A Probability Review

Outline:

- A probability review.

Shorthand notation: RV stands for random variable.

A Probability Review

Reading:

- Go over handouts 2–5 in EE 420x notes.

Basic probability rules:

(1) $\Pr\{\Omega\} = 1$, $\Pr\{\emptyset\} = 0$, $0 \leq \Pr\{A\} \leq 1$;
$\Pr\{\cup_{i=1}^{\infty} A_i\} = \sum_{i=1}^{\infty} \Pr\{A_i\}$ if $\underbrace{A_i \cap A_j}_{A_i \text{ and } A_j \text{ disjoint}} = \emptyset$ for all $i \neq j$;

(2) $\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\} - \Pr\{A \cap B\}$, $\Pr\{A^{\mathrm{c}}\} = 1 - \Pr\{A\}$;

(3) If $A \perp\!\!\!\perp B$, then $\Pr\{A \cap B\} = \Pr\{A\} \cdot \Pr\{B\}$;

(4)

$$\Pr\{A \mid B\} = \frac{\Pr\{A \cap B\}}{\Pr\{B\}} \quad \text{(conditional probability)}$$

or

$$\Pr\{A \cap B\} = \Pr\{A \mid B\} \cdot \Pr\{B\} \quad \text{(chain rule)};$$

(5)

$$\Pr\{A\} = \Pr\{A \mid B_1\} \Pr\{B_1\} + \cdots + \Pr\{A \mid B_n\} \Pr\{B_n\}$$

if $B_1, B_2, \ldots, B_n$ form a *partition* of the full space Ω;

(6) Bayes' rule:

$$\Pr\{A \mid B\} = \frac{\Pr\{B \mid A\} \Pr\{A\}}{\Pr\{B\}}.$$

Reminder: Independence, Correlation and Covariance

For simplicity, we state all the definitions for pdfs; the corresponding definitions for pmfs are analogous.

Two random variables X and Y are *independent* if

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

Correlation between real-valued random variables X and Y:

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

Covariance between real-valued random variables X and Y:

$$\begin{aligned}\mathrm{cov}_{X,Y}(X,Y) &= \mathrm{E}_{X,Y}\{(X-\mu_X)\,(Y-\mu_Y)\} \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mu_X)\,(y-\mu_Y)\,f_{X,Y}(x,y)\,dx\,dy\end{aligned}$$

where

$$\mu_X = \mathrm{E}_X(X) = \int_{-\infty}^{+\infty} x\, f_X(x)\, dx$$
$$\mu_Y = \mathrm{E}_Y(Y) = \int_{-\infty}^{+\infty} y\, f_Y(y)\, dy.$$

Uncorrelated random variables: Random variables X and Y are *uncorrelated* if

$$c_{X,Y} = \mathrm{cov}_{X,Y}(X, Y) = 0. \qquad (1)$$

If X and Y are real-valued RVs, then (1) can be written as

$$\mathrm{E}_{X,Y}\{X\,Y\} = \mathrm{E}_X\{X\}\,\mathrm{E}_Y\{Y\}.$$

Mean Vector and Covariance Matrix:

Consider a random vector

$$\boldsymbol{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}.$$

The *mean* of this random vector is defined as

$$\boldsymbol{\mu}_{\boldsymbol{X}} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{bmatrix} = \mathrm{E}_{\boldsymbol{X}}\{\boldsymbol{X}\} = \begin{bmatrix} \mathrm{E}_{X_1}[X_1] \\ \mathrm{E}_{X_2}[X_2] \\ \vdots \\ \mathrm{E}_{X_N}[X_N] \end{bmatrix}.$$

Denote the *covariance* between X_i and X_k, $\mathrm{cov}_{X_i,X_k}(X_i, X_k)$, by $c_{i,k}$; hence, the *variance* of X_i is $c_{i,i} = \mathrm{cov}_{X_i,X_k}(X_i, X_i) = \underbrace{\mathrm{var}_{X_i}(X_i) = \sigma^2_{X_i}}_{\text{more notation}}$. The *covariance matrix* of $\boldsymbol{X}$ is defined as

$$C_{\boldsymbol{X}} = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & \cdots & c_{1,N} \\ c_{2,1} & c_{2,2} & \cdots & \cdots & c_{2,N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{N,1} & c_{N,2} & \cdots & \cdots & c_{N,N} \end{bmatrix}.$$

The above definitions apply to both real and complex vectors $\boldsymbol{X}$.

Covariance matrix of a real-valued random vector $\boldsymbol{X}$:

$$\begin{aligned} C_{\boldsymbol{X}} &= \mathrm{E}_{\boldsymbol{X}}\{(\boldsymbol{X} - \mathrm{E}_{\boldsymbol{X}}[\boldsymbol{X}])\,(\boldsymbol{X} - \mathrm{E}_{\boldsymbol{X}}[\boldsymbol{X}])^T\} \\ &= \mathrm{E}_{\boldsymbol{X}}[\boldsymbol{X}\,\boldsymbol{X}^T] - \mathrm{E}_{\boldsymbol{X}}[\boldsymbol{X}]\,(\mathrm{E}_{\boldsymbol{X}}[\boldsymbol{X}])^T. \end{aligned}$$

For real-valued $\boldsymbol{X}$, $c_{i,k} = c_{k,i}$ and, therefore, $C_{\boldsymbol{X}}$ is a symmetric matrix.

Linear Transform of Random Vectors

Linear Transform. For real-valued $\boldsymbol{Y}, \boldsymbol{X}, A$,

$$\boldsymbol{Y} = g(\boldsymbol{X}) = A\,\boldsymbol{X}.$$

Mean Vector:

$$\boldsymbol{\mu}_{\boldsymbol{Y}} = \mathrm{E}_{\boldsymbol{X}}\{A\,\boldsymbol{X}\} = A\,\boldsymbol{\mu}_{\boldsymbol{X}}. \tag{2}$$

Covariance Matrix:

$$\begin{aligned}
C_{\boldsymbol{Y}} &= \mathrm{E}_{\boldsymbol{Y}}\{\boldsymbol{Y}\,\boldsymbol{Y}^T\} - \boldsymbol{\mu}_{\boldsymbol{Y}}\,\boldsymbol{\mu}_{\boldsymbol{Y}}^T \\
&= \mathrm{E}_{\boldsymbol{X}}\{A\,\boldsymbol{X}\,\boldsymbol{X}^T\,A^T\} - A\,\boldsymbol{\mu}_{\boldsymbol{X}}\,\boldsymbol{\mu}_{\boldsymbol{X}}^T\,A^T \\
&= A\Big(\underbrace{\mathrm{E}_{\boldsymbol{X}}\{\boldsymbol{X}\,\boldsymbol{X}^T\} - \boldsymbol{\mu}_{\boldsymbol{X}}\,\boldsymbol{\mu}_{\boldsymbol{X}}^T}_{C_{\boldsymbol{X}}}\Big)A^T \\
&= A\,C_{\boldsymbol{X}}\,A^T. \qquad (3)
\end{aligned}$$

Reminder: Iterated Expectations

In general, we can find $\mathrm{E}_{X,Y}[g(X,Y)]$ using *iterated expectations*:

$$\mathrm{E}_{X,Y}[g(X,Y)] = \mathrm{E}_{Y}\{\mathrm{E}_{X \mid Y}[g(X,Y) \mid Y]\} \tag{4}$$

where $\mathrm{E}_{X \mid Y}$ denotes the expectation with respect to $f_{X|Y}(x \mid y)$ and E_{Y} denotes the expectation with respect to $f_Y(y)$.

