Analytic Geometry 1

Liang Zheng

Australian National University

liang.zheng@anu.edu.au

3.1 Norms

- A norm on a vector space V is a function

$$\|\cdot\| : V \to \mathbb{R},$$

$$x \mapsto \|x\|, \text{ vector to real number}$$

which assigns each vector x its length $\|x\| \in \mathbb{R}$.

Examples

- The Manhattan norm on $\mathbb{R}^n$ is defined for $x \in \mathbb{R}^n$ as

$$\|x\|_1 := \sum_{i=1}^{n} |x_i|,$$

$\begin{bmatrix} x_1 \\ x_1 \\ \vdots \\ x_n \end{bmatrix}$

L1

>=0
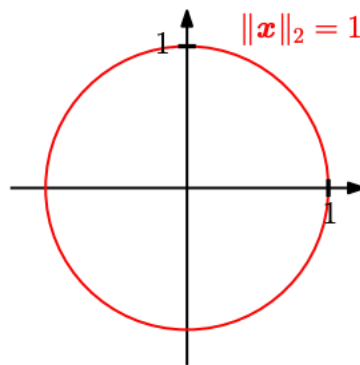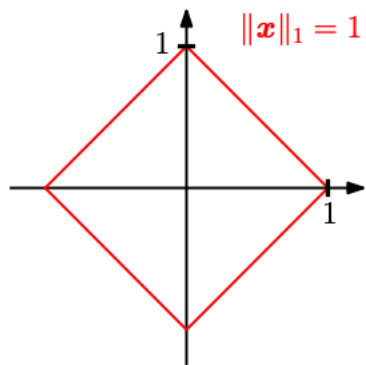
where $|\cdot|$ is the absolute value. It is also called ℓ_1 norm.

- The Euclidean norm of $x \in \mathbb{R}^n$ is defined as

$$\|x\|_2 := \sqrt{\sum_{i=1}^{n} x_i{}^2} = \sqrt{x^\mathrm{T} x}$$

It is the Euclidean distance of x from the origin; also called ℓ_2 norm
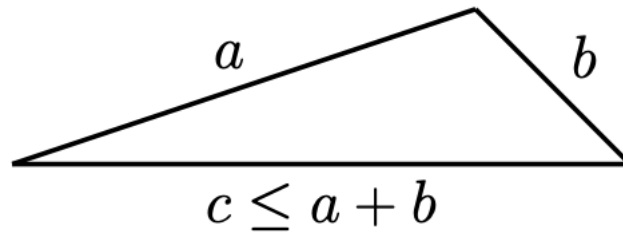
$\|x\|_1 = 1$

$\|x\|_2 = 1$

3

3.1 Norms

For all $\lambda \in \mathbb{R}$, and $\boldsymbol{x}, \boldsymbol{y} \in V$ the following holds:

- Absolutely homogeneous: $\|\lambda \boldsymbol{x}\| = |\lambda| \|\boldsymbol{x}\|$

- Triangle inequality: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$

- Positive definite: $\|\boldsymbol{x}\| \geq 0$ and $\|\boldsymbol{x}\| = 0 \iff \boldsymbol{x} = \boldsymbol{0}$

