Random Vectors and Multivariate Probability Models

Christ D. Richmond

ECE 581 Random Signals and Noise
Department of Electrical and Computer Engineering

October 3, 2024

Table of contents

Remarks

- So far we've considered probability models that allow us to characterize one or two (i.e. a pair of) random variables, e.g. the pair (X, Y).
- We now consider probability theory to handle multivariable observations for more than two random variables, e.g. a random vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ of n variables.
- The theory presented will be seen to be straightforward extensions of ideas we've already developed for joint PMFs, PDFs, and CDFs.

Multivariate Joint PMF, CDF, and PDF

- The joint PMF for a set of n discrete random variables represented collectively as the vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ is defined as

$$\begin{aligned} \Pr(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) &= P_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) \\ \Pr(\mathbf{X} = \mathbf{x}) &= P_{\mathbf{X}}(\mathbf{x}), \end{aligned}$$

  and the probability of event $B \subseteq \mathbb{R}^N$ in the support of $P_{\mathbf{X}}$, i.e. $S_{\mathbf{X}}$, is given by

$$\Pr(B) = \sum_{(x_1, x_2, \ldots, x_n) \in B} P_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = \sum_{\mathbf{x} \in B} P_{\mathbf{X}}(\mathbf{x}).$$

- The joint PDF for a set of n continuous random variables represented collectively as the vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ is a function $f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = f_{\mathbf{X}}(\mathbf{x})$ defined such that for any subset $B \subseteq \mathbb{R}^N$ in $S_{\mathbf{X}}$ the support $f_{\mathbf{X}}$ we have

$$\begin{aligned} \Pr(B) &= \int \int \cdots \int_{(x_1, x_2, \ldots, x_n) \in B} f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) dx_1 dx_2 \cdots dx_n \\ &= \int_{\mathbf{x} \in B} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}. \end{aligned}$$

  - Note that the probability of event B is given by the n-dimensional volume contained under the density in set B.

Multivariate Joint PMF, CDF, and PDF Cont.

- The joint CDF for a set of n random variables represented collectively as the vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ is defined as

$$\begin{aligned} \Pr(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) &= F_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) \\ \Pr(\mathbf{X} \leq \mathbf{x}) &= F_{\mathbf{X}}(\mathbf{x}). \end{aligned} \tag{1}$$

  - It is noteworthy that (1) is the definition of the CDF whether any, or all of the X_i are discrete, continuous, or even mixed/hybrid.

- If X_i are all continuous random variables, then

$$F_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f_{X_1,X_2,\ldots,X_n}(a_1, a_2, \ldots, a_n) da_1 da_2 \cdots da_n,$$

  and by the fundamental theorem of calculus we know

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

- By the axioms of probability we obtain for discrete r.v.'s
  1. $P_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = P_{\mathbf{X}}(\mathbf{x}) \geq 0,$
  2. $\displaystyle\sum_{x_1 \in S_{X_1}} \sum_{x_2 \in S_{X_2}} \cdots \sum_{x_n \in S_{X_n}} P_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = \sum_{\mathbf{x} \in S_{\mathbf{X}}} P_{\mathbf{X}}(\mathbf{x}) = 1.$

Multivariate Joint PMF, CDF, and PDF Cont.

- By the axioms of probability we obtain for continous r.v.'s:
  1. $f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) = f_{\mathbf{X}}(\mathbf{x}) \geq 0$,
  2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{X_1,X_2,\ldots,X_n}(a_1,a_2,\ldots,a_n)da_1da_2\cdots da_n = \int_{\mathbf{x}\in S_{\mathbf{X}}} f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} = 1.$
- For variables $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$, if we choose any partition such that $\mathbf{X}_1 = [X_{i_1}, X_{i_2}, \ldots, X_{i_m}]^T$ and $\mathbf{X}_2 = [X_{i_{m+1}}, X_{i_{m+2}}, \ldots, X_{i_n}]^T$ where each $i_k \in \{1, 2, \ldots, n\}$ is unique, then the marginals can be obtained via

$$P_{\mathbf{X}_1}(\mathbf{x}_1) = \sum_{\mathbf{x}_2\in S_{\mathbf{X}_2}} P_{\mathbf{X}_1,\mathbf{X}_2}(\mathbf{x}_1,\mathbf{x}_2), \qquad f_{\mathbf{X}_1}(\mathbf{x}_1) = \int_{\mathbf{x}_2\in S_{\mathbf{X}_2}} f_{\mathbf{X}_1,\mathbf{X}_2}(\mathbf{x}_1,\mathbf{x}_2)d\mathbf{x}_2$$

  for discrete and continuous r.v.'s respectively where it is implicit that $P_{\mathbf{X}} = P_{\mathbf{X}_1,\mathbf{X}_2}$ and $f_{\mathbf{X}} = f_{\mathbf{X}_1,\mathbf{X}_2}$, i.e. the concatenation of $\mathbf{X}_1$ and $\mathbf{X}_2$ constitutes the same set of random variables present in $\mathbf{X}$.
  - As an example, if $\mathbf{X} = [W, X, Y, Z]^T$, then some marginals are given by

$$P_{X,Y,Z}(x,y,z) = \sum_{w\in S_W} P_{W,X,Y,Z}(w,x,y,z), \quad P_{W,Z}(w,z) = \sum_{x\in S_X}\sum_{y\in S_Y} P_{W,X,Y,Z}(w,x,y,z)$$

for discrete r.v.'s; and

$$f_{X,Y,Z}(x,y,z) = \int_{w\in S_W} f_{W,X,Y,Z}(w,x,y,z)dw, \quad f_{W,Z}(w,z) = \int_{x\in S_X}\int_{y\in S_Y} f_{W,X,Y,Z}(w,x,y,z)dxdy.$$

for continuous r.v.'s.

Multivariate Joint PMF, CDF, and PDF Cont.

- Consider the pair of random vectors $\mathbf{X}$ and $\mathbf{Y}$ where $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ and $\mathbf{Y} = [Y_1, Y_2, \ldots, Y_m]^T$ and note the following:
  - The joint PMF of $\mathbf{X}$ and $\mathbf{Y}$ is given by

$$P_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) = P_{X_1,X_2,\ldots,X_n,Y_1,Y_2,\ldots,Y_m}(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m).$$

  - The joint PDF of $\mathbf{X}$ and $\mathbf{Y}$ is given by

$$f_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) = f_{X_1,X_2,\ldots,X_n,Y_1,Y_2,\ldots,Y_m}(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m).$$

  - The joint CDF of $\mathbf{X}$ and $\mathbf{Y}$ is given by

$$F_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) = F_{X_1,X_2,\ldots,X_n,Y_1,Y_2,\ldots,Y_m}(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m).$$

Independence of Random Variables

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

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

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

  for all $x_1, x_2, \ldots, x_n$, for discrete r.v.'s and continuous r.v.'s respectively.
- Random variables $X_1, X_2, \ldots, X_n$ are said to be independent identically distributed (i.i.d.) if

$$P_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = P_X(x_1) \cdot P_X(x_2) \cdots P_X(x_n)$$

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

  for all $x_1, x_2, \ldots, x_n$, for discrete r.v.'s and continuous r.v.'s respectively.
  - Note that $P_{X_i} = P_X$ for all X_i, $i = 1, 2, \ldots, n$ for the discrete r.v.'s.
  - Note that $f_{X_i} = f_X$ for all X_i, $i = 1, 2, \ldots, n$ for the continuous r.v.'s.
- A pair of random vectors $\mathbf{X}$ and $\mathbf{Y}$ is said to be independent if

$$P_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) = P_{\mathbf{X}}(\mathbf{x}) \cdot P_{\mathbf{Y}}(\mathbf{y}), \quad f_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}) = f_{\mathbf{X}}(\mathbf{x}) \cdot f_{\mathbf{Y}}(\mathbf{y}),$$

  for discrete r.v.'s and continuous r.v.'s respectively.
- **Quiz 5.10 Yates/Goodman**
  - To be discussed in class

Quiz 5.10

The random variables $Y_1, \ldots, Y_4$ have the joint PDF

$$f_{Y_1,\ldots,Y_4}(y_1, \ldots, y_4) = \begin{cases} 4 & 0 \le y_1 \le y_2 \le 1, 0 \le y_3 \le y_4 \le 1, \\ 0 & \text{otherwise.} \end{cases} \qquad (1)$$

Let C denote the event that $\max_i Y_i \le 1/2$. Find $P[C]$.

Mean of Scalar Valued Function

- We can define scalar valued functions of multiple random variables, i.e. functions of vectors, as $g(X_1, X_2, \ldots, X_n) = g(\mathbf{X})$.
  - Because $g(\mathbf{X})$ is a function of random variables, it is itself a random variable.
  - An example of such a function is $g(\mathbf{X}) = \mathbf{a}^T\mathbf{X} = a_1X_1 + a_2X_2 + \cdots + a_nX_n$.
- The random variable $g(\mathbf{X})$ has expected value defined as

$$E\{g(\mathbf{X})\} = \sum_{x_1 \in S_{X_1}} \sum_{x_2 \in S_{X_2}} \cdots \sum_{x_n \in S_{X_n}} g(x_1, x_2, \ldots, x_n) P_{\mathbf{X}}(x_1, x_2, \ldots, x_n) = \sum_{\mathbf{x} \in S_{\mathbf{X}}} g(\mathbf{x}) P_{\mathbf{X}}(\mathbf{x}).$$

$$E\{g(\mathbf{X})\} = \int_{x_1 \in S_{X_1}} \int_{x_2 \in S_{X_2}} \cdots \int_{x_n \in S_{X_n}} g(x_1, x_2, \ldots, x_n) f_{\mathbf{X}}(x_1, x_2, \ldots, x_n) dx_1 dx_2 \cdots dx_n$$

$$= \int_{\mathbf{x} \in S_{\mathbf{X}}} g(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad \text{for discrete r.v.'s and continuous r.v.'s respectively.}$$

  - Of course, there is an alternative way to find $E\{g(\mathbf{X})\}$.
  - If we let $W = g(\mathbf{X})$ and we can find its PDF $f_W(w)$ (or PMF $P_W(w)$), then note that

$$E\{g(\mathbf{X})\} = E\{W\} = \begin{cases} \int_{w \in S_W} w \cdot f_W(w) dw, & \text{(continuous r.v.)} \\ \sum_{w \in S_W} w \cdot P_W(w), & \text{(discrete r.v.).} \end{cases}$$

Mean of Scalar Valued Function Cont.

