Very short review of linear algebra.

1 Vector space

A vector space $\mathcal{V}$ is a set of objects - vectors - equipped with addition and scalar multiplication. That is, for vectors $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ and a scalar $a$, we have that $\mathbf{u} + \mathbf{v} \in \mathcal{V}$ and $a\mathbf{v} \in \mathcal{V}$ are also vectors. Thus, the vector space is closed under addition and scalar multiplication.

Further, a distribution law holds:

$$a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$$

Here, we consider only real vector spaces where $a \in \mathbb{R}$.

1.1 Basis

A basis of a vector space $\mathcal{V}$ is a collection of vectors $\mathbf{b}_1, \ldots, \mathbf{b}_N \in \mathcal{V}$ such that

1. the basis vectors are linearly independent, i.e.

$$\sum_{i=1}^{N} a_i \mathbf{b}_i = 0 \implies a_i = 0 \ \forall i = 1, \ldots, N$$

2. every vector $\mathbf{v} \in \mathcal{V}$ can be written as

$$\mathbf{v} = \sum_i c_i \mathbf{b}_i$$

with a unique set of scalars $c_1, \ldots, c_N \in \mathbb{R}$.

Example 1 The vectors $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ are linearly independent, while $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ -2 \end{pmatrix}$ are not.

$N$ is called the dimension of the vector space. Here, we only consider finite dimensional vector spaces.

Often, when a basis is chosen, the vectors are identified with their representation in terms of the basis vectors. That is, we write $\mathbf{v} = \begin{pmatrix} c_1 \\ \vdots \\ c_N \end{pmatrix}$. The basis vectors
themselves are then represented as the standard coordinate vectors $e_1, \ldots, e_N$, i.e.

$$b_1 = e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$
$$b_2 = e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$
$$\vdots$$
$$b_N = e_N = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

### 2 Matrices

A matrix $A$ represents a linear map from an $N$-dimensional vector space $V$ into an $M$-dimensional vector space $U$. It holds that

$$A(u + v) = Au + Av \text{ and } A(av) = aAv$$

Again, assuming a standard chose of basis for $V$ and $U$ a matrix is written as an $M \times N$ dimensional array of real numbers, i.e. $A \in \mathbb{R}^{M \times N}$.

Matrices can be added and multiplied:

- Consider two matrices $A \in \mathbb{R}^{M \times N}$ and $B \in \mathbb{R}^{M \times N}$. Then, $A + B$ is an $M \times N$ matrix with entries:

$$(A + B)_{ij} = A_{ij} + B_{ij}$$

- Consider two matrices $A \in \mathbb{R}^{M \times N}$ and $B \in \mathbb{R}^{N \times D}$. Then, the matrix product $A \cdot B$ is an $M \times D$ matrix with entries:

$$(AB)_{ij} = \sum_{k=1}^{N} A_{ik}B_{kj}$$

For the matrix product to exists, the inner dimensions of the matrices $A, B$ have to match. The matrix product has the following properties:

- **Associativity:** $(AB)C = A(BC)$
• **Non-commutative:** In general, $AB \neq BA$ even when the dimensions match, e.g. for square matrices.

Then, the action of the matrix $A$ on some $N$-dimensional vector $v$ can be computed as a special case of matrix multiplication:

$$
Av = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\
\vdots & \ddots & \vdots \\
a_{M1} & \cdots & a_{MN} \end{pmatrix} \begin{pmatrix} c_1 \\
\vdots \\
c_N \end{pmatrix} = \begin{pmatrix} d_1 \\
\vdots \\
d_M \end{pmatrix}
$$

where $d_i = \sum_{j=1}^{N} a_{ij} c_j$

Note that the (column-)vector $v$ is considered as an $N \times 1$ matrix.