$$c \leq a + b$$

Sparse representation Wright et al., TPAMI, 2009
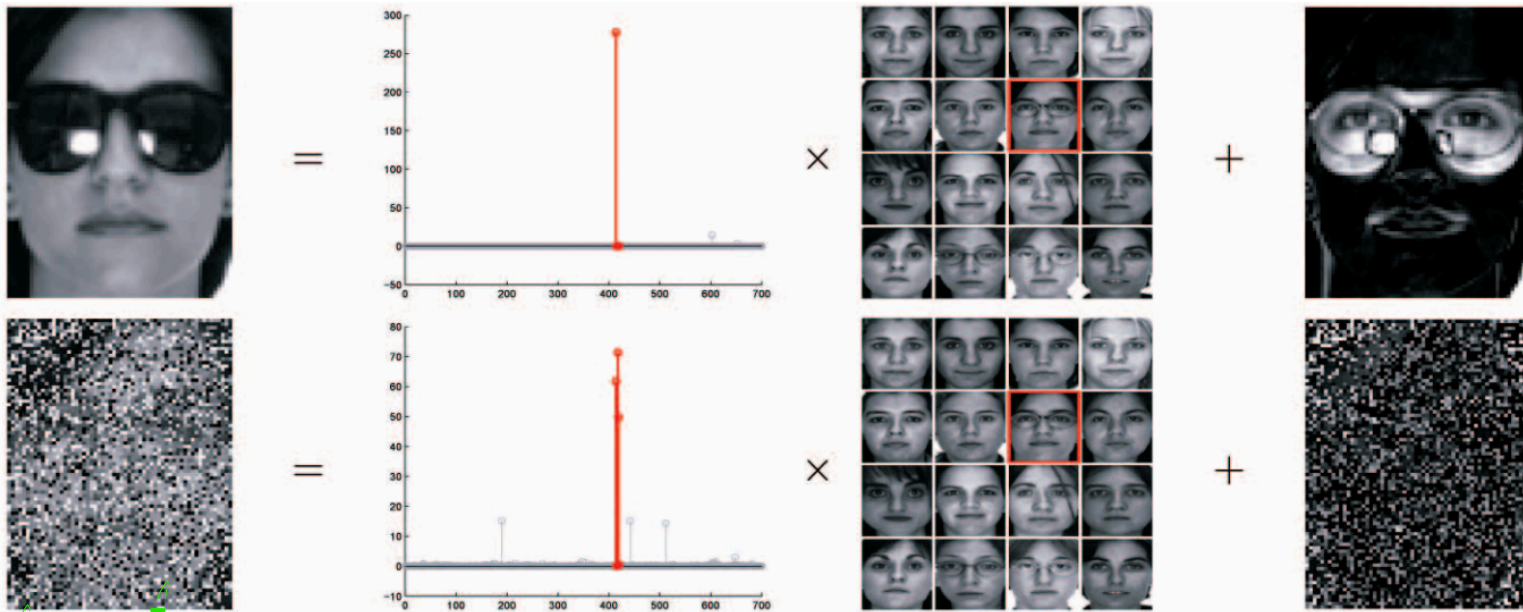
Input

$$y = \alpha_{i,1} v_{i,1} + \alpha_{i,2} v_{i,2} + \cdots + \alpha_{i,n_i} v_{i,n_i}$$

$$A \doteq [A_1, A_2, \ldots, A_k] = [v_{1,1}, v_{1,2}, \ldots, v_{k,n_k}]$$

$$y = A x_0 \quad \text{coefficients/weights}$$

$$\hat{x}_0 = \arg\min \|x\|_0 \quad \text{subject to} \quad Ax = y$$

L0 norm, the number of non-zero elements

5

3.2.1 Dot Product

- Scalar product/dot product in $\mathbb{R}^n$ is given by

$$x^{\mathrm{T}} y = \sum_{i=1}^{n} x_i y_i \qquad \mathsf{R}$$

$$\underline{1 \times n \quad n \times 1}$$

Bilinear mapping

- A bilinear mapping Ω is a mapping with two arguments, and it is linear in each argument. Consider a vector space V, for all $x, y, z \in V, \lambda, \varphi \in \mathbb{R}$,

$$\Omega(\lambda x + \varphi y, z) = \lambda\Omega(x, z) + \varphi\Omega(y, z)$$

Ω is linear in the first argument

$$\Omega(x, \lambda y + \varphi z) = \lambda\Omega(x, y) + \varphi\Omega(x, z).$$

Ω is linear in the second argument

Inner product

- Let V be a vector space and $\Omega: V \times V \to \mathbb{R}$ be a bilinear mapping.

- Ω is called symmetric if $\Omega(\boldsymbol{x}, \boldsymbol{y}) = \Omega(\boldsymbol{y}, \boldsymbol{x})$

- Ω is called positive definite if

$$\forall \boldsymbol{x} \in V \setminus \{\boldsymbol{0}\} : \Omega(\boldsymbol{x}, \boldsymbol{x}) > 0, \qquad \Omega(\boldsymbol{0}, \boldsymbol{0}) = 0 \qquad >=C$$

- A positive definite, symmetric bilinear mapping $\Omega: V \times V \to \mathbb{R}$ is called an inner product on V. We write $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ instead of $\Omega(\boldsymbol{x}, \boldsymbol{y})$.