- If $g(\mathbf{X}) = g_1(X_1)g_2(X_2)\cdots g_n(X_n)$ and $X_1, X_2, \ldots, X_n$ are independent continuous r.v.'s, then

$$
\begin{aligned}
E[g(\mathbf{X})] &= \int_{\mathbf{x}\in S_{\mathbf{X}}} g(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \\
&= \int_{x_1\in S_{X_1}}\int_{x_2\in S_{X_2}}\cdots\int_{x_n\in S_{X_n}} g_1(x_1)g_2(x_2)\cdots g_n(x_n) f_{X_1}(x_1) f_{X_2}(x_2)\cdots f_{X_n}(x_n) dx_1 dx_2\cdots dx_n \\
&= \int_{x_1\in S_{X_1}} g_1(x_1) f_{X_1}(x_1) dx_1 \int_{x_2\in S_{X_2}} g_2(x_2) f_{X_2}(x_2) dx_2 \cdots \int_{x_n\in S_{X_n}} g_n(x_n) f_{X_n}(x_n) dx_n \\
&= E[g_1(X_1)]\cdot E[g_2(X_2)]\cdots E[g_n(X_n)].
\end{aligned}
$$

Expected Value of a Vector and Matrix

- The expected value of the random vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ is the vector of expected values:

$$E[\mathbf{X}] = [E[X_1], E[X_2], \ldots, E[X_n]]^T \triangleq \boldsymbol{\mu}_{\mathbf{X}}.$$

- Expectation is a linear operator. Note that for $g(\mathbf{X}) = \mathbf{a}^T\mathbf{X}$

$$\begin{aligned} E[g(\mathbf{X})] &= E[\mathbf{a}^T\mathbf{X}] \\ &= E[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = a_1E[X_1] + a_2E[X_2] + \cdots + a_nE[X_n] \\ &= a_1\mu_{X_1} + a_2\mu_{X_2} + \cdots + a_n\mu_{X_n} \\ &= \mathbf{a}^T\boldsymbol{\mu}_{\mathbf{X}}. \end{aligned}$$

- Similarly, if $\mathbf{A}_{n\times n} = [\mathbf{a}_1|\mathbf{a}_2|\cdots|\mathbf{a}_n]$, then $n\times 1$ vector $\mathbf{A}^T\mathbf{X}$ has rows $\mathbf{a}_i^T\mathbf{X}$ and mean vector

$$E[\mathbf{A}^T\mathbf{X}] = [E[\mathbf{a}_1^T\mathbf{X}], E[\mathbf{a}_2^T\mathbf{X}], \ldots, E[\mathbf{a}_n^T\mathbf{X}]]^T = [\mathbf{a}_1^T\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{a}_2^T\boldsymbol{\mu}_{\mathbf{X}}, \ldots, \mathbf{a}_n^T\boldsymbol{\mu}_{\mathbf{X}}]^T = \mathbf{A}^T\boldsymbol{\mu}_{\mathbf{X}}.$$

- The expected value of a random $n\times m$ matrix $\mathbf{G}$ is the matrix of expected values:

$$\mathbf{G} = \begin{bmatrix} G_{1,1} & G_{1,2} & \ldots & G_{1,m} \\ G_{2,1} & G_{2,2} & \ldots & G_{2,m} \\ \vdots & \vdots & \vdots & \vdots \\ G_{n,1} & G_{n,2} & \ldots & G_{n,m} \end{bmatrix} \Longrightarrow E[\mathbf{G}] = \begin{bmatrix} E[G_{1,1}] & E[G_{1,2}] & \ldots & E[G_{1,m}] \\ E[G_{2,1}] & E[G_{2,2}] & \ldots & E[G_{2,m}] \\ \vdots & \vdots & \vdots & \vdots \\ E[G_{n,1}] & E[G_{n,2}] & \ldots & E[G_{n,m}] \end{bmatrix}. \quad (4)$$

Expected Value of a Vector and Matrix Cont.

- Recall that for a matrix transpose $[\mathbf{G}^T]_{i,k} = [\mathbf{G}]_{k,i}$.
- By the above definition it follows that

$$E[\mathbf{G}^T] = \begin{bmatrix} E[G_{1,1}] & E[G_{2,1}] & \dots & E[G_{n,1}] \\ E[G_{1,2}] & E[G_{2,2}] & \dots & E[G_{n,2}] \\ \vdots & \vdots & \vdots & \vdots \\ E[G_{1,m}] & E[G_{2,m}] & \dots & E[G_{n,m}] \end{bmatrix} = (E[\mathbf{G}])^T, \tag{5}$$

an $m \times n$ matrix.

  - This shows that the expected value of the transpose of a matrix is the transpose of the expected value of said matrix.

- Consider the random $n \times n$ matrix $\tilde{\mathbf{G}}$ and note for vector $\mathbf{a} = [a_1, a_2, \dots, a_n]^T$ that

$$\mathbf{a}^T\tilde{\mathbf{G}}\mathbf{a} = \sum_{i=1}^{n}\sum_{k=1}^{n} a_i a_k \cdot \tilde{G}_{i,k}.$$ Similarly, by (4) we have $$\mathbf{a}^T E[\tilde{\mathbf{G}}]\mathbf{a} = \sum_{i=1}^{n}\sum_{k=1}^{n} a_i a_k \cdot E[\tilde{G}_{i,k}].$$

Thus, by linearity of expectation we see that

$$\mathbf{a}^T E[\tilde{\mathbf{G}}]\mathbf{a} = E[\mathbf{a}^T\tilde{\mathbf{G}}\mathbf{a}]. \tag{6}$$

Correlation Matrix

- The correlation matrix of a random vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ is defined as $n \times n$ matrix

$$\mathbf{R_X} = E[\mathbf{X}\mathbf{X}^T] = \begin{bmatrix} E[X_1X_1] & E[X_1X_2] & \ldots & E[X_1X_n] \\ E[X_2X_1] & E[X_2X_2] & \ldots & E[X_2X_n] \\ \vdots & \vdots & \vdots & \vdots \\ E[X_nX_1] & E[X_nX_2] & \ldots & E[X_nX_n] \end{bmatrix}. \quad (7)$$

  - Note (7) is obtained with $\mathbf{G} = \mathbf{X}\mathbf{X}^T$ in (4), i.e. where $[\mathbf{G}]_{i,k} = G_{i,k} = X_iX_k$.
  - It follows from definition of expected value for a matrix.
  - $[\mathbf{R_X}]_{i,k}$ is the correlation between the pair of random variables X_i and X_k.
  - $\mathbf{R_X}$ is sometimes called the autocorrelation matrix, since the correlation of $\mathbf{X}$ with itself is being determined.

Correlation Matrix Cont.

- Properties of a correlation matrix include:
  1. Symmetric, i.e. $\mathbf{R_X} = \mathbf{R_X}^T$.
     - This follows from $\mathbf{R_X}^T = (E[\mathbf{X}\mathbf{X}^T])^T = E[(\mathbf{X}\mathbf{X}^T)^T] = E[(\mathbf{X}^T)^T(\mathbf{X})^T]$ $= E[\mathbf{X}\mathbf{X}^T] = \mathbf{R_X}$ where we've used (5).
  2. Positive semi-definite, i.e. $\mathbf{a}^T\mathbf{R_X}\mathbf{a} \geq 0$ for all vectors $\mathbf{a} \in \mathbb{R}^N$, $\mathbf{a} \neq \mathbf{0}$.
     - Note $\mathbf{a}^T\mathbf{R_X}\mathbf{a} = \mathbf{a}^T E[\mathbf{X}\mathbf{X}^T]\mathbf{a} = E[\mathbf{a}^T\mathbf{X}\mathbf{X}^T\mathbf{a}] = E[(\mathbf{a}^T\mathbf{X})^2]$ by (6).
     - Let $W = \mathbf{a}^T\mathbf{X}$, a real scalar random variable, and note that

$$E[(\mathbf{a}^T\mathbf{X})^2] = E[W^2] = \begin{cases} \displaystyle\int_{S_W} w^2 f_W(w)dw \geq 0, & \text{(continuous r.v.)} \\ \displaystyle\sum_{S_W} w^2 P_W(w) \geq 0, & \text{(discrete r.v.)} \end{cases}$$

       since sum of real non-negative terms yields a non-negative real number.
  3. Eigenvalues are real and non-negative, i.e. $\lambda_i(\mathbf{R_X}) \geq 0$.
     - If $\mathbf{q}_i$ is an eigenvector, then $\mathbf{R_X}\mathbf{q}_i = \lambda_i\mathbf{q}_i$.
     - Let $\mathbf{a} = \mathbf{q}_i/\|\mathbf{q}_i\|$ in (ii) above; note $\mathbf{a}^T\mathbf{R_X}\mathbf{a} = \dfrac{\mathbf{q}_i^T\mathbf{R_X}\mathbf{q}_i}{\|\mathbf{q}_i\|^2} = \dfrac{\mathbf{q}_i^T\mathbf{q}_i\lambda_i}{\|\mathbf{q}_i\|^2} = \lambda_i$.
     - Thus, by same arguments in (ii), it follows that λ_i is real and $\lambda_i \geq 0$.

