Linear Algebra and Matrices

Before embarking on a study of systems of differential equations we will first review, very quickly, some fundamental objects and operations in linear algebra.

1. Matrices

**Definition 1.1.** An \( n \times m \) matrix ("n by m matrix") is an arrangement of \( nm \) objects (usually numbers) into a rectangular array with \( n \) rows and \( m \) columns. We’ll typically denote the entry in the \( i \)th row and \( j \)th column of a matrix \( A \) as \( a_{ij} \) (and similarly, \( b_{ij} \) for the \( (ij) \)th entry of a matrix \( B \)). Thus, for example,

\[
A = \begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1m} \\
a_{21} & a_{22} & \ddots & a_{2m} \\
& \ddots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nm}
\end{pmatrix}
\]

An \( n \times 1 \) matrix

\[
v = \begin{pmatrix}
v_1 \\
v_2 \\
\vdots \\
v_n
\end{pmatrix}
\]

is called a **column vector** and a \( 1 \times m \) matrix

\[
v = [v_1, v_2, \ldots, v_m]
\]

is called a **row vector**.

1.1. Matrix Operations.

**Definition 1.2.** The **sum** of two \( n \times m \) matrices \( A \) and \( B \) is the \( n \times m \) matrix \( A + B \) with entries

\[
(A + B)_{ij} = a_{ij} + b_{ij}
\]

**Definition 1.3.** The **scalar product** of an \( n \times m \) matrix \( A \) with a number \( \lambda \) is the \( n \times m \) matrix \( \lambda A \) with entries

\[
(\lambda A)_{ij} = \lambda a_{ij}
\]

**Definition 1.4.** The **difference** of two \( n \times m \) matrices \( A \) and \( B \) is the \( n \times m \) matrix \( A - B \) with entries

\[
(A - B)_{ij} = a_{ij} - b_{ij} = (A + (-1)B)_{ij}
\]

**Definition 1.5.** The **transpose** of an \( n \times m \) matrix \( A \) is the \( m \times n \) matrix \( A^t \) with entries

\[
(A^t)_{ij} = A_{ji}
\]

(Note the \( i \)th row of \( A^t \) is simply the \( j \)th column of \( A \) written horizontally.)

**Definition 1.6.** An \( n \times n \) matrix \( A \) is said to be **symmetric** if

\[
A = A^t
\]
Note that the transpose of a row-vector is a column vector and the transpose of a column vector is a row vector.

**Definition 1.7.** The dot product of an n-dimensional row vector \([r_1, \ldots, r_n]\) with a n-dimensional column vector \[
\begin{bmatrix}
c_1 \\
\vdots \\
c_n
\end{bmatrix}
\]
is the number
\\
[r_1, \ldots, r_n] \cdot \begin{bmatrix} c_1 \\
\vdots \\
c_n \end{bmatrix} = r_1c_1 + \cdots + r(nc_n)
\\
(This coincides with the usual dot products of vectors when we think of vectors as ordered lists of numbers rather than special kinds of matrices.)

**Definition 1.8.** The matrix product of an \(n \times m\) matrix \(A\) with a \(m \times q\) matrix \(B\) is the \(n \times q\) matrix with entries
\\
(AB)_{ij} = \sum_{k=1}^{m} a_{ik}b_{kj}
\\
Note that the \((ij)\)th entry of the matrix product \(AB\) is the dot product of the \(i\)th row of \(A\) with the \(j\)th column of \(B\).

### 2. Complex Numbers

In what follows it will be sometime necessary to consider vectors and matrices with entries that are complex numbers. We recall here some basic facts about complex numbers.

**Definition 1.9.** A complex numbers are pairs \((x, y)\) of real numbers which satisfying the following addition and multiplication rules
\\
\[
(x, y) + (x', y') = (x + x', y + y')
\]
\\
\[
(x, y) \cdot (x', y') = (xx' - yy', xy' + yx')
\]

Usually though we write a complex number as
\\
\[z = x + iy\]

and then add and multiply complex numbers by employing the usual rules of arithmetic and the relation \(i^2 = -1\). Thus,
\\
\[\begin{align*}
(x + iy) (x' + iy') &= xx' + x(iy') + iy(x') + (iy)(iy') \\
&= xx' - yy' + i(xy' + yx')
\end{align*}\]

When we write
\\
\[z = x + iy\]

we say that \(x\) is the real part of \(z\) and \(y\) is the imaginary part of \(z\).

