Math 5305 Notes
Chapter 3

Jesse Crawford

Department of Mathematics
Tarleton State University
### Definition
- Suppose $n$ and $m$ are positive integers.
- The set of $n \times m$ matrices with real entries is denoted by $\mathbb{R}^{n \times m}$.

### Definition
- Suppose $A, B \in \mathbb{R}^{n \times m}$.
- Define $A + B \in \mathbb{R}^{n \times m}$ by

$$ (A + B)_{ij} = A_{ij} + B_{ij}, \text{ for } i = 1, \ldots, n \text{ and } j = 1, \ldots, m. $$
Definition

- Suppose $A \in \mathbb{R}^{I \times J}$ and $B \in \mathbb{R}^{J \times K}$.
- Define $AB \in \mathbb{R}^{I \times K}$ by

$$
(AB)_{ik} = \sum_{j=1}^{J} A_{ij}B_{jk}, \text{ for } i = 1, \ldots, I \text{ and } k = 1, \ldots, K.
$$

Proposition

Consider matrices $A \in \mathbb{R}^{I \times J}$, $B \in \mathbb{R}^{J \times K}$, $C \in \mathbb{R}^{K \times L}$, and $D \in \mathbb{R}^{L \times M}$. Then, for any $i = 1, \ldots, I$ and $m = 1, \ldots, M$,

$$
(ABCD)_{im} = \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{\ell=1}^{L} A_{ij}B_{jk}C_{k\ell}D_{\ell m}.
$$
Transpose and Trace

**Definition**

- Suppose \( A \in \mathbb{R}^{n \times m} \).
- Define \( A' \in \mathbb{R}^{m \times n} \) by
  \[
  (A')_{ji} = A_{ij}, \text{ for } i = 1, \ldots, n \text{ and } j = 1, \ldots, m.
  \]
- If \( A' = A \), then \( A \) is called *symmetric*.

**Definition**

- Suppose \( A \in \mathbb{R}^{n \times n} \).
- Define the trace of \( A \), \( \text{trace}(A) \) by
  \[
  \text{trace}(A) = \sum_{i=1}^{n} A_{ii}.
  \]
Definition

Given two vectors $u, v \in \mathbb{R}^n$, their *inner product* is

$$u \cdot v = u'v = u_1v_1 + \cdots + u_nv_n.$$  

Definition

The *norm*, *length*, or *magnitude* of a vector $u \in \mathbb{R}^n$ is

$$\|u\| = \sqrt{u'u} = \sqrt{u_1^2 + \cdots + u_n^2}.$$
Section 3.1: Introduction

Section 3.2: Determinants and Inverses

Section 3.3: Random Vectors

Section 3.4: Positive Definite Matrices

Section 3.5: The Normal Distribution
If $A$ is a square matrix, its *determinant* is denoted by $\det(A)$ or $|A|$.  

**Examples:**

\[
\begin{vmatrix} 1 & 2 \\ 5 & 3 \end{vmatrix} = 1 \cdot 3 - 5 \cdot 2 = -7
\]

\[
\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 0 & 1 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 3 & 1 \\ 1 & 1 \end{vmatrix} - 2 \cdot \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} + 3 \cdot \begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix}
\]

\[
= 1 \cdot 2 - 2 \cdot 2 + 3 \cdot 2 = 4
\]
Inverses and Kernels

**Definition**

An $n \times n$ matrix $A$ is **invertible** if there exists an $n \times n$ matrix $A^{-1}$, such that

$$AA^{-1} = A^{-1}A = I.$$ 

**Definition**

The **kernel** of an $n \times m$ matrix $A$ is

$$\ker(A) = \{ v \in \mathbb{R}^m \mid Av = 0 \}.$$
### Linear Independence and Rank

#### Definition
- Suppose $v_1, v_2, \ldots, v_k$ are vectors. They are **linearly independent** if, for any scalars $c_1, c_2, \ldots, c_k$,

  \[
  c_1 v_1 + c_2 v_2 + \cdots + c_k v_k = 0 \implies c_1 = c_2 = \cdots = c_k = 0.
  \]

#### Definition
- The **rank** of a matrix is the maximum number of linearly independent columns it has.
- If $X$ is an $n \times p$ matrix, and rank$(X) = p$, then $X$ has **full rank**.