Covariance Matrix

- Properties of a correlation matrix cont.:
  - (iv) Eigenvectors are orthonormal, i.e. $\mathbf{R}_\mathbf{X} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$ where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, $\mathbf{Q} = [\mathbf{q}_1|\mathbf{q}_2|\cdots|\mathbf{q}_n]$ with[1] $\mathbf{q}_i^T\mathbf{q}_k = \delta_{i,k}$.
    - The proof of this is involved. See Chapter 8 in [4] for details.
- The covariance matrix of a random vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ is defined as $n \times n$ matrix

$$\mathbf{C}_\mathbf{X} = E[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T] =$$

$$\begin{bmatrix} E[(X_1 - \mu_{X_1})(X_1 - \mu_{X_1})] & E[(X_1 - \mu_{X_1})(X_2 - \mu_{X_2})] & \ldots & E[(X_1 - \mu_{X_1})(X_n - \mu_{X_n})] \\ E[(X_2 - \mu_{X_2})(X_1 - \mu_{X_1})] & E[(X_2 - \mu_{X_2})(X_2 - \mu_{X_2})] & \ldots & E[(X_2 - \mu_{X_2})(X_n - \mu_{X_n})] \\ \vdots & \vdots & \vdots & \vdots \\ E[(X_n - \mu_{X_n})(X_1 - \mu_{X_1})] & E[(X_n - \mu_{X_n})(X_2 - \mu_{X_2})] & \ldots & E[(X_n - \mu_{X_n})(X_n - \mu_{X_n})] \end{bmatrix}. \quad (8)$$

- The covariance matrix can also be expressed as

$$\begin{aligned} \mathbf{C}_\mathbf{X} &= E[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T] = E[\mathbf{X}\mathbf{X}^T - \mathbf{X}\boldsymbol{\mu}_\mathbf{X}^T - \boldsymbol{\mu}_\mathbf{X}\mathbf{X}^T + \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}_\mathbf{X}^T] \\ &= E[\mathbf{X}\mathbf{X}^T] - E[\mathbf{X}\boldsymbol{\mu}_\mathbf{X}^T] - E[\boldsymbol{\mu}_\mathbf{X}\mathbf{X}^T] + E[\boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}_\mathbf{X}^T] \\ &= E[\mathbf{X}\mathbf{X}^T] - E[\mathbf{X}]\boldsymbol{\mu}_\mathbf{X}^T - \boldsymbol{\mu}_\mathbf{X}E[\mathbf{X}^T] + \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}_\mathbf{X}^T \\ &= \mathbf{R}_\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}_\mathbf{X}^T - \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}_\mathbf{X}^T + \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}_\mathbf{X}^T \\ &= \mathbf{R}_\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}\boldsymbol{\mu}_\mathbf{X}^T. \end{aligned} \quad (9)$$

[1] $\delta_{i,k}$ is the discrete delta function defined to equal 1 when $i = k$ and 0 when $i \neq k$.

Covariance Matrix Cont.

- This $n \times n$ matrix is the multivariate generalization of the univariate $\text{Var}[X] = E[X^2] - E^2[X]$ and bivariate $\text{Cov}[X_i, X_k] = E[X_i X_k] - E[X_i]E[X_k]$.
- Clearly, $[\mathbf{C_X}]_{i,i} = \text{Var}[X_i] = \sigma^2_{X_i}$, and $[\mathbf{C_X}]_{i,k} = \text{Cov}[X_i, X_k] = \sigma_{X_i, X_k}$.
- $\mathbf{C_X}$ is sometimes called the autocovariance matrix, since the covariance of $\mathbf{X}$ with itself is being determined.
- Note from (9) that $\mathbf{R_X} = \mathbf{C_X} + \boldsymbol{\mu}_\mathbf{X} \boldsymbol{\mu}_\mathbf{X}^T$, and if $E[\mathbf{X}] = \boldsymbol{\mu}_\mathbf{X} = \mathbf{0}$, then $\mathbf{R_X} = \mathbf{C_X}$.
- Properties of the covariance matrix are the same as the correlation matrix:
  1. Symmetric, i.e. $\mathbf{C_X} = \mathbf{C_X}^T$.
  2. Positive semi-definite, i.e. $\mathbf{a}^T \mathbf{C_X} \mathbf{a} \geq 0$ for all vectors $\mathbf{a} \in \mathbb{R}^N$, $\mathbf{a} \neq \mathbf{0}$.
  3. Eigenvalues are real and non-negative, i.e. $\tilde{\lambda}_i(\mathbf{C_X}) \geq 0$.
  4. Eigenvectors are orthonormal, i.e. $\mathbf{C_X} = \tilde{\mathbf{Q}} \tilde{\boldsymbol{\Lambda}} \tilde{\mathbf{Q}}^T$ where $\tilde{\boldsymbol{\Lambda}} = \text{diag}(\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_n)$, $\tilde{\mathbf{Q}} = [\tilde{\mathbf{q}}_1 | \tilde{\mathbf{q}}_2 | \cdots | \tilde{\mathbf{q}}_n]$ with $\tilde{\mathbf{q}}_i^T \tilde{\mathbf{q}}_k = \delta_{i,k}$.
- Note vector $\mathbf{X}_0 = \mathbf{X} - \boldsymbol{\mu}_\mathbf{X}$ has correlation matrix

$$\mathbf{R}_{\mathbf{X}_0} = E[\mathbf{X}_0 \mathbf{X}_0^T] = E[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T] = \mathbf{C_X}.$$

  - Thus, a covariance matrix is a special case of a correlation matrix, i.e. a covariance matrix is the correlation matrix for centered variables (i.e. having zero mean).

Cross Correlation Matrix

- The cross correlation matrix between vectors $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ and $\mathbf{Y} = [Y_1, Y_2, \ldots, Y_m]^T$ is defined as the $n \times m$ matrix

$$\mathbf{R_{XY}} = E[\mathbf{XY}^T] = \begin{bmatrix} E[X_1Y_1] & E[X_1Y_2] & \ldots & E[X_1Y_m] \\ E[X_2Y_1] & E[X_2Y_2] & \ldots & E[X_2Y_m] \\ \vdots & \vdots & \vdots & \vdots \\ E[X_nY_1] & E[X_nY_2] & \ldots & E[X_nY_m] \end{bmatrix}. \quad (10)$$

  - Note (10) is obtained with $\mathbf{G} = \mathbf{XY}^T$ in (4), i.e. where $[\mathbf{G}]_{i,k} = G_{i,k} = X_iY_k$.
  - It follows from definition of expected value for a matrix.
  - $[\mathbf{R_{XY}}]_{i,k}$ is the correlation between the pair of random variables X_i and Y_k.
- Note that $\mathbf{R_{YX}} = \mathbf{R}_{\mathbf{XY}}^T$, i.e. $E[\mathbf{YX}^T] = E[(\mathbf{XY}^T)^T] = (E[\mathbf{XY}^T])^T$ where (5) is used.
- Note also if $\mathbf{Y} = \mathbf{X}$, then $\mathbf{R_{XX}} = \mathbf{R_X}$.

Cross Covariance Matrix

- The cross covariance matrix between vectors $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ and $\mathbf{Y} = [Y_1, Y_2, \ldots, Y_m]^T$ is defined as the $n \times m$ matrix

$$\mathbf{C_{XY}} = E[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})(\mathbf{Y} - \boldsymbol{\mu}_\mathbf{Y})^T] =$$

$$\begin{bmatrix} E[(X_1 - \mu_{X_1})(Y_1 - \mu_{Y_1})] & E[(X_1 - \mu_{X_1})(Y_2 - \mu_{Y_2})] & \ldots & E[(X_1 - \mu_{X_1})(Y_m - \mu_{Y_m})] \\ E[(X_2 - \mu_{X_2})(Y_1 - \mu_{Y_1})] & E[(X_2 - \mu_{X_2})(Y_2 - \mu_{Y_2})] & \ldots & E[(X_2 - \mu_{X_2})(Y_m - \mu_{Y_m})] \\ \vdots & \vdots & \vdots & \vdots \\ E[(X_n - \mu_{X_n})(Y_1 - \mu_{Y_1})] & E[(X_n - \mu_{X_n})(Y_2 - \mu_{Y_2})] & \ldots & E[(X_n - \mu_{X_n})(Y_m - \mu_{Y_m})] \end{bmatrix}.$$

  - This matrix is the multivariate generalization of the bivariate $\text{Cov}[X, Y] = E[XY] - E[X]E[Y]$.
  - Clearly, $[\mathbf{C_{XY}}]_{i,k} = \text{Cov}[X_i, Y_k] = \sigma_{X_i,Y_k} = E[X_i Y_k] - E[X_i]E[Y_k]$.
- Note that $\mathbf{C_{YX}} = \mathbf{C}_\mathbf{XY}^T$.
- Note also if $\mathbf{Y} = \mathbf{X}$, then $\mathbf{C_{XX}} = \mathbf{C_X}$.
- Vectors $\mathbf{X}$ and $\mathbf{Y}$ are said to be uncorrelated if $\mathbf{C_{XY}} = \mathbf{0}_{n \times m}$.

Linear Function of Correlated Random Vector Pair

- Given pair of random vectors $\mathbf{X}$ and $\mathbf{Y}$, consider $\mathbf{Z} = \mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{Y} + \mathbf{d}$:
  - $\mathbf{Z}$ has mean

$$E\{\mathbf{Z}\} = E\{\mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{Y} + \mathbf{d}\} = \mathbf{A}E\{\mathbf{X}\} + \mathbf{B}E\{\mathbf{Y}\} + \mathbf{d} \triangleq \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{B}\boldsymbol{\mu}_{\mathbf{Y}} + \mathbf{d} \triangleq \boldsymbol{\mu}_{\mathbf{Z}}.$$

  - Noting $\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{Z}} = \mathbf{A}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) + \mathbf{B}(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}}) \Longrightarrow$

$$(\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{Z}})(\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{Z}})^T = \mathbf{A}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T\mathbf{A}^T + \mathbf{A}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})^T\mathbf{B}^T$$
$$+\mathbf{B}(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T\mathbf{A}^T + \mathbf{B}(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})^T\mathbf{B}^T.$$

  - Thus, $\mathbf{Z}$ has covariance

$$\mathbf{C}_{\mathbf{Z}} = E\{(\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{Z}})(\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{Z}})^T\} = \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}^T + \mathbf{A}\mathbf{C}_{\mathbf{XY}}\mathbf{B}^T + \mathbf{B}\mathbf{C}_{\mathbf{YX}}\mathbf{A}^T + \mathbf{B}\mathbf{C}_{\mathbf{Y}}\mathbf{B}^T.$$

    - If $\mathbf{X}$ and $\mathbf{Y}$ are uncorrelated, then $\mathbf{C}_{\mathbf{Z}} = \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}^T + \mathbf{B}\mathbf{C}_{\mathbf{Y}}\mathbf{B}^T$.
  - If $\mathbf{B} = \mathbf{0}$, then $\mathbf{Z} = \mathbf{A}\mathbf{X} + \mathbf{d}$, $\boldsymbol{\mu}_{\mathbf{Z}} = \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{d}$, and $\mathbf{C}_{\mathbf{Z}} = \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}^T$.

