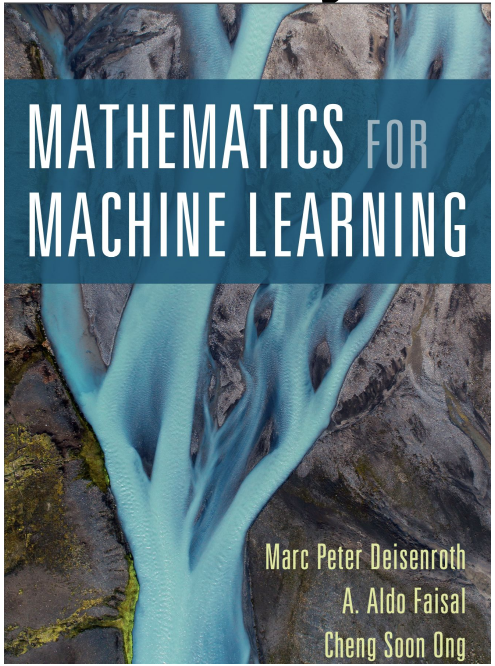
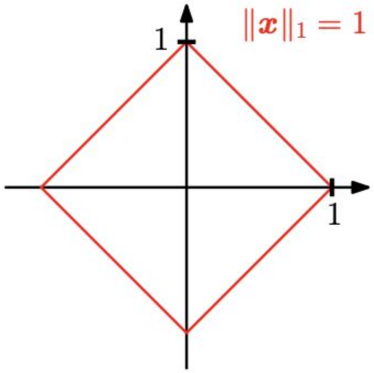
Chapter 3: Analytic Geometry

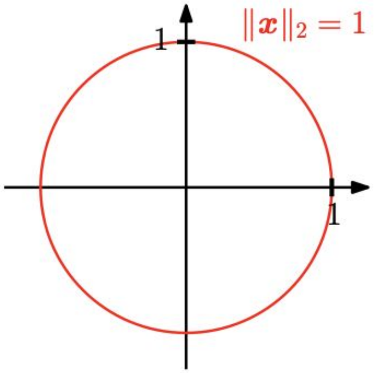


San Diego Machine Learning
Ryan Chesler

3.1 Norms

- A measure of the size or magnitude of a vector
- Manhattan/L1 Norm
- Euclidean/L2 Norm





3.1 Properties of Norms

- Absolutely homogeneous: $// \lambda x // = |\lambda| // x //$
- Triangle inequality: $// x + y // \leqslant // x // + // y //$
- Positive definite: $// x // \geqslant 0$ and $// x // = 0 \Longleftrightarrow x=0$

3.2 Inner Products

- Measures alignment (angle, projection) between two vectors
- Generalization of dot product

$$\boldsymbol{x}^\top \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i \,.$$

Example 3.4 (Symmetric, Positive Definite Matrices)
Consider the matrices

$$\boldsymbol{A}_1 = \begin{bmatrix} 9 & 6 \\ 6 & 5 \end{bmatrix}, \quad \boldsymbol{A}_2 = \begin{bmatrix} 9 & 6 \\ 6 & 3 \end{bmatrix}. \tag{3.12}$$

$\boldsymbol{A}_1$ is positive definite because it is symmetric and

$$\begin{aligned} \boldsymbol{x}^\top \boldsymbol{A}_1 \boldsymbol{x} &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 9 & 6 \\ 6 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} & \text{(3.13a)} \\ &= 9x_1^2 + 12x_1x_2 + 5x_2^2 = (3x_1 + 2x_2)^2 + x_2^2 > 0 & \text{(3.13b)} \end{aligned}$$

for all $\boldsymbol{x} \in V \backslash \{\boldsymbol{0}\}$. In contrast, $\boldsymbol{A}_2$ is symmetric but not positive definite because $\boldsymbol{x}^\top \boldsymbol{A}_2 \boldsymbol{x} = 9x_1^2 + 12x_1x_2 + 3x_2^2 = (3x_1 + 2x_2)^2 - x_2^2$ can be less than 0, e.g., for $\boldsymbol{x} = [2, -3]^\top$.

