# Contents

1  **A Review of the Basics** .................................................. 3  
   1.1 Matrices and Vectors .................................................. 3  
   1.2 Determinants ............................................................ 6  
   1.3 Systems of Equations .................................................. 7  
   1.4 Linear Independence ................................................... 9  
   1.5 Vector Spaces and Bases .............................................. 10  
   1.6 Orthogonality and Orthonormality ................................. 12  
   1.7 Diagonalizable Matrices .............................................. 13
Chapter 1

A Review of the Basics

Here is a brief summary of the critical definitions and theorems necessary. No proofs are provided for theorems - check any introductory linear algebra text. We assume that all vectors are real, all scalars are real, etc. because that’s all that is necessary for the remainder of the text.

1.1 Matrices and Vectors

Definition 1.1.0.1. A matrix is a rectangular array of numbers. We say it is \( n \times m \) if it has \( n \) rows and \( m \) columns, and that its dimensions are \( n \times m \).

Example 1.1. The following is a \( 3 \times 5 \) matrix:

\[
\begin{bmatrix}
1 & 0 & -1 & 2 & 3 \\
0 & 1.2 & 8 & 0 & 1 \\
-10 & 0 & 1 & 1 & 4
\end{bmatrix}
\]

Definition 1.1.0.2. A vector is an \( n \times 1 \) matrix.

Matrices are generally denoted by upper-case letters \( A, B, \) etc. while vectors are generally denoted by lower-case letters with a bar over them \( \bar{b}, \bar{x}, \) etc.

Example 1.2. We might write:

\[
A = \begin{bmatrix}
1 & 0 & -1 & 2 & 3 \\
0 & 1.2 & 8 & 0 & 1 \\
-10 & 0 & 1 & 1 & 4
\end{bmatrix}
\quad \text{and} \quad
\bar{v} = \begin{bmatrix}
1 \\
-1 \\
0 \\
3
\end{bmatrix}
\]
Definition 1.1.0.3. A matrix is square if $n = m$.

If a matrix is denoted by $A$ then the entry in row $i$ and column $j$ will typically be denoted $a_{ij}$ or $a_{(i,j)}$.

If a vector is denoted by $\vec{b}$ then the entry in row $i$ will typically be denoted by $b_i$.

Definition 1.1.0.4. If $A$ is an $n \times m$ matrix and $\vec{b}$ is an $m \times 1$ vector then we may define $A\vec{b}$ by

$$A\vec{b} = \begin{bmatrix} a_{11}b_1 + a_{12}b_2 + \ldots + a_{1m}b_m \\ a_{21}b_1 + a_{22}b_2 + \ldots + a_{2m}b_m \\ \vdots \\ a_{n1}b_1 + a_{n2}b_2 + \ldots + a_{nm}b_m \end{bmatrix}$$

Example 1.3. We have:

$$\begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -2 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix} = \begin{bmatrix} (1)(5) + (3)(6) + (-1)(7) \\ (0)(5) + (2)(6) + (-2)(7) \end{bmatrix} = \begin{bmatrix} 16 \\ -2 \end{bmatrix}$$

Definition 1.1.0.5. If $A$ is an $n \times m$ matrix and $B$ is an $m \times p$ matrix then if the columns of $B$ are denoted $\vec{b}_1, \ldots, \vec{b}_p$ then we may define $AB$ by

$$AB = [A\vec{b}_1 \ A\vec{b}_2 \ \ldots \ \ A\vec{b}_p]$$

Example 1.4. We have:

$$\begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -2 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 6 & -2 \\ 7 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -2 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -1 \\ 0 & 2 & -2 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 16 & -7 \\ -2 & -6 \end{bmatrix}$$

Notice that $AB$ may be defined when $BA$ is not (because of the dimensions). Even if both $AB$ and $BA$ are defined they may be different sizes. Even when they’re the same size they may be have different entries.

Definition 1.1.0.6. The main diagonal of a matrix $A$ is the entries $a_{11}, a_{22}, \ldots, a_{nn}$.

Definition 1.1.0.7. The identity matrix $I_n$ is the $n \times n$ matrix with 1s on the
main diagonal and 0s elsewhere. When the size is clear or implied we simply write $I$.

Definition 1.1.0.8. If $A$ is a square matrix then the *transpose* of $A$, denoted $A^T$, is the matrix whose $(i, j)$ entry equals $a_{ji}$. That is, it is obtained by reflecting $A$ in the main diagonal.