- The pair $(V, \langle \cdot, \cdot \rangle)$ is called is called an inner product vector space. If we use the dot product, we call $(V, \langle \cdot, \cdot \rangle)$ a Euclidean vector space.

Example

- Consider $V = \mathbb{R}^2$. If we define

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := x_1 y_1 - (x_1 y_2 + x_2 y_1) + 2 x_2 y_2$$

- then $\langle \cdot, \cdot \rangle$ is an inner product but different from the dot product.

This mapping is symmetric: it is easy to derive $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$

Is it positive definite?

$\forall \boldsymbol{x} \in V \backslash \{\boldsymbol{0}\}, \langle \boldsymbol{x}, \boldsymbol{x} \rangle = x_1^2 - (x_1 x_2 + x_2 x_1) + 2 x_2^2 = (x_1 - x_2)^2 + x_2^2 > 0$

3.2.3 Symmetric, Positive Definite Matrices

- Consider an n-dimensional vector space V with an inner product $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$, and a basis $B = (\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n)$ of V.

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \left\langle \sum_{i=1}^{n} \varphi_i \boldsymbol{b}_i, \sum_{j=1}^{n} \lambda_j \boldsymbol{b}_j \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \varphi_i \langle \boldsymbol{b}_i, \boldsymbol{b}_j \rangle \lambda_j = \widehat{\boldsymbol{x}}^T A \widehat{\boldsymbol{y}}$$

Basis vectors

where $A_{ij} := \langle \boldsymbol{b}_i, \boldsymbol{b}_j \rangle$ and $\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}$ are the coordinates of $\boldsymbol{x}, \boldsymbol{y}$ with respect to the basis B.

- The inner product $\langle \cdot, \cdot \rangle$ is uniquely determined through A. The symmetry of the inner product also means that A is symmetric.

- The positive definiteness of the inner product implies that

$$\forall \boldsymbol{x} \in V \setminus \{\boldsymbol{0}\} : \langle \boldsymbol{x}, \boldsymbol{x} \rangle = \boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} > 0$$

3.2.3 Symmetric, Positive Definite Matrices

- A symmetric matrix $A \in \mathbb{R}^{n \times n}$ that satisfies $\forall x \in V \setminus \{0\} : x^T A x > 0$ is called symmetric, positive definite, or just positive definite. If only $\geq$ holds, then A is called symmetric, positive semidefinite.

- Example

$$A_1 = \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 1 & 3 \\ 3 & 3 \end{bmatrix}$$

- A_1 is positive definite because it is symmetric and

$$x^T A_1 x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 3x_1^2 + 2x_1 x_2 + 4x_2^2 = (x_1 + x_2)^2 + 2x_1^2 + 3x_2^2 > 0$$

for all $x \in V \setminus \{0\}$.

- A_2 is symmetric but not positive definite

$$x^T A_2 x = x_1^2 + 6x_1 x_2 + 3x_2^2 = (x_1 + 3x_2)^2 - 6x_2^2 \text{ can be less than 0}$$

3.2.3 Symmetric, Positive Definite Matrices

- For a real-valued, finite-dimensional vector space V and a basis B of V, it holds that $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ is an inner product if and only if there exists a symmetric, positive definite matrix $A \in \mathbb{R}^{n \times n}$ with

$$\langle x, y \rangle = \widehat{x}^T A \widehat{y}$$

- If $A \in \mathbb{R}^{n \times n}$ is symmetric and positive definite,

  the diagonal elements a_{ii} of A are positive because $a_{ii} = e_i^{\mathrm{T}} A e_i = \langle e_i, e_i \rangle > 0$, where e_i is the ith vector of the standard basis in $\mathbb{R}^n$.

