Lecture 3 Review:

Random vectors: vectors of random variables.

- The expectation of a random vector is just the vector of expectations.
- $\text{cov}(\mathbf{X}, \mathbf{Y})$ is a matrix with i, j entry $\text{cov}(X_i, Y_j)$
- $\text{cov}(\mathbf{AX}, \mathbf{BY}) = \mathbf{A}\text{cov}(\mathbf{X}, \mathbf{Y})\mathbf{B}'$
- We introduced quadratic forms – $\mathbf{X}'\mathbf{AX}$, where $\mathbf{X}$ is a random vector and $\mathbf{A}$ is a matrix. More to come ...

4.1 Definition of the Multivariate Normal Distribution

The following are equivalent definitions of the multivariate normal distribution (MVN). Given a vector $\boldsymbol{\mu}$ and p.s.d. matrix $\boldsymbol{\Sigma}$, $\mathbf{Y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ if:

Definition 1: For p.d. $\boldsymbol{\Sigma}$, the density function of $\mathbf{Y}$ is

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2} \exp\{-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})\}.$$

Definition 2: The moment generating function (m.g.f.) of $\mathbf{Y}$ is

$$M_{\mathbf{Y}}(\mathbf{t}) \equiv E[e^{\mathbf{t}'\mathbf{Y}}] = \exp\{\boldsymbol{\mu}'\mathbf{t} + \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\}.$$

Definition 3: $\mathbf{Y}$ has the same distribution as

$$\mathbf{AZ} + \boldsymbol{\mu},$$

where $\mathbf{Z} = (Z_1, \ldots, Z_k)$ are independent $N(0, 1)$ random variables and $\mathbf{A}_{n \times k}$ satisfies $\mathbf{AA}' = \boldsymbol{\Sigma}$.

COMMENT: You may be inclined to focus on definition 1, but the others are more useful.

Theorem: Definitions 1, 2, and 3 are equivalent for $\mathbf{\Sigma} > 0$. Definitions 2 and 3 are equivalent for $\mathbf{\Sigma} \geq 0$

Proof of Def 3 $\Rightarrow$ Def 2:

For $Z_i \sim N(0,1)$,

$$M_{Z_i}(t_i) = E[e^{t_i Z_i}] = \int_{-\infty}^{\infty} e^{z_i t_i} \frac{e^{-z_i^2/2}}{\sqrt{2\pi}} dz_i = e^{t_i^2/2} \int_{-\infty}^{\infty} \frac{e^{-(z_i-t_i)^2/2}}{\sqrt{2\pi}} dz_i = e^{t_i^2/2}.$$

If $\mathbf{Z} = (Z_1, \ldots, Z_k)$ is a random sample from $N(0,1)$, then

$$M_{\mathbf{Z}}(\mathbf{t}) = E[e^{\Sigma_i z_i t_i}] = E[\prod_{i=1}^{k} e^{z_i t_i}] \stackrel{\text{ind}}{=} \prod_{i=1}^{k} E[e^{z_i t_i}] = \prod_{i=1}^{k} M_{Z_i}(t_i) = \exp\{\sum_{i=1}^{k} t_i^2/2\} = \exp\{\mathbf{t}'\mathbf{t}/2\}.$$

If $\mathbf{Y} = \mathbf{A}\mathbf{Z} + \boldsymbol{\mu}$,

$$\begin{aligned}
M_{\mathbf{Y}}(\mathbf{t}) &\equiv E[\exp\{\mathbf{Y}'\mathbf{t}\}] \\
&= E[\exp\{(\mathbf{A}\mathbf{Z} + \boldsymbol{\mu})'\mathbf{t}\}] \\
&= \exp\{\boldsymbol{\mu}'\mathbf{t}\} E[\exp\{(\mathbf{A}\mathbf{Z})'\mathbf{t}\}] \\
&= \exp\{\boldsymbol{\mu}'\mathbf{t}\} M_{\mathbf{Z}}(\mathbf{A}'\mathbf{t}) \\
&= \exp\{\boldsymbol{\mu}'\mathbf{t}\} \exp\{\frac{1}{2}(\mathbf{A}'\mathbf{t})'(\mathbf{A}'\mathbf{t})\} \\
&= \exp\{\boldsymbol{\mu}'\mathbf{t}\} \exp\{\frac{1}{2}\mathbf{t}'(\mathbf{A}\mathbf{A}')\mathbf{t}\} \\
&= \exp\{\boldsymbol{\mu}'\mathbf{t} + \frac{1}{2}\mathbf{t}'\mathbf{\Sigma}\mathbf{t}\}.
\end{aligned}$$