The **complex conjugate** of \(z = x + iy\) is the complex number \(\bar{z}\) obtained by switching the sign of it imaginary part
\\
\[\bar{z} = x - iy\]

We have
\\
\[
Re (z) = x = \frac{z + \bar{z}}{2}
\]
\\
\[
Im (z) = y = \frac{z - \bar{z}}{2i}
\]
\\
\[Re ()\]
We will also sometimes consider functions of a complex variable. Here we only state the very fundamental Euler Formula
\[ e^{x+iy} = e^x (\cos (y) + i \sin (y)) \]
We note that
\[ \cos (y) = Re (e^{iy}) \]
\[ \sin (y) = Im (e^{iy}) \]

**Definition 1.10.** A **complex matrix** is a matrix whose entries are complex numbers. Similarly, a **complex vector** is a vector whose components are complex numbers. The **hermitian adjoint** of a complex matrix \( A \) is the matrix \( A^\dagger \) with entries
\[ (A^\dagger)_{ij} = \overline{a_{ji}} \]
Equivalently,
\[ A^\dagger = (\overline{A})^t = (\overline{A^t}) \]

**Example 1.11.**
\[ A = \begin{pmatrix} 1 + i & 1 - 2i \\ 3 + i & 3 \end{pmatrix} \Rightarrow A^\dagger = \begin{pmatrix} 1 - i & 3 - i \\ 1 + 2i & 3 \end{pmatrix} \]

**Definition 1.12.** An \( n \times n \) matrix is **hermitian** (or **self-adjoint**) if
\[ A = A^\dagger \]

**Definition 1.13.** Let \( x \) and \( y \) be two complex (column) vectors. The (hermitian) inner product \( \langle x, y \rangle \) of \( x \) and \( y \) is the complex number
\[ x^t y = [x_1, \cdots, x_n] \begin{bmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{bmatrix} = x_1 \overline{y_1} + \cdots + x_n \overline{y_n} \]

**Remark 1.14.** If \( A \) is a real hermitian matrix (that is, all of its entries are real numbers), then
\[ A^\dagger = (\overline{A})^t = (\overline{A^t}) = A^t \]
so a real hermitian matrix is just a real symmetric matrix. If \( x \) and \( y \) are two real vectors then the hermitian inner product \( \langle x, y \rangle \) coincides with the usual dot product of real vectors.

### 3. Determinants

**Definition 1.15.** The \( (ij)^{th} \) **minor** of an \( n \times n \) matrix \( A \) is the \( (n-1) \times (n-1) \) matrix \( M_{ij} \) obtained by deleting the \( i^{th} \) row and \( j^{th} \) column from \( A \).

**Definition 1.16.** The determinant of an \( n \times n \) matrix \( A \) is the number \( \det A \) determined by the following recursive algorithm:

- If \( A \) is a \( 1 \times 1 \) matrix \([a_{11}]\), then \( \det A = a_{11} \)

- If \( A \) is an \( n \times n \) matrix then
\[ \det A : = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det M_{ij} \quad (i = \text{any fixed row index}) \]
\[ : = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det M_{ij} \quad (j = \text{any fixed column index}) \]
The following formulas for the determinants of $2 \times 2$ and $3 \times 3$ matrices follow easily from the above definition:

\[
\begin{vmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{vmatrix} = (-1)^{1+1} a_{11} \det(M_{11}) + (-1)^{1+2} a_{12} \det(M_{12})
= a_{11} \det[a_{22}] - a_{12} \det[a_{21}]
= a_{11}a_{22} - a_{12}a_{21}
\]

\[
\begin{vmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{vmatrix} = (-1)^{1+1} a_{11} \det\begin{vmatrix}
a_{22} & a_{23} \\
a_{32} & a_{33}
\end{vmatrix} + (-1)^{1+2} a_{12} \det\begin{vmatrix}
a_{21} & a_{23} \\
a_{31} & a_{33}
\end{vmatrix}
+ (-1)^{1+3} a_{13} \det\begin{vmatrix}
a_{21} & a_{22} \\
a_{31} & a_{32}
\end{vmatrix}
= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})
\]

