Àlgebra lineal i geometria
4. Ortogonalitat

Grau en Enginyeria Física

2023-24

Universitat Politècnica de Catalunya
Departament de Matemàtiques

Marta Casanellas

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Outline

Outline

The Euclidean scalar product

Definition

The **Euclidean scalar product** (or **dot product**) $< u, v >$ of two vectors $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, v = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ is

$$< u, v > := u^t v = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Properties:

1. $< u, u > \geq 0 \; \forall u$ and $< u, u >= 0 \Leftrightarrow u = 0$ (*positive definite*)
2. $< u, v >=< v, u >$ (*symmetric*).
3. *bilineal*:
   - $< a_1 u_1 + a_2 u_2, v >= a_1 < u_1, v > + a_2 < u_2, v >$,
   - $< u, a_1 v_1 + a_2 v_2 >= a_1 < u, v_1 > + a_2 < u, v_2 >$.

Any function $\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ that satisfies these properties is called a *scalar product*.

The Euclidean scalar product

Definition

The **Euclidean scalar product** (or **dot product**) $< u, v >$ of two vectors $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, v = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ is

$$< u, v > := u^t v = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Properties:

1. $< u, u > \geq 0 \ \forall u$ and $< u, u >= 0 \Leftrightarrow u = 0$ (*positive definite*)
2. $< u, v >=< v, u >$ (*symmetric*).
3. *bilineal*:
   - $< a_1 u_1 + a_2 u_2, v >= a_1 < u_1, v > + a_2 < u_2, v >$;
   - $< u, a_1 v_1 + a_2 v_2 >= a_1 < u, v_1 > + a_2 < u, v_2 >$.

Any function $\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ that satisfies these properties is called a *scalar product*.

The Euclidean scalar product

Definition

The **Euclidean scalar product** (or **dot product**) $< u, v >$ of two vectors $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, v = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ is

$$< u, v > := u^t v = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n.$$

Properties:

1. $< u, u > \geq 0 \ \forall u$ and $< u, u > = 0 \Leftrightarrow u = 0$ (*positive definite*)
2. $< u, v > = < v, u >$ (*symmetric*).
3. *bilineal*:
   - $< a_1 u_1 + a_2 u_2, v > = a_1 < u_1, v > + a_2 < u_2, v >$;
   - $< u, a_1 v_1 + a_2 v_2 > = a_1 < u, v_1 > + a_2 < u, v_2 >$.

Any function $\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ that satisfies these properties is called a *scalar product*.

The Euclidean scalar product

Definition

The **Euclidean scalar product** (or **dot product**) $< u, v >$ of two vectors $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, v = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ is

$$< u, v > := u^t v = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n .$$

Properties:

1. $< u, u > \geq 0 \ \forall u$ and $< u, u >= 0 \Leftrightarrow u = 0$ (*positive definite*)
2. $< u, v >=< v, u >$ (*symmetric*).
3. *bilineal*:
   - $< a_1 u_1 + a_2 u_2, v >= a_1 < u_1, v > + a_2 < u_2, v >$;
   - $< u, a_1 v_1 + a_2 v_2 >= a_1 < u, v_1 > + a_2 < u, v_2 >$.

Any function $\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ that satisfies these properties is called a *scalar product*.

Bilinear forms

Let E be an $\mathbb{R}$-e.v. A **bilinear form** on E is a map $\varphi : E \times E \longrightarrow \mathbb{R}$ such that, $\forall\ u, v, w \in E$ and $\lambda \in \mathbb{R}$:

(a) $\varphi(u+v, w) = \varphi(u, w) + \varphi(v, w)\ \varphi(\lambda u, w) = \lambda\varphi(u, w)$,

(b) $\varphi(w, u+v) = \varphi(w, u) + \varphi(w, v)\ \varphi(w, \lambda u) = \lambda\varphi(w, u)$.

If $\mathbf{u} = \{u_1, \dots, u_n\}$ is a basis of E, then the **matrix of φ in the basis u** is defined as

$$M_{\mathbf{u}}(\varphi) = \begin{pmatrix} \varphi(u_1, u_1) & \cdots & \varphi(u_1, u_n) \\ \vdots & & \vdots \\ \varphi(u_n, u_1) & \cdots & \varphi(u_n, u_n) \end{pmatrix}.$$

Bilinear forms

Let E be an $\mathbb{R}$-e.v. A **bilinear form** on E is a map $\varphi : E \times E \longrightarrow \mathbb{R}$ such that, $\forall\ u, v, w \in E$ and $\lambda \in \mathbb{R}$:

(a) $\varphi(u+v, w) = \varphi(u, w) + \varphi(v, w)\ \varphi(\lambda u, w) = \lambda\varphi(u, w)$,

(b) $\varphi(w, u+v) = \varphi(w, u) + \varphi(w, v)\ \varphi(w, \lambda u) = \lambda\varphi(w, u)$.

