a_2)a_1^{x_1}a_2^{x_2}. \end{aligned}$$

The pmf of X_1 is

$$\begin{aligned} p_{X_1}(x_1) &= (1-a_1)(1-a_2)\sum_{x_2=0}^{\infty} a_1^{x_1} a_2^{x_2} \\ &= (1-a_1)a_1^{x_1}. \end{aligned}$$

- Random variables $X_1, X_2, \ldots, X_n$ are jointly continuous random variables if the probability of any n-dimensional event A is given by an n-dimensional integral of a probability density function:

$$P[(X_1, \ldots, X_n) \in A] = \int \cdots \int_{\boldsymbol{x} \in A} f_{X_1,\ldots,X_n}(x'_1, \ldots, x'_n) dx' \ldots dx'_n,$$

where $f_{X_1,\ldots,X_n}(x'_1, \ldots, x'_n)$ is the **joint probability density function**.

- Joint cdf of $\boldsymbol{X}$ is obtained from the joint pdf:

$$\begin{aligned} & F_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) \\ = & \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{X_1,\ldots,X_n}(x'_1, \ldots, x'_n) dx' \ldots dx'_n. \end{aligned}$$

- Joint pdf (if the derivative exists) is given by

$$f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)=\frac{\partial^n}{\partial x_1\cdots\partial x_n}F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n).$$

- The marginal pdf for a subset of random variables is obtained by integrating the other variables out. For example, the marginal pdf of X_1 is

$$f_{X_1}(x_1) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_{X_1,X_2,\ldots,X_n}(x_1, x_2', \ldots, x_n') dx_2' \cdots dx_n'.$$

- The marginal pdf for $X_1, \ldots, X_{n-1}$ is given by

$$f_{X_1,\ldots,X_{n-1}}(x_1, \ldots, x_{n-1}) = \int_{-\infty}^{+\infty} f_{X_1,\ldots,X_n}(x_1, \ldots, x_{n-1}, x_n') dx_n'.$$

- Conditional pdf is given by

$$f_{X_n}(x_n|x_1,\ldots,x_{n-1}) = \frac{f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)}{f_{X_1,\ldots,X_{n-1}}(x_1,\ldots,x_{n-1})}.$$

-

$$\begin{aligned}&f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)\\&= f_{X_n}(x_n|x_1,\ldots,x_{n-1})\\&\quad\times f_{X_{n-1}}(x_{n-1}|x_1,\ldots,x_{n-2})\cdots f_{X_2}(x_2|x_1)f_{X_1}(x_1).\end{aligned}$$

Example: The random variables X_1, X_2, and X_3 have the joint Gaussian pdf:

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3) = \frac{e^{-(x_1^2+x_2^2-\sqrt{2}x_1x_2+\frac{1}{2}x_3^2)}}{2\pi\sqrt{\pi}}.$$

Find the marginal pdf of X_1 and X_3.

Sol: The marginal pdf for the pair X_1 and X_3 is

$$f_{X_1,X_3}(x_1,x_3) = \frac{e^{-x_3^2/2}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{e^{-(x_1^2+x_2^2-\sqrt{2}x_1x_2)}}{2\pi/\sqrt{2}} dx_2.$$

The above integral gives

$$f_{X_1,X_3}(x_1,x_3) = \frac{e^{-x_3^2/2}}{\sqrt{2\pi}} \frac{e^{-x_1^2/2}}{\sqrt{2\pi}}.$$

Therefore, X_1 and X_3 are independent zero-mean, unit-variance Gaussian random variables.

Independence

- $X_1, \ldots, X_n$ are independent if

$$P[X_1 \in A_1, \ldots, X_n \in A_n] = P[X_1 \in A_1] \ldots P[X_n \in A_n].$$

- $X_1, \ldots, X_n$ are independent if and only if

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n) \quad \forall\, x_1, \ldots, x_n.$$

- If the random variables are discrete, then the above equation is equivalent to

$$p_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = p_{X_1}(x_1) \ldots p_{X_n}(x_n) \quad \forall\, x_1, \ldots, x_n;$$

If the random variables are jointly continuous, then the above equation is equivalent to

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

Example: The n samples $X_1, X_2, \ldots, X_n$ of a "white noise" signal have joint pdf given by

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \frac{e^{-(x_1^2+\cdots+x_n^2)/2}}{(2\pi)^{n/2}} \quad \forall\ x_1,\ldots,x_n.$$

It is clear that the above is the product of n one-dimensional Gaussian pdf's. Thus $X_1,\ldots,X_n$ are independent Gaussian random variables.

4.6 Functions of Several Random Variables

One function of several random variables

- Let Z be defined as a function of several random variables:

$$Z = g(X_1, X_2, \ldots, X_n).$$

- The cdf of Z is $P[Z \leq z] = P[R_z = \{\boldsymbol{x} = (x_1, \ldots, x_n) : g(\boldsymbol{x}) \leq z\}]$, and

$$\begin{aligned} F_Z(z) &= P[\boldsymbol{X} \in R_z] \\ &= \int \cdots \int_{\boldsymbol{x} \in R_z} f_{X_1,\ldots,X_n}(x'_1, \ldots, x'_n) dx'_1 \ldots dx'_n. \end{aligned}$$

- The pdf of Z is then found by taking the derivative of $F_Z(z)$.

Example: Let $Z = X + Y$. Find $F_Z(z)$ and $f_Z(z)$ in terms of the joint pdf of X and Y.

Sol: The cdf of Z is

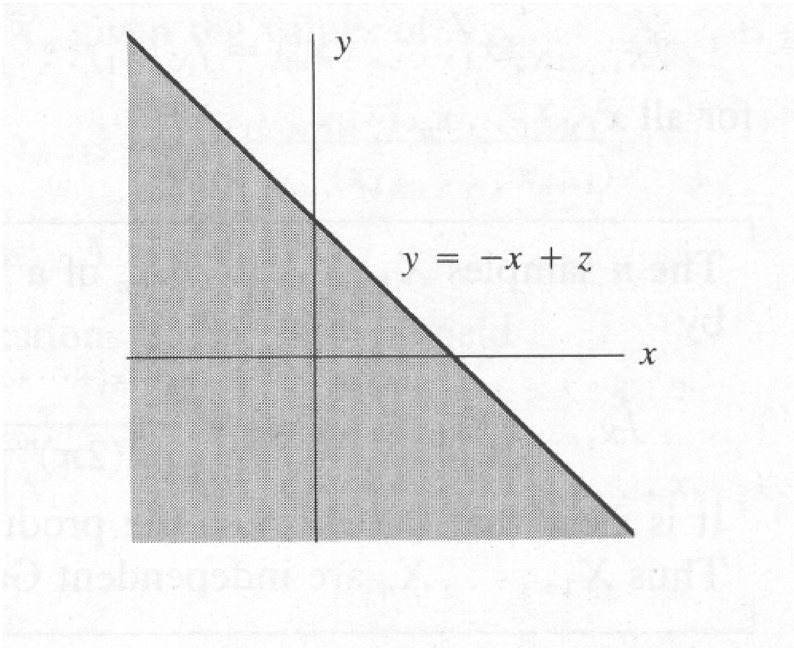
$$F_Z(z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{z-x'} f_{X,Y}(x', y') dy' dx'.$$

The pdf of Z is

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \int_{-\infty}^{+\infty} f_{X,Y}(x', z - x') dx'.$$

If X and Y are independent random variables, then

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x') f_Y(z - x') dx' \quad -- \quad \text{convolution integral.}$$


y
y = −x + z
x

Example: Find the pdf of the sum $Z = X + Y$ of two zero-mean, unit-variance Gaussian random variables with correlation coefficient $\rho = -1/2$.