Mean of Scalar Valued Function of Random Vector Pair

- Given pair $[\mathbf{X};\mathbf{Y}] \sim f_{\mathbf{X},\mathbf{Y}}$ the expected value of a scalar valued function $g(\mathbf{X},\mathbf{Y})$ is

$$\begin{aligned} E\{g(\mathbf{X},\mathbf{Y})\} &= \iint_{[\mathbf{x};\mathbf{y}]\in S_{\mathbf{X},\mathbf{Y}}} g(\mathbf{x},\mathbf{y}) \cdot f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) d\mathbf{x} d\mathbf{y} \\ E\{g(\mathbf{X},\mathbf{Y})\} &= \sum_{[\mathbf{x};\mathbf{y}]\in S_{\mathbf{X},\mathbf{Y}}} g(\mathbf{x},\mathbf{y}) \cdot P_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) \end{aligned}$$

  for continuous and discrete r.v.'s respectively.
- Thus, we can find mean of $W = g(\mathbf{X},\mathbf{Y})$ without knowing PDF $f_W(w)$ or PMF $P_W(w)$.
- Note if we choose $g(\mathbf{X},\mathbf{Y}) = g_0(\mathbf{X})$, i.e. strictly a function of $\mathbf{X} \Longrightarrow$

$$\begin{aligned} E\{g_0(\mathbf{X})\} &= \iint_{[\mathbf{x};\mathbf{y}]\in S_{\mathbf{X},\mathbf{Y}}} g_0(\mathbf{x}) \cdot f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) d\mathbf{x} d\mathbf{y} \\ &= \int_{\mathbf{x}\in S_{\mathbf{X}}} g_0(\mathbf{x}) \cdot \left[\int_{\mathbf{y}\in S_{\mathbf{Y}}} f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) d\mathbf{y} \right] d\mathbf{x} = \int_{\mathbf{x}\in S_{\mathbf{X}}} g_0(\mathbf{x}) \cdot f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \end{aligned}$$

- Similarly, we can choose $g(\mathbf{X},\mathbf{Y}) = g_0(\mathbf{Y})$ to obtain $E\{g_0(\mathbf{Y})\}$.

Chain Rule of Expectation: Random Vector Pair

- Recall from Bayes theorem that $f_{\mathbf{X},\mathbf{Y}} = f_{\mathbf{X}} \cdot f_{\mathbf{Y}|\mathbf{X}} = f_{\mathbf{Y}} \cdot f_{\mathbf{X}|\mathbf{Y}} \Longrightarrow$

$$E_{\mathbf{X},\mathbf{Y}}\{g(\mathbf{X},\mathbf{Y})\} = \iint g(\mathbf{x},\mathbf{y}) \cdot f_{\mathbf{Y}}(\mathbf{y}) f_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) d\mathbf{x} d\mathbf{y} = \int f_{\mathbf{Y}}(\mathbf{y}) \left[\int g(\mathbf{x},\mathbf{y}) \cdot f_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) d\mathbf{x} \right] d\mathbf{y}$$
$$= \int f_{\mathbf{Y}}(\mathbf{y}) \cdot E\{g(\mathbf{X},\mathbf{Y})|\mathbf{Y}=\mathbf{y}\} d\mathbf{y} = E_{\mathbf{Y}}\left\{ E_{\mathbf{X}|\mathbf{Y}}\{g(\mathbf{X},\mathbf{Y})|\mathbf{Y}\}\right\}$$

  - Similarly, $E_{\mathbf{X},\mathbf{Y}}\{g(\mathbf{X},\mathbf{Y})\} = E_{\mathbf{X}}\left\{ E_{\mathbf{Y}|\mathbf{X}}\{g(\mathbf{X},\mathbf{Y})|\mathbf{X}\}\right\}$.
  - These are multivariate extensions of the chain rule of expectation, or sometimes called the law of total expectation.

On Multivariate Moment Generating Function

- The MGF for a random vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ is defined as[2]

$$\phi_{\mathbf{X}}(\mathbf{s}) = E\{e^{\mathbf{s}^T\mathbf{X}}\} = E\left\{\exp\left(\sum_{i=1}^{n} s_i X_i\right)\right\} = \begin{cases} \int_{\mathbf{x}\in S_{\mathbf{X}}} e^{\mathbf{s}^T\mathbf{x}} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, & \text{(continuous r.v.)} \\ = \sum_{\mathbf{x}\in S_{\mathbf{X}}} e^{\mathbf{s}^T\mathbf{x}} P_{\mathbf{X}}(\mathbf{x}), & \text{(discrete r.v.).} \end{cases}$$

- Note $\phi_{\mathbf{X}}(\mathbf{0}) = \int_{\mathbf{x}\in S_{\mathbf{X}}} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = 1$, $\phi_{\mathbf{X}}(\mathbf{0}) = \sum_{\mathbf{x}\in S_{\mathbf{X}}} P_{\mathbf{X}}(\mathbf{x}) = 1$ for continuous/discrete case.

- Mixed $m-$th moments can be obtained for arbitrary indices $i_m \in \{1, 2, \ldots, n\}$:

$$E\left\{X_{i_1} X_{i_2} \cdots X_{i_m}\right\} = \left.\frac{\partial^m \phi_{\mathbf{X}}(\mathbf{s})}{\partial s_{i_1} \partial s_{i_2} \cdots \partial s_{i_m}}\right|_{\mathbf{s}=\mathbf{0}}. \tag{11}$$

[2] The MGF is essentially the n-dimensional Laplace transform for the multivariate joint PDF $f_{\mathbf{X}}$ of n random variables X_i, $i = 1, 2, \ldots, n$.

The Multivariate Gaussian Distribution

- A very widely used model for random multivariate data is the multivariate Gaussian PDF (sometimes referred as the Normal distribution).
- Recall that the PDF and MGF for a Gaussian scalar random variable are given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \stackrel{\mathcal{L}}{\longleftrightarrow} e^{s\mu + \frac{1}{2}s^2\sigma^2} = \phi_X(s).$$

- The multivariate extension to an $n \times 1$ vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ is given by

$$f_{\mathbf{X}}(\mathbf{x}) = |2\pi \mathbf{C}_{\mathbf{X}}|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{x} - \boldsymbol{\mu}_{\mathbf{X}})\right] \tag{12}$$

  where $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$, $\boldsymbol{\mu}_{\mathbf{X}} = E[\mathbf{X}]$, and $\mathbf{C}_{\mathbf{X}} = E[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T]$. The support is $S_{\mathbf{X}} = \{\mathbf{x} | \mathbf{x} \in \mathbb{R}^n\}$, i.e. all of Euclidean n-space.
- The multivariate Gaussian PDF has an analogous form to the univariate Gaussian PDF and similar interpretations:
  - $n \times 1$ vector $\boldsymbol{\mu}_{\mathbf{X}}$ is the mean vector for random vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$.
  - $n \times n$ matrix $\mathbf{C}_{\mathbf{X}}$ is the covariance matrix of random vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$.
  - We denote this distribution as $\mathbf{X} \sim N(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$, and sometimes as $\mathbf{X} \sim N_{n\times 1}(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$ where dimensions of $\mathbf{X}$ are specified.

MGF for Multivariate Gaussian

- The MGF for $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ is $\phi_\mathbf{X}(\mathbf{s}) = E\{e^{\mathbf{s}^T\mathbf{X}}\} = E\{e^{s_1X_1+s_2X_2+\cdots+s_nX_n}\}$ and is given by

$$\phi_\mathbf{X}(\mathbf{s}) = \exp\left[\mathbf{s}^T\boldsymbol{\mu}_\mathbf{X} + \frac{1}{2}\mathbf{s}^T\mathbf{C}_\mathbf{X}\mathbf{s}\right] \qquad (13)$$

for all $\mathbf{s} \in \mathbb{C}^n$ including each real axis $-\infty < s_i < \infty$, $i = 1, 2, \ldots, n$.

  - The proof of this resulting MGF is exactly the same as the one given for the bivariate Gaussian case given in Lecture Notes 5.
  - All the 2×1 vector quantities appearing in the proof of the bivariate Gaussian MGF simply can be replaced with $n \times 1$ vector quantities:
    - $f_\mathbf{X} \longrightarrow f_\mathbf{W}$
    - $\mathbf{X} \longrightarrow \mathbf{W}$ and $\mathbf{x} \longrightarrow \mathbf{w}$
    - $\boldsymbol{\mu}_\mathbf{X} \longrightarrow \boldsymbol{\mu}_W$ and $\mathbf{C}_\mathbf{X} \longrightarrow \mathbf{C}_W$
    - $\mathbf{s} = [s_1, s_2, \ldots, s_n]^T \longrightarrow \mathbf{s} = [s_x, s_y]^T$.

    and the proof follows exactly the same steps.

- Note that the multivariate Gaussian distribution is completely specified by its mean $\boldsymbol{\mu}_\mathbf{X}$ and covariance $\mathbf{C}_\mathbf{X}$.

Gaussians Regenerate Under Linear Transformations

- Recall if univariate r.v. $X \sim N(\mu, \sigma^2)$, then $Y = aX + b \sim N(a\mu + b, a^2\sigma^2)$.
- Let $\mathbf{X} \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ and consider the affine (linear) transformation $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$.
  - Note that the MGF of the random vector $\mathbf{Y}$ is given by

$$
\begin{aligned}
\phi_\mathbf{Y}(\mathbf{s}_y) &= E\{e^{\mathbf{s}_y^T\mathbf{Y}}\} = E\{e^{\mathbf{s}_y^T(\mathbf{A}\mathbf{X}+\mathbf{b})}\} = E\{e^{\mathbf{s}_y^T\mathbf{A}\mathbf{X}} \cdot e^{\mathbf{s}_y^T\mathbf{b}}\} = e^{\mathbf{s}_y^T\mathbf{b}} \cdot E\{e^{\mathbf{s}_y^T\mathbf{A}\mathbf{X}}\} \\
&= e^{\mathbf{s}_y^T\mathbf{b}} \cdot \phi_\mathbf{X}(\mathbf{s}_x)\big|_{\mathbf{s}_x = \mathbf{A}^T\mathbf{s}_y} = e^{\mathbf{s}_y^T\mathbf{b}} \cdot \exp\left[\mathbf{s}_y^T\mathbf{A}\boldsymbol{\mu}_\mathbf{X} + \frac{1}{2}\mathbf{s}_y^T\mathbf{A}\mathbf{C}_\mathbf{X}\mathbf{A}^T\mathbf{s}_y\right] \\
&= \exp\left[\mathbf{s}_y^T(\mathbf{A}\boldsymbol{\mu}_\mathbf{X} + \mathbf{b}) + \frac{1}{2}\mathbf{s}_y^T\mathbf{A}\mathbf{C}_\mathbf{X}\mathbf{A}^T\mathbf{s}_y\right] \\
&\triangleq \exp\left[\mathbf{s}_y^T\boldsymbol{\mu}_\mathbf{Y} + \frac{1}{2}\mathbf{s}_y^T\mathbf{C}_\mathbf{Y}\mathbf{s}_y\right] = \phi_\mathbf{Y}(\mathbf{s}_y).
\end{aligned}
$$

  - The Laplace transform is a unique one-to-one transformation.
  - Hence, we have that

$$
\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \Longrightarrow \mathbf{Y} \sim N(\mathbf{A}\boldsymbol{\mu}_\mathbf{X} + \mathbf{b}, \mathbf{A}\mathbf{C}_\mathbf{X}\mathbf{A}^T). \tag{14}
$$

    i.e. affine/linear transformations of Gaussian random vectors result in Gaussian distributed random vectors.