If $\mathbf{u} = \{u_1, \ldots, u_n\}$ is a basis of E, then the **matrix of φ in the basis $\mathbf{u}$** is defined as

$$M_{\mathbf{u}}(\varphi) = \begin{pmatrix} \varphi(u_1, u_1) & \cdots & \varphi(u_1, u_n) \\ \vdots & & \vdots \\ \varphi(u_n, u_1) & \cdots & \varphi(u_n, u_n) \end{pmatrix}.$$

Matrix of a bilinear form

Properties:

1. If $v_{\mathbf{u}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, w_{\mathbf{u}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \Rightarrow$

   $\varphi(v, w) = (x_1 \dots x_n) M_{\mathbf{u}}(\varphi) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ and $M_{\mathbf{u}}(\varphi)$ is the unique matrix that satisfies this.

2. If $\mathbf{v}$ is another basis, then

$$M_{\mathbf{v}}(\varphi) = A^t_{\mathbf{v} \to \mathbf{u}} M_{\mathbf{u}}(\varphi) A_{\mathbf{v} \to \mathbf{u}}$$

A bilinear form φ is **symmetric** if $\varphi(u, v) = \varphi(v, u)$ for all u, v.
A bilinear form is symmetric $\Leftrightarrow M_{\mathbf{u}}(\varphi)$ is a symmetric matrix for any basis $\mathbf{u}$.

Matrix of a bilinear form

Properties:

1. If $v_{\mathbf{u}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, w_{\mathbf{u}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \Rightarrow$

   $\varphi(v, w) = (x_1 \dots x_n) M_{\mathbf{u}}(\varphi) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ and $M_{\mathbf{u}}(\varphi)$ is the unique matrix that satisfies this.
2. If $\mathbf{v}$ is another basis, then

$$M_{\mathbf{v}}(\varphi) = A^t_{\mathbf{v}\to\mathbf{u}} M_{\mathbf{u}}(\varphi) A_{\mathbf{v}\to\mathbf{u}}$$

A bilinear form φ is **symmetric** if $\varphi(u, v) = \varphi(v, u)$ for all u, v.
A bilinear form is symmetric $\Leftrightarrow M_{\mathbf{u}}(\varphi)$ is a symmetric matrix for any basis $\mathbf{u}$.

Matrix of a bilinear form

Properties:

1. If $v_{\mathbf{u}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, w_{\mathbf{u}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \Rightarrow$

   $\varphi(v, w) = (x_1 \dots x_n) M_{\mathbf{u}}(\varphi) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ and $M_{\mathbf{u}}(\varphi)$ is the unique matrix that satisfies this.
2. If $\mathbf{v}$ is another basis, then

$$M_{\mathbf{v}}(\varphi) = A^t_{\mathbf{v}\to\mathbf{u}} M_{\mathbf{u}}(\varphi) A_{\mathbf{v}\to\mathbf{u}}$$

A bilinear form φ is **symmetric** if $\varphi(u, v) = \varphi(v, u)$ for all u, v.
A bilinear form is symmetric $\Leftrightarrow M_{\mathbf{u}}(\varphi)$ is a symmetric matrix for any basis $\mathbf{u}$.

Scalar products

Let E be an $\mathbb{R}$-e.v. and φ a bilinear form on E. One says that φ is **positive definite** if $\varphi(u, u) \geq 0$ with equality only when $u = 0$.

Definition

A **scalar product** on E is a symmetric, positive definite bilinear form $<,>: E \times E \longrightarrow \mathbb{R}$. An $\mathbb{R}$-e.v together with a scalar product is called a **Euclidean** vector space.

Examples:

- The Euclidean scalar product
- $E = \mathcal{F}([a, b], \mathbb{R}) = \{$ continuous real functions defined on $[a, b]\}$, then the following defines a scalar product on E:

$$< f, g >:= \int_a^b f(x)g(x)dx.$$

Scalar products

Let E be an $\mathbb{R}$-e.v. and φ a bilinear form on E. One says that φ is **positive definite** if $\varphi(u, u) \geq 0$ with equality only when $u = 0$.