Definition 1.1.0.9. A square matrix $A$ is *symmetric* if $A^T = A$.

Definition 1.1.0.10. If $A$ is an $n \times n$ square matrix then $A$ is *invertible* if there is another $n \times n$ matrix, denoted $A^{-1}$, such that $AA^{-1} = I$ and $A^{-1}A = I$.

Example 1.5. Observe that
\[
\begin{bmatrix}
  5 & 7 \\
  2 & 3
\end{bmatrix}
\begin{bmatrix}
  3 & 7 \\
-2 & 5
\end{bmatrix} =
\begin{bmatrix}
  1 & 0 \\
  0 & 1
\end{bmatrix}
\]
so that
\[
\begin{bmatrix}
  5 & 7 \\
  2 & 3
\end{bmatrix}^{-1} = \begin{bmatrix}
  3 & 7 \\
-2 & 5
\end{bmatrix} \text{ and } \begin{bmatrix}
  3 & 7 \\
-2 & 5
\end{bmatrix}^{-1} = \begin{bmatrix}
  5 & 7 \\
-2 & 5
\end{bmatrix}
\]
and both matrices are invertible.

Most matrix inverses are not nearly this pretty, nor are inverses easy to find.
The exception is the $2 \times 2$ case.

Theorem 1.1.0.1. If
\[
A = \begin{bmatrix}
  a & b \\
  c & d
\end{bmatrix}
\]
then $A$ is invertible iff $ad - bc \neq 0$ in which case
\[
\begin{bmatrix}
  a & b \\
  c & d
\end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix}
  d & -b \\
- c & a
\end{bmatrix}
\]

Proof. Omitted. \qed

Not all matrices are invertible but “most” are, where “most” means something rigorous and meaningful.

Definition 1.1.0.11. If $\vec{v}, \vec{w} \in \mathbb{R}^n$ then we define the *dot product*
\[
\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w} = v_1 w_1 + ... + v_n w_n
\]
Definition 1.1.0.12. If \( \vec{v} \in \mathbb{R}^n \) then the magnitude or norm or length of \( \vec{v} \) is defined by
\[
||\vec{v}|| = \sqrt{v_1^2 + ... + v_n^2}
\]

Definition 1.1.0.13. A diagonal matrix is a square matrix \( A \) such that \( a_{ij} = 0 \) for \( i \neq j \).

1.2 Determinants

Definition 1.2.0.1. If \( A \) is an \( n \times m \) matrix then the matrix minor denoted by \( A_{ij} \) is the \((n-1) \times (m-1)\) matrix obtained by removing row \( i \) and column \( j \) from \( A \).

Definition 1.2.0.2. The determinant of an square matrix \( A \) denoted \( \det(A) \) or just \( \det A \) is defined recursively as follows:
- If \( A \) is \( 2 \times 2 \) then
  \[
  \det(A) = a_{11}a_{22} - a_{12}a_{21}
  \]
- If \( A \) is larger than \( 2 \times 2 \) then
  \[
  \det(A) = +a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13}) - \ldots \pm a_{1n}\det(A_{1n})
  \]

Example 1.6. For a \( 2 \times 2 \)
\[
\det \begin{bmatrix} 5 & 3 \\ -2 & 6 \end{bmatrix} = (5)(6) - (3)(-2) = 36
\]

Example 1.7. For a \( 3 \times 3 \)
\[
\det \begin{bmatrix} 1 & 2 & -3 \\ 0 & 5 & 1 \\ -2 & 4 & 7 \end{bmatrix} = +1 \det \begin{bmatrix} 5 & 1 \\ 4 & 7 \end{bmatrix} - 2 \det \begin{bmatrix} 0 & 1 \\ -2 & 7 \end{bmatrix} + (-3) \det \begin{bmatrix} 0 & 5 \\ -2 & 4 \end{bmatrix}
\]
\[
= +1(31) - 2(2) + (-3)(10)
\]
\[
= -3
\]

Theorem 1.2.0.1. A square matrix \( A \) is invertible iff \( \det(A) \neq 0 \).

Proof. Omitted.

Example 1.7 Revisited. The matrix:
\[
\begin{bmatrix} 1 & 2 & -3 \\ 0 & 5 & 1 \\ -2 & 4 & 7 \end{bmatrix}
\]
1.3. SYSTEMS OF EQUATIONS

has determinant $-3 \neq 0$ and hence is invertible.