Proof.

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

□

Reminder: Law of Conditional Variances

Define the conditional variance of X given $Y = y$ to be the variance of X with respect to $f_{X\,|\,Y}(x\,|\,y)$, i.e.

$$\begin{aligned}\mathrm{var}_{X\,|\,Y}(X\,|\,Y=y) &= \mathrm{E}_{X\,|\,Y}\Big[(X-\mathrm{E}_{X|Y}[X\,|\,y])^2\,|\,y\Big]\\ &= \mathrm{E}_{X\,|\,Y}[X^2\,|\,y]-(\mathrm{E}_{X|Y}[X\,|\,y])^2.\end{aligned}$$

The random variable $\mathrm{var}_{X\,|\,Y}(X\,|\,Y)$ is a function of Y only, taking values $\mathrm{var}(X\,|\,Y=y)$. Its expected value with respect to Y is

$$\begin{aligned}\mathrm{E}_Y\{\mathrm{var}_{X\,|\,Y}(X\,|\,Y)\} &= \mathrm{E}_Y\Big\{\mathrm{E}_{X\,|\,Y}[X^2\,|\,Y]-(\mathrm{E}_{X|Y}[X\,|\,Y])^2\Big\}\\ &\overset{\text{iterated exp.}}{=} \mathrm{E}_{X,Y}[X^2]-\mathrm{E}_Y\{(\mathrm{E}_{X|Y}[X\,|\,Y])^2\}\\ &= \mathrm{E}_X[X^2]-\mathrm{E}_Y\{(\mathrm{E}_{X|Y}[X\,|\,Y])^2\}.\end{aligned}$$

Since $\mathrm{E}_{X|Y}[X\,|\,Y]$ is a random variable (and a function of Y only), it has variance:

$$\begin{aligned}\mathrm{var}_Y\{\mathrm{E}_{X|Y}[X\,|\,Y]\} &= \mathrm{E}_Y\{(\mathrm{E}_{X|Y}[X\,|\,Y])^2\}-(\mathrm{E}_Y\{\mathrm{E}_{X|Y}[X\,|\,Y]\})^2\\ &\overset{\text{iterated exp.}}{=} \mathrm{E}_Y\{(\mathrm{E}_{X|Y}[X\,|\,Y])^2\}-(\mathrm{E}_{X,Y}[X])^2\\ &= \mathrm{E}_Y\{\mathrm{E}_{X|Y}[X\,|\,Y]^2\}-(\mathrm{E}_X[X])^2.\end{aligned}$$

Adding the above two expressions yields the *law of conditional variances*:

$$\mathrm{E}_{Y}\{\mathrm{var}_{X \mid Y}(X \mid Y)\} + \mathrm{var}_{Y}\{\mathrm{E}_{X|Y}[X \mid Y]\} = \mathrm{var}_{X}(X). \quad (5)$$

Note: (4) and (5) hold for both real- and complex-valued random variables.

Useful Expectation and Covariance Identities for Real-valued Random Variables and Vectors

$$\begin{aligned} \mathrm{E}_{X,Y}[a\,X + b\,Y + c] &= a \cdot \mathrm{E}_X[X] + b\,\cdot \mathrm{E}_Y[Y] + c \\ \mathrm{var}_{X,Y}(a\,X + b\,Y + c) &= a^2\,\mathrm{var}_X(X) + b^2\,\mathrm{var}_Y(Y) \\ &\quad +2\,a\,b \cdot \mathrm{cov}_{X,Y}(X,Y) \end{aligned}$$

where a, b, and c are constants and X and Y are random variables. A vector/matrix version of the above identities:

$$\begin{aligned} &\mathrm{E}_{\boldsymbol{X},\boldsymbol{Y}}[A\,\boldsymbol{X} + B\,\boldsymbol{Y} + \boldsymbol{c}] = A\,\mathrm{E}_{\boldsymbol{X}}[\boldsymbol{X}] + B\,\mathrm{E}_{\boldsymbol{Y}}[\boldsymbol{Y}] + \boldsymbol{c} \\ &\mathrm{cov}_{\boldsymbol{X},\boldsymbol{Y}}(A\,\boldsymbol{X} + B\,\boldsymbol{Y} + \boldsymbol{c}) = A\,\mathrm{cov}_{\boldsymbol{X}}(\boldsymbol{X})\,A^T + B\,\mathrm{cov}_{\boldsymbol{Y}}(\boldsymbol{Y})\,B^T \\ &+A\,\mathrm{cov}_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{X},\boldsymbol{Y})\,B^T + B\,\mathrm{cov}_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{Y},\boldsymbol{X})\,A^T \end{aligned}$$

where "T" denotes a transpose and

$$\mathrm{cov}_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{X},\boldsymbol{Y}) = \mathrm{E}_{\boldsymbol{X},\boldsymbol{Y}}\{(\boldsymbol{X} - \mathrm{E}_{\boldsymbol{X}}[\boldsymbol{X}])\,(\boldsymbol{Y} - \mathrm{E}_{\boldsymbol{Y}}[\boldsymbol{Y}])^T\}.$$

Useful properties of crosscovariance matrices:

- $$\mathrm{cov}_{\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}}(\boldsymbol{X},\boldsymbol{Y}+\boldsymbol{Z}) = \mathrm{cov}_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{X},\boldsymbol{Y}) + \mathrm{cov}_{\boldsymbol{X},\boldsymbol{Z}}(\boldsymbol{X},\boldsymbol{Z}).$$

- $$\mathrm{cov}_{\boldsymbol{Y},\boldsymbol{X}}(\boldsymbol{Y},\boldsymbol{X}) = [\mathrm{cov}_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{X},\boldsymbol{Y})]^T.$$

- $$\begin{aligned}\mathrm{cov}_{\boldsymbol{X}}(\boldsymbol{X}) &= \mathrm{cov}_{\boldsymbol{X}}(\boldsymbol{X},\boldsymbol{X}) \\ \mathrm{var}_X(X) &= \mathrm{cov}_X(X,X).\end{aligned}$$

- $$\mathrm{cov}_{\boldsymbol{X},\boldsymbol{Y}}(A\,\boldsymbol{X}+\boldsymbol{b}, P\,\boldsymbol{Y}+\boldsymbol{q}) = A\,\mathrm{cov}_{\boldsymbol{X},\boldsymbol{Y}}(\boldsymbol{X},\boldsymbol{Y})\,P^T.$$

(To refresh memory about covariance and its properties, see p. 12 of handout 5 in EE 420x notes. For random vectors, see handout 7 in EE 420x notes, particularly pp. 1–15.)

