This module will
1) Introduce Gaussian random vectors.
2) Show how characteristic functions can be applied to random vectors.

Reading material:
sections 5.6 and 5.7 (either 3rd or 4th edition of Stark & Woods).

Gaussian random Vectors
Recall the scalar Gaussian random variable,
\[ f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right\} \]

Consider a random vector with \( N \) independent components,
\[ f_X(x_1, \ldots, x_N) = \frac{1}{(2\pi)^{N/2} \sigma_1 \cdots \sigma_N} \exp\left\{ -\frac{1}{2} \sum_{i=1}^{N} \left( \frac{x_i - \mu_i}{\sigma_i} \right)^2 \right\} \]

Here, \( \mu_i \) and \( \sigma_i \) denote the expected value and standard deviation of component \( X_i \), respectively.
Example (5.6 - 2 from Stark & Woods)
consider a vector $X = (X_1, X_2)$ where $X_1, X_2$ are each individually $N(0,1)$ but they're correlated.

We have $f_X(x_1, x_2) = \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left\{ -\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)} \right\}$

Active learning component (part 1)
(1) Take $\rho = 0$, what's $f_X(x_1, x_2)$?

(2) Take $\rho = 0.5$, what's $f_X(x_1, x_2)$?

(3) What's the covariance matrix, $K(X)$?
Interestingly, the pdf for the fully general random Gaussian vector case can be written:

\[ f_X(x) = \frac{1}{(2\pi)^{n/2} \left| \text{det}(K(x)) \right|^{1/2}} \exp \left\{ -\frac{1}{2} (x - \mu)^{\top} K^{-1} (x - \mu) \right\} \]

Note \((x - \mu)^{\top} K^{-1} (x - \mu) = \square\)

\[
\begin{array}{c}
\text{row} \\
\text{matrix}
\end{array}
\]

Active learning exercise (part 2)

Please revisit the previous example where \(\mu = [0]^{\top}\), and verify equation \(\heartsuit\).
But what is this all good for? What are the useful properties of Gaussian random vectors?

**Theorem 5.6-2**

Let \( x \in \mathbb{R}^N \) be a Gaussian RVec with positive definite matrix \( K(x) \) and mean vector \( \mu(x) \). Let \( A \in \mathbb{R}^{M \times N} \) be an \( M \times N \) matrix of rank \( M \).

Then \( Y = AX \) is an \( M \)-dimensional Gaussian RVec with positive definite covariance matrix \( K(Y) = A \cdot K(x) \cdot A^T \) and mean vector \( \beta = E[Y] = A \cdot \mu \).

**Notes**

1) "Full rank" means that (assuming \( M \leq N \)) the space of outputs of this transformation of the form \( \beta = AV, V \in \mathbb{R}^N \), is \( M \)-dimensional.

2) For \( M = N \), full rank means that \( A \) is invertible.

3) Let \( M = \frac{1}{2} N \), then

\[
K(Y) = \begin{pmatrix} \ast & \ast & \ast & \ast \\
A & \ast & \ast & \ast \\
N & A^T & \ast & \ast \\
M & M & M & M \\
\end{pmatrix}
\]

\( K(Y) \) is also a square matrix.
Example 5.6-1
Consider \( X = (x_1, x_2) \) with \( M_X = (8) \) and 
\[ k(x) = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}. \]
we'll find \( Y = CX \) that contains uncorrelated components of unit variance.

We want \( k(Y) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \).

what we're really doing is whitening

\[ X \rightarrow \text{diagonalize} \rightarrow \text{whiten} \rightarrow Y \]

Let's compute e-values and e-vectors.
\[ \det (k(X) - \lambda I) = \det \begin{bmatrix} 3-\lambda & -1 \\ -1 & 3-\lambda \end{bmatrix} \]
\[ = (3-\lambda)^2 - (-1)^2 \]
\[ = (3-\lambda + 1)(3-\lambda - 1) = 0 \]
\[ \lambda_1 = 4 \]
\[ \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad v_1^2 + v_2^2 = 1 \]
\[ \Rightarrow v_1 = \frac{1}{\sqrt{2}}, \quad v_2 = -\frac{1}{\sqrt{2}} \]
\[ \lambda_2 = 2 \]
\[ \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad v_3^2 + v_4^2 = 1 \]
\[ \Rightarrow v_3 = v_4 = \frac{1}{\sqrt{2}} \]

Therefore, \( U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} v_1 & v_3 \\ v_2 & v_4 \end{bmatrix} \)
and \( \Lambda = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \).
(example continued...)  
First we apply \( z = U^T x \)  
\[
\begin{bmatrix}
\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\
\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} 
\end{bmatrix} x
\]

Then we apply \( y = A^{-\frac{1}{2}} z = \left( \frac{1}{\sqrt{2}} \ 0 \right) \left( \frac{1}{\sqrt{2}} \ -\frac{1}{\sqrt{2}} \right) x \)  
\[
= \left( \frac{1}{\sqrt{2}} \ 0 \right) \frac{1}{\sqrt{2}} \left( \frac{1}{\sqrt{2}} \ -\frac{1}{\sqrt{2}} \right) (x_1)
\]

