Chapter 2 Vector Spaces and Matrices
Vectors and Matrices

- $n$-dimensional column vector and row vector

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad a^T = [a_1, a_2, \ldots, a_n]$$

- Properties

\[ a + b = b + a \]
\[ (a + b) + c = a + (b + c) \]
\[ 0 = (0, 0, \ldots, 0) \]
\[ a + 0 = 0 + a = a \]
\[ \alpha(a + b) = \alpha a + \alpha b \]
\[ \alpha(\beta a) = (\alpha \beta) a \]
\[ 1a = a \]
\[ (-1)a = -a \]
\[ \alpha 0 = 0 = 0a \]
Linearly Independent

- A set of vectors \( \{a_1, ..., a_k\} \) is said to be \textit{linearly independent} if the equality \( \alpha_1 a_1 + \alpha_2 a_2 + ... + \alpha_k a_k = 0 \) implies that all coefficients \( \alpha_i, i = 1, ..., k \), are equal to zero.
- Any set of vectors containing the vector \( 0 \) is \textit{linearly dependent}.
- A set of vectors \( \{a_1, ..., a_k\} \) is linearly dependent if and only if one of the vectors from the set is a \textit{linear combination} of the remaining vectors.
Subspace

- A subset $\mathcal{V}$ of $\mathbb{R}^n$ is called a **subspace** of $\mathbb{R}^n$ if $\mathcal{V}$ is closed under addition and closed under scalar multiplication.
  - If $\mathbf{a}$ and $\mathbf{b}$ are vectors in $\mathcal{V}$, then the vectors $\mathbf{a} + \mathbf{b}$ and $\alpha \mathbf{a}$ are also in $\mathcal{V}$ for every scalar $\alpha$.

- Let $\mathbf{a}_1, \ldots, \mathbf{a}_k$ be arbitrary vectors in $\mathbb{R}^n$. The set of all their linear combinations is called the **span** of $\mathbf{a}_1, \ldots, \mathbf{a}_k$.

- Given a subspace $\mathcal{V}$, any set of linearly independent vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\} \subset \mathcal{V}$ such that $\mathcal{V} = \text{span}[\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k]$ is referred to as a **basis** of the subspace $\mathcal{V}$.

- All bases of a subspace $\mathcal{V}$ contain the same number of vectors. This number is called the **dimension** of $\mathcal{V}$. 
If \( \{a_1, \ldots, a_k\} \) is a basis of \( V \), then any vector \( a \) of \( V \) can be represented uniquely as \( a = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_k a_k \), where \( \alpha_i \in \mathbb{R}, i = 1, 2, \ldots, k \).

The coefficients \( \alpha_i, i = 1, 2, \ldots, k \), are called the coordinates of \( a \) with respect to the basis \( \{a_1, \ldots, a_k\} \).
Rank of A Matrix

Consider the \( m \times n \) matrix

\[
A = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\]

The maximal number of linearly independent columns of \( A \) is called the rank of \( A \), denoted by \( \text{rank}(A) \). The \( \text{rank}(A) \) is the dimension of \( \text{span}[a_1, a_2, \ldots, a_k] \)

The rank of a matrix \( A \) is invariant under the following operations:

- Multiplication of the columns of \( A \) by nonzero scalars.
- Interchange of the columns.
- Addition to a given column a linear combination of other columns.
The determinant of the matrix $A = [a_1, a_2, \ldots, a_n]$ is a function of its columns and has the following properties:

- $\det[a_1, \ldots, a_{k-1}, \alpha a_k, \ldots, a_n] = \alpha \det[a_1, \ldots, a_{k-1}, a_k, \ldots, a_n]$
- If we have $a_k = a_{k+1}$, $\det(A) = 0$
- Determinant of an identity matrix is 1.
- If one of the columns is 0, then the determinant is equal to zero.
- The determinant does not change if we add to a column another column multiplied by a scalar.
- The determinant changes its sign if we interchange columns.
Determinant

- A *pth-order minor* of an $m \times n$ matrix $A$, with $p \leq \min\{m, n\}$, is the determinant of a $p \times p$ matrix obtained from $A$ by deleting $m-p$ rows and $n-p$ columns.

- If an $m \times n$ matrix $A \ (m \geq n)$ has a nonzero $n$th-order minor, then the columns of $A$ are linearly independent; that is, $\text{rank}(A)=n$.

- The rank of a matrix is equal to the highest order of its nonzero minor(s).