Sol:We have

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{+\infty} f_{X,Y}(x', z - x')dx' \\ &= \frac{1}{2\pi(1-\rho^2)^{1/2}} \int_{-\infty}^{+\infty} e^{-[(x')^2 - 2\rho x'(z-x') + (z-x')^2]/2(1-\rho^2)} dx' \\ &= \frac{1}{2\pi(3/4)^{1/2}} \int_{-\infty}^{+\infty} e^{-((x')^2 - x'z + z^2)/2(3/4)} dx'. \end{aligned}$$

After completing the square of the argument in the exponent we have

$$f_Z(z) = \frac{e^{-z^2/2}}{\sqrt{2\pi}}.$$

Thus, the sum of two nonindependent Gaussian random variables is also a Gaussian random variable.

- Find pdf of a function from conditional pdf

$$f_Z(z) = \int_{-\infty}^{+\infty} f_Z(z|y')f_Y(y')dy'$$

Example: Let $Z = X/Y$. Find the pdf of Z if X and Y are independent and both exponential distributed with mean one.
Sol: Assume $Y = y$, then $Z = X/y$ is a scaled version of X. Therefore

$$f_Z(z|y) = |y|f_X(yz|y).$$

The pdf of Z is

$$f_Z(z) = \int_{-\infty}^{+\infty} |y'|f_X(y'z|y')f_Y(y')dy' = \int_{-\infty}^{+\infty} |y'|f_{X,Y}(y'z, y')dy'.$$

Since X and Y are independent and exponentially distributed with mean one, we have

$$\begin{aligned} f_Z(z) &= \int_0^\infty y' f_X(y'z) f_Y(y') dy' \qquad z > 0 \\ &= \int_0^\infty y' e^{-y'z} e^{-y'} dy' \\ &= \frac{1}{(1+z)^2} \qquad z > 0. \end{aligned}$$

Transformations of Random Variables

- Let $X_1, X_2, \ldots, X_n$ be random variables.
- Let random variables $Z_1, Z_2, \ldots, Z_n$ be defined as

$$Z_1 = g_1(\boldsymbol{X}), \quad Z_2 = g_2(\boldsymbol{X}), \quad \cdots, \quad Z_n = g_n(\boldsymbol{X})$$

- How to find the joint cdf and pdf of $Z_1, \ldots, Z_n$?
- Joint cdf of $Z_1, \ldots, Z_n$ is

$$F_{Z_1,\ldots,Z_n}(z_1, \ldots, z_n) = P[g_1(\boldsymbol{X}) \le z_1, \ldots, g_n(\boldsymbol{X}) \le z_n].$$

- If $X_1, \ldots, X_n$ have a joint pdf, then

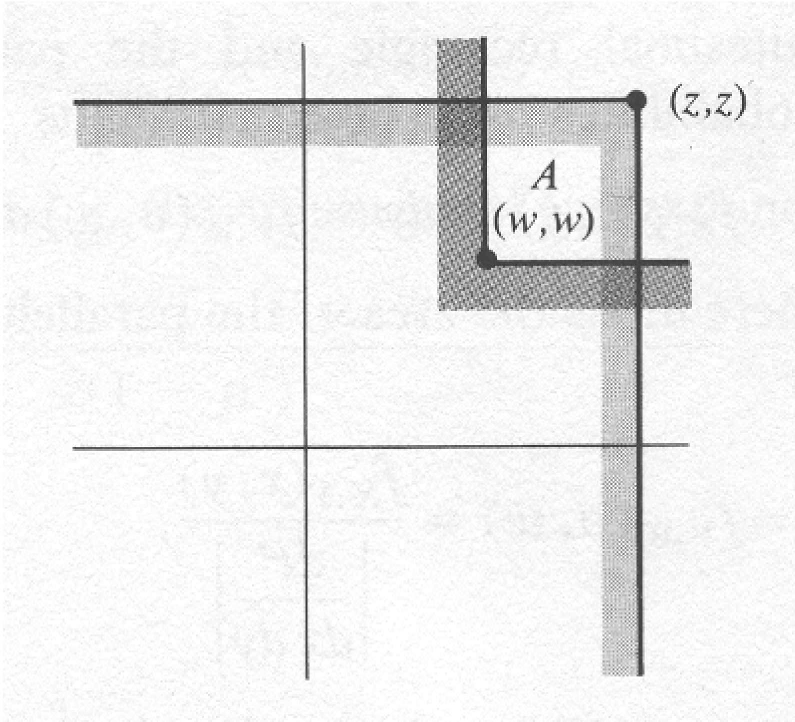
$$F_{Z_1,\ldots,Z_n}(z_1, \ldots, z_n) = \int \cdots \int_{\boldsymbol{x}':g_k(\boldsymbol{x}')\le z_k} f_{X_1,\ldots,X_n}(x_1', \ldots, x_n') dx_1' \cdots dx_n'.$$

Example: Let random variables W and Z be defined as

$$W = \min(X, Y) \text{ and } Z = \max(X, Y).$$

Find the joint cdf of W and Z in terms of the joint cdf of X and Y.
Sol:

$$F_{W,Z}(w, z) = P[\{\min(X, Y) \le w\} \cap \{\max(X, Y) \le z\}].$$



If $z > w$,

$$\begin{aligned} F_{W,Z}(w,z) &= F_{X,Y}(z,z) - P[A] \\ &= F_{X,Y}(z,z) \\ &\quad - \{F_{X,Y}(z,z) - F_{X,Y}(w,z) - F_{X,Y}(z,w) + F_{X,Y}(w,w)\} \end{aligned}$$