4. Inverses

**Definition 1.17.** The *inverse* of an $n \times n$ matrix $A$ is the $n \times n$ matrix $A^{-1}$ such that

\[A^{-1}A = I = AA^{-1}\]

(N.B. matrix inverses do not always exist.)

There are two basic ways of computing matrix inverses:

- **Row reduction:**

  \[
  \begin{bmatrix} A & I \end{bmatrix} \xrightarrow{\text{row reduction}} \begin{bmatrix} I & A^{-1} \end{bmatrix}
  \]

- **Cofactor Method:**

  \[A^{-1} = \frac{1}{\det A} C'\]

  where $C$ is the cofactor matrix of $A$ whose entries are $\pm 1$ times the determinants of the $(n - 1) \times (n - 1)$ minors of $A$.

  \[(C)_{ij} = (-1)^{i+j} \det(M_{ij})\]

5. Eigenvalues and Eigenvectors

**Definition 1.18.** Let $A$ be an $n \times n$ matrix. If there exists a number $\lambda$ and an $n$-dimensional (column) vector $v$ such that

\[Av = \lambda v\]

then $v$ is said to be an eigenvector of $A$ and $\lambda$ is said to be the eigenvalue of $A$ corresponding to the eigenvector $v$.

The following algorithm determines all the eigenvectors and eigenvalues of an $n \times n$ matrix $A$.

- **Set** $p_A = \det(A - \lambda I)$. Here $\lambda$ is regarded as a variable, and $\lambda I$ is the matrix obtained by scalar multiplying the $n \times n$ identity matrix by $\lambda$. $A - \lambda I$ is thus the matrix obtained by subtracting $\lambda$ from each of the diagonal entries of $A$. Upon computation, $\det(A - \lambda I)$ will yield a polynomial of degree $n$ in $\lambda$. The solutions of $p_A(\lambda) = 0$ will be the eigenvalues of $A$. 

• For each root $\lambda = r$ of $p_A(\lambda) = 0$, find the general solution of 

$$(A - rI)x = 0$$

and expand that solution in terms of the free parameters of the solution. The eigenvectors corresponding to the eigenvalue $r$ of $A$ will be the constant vectors in that expansion.

**Example 1.19.** Find the eigenvalues and eigenvectors of $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.

• We have

$$p_A(\lambda) := \det \begin{bmatrix} 1 - \lambda & 2 \\ 2 & 1 - \lambda \end{bmatrix} = (1 - \lambda)^2 - 4 = \lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1)$$

Since $\lambda = 3, -1$ are the solutions of $p_A(\lambda) = 0$, these are the eigenvalues of $A$.

• Now we look for the eigenvectors corresponding to $\lambda = 3$.

$$\begin{bmatrix} 1 - 3 & 2 \\ 2 & 1 - 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The reduced row echelon form of the coefficient matrix is

$$\begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \Rightarrow \begin{cases} x_1 - x_2 = 0 \\ 0 = 0 \end{cases} \Rightarrow x_1 = x_2$$

so

$$v_{\lambda=3} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

• Similarly, we look for the eigenvectors corresponding to $\lambda = -1$:

$$\begin{bmatrix} 1 - (-1) & 2 \\ 2 & 1 - (-1) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x_1 = -x_2$$

$$\Rightarrow x = \begin{bmatrix} -x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

and so

$$v_{\lambda=-1} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

**Example 1.20.** Suppose

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

Find the eigenvectors and eigenvalues of $A$.