- A *nonsingular* (or *invertible*) matrix is a square matrix whose determinant is nonzero.
Given $m$ equations in $n$ unknowns of the form

\[
\begin{align*}
  a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\
  a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\
  \vdots & \\
  a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. 
\end{align*}
\]

We can represent the system of equations as $Ax = b$

\[
A = \begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix} = [a_1, a_2, \ldots, a_n], \quad b = \begin{bmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_m
\end{bmatrix}, \quad x = \begin{bmatrix}
  x_1 \\
  x_2 \\
  \vdots \\
  x_n
\end{bmatrix}
\]

The system of equations $Ax = b$ has a solution if and only if $\text{rank}(A) = \text{rank}([A, b])$

If $\text{rank}(A)=m$, a solution to $Ax = b$ can be obtained by assigning arbitrary values for $n-m$ variables and solving for the remaining ones.
General and Particular Solutions

- **Theorem 4.7.2**
  - If $x_0$ denotes any single solution of a consistent linear system $Ax = b$, and if $v_1, v_2, \ldots, v_k$ form a basis for the null space of $A$, (that is, the solution space of the homogeneous system $Ax = 0$), then every solution of $Ax = b$ can be expressed in the form
    \[ x = x_0 + c_1v_1 + c_2v_2 + \cdots + c_kv_k \]
  - Conversely, for all choices of scalars $c_1, c_2, \ldots, c_k$, the vector $x$ in this formula is a solution of $Ax = b$. 

Example (General Solution of $A\mathbf{x} = \mathbf{b}$)

- The solution to the nonhomogeneous system

\[
\begin{align*}
  x_1 + 3x_2 - 2x_3 + 2x_5 &= 0 \\
  2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= -1 \\
  5x_3 + 10x_4 + 15x_6 &= 5 \\
  2x_1 + 5x_2 + 8x_4 + 4x_5 + 18x_6 &= 6
\end{align*}
\]

is

\[
\begin{align*}
  x_1 &= -3r - 4s - 2t, \\
  x_3 &= -2s, \\
  x_4 &= s, \\
  x_5 &= t, \\
  x_6 &= 1/3
\end{align*}
\]

which is the general solution.

- The vector $\mathbf{x}_0$ is a **particular solution** of nonhomogeneous system, and the linear combination $\mathbf{x}$ is the **general solution** of the homogeneous system.
Inner Products and Norms

For $x, y \in \mathbb{R}^n$, their **Euclidean inner product** is

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^T y$$

**Properties:**

- **Positivity:** $\langle x, x \rangle \geq 0$, $\langle x, x \rangle = 0$ if and only if $x = 0$
- **Symmetry:** $\langle x, y \rangle = \langle y, x \rangle$
- **Additivity:** $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- **Homogeneity:** $\langle r \cdot x, y \rangle = r \cdot \langle x, y \rangle$ for every $r \in \mathbb{R}$

The vectors $x$ and $y$ are said to be **orthogonal** if $\langle x, y \rangle = 0$

**The Euclidean norm** of a vector $x$ is defined as

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{x^T x}$$
Inner Products and Norms

- **Cauchy-Schwarz Inequality**: For any two vectors \( x \) and \( y \) in \( \mathbb{R}^n \), the Cauchy-Schwarz inequality
\[
|\langle x, y \rangle| \leq \|x\| \|y\|
\]
holds. Furthermore, equality holds if and only if \( x = \alpha y \) for some \( \alpha \in \mathbb{R} \).

- The Euclidean norm of a vector has the following properties:
  - Positivity: \( \|x\| \geq 0 \), \( \|x\| = 0 \) if and only if \( x = 0 \)
  - Homogeneity: \( \|rx\| = |r|\|x\|, r \in \mathbb{R} \)
  - Triangle inequality: \( \|x + y\| \leq \|x\| + \|y\| \)
Inner Products and Norms

- The Euclidean norm is an example of a general vector norm, which is any function satisfying the three properties of positivity, homogeneity, and triangle inequality.

- Other vector norms:
  - 1-norm: $\|x\|_1 = |x_1| + \cdots + |x_n|$  
  - $\infty$-norm: $\|x\|_{\infty} = \max_i |x_i|$  
  - The Euclidean norm is often referred as the 2-norm, denoted by $\|x\|_2$  
  - $p$-norm: $\|x\|_p = \begin{cases} (|x_1|^p + \cdots + |x_n|^p)^{1/p} & \text{if } 1 \leq p < \infty \\ \max\{|x_1|, \ldots, |x_n|\} & \text{if } p = \infty \end{cases}$
A function \( f : \mathbb{R}^n \rightarrow \mathbb{R}^m \) is **continuous** at \( x \) if for all \( \varepsilon > 0 \), there exists \( \delta > 0 \) such that \( \| y - x \| < \delta \implies \| f(y) - f(x) \| < \varepsilon \).

If the function \( f \) is continuous at every point in \( \mathbb{R}^n \), we say that it is continuous on \( \mathbb{R}^n \).
For the complex vector space $C^n$, we define an inner product
\[ \langle x, y \rangle = \sum_{i=1}^{n} x_i \bar{y}_i, \]
where the bar denotes complex conjugation.

Properties:
\[ \langle x, x \rangle \geq 0, \quad \langle x, x \rangle = 0 \text{ if and only if } x = 0 \]
\[ \langle x, y \rangle = \overline{\langle y, x \rangle} \]
\[ \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \]
\[ \langle r x, y \rangle = r \langle x, y \rangle, \text{ where } r \in C \]