Mathematically speaking since the chances of having $\text{det}(A) = 0$ are very small it is in this sense that we can say that “most” matrices are invertible since “most” matrices have nonzero determinant.

1.3 Systems of Equations

Definition 1.3.0.1. A linear system of $m$ equations in the variables $x_1, ..., x_n$ given by

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

may be represented by the matrix equation

\[ A\vec{x} = \vec{b} \]

Example 1.8. The system of equations

\[
\begin{align*}
  2x_1 + 3x_2 - 1x_3 &= 7 \\
  -1x_1 + 7x_2 + 4x_3 &= -2
\end{align*}
\]

may be rewritten as

\[
\begin{bmatrix}
  2 & 3 & -1 \\
  -1 & 7 & 4
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2 \\
  x_3
\end{bmatrix}
= 
\begin{bmatrix}
  7 \\
  -2
\end{bmatrix}
\]

Theorem 1.3.0.1. The matrix equation $A\vec{x} = \vec{b}$ has either no solutions, one solution, or infinitely many solutions. There is one solution iff $A$ is invertible and in that case the solution is given by $\vec{x} = A^{-1}\vec{b}$.

Proof. Omitted.

Example 1.9. The matrix equation

\[
\begin{bmatrix}
  5 & 7 \\
  2 & 3
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2
\end{bmatrix}
= 
\begin{bmatrix}
  2 \\
  -1
\end{bmatrix}
\]

has exactly one solution because the matrix is invertible as we saw earlier. The solution is given by
\[
\bar{x} = \begin{bmatrix}
5 & 7 \\
2 & 3 \\
\end{bmatrix}^{-1} \begin{bmatrix}
2 \\
-1 \\
\end{bmatrix} = \begin{bmatrix}
3 & -7 \\
-2 & 5 \\
\end{bmatrix} \begin{bmatrix}
2 \\
-1 \\
\end{bmatrix} = \begin{bmatrix}
13 \\
-9 \\
\end{bmatrix}
\]

**Definition 1.3.0.2.** If \( A \) is an \( n \times n \) matrix then \( \lambda \) is an *eigenvalue* for \( A \) if there is a nonzero vector \( \bar{v} \) such that \( A\bar{v} = \lambda\bar{v} \). In this case \( \bar{v} \) is the corresponding *eigenvector*. The set of all eigenvectors for a given eigenvalue is called the *eigenspace* of that eigenvalue.

Notice that any multiple of an eigenvector is also an eigenvector.

**Example 1.10.** The matrix

\[
A = \begin{bmatrix}
4 & 7 \\
1 & -2 \\
\end{bmatrix}
\]

has two eigenvalues. One is \( \lambda_1 = 5 \) with eigenvector \( \begin{bmatrix} 7 \\ 1 \end{bmatrix} \) because:

\[
\begin{bmatrix}
4 & 7 \\
1 & -2 \\
\end{bmatrix} \begin{bmatrix} 7 \\ 1 \end{bmatrix} = 5 \begin{bmatrix} 7 \\ 1 \end{bmatrix}
\]

The other eigenvalue is \( \lambda_2 = -3 \) with eigenvector \( \begin{bmatrix} -1 \\ -1 \end{bmatrix} \). This can be easily checked.

**Definition 1.3.0.3.** If \( A \) is an \( n \times n \) matrix then the *characteristic polynomial* of \( A \) denoted \( \text{char}(A) \), is defined by

\[
\text{char}(A) = \det(\lambda I - A)
\]

Note: Some authors define the characteristic polynomial to be:

\[
\text{char}(A) = \det(A - \lambda I)
\]

These two definitions differ by a factor of \((-1)^n\).

**Theorem 1.3.0.2.** The eigenvalues of a matrix \( A \) are the roots of the characteristic polynomial.

**Proof.** Omitted. \( \square \)

Since the characteristic polynomial has degree \( n \) this tells us that an \( n \times n \) matrix has \( n \) eigenvalues, counting multiplicity.
Example 1.11. If we have:

\[
A = \begin{bmatrix}
1 & 2 & -1 \\
1 & 2 & 3 \\
0 & 4 & 3
\end{bmatrix}
\]

then

\[
\text{char}(A) = \det(\lambda I - A)
= \det \begin{bmatrix}
\lambda - 1 & -2 & 1 \\
-1 & \lambda - 2 & -3 \\
0 & -4 & \lambda - 3
\end{bmatrix}
= \lambda^3 - 6\lambda^2 - 3\lambda + 16
\]

The eigenvalues of the matrix are roots of this, \(\lambda_1 \approx 6.0593, \lambda_2 \approx -1.6549\) and \(\lambda_3 \approx 1.5956\).