Proof of Def 2 $\Rightarrow$ Def 3:

Since $\boldsymbol{\Sigma} \geq 0$ (and $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}'$), there exists an orthogonal matrix, $\mathbf{T}^{n\times n}$, such that $\mathbf{T}'\boldsymbol{\Sigma}\mathbf{T} = \boldsymbol{\Lambda}$, where $\boldsymbol{\Lambda}$ is diagonal with non-negative elements. Therefore,

$$\begin{aligned} \boldsymbol{\Sigma} &= \mathbf{T}\boldsymbol{\Lambda}\mathbf{T}' \\ &= \mathbf{T}\boldsymbol{\Lambda}^{1/2}\boldsymbol{\Lambda}^{1/2}\mathbf{T}' \\ &= (\mathbf{T}\boldsymbol{\Lambda}^{1/2})(\mathbf{T}\boldsymbol{\Lambda}^{1/2})' \\ &= \mathbf{A}\mathbf{A}'. \end{aligned}$$

In other words, let $\mathbf{A} = \mathbf{T}\boldsymbol{\Lambda}^{1/2}$. Now, in the previous proof we showed the m.g.f. of $\mathbf{AZ} + \boldsymbol{\mu}$ is

$$\exp\{\boldsymbol{\mu}'\mathbf{t} + \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\},$$

the same as $\mathbf{Y}$. Because the m.g.f. uniquely determines the distribution (when the m.g.f. exists in a neighbourhood of $\mathbf{t} = \mathbf{0}$), $\mathbf{Y}$ has the same distribution as $\mathbf{AZ} + \boldsymbol{\mu}$.

Proof of Def 3 ⇒ Def 1: (for p.d. $\mathbf{\Sigma}$).

Because $\mathbf{\Sigma}$ is positive definite, there is a non-singular $\mathbf{A}_{n\times n}$ such that $\mathbf{A}\mathbf{A}' = \mathbf{\Sigma}$ (lecture notes # 2, page 10). Let $\mathbf{Y} = \mathbf{A}\mathbf{Z} + \boldsymbol{\mu}$, where $\mathbf{Z} = (Z_1, \ldots, Z_n)$ is a random sample from $N(0, 1)$. The density of $\mathbf{Z}$ is

$$f_{\mathbf{Z}}(\mathbf{z}) = \prod_{i=1}^{n}(2\pi)^{-1/2}\exp\{-\frac{1}{2}Z_i^2\} = (2\pi)^{-n/2}\exp\{-\frac{1}{2}\mathbf{Z}'\mathbf{Z}\}.$$

The density function of $\mathbf{Y}$ is

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{Z}}(\mathbf{z}(\mathbf{y}))|J|,$$

where J is the Jacobian

$$J = \left|\left(\frac{\partial Z_i}{\partial Y_j}\right)\right| = |\mathbf{A}^{-1}| = |\mathbf{A}|^{-1},$$

because $\mathbf{Z} = \mathbf{A}^{-1}(\mathbf{Y} - \boldsymbol{\mu})$. Therefore,

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= f_{\mathbf{Z}}(\mathbf{A}^{-1}(\mathbf{y} - \boldsymbol{\mu}))|\mathbf{A}|^{-1} \\ &= (2\pi)^{-n/2}|\mathbf{A}|^{-1}\exp\{-\frac{1}{2}[\mathbf{A}^{-1}(\mathbf{y} - \boldsymbol{\mu})]'[\mathbf{A}^{-1}(\mathbf{y} - \boldsymbol{\mu})]\} \\ &= (2\pi)^{-n/2}|\mathbf{A}\mathbf{A}'|^{-1/2}\exp\{-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'(\mathbf{A}\mathbf{A}')^{-1}(\mathbf{y} - \boldsymbol{\mu})\} \\ &= (2\pi)^{-n/2}|\mathbf{\Sigma}|^{-1/2}\exp\{-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{y} - \boldsymbol{\mu})\} \end{aligned}$$

(Using: $|\mathbf{A}|^{-1} = |\mathbf{A}|^{-\frac{1}{2}}|\mathbf{A}|^{-\frac{1}{2}} = |\mathbf{A}|^{-\frac{1}{2}}|\mathbf{A}'|^{-\frac{1}{2}} = (|\mathbf{A}||\mathbf{A}'|)^{-\frac{1}{2}} = |\mathbf{A}\mathbf{A}'|^{-\frac{1}{2}}$)