3.3 Lengths and distances

- We can measure the length of vectors with inner products
- Can also measure the distance between two vectors

Example 3.5 (Lengths of Vectors Using Inner Products)
In geometry, we are often interested in lengths of vectors. We can now use an inner product to compute them using (3.16). Let us take $\boldsymbol{x} = [1, 1]^\top \in \mathbb{R}^2$. If we use the dot product as the inner product, with (3.16) we obtain

$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^\top \boldsymbol{x}} = \sqrt{1^2 + 1^2} = \sqrt{2} \tag{3.18}$$

as the length of $\boldsymbol{x}$. Let us now choose a different inner product:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^\top \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \boldsymbol{y} = x_1 y_1 - \frac{1}{2}(x_1 y_2 + x_2 y_1) + x_2 y_2 \,. \tag{3.19}$$

If we compute the norm of a vector, then this inner product returns smaller values than the dot product if x_1 and x_2 have the same sign (and $x_1 x_2 > 0$); otherwise, it returns greater values than the dot product. With this inner product, we obtain

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = x_1^2 - x_1 x_2 + x_2^2 = 1 - 1 + 1 = 1 \implies \|\boldsymbol{x}\| = \sqrt{1} = 1 \,, \tag{3.20}$$

such that $\boldsymbol{x}$ is "shorter" with this inner product than with the dot product.

The mapping of vectors to a distance can be called a metric

Connected to interesting field of metric learning

3.4 Angles and Orthogonality

- With inner products we can also measure the angle between vectors
- This tells us how similar their orientation is
- Vectors are orthogonal(generalization of perpendicular) if inner product is 0
- If norms of x and y are 1 then it is also orthonormal

3.5 Orthonormal Basis

- Basis vectors are orthogonal and have length of 1
- We can use gaussian elimination on linearly independent vectors to find an orthonormal basis
  - Called Gram-Schmidt
- [1, 0, 0] [0, 1, 0] [0, 0, 1]

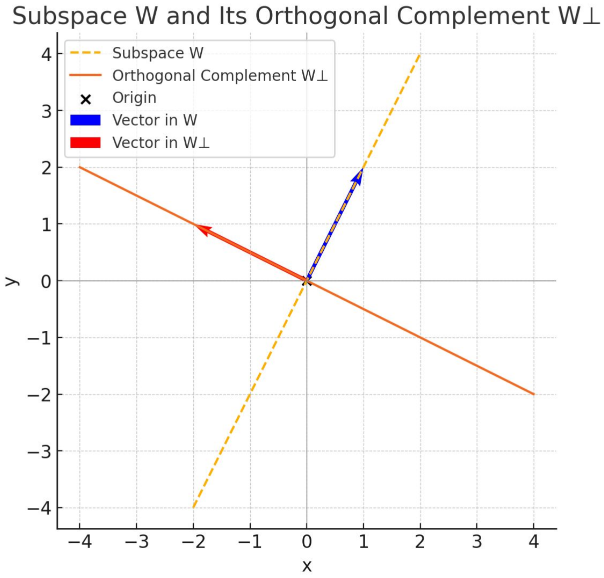
Example 3.8 (Orthonormal Basis)
The canonical/standard basis for a Euclidean vector space $\mathbb{R}^n$ is an orthonormal basis, where the inner product is the dot product of vectors.

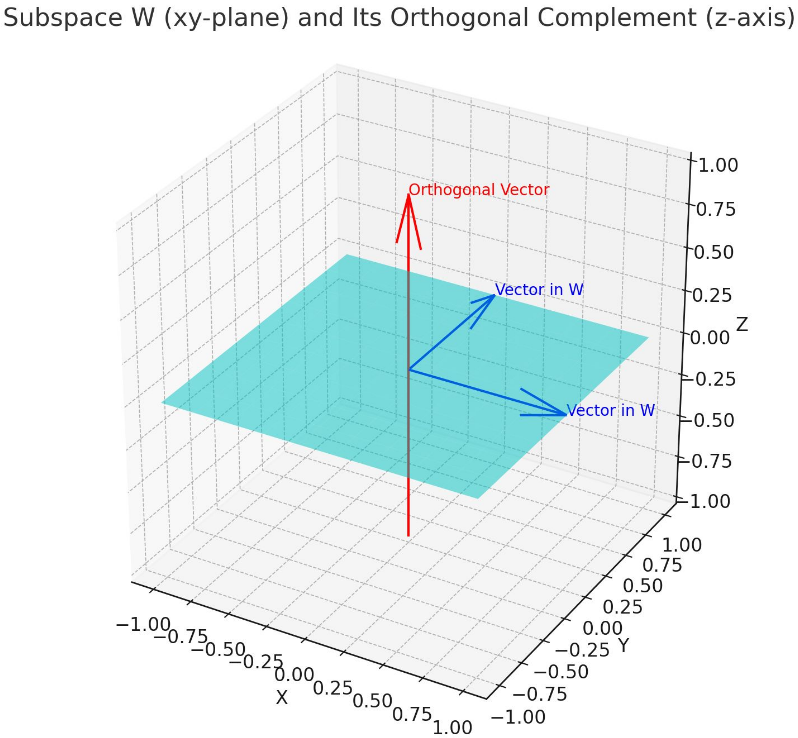
In $\mathbb{R}^2$, the vectors

$$\boldsymbol{b}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \quad \boldsymbol{b}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} \tag{3.35}$$

form an orthonormal basis since $\boldsymbol{b}_1^\top \boldsymbol{b}_2 = 0$ and $\|\boldsymbol{b}_1\| = 1 = \|\boldsymbol{b}_2\|$.