Useful theorems:

(1) (handout 5 in EE 420x notes)

$$\mathrm{E}_X(X) = \mathrm{E}_Y[\mathrm{E}_{X\,|\,Y}(X\,|\,Y)] \quad \text{shown on p. 8}$$

$$\mathrm{E}_{X\,|\,Y}[g(X)\cdot h(Y)\,|\,y] = h(y)\cdot \mathrm{E}_{X\,|\,Y}[g(X)\,|\,y]$$

$$\mathrm{E}_{X,Y}[g(X)\cdot h(Y)] = \mathrm{E}_Y\{h(Y)\cdot \mathrm{E}_{X\,|\,Y}[g(X)\,|\,Y]\}.$$

The vector version of (1) is the same — just put bold letters.

(2)

$$\mathrm{var}_X(X) = \mathrm{E}_Y[\mathrm{var}_{X \mid Y}(X \mid Y)] + \mathrm{var}_Y(\mathrm{E}_{X \mid Y}[X|Y]);$$

and the vector/matrix version is

$$\underbrace{\mathrm{cov}_{\boldsymbol{X}}(\boldsymbol{X})}_{\text{variance/covariance matrix of } \boldsymbol{X}} = \mathrm{E}_{\boldsymbol{Y}}[\mathrm{cov}_{\boldsymbol{X} \mid \boldsymbol{Y}}(\boldsymbol{X} \mid \boldsymbol{Y})]$$

$$+\mathrm{cov}_{\boldsymbol{Y}}(\mathrm{E}_{\boldsymbol{X} \mid \boldsymbol{Y}}[\boldsymbol{X}|\boldsymbol{Y}])$$ shown on p. .

(3) Generalized law of conditional variances:

$$\mathrm{cov}_{X,Y}(X, Y) = \mathrm{E}_Z[\mathrm{cov}_{X,Y \mid Z}(X, Y \mid Z)]$$
$$+\mathrm{cov}_Z(\mathrm{E}_{X \mid Z}[X \mid Z], \mathrm{E}_{Y \mid Z}[Y \mid Z]).$$

(4) Transformation:

$$\boldsymbol{Y} = \boldsymbol{g}(\boldsymbol{X}) \quad \overset{\text{one-to-one}}{\Longleftrightarrow} \quad \begin{array}{c} Y_1 = g_1(X_1, \ldots, X_n) \\ \vdots \\ Y_n = g_n(X_1, \ldots, X_n) \end{array}$$

then

$$f_Y(\boldsymbol{y}) = f_X(h_1(y_1), \ldots, h_n(y_n)) \cdot |J|$$

where $h(\cdot)$ is the unique inverse of $g(\cdot)$ and

$$J = \left| \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{y}^T} \right| = \left| \begin{array}{ccc} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \cdots & \frac{\partial x_n}{\partial y_n} \end{array} \right|$$

Print and read the handout PROBABILITY DISTRIBUTIONS from the Course readings section on `WebCT`. Bring it with you to the midterm exams.

Jointly Gaussian Real-valued RVs

Scalar Gaussian random variables:

$$f_X(x) \quad = \quad \frac{1}{\sqrt{2\,\pi\,\sigma_X^2}} \exp\Big[-\frac{(x-\mu_X)^2}{2\,\sigma_X^2}\Big].$$

Definition. *Two real-valued RVs X and Y are jointly Gaussian if their joint pdf is of the form*

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{X,Y}^2}} \cdot \exp\Big\{-\frac{1}{2\,(1-\rho_{X,Y}^2)} \cdot \Big[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - 2\,\rho_{X,Y}\frac{(x-\mu_X)\,(y-\mu_Y)}{\sigma_X\sigma_Y}\Big]\Big\}. \tag{6}$$

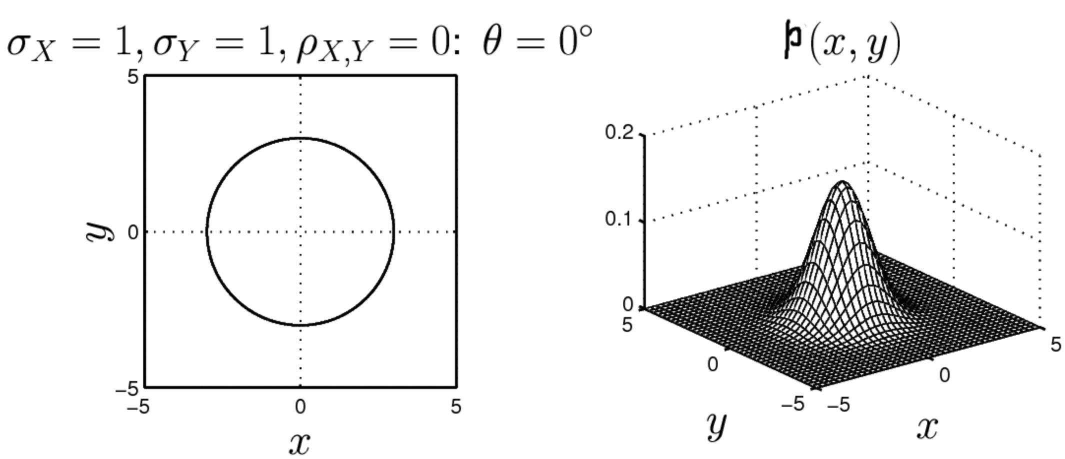
This pdf is parameterized by $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$, and $\rho_{X,Y}$. Here, $\sigma_X = \sqrt{\sigma_X^2}$ and $\sigma_Y = \sqrt{\sigma_Y^2}$.

Note: We will soon define a more general multivariate Gaussian pdf.