Definition

A **scalar product** on E is a symmetric, positive definite bilinear form $<, >: E \times E \longrightarrow \mathbb{R}$. An $\mathbb{R}$-e.v together with a scalar product is called a **Euclidean** vector space.

Examples:

- The Euclidean scalar product
- $E = \mathcal{F}([a, b], \mathbb{R}) = \{$ continuous real functions defined on $[a, b]\}$, then the following defines a scalar product on E:

$$< f, g >:= \int_a^b f(x)g(x)dx.$$

Scalar products

Let E be an $\mathbb{R}$-e.v. and φ a bilinear form on E. One says that φ is **positive definite** if $\varphi(u,u) \geq 0$ with equality only when $u = 0$.

Definition

A **scalar product** on E is a symmetric, positive definite bilinear form $<,>: E \times E \longrightarrow \mathbb{R}$. An $\mathbb{R}$-e.v together with a scalar product is called a **Euclidean** vector space.

Examples:

- The Euclidean scalar product
- $E = \mathcal{F}([a,b],\mathbb{R}) = \{$ continuous real functions defined on $[a,b]\}$, then the following defines a scalar product on E:

$$< f, g > := \int_a^b f(x)g(x)dx.$$

Norm and distance

Let E be an $\mathbb{R}$-e.v. with scalar product $<,>$. The **norm** of $u \in E$ is $\|u\| = \sqrt{<u,u>}$.
If $<,>$ is the Euclidean product, the norm is called the *standard, Euclidean, or 2-norm* and is also denoted as $||u||_2$.
Properties: for any $u, v \in E$ and $c \in \mathbb{R}$

1. $\|u\| \geq 0 \ \forall u$ and $\|u\| = 0 \Leftrightarrow u = 0$;
2. $\|cu\| = |c|\|u\| \ c \in \mathbb{R}$;
3. $| < u, v > | \leq \|u\|\|v\|$ (Cauchy-Schwarz inequality)
4. $\|u + v\| \leq \|u\| + \|v\|$ (triangular inequality);

Any function $f : E \longrightarrow \mathbb{R}$ that satisfies properties 1,2,4 is called a *norm* (and is not necessarily defined through a scalar product).

Norm and distance

Let E be an $\mathbb{R}$-e.v. with scalar product $<,>$. The **norm** of $u \in E$ is $\|u\| = \sqrt{< u, u >}$.
If $<,>$ is the Euclidean product, the norm is called the *standard, Euclidean, or 2-norm* and is also denoted as $||u||_2$.
Properties: for any $u, v \in E$ and $c \in \mathbb{R}$

1. $\|u\| \geq 0 \ \forall u$ and $\|u\| = 0 \Leftrightarrow u = 0$;
2. $\|cu\| = |c|\|u\| \ c \in \mathbb{R}$;
3. $| < u, v > | \leq \|u\|\|v\|$ (Cauchy-Schwarz inequality)
4. $\|u + v\| \leq \|u\| + \|v\|$ (triangular inequality);

Any function $f : E \longrightarrow \mathbb{R}$ that satisfies properties 1,2,4 is called a *norm* (and is not necessarily defined through a scalar product).

Norm and distance

Let E be an $\mathbb{R}$-e.v. with scalar product $<,>$. The **norm** of $u \in E$ is $\|u\| = \sqrt{<u,u>}$.
If $<,>$ is the Euclidean product, the norm is called the *standard, Euclidean, or 2-norm* and is also denoted as $||u||_2$.
Properties: for any $u, v \in E$ and $c \in \mathbb{R}$

1. $\|u\| \geq 0 \ \forall u$ and $\|u\| = 0 \Leftrightarrow u = 0$;
2. $\|cu\| = |c|\|u\| \ c \in \mathbb{R}$;
3. $|<u,v>| \leq \|u\|\|v\|$ (Cauchy-Schwarz inequality)
4. $\|u+v\| \leq \|u\| + \|v\|$ (triangular inequality);

Any function $f : E \longrightarrow \mathbb{R}$ that satisfies properties 1,2,4 is called a *norm* (and is not necessarily defined through a scalar product).

