Last Time:
\[ K_{\hat{x}} = E \left[ (\hat{x} - \hat{M}_x) (x - M_x)^T \right] \]

Properties:
1) \( K_{\hat{x}} \) is symmetric
2) \( K_{\hat{x}} \) is positive semi-definite

(Form: \( \hat{u}^T K_{\hat{x}} \hat{u} \geq 0 \))

Random vectors \( \hat{x}, \hat{y} \)
\[ K_{\hat{x}\hat{y}} = E \left[ (\hat{x} - \hat{M}_x) (\hat{y} - \hat{M}_y)^T \right] \]

This is called a cross-covariance matrix

---

Moment Generating Function
\[ g_{\hat{x}}(s) = E[e^{s^T \hat{x}}] \]
\[ E \left[ e^{s^T \hat{x}} \right] = \int e^{s^T x} f_{\hat{x}}(x) \, dx \]

---

Functions of a Random Vector

Suppose \( \hat{x} \) has a density \( f_{\hat{x}} \)
and \( h: \mathbb{R}^n \to \mathbb{R}^n \) and is one-to-one
let \( \hat{z} := h(\hat{x}) \)

What is the density of \( \hat{z} \)
\[ p \left( \frac{\tilde{z}}{z} \in B(z) \right) = \int_{B(z)} f_{\tilde{z}}(\tilde{z}) \, d\tilde{z} = f_{\tilde{z}}(\tilde{z}) \cdot \text{area} \left( B(z) \right) \]

\[ = p \left( \tilde{z} \in h^{-1}(B(z)) \right) \]

\[ = p \left( \tilde{x} \in h^{-1}\left( B(z) \right) \right) \]

\[ \equiv f_{\tilde{x}}(h^{-1}(z)) \cdot \frac{\text{area} \left( h^{-1}(B(z)) \right)}{\text{area} \left( B(z) \right)} \]

**Note**

\[ h^{-1}(\tilde{z} + \epsilon) = \left[ \frac{\partial h^{-1}(\tilde{z})}{\partial \tilde{z}} \right] \epsilon + h^{-1}(\tilde{z}) \]

\[ h^{-1}(\tilde{z} + \epsilon) = \left[ \frac{\partial h^{-1}(\tilde{z})}{\partial \tilde{z}} \cdots \frac{\partial h^{-1}(\tilde{z})}{\partial \tilde{z}_n} \right] \epsilon + h^{-1}(\tilde{z}) \]
\( J := \text{Jacobian matrix of } h^{-1}(c) \)

\[ f_\hat{c}(z) = f_x(h^{-1}(z)) \left| \det(J(\hat{c})) \right| \]

\[ \uparrow \quad J \text{ is Jacobian of } h^{-1}(c) \]

Suppose \( J^* \) is Jacobian of \( h(c) \)

We can show this:

\[ f_\hat{c}(z) = f_x(h^{-1}(z)) \left| \det(J^*(h^{-1}(z))) \right|^{-1} \]

---

**Gaussian Random Vectors**

\( W = N(0, 1) \) means that

\[ f_W(w) = \frac{1}{\sqrt{2\pi}} e^{-w^2/2} \]

Why give attention to Gaussian:

1) They are common
2) They are easy to work with
   - Easy conditional density and expectation calculations
   - Preserved by linear systems
3) Elegant solutions to Kalman and Wiener Filter problems.
why common?

**Central Limit Theorem**

Roughly, suppose $X_i = 1, \ldots, n$ are independent and identically distributed (i.i.d.)

Then

$$
\frac{1}{\sigma \cdot \sqrt{n}} \sum_{i=1}^{n} (X_i - \mu) \sim N(0, 1) \quad N = E[X_i]
$$

$$
\frac{1}{n} \sum_{i=1}^{n} (X_i - \mu) \to 0
$$

$\sigma^2 = var(X_i)$

---

In general, let

$$X = N(\mu, \sigma^2)$$

Mean:

$$f_X(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

---

A collection $\{X_1, \ldots, X_n\}$ is said to be *jointly Gaussian* if

$$\text{for all } \mathbf{a} \in \mathbb{R}^n, \sum_{i=1}^{n} a_i X_i \text{ is Gaussian}$$

Example of individually but not jointly Gaussian

Let $X$ be $N(0, 1)$

Let $Y$ be $\begin{cases} X & \text{if } |X| \leq 1 \\ -X & \text{if } |X| > 1 \end{cases}$, $Y$ is Gaussian
But

\[ \frac{1}{2} = X + Y = \begin{cases} 2X & \text{if } |X| \leq 1 \\ 0 & \text{if } |X| > 1 \end{cases} \]

Thus \( X + Y \) is not \( JG \)

Let \( \tilde{X} \in \mathbb{R}^n \) be vector with iid components \( X_i \sim N(0, 1) \). Then \( \tilde{X} \) is a \( JG \) random vector.