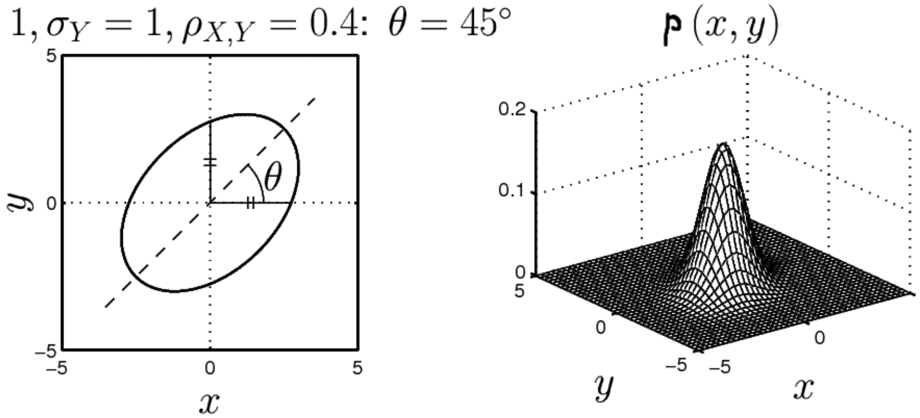
If X and Y are jointly Gaussian, contours of equal joint pdf are ellipses defined by the quadratic equation

$$\frac{(x-\mu_X)^2}{\sigma_X^2}+\frac{(y-\mu_Y)^2}{\sigma_Y^2}-2\,\rho_{X,Y}\frac{(x-\mu_X)\,(y-\mu_Y)}{\sigma_X\sigma_Y} = \text{const} \geq 0.$$

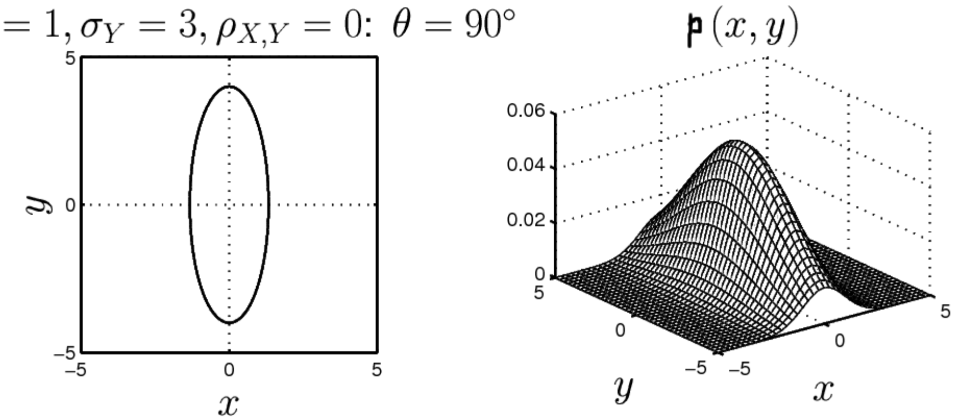
Examples: In the following examples, we plot contours of the joint pdf $f_{X,Y}(x,y)$ for zero-mean jointly Gaussian RVs for various values of σ_X, σ_Y, and $\rho_{X,Y}$.



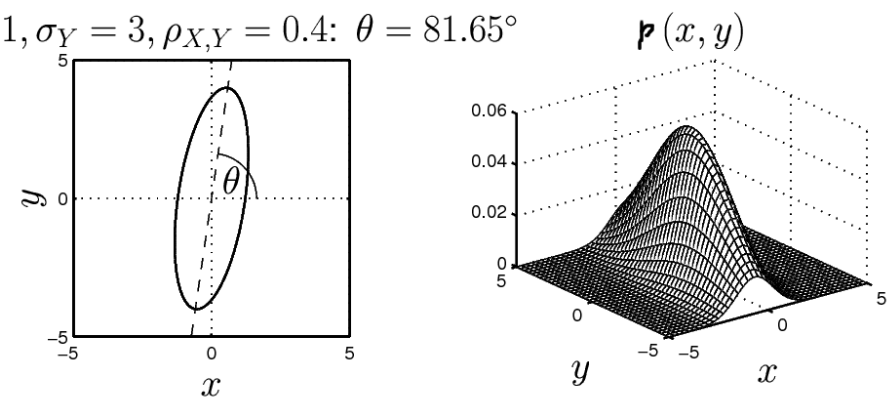
$\sigma_X = 1, \sigma_Y = 1, \rho_{X,Y} = 0.4$: $\theta = 45°$



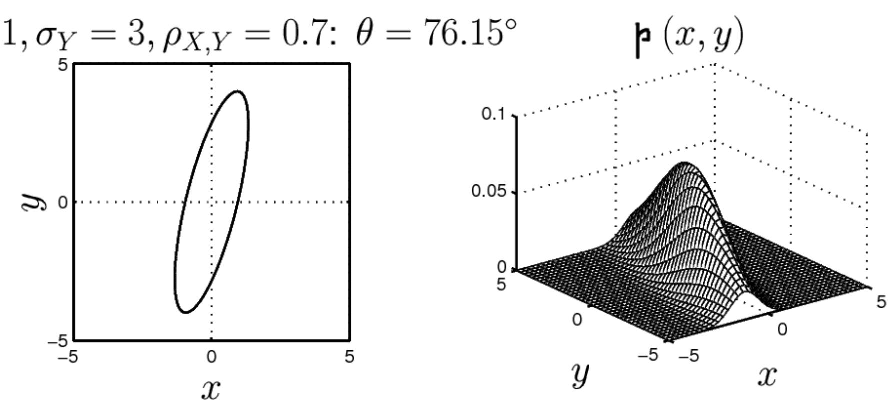
$\sigma_X = 1, \sigma_Y = 3, \rho_{X,Y} = 0$: $\theta = 90°$