Norm and distance

Let E be an $\mathbb{R}$-e.v. with scalar product $<,>$. The **norm** of $u \in E$ is $\|u\| = \sqrt{< u, u >}$.
If $<,>$ is the Euclidean product, the norm is called the *standard, Euclidean, or 2-norm* and is also denoted as $||u||_2$.
Properties: for any $u, v \in E$ and $c \in \mathbb{R}$

1. $\|u\| \geq 0 \; \forall u$ and $\|u\| = 0 \Leftrightarrow u = 0$;
2. $\|cu\| = |c|\|u\| \; c \in \mathbb{R}$;
3. $| < u, v > | \leq \|u\|\|v\|$ (Cauchy-Schwarz inequality)
4. $\|u + v\| \leq \|u\| + \|v\|$ (triangular inequality);

Any function $f : E \longrightarrow \mathbb{R}$ that satisfies properties 1,2,4 is called a *norm* (and is not necessarily defined through a scalar product).

Norm and distance

Let E be an $\mathbb{R}$-e.v. with scalar product $<,>$. The **norm** of $u \in E$ is $\|u\| = \sqrt{< u, u >}$.
If $<,>$ is the Euclidean product, the norm is called the *standard, Euclidean, or 2-norm* and is also denoted as $||u||_2$.
Properties: for any $u, v \in E$ and $c \in \mathbb{R}$

1. $\|u\| \geq 0 \; \forall u$ and $\|u\| = 0 \Leftrightarrow u = 0$;
2. $\|cu\| = |c|\|u\| \; c \in \mathbb{R}$;
3. $| < u, v > | \leq \|u\|\|v\|$ (Cauchy-Schwarz inequality)
4. $\|u + v\| \leq \|u\| + \|v\|$ (triangular inequality);

Any function $f : E \longrightarrow \mathbb{R}$ that satisfies properties 1,2,4 is called a *norm* (and is not necessarily defined through a scalar product).

Other norms

If $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, one defines:

1. The **1-norm** (also called taxicab or Manhattan norm):

$$||x||_1 = |x_1| + \cdots + |x_n|.$$

2. The **maximum norm** (also called infinite norm):

$$||x||_\infty = \max(|x_1|, \ldots, |x_n|).$$

Other norms

If $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, one defines:

1. The **1-norm** (also called taxicab or Manhattan norm):

$$||x||_1 = |x_1| + \cdots + |\mathrm{x}_n|.$$

2. The **maximum norm** (also called infinite norm):

$$||x||_\infty = \max(|x_1|, \dots, |\mathrm{x}_n|).$$

Distances and Angles

Let E be an $\mathbb{R}$-e.v. with scalar product $<,>$.

- A vector u is called a **unit** vector if $\|u\| = 1$. Given a vector $v \neq 0$, we can always find a unit vector in its direction: $v/\|v\|$ (we say that we have **normalized** v).
- The **distance** between two vectors $u, v \in E$, is $d(u,v) = \|u - v\|$.
- The (unoriented) **angle** between two vectors $u \neq 0, v \neq 0 \in E$ is the unique $\alpha \in [0, \pi]$ such that $\cos(\alpha) = \frac{<u,v>}{\|u\| \cdot \|v\|}$ (the sign of $\widehat{uv}$ depends on the orientation we choose).
- Two vectors u, v are **orthogonal** (also denoted $u \perp v$) if $< u, v >= 0$.
- Two orthogonal vectors have $\widehat{uv} = \pm\frac{\pi}{2}$.
- If $u \perp v$ and $u, v \neq 0 \Rightarrow u, v$ are l.i.

Distances and Angles

Let E be an $\mathbb{R}$-e.v. with scalar product $<,>$.

- A vector u is called a **unit** vector if $\|u\| = 1$. Given a vector $v \neq 0$, we can always find a unit vector in its direction: $v/\|v\|$ (we say that we have **normalized** v).
- The **distance** between two vectors u, $v \in E$, is $d(u, v) = \|u - v\|$.
- The (unoriented) **angle** between two vectors $u \neq 0, v \neq 0 \in E$ is the unique $\alpha \in [0, \pi]$ such that $\cos(\alpha) = \frac{<u,v>}{\|u\| \cdot \|v\|}$ (the sign of $\widehat{uv}$ depends on the orientation we choose).
- Two vectors u, v are **orthogonal** (also denoted $u \perp v$) if $< u, v >= 0$.
- Two orthogonal vectors have $\widehat{uv} = \pm\frac{\pi}{2}$.
- If $u \perp v$ and $u, v \neq 0 \Rightarrow u, v$ are l.i.

