1. Vector Spaces (Sections 1.1–1.6)

Important things you need to know include the definition of a vector space, a subspace, a spanning set, linear (in)dependence, a basis and the dimension of a vector space. You should also be able to solve a system of linear equations using row operations and find the solution set of a given linear system. Review chapter 1 of the textbook and/or the class notes if you feel uncomfortable with any of these concepts/terminology.

2. Linear transformations and matrices (Sections 2.1–2.5 & 3.2)

Throughout, let $V$ and $W$ be vector spaces over $\mathbb{R}$.

- A function $T : V \to W$ is linear if $T(cu + v) = cT(u) + v$ for every $u, v \in V$ and $c \in \mathbb{R}$.
- Let $T : V \to W$ be linear. The **null space** of $T$ is
  \[ \mathcal{N}(T) := \{ v \in V \mid T(v) = 0 \} \]
  which is a subspace of $V$ and its dimension is called the **nullity** of $T$. The **range** of $T$ is
  \[ \mathcal{R}(T) := \{ T(v) \mid v \in V \}, \]
  which is a subspace of $W$ and its dimension is called the **rank** of $T$.
- **(Dimension Theorem)** If $V$ and $W$ are finite dimensional and $T : V \to W$ is linear, then
  \[ \text{nullity}(T) + \text{rank}(T) = \dim(V). \]
- A linear transformation $T : V \to W$ is an **isomorphism** if it is invertible (which is equivalent to one-to-one and onto).
- $V$ and $W$ are said to be isomorphic if there exists an isomorphism $T : V \to W$.
- If $V$ and $W$ are finite dimensional, then $V$ and $W$ are isomorphic if and only if $\dim(V) = \dim(W)$.
- Suppose $V$ is finite dimensional and $\beta = \{ v_1, v_2, \ldots, v_n \}$ is an ordered basis for $V$. If $v \in V$ satisfies
  \[ v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n \]
  for some $a_1, a_2, \ldots, a_n \in \mathbb{R}$, then the coordinate vector of $v$ with respect to $\beta$ is
  \[ [v]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}. \]
Let $V$ and $W$ be finite dimensional vector spaces with bases $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_m\}$, respectively and let $T : V \to W$ be linear. The matrix of $T$ in the bases $\beta$ and $\gamma$ is

$$[T]_\beta^\gamma = \begin{pmatrix} [T(v_1)]_\gamma & [T(v_2)]_\gamma & \cdots & [T(v_n)]_\gamma \end{pmatrix} \in M_{m \times n}(\mathbb{R}).$$

For every $v \in V$, we have

$$[T(v)]_\gamma = [T]_\beta^\gamma [v]_\beta.$$

Define

$$\mathcal{L}(V,W) = \{ T : V \to W \mid T \text{ is linear} \},$$

which is a vector space under the “natural” addition and scalar multiplication. We define $\mathcal{L}(V) = \mathcal{L}(V,V)$.

Let $X, Y, Z$ be finite dimensional vector spaces with bases $\alpha, \beta, \gamma$ and let $S \in \mathcal{L}(X,Y)$ and $T \in \mathcal{L}(Y,Z)$. Then $T \circ S \in \mathcal{L}(X,Z)$ and

$$[T \circ S]_\alpha^\gamma = [T]_\beta^\gamma [S]_\alpha^\beta.$$

A matrix $A \in M_{n \times n}(\mathbb{R})$ is invertible if there exists $B \in M_{n \times n}(\mathbb{R})$ such that $AB = I_n = BA$.

Let $V$ and $W$ are finite dimensional vector spaces with bases $\beta$ and $\gamma$. Suppose that $\dim(V) = n$ and $\dim(W) = m$, then the function

$$\Phi : \mathcal{L}(V,W) \to M_{n \times n}(\mathbb{R})$$

$$T \mapsto [T]_\beta^\gamma$$

is an isomorphism. In particular, $\dim(\mathcal{L}(V,W)) = mn$.

Let $V, W, \beta, \gamma$ be given as above. Then $T \in \mathcal{L}(V,W)$ is invertible if and only if $[T]_\beta^\gamma$ is invertible. Furthermore, $([T]_\beta^\gamma)^{-1} = [T^{-1}]_\gamma^\beta$.