Standardized Multivariate Gaussian

- If $\mathbf{Z} = [Z_1, Z_2, \ldots, Z_n]^T \sim N(\mathbf{0}, \mathbf{I}_n)$, i.e. where $\boldsymbol{\mu}_\mathbf{Z} = \mathbf{0}$ and $\mathbf{C}_\mathbf{Z} = \mathbf{I}_n$, then $\mathbf{Z}$ is said to be a $n \times 1$ standardized multivariate Gaussian random vector.
- Note from (12) that the PDF of $\mathbf{Z}$ is given by

$$
\begin{aligned}
f_\mathbf{Z}(\mathbf{z}) &= |2\pi\mathbf{I}_n|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{z}-\mathbf{0})^T\mathbf{I}_n^{-1}(\mathbf{z}-\mathbf{0})\right] = (2\pi)^{-n/2}|\mathbf{I}_n|^{-1/2}\exp\left[-\frac{1}{2}\|\mathbf{z}\|^2\right] \\
&= (2\pi)^{-n/2}\exp\left[-\frac{1}{2}\sum_{i=1}^{n} z_i^2\right] = \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z_i^2} = f_{Z_1}(z_1)\cdot f_{Z_2}(z_2)\cdots f_{Z_n}(z_n)
\end{aligned}
\tag{15}
$$

  - where $Z_i \sim f_{Z_i}(z_i) = f_Z(z_i) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z_i^2}$,
  - i.e. each $Z_i \sim N(0,1)$ is i.i.d. standardized univariate normal.
- $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n)$ is sometimes called a white Gaussian random vector.
  - When covariance is $\mathbf{C}_\mathbf{Z} \neq \mathbf{I}_n$, then $\mathbf{Z}$ is described as colored Gaussian.
- Note that if $\tilde{\mathbf{Z}} = \mathbf{Q}\mathbf{Z}$ (linear transformation of $\mathbf{Z}$) where $\mathbf{Q}$ is a $n \times n$ orthogonal matrix, i.e. $\mathbf{Q}^T\mathbf{Q} = \mathbf{Q}\mathbf{Q}^T = \mathbf{I}_n$, then $\tilde{\mathbf{Z}}$ is also Gaussian. Specifically, by (14) we know that
$$\tilde{\mathbf{Z}} \sim N(\mathbf{Q}\mathbf{0}, \mathbf{Q}\mathbf{I}_n\mathbf{Q}^T) = N(\mathbf{0}, \mathbf{Q}\mathbf{Q}^T) = N(\mathbf{0}, \mathbf{I}_n).$$
  - Note that $\mathbf{Z}$ and $\tilde{\mathbf{Z}}$ have the same PDF.
  - We say that $\mathbf{Z}$ and $\tilde{\mathbf{Z}}$ are identically distributed.
  - We sometimes denote this as $\mathbf{Z} \stackrel{d}{=} \tilde{\mathbf{Z}}$.

Standardized Multivariate Gaussian Cont.

- This demonstrates that white Gaussian random vectors have PDFs that are invariant to orthogonal linear transformation.
  - If $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n)$, then $\mathbf{QZ} \sim N(\mathbf{0}, \mathbf{I}_n)$ for any $n \times n$ orthogonal matrix $\mathbf{Q}$.
- Consider $n \times 1$ vector $\mathbf{X} \sim N[\mathbf{0}, \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)]$.
  - This is not standardized Gaussian nor white in general.
  - Each element X_i, however, is uncorrelated with X_j for $i \neq j$.
  - Note by (12) the PDF of this $\mathbf{X}$ is given by

$$
\begin{aligned}
f_{\mathbf{X}}(\mathbf{x}) &= |2\pi \text{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2)|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{0})^T \text{diag}(1/\sigma_1^2, 1/\sigma_2^2, \ldots, 1/\sigma_n^2)(\mathbf{x}-\mathbf{0})\right] \\
&= (2\pi)^{-n/2} \left(\prod_{i=1}^n \sigma_i^2\right)^{-1/2} \exp\left[-\frac{1}{2}\sum_{i=1}^n \frac{1}{\sigma_i^2} x_i^2\right] \qquad (16) \\
&= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{x_i^2}{2\sigma_i^2}} = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdots f_{X_n}(x_n)
\end{aligned}
$$

  - where $X_i \sim f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{x_i^2}{2\sigma_i^2}}$, $i = 1, 2, \ldots, n$ are clearly independent random variables.

A Transformation Resulting in Uncorrelated R.V.'s

- Let $\mathbf{X} \sim N(\mathbf{0}, \mathbf{C_X})$ where $\mathbf{C_X} = \mathbf{Q\Lambda Q}^T$ and consider $\mathbf{Y} = \mathbf{Q}^T\mathbf{X}$, i.e. a linear transformation.
- We know that $\mathbf{Y}$ is also a Gaussian random vector.
- Specifically, we note by (14) that

$$\mathbf{Y} \sim N(\mathbf{Q}^T\mathbf{0}, \mathbf{Q}^T\mathbf{C_X}\mathbf{Q}) = N(\mathbf{0}, \mathbf{Q}^T\mathbf{Q\Lambda Q}^T\mathbf{Q}) = N(\mathbf{0}, \mathbf{I\Lambda I}) = N(\mathbf{0}, \mathbf{\Lambda}). \quad (17)$$

- Hence, $\mathbf{Y} \sim N[\mathbf{0}, \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)]$.
- Similar to analysis of (16) it follows that the elements of $\mathbf{Y} = [Y_1, Y_2, \ldots, Y_n]^T$ are all uncorrelated and independent where $Y_i \sim N(0, \lambda_i)$, $i = 1, 2, \ldots, n$.

Uncorrelated Gaussian R.V.'s are Independent

- Let $n \times 1$ vector $\mathbf{X} \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ and consider the partition

$$\mathbf{X} = \begin{bmatrix} (\mathbf{X}_1)_{m\times 1} \\ (\mathbf{X}_2)_{(n-m)\times 1} \end{bmatrix}, \quad \boldsymbol{\mu}_\mathbf{X} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \quad \mathbf{C}_\mathbf{X} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}.$$

  - If $\mathbf{C}_{12} = \mathbf{C}_{21}^T = \mathbf{0}$, i.e. $\mathbf{X}_1$ and $\mathbf{X}_2$ are uncorrelated, then

$$\mathbf{C}_\mathbf{X}^{-1} = \begin{bmatrix} \mathbf{C}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{22}^{-1} \end{bmatrix}, \text{ and } |\mathbf{C}_\mathbf{X}| = |\mathbf{C}_{11}| \cdot |\mathbf{C}_{22}| \implies \text{(12) becomes}$$

$$\begin{aligned} f_\mathbf{X} &= (2\pi)^{-n/2}|\mathbf{C}_{11}| \cdot |\mathbf{C}_{22}| \exp\left[-\frac{1}{2}(\mathbf{x}_1 - \boldsymbol{\mu}_1)^T\mathbf{C}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1) - \frac{1}{2}(\mathbf{x}_2 - \boldsymbol{\mu}_2)^T\mathbf{C}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)\right] \\ &= (2\pi)^{-m/2}|\mathbf{C}_{11}| \exp\left[-\frac{1}{2}(\mathbf{x}_1 - \boldsymbol{\mu}_1)^T\mathbf{C}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1)\right] \times \\ &\quad (2\pi)^{-(n-m)/2}|\mathbf{C}_{22}| \exp\left[-\frac{1}{2}(\mathbf{x}_2 - \boldsymbol{\mu}_2)^T\mathbf{C}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)\right] \\ &= f_{\mathbf{X}_1} \cdot f_{\mathbf{X}_2}. \end{aligned}$$

- Thus, if Gaussian $\mathbf{X}_1$ and $\mathbf{X}_2$ are uncorrelated, then they are also independent.

Uncorrelated Gaussian R.V.'s are Independent Cont.

- Consider the bivariate Gaussian case, i.e. where $n = 2$ and $m = 1$. Recall from equation (17) in Lecture Notes 5 that $[\mathbf{C}_W]_{1,2} = C_{12} = 0$ means $\rho_{X,Y}\sigma_X\sigma_Y = 0$, or simply that $\rho_{X,Y} = 0$. The joint bivariate Gaussian on slide 29 of Lecture Notes 5 becomes

$$f_{\mathbf{W}}(\mathbf{w}) = f_{X,Y}(x,y) = \frac{\exp\left[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}{2\pi\sigma_X\sigma_Y}$$

$$= \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{\frac{-(x-\mu_X)^2}{2\sigma_X^2}} \cdot \frac{1}{\sqrt{2\pi\sigma_Y^2}} e^{\frac{-(y-\mu_Y)^2}{2\sigma_Y^2}} = f_X \cdot f_Y.$$

- In general, however, when $\mathbf{C}_{12} = \mathbf{C}_{21}^T \neq \mathbf{0}$, i.e. $\mathbf{X}_1$ and $\mathbf{X}_2$ correlated, it follows that
  - The marginal joint PDF for $\mathbf{X}_1$ is $N(\boldsymbol{\mu}_1, \mathbf{C}_{11})$.
  - The marginal joint PDF for $\mathbf{X}_2$ is $N(\boldsymbol{\mu}_2, \mathbf{C}_{22})$.
  - The conditional distribution of $\mathbf{X}_1$ given $\mathbf{X}_2$ is given by

$$\mathbf{X}_1 | \mathbf{X}_2 \sim N[\boldsymbol{\mu}_1 + \mathbf{C}_{12}\mathbf{C}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2), \mathbf{C}_{11.2}] \quad \text{where} \;\; \mathbf{C}_{11.2} \triangleq \mathbf{C}_{11} - \mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{C}_{21}. \tag{18}$$

  - You will prove this in your homework.