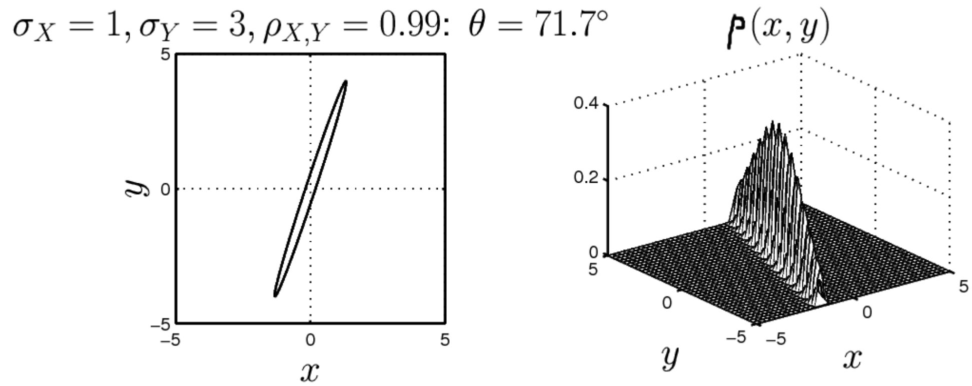


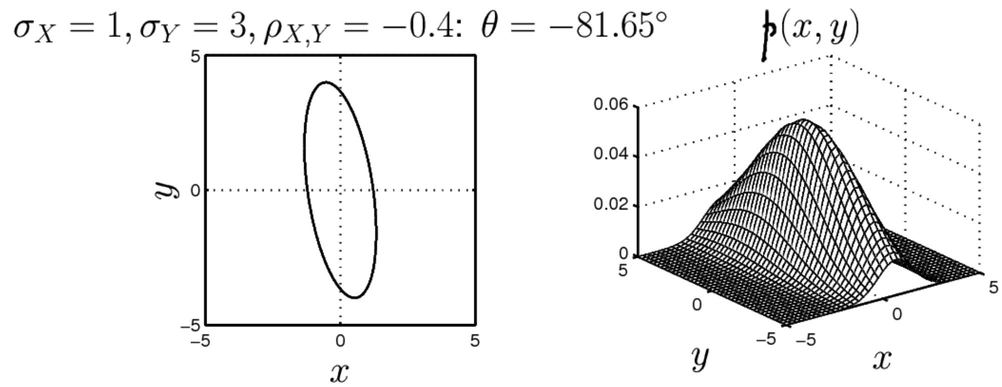
$\sigma_X = 1, \sigma_Y = 3, \rho_{X,Y} = 0.4$: $\theta = 81.65°$



$\sigma_X = 1, \sigma_Y = 3, \rho_{X,Y} = 0.7$: $\theta = 76.15°$







If X and Y are jointly Gaussian, the conditional pdfs are Gaussian, e.g.

$$X \mid \{Y = y\} \sim \mathcal{N}\Big(\rho_{X,Y} \cdot \sigma_X \cdot \frac{y - \mathrm{E}_Y[Y]}{\sigma_Y} + \mathrm{E}_X[X], (1 - \rho_{X,Y}^2) \cdot \sigma_X^2\Big). \quad (7)$$

If X and Y are jointly Gaussian and uncorrelated, i.e. $\rho_{X,Y} = 0$, they are also independent.

Gaussian Random Vectors

Real-valued Gaussian random vectors:

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\,\pi)^{N/2}\,|C_{\boldsymbol{X}}|^{1/2}} \exp\big[-\tfrac{1}{2}\,(\boldsymbol{x}-\boldsymbol{\mu}_{\boldsymbol{X}})^T C_{\boldsymbol{X}}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_{\boldsymbol{X}})\big].$$

Complex-valued Gaussian random vectors:

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \frac{1}{\pi^N |C_{\boldsymbol{Z}}|} \exp\big[-(\boldsymbol{z}-\boldsymbol{\mu}_{\boldsymbol{Z}})^H\,C_{\boldsymbol{Z}}^{-1}\,(\boldsymbol{z}-\boldsymbol{\mu}_{\boldsymbol{Z}})\big].$$

Notation for *real-* and *complex-valued* Gaussian random vectors:

$\boldsymbol{X} \sim \mathcal{N}_{\rm r}(\boldsymbol{\mu}_{\boldsymbol{X}}, C_{\boldsymbol{X}})$ [or simply $\mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{X}}, C_{\boldsymbol{X}})$] real

$\boldsymbol{X} \sim \mathcal{N}_{\rm c}(\boldsymbol{\mu}_{\boldsymbol{X}}, C_{\boldsymbol{X}})$ complex.

An *affine transform* of a Gaussian vector is also a Gaussian random vector, i.e. if

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

then

$\boldsymbol{Y} \sim \mathcal{N}_{\rm r}(A\,\boldsymbol{\mu}_{\boldsymbol{X}} + \boldsymbol{b}, A\,C_{\boldsymbol{X}}\,A^T)$ real

$\boldsymbol{Y} \sim \mathcal{N}_{\rm c}(A\,\boldsymbol{\mu}_{\boldsymbol{X}} + \boldsymbol{b}, A\,C_{\boldsymbol{X}}\,A^H)$ complex.

The Gaussian random vector $\boldsymbol{W} \sim \mathcal{N}_\mathrm{r}(\mathbf{0}, \sigma^2 I_n)$ (where I_n denotes the identity matrix of size n) is called *white*; pdf contours of a white Gaussian random vector are spheres centered at the origin. Suppose that $W[n]$, $n = 0, 1, \ldots, N-1$ are independent, identically distributed (i.i.d.) zero-mean univariate Gaussian $\mathcal{N}(0, \sigma^2)$. Then, for $\boldsymbol{W} = [W[0], W[1], \ldots, W[N-1]]^T$,

$$f_{\boldsymbol{W}}(\boldsymbol{w}) = \mathcal{N}(\boldsymbol{w} \,|\, \mathbf{0}, \sigma^2 I).$$

Suppose now that, for these $W[n]$,

$$Y[n] = \theta + W[n]$$

where θ is a constant. What is the joint pdf of $Y[0], Y[1], \ldots$, and $Y[N-1]$? This pdf is the pdf of the vector $\boldsymbol{Y} = [Y[0], Y[1], \ldots, Y[N-1]]^T$:

$$\boldsymbol{Y} = \mathbf{1}\,\theta + \boldsymbol{W}$$

where $\mathbf{1}$ is an $N \times 1$ vector of ones. Now,

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \mathcal{N}(\boldsymbol{y} \,|\, \mathbf{1}\,\theta, \sigma^2 I).$$

Since θ is a constant,

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = f_{\boldsymbol{Y} \,|\, \theta}(\boldsymbol{y} \,|\, \theta).$$

Gaussian Random Vectors

A real-valued random vector $\boldsymbol{X} = [X_1, X_2, \ldots, X_n]^T$ with

- mean $\boldsymbol{\mu}$ and
- covariance matrix Σ with determinant $|\Sigma| > 0$ (i.e. Σ is positive definite)

is a *Gaussian random vector* (or $X_1, X_2, \ldots, X_n$ are *jointly Gaussian RVs*) if and only if its joint pdf is

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{|2\pi\Sigma|^{1/2}} \exp[-\tfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})]. \tag{8}$$

Verify that, for $n = 2$, this joint pdf reduces to the two-dimensional pdf in (6).

Notation: We use $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ to denote a Gaussian random vector. Since Σ is positive definite, Σ^{-1} is also positive definite and, for $\boldsymbol{x} \neq \boldsymbol{\mu}$,

$$(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) > 0$$

which means that the contours of the multivariate Gaussian pdf in (8) are ellipsoids.