Now define \( C = \frac{1}{\sqrt{2}} \left( \frac{1}{\sqrt{2}} \ -\frac{1}{\sqrt{2}} \right) \)  
we saw from the theorem that \( KL(y) = A \cdot K(x) \cdot A^T \)  
\[
= \frac{1}{\sqrt{2}} \left( \frac{1}{\sqrt{2}} \ -\frac{1}{\sqrt{2}} \right) \left( \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \right) \frac{1}{\sqrt{2}} \left( \frac{1}{\sqrt{2}} \ \frac{1}{\sqrt{2}} \right) 
\]

verified in matlab  
\[
\begin{bmatrix}
\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\
\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} 
\end{bmatrix} \left( \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \right) \frac{1}{\sqrt{2}} \left( \frac{1}{\sqrt{2}} \ \frac{1}{\sqrt{2}} \right) 
\]

\[
= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]

This is what we wanted!
Theorem. Let $X$ be a Gaussian zero mean (for convenience) RVec with covariance matrix $K(X)$. Then there exists $C \in \mathbb{R}^{N \times N}$ such that $Y = C X$, and components of $Y$ have unit variance and are independent.

Proof. Take $C = \Lambda^{-\frac{1}{2}} U^T$. Then

$$K(Y) = C \cdot K(X) \cdot C^T = \Lambda^{-\frac{1}{2}} U^T K(X) (\Lambda^{-\frac{1}{2}} U^T)^T = \Lambda^{-\frac{1}{2}} U^T K(X) \Lambda^{-\frac{1}{2}}$$

$$(U^T)^T = U$$

$$(\Lambda^{-\frac{1}{2}})^T = \Lambda^{-\frac{1}{2}}$$

$$U^T K U = \Lambda \rightarrow \Lambda^{-\frac{1}{2}} \Lambda \Lambda^{-\frac{1}{2}} = I$$

Note. For Gaussian RVecs, correlation and dependence are the same. Therefore, a Gaussian RVec with diagonal covariance matrix has independent components.

More about this later.
Characteristic functions of RVecs

\[ \phi_X(w) = \mathbb{E}\left[ e^{jw^Tx} \right] \]

Notes:
1) \( \begin{bmatrix} w^* \\ \text{row} \end{bmatrix} \begin{bmatrix} x^T \end{bmatrix} = \begin{bmatrix} x \end{bmatrix} \]

2) For continuous valued RVecs,
\[ \phi_X(w) = \int_{-\infty}^{\infty} f_X(x) e^{jw^Tx} dx \]

_pdf of vector

This is similar to a Fourier transform, and we also have
\[ f_X(x) = \frac{1}{(2\pi)^N} \int_{-\infty}^{\infty} \phi_X(w) e^{-jw^Tx} dw \]

Integration over all \( N \)-dimensional omegas
Similar to before, characteristic functions can be used to compute moments,
\[ E[X_1^{k_1} \cdots X_N^{k_N}] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1^{k_1} \cdots x_N^{k_N} \phi_X(x) \, dx_1 \cdots dx_N \]

**Properties**

1) \( |\phi_X(w)| \leq |\phi_X(w=0)| = 1 \)

   In fact, \( \phi_X(0) = E[e^{jwx}] \)
   
   \[ = E[e^{j0}] = E[1] = 1 \]

   no absolute value

2) \( \phi_X^*(w) = \phi_X(-w) \) (conjugate)

3) Characteristic functions of subsets of components of \( X \) are readily computed.

**Example**

\( N=3 \), \( X = (X_1, X_2, X_3)^T \)

\( \phi_{X_1,X_2}(w_1,w_2) = \phi_{X_1,X_2,X_3}(w_1, w_2, 0) \).
Example 5.7-2

Suppose that $X = (X_1, \ldots, X_N)^T$ contains Poisson components, each with parameter $\lambda$. Let $Z = X_1 + X_2 + \ldots + X_N$.

Compute the PMF of $Z$ for independent $X_i$.

We know $P_{X_i}(k) = \frac{\lambda^k e^{-\lambda}}{k!}$.

The CF is $\Phi_{X_i}(w) = E[e^{jwX_i}] = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} e^{jw_k}$

$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(ae^{jw})^k}{k!}$

$= e^{-\lambda} e^{ae^{jw}}$

$= e^{\lambda (e^{jw} - 1)}$.

Because $X_i$ are independent,

$\Phi_Z(w) = \prod_{i=1}^{N} e^{\lambda (e^{jw} - 1)}$

$= e^{N\lambda (e^{jw} - 1)}$.

This is the CF of Poisson $(N, \lambda)$.

Therefore, $Z \sim \text{Poisson}(N, \lambda)$.

Note Most distributions of various parametric form don't add up so nicely.
characteristic functions of Gaussian RVecs

It can be shown (see Section 5.7) that

\[ \Phi_X(w) = \exp\{jw^T\mu - \frac{1}{2} w^T K(x) w\} . \]

This immediately yields:

Properties

1) \( K(x) \) diagonal \( \Rightarrow \Phi_X(w) = \exp\left\{ \sum_{i=1}^{N} jw_i \mu_i - \frac{1}{2} w_i \sigma_i^2 \right\} \]

\[ = \prod_{i=1}^{N} \exp\{jw_i \mu_i - \frac{1}{2} w_i \sigma_i^2 \} \]

and the \( X_i \)'s are independent, as we mentioned earlier.

2) Apply any linear transformation to \( X \), i.e., \( Y = AX \). Then \( Y \) has the same Gaussian form.

This property also appeared earlier (Page 4); the transform of a Gaussian RVec is another Gaussian RVec.