Distances and Angles

Let E be an $\mathbb{R}$-e.v. with scalar product $<,>$.

- A vector u is called a **unit** vector if $\|u\| = 1$. Given a vector $v \neq 0$, we can always find a unit vector in its direction: $v/\|v\|$ (we say that we have **normalized** v).
- The **distance** between two vectors u, $v \in E$, is $d(u, v) = \|u - v\|$.
- The (unoriented) **angle** between two vectors $u \neq 0, v \neq 0 \in E$ is the unique $\alpha \in [0, \pi]$ such that $\cos(\alpha) = \frac{<u,v>}{||u||\cdot||v||}$ (the sign of $\widehat{uv}$ depends on the orientation we choose).
- Two vectors u, v are **orthogonal** (also denoted $u \perp v$) if $< u, v >= 0$.
- Two orthogonal vectors have $\widehat{uv} = \pm\frac{\pi}{2}$.
- If $u \perp v$ and $u, v \neq 0 \Rightarrow u, v$ are l.i.

Distances and Angles

Let E be an $\mathbb{R}$-e.v. with scalar product $<,>$.

- A vector u is called a **unit** vector if $\|u\| = 1$. Given a vector $v \neq 0$, we can always find a unit vector in its direction: $v/\|v\|$ (we say that we have **normalized** v).
- The **distance** between two vectors u, $v \in E$, is $d(u,v) = \|u - v\|$.
- The (unoriented) **angle** between two vectors $u \neq 0, v \neq 0 \in E$ is the unique $\alpha \in [0,\pi]$ such that $\cos(\alpha) = \frac{<u,v>}{||u||\cdot||v||}$ (the sign of $\widehat{uv}$ depends on the orientation we choose).
- Two vectors u, v are **orthogonal** (also denoted $u \perp v$) if $< u, v >= 0$.
- Two orthogonal vectors have $\widehat{uv} = \pm\frac{\pi}{2}$.
- If $u \perp v$ and $u, v \neq 0 \Rightarrow u, v$ are l.i.

Distances and Angles

Let E be an $\mathbb{R}$-e.v. with scalar product $<,>$.

- A vector u is called a **unit** vector if $\|u\| = 1$. Given a vector $v \neq 0$, we can always find a unit vector in its direction: $v/\|v\|$ (we say that we have **normalized** v).
- The **distance** between two vectors u, $v \in E$, is $d(u, v) = \|u - v\|$.
- The (unoriented) **angle** between two vectors $u \neq 0, v \neq 0 \in E$ is the unique $\alpha \in [0, \pi]$ such that $\cos(\alpha) = \frac{<u,v>}{||u|| \cdot ||v||}$ (the sign of $\widehat{uv}$ depends on the orientation we choose).
- Two vectors u, v are **orthogonal** (also denoted $u \perp v$) if $< u, v >= 0$.
- Two orthogonal vectors have $\widehat{uv} = \pm\frac{\pi}{2}$.
- If $u \perp v$ and $u, v \neq 0 \Rightarrow u, v$ are l.i.

Distances and Angles

Let E be an $\mathbb{R}$-e.v. with scalar product $<,>$.

- A vector u is called a **unit** vector if $\|u\| = 1$. Given a vector $v \neq 0$, we can always find a unit vector in its direction: $v/\|v\|$ (we say that we have **normalized** v).
- The **distance** between two vectors u, $v \in E$, is $d(u, v) = \|u - v\|$.
- The (unoriented) **angle** between two vectors $u \neq 0, v \neq 0 \in E$ is the unique $\alpha \in [0, \pi]$ such that $\cos(\alpha) = \frac{<u,v>}{||u|| \cdot ||v||}$ (the sign of $\widehat{uv}$ depends on the orientation we choose).
- Two vectors u, v are **orthogonal** (also denoted $u \perp v$) if $< u, v >= 0$.
- Two orthogonal vectors have $\widehat{uv} = \pm\frac{\pi}{2}$.
- If $u \perp v$ and $u, v \neq 0 \Rightarrow u, v$ are l.i.