Proof of Def 1 ⇒ Def 2 (for p.d. $\mathbf{\Sigma}$): *Exercise: Use pdf in Def 1 and solve directly for mgf.*

4.2 Properties of the Multivariate Normal Distribution

1. $E[\mathbf{Y}] = \boldsymbol{\mu}$, $\text{cov}(\mathbf{Y}) = \boldsymbol{\Sigma}$ (verify using Definition 3 and properties of means and covariances of random vectors)

2. If $\mathbf{Z} = (Z_1, \ldots, Z_n)$ is a random sample from $N(0, 1)$ then $\mathbf{Z}$ has the $N_n(\mathbf{0_n}, \mathbf{I_{n\times n}})$ distribution (use Definition 3).

3. If $\boldsymbol{\Sigma}$ is not p.d. then $\mathbf{Y}$ has a *singular MVN* distribution and no density function exists.

Example: A singular MVN distribution. Let $\mathbf{Z} = (Z_1, Z_2)' \sim N_2(\mathbf{0}, \mathbf{I})$, and let $\mathbf{A}$ be the linear transformation matrix $\mathbf{A} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

Let $\mathbf{Y} = (Y_1, Y_2)'$ be the linear transformation

$$\mathbf{Y} = \mathbf{AZ} = \begin{pmatrix} (Z_1 - Z_2)/2 \\ (Z_2 - Z_1)/2 \end{pmatrix}.$$

By Definition 3, $\mathbf{Y} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \mathbf{AA}'$.

$\boldsymbol{\Sigma} = \mathbf{AA}' = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$

corr$=\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$. Makes sense!

4.3 Linear Transformations of MVN Vectors

1. If $\mathbf{Y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\mathbf{C}_{p\times n}$ is a matrix of rank p, then $\mathbf{CY} \sim N_p(\mathbf{C}\boldsymbol{\mu}, \mathbf{C\Sigma C}')$.

   Proof: By Def 3, $\mathbf{Y} = \mathbf{AZ} + \boldsymbol{\mu}$, where $\mathbf{AA}' = \boldsymbol{\Sigma}$. Then

$$\begin{aligned}
\mathbf{CY} &= \mathbf{C}(\mathbf{AZ} + \boldsymbol{\mu}) \\
&= \mathbf{CAZ} + \mathbf{C}\boldsymbol{\mu} \\
&\sim N(\mathbf{C}\boldsymbol{\mu}, \mathbf{CA}(\mathbf{CA})') \quad \text{(by Def 3)} \\
&= N(\mathbf{C}\boldsymbol{\mu}, \mathbf{C}(\mathbf{AA}')\mathbf{C}) \\
&= N(\mathbf{C}\boldsymbol{\mu}, \mathbf{C\Sigma C}').
\end{aligned}$$

2. $\mathbf{Y}$ is MVN if and only if $\mathbf{a}'\mathbf{Y}$ is normally distributed for all non-zero vectors $\mathbf{a}$.

   Proof: If $\mathbf{Y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ then $\mathbf{a}'\mathbf{Y} \sim N(\mathbf{a}'\boldsymbol{\mu}, \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a})$ by 4.3.1 (above).

   Conversely, assume that $X = \mathbf{a}'\mathbf{Y}$ is univariate normal for all non-zero $\mathbf{a}$. In other words, $X \sim N(\mathbf{a}'\boldsymbol{\mu}, \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a})$, where $\boldsymbol{\mu} = E[\mathbf{Y}]$ and $\boldsymbol{\Sigma} = \text{cov}(\mathbf{Y})$. Using the form of the m.g.f. of a univariate normal random variable, the m.g.f. of X is

$$E[\exp(Xt)] = M_X(t) = \exp\{(\mathbf{a}'\boldsymbol{\mu})t + \frac{1}{2}(\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a})t^2\}$$

   for all t. Setting $t = 1$ in $M_X(t)$ gives $M_{\mathbf{Y}}$:

$$E[\exp(\mathbf{a}'\mathbf{Y})] = M_X(t=1) = \exp\{(\mathbf{a}'\boldsymbol{\mu}) + \frac{1}{2}(\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a})\} = M_{\mathbf{Y}}(\mathbf{a}),$$

   which is the m.g.f. of $N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Therefore, $\mathbf{Y} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ by Def 2.

   In words, a random vector is MVN iff every linear combination of its random variable components is a normal random variable.