$$= \quad F_{X,Y}(w,z) + F_{X,Y}(z,w) - F_{X,Y}(w,w).$$

If $z \leq w$, then

$$F_{W,Z}(w,z) = F_{X,Y}(z,z).$$

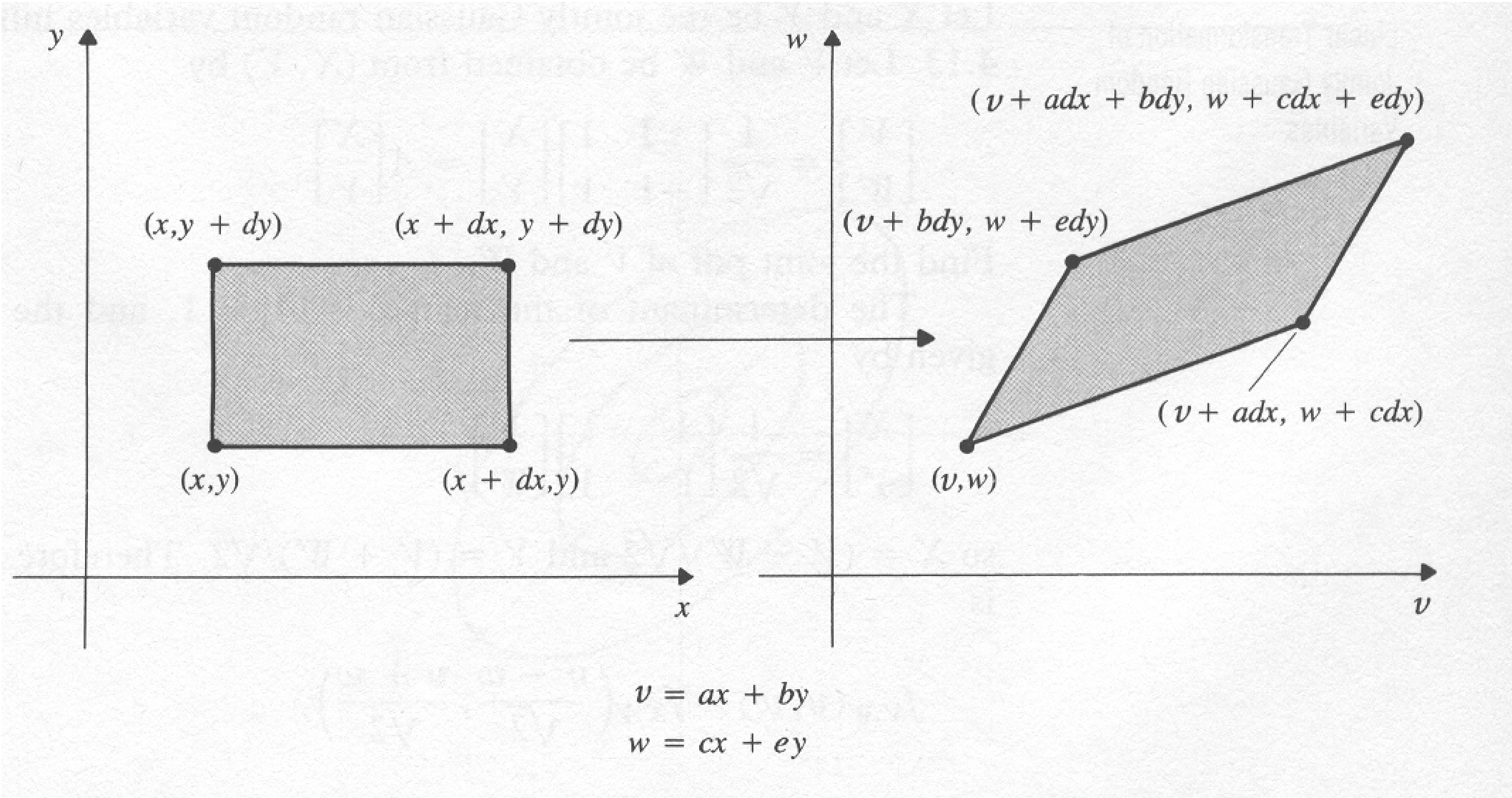
pdf of Linear Transformations

- Consider the **linear transformation** of two random variables:

$$V = aX + bY \qquad W = cX + eY \qquad \text{or} \qquad \begin{bmatrix} V \\ W \end{bmatrix} = \underbrace{\begin{bmatrix} a & b \\ c & e \end{bmatrix}}_{A} \begin{bmatrix} X \\ Y \end{bmatrix}$$

- Assume that A is invertible, that is,

$$\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1} \begin{bmatrix} v \\ w \end{bmatrix}.$$



- Rectangle → parallelogram

$$f_{X,Y}(x,y)dxdy \approx f_{V,W}(v,w)dP,$$

where dP is the area of the parallelogram.

- The joint pdf of V and W is

$$f_{V,W}(v,w) = \frac{f_{X,Y}(x,y)}{\left|\frac{dP}{dxdy}\right|}.$$

- $dP/dxdy$ is called "Stretch factor." It can be shown $dP = (|ae - bc|)dxdy$, so

$$\left|\frac{dP}{dxdy}\right| = \frac{|ae-bc|(dxdy)}{dxdy} = |ae-bc| = |A|,$$

where $|A|$ is the determinant of A.

- Let the n-dimensional vector $\boldsymbol{Z}$ be

$$\boldsymbol{Z} = A\boldsymbol{X}, \qquad \text{where } A \text{ is an } n \times n \text{ invertable matrix.}$$

- The joint pdf of $\boldsymbol{Z}$ is then

$$\begin{aligned} f_{\boldsymbol{Z}}(\boldsymbol{z}) &= f_{Z_1,\ldots,Z_n}(z_1,\ldots,z_n) = \left.\frac{f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)}{|A|}\right|_{\boldsymbol{x}=A^{-1}\boldsymbol{z}} \\ &= \frac{f_{\boldsymbol{X}}(A^{-1}\boldsymbol{z})}{|A|}. \end{aligned}$$

Example: Let X and Y be the jointly Gaussian random variables with the pdf

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-(x^2-2\rho xy+y^2)/2(1-\rho^2)} \qquad -\infty < x, y < \infty.$$

Let V and W be obtained from (X, Y) by

$$\begin{bmatrix} V \\ W \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = A \begin{bmatrix} X \\ Y \end{bmatrix}.$$

Find the joint pdf of V and W.

Sol: $|A|=1$ and the inverse mapping is given by

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} V \\ W \end{bmatrix}.$$

Hence, $X = (V - W)/\sqrt{2}$ and $Y = (V + W)/\sqrt{2}$. Therefore, the joint pdf of V and W is

$$f_{V,W}(v, w) = f_{X,Y}\left(\frac{v - w}{\sqrt{2}}, \frac{v + w}{\sqrt{2}}\right).$$