3.3 Lengths and Distances

- Any inner product induces a norm

$$\|x\| := \sqrt{\langle x, x \rangle}$$

- Cauchy-Schwarz Inequality

- For an inner product vector space $(V, \langle \cdot, \cdot \rangle)$ the induced norm $\|\cdot\|$ satisfies the Cauchy-Schwarz inequality

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

Example - Lengths of Vectors Using Inner Products

- We can now use an inner product to compute vector lengths, using $\|x\| :=$
  $\sqrt{\langle x, x \rangle}$. Consider $x = [1,1]^{\mathrm{T}} \in \mathbb{R}^2$. If we use the dot product as the inner
  product, we obtain

$$\|x\| = \sqrt{x^T x} = \sqrt{1^2 + 1^2} = \sqrt{2}$$

$$x^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} y \text{ is dot product}$$

as the length of x. Let us now choose a different inner product:

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

With this inner product, we obtain

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

x is "shorter" with this inner product than with the dot product.

14

3.3 Lengths and Distances

- Consider an inner product space $(V, \langle \cdot, \cdot \rangle)$, then

$$d(\boldsymbol{x}, \boldsymbol{y}) := \|\boldsymbol{x} - \boldsymbol{y}\| = \sqrt{\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle}$$

is called the distance between $\boldsymbol{x}$ and $\boldsymbol{y}$ for $\boldsymbol{x}, \boldsymbol{y} \in V$.

- If we use the dot product as the inner product, then the distance is called Euclidean distance.

3.3 Lengths and Distances

- The mapping

$$d : V \times V \to \mathbb{R}$$
$$(\boldsymbol{x}, \boldsymbol{y}) \mapsto d(\boldsymbol{x}, \boldsymbol{y})$$

  is called a metric.

- A metric d satisfies the following:

- d is positive definite, i.e., $d(\boldsymbol{x}, \boldsymbol{y}) \geq 0$ for all $\boldsymbol{x}, \boldsymbol{y} \in V$ and $d(\boldsymbol{x}, \boldsymbol{y}) = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{y}$

- d is symmetric, i.e., $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$ for all $\boldsymbol{x}, \boldsymbol{y} \in V$

- Triangle inequality: $d(\boldsymbol{x}, \boldsymbol{z}) \leq d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z})$ for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$

- Very similar $\boldsymbol{x}$ and $\boldsymbol{y}$ will result in a large value for the inner product and a small value for the metric.
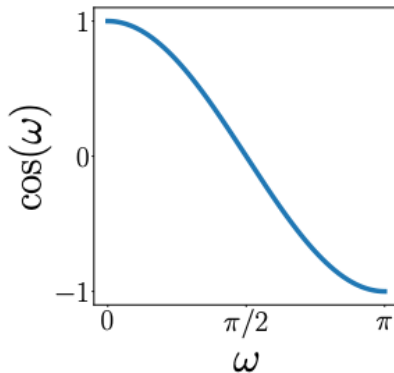
3.4 Angles and Orthogonality

$$|\langle x, y \rangle| \leq \|x\|\|y\|$$

- According to Cauchy-Schwarz inequality, assume $x \neq 0$, $y \neq 0$. Then,

$$-1 \leq \frac{\langle x, y \rangle}{\|x\|\|y\|} \leq 1$$

Therefore, there exists a unique $\omega \in [0, \pi]$, with

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

The number ω is the angle between the vectors x and y.
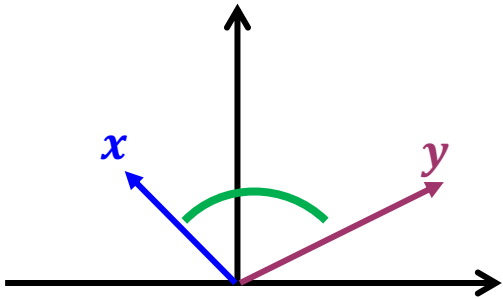
- The angle between two vectors tells us how similar their orientations are.