**Example 2**

$$
\begin{pmatrix} 1 & 2 & 3 \\
4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 3 \\
2 \\
1 \end{pmatrix} = \begin{pmatrix} 10 \\
28 \end{pmatrix}
$$

**2.0.1 Inverse**

The inverse of a square matrix $A \in \mathbb{R}^{N \times N}$ is the matrix $A^{-1}$ with the property that $AA^{-1} = A^{-1}A = I_N$ where $I_N$ denotes the $N$-dimensional identity matrix with $I_{ij} = 1$ if $i = j$ and 0 otherwise.

Not every square matrix is invertible. We will see an example below.

**Example 3** The inverse of $A = \begin{pmatrix} 3 & 2 \\
1 & 4 \end{pmatrix}$, is given by

$$
A^{-1} = \frac{1}{10} \begin{pmatrix} 4 & -2 \\
-1 & 3 \end{pmatrix}
$$

We can check this by matrix multiplication:

$$
\begin{pmatrix} 3 & 2 \\
1 & 4 \end{pmatrix} \frac{1}{10} \begin{pmatrix} 4 & -2 \\
-1 & 3 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 10 & 0 \\
0 & 10 \end{pmatrix}
$$

$$
\frac{1}{10} \begin{pmatrix} 4 & -2 \\
-1 & 3 \end{pmatrix} \begin{pmatrix} 3 & 2 \\
1 & 4 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 10 & 0 \\
0 & 10 \end{pmatrix}
$$

Some useful properties of inverses:

- $(A^{-1})^{-1} = A$
- $(AB)^{-1} = B^{-1}A^{-1}$
2.0.2 Transpose

Transposing a matrix means to exchange row and columns. When \( A \in \mathbb{R}^{M \times N} \) its transpose, written as \( A^T \) is an \( N \times M \) matrix. Its entries are \((A^T)_{ij} = A_{ji}\).

Some useful properties of transposes:

- \((A^T)^T = A\)
- \((A + B)^T = A^T + B^T\)
- \((AB)^T = B^T A^T\)
- \((A^{-1})^T = (A^T)^{-1}\)

2.0.3 Trace

The trace \( \text{tr}(A) \) of a square matrix \( A \in \mathbb{R}^{N \times N} \) is defined as the sum of its diagonal elements, i.e. \( \text{tr}(A) = \sum_{i=1}^{N} A_{ii} \).

Some useful properties of the trace:

- \( \text{tr}(A^T) = \text{tr}(A)\)
- \( \text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)\)
- \( \text{tr}(AB) = \text{tr}(BA)\)

From the last property it follows that the trace is cyclic, i.e.

\[ \text{tr}(ABC) = \text{tr}(CAB) = \text{tr}(BCA) \]

2.1 Determinant

Determinants play an important role in computing eigenvalues and inverses. The determinant of a square matrix \( A \) is defined recursively:

1. The determinant of a \( 2 \times 2 \) matrix is given by

\[
\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc
\]

2. The determinant of an \( N \times N \) matrix \( A \), \( N > 2 \), is defined as

\[
\det A = \sum_{i=1}^{N} (-1)^{i+j} A_{ij} \det A_{-i,-j} = \sum_{j=1}^{N} (-1)^{i+j} A_{ij} \det A_{-i,-j}
\]

where \( A_{-i,-j} \) denotes the \((N-1) \times (N-1)\) matrix obtained by removing row \( i \) and column \( j \) from \( A \).
Example 4

\[
\det \begin{pmatrix}
1 & 2 & 3 \\
3 & 2 & 1 \\
4 & 2 & 0
\end{pmatrix} = 1 \det \begin{pmatrix}
2 & 1 \\
4 & 0
\end{pmatrix} - 2 \det \begin{pmatrix}
3 & 1 \\
4 & 0
\end{pmatrix} + 3 \det \begin{pmatrix}
3 & 2 \\
4 & 2
\end{pmatrix}
\]

\[
= 1(0 - 2) - 2(0 - 4) + 3(6 - 8) = -2 + 8 - 6 = 0
\]

This matrix cannot be invertible due to the following theorem:

A matrix \( A \) is invertible if and only if its determinant is non-zero, i.e. \( \det A \neq 0 \).