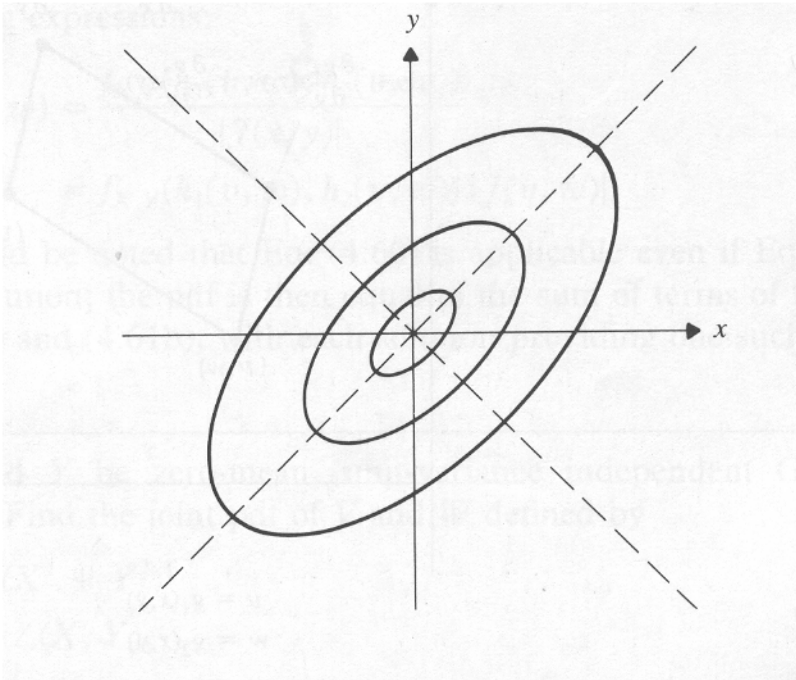
The argument of the exponent becomes

$$\frac{(v - w)^2/2 - 2\rho(v - w)(v + w)/2 + (v + w)^2/2}{2(1 - \rho^2)}$$
$$= \frac{v^2}{2(1 + \rho)} + \frac{w^2}{2(1 - \rho)}.$$

Thus

$$f_{V,W}(v, w) = \frac{1}{2\pi(1 - \rho^2)^{1/2}} e^{-\{[v^2/2(1+\rho)]+[w^2/2(1-\rho)]\}}.$$

Therefore, V and W are independent.


y
x

pdf of General Transformations

- Let V and W be defined by two nonlinear functions of X and Y:

$$V = g_1(X, Y) \qquad \text{and} \qquad W = g_2(X, Y)$$

- Assume that $g_1(x, y)$ and $g_2(x, y)$ are invertible, that is,

$$x = h_1(v, w) \qquad \text{and} \qquad y = h_2(v, w)$$

- The approximation is

$$g_k(x + dx, y) \approx g_k(x, y) + \frac{\partial}{\partial x} g_k(x, y) dx \qquad k = 1, 2.$$

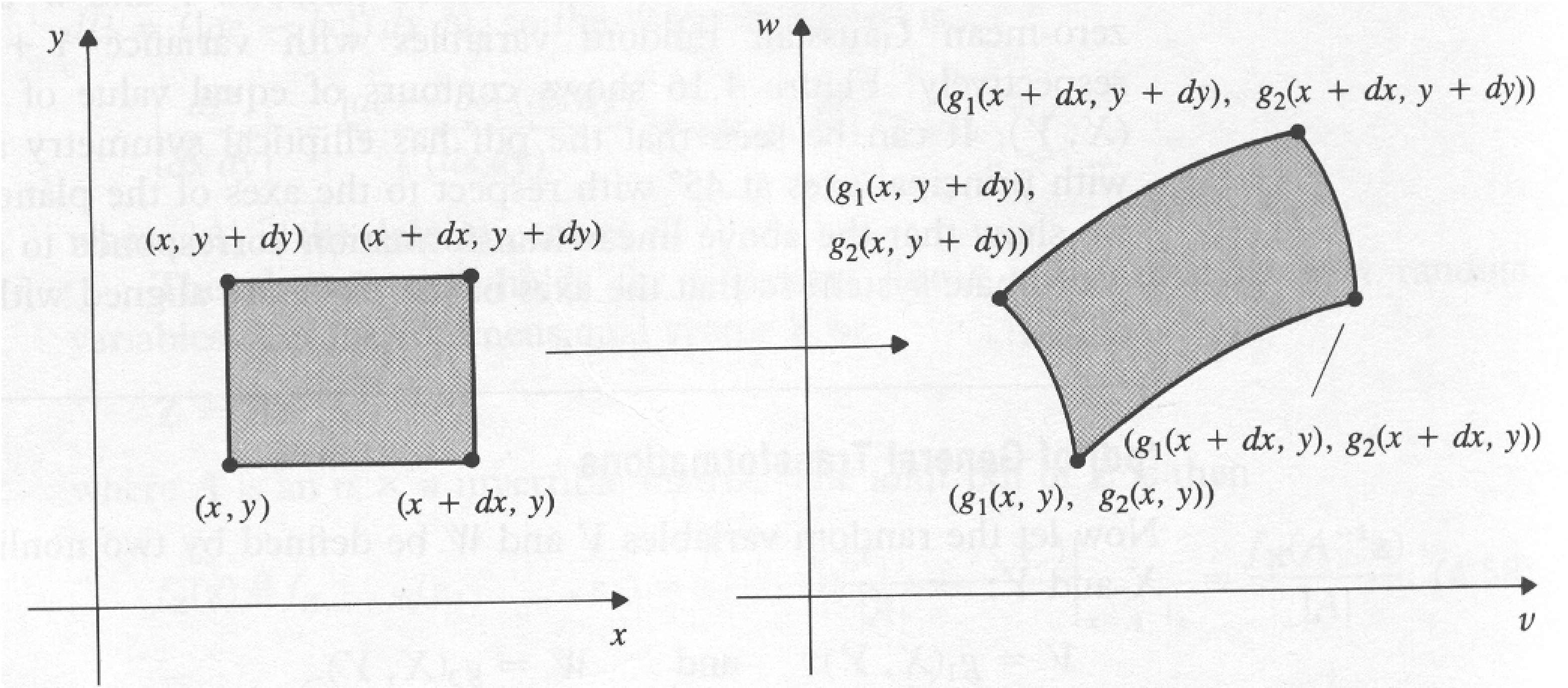
- The probability of the infinitesimal rectangle and the parallelogram are approximately equal