Transformation to a Standardized Gaussian

- Recall if univariate $X \sim N(\mu, \sigma^2)$, then $Z = (X - \mu)/\sigma \sim N(0, 1)$, i.e. standardized normal.
- Let $\mathbf{X} \sim N(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$ and consider $\mathbf{Z} = \mathbf{C}_{\mathbf{X}}^{-1/2}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})$.
- Note that $\mathbf{Z} = \mathbf{A}\mathbf{X} + \mathbf{b}$ where $\mathbf{A} = \mathbf{C}_{\mathbf{X}}^{-1/2}$ and $\mathbf{b} = -\mathbf{C}_{\mathbf{X}}^{-1/2}\boldsymbol{\mu}_{\mathbf{X}}$.
- Thus, we know from property of Gaussians under linear transformations that

$$\mathbf{Z} \sim N(\mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{b}, \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}^T) = N(\mathbf{C}_{\mathbf{X}}^{-1/2}\boldsymbol{\mu}_{\mathbf{X}} - \mathbf{C}_{\mathbf{X}}^{-1/2}\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}}^{-1/2}\mathbf{C}_{\mathbf{X}}\mathbf{C}_{\mathbf{X}}^{-1/2}) = N(\mathbf{0}, \mathbf{I}_n),$$

  i.e. $\mathbf{Z}$ is a multivariate normal with zero mean and identity covariance.
- Clearly, subtracting $\boldsymbol{\mu}_{\mathbf{X}}$ from $\mathbf{X}$ (i.e. centering it) yields a zero mean vector.
- Multiplication by $\mathbf{C}_{\mathbf{X}}^{-1/2}$ is often referred to as a whitening transformation.
  - If $\mathbf{X}_0 \sim N(\mathbf{0}, \mathbf{C}_{\mathbf{X}})$, then by (14) $\Longrightarrow$ $\mathbf{C}_{\mathbf{X}}^{-1/2}\mathbf{X}_0 \sim N(\mathbf{0}, \mathbf{C}_{\mathbf{X}}^{-1/2}\mathbf{C}_{\mathbf{X}}\mathbf{C}_{\mathbf{X}}^{-1/2}) = N(\mathbf{0}, \mathbf{I}_n)$.
  - i.e. multiplication by $\mathbf{C}_{\mathbf{X}}^{-1/2}$ transform vector $\mathbf{X}_0$ into a white Gaussian.

A Stochastic Representation of Multivariate Gaussian

- Recall if univariate r.v. $Z \sim N(0,1)$, then $X \stackrel{d}{=} \sigma \cdot Z + \mu \sim N(\mu, \sigma^2)$.
- Let $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n)$ and consider $\mathbf{H} = \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z} + \boldsymbol{\mu}_{\mathbf{X}}$, i.e. a linear transformation.
- We know that $\mathbf{H}$ is also a Gaussian random vector.
- Specifically, note by (14) that

$$\mathbf{H} \sim N(\mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{0} + \boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{I}_n\mathbf{C}_{\mathbf{X}}^{1/2}) = N(\mathbf{0} + \boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{C}_{\mathbf{X}}^{1/2}) = N(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}}).$$

- Clearly, $\mathbf{H} \stackrel{d}{=} \mathbf{X}$ where $\mathbf{X} \sim N(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$, i.e. $\mathbf{X} \stackrel{d}{=} \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z} + \boldsymbol{\mu}_{\mathbf{X}}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n)$.
- This is a very useful way of numerically simulating an arbitrary multivariate Gaussian random vector; namely, if a sample vector from distribution $N(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$ is desired, then:
  - Generate a sample of a standardized Gaussian random vector $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n)$.
  - Use it to form $\mathbf{H} = \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z} + \boldsymbol{\mu}_{\mathbf{X}}$.
  - This vector $\mathbf{H}$ is now a sample vector from PDF $N(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$.

Useful Moment Theorem for Zero Mean Gaussians

- Let $\mathbf{X} \sim N(\mathbf{0}, \mathbf{C_X})$ and note useful moment theorem for zero mean Gaussians:

$$E\{X_1 \cdot X_2 \cdots X_m\} = \begin{cases} 0, & m \text{ odd} \\ \displaystyle\sum_{\substack{\text{All} \\ \text{distinct} \\ \text{pairs}}} \prod E\{X_i X_j\}, & m \text{ even} \end{cases}$$

For example, if $m = 4 \Longrightarrow$

$$\begin{aligned} E\{X_1X_2X_3X_4\} &= E\{X_1X_2\}E\{X_3X_4\} + E\{X_1X_3\}E\{X_2X_4\} + E\{X_1X_4\}E\{X_2X_3\} \\ &= [\mathbf{C_X}]_{1,2}[\mathbf{C_X}]_{3,4} + [\mathbf{C_X}]_{1,3}[\mathbf{C_X}]_{2,4} + [\mathbf{C_X}]_{1,4}[\mathbf{C_X}]_{2,3}. \end{aligned} \tag{19}$$

  - For proof see [2] page 258.
  - This can likewise be established via MGF as described in (11).
  - Note that if we choose $X_1 = X_2 = X_3 = X_4 \Longrightarrow E\{X_1^4\} = 3[\mathbf{C_X}]_{1,1}^2 = 3\sigma_{X_1}^4$.
    - This is consistent with HW3 problem 12-A.

Useful Integral Theorem for Norm Squared of R.V.

- Next we present a multivariate integral theorem that is often useful when dealing with spherically symmetric distributions.
  - If a PDF is strictly a function of the norm squared of it's argument, then it is said to be spherically symmetric.
  - For example, in (15) note that for $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) = f_{\mathbf{Z}}(\mathbf{z}) = g(\|\mathbf{z}\|^2) = g(\mathbf{z}^T\mathbf{z})$.
  - Thus, the PDF of $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n)$ is spherically symmetric.
- A proof of the following theorem can be found in [5].
- **Theorem (Real Random Variables):** Integrating a well-behaved function $q(\mathbf{a}^T\mathbf{a})$ over all $\mathbf{a} \in \mathbb{R}^n$ has the equivalent integral representation

$$\int_{\mathbf{a}\in\mathbb{R}^n} q(\mathbf{a}^T\mathbf{a})d\mathbf{a} = \frac{\pi^{n/2}}{\Gamma(n/2)} \int_0^\infty r^{\frac{n}{2}-1} q(r)dr \tag{20}$$

  where $\Gamma(z)$ is the univariate Gamma function that has the property that $\Gamma(n+1) = n!$ for n that is a non-negative integer.
  - Generally speaking, the proof is established by a change of variables to polar coordinates, and then simply integrating over all the angle variables. See [5] for details.
  - We will prove this for the simpler $n = 2$ case later.

PDF of Norm Squared of Spherically Symmetric R.V.

- Consider any random vector $\mathbf{Z} = [Z_1, Z_2, \ldots, Z_n]^T \sim f_{\mathbf{Z}}(\mathbf{z}) = g(\|\mathbf{z}\|^2)$, i.e. any random vector with a spherically symmetric distribution.
  - Define r.v. $\rho = \|\mathbf{Z}\|^2 = \sum_{i=1}^n Z_i^2$ and note mean of function $h(\mathbf{Z}^T\mathbf{Z})$ is

$$\begin{aligned} E\{h(\mathbf{Z}^T\mathbf{Z})\} &= \int_{\mathbf{z}\in\mathbb{R}^n} h(\mathbf{z}^T\mathbf{z})g(\mathbf{z}^T\mathbf{z})d\mathbf{z} = \frac{\pi^{n/2}}{\Gamma(n/2)}\int_0^\infty r^{\frac{n}{2}-1}h(r)g(r)dr \\ &= E\{h(\rho)\} = \int_0^\infty h(r)f_\rho(r)dr \end{aligned} \tag{21}$$

  - where first equality is by definition of expectation;
  - second equality follows from (20);
  - the third equality must hold from basic probability, i.e. treating $h(\cdot)$ as function of $\rho \sim f_\rho$;
  - and the last equality is by definition of expectation (averaging over $\rho \sim f_\rho$).
- Comparing last two integrals in (21), we can conclude that the pdf of ρ must be

$$f_\rho(r) = \frac{\pi^{n/2}}{\Gamma(n/2)} r^{\frac{n}{2}-1} g(r), \quad r \geq 0.$$

Central Chi-Squared Distribution

- Consider $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n)$, i.e. a standardized multivariate Gaussian vector.
- The PDF of $\rho = \|\mathbf{Z}\|^2 = \sum_{i=1}^{n} Z_i^2$ is given by a central chi-squared of n degrees-of-freedom:

$$\rho \sim f_\rho(r) = \frac{1}{2^{n/2}\Gamma(n/2)} r^{\frac{n}{2}-1} e^{-r/2}, \quad r \geq 0. \tag{22}$$

Proof: Note that if f_ρ is pdf of $\rho = \|\mathbf{Z}\|^2$, then any function $h(\rho)$ has average value

$$E\{h(\rho)\} = \int_0^\infty h(a) f_\rho(a) da. \tag{23}$$

This average likewise follows when $h(\cdot)$ is treated as a function of $\mathbf{Z}$, i.e.

$$\begin{aligned} E\{h(\|\mathbf{Z}\|^2)\} &= \int_{\mathbb{R}^n} h(\|\mathbf{b}\|^2) f_{\mathbf{z}}(\mathbf{b}) d\mathbf{b} = \int_{\mathbb{R}^n} h(\|\mathbf{b}\|^2)(2\pi)^{-(n/2)} \exp\left[-\frac{1}{2}\|\mathbf{b}\|^2\right] d\mathbf{b} \\ &= \frac{\pi^{n/2}}{\Gamma(n/2)} \int_0^\infty a^{\frac{n}{2}-1} h(a)(2\pi)^{-(n/2)} \exp\left[-\frac{1}{2}a\right] da \end{aligned} \tag{24}$$

  - where the last equality follows from the previously discussed integral theorem.
  - Comparing (23) and (24) that must be equal, it follows that f_ρ is given by (22). ■

- Central chi-squared distribution of n dofs sometimes denoted by notation $\rho \overset{d}{=} \chi_n^2$.

Moments of a Central Chi-Squared Random Variable

- Since (22) is a pdf, it has unit area:

$$\int_0^\infty f_\rho(r)dr = 1 \Longrightarrow \int_0^\infty \frac{1}{2^{n/2}\Gamma(n/2)} r^{\frac{n}{2}-1} e^{-r/2} dr = 1 \Longrightarrow$$
$$\int_0^\infty r^{\frac{n}{2}-1} e^{-r/2} dr = 2^{n/2}\Gamma(n/2) \tag{25}$$

where last equality is a useful integral identity.