The Gaussian random vector $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 I_n)$ (where I_n denotes the identity matrix of size n) is called *white* — contours of the pdf of a white Gaussian random vector are spheres centered at the origin.

Properties of Real-valued Gaussian Random Vectors

Property 1: For a Gaussian random vector, "uncorrelation" implies independence.

This is easy to verify by setting $\Sigma_{i,j} = 0$ for all $i \neq j$ in the joint pdf, then Σ becomes diagonal and so does Σ^{-1}; then, the joint pdf reduces to the product of marginal pdfs $f_{X_i}(x_i) = \mathcal{N}(\mu_i, \Sigma_{i,i}) = \mathcal{N}(\mu_i, \sigma^2_{X_i})$. Clearly, this property holds for blocks of RVs (subvectors) as well.

Property 2: A linear transform of a Gaussian random vector $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{X}}, \Sigma_{\boldsymbol{X}})$ yields a Gaussian random vector:

$$\boldsymbol{Y} = A\,\boldsymbol{X} \sim \mathcal{N}(A\,\boldsymbol{\mu}_{\boldsymbol{X}}, A\,\Sigma_{\boldsymbol{X}}\,A^T).$$

It is easy to show that $\mathrm{E}_{\boldsymbol{Y}}[\boldsymbol{Y}] = A\,\boldsymbol{\mu}_{\boldsymbol{X}}$ and $\mathrm{cov}_{\boldsymbol{Y}}(\boldsymbol{Y}) = \Sigma_{\boldsymbol{Y}} = A\,\Sigma_{\boldsymbol{X}}\,A^T$. So

$$\mathrm{E}_{\boldsymbol{Y}}[\boldsymbol{Y}] = \mathrm{E}_{\boldsymbol{X}}[A\,\boldsymbol{X}] = A\,\mathrm{E}_{\boldsymbol{X}}[\boldsymbol{X}] = A\,\boldsymbol{\mu}_{\boldsymbol{X}}$$

and

$$\begin{aligned}
\Sigma_{\boldsymbol{Y}} &= \mathrm{E}_{\boldsymbol{Y}}[(\boldsymbol{Y} - \mathrm{E}_{\boldsymbol{Y}}[\boldsymbol{Y}])\,(\boldsymbol{Y} - \mathrm{E}_{\boldsymbol{Y}}[\boldsymbol{Y}])^T] \\
&= \mathrm{E}_{\boldsymbol{X}}[(A\,\boldsymbol{X} - A\,\boldsymbol{\mu}_{\boldsymbol{X}})\,(A\,\boldsymbol{X} - A\,\boldsymbol{\mu}_{\boldsymbol{X}})^T] \\
&= A\,\mathrm{E}_{\boldsymbol{X}}[(\boldsymbol{X} - \boldsymbol{\mu}_{\boldsymbol{X}})\,(\boldsymbol{X} - \boldsymbol{\mu}_{\boldsymbol{X}})^T]\,A^T = A\,\Sigma_{\boldsymbol{X}}\,A^T.
\end{aligned}$$

Of course, if we use the definition of a Gaussian random vector in (8), we cannot yet claim that $\boldsymbol{Y}$ is a Gaussian random vector. (For a different definition of a Gaussian random vector, we would be done right here.)

Proof. We need to verify that the joint pdf of $\boldsymbol{Y}$ indeed has the right form. Here, we decide to take the equivalent (easier) task and verify that the *characteristic function* of $\boldsymbol{Y}$ has the right form.

Definition. *Suppose $\boldsymbol{X} \sim f_{\boldsymbol{X}}(\boldsymbol{X})$. Then the characteristic function of $\boldsymbol{X}$ is given by*

$$\Phi_{\boldsymbol{X}}(\boldsymbol{\omega}) = \mathrm{E}_{\boldsymbol{X}}[\exp(j\,\boldsymbol{\omega}^T\,\boldsymbol{X})]$$

where $\boldsymbol{\omega}$ is an n-dimensional real-valued vector and $j = \sqrt{-1}$.

Thus

$$\Phi_{\boldsymbol{X}}(\boldsymbol{\omega}) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_{\boldsymbol{X}}(\boldsymbol{x}) \exp(j\,\boldsymbol{\omega}^T\,\boldsymbol{x})\,d\boldsymbol{x}$$

proportional to the inverse multi-dimensional Fourier transform of $f_{\boldsymbol{X}}(\boldsymbol{x})$; therefore, we can find $f_{\boldsymbol{X}}(\boldsymbol{x})$ by taking the Fourier transform (with the appropriate proportionality factor):

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{(2\pi)^n} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \Phi_{\boldsymbol{X}}(\boldsymbol{\omega}) \exp(-j\,\boldsymbol{\omega}^T\,\boldsymbol{x})\,d\boldsymbol{x}$$

Example: The characteristic function for $X \sim \mathcal{N}(\mu, \sigma^2)$ is given by

$$\Phi_X(\omega) = \exp(-\tfrac{1}{2}\,\omega^2\,\sigma^2 + j\,\mu\,\omega) \tag{9}$$

and for a Gaussian random vector $\boldsymbol{Z} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$,

$$\Phi_{\boldsymbol{Z}}(\boldsymbol{\omega}) = \exp(-\tfrac{1}{2}\,\boldsymbol{\omega}^T\,\Sigma\boldsymbol{\omega} + j\,\boldsymbol{\omega}^T\,\boldsymbol{\mu}). \tag{10}$$

Now, go back to our proof: the characteristic function of $\boldsymbol{Y} = A\boldsymbol{X}$ is

$$\begin{aligned}
\Phi_{\boldsymbol{Y}}(\boldsymbol{\omega}) &= \mathrm{E}_{\boldsymbol{Y}}[\exp(j\,\boldsymbol{\omega}^T\,\boldsymbol{Y})] \\
&= \mathrm{E}_{\boldsymbol{X}}[\exp(j\,\boldsymbol{\omega}^T\,A\,\boldsymbol{X})] \\
&= \exp(-\tfrac{1}{2}\,\boldsymbol{\omega}^T\,A\,\Sigma_{\boldsymbol{X}}\,A^T\,\boldsymbol{\omega} + j\,\boldsymbol{\omega}^T\,A\,\boldsymbol{\mu}_{\boldsymbol{X}}).
\end{aligned}$$

Thus

$$\boldsymbol{Y} = A\,\boldsymbol{X} \sim \mathcal{N}(A\,\boldsymbol{\mu}_{\boldsymbol{X}}, A\Sigma_{\boldsymbol{X}}A^T).$$

Example: Let

$$\mathbf{X} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\right)$$

Find the joint pdf of

$$\mathbf{Y} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\mathbf{X}$$

Solution: From Property 2, we conclude that

$$\mathbf{Y} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\right) = \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} 7 & 3 \\ 3 & 2 \end{bmatrix}\right)$$