Some useful properties of determinants:

- \( \det(AB) = \det(A) \det(B) \) if \( A \) and \( B \) are invertible
- \( \det(A^{-1}) = \frac{1}{\det A} \)

Determinants are also used in computing eigenvalues.

2.2 Eigenvectors and eigenvalues

A square matrix \( A \in \mathbb{R}^{N \times N} \) maps from an \( N \)-dimensional vector space into itself.

Eigenvectors identify special directions, as vectors, which are just scaled by the action of \( A \):

\[ Av = \lambda v \]

We say that \( v \) is an eigenvector with eigenvalue \( \lambda \).

2.2.1 How to find eigenvectors and -values?

Note that an eigenvalue \( \lambda \) has to solve the equation \( \det(A - \lambda I_N) = 0 \). Thus, eigenvalues are roots of the so called characteristic polynomial of \( A \).

To see this, write the eigenvector equation as \( (A - \lambda I_N)v = 0 \). Since \( v \) is non-zero the matrix \( A - \lambda I_N \) cannot be invertible and thus, its determinant has to vanish.

Example 5

Compute the eigenvalues of \( A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \).

\[
\det(A - \lambda I_2) = \det \begin{pmatrix}
3 - \lambda & 2 \\
1 & 4 - \lambda
\end{pmatrix} = (3 - \lambda)(4 - \lambda) - 2 \cdot 1 = \lambda^2 - 7\lambda + 10 = (\lambda - 5)(\lambda - 2)
\]

Having found the eigenvalues \( \lambda_1 = 5, \lambda_2 = 2 \) we can solve for the eigenvectors:

\[
\begin{pmatrix}
3 & 2 \\
1 & 4
\end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 5 \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}
\]

\[
\begin{pmatrix}
3v_1 + 2v_2 - 5v_1 \\
v_1 + 4v_2 - 5v_2
\end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\]
Solving this, we find that \( \begin{pmatrix} 1 \\ 1 \end{pmatrix} \) is an eigenvector to the eigenvalue \( \lambda_1 = 5 \).

Similarly, \( \begin{pmatrix} 2 \\ -1 \end{pmatrix} \) is an eigenvector to \( \lambda_2 = 2 \).

A diagonalization of \( A \) can be obtained from its eigenvectors:

- Let \( U \) be a matrix, containing the eigenvectors of \( A \) as its columns. Then, \( AU = UD \) where \( D \) is a diagonal matrix containing the corresponding eigenvalues along its diagonal\(^1\).

- Multiplying with \( U^{-1} \) yields: \( A = UDU^{-1} \)

**Example 6** Diagonalize \( A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \):

\[
\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 5 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 2 \end{pmatrix}
\]

And thus,

\[
\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1/3 & 2/3 \\ 1/3 & -1/3 \end{pmatrix}
\]

In general, eigenvectors of a matrix can be complex. When \( A \) is symmetric, i.e. \( A_{ij} = A_{ji} \), all its eigenvalues are real.

In this case, \( U \) in the above diagonalization can be chosen as an orthogonal matrix, i.e. \( U^{-1} = U^T \):

- Normalize all eigenvectors \( v \) to unit length, i.e. \( v^Tv = 1 \).

- Eigenvectors \( v, w \) (for different eigenvalues \( \lambda_v \neq \lambda_w \)) are orthogonal:

\[
\lambda_v v^T w = (Av)^T w = v^T A^T w = v^T (Aw) = \lambda_w v^T w
\]

and thus, \( v^T w = 0 \).

For same eigenvalues just choose an orthogonal basis for the corresponding subspace.

Let \( \lambda_1, \ldots, \lambda_N \) denote the eigenvalues of \( A \). Then,

- \( \text{tr}(A) = \sum_{i=1}^N \lambda_i \)

- \( \text{det}(A) = \prod_{i=1}^N \lambda_i \)

---

\(^1\)Multiplying \( A \) with a diagonal matrix from the left scales the rows of \( A \), while multiplying from the right scales the columns.