(Computing $A^{-1}$) Let $A \in M_{n \times n}(\mathbb{R})$. Apply row operations to $(A \mid I_n)$ to obtain $(\text{RREF}(A) \mid B)$, where $\text{RREF}(A)$ is a matrix in reduced row echelon form. If $\text{RREF}(A) \neq I_n$, then $A$ is not invertible; otherwise, $B = A^{-1}$.

(The rank of a matrix) Let $A \in M_{m \times n}(\mathbb{R})$. Then the row space and column space of $A$ are

$$\text{RowSp}(A) = \text{span}(\text{rows of } A) \subseteq \mathbb{R}^n,$$

$$\text{ColSp}(A) = \text{span}(\text{columns of } A) \subseteq \mathbb{R}^m.$$ 

We have $\dim(\text{RowSp}(A)) = \dim(\text{ColSp}(A)) = \text{rank}(L_A)$; this quantity is called the rank of $A$. Note that elementary row operations do not change the row space. Hence if $B$ is a matrix in row echelon form obtained from $A$ using row operations, then $\text{rank}(A) = \text{rank}(B) = \text{number of leading entries of } B$.

(Change of Coordinates) Let $V$ be an $n$–dimensional vector space and let $\beta' = \{v_1, v_2, \ldots, v_n\}$ and $\beta = \{w_1, w_2, \ldots, w_n\}$ be bases for $V$. The change of coordinate
matrix that changes the \( \beta' \)-coordinates into \( \beta \)-coordinates is
\[
[I_V]_{\beta'}^\beta = \left( \begin{array}{c|c|c|c}
[v_1]_\beta & [v_2]_\beta & \cdots & [v_n]_\beta \\
\end{array} \right) \in \mathcal{M}_{n \times n}(\mathbb{R}).
\]

For every \( v \in V \), we have
\[
[v]_\beta = [I_V]_{\beta'}^\beta [v]_{\beta'}.
\]

If \( T : V \to V \) is linear, then
\[
[T]_{\beta'} = Q^{-1}[T]_\beta Q
\]
where \( Q = [I_V]_{\beta'}^\beta \) and \( Q^{-1} = [I_V]_{\beta}^{\beta'} \).

\section*{3. Determinants (Sections 4.1–4.4)}

- The determinant of \( (A_{11}) \in \mathcal{M}_{1 \times 1}(\mathbb{R}) \) is \( \det((A_{11})) = A_{11} \) and the determinant of \( A \in \mathcal{M}_{n \times n}(\mathbb{R}), n \geq 2, \) is defined recursively as
\[
\det(A) = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \det(\bar{A}_{1j})
\]
where \( \bar{A}_{1j} \) is the \((n-1) \times (n-1)\) matrix obtained from \( A \) by deleting the \( i \)th row and the \( j \)th column.

- One can also compute \( \det(A) \) by taking the expansion above along any row: for any \( 1 \leq r \leq n \), we have
\[
\det(A) = \sum_{j=1}^{n} (-1)^{r+j} A_{rj} \det(\bar{A}_{rj}).
\]

- \textbf{(Multi-linearity of determinants)} Let \( n \in \mathbb{N}, 1 \leq r \leq n, \) and let \( k \in \mathbb{R} \). Then
\[
\det \begin{pmatrix}
  a_1 \\
  \vdots \\
  a_{r-1} \\
  b_r + kc_r \\
  a_{r+1} \\
  \vdots \\
  a_n
\end{pmatrix} = \det \begin{pmatrix}
  a_1 \\
  \vdots \\
  a_{r-1} \\
  b_r \\
  a_{r+1} \\
  \vdots \\
  a_n
\end{pmatrix} + k \det \begin{pmatrix}
  a_1 \\
  \vdots \\
  a_{r-1} \\
  c_r \\
  a_{r+1} \\
  \vdots \\
  a_n
\end{pmatrix}.
\]

- \textbf{(Determinants and row operations)} Let \( A \in \mathcal{M}_{n \times n} \) and let \( R_i \) be the \( i \)th row of \( A \). Then, we have
\[
A \xrightarrow{R_i \leftrightarrow R_j} B \implies \det(B) = -\det(A),
\]
\[
A \xrightarrow{R_i \rightarrow cR_j} B \implies \det(B) = c \det(A),
\]
\[
A \xrightarrow{R_i \rightarrow R_i + cR_j} B \implies \det(B) = \det(A).
\]
• **(Computing \( \det(A) \) using row operations)** Let \( A \in M_{n \times n}(\mathbb{R}) \). Suppose we obtain a matrix \( B \in M_{n \times n}(\mathbb{R}) \) in row echelon form (hence upper triangular) after performing a sequence of row operations on \( A \). Then \( \det(B) = \) the product of the diagonal entries of \( B \) and we can recover \( \det(A) \) using the theorem above.