#### Proposition
The rank of a matrix is the number of nonzero rows it has in reduced row echelon form.
Theorem

For an $n \times n$ matrix $A$, the following are equivalent:

- $\det(A) \neq 0$
- $A$ is invertible
- $\ker(A) = \{0\}$
- For any $c \in \mathbb{R}^n$, $Ac = 0$ implies $c = 0$
- All of the columns of $A$ are linearly independent
- $\text{rank}(A) = n$
- $A$ has full rank
The Big Theorem for Nonsquare Matrices

**Theorem**

For an \( n \times p \) matrix \( X \), the following are equivalent:

- \( \ker(X) = \{0\} \)
- For any \( c \in \mathbb{R}^p \), \( Xc = 0 \) implies \( c = 0 \)
- All of the columns of \( X \) are linearly independent
- \( \text{rank}(X) = p \)
- \( X \) has full rank
1. Section 3.1: Introduction
2. Section 3.2: Determinants and Inverses
3. Section 3.3: Random Vectors
4. Section 3.4: Positive Definite Matrices
5. Section 3.5: The Normal Distribution
Covariance Between Two Random Variables

Definition

Let $X$ and $Y$ be two random variables.

The covariance between $X$ and $Y$ is

$$\text{cov}(X, Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$$

The covariance measures the strength of the association between $X$ and $Y$, and the sign indicates whether the relationship is positive or negative.
Correlation Between Two Random Variables

Definition

The correlation coefficient between $X$ and $Y$ is

$$\rho = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}.$$ 

$-1 \leq \rho \leq 1$

Values of $\rho$ near 1 indicate a strong positive relationship.

Values of $\rho$ near $-1$ indicate a strong negative relationship.

Values of $\rho$ near 0 indicate a weak or nonlinear relationship.
Strong Positive Correlation

\[ \text{corr}(X, Y) = 0.9 \]
Strong Negative Correlation

corr(x, y) = -0.9
Virtually No Correlation

$\text{corr}(X, Y) = 0.06$
Definition

- A random vector is a vector whose components are random variables.
- If $U_1, \ldots, U_n$ are random variables, then

$$ U = \begin{pmatrix} U_1 \\ \vdots \\ U_n \end{pmatrix} $$

is a random vector.
Expected Value of a Random Vector

**Definition**

- Given a random vector

\[
U = \begin{pmatrix}
U_1 \\
\vdots \\
U_n
\end{pmatrix}
\]

the *expected value* of \( U \) is

\[
E(U) = \begin{pmatrix}
E(U_1) \\
\vdots \\
E(U_n)
\end{pmatrix}
\]

\[
[E(U)]_i = E(U_i), \text{ for every } i
\]
Expected Value of a Random Matrix

Definition

Given a random matrix

\[ U = \begin{pmatrix}
  U_{11} & \cdots & U_{1m} \\
  \vdots & \ddots & \vdots \\
  U_{n1} & \cdots & U_{nm}
\end{pmatrix} \]

the expected value of \( U \) is

\[ E(U) = \begin{pmatrix}
  E(U_{11}) & \cdots & E(U_{1m}) \\
  \vdots & \ddots & \vdots \\
  E(U_{n1}) & \cdots & E(U_{nm})
\end{pmatrix} \]

\[ [E(U)]_{ij} = E(U_{ij}), \text{ for every } i, j \]
Covariance Matrix of a Random Vector

**Definition**

- Given a random vector \( U = \begin{pmatrix} U_1 \\ \vdots \\ U_n \end{pmatrix} \)

the covariance matrix of \( U \) is

\[
\text{cov}(U) = E \left\{ \begin{pmatrix} U_1 - E(U_1) \\ \vdots \\ U_n - E(U_n) \end{pmatrix} (U_1 - E(U_1), \ldots, U_n - E(U_n)) \right\}
\]
More on Covariance

\[
\text{cov}(U) = E[(U - E(U))(U - E(U))'] = E(UU') - E(U)E(U)'
\]

- The \(i\)th diagonal element of \(\text{cov}(U)\) is \(\text{Var}(U_i)\).
- The \((i, j)\) entry of \(\text{cov}(U)\) is \(\text{cov}(U_i, U_j)\).
Section 3.1: Introduction

Section 3.2: Determinants and Inverses

Section 3.3: Random Vectors

Section 3.4: Positive Definite Matrices

Section 3.5: The Normal Distribution
Definition

An $n \times n$ matrix $G$ is *non-negative definite* if
- $G$ is symmetric, and
- $x'Gx \geq 0$, for all $x \in \mathbb{R}^n$.

Definition

An $n \times n$ matrix $G$ is *positive definite* if
- $G$ is symmetric, and
- $x'Gx > 0$, for all nonzero $x \in \mathbb{R}^n$.

