Contents

1 Linear Algebra Tools

This chapter introduces inner product to give geometric meaning to vectors and vector spaces, enabling calculations of length, distance, and angles.

Definition (Symmetric Positive Definitive Matrix). *A symmetric matrix $A \in \mathbb{R}^{n\times n}$ that satisfies*

$$\text{for every nonzero vector } x : x^T Ax > 0 \tag{1.1}$$

*is called **positive definite**. If only $\geq$ holds in 1.1, then A is called **positive semidefinite**.*

These properties helps in identifying positive definite matrices without having to check the definition explicitly:

1. The null space of A contains only the null vector;
2. The diagonal elements a_{ii} of A are positive;
3. The eigenvalues of A are real and positive.

1.1 Angles and Orthogonality

The angle ω between vectors x and y is computed as:

$$\cos\omega = \frac{\langle x, y\rangle}{\|x\|_2\|y\|_2}$$

Here, $\langle x, y\rangle$ denotes the inner product between x and y.
This angle indicated the vectors' similarity in orientation.

Definition (Orthogonal vectors). *Two vectors are orthogonal if $\langle x, y\rangle = 0$. If additionally $\|x\| = 1 = \|y\|$, then x and y are orthonormal.*

Definition (Orthogonal matrix). *A square matrix is an orthogonal matrix if and only if <u>its columns are orthonormal</u> so that*

$$AA^T = I = A^T A$$

which implies that

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

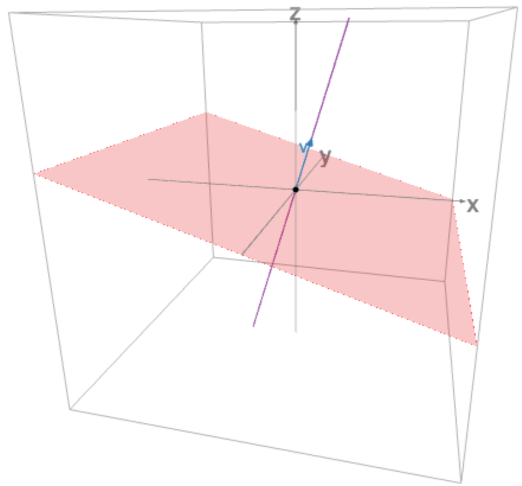
The length of a vector x is not changed when transforming it using an orthogonal matrix A.

$$\|Ax\|_2^2 = \|x\|_2^2$$

Moreover, the angle between any two vectors x, y is also unchanged when transforming both of them using an orthogonal matrix A.

Definition (Orthonormal Basis). *In an n-dimensional vector space V with a basis set $\{b_1, \ldots, b_n\}$, if all the basis vectors are orthogonal to each other, the basis is called as an **orthogonal basis**. Additionally, if the length of each basis vector is 1, the basis is referred to as an **orthonormal basis**.*

We can also have vector spaces that are orthogonal to each other. Given a vector space V of dimension D, let's consider a subspace U of dimension M such that $U \subseteq V$. Then its **orthogonal complement** $U^\perp$ is a $D - M$ dimensional subspace V and contains all vectors in V that are orthogonal to every vector in U.



1.1.1 Orthogonal Projections

Projections are key linear transformations in machine learning and are particularly useful for handling high-dimensional data. Often, only a few dimensions in such data are essential for capturing the most relevant information. By projecting the original high-dimensional data onto a lower dimensional feature space, we can work more efficiently to learn about the dataset and extract significant patterns.

Definition (Projection). *Let V be a vector space and $U \subseteq V$ a subspace of V. A linear mapping $\pi : V \to U$ is called **projection** if it satisfies $\pi^2 = \pi \circ \pi = \pi$.*

Given that linear mappings can be represented by transformation matrices, the above definition extends naturally to *projection matrices* P_π. These matrices exhibit the property that $P_\pi^2 = P_\pi$.
The projection $\pi_U(x)$ of a vector $x \in \mathbb{R}^n$ onto a subspace U is the closest point necessarily in U to x.

2 Matrix Decompositions

2.1 Eigenvalues and Eigenvectors

Eigenanalysis helps us understand linear transformations represented by a matrix A. Eigenvectors x are special vectors that only get scaled, not rotated, when multiplied by A. The scaling factor is the eigenvalue λ, which indicated how much x is stretched or shrunk. λ can also be zero.

Definition (Eigenvalue and Eigenvector)**.** *Let $A \in \mathbb{R}^{n\times n}$ be a square matrix. Then $\lambda \in \mathbb{R}$ is an* ***eigenvalue*** *of A and nonzero vector x is the corresponding* ***eigenvector*** *of A if*

$$Ax = \lambda x \tag{2.1}$$

We call 2.1 the ***eigenvalue equation****.*

The following statements are equivalent:

- λ is an eigenvalue of $A \in \mathbb{R}^{n\times n}$.
- A nonzero vector x exists such that $Ax = \lambda x$ or, equivalently, $(A - \lambda I_n)x = 0$ for $x \neq 0$.
- Then $A - \lambda I$ is a **singular matrix** and its determinant is **zero**.

Each eigenvector x has one unique eigenvalue λ, but each λ can have multiple eigenvectors.

Definition (Eigenspace and Eigenspectrum)**.** *For $A \in \mathbb{R}^{n\times n}$, the set of all eigenvectors of A associated with an eigenvalue λ spans a subspace of $\mathbb{R}^n$, which is called the* ***eigenspace*** *of A with respect to λ and is denoted by E_λ. The set of all eigenvalues of A is called the* ***eigenspectrum*** *of A.*

Definition. *Let λ_i be an eigenvalue of a square matrix A. Then the* ***geometric multiplicity*** *of λ_i is the number of linearly independent eigenvectors associated with λ_i. In other words, it is the dimensionality of the eigenspace spanned by the eigenvectors associated with λ_i.*

Theorem. *The eigenvectors $x_1, \ldots, x_n$ of a matrix $A \in \mathbb{R}^{n\times n}$ with n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ are linearly independent.*

This theorem states that eigenvectors of a matrix with n distinct eigenvalues form a basis of $\mathbb{R}^n$.