- Using the dot product, the angle between x and $y = 4x$ is 0, so their orientation is the same.

$$\cos\omega = \frac{\langle x, 4x \rangle}{\|x\|\|4x\|} = \frac{4\langle x, x \rangle}{\sqrt{x^T x}\sqrt{(4x)^T(4x)}} = \frac{4\langle x, x \rangle}{4\|x\|\|x\|} = \frac{\langle x, x \rangle}{\|x\|\|x\|} = 1$$

Example (Angle between Vectors)

- Let us compute the angle between $x = [-1,1]^{\mathrm{T}} \in \mathbb{R}^2$ and $y = [2,1]^{\mathrm{T}} \in \mathbb{R}^2$. We use the dot product as the inner product. We get

$$cos\omega = \frac{\langle x, y \rangle}{\sqrt{\langle x, x \rangle \langle y, y \rangle}} = \frac{x^T y}{\sqrt{x^T x y^T y}} = \frac{-1}{\sqrt{10}}$$

- and the angle between the two vectors is $\arccos\left(\frac{-1}{\sqrt{10}}\right) \approx 1.89\mathrm{rad}$, which corresponds to about 108.4°.

- We then use inner product to characterize orthogonality.
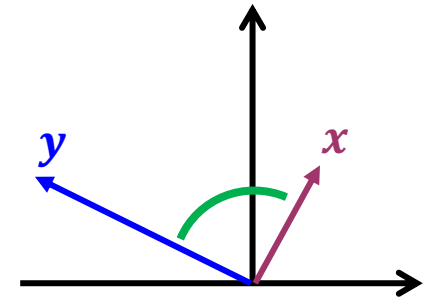
3.4 Angles and Orthogonality

- Two vectors x and y are orthogonal **if and only if** $\langle x, y \rangle = 0$, and we write $x \perp y$. If additionally $\|x\| = \|y\| = 1$, i.e., the vectors are unit vectors, then x and y are orthonormal.

- 0-vector is orthogonal to every vector in the vector space

- Example (Orthogonal Vectors)

- Consider $x = [1,2]^{\mathrm{T}}$ and $y = [-4,2]^{\mathrm{T}}$

- Using dot product as inner product, we have

  - $\langle x, y \rangle = 0$, so $x \perp y$.

- if we choose the inner product

$$\langle x, y \rangle = x^{T} \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix} y$$

- the angle ω between x and y is given by

$$cos\omega = \frac{\langle x, y \rangle}{\|x\|\|y\|} = -\frac{2}{\sqrt{17 \times 12}} \quad \Longrightarrow \quad \omega \approx 1.43\mathrm{rad} \approx 81.95°$$

19

3.4 Angles and Orthogonality

- A square matrix $A \in \mathbb{R}^{n \times n}$ is an <u>orthogonal matrix</u> if and only if its <u>columns are</u> ==orthonormal==, such that

$$A\,A^T = I = A^T A$$

which implies that

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

i.e., the inverse is obtained by simply transposing the matrix

Properties - length

- The length of a vector x is not changed when transforming it using an orthogonal matrix A. For dot product, we obtain

$$\|Ax\|^2 = (Ax)^T(Ax) = x^T A^T A x = x^T I x = x^T x = \|x\|^2$$

Properties - angle

- The ==angle== between any two vectors x and y as measured by their inner product, is also <u>unchanged</u> when transforming <u>both of</u> <u>them</u> using an <u>orthogonal matrix A</u>. We use the dot product as inner product

$$cos\omega = \frac{(Ax)^T(Ay)}{\|Ax\|\|Ay\|} = \frac{x^T A^T A y}{\sqrt{x^T A^T A x y^T A^T A y}} = \frac{x^T y}{\|x\|\|y\|}$$

x, y

- Orthogonal matrices A with $A^{-1} = A^T$ preserve both angles and distances.

- Orthogonal matrices define transformations that are rotations

3.5 Orthonormal Basis

- Consider an n-dimensional vector space V and a basis $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ of V. For all $i, j = 1, \cdots, n$, if

$$\langle \boldsymbol{b}_i, \boldsymbol{b}_j \rangle = 0 \quad \text{for} \quad i \neq j \qquad (1)$$

$$\langle \boldsymbol{b}_i, \boldsymbol{b}_i \rangle = 1$$

then the basis is called an orthonormal basis (ONB).

If only (1) is satisfied, the basis is called an orthogonal basis.