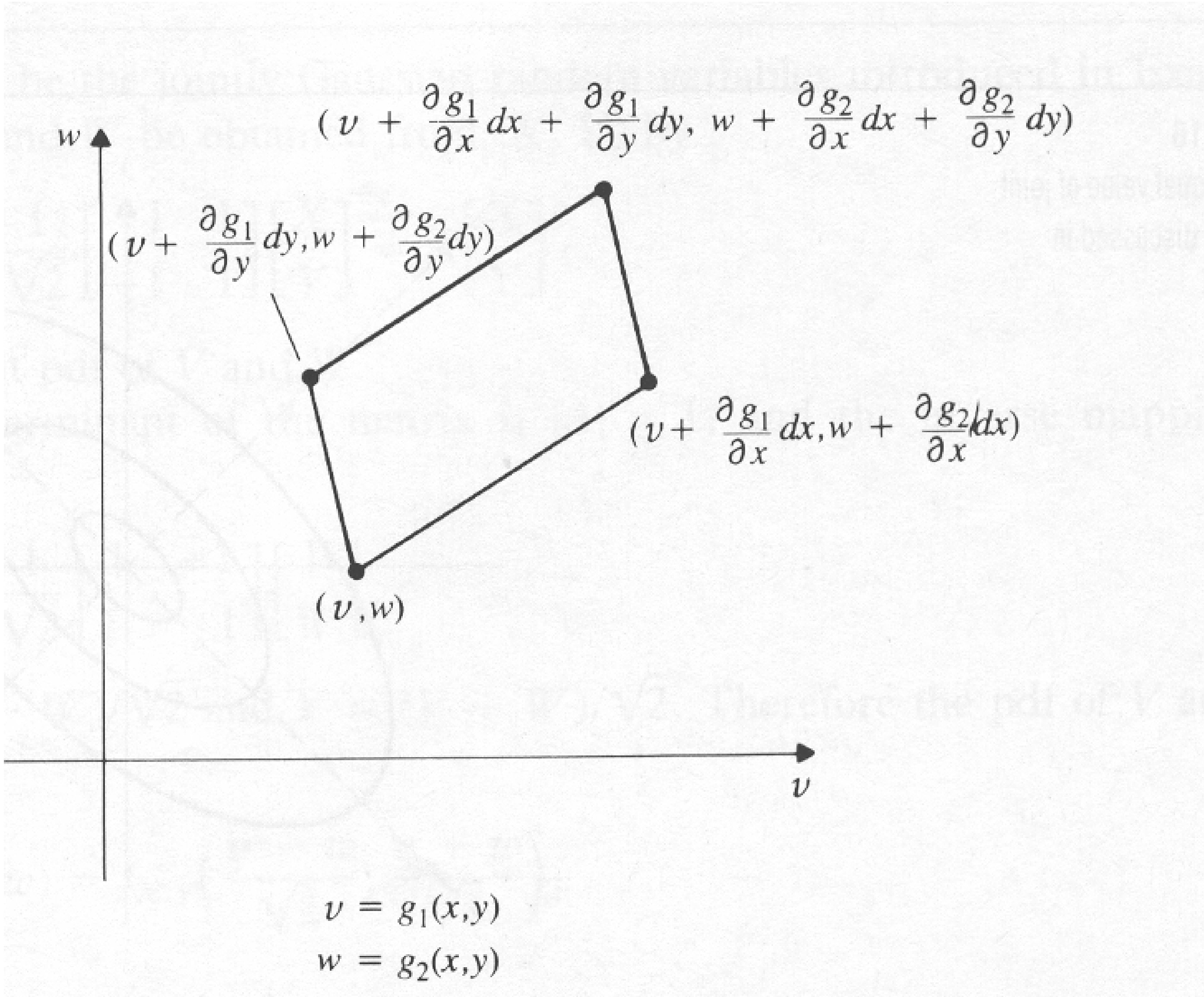
$$f_{X,Y}(x,y)dxdy = f_{V,W}(v,w)dP$$

and

$$f_{V,W}(v,w) = \frac{f_{X,Y}(h_1(v,w), h_2(v,w))}{|\frac{dP}{dxdy}|},$$

where dP is the area of the parallelogram.


y
(x, y + dy)
(x + dx, y + dy)
(x, y)
(x + dx, y)
x
w
(g1(x + dx, y + dy), g2(x + dx, y + dy))
(g1(x, y + dy),
g2(x, y + dy))
(g1(x + dx, y), g2(x + dx, y))
(g1(x, y), g2(x, y))
v


$(v + \frac{\partial g_1}{\partial x} dx + \frac{\partial g_1}{\partial y} dy,\ w + \frac{\partial g_2}{\partial x} dx + \frac{\partial g_2}{\partial y} dy)$
w
$(v + \frac{\partial g_1}{\partial y} dy, w + \frac{\partial g_2}{\partial y} dy)$
$(v + \frac{\partial g_1}{\partial x} dx, w + \frac{\partial g_2}{\partial x} dx)$
(v, w)
v
$v = g_1(x,y)$
$w = g_2(x,y)$

- Stretch factor – **Jacobian** of the transformation:

$$\mathcal{J}(x,y) = \det \begin{bmatrix} \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{bmatrix}.$$

- Jacobian of the inverse transformation is given by

$$\mathcal{J}(v,w) = \det \begin{bmatrix} \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \end{bmatrix}.$$

- It can be shown that

$$|\mathcal{J}(v,w)| = \frac{1}{|\mathcal{J}(x,y)|}.$$

- Joint pdf of V and W is then

$$\begin{aligned} f_{V,W}(v,w) &= \frac{f_{X,Y}(h_1(v,w),h_2(v,w))}{|\mathcal{J}(x,y)|} \\ &= f_{X,Y}(h_1(v,w),h_2(v,w))|\mathcal{J}(v,w)|. \end{aligned}$$

Example: Let X and Y be zero-mean, unit-variance independent Gaussian random variables. Find the joint pdf of V and W defined by

$$\begin{aligned} V &= (X^2 + Y^2)^{1/2} \\ W &= \angle(X, Y), \end{aligned}$$

where $\angle\theta$ denotes the angle in the range $(0, 2\pi)$ that is defined by the point (x, y).

Sol: Changing from Cartesian to polar coordinates. The inverse transformation is given by

$$x = v\cos w \qquad \text{and} \qquad y = v\sin w.$$

The Jacobian is given by

$$\mathcal{J}(v, w) = \begin{vmatrix} \cos w & -v\sin w \\ \sin w & v\cos w \end{vmatrix} = v.$$

Thus,

$$\begin{aligned} f_{V,W}(v,w) &= \frac{v}{2\pi} e^{-[v^2\cos^2(w)+v^2\sin^2(w)]/2} \\ &= \frac{1}{2\pi} v e^{-v^2/2} \qquad 0 \le v, 0 \le w < 2\pi. \end{aligned}$$

The pdf of a **Rayleigh random variable** is given by

$$f_V(v) = v e^{-v^2/2} \qquad v \ge 0.$$

Therefore, radius V and angle W are independent random variables and
V: Rayleigh random variable;
W: uniformly distributed $(0, 2\pi)$.

4.7 Expected Value of Functions of Random Variables

- The expected value of $Z = g(X, Y)$ is given by

$$E[Z] = \begin{cases} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y) f_{X,Y}(x,y)dxdy & X, Y \text{ are jointly continuous} \\ \sum_i \sum_n g(x_i, y_n) p_{X,Y}(x_i, y_n) & X, Y \text{ are discrete} \end{cases}$$

Example: Let $Z = X + Y$. Find $E[Z]$.

Sol:

$$\begin{aligned} E[Z] &= E[X+Y] \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x'+y')f_{X,Y}(x',y')dx'dy' \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x' f_{X,Y}(x',y')dx'dy' + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y' f_{X,Y}(x',y')dx'dy' \end{aligned}$$

$$= \int_{-\infty}^{+\infty} x' f_X(x') dx' + \int_{-\infty}^{+\infty} y' f_Y(y') dy' = E[X] + E[Y].$$

- Expected value of a sum of n random variables is

$$E[X_1 + X_2 + \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n].$$

Example: Suppose that X and Y are independent random variables, and let $g(X,Y) = g_1(X)g_2(Y)$. Show that $E[g(X,Y)] = E[g_1(X)g_2(Y)] = E[g_1(X)]E[g_2(X)]$.