Note that any positive definite matrix is non-negative definite.
We will denote the set of $n \times n$ positive definite matrices by $\text{PD}(n)$. 
Diagonal and Orthogonal Matrices

Definition
A matrix $D$ is diagonal if all of its entries off the diagonal are zero,

$$D_{ij} = 0 \text{ when } i \neq j.$$  

Definition
- A matrix $R$ is orthogonal if $R' R = I$.
- If $R$ is orthogonal, $R^{-1} = R'$, and $RR' = I$. 

(Tarleton State University)
Diagonalizing a Positive Definite Matrix

**Theorem**

- **G is non-negative definite iff** there exists
- **a diagonal matrix D whose diagonal entries are non-negative and**
- **an orthogonal matrix R,**
- **such that** \( G = RDR' \)

- This theorem also holds if we replace both instances of “non-negative” with “positive”.
- The columns of \( R \) are the **eigenvectors** of \( G \).
- The diagonal entries of \( D \) are the **eigenvalues** of \( G \).
Eigenvalues and Eigenvectors

Definition

Suppose $A \in \mathbb{R}^{n \times n}$, $x \in \mathbb{R}^n$ is nonzero, and $\lambda \in \mathbb{R}$, such that

$$Ax = \lambda x.$$

- Then $\lambda$ is an *eigenvalue* of $A$, and
- $x$ is an *eigenvector* of $A$ corresponding to $\lambda$. 
**Definition (Univariate Normal Distribution)**

- The *normal distribution* on $\mathbb{R}$ with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$ is given by the p.d.f.

\[
f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left( -\frac{(x - \mu)^2}{2\sigma^2} \right), \text{ for } x \in \mathbb{R}.
\]

- Denoted by $N(\mu, \sigma^2)$

- If $X \sim N(\mu, \sigma^2)$, then $E(X) = \mu$, and $\text{Var}(X) = \sigma^2$, so these parameters deserve their names.

**Proposition**

- If $X \sim N(\mu, \sigma^2)$, then

\[
Z = \frac{X - \mu}{\sigma} \sim N(0, 1).
\]

- The distribution $N(0, 1)$ is called the *standard normal distribution*.
Definition (Multivariate Normal Distribution)

- The *multivariate normal distribution* on $\mathbb{R}^n$ with mean $\mu \in \mathbb{R}^n$ and covariance matrix $\Sigma \in \text{PD}(n)$ is given by the p.d.f.

\[
f(x) = \left( \frac{1}{\sqrt{2\pi}} \right)^n \frac{1}{\sqrt{\det \Sigma}} \exp \left[ -\frac{1}{2} (x - \mu)'\Sigma^{-1}(x - \mu) \right], \text{ for } x \in \mathbb{R}^n.
\]

- Denoted by $N(\mu, \Sigma)$.

- If $X \sim N(\mu, \Sigma)$, then $E(X) = \mu$, and $\text{cov}(X) = \Sigma$, so these parameters deserve their names.

- The random variables $X_1, \ldots, X_n$ are *jointly normal*. 
Covariance and Independence

Definition

- Two random variables $X$ and $Y$ with $\text{cov}(X, Y) = 0$ are said to be uncorrelated.
- In general, if $X$ and $Y$ are independent, then $X$ and $Y$ are uncorrelated:

  $$X \text{ and } Y \text{ independent} \Rightarrow \text{cov}(X, Y) = 0.$$ 

- The converse is generally not true. There are examples of uncorrelated random variables that are dependent.
- For jointly normal random variables, independence is equivalent to being uncorrelated.
**Proposition**

Suppose $Z_1, \ldots, Z_n$ are IID $N(0, 1)$ random variables. Then $Z = (Z_1, \ldots, Z_n)' \sim N(0, I)$.

**Proposition**

Suppose $\mu \in \mathbb{R}^n$ and $\Sigma$ is an $n \times n$ non-negative definite matrix. Note that $\Sigma$ has a non-negative definite square root $\Sigma^{\frac{1}{2}}$. Then $X \sim N(\mu, \Sigma)$ iff there exists a random vector $Z \sim N(0, I)$, such that

$$X = \mu + \Sigma^{\frac{1}{2}} Z.$$
Proposition

Let $X \sim N(\mu, \Sigma)$, $A \in \mathbb{R}^{m \times n}$, and $c, d \in \mathbb{R}^n$.

$$AX \sim N(A\mu, A\Sigma A')$$

That is, $AX$ has a multivariate normal distribution, and

$$E(AX) = AE(X) \text{ and } \text{cov}(AX) = A\text{cov}(X)A'.$$

The covariance between $c'X$ and $d'X$ is

$$\text{cov}(c'X, d'X) = c'\Sigma d.$$