• The characteristic polynomial is

$$p_A(\lambda) = \det \begin{bmatrix} 1 - \lambda & 1 \\ -1 & 1 - \lambda \end{bmatrix} = (1 - \lambda)^2 + 1 = \lambda^2 - 2\lambda + 2$$

To solve $p_A(\lambda) = 0$, we apply the quadratic formula

$$ax^2 + bx + c = 0 \implies x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

and find

$$p_A(\lambda) = 0 \implies \lambda = \frac{2 \pm \sqrt{4 - 8}}{2} = \frac{2 \pm \sqrt{-4}}{2} = \frac{2 \pm \sqrt{-i}4}{2} = \frac{2 \pm 2i}{2} = 1 \pm i$$

Thus we have two complex eigenvalues.
6. Diagonalization of Matrices

Recall that a diagonal matrix is a square $n \times n$ matrix with non-zero entries only along the diagonal from the under left to the lower right (the main diagonal).

Diagonal matrices are particularly convenient for eigenvalue problems since the eigenvalues of a diagonal matrix

$$ A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{bmatrix} $$

coincide with the diagonal entries $\{a_{ii}\}$ and the eigenvector corresponding the eigenvalue $a_{ii}$ is just the $i^{th}$ coordinate vector. That is, if $A$ is of the above form, we always have

$$ Ae_i = a_{ii}e_i $$

**Definition 1.21.** An $n \times n$ matrix $A$ is **diagonalizable** if there is an invertible $n \times n$ matrix $C$ such that $C^{-1}AC$ is a diagonal matrix. The matrix $C$ is said to **diagonalize** $A$.

**Lemma 1.22.** Let $A$ be a real (or complex) $n \times n$ matrix, let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be a set of $n$ real (respectively, complex) scalars, and let $v_1, v_2, \ldots, v_n$ be a set of $n$ vectors in $\mathbb{R}^n$ (respectively, $\mathbb{C}^n$). Let $C$ be the $n \times n$ matrix formed by using $v_j$ for $j^{th}$ column vector, and let $D$ be the $n \times n$ diagonal matrix whose diagonal entries are $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then

$$ AC = CD $$

if and only if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of $A$ and each $v_j$ is an eigenvector of $A$ corresponding the eigenvalue $\lambda_j$.

Now suppose $AC = CD$, and the matrix $C$ is invertible. Then we can write

$$ D = C^{-1}AC. $$

And so we can think of the matrix $C$ as converting $A$ into a diagonal matrix.

**Theorem 1.23.** An $n \times n$ matrix $A$ is diagonalizable if and only if it has $n$ linearly independent eigenvectors.

**Example 1.24.** Find the matrix that diagonalizes

$$ A = \begin{bmatrix} 2 & 6 \\ 0 & -1 \end{bmatrix} $$
First we’ll find the eigenvalues and eigenvectors of $A$.

$$0 = \det (A - \lambda I) = \det \begin{bmatrix} 2 - \lambda & 6 \\ 0 & -1 - \lambda \end{bmatrix} = (2 - \lambda)(-1 - \lambda) \implies \lambda = 2, -1$$

The eigenvectors corresponding to the eigenvalue $\lambda = 2$ are solutions of $(A - (2)I)x = 0$ or

$$\begin{bmatrix} 0 & 6 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies 6x_2 = 0, -3x_2 = 0 \implies x_2 = 0 \implies x = r \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

The eigenvectors corresponding to the eigenvalue $\lambda = -1$ are solutions of $(A - (-1)I)x = 0$ or

$$\begin{bmatrix} 3 & 6 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies 3x_1 + 6x_2 = 0, 0 = 0 \implies x_1 = -2x_2 \implies x = r \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

So the vectors $v_1 = [1, 0]$ and $v_2 = [-2, 1]$ will be eigenvectors of $A$. We now arrange these two vectors as the column vectors of the matrix $C$.

$$C = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

In order to compute the diagonalization of $A$ we also need $C^{-1}$. This we compute using the technique of Section 1.5:

$$\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \implies C^{-1} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Finally,

$$D = C^{-1}AC = C^{-1}(AC) = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 6 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$