Property 3: Marginals of a Gaussian random vector are Gaussian, i.e. if $\boldsymbol{X}$ is a Gaussian random vector, then, for any $\{i_1, i_2, \ldots, i_k\} \subset \{1, 2, \ldots, n\}$,

$$\boldsymbol{Y} = \begin{bmatrix} X_{i_1} \\ X_{i_2} \\ \vdots \\ X_{i_k} \end{bmatrix}$$

is a Gaussian random vector. To show this, we use Property 2. Here is an example with $n = 3$ and $\boldsymbol{Y} = \begin{bmatrix} X_1 \\ X_3 \end{bmatrix}$. We set

$$\boldsymbol{Y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

thus

$$\boldsymbol{Y} \sim \mathcal{N}\Big(\begin{bmatrix} \mu_1 \\ \mu_3 \end{bmatrix}, \begin{bmatrix} \Sigma_{1,1} & \Sigma_{1,3} \\ \Sigma_{3,1} & \Sigma_{3,3} \end{bmatrix} \Big).$$

Here

$$\mathrm{E}_{\boldsymbol{X}} \left\{ \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \right\} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}$$

and

$$\mathrm{cov}_{\boldsymbol{X}} \left\{ \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \right\} = \begin{bmatrix} \Sigma_{1,1} & \Sigma_{1,2} & \Sigma_{1,3} \\ \Sigma_{2,1} & \Sigma_{2,2} & \Sigma_{2,3} \\ \Sigma_{3,1} & \Sigma_{3,2} & \Sigma_{3,3} \end{bmatrix}$$

and note that

$$\begin{bmatrix} \mu_1 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}$$

and

$$\begin{bmatrix} \Sigma_{1,1} & \Sigma_{1,3} \\ \Sigma_{3,1} & \Sigma_{3,3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Sigma_{1,1} & \Sigma_{1,2} & \Sigma_{1,3} \\ \Sigma_{2,1} & \Sigma_{2,2} & \Sigma_{2,3} \\ \Sigma_{3,1} & \Sigma_{3,2} & \Sigma_{3,3} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

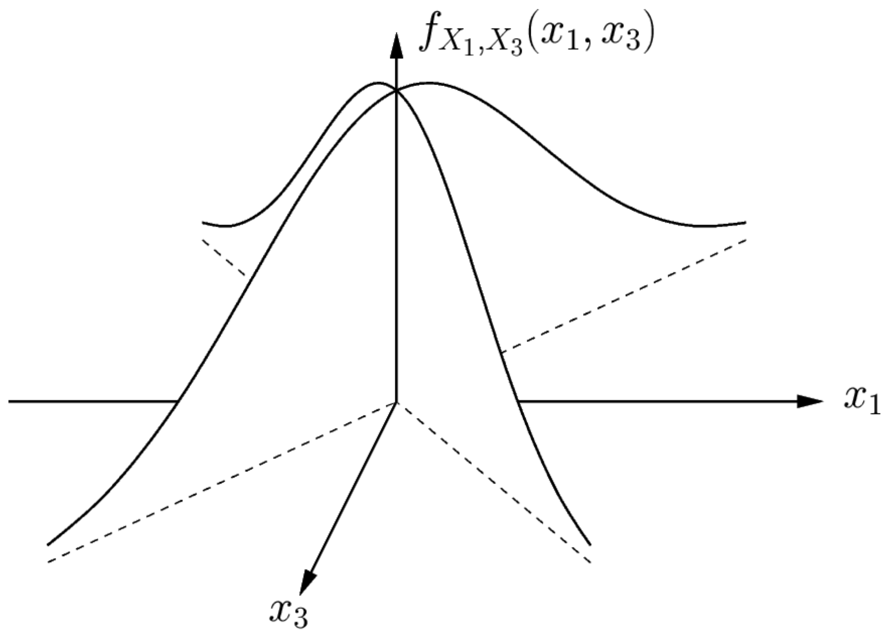
The converse of Property 3 does not hold in general; here is a counterexample:

Example: Suppose $X_1 \sim \mathcal{N}(0,1)$ and

$$X_2 = \begin{cases} 1, & \text{w.p. } \frac{1}{2} \\ -1, & \text{w.p. } \frac{1}{2} \end{cases}$$

are independent RVs and consider $X_3 = X_1 \, X_2$. Observe that

- $X_3 \sim \mathcal{N}(0,1)$ and
- $f_{X_1,X_3}(x_1, x_3)$ is *not* a jointly Gaussian pdf.



□

Property 4: Conditionals of Gaussian random vectors are Gaussian, i.e. if

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix} \sim \mathcal{N}\Big(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{bmatrix} \Big)$$

then

$$\{\boldsymbol{X}_2 \,|\, \boldsymbol{X}_1 = \boldsymbol{x}_1\} \sim \mathcal{N}\Big(\Sigma_{2,1}\Sigma_{1,1}^{-1}(\boldsymbol{x}_1 - \boldsymbol{\mu}_1) + \boldsymbol{\mu}_2, \, \Sigma_{2,2} - \Sigma_{2,1}\Sigma_{1,1}^{-1}\Sigma_{1,2} \Big)$$

and

$$\{\boldsymbol{X}_1 \,|\, \boldsymbol{X}_2 = \boldsymbol{x}_2\} \sim \mathcal{N}\Big(\Sigma_{1,2}\Sigma_{2,2}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2) + \boldsymbol{\mu}_1, \, \Sigma_{1,1} - \Sigma_{1,2}\Sigma_{2,2}^{-1}\Sigma_{2,1} \Big).$$

Example: Compare this result to the case of $n = 2$ in (7):

$$\{X_2 \,|\, X_1 = x_1\} \sim \mathcal{N}\Big(\frac{\Sigma_{2,1}}{\Sigma_{1,1}}\,(x_1 - \mu_1) + \mu_2, \,\Sigma_{2,2} - \frac{\Sigma_{1,2}^2}{\Sigma_{1,1}}\Big).$$

In particular, having $X = X_2$ and $Y = X_1, y = x_1$, this result becomes:

$$\{X \,|\, Y = y\} \sim \mathcal{N}\Big(\frac{\sigma_{X,Y}}{\sigma_Y^2}\,(y - \mu_Y) + \mu_X, \sigma_X^2 - \frac{\sigma_{X,Y}^2}{\sigma_Y^2}\Big)$$

where $\sigma_{X,Y} = \mathrm{cov}_{X,Y}(X, Y), \sigma_X^2 = \mathrm{cov}_{X,X}(X, X) = \mathrm{var}_X(X)$, and $\sigma_Y^2 = \mathrm{cov}_{Y,Y}(Y, Y) = \mathrm{var}_Y(Y)$. Now, it is clear that

$$\rho_{X,Y} = \frac{\sigma_{X,Y}}{\sigma_X\,\sigma_Y}$$

where $\sigma_X = \sqrt{\sigma_X^2} > 0$ and $\sigma_Y = \sqrt{\sigma_Y^2} > 0$.