1.4 Linear Independence

**Definition 1.4.0.1.** A set of vectors \(\{\vec{v}_1, ..., \vec{v}_n\}\) is linearly independent if \(a_1\vec{v}_1 + \ldots + a_n\vec{v}_n = \vec{0}\) implies \(a_1 = \ldots = a_n = 0\).

As a consequence of this a set of just two vectors is linearly independent iff no\(ther\) is a multiple of the other.

**Example 1.12.** The set of vectors

\[
\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix} \right\}
\]

is a linearly independent set.

A classic way of thinking of linear independence is that it is impossible to write any one of the vectors as a linear combination of the other vectors.

One of the consequences of having a linearly independent set is that if some vector \(\vec{v}\) is a linear combination of that set then only that specific linear combination works.

**Definition 1.4.0.2.** If a set of vectors is not linearly independent then it is linearly dependent.
A classic way of thinking of linear dependence is that one vector may be written as a linear combination of the others.

**Example 1.13.** The set of vectors

\[
\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}
\]

is linearly dependent. Observe that

\[
\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}
\]

### 1.5 Vector Spaces and Bases

**Definition 1.5.0.1.** A *vector space* is a nonempty set $V$ of vectors such that the following properties hold:

1. $\vec{0} \in V$.
2. If $\vec{u}, \vec{v} \in V$ then $\vec{u} + \vec{v} \in V$.
3. $\vec{u} \in V$ then $-\vec{u} \in V$.
4. If $\vec{u} \in V$ and $c \in \mathbb{R}$ then $c\vec{u} \in V$.

Note that 1 actually follows from 2 and 3 together but it’s worth listing on its own anyway.

**Definition 1.5.0.2.** Given a set of vectors $S = \{\vec{v}_1, \ldots, \vec{v}_n\}$ then the *span* of $S$ denoted $\text{span}(S)$ is the set of all linear combinations of vectors in $S$. More rigorously

\[
\text{span}(S) = \left\{ a_1\vec{v}_1 + \ldots + a_n\vec{v}_n \mid a_1, \ldots, a_n \in \mathbb{R} \right\}
\]

**Example 1.14.** If

\[
S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix} \right\}
\]

Then

\[
\text{span}(S) = \left\{ a_1 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix} \mid a_1, a_2 \in \mathbb{R} \right\}
\]
1.5. VECTOR SPACES AND BASES

**Theorem 1.5.0.1.** Given a set of vectors \( S = \{ \vec{v}_1, ..., \vec{v}_n \} \), the span of \( S \) is a vector space.

*Proof. Omitted.* □

**Definition 1.5.0.3.** If \( V \) is a vector space then a *basis* for \( V \) is a linearly independent set \( B \) of vectors such that \( V = \text{span}(B) \).

**Example 1.15.** The set

\[
B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \right\}
\]

is a basis for \( \mathbb{R}^3 \).

Essentially a basis for a vector space \( V \) is a set of building blocks \( B \) such that each vector in \( V \) can be written uniquely as a linear combination of vectors in \( B \).

**Definition 1.5.0.4.** If \( A \) is an \( m \times n \) matrix then the *column space of \( A \)* denoted \( \text{col}(A) \) is the span of the columns of \( A \).

**Theorem 1.5.0.2.** Every vector space has a basis and the number of vectors in a basis of a vector space is independent of the choice of basis. That is, every basis has exactly the same number of vectors as every other basis.

*Proof. Omitted.* □

**Definition 1.5.0.5.** For a vector space \( V \) the *dimension of \( V \)* denoted \( \text{dim}(V) \), is defined as the number of vectors in a basis of \( V \).

**Theorem 1.5.0.3.** If \( A \) is an \( n \times n \) matrix and \( \lambda \) is an eigenvalue with eigenspace \( V \) then the dimension of the eigenspace is less than or equal to the multiplicity of \( \lambda \) as a root of the characteristic polynomial.