Orthonormal basis

Definition

A basis $\{v_1, \ldots, v_d\}$ of a subspace $F \subseteq E$ is an **orthonormal basis** (b.o.n) if its vectors are

- pairwise orthogonal: $< v_i, v_j >= 0$ if $i \neq j$
- and normalized: $\|v_i\| = 1$ for $i = 1, 2, \ldots, d$.

- called **orthogonal** if pairwise orthogonal but not normalized.
- Ex: the standard basis is a b.o.n of $\mathbb{R}^n$ for Euclidean product.
- If $\mathbf{v} = \{v_1, \ldots, v_n\}$ is b.o.n. of $E \Rightarrow$ the coordinates of v in basis $\mathbf{v}$ are

$$(< v, v_1 >, \ldots, < v, v_n >).$$

- If $\mathbf{v}$ is b.o.n. of E and $\mathbf{u}$ basis of E then,

$$\mathbf{u} \text{ is b.o.n} \quad \Leftrightarrow \quad A^t_{\mathbf{u}\to\mathbf{v}} A_{\mathbf{u}\to\mathbf{v}} = I.$$

- For any basis $\mathbf{u}$ of E, $\exists$ scalar product such that $\mathbf{u}$ is b.o.n for it.

Orthonormal basis

Definition

A basis $\{v_1, \ldots, v_d\}$ of a subspace $F \subseteq E$ is an **orthonormal basis** (b.o.n) if its vectors are

- pairwise orthogonal: $< v_i, v_j >= 0$ if $i \neq j$
- and normalized: $\|v_i\| = 1$ for $i = 1, 2, \ldots, d$.

- called **orthogonal** if pairwise orthogonal but not normalized.
- Ex: the standard basis is a b.o.n of $\mathbb{R}^n$ for Euclidean product.
- If $\mathbf{v} = \{v_1, \ldots, v_n\}$ is b.o.n. of $E \Rightarrow$ the coordinates of v in basis $\mathbf{v}$ are

$$(< v, v_1 >, \ldots, < v, v_n >).$$

- If $\mathbf{v}$ is b.o.n. of E and $\mathbf{u}$ basis of E then,

$$\mathbf{u} \text{ is b.o.n} \quad \Leftrightarrow \quad A^t_{\mathbf{u}\rightarrow\mathbf{v}} A_{\mathbf{u}\rightarrow\mathbf{v}} = I.$$

- For any basis $\mathbf{u}$ of E, $\exists$ scalar product such that $\mathbf{u}$ is b.o.n for it.

Orthonormal basis

Definition

A basis $\{v_1, \ldots, v_d\}$ of a subspace $F \subseteq E$ is an **orthonormal basis** (b.o.n) if its vectors are

- pairwise orthogonal: $< v_i, v_j >= 0$ if $i \neq j$
- and normalized: $\|v_i\| = 1$ for $i = 1, 2, \ldots, d$.

- called **orthogonal** if pairwise orthogonal but not normalized.
- Ex: the standard basis is a b.o.n of $\mathbb{R}^n$ for Euclidean product.
- If $\mathbf{v} = \{v_1, \ldots, v_n\}$ is b.o.n. of $E \Rightarrow$ the coordinates of v in basis $\mathbf{v}$ are

$$(< v, v_1 >, \ldots, < v, v_n >).$$

- If $\mathbf{v}$ is b.o.n. of E and $\mathbf{u}$ basis of E then,

$$\mathbf{u} \text{ is b.o.n} \quad \Leftrightarrow \quad A^t_{\mathbf{u}\to\mathbf{v}} A_{\mathbf{u}\to\mathbf{v}} = I.$$

- For any basis $\mathbf{u}$ of E, $\exists$ scalar product such that $\mathbf{u}$ is b.o.n for it.

Orthonormal basis

Definition

A basis $\{v_1, \dots, v_d\}$ of a subspace $F \subseteq E$ is an **orthonormal basis** (b.o.n) if its vectors are

- pairwise orthogonal: $< v_i, v_j >= 0$ if $i \neq j$
- and normalized: $\|v_i\| = 1$ for $i = 1, 2, \dots, d$.