Sol:

$$\begin{aligned} E[g_1(X)g_2(X)] &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g_1(x')g_2(y')f_X(x')f_Y(y')dx'dy' \\ &= \left\{\int_{-\infty}^{+\infty} g_1(x')f_X(x')dx'\right\}\left\{\int_{-\infty}^{+\infty} g_2(y')f_Y(y')dy'\right\} \\ &= E[g_1(X)]E[g_2(X)]. \end{aligned}$$

In general, if $X_1,\ldots,X_n$ are independent random variables, then

$$E[g_1(X_1)g_2(X_2)\cdots g_n(X_n)] = E[g_1(X_1)]E[g_2(X_2)]\cdots E[g_n(X_n)].$$

Correlation and Covariance of Two Random Variables

- Joint moment of X and Y is

$$E[X^jY^k] = \begin{cases} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^j y^k f_{X,Y}(x,y)dxdy & X, Y \text{ jointly continuous} \\ \sum_i \sum_n x_i^j y_n^k p_{X,Y}(x_i, y_n) & X, Y \text{ discrete} \end{cases}.$$

- If $j = 0$, then we obtain moments of Y, and if $k = 0$, then we obtain the moments of X.
- **Correlation of** X **and** Y is defined as $E[XY]$.
- If $E[XY] = 0$, then X and Y are **orthogonal**.
- The jkth **central moment of** X **and** Y is defined as

$$E[(X - E[X])^j(Y - E[Y])^k].$$

- **Covariance of X and Y** is defined as the $j = k = 1$ central moment:

$$\text{COV}(X, Y) = E[(X - E[X])(Y - E[Y])].$$

-

$$\begin{aligned}
\text{COV}(X, Y) &= E[XY - XE[Y] - YE[X] + E[X]E[Y]] \\
&= E[XY] - 2E[X]E[Y] + E[X]E[Y] \\
&= E[XY] - E[X]E[Y].
\end{aligned}$$

Example: Let X and Y be independent random variables. Find their covariance.

$$\begin{aligned} \text{COV}(X, Y) &= E[(X - E[X])(Y - E[Y])] \\ &= E[X - E[X]]E[Y - E[Y]] \\ &= 0. \end{aligned}$$

Pairs of independent random variables have covariance zero.

- **Correlation Coefficient of X and Y**

$$\rho_{X,Y} = \frac{\text{COV}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[XY] - E[X]E[Y]}{\sigma_X \sigma_Y},$$

  where $\sigma_X = \sqrt{\text{VAR}(X)}$ and $\sigma_Y = \sqrt{\text{VAR}(Y)}$.

- $\rho_{X,Y}$ is at most 1 in magnitude, that is,

$$-1 \le \rho_{X,Y} \le 1.$$

  This result is from the fact that the expected value of the square of a random variable is nonnegative:

$$\begin{aligned} 0 &\le E\left\{\left(\frac{X - E[X]}{\sigma_X} \pm \frac{Y - E[Y]}{\sigma_Y}\right)^2\right\} \\ &= 1 \pm 2\rho_{X,Y} + 1 = 2(1 \pm \rho_{X,Y}). \end{aligned}$$

- The extreme values of $\rho_{X,Y}$ are achieved when X and Y are related linearly, $Y = aX + b$:

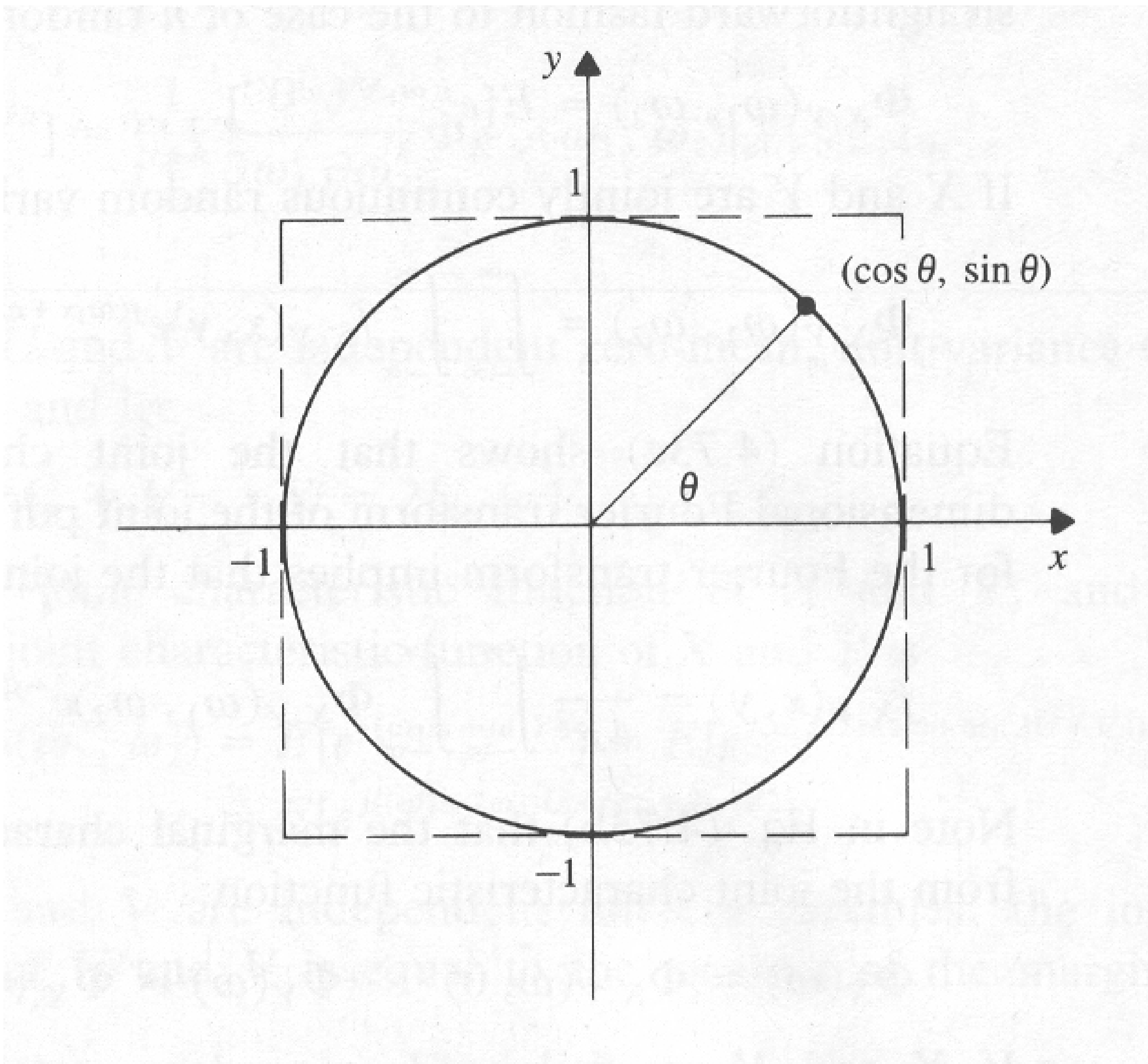
$$\rho_{X,Y} = \begin{cases} 1 & a > 0 \\ -1 & a < 0 \end{cases} .$$