Example: Suppose

$$\begin{bmatrix} X_1 \\ -- \\ X_2 \\ X_3 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 1 \\ -- \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 & | & 2 & 1 \\ -- & | & -- & -- \\ 2 & | & 5 & 2 \\ 1 & | & 2 & 9 \end{bmatrix}\right)$$

From Property 4, it follows that

$$E_{X_2|X_1}(\mathbf{X}_2|X_1 = x) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}(x-1) + \underbrace{\begin{bmatrix} 2 \\ 2 \end{bmatrix}}_{\mu_2} = \begin{bmatrix} 2x \\ x+1 \end{bmatrix}$$

$$\Sigma_{\{\mathbf{X}_2|X_1=x\}} = \begin{bmatrix} 5 & 2 \\ 2 & 9 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 8 \end{bmatrix}$$

Property 5: If $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ then

$(\boldsymbol{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) \sim \chi^2_d$ (Chi-square in your distr. table).

Additive Gaussian Noise Channel

Consider a communication channel with input

$$X \sim \mathcal{N}(\mu_X, \tau_X^2)$$

and noise

$$W \sim \mathcal{N}(0, \sigma^2)$$

where X and W are independent and the measurement Y is

$$Y = X + W.$$

Since X and W are independent, we have

$$f_{X,W}(x, w) = f_X(x)\, f_W(x)$$

and

$$\begin{bmatrix} X \\ W \end{bmatrix} \sim \mathcal{N}\Big(\begin{bmatrix} \mu_X \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_X^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} \Big).$$

What is $f_{Y \,|\, X}(y \,|\, x)$? Since

$$\{Y \,|\, X = x\} = x + W \sim \mathcal{N}(x, \sigma^2)$$

we have

$$f_{Y \,|\, X}(y \,|\, x) = \mathcal{N}(y \,|\, x, \sigma^2).$$

How about $f_Y(y)$? Construct the joint pdf $f_{X,Y}(x,y)$ of X and Y: since

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} X \\ W \end{bmatrix}$$

then

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathcal{N}\Big(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \mu_X \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \tau_X^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \Big)$$

yielding

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathcal{N}\Big(\begin{bmatrix} \mu_X \\ \mu_X \end{bmatrix}, \begin{bmatrix} \tau_X^2 & \tau_X^2 \\ \tau_X^2 & \tau_X^2 + \sigma^2 \end{bmatrix} \Big).$$

Therefore,

$$f_Y(y) = \mathcal{N}\Big(y \,\Big|\, \mu_X, \tau_X^2 + \sigma^2\Big).$$

Complex Gaussian Distribution

Consider joint pdf of real and imaginary part of an $n \times 1$ complex vector $\boldsymbol{Z}$

$$\boldsymbol{Z} = \boldsymbol{U} + j\,\boldsymbol{V}.$$

Assume

$$\boldsymbol{X} = \left[\begin{array}{c} \boldsymbol{U} \\ \boldsymbol{Y} \end{array} \right].$$

The $2\,n$-variate Gaussian pdf of the (real!) vector $\boldsymbol{X}$ is

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{\sqrt{(2\,\pi)^{2\,n} |\Sigma_{\boldsymbol{X}}|}} \exp\left[-\tfrac{1}{2}\,(\boldsymbol{z} - \boldsymbol{\mu}_{\boldsymbol{X}})^T\,\Sigma_{\boldsymbol{X}}^{-1}(\boldsymbol{z} - \boldsymbol{\mu}_{\boldsymbol{X}})\right]$$

where

$$\boldsymbol{\mu}_{\boldsymbol{X}} \quad = \quad \left[\begin{array}{c} \boldsymbol{\mu}_{\boldsymbol{U}} \\ \boldsymbol{\mu}_{\boldsymbol{V}} \end{array} \right], \quad \Sigma_{\boldsymbol{X}} = \left[\begin{array}{cc} \Sigma_{\boldsymbol{UU}} & \Sigma_{\boldsymbol{UV}} \\ \Sigma_{\boldsymbol{VU}} & \Sigma_{\boldsymbol{VV}} \end{array} \right].$$

Therefore,

$$\Pr\{\boldsymbol{x} \in A\} = \int_{\boldsymbol{x} \in A} f_{\boldsymbol{X}}(\boldsymbol{x})\,d\boldsymbol{x}.$$

Complex Gaussian Distribution (cont.)

Assume that $\Sigma_{\boldsymbol{X}}$ has a special structure:

$$\Sigma_{\boldsymbol{UU}} = \Sigma_{\boldsymbol{VV}} \quad \text{and} \quad \Sigma_{\boldsymbol{UV}} = -\Sigma_{\boldsymbol{VU}}.$$

[Note that $\Sigma_{\boldsymbol{UV}} = \Sigma_{\boldsymbol{VU}}^T$ has to hold as well.] Then, we can define a complex Gaussian pdf of $\boldsymbol{Z}$ as

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \frac{1}{\pi^n \, |\Sigma_{\boldsymbol{X}}|} \exp\left[-(\boldsymbol{z} - \boldsymbol{\mu}_{\boldsymbol{Z}})^H \Sigma_{\boldsymbol{Z}}^{-1} (\boldsymbol{z} - \boldsymbol{\mu}_{\boldsymbol{Z}})\right]$$

where

$$\begin{aligned}
\boldsymbol{\mu}_{\boldsymbol{Z}} &= \boldsymbol{\mu}_{\boldsymbol{U}} + j\,\boldsymbol{\mu}_{\boldsymbol{V}} \\
\Sigma_{\boldsymbol{Z}} &= \mathrm{E}_{\boldsymbol{Z}}\{(\boldsymbol{Z} - \boldsymbol{\mu}_{\boldsymbol{Z}})\,(\boldsymbol{Z} - \boldsymbol{\mu}_{\boldsymbol{Z}})^H\} = 2\,(\Sigma_{\boldsymbol{UU}} + j\,\Sigma_{\boldsymbol{VU}}) \\
\boldsymbol{0} &= \mathrm{E}_{\boldsymbol{Z}}\{(\boldsymbol{Z} - \boldsymbol{\mu}_{\boldsymbol{Z}})\,(\boldsymbol{Z} - \boldsymbol{\mu}_{\boldsymbol{Z}})^T\}.
\end{aligned}$$