- (25) can be used to find the *m*-th moment of χ_n^2:

$$E\{(\chi_n^2)^m\} = \int_0^\infty \frac{1}{2^{n/2}\Gamma(n/2)} r^m \cdot r^{\frac{n}{2}-1} e^{-r/2} dr = \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^\infty r^{\left(m+\frac{n}{2}\right)-1} e^{-r/2} dr$$
$$\Longrightarrow E\{(\chi_n^2)^m\} = \frac{2^{\left(m+\frac{n}{2}\right)}\Gamma\left(m+\frac{n}{2}\right)}{2^{n/2}\Gamma(n/2)} = \frac{2^m\Gamma\left(m+\frac{n}{2}\right)}{\Gamma(n/2)}.$$

- Since $\Gamma(z+m) = (z+m-1)(z+m-2)\cdots(z+1)z\Gamma(z)$ for positive integers m [6], it follows that $\Gamma\left(m+\frac{n}{2}\right) = \left(\frac{n}{2}+m-1\right)\left(\frac{n}{2}+m-2\right)\cdots\left(\frac{n}{2}+1\right)\frac{n}{2}\Gamma(n/2) \Longrightarrow$

$$E\{(\chi_n^2)^m\} = 2^m\left(\frac{n}{2}+m-1\right)\left(\frac{n}{2}+m-2\right)\cdots\left(\frac{n}{2}+1\right)\frac{n}{2}. \tag{27}$$

Moments of a Central Chi-Squared Random Variable Cont.

- Thus, the mean and variance of χ_n^2 are given by

$$\left.\begin{aligned} E\{\chi_n^2\} &= 2 \cdot \frac{n}{2} = n \\ E\{(\chi_n^2)^2\} &= 2^2 \left(\frac{n}{2}+1\right)\frac{n}{2} = (n+2)n \end{aligned}\right\} \Longrightarrow \begin{aligned} \sigma_{\chi_n^2}^2 &= E\{(\chi_n^2)^2\} - E^2\{\chi_n^2\} \\ &= (n+2)n - n^2 \\ &= 2n. \end{aligned} \tag{28}$$

- Note from mean and variance that PDF is more concentrated near origin for small n, and moves away from origin and spreads out as n increases.
- The mean and variance of χ_n^2 can also be found by recalling $\chi_n^2 \stackrel{d}{=} \sum_{i=1}^n Z_i^2$ where $Z_i \sim \mathcal{N}(0,1)$. The Gaussian fourth order moment formula (19) would have to be used.
- Interestingly, we can likewise find the inverse m-th moment of χ_n^2:

$$E\left\{\frac{1}{(\chi_n^2)^m}\right\} = \int_0^\infty \frac{1}{2^{n/2}\Gamma(n/2)} r^{-m} \cdot r^{\frac{n}{2}-1} e^{-r/2} dr = \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^\infty r^{\left(\frac{n}{2}-m\right)-1} e^{-r/2} dr$$

$$\Longrightarrow E\left\{\frac{1}{(\chi_n^2)^m}\right\} = \frac{2^{\left(\frac{n}{2}-m\right)}\Gamma\left(\frac{n}{2}-m\right)}{2^{n/2}\Gamma(n/2)} = \frac{\Gamma\left(\frac{n}{2}-m\right)}{2^m\Gamma(n/2)}$$

where again (25) has been used. This is valid for $\frac{n}{2} - m \geq 1$.

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n)$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n)$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \le \zeta]$?
  - Consider that $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \le \zeta]$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta]$?
  - Consider that $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2} \mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$\Pr[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta]$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
    - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
    - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
    - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
    - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta]$?
  - Consider that $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta]$$

$$= \Pr[(\mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z}) \leq \zeta]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta]$$
$$= \Pr[(\mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z}) \leq \zeta]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta]$?
  - Consider that $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2} \mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2} \mathbf{Z} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2} \mathbf{Z} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta]$$

$$= \Pr[(\mathbf{C}_{\mathbf{X}}^{1/2} \mathbf{Z})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{C}_{\mathbf{X}}^{1/2} \mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_{\mathbf{X}}^{1/2} \mathbf{C}_{\mathbf{X}}^{-1} \mathbf{C}_{\mathbf{X}}^{1/2} \mathbf{Z} \leq \zeta]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \le \zeta]$?
  - Consider that $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \le \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \le \zeta]$$

$$= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \le \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2} \mathbf{C}_\mathbf{X}^{-1} \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z} \le \zeta]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$$

$$= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2} \mathbf{C}_\mathbf{X}^{-1} \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$$

$$= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2} \mathbf{C}_\mathbf{X}^{-1} \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \le \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \le \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \le \zeta]$$

$$= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \le \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2} \mathbf{C}_\mathbf{X}^{-1} \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z} \le \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \le \zeta]$$

$$= \Pr[\mathbf{Z}^T \mathbf{Z} \le \zeta]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$$

$$= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2} \mathbf{C}_\mathbf{X}^{-1} \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta]$$

$$= \Pr[\mathbf{Z}^T \mathbf{Z} \leq \zeta]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \dots, X_n]^T \sim N(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta]$?
  - Consider that $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta]$$

$$= \Pr[(\mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_{\mathbf{X}}^{1/2} \mathbf{C}_{\mathbf{X}}^{-1} \mathbf{C}_{\mathbf{X}}^{1/2} \mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta]$$

$$= \Pr[\mathbf{Z}^T \mathbf{Z} \leq \zeta] = \Pr[\|\mathbf{Z}\|^2 \leq \zeta]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \le \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \le \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \le \zeta]$$

$$= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \le \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2} \mathbf{C}_\mathbf{X}^{-1} \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z} \le \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \le \zeta]$$

$$= \Pr[\mathbf{Z}^T \mathbf{Z} \le \zeta] = \Pr[\|\mathbf{Z}\|^2 \le \zeta]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$$

$$= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2} \mathbf{C}_\mathbf{X}^{-1} \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta]$$

$$= \Pr[\mathbf{Z}^T \mathbf{Z} \leq \zeta] = \Pr[\|\mathbf{Z}\|^2 \leq \zeta] = \Pr\left[\sum_{i=1}^{n} Z_i^2 \leq \zeta\right]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1}(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1}(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1}(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$$

$$= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1}(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2}\mathbf{C}_\mathbf{X}^{-1}\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta]$$

$$= \Pr[\mathbf{Z}^T\mathbf{Z} \leq \zeta] = \Pr[\|\mathbf{Z}\|^2 \leq \zeta] = \Pr\left[\sum_{i=1}^{n} Z_i^2 \leq \zeta\right]$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$$

$$= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2} \mathbf{C}_\mathbf{X}^{-1} \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta]$$

$$= \Pr[\mathbf{Z}^T \mathbf{Z} \leq \zeta] = \Pr[\|\mathbf{Z}\|^2 \leq \zeta] = \Pr\left[\sum_{i=1}^{n} Z_i^2 \leq \zeta\right] = \Pr(\chi_n^2 \leq \zeta)$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1}(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\begin{aligned}\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1}(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] &= \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1}(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] \\ &= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1}(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2}\mathbf{C}_\mathbf{X}^{-1}\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta] \\ &= \Pr[\mathbf{Z}^T\mathbf{Z} \leq \zeta] = \Pr[\|\mathbf{Z}\|^2 \leq \zeta] = \Pr\left[\sum_{i=1}^{n} Z_i^2 \leq \zeta\right] = \Pr(\chi_n^2 \leq \zeta)\end{aligned}$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$$

$$= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2} \mathbf{C}_\mathbf{X}^{-1} \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta]$$

$$= \Pr[\mathbf{Z}^T \mathbf{Z} \leq \zeta] = \Pr[\|\mathbf{Z}\|^2 \leq \zeta] = \Pr\left[\sum_{i=1}^{n} Z_i^2 \leq \zeta\right] = \Pr(\chi_n^2 \leq \zeta)$$

$$F_{\chi_n^2}(\zeta) = \Pr(\chi_n^2 \leq \zeta)$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_{\mathbf{X}}, \mathbf{C}_{\mathbf{X}})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$
\begin{aligned}
\Pr[(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta] &= \Pr[(\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1} (\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_{\mathbf{X}}) \leq \zeta] \\
&= \Pr[(\mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z})^T \mathbf{C}_{\mathbf{X}}^{-1} (\mathbf{C}_{\mathbf{X}}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_{\mathbf{X}}^{1/2} \mathbf{C}_{\mathbf{X}}^{-1} \mathbf{C}_{\mathbf{X}}^{1/2} \mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta] \\
&= \Pr[\mathbf{Z}^T \mathbf{Z} \leq \zeta] = \Pr[\|\mathbf{Z}\|^2 \leq \zeta] = \Pr\left[\sum_{i=1}^{n} Z_i^2 \leq \zeta\right] = \Pr(\chi_n^2 \leq \zeta)
\end{aligned}
$$

$$
F_{\chi_n^2}(\zeta) = \Pr(\chi_n^2 \leq \zeta)
$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$$

$$= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2} \mathbf{C}_\mathbf{X}^{-1} \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta]$$

$$= \Pr[\mathbf{Z}^T \mathbf{Z} \leq \zeta] = \Pr[\|\mathbf{Z}\|^2 \leq \zeta] = \Pr\left[\sum_{i=1}^{n} Z_i^2 \leq \zeta\right] = \Pr(\chi_n^2 \leq \zeta)$$

$$F_{\chi_n^2}(\zeta) = \Pr(\chi_n^2 \leq \zeta) = \int_0^{\zeta} \frac{1}{2^{n/2}\Gamma(n/2)} r^{\frac{n}{2}-1} e^{-r/2} dr$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$
\begin{aligned}
\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] &= \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] \\
&= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2} \mathbf{C}_\mathbf{X}^{-1} \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta] \\
&= \Pr[\mathbf{Z}^T\mathbf{Z} \leq \zeta] = \Pr[\|\mathbf{Z}\|^2 \leq \zeta] = \Pr\left[\sum_{i=1}^n Z_i^2 \leq \zeta\right] = \Pr(\chi_n^2 \leq \zeta)
\end{aligned}
$$

$$
F_{\chi_n^2}(\zeta) = \Pr(\chi_n^2 \leq \zeta) = \int_0^\zeta \frac{1}{2^{n/2}\Gamma(n/2)} r^{\frac{n}{2}-1} e^{-r/2} dr
$$

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\Pr[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta] = \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1} (\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \leq \zeta]$$

$$= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1} (\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2} \mathbf{C}_\mathbf{X}^{-1} \mathbf{C}_\mathbf{X}^{1/2} \mathbf{Z} \leq \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \leq \zeta]$$

$$= \Pr[\mathbf{Z}^T\mathbf{Z} \leq \zeta] = \Pr[\|\mathbf{Z}\|^2 \leq \zeta] = \Pr\left[\sum_{i=1}^{n} Z_i^2 \leq \zeta\right] = \Pr(\chi_n^2 \leq \zeta) \tag{30}$$

$$F_{\chi_n^2}(\zeta) = \Pr(\chi_n^2 \leq \zeta) = \int_0^{\zeta} \frac{1}{2^{n/2}\Gamma(n/2)} r^{\frac{n}{2}-1} e^{-r/2} dr = \frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)},$$

where $\gamma(m, w) = \int_0^w t^{m-1} e^{-t} dt$ is lower incomplete Gamma function [6]; change of variables $u = r/2$, $du = dr/2$ was used.