Proof: Let \( Y = a^T \tilde{X} \)

\[ g_Y(s) = E \left[ e^{s^T \tilde{X}} \right] \]

\[ = \prod_{i=1}^{n} E \left[ e^{s_i X_i} \right] \]

\[ = \prod_{i=1}^{n} \frac{e^{s_i^2 \sigma^2 / 2}}{\sqrt{2\pi \sigma^2}} \]

\[ = e^{\frac{1}{2} s^2 \sigma^2} \]

\[ \sigma^2 = \frac{\sum a_i^2}{m} \]

Yes \( Y \) is Gaussian

Thus \( \tilde{X} \) is \( JG \)
Definition: We say \( \hat{X} = N(\hat{\mu}, K\hat{X}) \) if \( \hat{X} \) is \( \mathcal{C} \)-
with covariance \( K\hat{X} \), and mean \( \hat{\mu} \).

Fact: \( g\hat{X}(s) = \exp(s^T\hat{\mu} + \frac{1}{2} s^TK\hat{X}s) \)

Proof: \( g\hat{X}(s) = E[\exp(s^T\hat{X})] \)

let \( Y^T = s^T\hat{X} \)

\[ = E[\exp(Y)] \]

\[ = g_Y(1) = C[\frac{\hat{\mu}^T Y + \frac{1}{2}s^T\sigma_Y^2}{s^T}] \]

\( \sigma_Y^2 \): \( \text{var}(Y) = \text{var}(s^T\hat{X}) \)

\[ = \sqrt{E[s^T\hat{X}X^T]} \]

\[ = E[st(\hat{X} - \hat{\mu})(X - \hat{\mu})^T] \]

\[ = s^TK\hat{X}s \]

\( \hat{Y} = E[s^T\hat{X}] = s^T\hat{\mu} \)

Substitute \( g\hat{X}(s) = \exp(s^T\hat{\mu} + \frac{1}{2} s^TK\hat{X}s) \)

Corollary: Joint distribution of \( \mathbf{U} = \sqrt{\frac{1}{2}}\hat{X} \) is specified completely by \( \hat{\mu} \) and \( K\hat{X} \).

Let \( \hat{w} = [w_1, \ldots, w_m]^T \) consist of iid Gaussian r.v.s with distribution \( NCO, 1) \)
Then \( f_w(\tilde{w}) = \frac{1}{(2\pi)^{m/2}} \exp \left[ -\frac{\tilde{w}^T \tilde{w}}{2} \right] \) is joint pdf.

Suppose \( \tilde{z} = A \tilde{w} + \tilde{\nu} \) then \( \tilde{z} \sim N(\nu, AA^T) \).

Proof: \( \tilde{z} \) is JG because its elements are lin. combs. of elements of \( \tilde{w} \):

\[
E[\tilde{z}] = E[A \tilde{w} + \tilde{\nu}] = A \cdot E[\tilde{w}] + E[\tilde{\nu}] = \nu
\]

\[
K_{\tilde{z}} = E[(A \tilde{w})(A \tilde{w})^T] = A \cdot E[\tilde{w} \tilde{w}^T] A^T
\]

\[
= A \cdot E[w_i^2 w_i w_i] A^T = A \cdot E[w_i^2] E[w_i] A^T = A \cdot I \cdot A^T
\]

\[
= AA^T
\]
Theorem

As a zero mean with arbitrary covariance matrix \( k_z \) and mean \( \tilde{w}_z \) can be constructed by taking

\[
\tilde{z} = A\tilde{w} + N
\]

for some \( A \), and \( N \)

---

Proof: Let

\[
K_z = P \Lambda P^*
\]

\( P = \) matrix of orthonormal eigenvectors,
\( \Lambda = \) diagonal matrix of
\( \) eigenvalues.

Let \( A = P \Lambda \sqrt{\tilde{w}_z} \)

Covariance \( (A\tilde{w} + N) = E[A\tilde{w}\tilde{w}^T A^T]\)

\[
= A E[\tilde{w}\tilde{w}^T] A^T
\]

\[
= P \Lambda^{1/2} I \Lambda^{1/2} P^T
\]

\[
= K_z
\]

\[
f_\tilde{z}(z) = \left| \det(CA^{-1}) \right| f_\tilde{w}(CA^{-1}(z - N)) \quad A^T \hat{A}
\]

\[
= \frac{1}{(2\pi)^{N/2} \det(K_z)^{1/2}} \exp \left[ -\frac{1}{2} (z - \hat{A})^T K_z^{-1} (z - \hat{A}) \right]
\]