- X, Y are said to be **uncorrelated** if $\rho_{X,Y} = 0$

$$\begin{array}{rcl} X, Y \text{ are independent} & \rightarrow & X, Y \text{ are uncorrelated} \\ X, Y \text{ are uncorrelated} & \overrightarrow{\text{NOT}} & X, Y \text{ are independent} \\ X, Y \text{ are uncorrelated jointly Gaussian} & \rightarrow & X, Y \text{ are independent} \end{array}$$

Example: Let Θ be uniformly distributed in the interval $(0, 2\pi)$. Let

$$X = \cos\Theta \text{ and } Y = \sin\Theta.$$


y
1
(cos θ, sin θ)
θ
−1
1
x
−1

X and Y are dependent.

$$\begin{aligned} E[XY] &= E[\sin\Theta\cos\Theta] = \frac{1}{2\pi}\int_0^{2\pi}\sin\phi\cos\phi d\phi \\ &= \frac{1}{4\pi}\int_0^{2\pi}\sin 2\phi d\phi = 0. \end{aligned}$$

Since $E[X] = E[Y] = 0$, it implies that X and Y are uncorrelated.

Example: Let X and Y be random variables with

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-y} & 0 \le y \le x < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

Find $E[XY]$, COV(X,Y), and $\rho_{X,Y}$.

Sol: First, find the mean, variance, and correlation of X and Y. We have $E[X] = 3/2$, $\text{VAR}[X] = 5/4$, $E[Y] = 1/2$ and $\text{VAR}[Y] = 1/4$. The correlation of X and Y is

$$\begin{aligned} E[XY] &= \int_0^\infty \int_0^x xy2e^{-x}e^{-y}dydx \\ &= \int_0^\infty 2xe^{-x}(1-e^{-x}-xe^{-x})dx = 1. \end{aligned}$$

Thus, the correlation coefficient is given by

$$\rho_{X,Y} = \frac{1 - \frac{3}{2}\frac{1}{2}}{\sqrt{\frac{5}{4}}\sqrt{\frac{1}{4}}} = \frac{1}{\sqrt{5}}.$$

4.8 Jointly Gaussian Random Variables

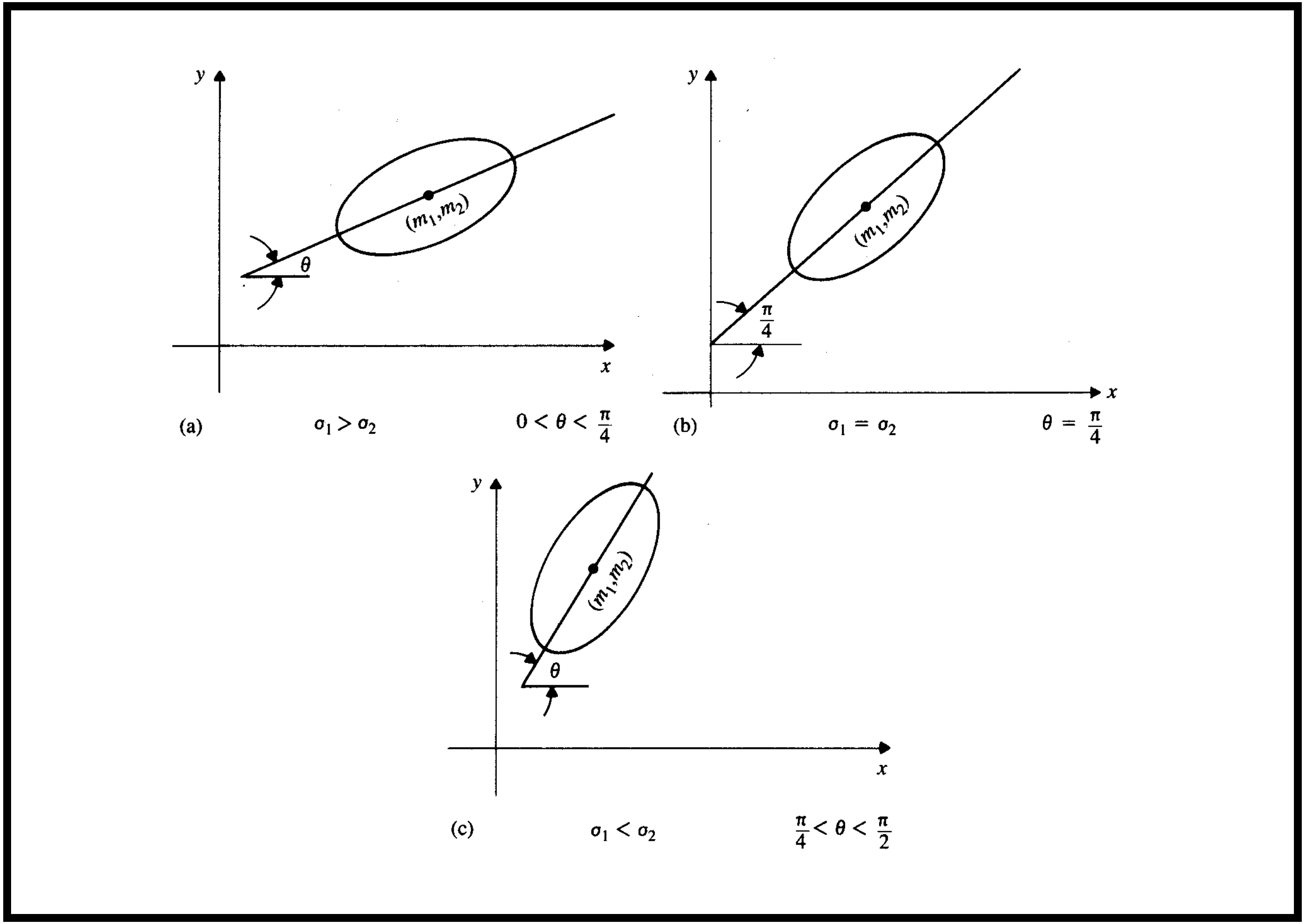
- X and Y are said to be jointly Gaussian if the joint pdf has the form

$$f_{X,Y}(x,y) = \frac{\exp\left\{\frac{-1}{2(1-\rho_{X,Y}^2)}\left[\left(\frac{x-m_1}{\sigma_1}\right)^2 - 2\rho_{X,Y}\left(\frac{x-m_1}{\sigma_1}\right)\left(\frac{y-m_2}{\sigma_2}\right) + \left(\frac{y-m_2}{\sigma_2}\right)^2\right]\right\}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{X,Y}^2}}.$$

- Bell shape
- Equal-pdf contour

- Marginal pdfs of X and Y are

$$f_X(x) = \frac{e^{-(x-m_1)^2/2\sigma_1^2}}{\sqrt{2\pi}\sigma_1}, \qquad f_Y(y) = \frac{e^{-(y-m_2)^2/2\sigma_2^2}}{\sqrt{2\pi}\sigma_2}.$$


y
x
θ
(m_1, m_2)
(a)
$\sigma_1 > \sigma_2$
$0 < \theta < \frac{\pi}{4}$
y
x
$\frac{\pi}{4}$
(m_1, m_2)
(b)
$\sigma_1 = \sigma_2$
$\theta = \frac{\pi}{4}$
y
x
θ
(m_1, m_2)
(c)
$\sigma_1 < \sigma_2$
$\frac{\pi}{4} < \theta < \frac{\pi}{2}$