Confidence Ellipse for Multivariate Gaussian

- Recall from (12) that the multivariate Gaussian has elliptical symmetry.
  - Indeed, recall in our discussions of the bivariate Gaussian we plotted lines of constant density and observed these to be given by ellipses in the (X, Y) plane.
  - If we repeat this exercise for the multivariate Gaussian in (12) then we'd find that regions of constant density are given by ellipsoids in n-dimensions.
- It is interesting to ask what is the probability that a random observation $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T \sim N(\boldsymbol{\mu}_\mathbf{X}, \mathbf{C}_\mathbf{X})$ will fall within a specific ellipsoidal region of constant density?
  - Specifically what is $\Pr[(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1}(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X}) \le \zeta]$?
  - Consider that $\mathbf{X} \overset{d}{=} \boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}$ where $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{I}_n) \Longrightarrow$

$$\begin{aligned}
\Pr[(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1}(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X}) \le \zeta] &= \Pr[(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1}(\boldsymbol{\mu}_\mathbf{X} + \mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} - \boldsymbol{\mu}_\mathbf{X}) \le \zeta] \\
&= \Pr[(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z})^T \mathbf{C}_\mathbf{X}^{-1}(\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z}) \le \zeta] = \Pr[\mathbf{Z}^T \mathbf{C}_\mathbf{X}^{1/2}\mathbf{C}_\mathbf{X}^{-1}\mathbf{C}_\mathbf{X}^{1/2}\mathbf{Z} \le \zeta] = \Pr[\mathbf{Z}^T \mathbf{I}_n \mathbf{Z} \le \zeta] \\
&= \Pr[\mathbf{Z}^T\mathbf{Z} \le \zeta] = \Pr[\|\mathbf{Z}\|^2 \le \zeta] = \Pr\left[\sum_{i=1}^n Z_i^2 \le \zeta\right] = \Pr(\chi_n^2 \le \zeta) \qquad (30)
\end{aligned}$$

$$F_{\chi_n^2}(\zeta) = \Pr(\chi_n^2 \le \zeta) = \int_0^\zeta \frac{1}{2^{n/2}\Gamma(n/2)} r^{\frac{n}{2}-1} e^{-r/2} dr = \frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)},$$

where $\gamma(m, w) = \int_0^w t^{m-1} e^{-t} dt$ is lower incomplete Gamma function [6]; change of variables $u = r/2$, $du = dr/2$ was used.

Confidence Ellipse for Multivariate Gaussian Cont.

- Note that when $n = 2$ we have that

$$\frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)}$$

Confidence Ellipse for Multivariate Gaussian Cont.

- Note that when $n = 2$ we have that

$$\frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)} = \frac{\gamma(1, \zeta/2)}{\Gamma(1)}$$

Confidence Ellipse for Multivariate Gaussian Cont.

- Note that when $n = 2$ we have that

$$\frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)} = \frac{\gamma(1, \zeta/2)}{\Gamma(1)}$$

Confidence Ellipse for Multivariate Gaussian Cont.

- Note that when $n = 2$ we have that

$$\frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)} = \frac{\gamma(1, \zeta/2)}{\Gamma(1)} = \frac{\gamma(1, \zeta/2)}{0!}$$

Confidence Ellipse for Multivariate Gaussian Cont.

- Note that when $n = 2$ we have that

$$\frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)} = \frac{\gamma(1, \zeta/2)}{\Gamma(1)} = \frac{\gamma(1, \zeta/2)}{0!}$$

Confidence Ellipse for Multivariate Gaussian Cont.

- Note that when $n = 2$ we have that

$$\frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)} = \frac{\gamma(1, \zeta/2)}{\Gamma(1)} = \frac{\gamma(1, \zeta/2)}{0!} = \int_0^{\zeta/2} e^{-t} dt = 1 - e^{-\zeta/2}.$$

Confidence Ellipse for Multivariate Gaussian Cont.

- Note that when $n = 2$ we have that

$$\frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)} = \frac{\gamma(1, \zeta/2)}{\Gamma(1)} = \frac{\gamma(1, \zeta/2)}{0!} = \int_0^{\zeta/2} e^{-t} dt = 1 - e^{-\zeta/2}.$$

Confidence Ellipse for Multivariate Gaussian Cont.

- Note that when $n = 2$ we have that

$$\frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)} = \frac{\gamma(1, \zeta/2)}{\Gamma(1)} = \frac{\gamma(1, \zeta/2)}{0!} = \int_0^{\zeta/2} e^{-t} dt = 1 - e^{-\zeta/2}.$$

- Thus, recall first bivariate example we considered when plotting regions of constant density:

$$\boldsymbol{\mu}_W = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \; \mathbf{C}_W = \begin{bmatrix} 5 & -10.4013 \\ -10.4013 & 30 \end{bmatrix}, \text{ where } \rho_{X,Y} = -0.85.$$

Confidence Ellipse for Multivariate Gaussian Cont.

- Note that when $n = 2$ we have that

$$\frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)} = \frac{\gamma(1, \zeta/2)}{\Gamma(1)} = \frac{\gamma(1, \zeta/2)}{0!} = \int_0^{\zeta/2} e^{-t} dt = 1 - e^{-\zeta/2}.$$

- Thus, recall first bivariate example we considered when plotting regions of constant density:

$$\boldsymbol{\mu}_W = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \ \mathbf{C}_W = \begin{bmatrix} 5 & -10.4013 \\ -10.4013 & 30 \end{bmatrix}, \text{ where } \rho_{X,Y} = -0.85.$$

Confidence Ellipse for Multivariate Gaussian Cont.

- Note that when $n = 2$ we have that

$$\frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)} = \frac{\gamma(1, \zeta/2)}{\Gamma(1)} = \frac{\gamma(1, \zeta/2)}{0!} = \int_0^{\zeta/2} e^{-t} dt = 1 - e^{-\zeta/2}.$$

- Thus, recall first bivariate example we considered when plotting regions of constant density:

$$\boldsymbol{\mu}_W = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \ \mathbf{C}_W = \begin{bmatrix} 5 & -10.4013 \\ -10.4013 & 30 \end{bmatrix}, \text{ where } \rho_{X,Y} = -0.85.$$

- We can compute these probability for $\zeta = -2 \ln \left[|2\pi \mathbf{C}_W|^{1/2} \cdot A \right] \Longrightarrow$:

$$\Pr\left[(\mathbf{w} - \boldsymbol{\mu}_W)^T \mathbf{C}_W^{-1} (\mathbf{w} - \boldsymbol{\mu}_W) \leq \zeta \right] = 1 - \exp\left[\ln\left(|2\pi \mathbf{C}_W|^{1/2} \cdot A \right) \right]$$
$$= 0.9, 0.5, \text{ and } 0.1 \text{ for } A = 0.0025, 0.0123, \text{ and } 0.0222 \text{ respectively.}$$

Confidence Ellipse for Multivariate Gaussian Cont.

- Note that when $n = 2$ we have that

$$\frac{\gamma(n/2, \zeta/2)}{\Gamma(n/2)} = \frac{\gamma(1, \zeta/2)}{\Gamma(1)} = \frac{\gamma(1, \zeta/2)}{0!} = \int_0^{\zeta/2} e^{-t} dt = 1 - e^{-\zeta/2}.$$

- Thus, recall first bivariate example we considered when plotting regions of constant density:

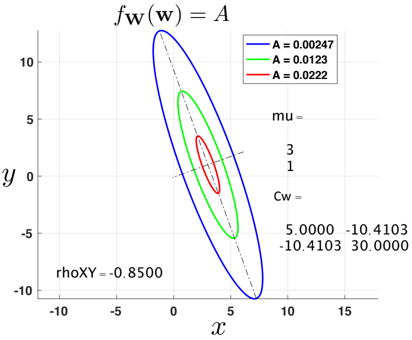
$$\boldsymbol{\mu}_W = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \ \mathbf{C}_W = \begin{bmatrix} 5 & -10.4013 \\ -10.4013 & 30 \end{bmatrix}, \text{ where } \rho_{X,Y} = -0.85.$$

- We can compute these probability for $\zeta = -2 \ln \left[|2\pi \mathbf{C}_W|^{1/2} \cdot A \right] \Longrightarrow$:

$$\Pr\left[(\mathbf{w} - \boldsymbol{\mu}_W)^T \mathbf{C}_W^{-1} (\mathbf{w} - \boldsymbol{\mu}_W) \le \zeta \right] = 1 - \exp\left[\ln\left(|2\pi \mathbf{C}_W|^{1/2} \cdot A \right) \right]$$
$$= 0.9,\ 0.5, \text{ and } 0.1 \text{ for } A = 0.0025,\ 0.0123, \text{ and } 0.0222 \text{ respectively.}$$

Confidence Ellipse for Multivariate Gaussian Cont.

- See figure below copied from Lecture Notes 5



References

[1] R. D. Yates and D. J. Goodman, *Probability and Stochastic Processes*, Third Edition, John Wiley & Sons, Inc., Hoboken, NJ, 2014.

[2] A. Papoulis, and S. U. Pillai, *Probability, random variables, and stochastic processes*. Boston, Mass: McGraw-Hill, 2014.

[3] S. M. Kay, *Intuitive probability and random processes using MATLAB*. New York: Springer, 2006.

[4] B. Noble and J. W. Daniel, *Applied Linear Algebra, 3rd Edition*, Englewood Cliffs, NJ: Prentice-Hall, 1988.

[5] Fang Kai-Tai and Zhang Yao-Ting, *Generalized Multivariate Analysis*, Springer-Verlag, New York, 1990.

[6] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, San Diego, CA: Academic Press, Inc., 1980.