3.6 Orthogonal Complement





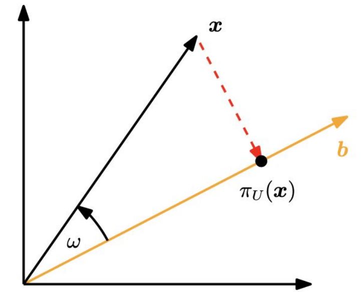
3.7 Inner Product of Functions

- Beyond just doing inner products on fixed size vectors we can do inner products of functions
  - Can think of the function like an infinitely long vector

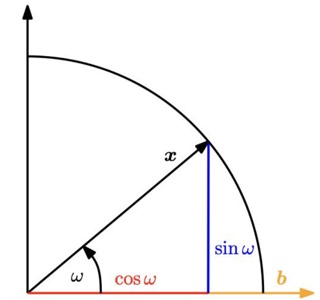
Example 3.9 (Inner Product of Functions)
If we choose $u = \sin(x)$ and $v = \cos(x)$, the integrand $f(x) = u(x)v(x)$ of (3.37), is shown in Figure 3.8. We see that this function is odd, i.e., $f(-x) = -f(x)$. Therefore, the integral with limits $a = -\pi, b = \pi$ of this product evaluates to 0. Therefore, $\sin$ and $\cos$ are orthogonal functions.

3.8 Orthogonal Projections

- Orthogonal projections are very important to many parts of machine learning
- PCA, Neural networks, linear regression - can be viewed from a dimension reduction light
- Can be thought of as finding the shortest path to a subspace
    - Depends on our inner product
- There exists a projection matrix that will project all points to the new subspace



(a) Projection of $x \in \mathbb{R}^2$ onto a subspace U with basis vector b.



(b) Projection of a two-dimensional vector x with $\|x\| = 1$ onto a one-dimensional subspace spanned by b.

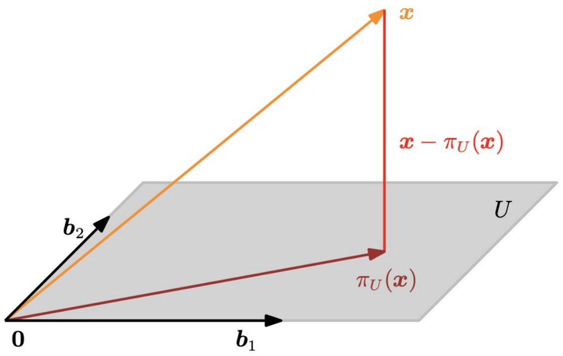
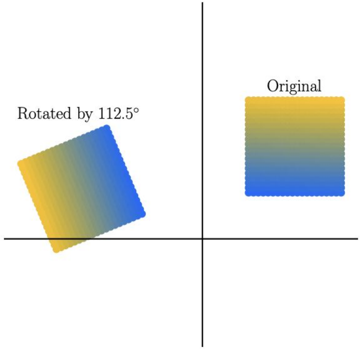


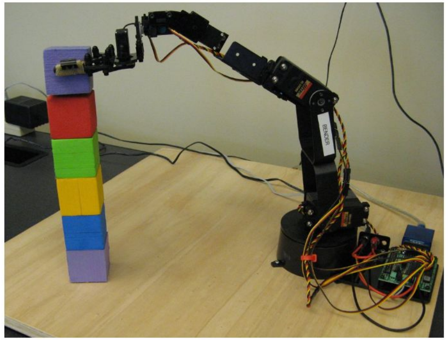
Figure 3.11 Projection onto a two-dimensional subspace U with basis b_1, b_2. The projection $\pi_U(x)$ of $x \in \mathbb{R}^3$ onto U can be expressed as a linear combination of b_1, b_2 and the displacement vector $x - \pi_U(x)$ is orthogonal to both b_1 and b_2.

3.9 Rotations

- Rotations preserve angles and distances
- Can apply a rotation matrix to get the desired outcome

$$:\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \quad : \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \quad \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$





Next Month - Matrix Decompositions