Example (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

$$\text{For } \mathbb{R}^3: \quad \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

- In $\mathbb{R}^2$, the vectors $\boldsymbol{b}_1 = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \boldsymbol{b}_2 = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

  form an orthonormal basis since $b_1^\top b_2 = 0$ and $\|b_1\| = 1 = \|b_2\|$

3.6 Orthogonal Complement

- A hyperplane U in a three-dimensional vector space can be described by its normal vector, which spans its orthogonal complement $U^{\perp}$
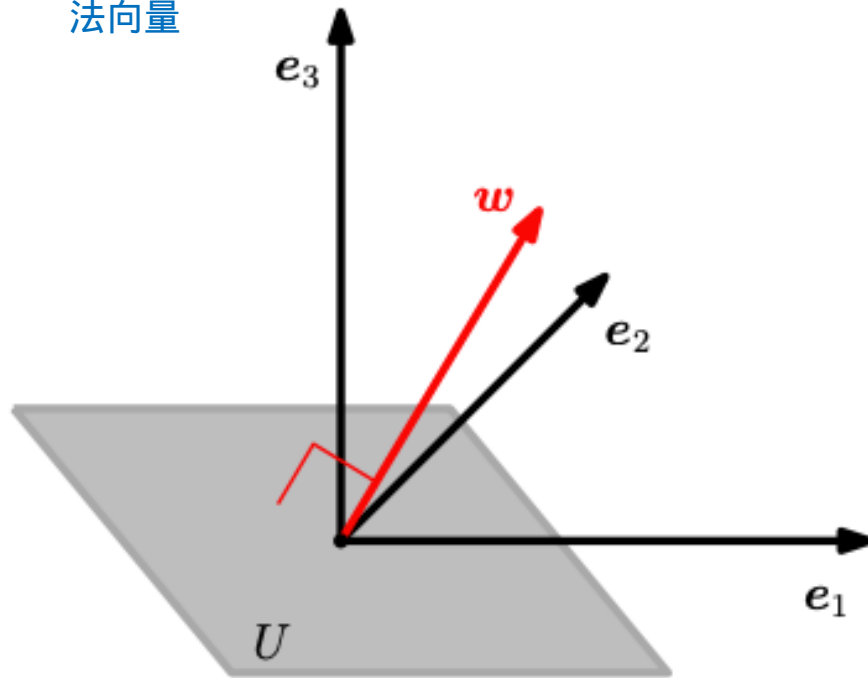
- Generally, orthogonal complements can be used to describe hyperplanes in n-dimensional vector and affine spaces

$(n-1)$ dimensional

3.6 Orthogonal Complement

- We now look at vector spaces that are orthogonal to each other

- Consider a D-dimensional vector space V and an M-dimensional subspace $U \subseteq V$. The orthogonal complement $U^\perp$ is a $(D-M)$-dimensional subspace of V and contains all vectors in V that are orthogonal to every vector in U.

- $U \cap U^\perp = \{\mathbf{0}\}$ so that any vector $\boldsymbol{x} \in V$ can be uniquely decomposed into

$$x = \sum_{m=1}^{M} \lambda_m \, \boldsymbol{b}_m + \sum_{j=1}^{D-M} \psi_j \boldsymbol{b}_j^\perp \,, \lambda_m, \psi_j \in \mathbb{R}$$

- Where $(\boldsymbol{b}_1, \dots, \boldsymbol{b}_M)$ is a basis of U and $(\boldsymbol{b}_1^\perp, \dots, \boldsymbol{b}_{D-M}^\perp)$ is a basis of $U^\perp$.

i.e. U+(U) = original vector space V
        b (b)   V

Check your understanding

(A) Norm characterises the length of a vector. ✓

(B) The norm of a vector can be a ~~complex~~ *Real* number. ✗

(C) The inner product takes ~~one~~ 2 vector as input and outputs a real number. ✗

(D) A metric characterises the similarity between two vectors. ✓

(E) Any bilinear mapping introduces an inner product ✗

(F) Any inner product introduces a norm ✓

(G) Any vector in $U^{\perp}$ is orthogonal to any vector in U. ✓

(H) In $\mathbb{R}^2$ there can be infinitely many bases, but only a finite number of orthogonal / orthonormal bases ✗