*Proof. Omitted.* □

**Example 1.16.** The matrix

\[
A = \begin{bmatrix} 2 & 0 & 0 \\ 10 & 12 & -30 \\ 5 & 5 & -13 \end{bmatrix}
\]
Has characteristic polynomial
\[ \lambda^3 - \lambda^2 - 8\lambda + 12 = (\lambda - 2)^2(\lambda + 3) \]
Consequently the eigenspace for \( \lambda_1 = 2 \) has dimension either 1 or 2. In this case it’s 2 but that’s a bit more work.

1.6 Orthogonality and Orthonormality

**Definition 1.6.0.1.** Two vectors are *orthogonal* if their dot product equals zero. A set of vectors \( \{\vec{v}_1, \ldots, \vec{v}_n\} \) is orthogonal if \( \vec{v}_i \cdot \vec{v}_j = 0 \) for all \( i \neq j \).

**Example 1.17.** The set of vectors
\[
\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} -5 \\ 2 \\ 1 \end{bmatrix} \right\}
\]
is an orthogonal set of vectors.

**Definition 1.6.0.2.** A set of vectors \( \{\vec{v}_1, \ldots, \vec{v}_n\} \) is *orthonormal* if \( \vec{v}_i \cdot \vec{v}_j = 0 \) for all \( i \neq j \) and \( \|\vec{v}_i\| = 1 \) for all \( i \).

**Example 1.18.** The set of vectors in the previous example is orthonormal if each vector is divided by its magnitude:
\[
\left\{ \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \begin{bmatrix} 0 \\ 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}, \begin{bmatrix} -5/\sqrt{30} \\ 2/\sqrt{30} \\ 1/\sqrt{30} \end{bmatrix} \right\}
\]
is an orthonormal set of vectors.

**Definition 1.6.0.3.** A square matrix is *orthogonal* if the column vectors form an orthonormal set.

**Example 1.19.** The matrix
\[
A = \begin{bmatrix}
1/\sqrt{6} & 0 & -5/\sqrt{30} \\
2/\sqrt{6} & 1/\sqrt{5} & 2/\sqrt{30} \\
1/\sqrt{6} & -2/\sqrt{5} & 1/\sqrt{30}
\end{bmatrix}
\]
is orthogonal.

**Theorem 1.6.0.1.** A square matrix \( A \) is orthogonal iff \( A^T A = A A^T = I \). That is, if \( A^T = A^{-1} \).
1.7. DIAGONALIZABLE MATRICES

Proof. Omitted. □

Orthogonal matrices are great simply because $A^{-1} = A^T$ and so the inverse is really convenient.

1.7 Diagonalizable Matrices

Definition 1.7.0.1. An $n \times n$ matrix $A$ is diagonalizable if there exists an $n \times n$ invertible matrix $P$ and an $n \times n$ diagonal matrix $D$ such that

$$A = PDP^{-1}$$

Theorem 1.7.0.1. A matrix $A$ is diagonalizable iff the dimension of each eigenspace equals the multiplicity of the corresponding eigenvalue in the characteristic polynomial.

Proof. Omitted. □

Theorem 1.7.0.2. If $A$ is diagonalizable then the invertible matrix $P$ is formed using the eigenvectors of $A$ and the diagonal matrix $D$ is formed using the eigenvalues of $A$. The eigenvector in column $i$ corresponds to the eigenvalue in column $i$.

Proof. Omitted. □

Example 1.20. If

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 10 & 12 & -30 \\ 5 & 5 & -13 \end{bmatrix}$$

Then

$$A = PDP^{-1}$$

where

$$P = \begin{bmatrix} 0 & 0 & 0.4016 \\ -0.8944 & -0.9487 & -0.9006 \\ -0.4472 & -0.3162 & -0.1663 \end{bmatrix}$$

and

$$D = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$
In this case the first column of $P$ is an eigenvector corresponding to the eigenvalue $\lambda_1 = -3$ and the second and third columns of $P$ are eigenvectors corresponding to the eigenvalue $\lambda_2 = 2$ which has multiplicity 2 and for which the dimension of the eigenspace is also 2.

**Definition 1.7.0.2.** An $n \times n$ matrix $A$ is **orthogonally diagonalizable** if there exists an $n \times n$ orthogonal matrix $Q$ and an $n \times n$ diagonal matrix $D$ such that

$$A = QDQ^T$$

**Theorem 1.7.0.3.** A matrix $A$ is orthogonally diagonalizable iff it is symmetric.

*Proof.* Omitted. \(\square\)

**Example 1.21.** The matrix

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ -1 & 3 & 4 \end{bmatrix}$$

is symmetric hence orthogonally diagonalizable.