• **(Properties of determinants)** Let \( A, B \in M_{n \times n}(\mathbb{R}) \). Then
  1. \( A \) is invertible if and only if \( \det(A) \neq 0 \).
  2. \( \det(AB) = \det(A) \det(B) \).
  3. If \( A \) is invertible, then \( \det(A^{-1}) = \frac{1}{\det(A)} \).

4. **Elementary Matrices** *(Sections 3.1)*

• An \( n \times n \) elementary matrix is obtained by performing an elementary row operation on the identity matrix \( I_n \). The elementary matrix is said to be of type 1, 2, or 3 according to whether the elementary row operation performed on \( I_n \) is a type 1, 2, or 3 operation, respectively.

• Let \( A \in M_{m \times n}(\mathbb{R}) \), and suppose that \( B \) is obtained from \( A \) by performing an elementary row operation. Let \( E \) be the elementary matrix obtained from \( I_n \) by performing the same elementary row operation as that which was performed on \( A \) to obtain \( B \). Then \( B = EA \).

• Elementary matrices are invertible and the inverse is the elementary matrix obtained by applying the inverse row operation to \( I_n \).

• Any invertible matrix can be written as a product of elementary matrices.

5. **Diagonalization** *(Sections 5.1–5.2)*

Let \( V \) be an \( n \)-dimensional vector spaces over \( \mathbb{R} \).

• A linear transformation \( T \in \mathcal{L}(V) \) is **diagonalizable** if there exists a basis \( \beta \) for \( V \) for which \( [T]_\beta \) is a diagonal matrix. A matrix \( A \in M_{n \times n}(\mathbb{R}) \) is diagonalizable if \( L_A \in \mathcal{L}(\mathbb{R}^n) \) is diagonalizable.

• A vector \( v \in V \) is an **eigenvector** of \( T \in \mathcal{L}(V) \) if \( v \neq 0 \) and \( T(v) = \lambda v \) for some \( \lambda \in \mathbb{R} \) and \( \lambda \) is the corresponding **eigenvalue**. Eigenvectors and eigenvalues of \( A \in M_{n \times n}(\mathbb{R}) \) are those of \( L_A \).

• Let \( \gamma \) be a basis for \( V \). Then \( T \) is diagonalizable if and only if \( [T]_\gamma \) is diagonalizable.

• The **characteristic polynomial** of \( A \in M_{n \times n}(\mathbb{R}) \) is \( f(t) = \det(A - tI_n) \).

• **(Eigenvalue test)** Let \( A \in M_{n \times n}(\mathbb{R}) \) with characteristic polynomial \( f(t) \). Then, \( \lambda \) is an eigenvalue of \( A \) if and only if \( f(\lambda) = 0 \); i.e., is a root of \( f(\cdot) \).

• Let \( \lambda \) be an eigenvalue of \( A \in M_{n \times n}(\mathbb{R}) \). The **eigenspace** of \( \lambda \) is \( E_\lambda = \mathcal{N}(A - \lambda I_n) \) and the dimension of \( E_\lambda \) is called the **geometric multiplicity** of \( \lambda \). The **algebraic multiplicity** of \( \lambda \) is the highest power of \( t - \lambda \) that divides the characteristic polynomial of \( A \).

• For any eigenvalue \( \lambda \), geometric multiplicity of \( \lambda \) \( \leq \) algebraic multiplicity of \( \lambda \).

• **(Diagonalization criteria)** Let \( A \in M_{n \times n}(\mathbb{R}) \).
– A is diagonalizable if and only if there exists a basis $\beta$ for $\mathbb{R}^n$ consisting of eigenvectors of $A$.

Comment: This criterion is normally used for theoretical purposes.

– A is diagonalizable if and only if there exist an invertible matrix $Q \in M_{n \times n}(\mathbb{R})$ and a diagonal matrix $D$ such that $D = Q^{-1}AQ$. Moreover, we can choose $Q = [I_{\mathbb{R}^n}]_\beta$, where $\beta$ is a basis for $\mathbb{R}^n$ consisting of eigenvectors of $A$ and $\gamma$ is the standard basis for $\mathbb{R}^n$, and the diagonal entries of $D$ are the eigenvalues corresponding to the eigenvectors in $\beta$.