- The conditional pdf $f_X(x|y)$ ($f_Y(y|x)$) is

$$\begin{aligned} f_X(x|y) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \\ &= \frac{\exp\left\{\frac{-1}{2(1-\rho_{X,Y}^2)\sigma_1^2}\left[x-\rho_{X,Y}\frac{\sigma_1}{\sigma_2}(y-m_2)-m_1\right]^2\right\}}{\sqrt{2\pi\sigma_1^2(1-\rho_{X,Y}^2)}}. \end{aligned}$$

- Conditional mean is $m_1 + \rho_{X,Y}(\sigma_1/\sigma_2)(y - m_2)$ and conditional variance $\sigma_1^2(1 - \rho_{X,Y}^2)$.
- Show that $\rho_{X,Y}$ is the correlation coefficient between X and Y.

  Sol: We have

$$\text{COV}(X,Y) = E[(X-m_1)(Y-m_2)]$$

$$= E[E[(X - m_1)(Y - m_2)|Y]].$$

The conditional expectation of $(X - m_1)(Y - m_2)$ given $Y = y$ is

$$\begin{aligned} E[(X - m_1)(Y - m_2)|Y = y] &= (y - m_2)E[X - m_1|Y = y] \\ &= (y - m_2)\left(E[X|Y = y] - m_1\right) \\ &= (y - m_2)\left(\rho_{X,Y}\frac{\sigma_1}{\sigma_2}(y - m_2)\right). \end{aligned}$$

Therefore,

$$E[(X - m_1)(Y - m_2)|Y] = \rho_{X,Y}\frac{\sigma_1}{\sigma_2}(Y - m_2)^2$$

and

$$\begin{aligned} \text{COV}(X, Y) &= E[E[(X - m_1)(Y - m_2)|Y]] = \rho_{X,Y}\frac{\sigma_1}{\sigma_2}E[(Y - m_2)^2] \\ &= \rho_{X,Y}\sigma_1\sigma_2. \end{aligned}$$

n jointly Gaussian Random Variables

- $X_1, X_2, \ldots, X_n$ are jointly Gaussian if the pdf is given by

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{x}) &= f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) \\ &= \frac{\exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})^T K^{-1}(\boldsymbol{x}-\boldsymbol{m})\right\}}{(2\pi)^{n/2}|K|^{1/2}}, \end{aligned}$$

where $\boldsymbol{x}$ and $\boldsymbol{m}$ are column vectors defined by

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \boldsymbol{m} = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{bmatrix} = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_n] \end{bmatrix}$$

and K is the **covariance matrix** that is defined by

$$K = \begin{bmatrix} \mathrm{VAR}(X_1) & \mathrm{COV}(X_1, X_2) & \cdots & \mathrm{COV}(X_1, X_n) \\ \mathrm{COV}(X_2, X_1) & \mathrm{VAR}(X_2) & \cdots & \mathrm{COV}(X_2, X_n) \\ \vdots & \vdots & & \vdots \\ \mathrm{COV}(X_n, X_1) & \cdots & & \mathrm{VAR}(X_n) \end{bmatrix}.$$

Example: Verify the two-dimensional Gaussian pdf.

Sol: The covariance matrix is

$$K = \begin{bmatrix} \sigma_1^2 & \rho_{X,Y}\sigma_1\sigma_2 \\ \rho_{X,Y}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \text{ and}$$

$$K^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho_{X,Y}^2)} \begin{bmatrix} \sigma_2^2 & -\rho_{X,Y}\sigma_1\sigma_2 \\ -\rho_{X,Y}\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix}.$$

The term in the exponential is

$$\frac{1}{\sigma_1^2\sigma_2^2(1-\rho_{X,Y}^2)}[x-m_1, y-m_2] \begin{bmatrix} \sigma_2^2 & -\rho_{X,Y}\sigma_1\sigma_2 \\ -\rho_{X,Y}\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix} \begin{bmatrix} x-m_1 \\ y-m_2 \end{bmatrix}$$

$$= \frac{((x-m_1)/\sigma_1)^2 - 2\rho_{X,Y}((x-m_1)/\sigma_1)((y-m_2)/\sigma_2) + ((y-m_2)/\sigma_2)^2}{(1-\rho_{X,Y}^2)}$$

.

Linear Transformation of Gaussian Random Variables

- Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be jointly Gaussian, and

$$\boldsymbol{Y} = A\boldsymbol{X},$$

where A is an $n \times n$ invertible matrix.

- The pdf of $\boldsymbol{Y}$ is

$$\begin{aligned} f_{\boldsymbol{Y}}(\boldsymbol{y}) &= \frac{f_{\boldsymbol{X}}(A^{-1}\boldsymbol{y})}{|A|} \\ &= \frac{\exp\left\{-\frac{1}{2}(A^{-1}\boldsymbol{y}-\boldsymbol{m})^T K^{-1}(A^{-1}\boldsymbol{y}-\boldsymbol{m})\right\}}{(2\pi)^{n/2}|A||K|^{1/2}}. \end{aligned}$$

Since

$$(A^{-1}\boldsymbol{y} - \boldsymbol{m}) = A^{-1}(\boldsymbol{y} - A\boldsymbol{m})$$

and

$$(A^{-1}\boldsymbol{y} - \boldsymbol{m})^T = (\boldsymbol{y} - A\boldsymbol{m})^T A^{-1T},$$

the argument of the exponential is

$$(\boldsymbol{y} - A\boldsymbol{m})^T A^{-1T} K^{-1} A^{-1}(\boldsymbol{y} - A\boldsymbol{m}) = (\boldsymbol{y} - A\boldsymbol{m})^T (AKA^T)^{-1}(\boldsymbol{y} - A\boldsymbol{m}).$$

Let $C = AKA^T$ $\boldsymbol{n} = A\boldsymbol{m}$. Noting that $\det(C) = \det(AKA^T) = \det(A)^2 \det(K)$ and we have

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \frac{e^{-(1/2)(\boldsymbol{y}-\boldsymbol{n})^T C^{-1}(\boldsymbol{y}-\boldsymbol{n})}}{(2\pi)^{n/2}|C|^{1/2}}.$$

Therefore, $\boldsymbol{Y}$ are jointly Gaussian with mean $\boldsymbol{n}$ and covariance C:

$$\boldsymbol{n} = A\boldsymbol{m} \quad \text{and} \quad C = AKA^T.$$

- It is possible to transform a vector of jointly Gaussian random variables into a vector of independent Gaussian random variables since it is always possible to find a matrix A such that $AKA^T = \Lambda$, where Λ is a diagonal matrix, due to the symmetry of A.