Comment: The expression $D = Q^{-1}AQ$ is useful for computing things like $A^k$, since $A^k = (QDQ^{-1})^k = QD^kQ^{-1}$.

– A is diagonalizable if and only if the characteristic polynomial of $A$ splits over $\mathbb{R}$ and for each eigenvalue $\lambda$ of $A$, the geometric multiplicity of $\lambda$ equals the algebraic multiplicity of $\lambda$.

Comment: This is the diagonalization test that we use in practice.

6. Markov Chains and Matrix Limits (Sections 5.3)

• Let $A, L$ be $n \times n$ real matrices. We say that $\lim_{n \to \infty} A^n = L$ if for all $1 \leq i, j \leq n$, we have $\lim_{n \to \infty} (A^n)_{ij} = (L)_{ij}$.

• If $A$ is diagonalizable over $\mathbb{R}$, i.e., $A = QDQ^{-1}$ for a real diagonal matrix $D$ and an invertible matrix $Q$, then $A^n = QD^nQ^{-1}$, then $\lim_{n \to \infty} A^n$ exists if and only if all diagonal entries of $D$ is in the interval $(-1, 1]$.

• A vector $u \in \mathbb{R}^n$ is a probability vector if all entries of $u$ are non-negative and their sum is one. e.g., $u = (0.5, 0, 0.5)$ is a probability vector, but $v = (-0.5, 1, 0.5)$ is not.

• A matrix $P \in M_{n \times n}(\mathbb{R})$ is called a transition matrix iff all entries of $P$ are nonnegative and each columns of $P$ sums to one.

• A transition matrix $P$ is regular if some power of $P$ contains only positive entries.

• Suppose $P$ is a transition matrix. Then 1 is an eigenvalue for $A$ and any other eigenvalue $\lambda$ has $|\lambda| < 1$.

• Suppose $P$ is a regular transition matrix. Then $\dim(E_1) = 1$, $E_1$ can be spanned by a probability vector $u$ and $\lim_{n \to \infty} A^n = [u | u | \cdots | u]$.

7. Inner Product Spaces (Sections 6.1–6.5)

Let $V$ be a vector spaces over $\mathbb{F}(= \mathbb{R} \text{ or } \mathbb{C})$.

• An inner product on $V$ is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ such that for all $x, y, z \in V$ and $c \in \mathbb{F}$, the following conditions hold:
  (i) $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$
  (ii) $\langle cx, y \rangle = c\langle x, y \rangle$
  (iii) $\langle x, y \rangle = \langle y, x \rangle$
  (iv) $\langle x, x \rangle > 0$ if $x \neq 0$.

• $V$ is an inner product space if $V$ is equipped with an (fixed) inner product.

Let $V$ be an inner product space.
• The following conditions hold for all $x, y, z \in V$ and $c \in F$:
  (i) $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$
  (ii) $\langle x, cy \rangle = \bar{c}\langle x, y \rangle$
  (iii) $\langle x, 0 \rangle = \langle 0, x \rangle = 0$
  (iv) $\langle x, x \rangle > 0$ if and only if $x \neq 0$.
• The norm of $v \in V$ is $\|v\| = \sqrt{\langle v, v \rangle}$.
• For all $x, y \in V$ and $c \in F$, we have
  (i) $\|cx\| = |c|\|x\|$
  (ii) $\|x\| = 0$ if and only if $x = 0$
  (iii) $|\langle x, y \rangle| \leq \|x\|\|y\|$ (Cauchy-Schwarz)
  (iv) $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality)
• A subset $S$ of $V$ is orthogonal if $\langle x, y \rangle = 0$ for all distinct $x, y \in S$ and $S$ is orthonormal if $S$ is orthogonal and $\|x\| = 1$ for all $x \in S$.
• If $S = \{v_1, v_2, \ldots, v_k\}$ is an orthonormal subset of $V$, then for every $y \in \text{span}(S)$, we have
  $$y = \sum_{i=1}^{k} \langle y, v_i \rangle v_i$$
and the scalars $\langle y, v_i \rangle$ are called the Fourier coefficients of $y$ relative to $S$. This can also be extended to the case when $S$ is infinite.
• If $S = \{v_1, v_2, \ldots, v_k\}$ is an orthonormal subset of $V$, then $S$ is linearly independent.