- called **orthogonal** if pairwise orthogonal but not normalized.
- Ex: the standard basis is a b.o.n of $\mathbb{R}^n$ for Euclidean product.
- If $\mathbf{v} = \{v_1, \dots, v_n\}$ is b.o.n. of $E \Rightarrow$ the coordinates of v in basis $\mathbf{v}$ are

$$(< v, v_1 >, \dots, < v, v_n >).$$

- If $\mathbf{v}$ is b.o.n. of E and $\mathbf{u}$ basis of E then,

$$\mathbf{u} \text{ is b.o.n} \quad \Leftrightarrow \quad A^t_{\mathbf{u}\to\mathbf{v}} A_{\mathbf{u}\to\mathbf{v}} = I.$$

- For any basis $\mathbf{u}$ of E, $\exists$ scalar product such that $\mathbf{u}$ is b.o.n for it.

Orthonormal basis

Definition

A basis $\{v_1, \ldots, v_d\}$ of a subspace $F \subseteq E$ is an **orthonormal basis** (b.o.n) if its vectors are

- pairwise orthogonal: $< v_i, v_j >= 0$ if $i \neq j$
- and normalized: $\|v_i\| = 1$ for $i = 1, 2, \ldots, d$.

- called **orthogonal** if pairwise orthogonal but not normalized.
- Ex: the standard basis is a b.o.n of $\mathbb{R}^n$ for Euclidean product.
- If $\mathbf{v} = \{v_1, \ldots, v_n\}$ is b.o.n. of $E \Rightarrow$ the coordinates of v in basis $\mathbf{v}$ are
$$(< v, v_1 >, \ldots, < v, v_n >).$$
- If $\mathbf{v}$ is b.o.n. of E and $\mathbf{u}$ basis of E then,
$$\mathbf{u} \text{ is b.o.n} \quad \Leftrightarrow \quad A^t_{\mathbf{u}\to\mathbf{v}} A_{\mathbf{u}\to\mathbf{v}} = I.$$
- For any basis $\mathbf{u}$ of E, $\exists$ scalar product such that $\mathbf{u}$ is b.o.n for it.

Orthonormal basis

Definition

A basis $\{v_1, \dots, v_d\}$ of a subspace $F \subseteq E$ is an **orthonormal basis** (b.o.n) if its vectors are

- pairwise orthogonal: $< v_i, v_j >= 0$ if $i \neq j$
- and normalized: $\|v_i\| = 1$ for $i = 1, 2, \dots, d$.

- called **orthogonal** if pairwise orthogonal but not normalized.
- Ex: the standard basis is a b.o.n of $\mathbb{R}^n$ for Euclidean product.
- If $\mathbf{v} = \{v_1, \dots, v_n\}$ is b.o.n. of $E \Rightarrow$ the coordinates of v in basis $\mathbf{v}$ are

$$(< v, v_1 >, \dots, < v, v_n >).$$

- If $\mathbf{v}$ is b.o.n. of E and $\mathbf{u}$ basis of E then,

$$\mathbf{u} \text{ is b.o.n} \quad \Leftrightarrow \quad A^t_{\mathbf{u}\to\mathbf{v}} A_{\mathbf{u}\to\mathbf{v}} = I.$$

- For any basis $\mathbf{u}$ of E, $\exists$ scalar product such that $\mathbf{u}$ is b.o.n for it.

Orthonormal basis

Definition

A basis $\{v_1, \ldots, v_d\}$ of a subspace $F \subseteq E$ is an **orthonormal basis** (b.o.n) if its vectors are

- pairwise orthogonal: $< v_i, v_j >= 0$ if $i \neq j$
- and normalized: $\|v_i\| = 1$ for $i = 1, 2, \ldots, d$.

- called **orthogonal** if pairwise orthogonal but not normalized.
- Ex: the standard basis is a b.o.n of $\mathbb{R}^n$ for Euclidean product.
- If $\mathbf{v} = \{v_1, \ldots, v_n\}$ is b.o.n. of $E \Rightarrow$ the coordinates of v in basis $\mathbf{v}$ are

$$(< v, v_1 >, \ldots, < v, v_n >).$$

- If $\mathbf{v}$ is b.o.n. of E and $\mathbf{u}$ basis of E then,

$$\mathbf{u} \text{ is b.o.n} \quad \Leftrightarrow \quad A^t_{\mathbf{u}\to\mathbf{v}} A_{\mathbf{u}\to\mathbf{v}} = I.$$

- For any basis $\mathbf{u}$ of E, $\exists$ scalar product such that $\mathbf{u}$ is b.o.n for it.

Orthonormal basis

Definition

A basis $\{v_1, \dots, v_d\}$ of a subspace $F \subseteq E$ is an **orthonormal basis** (b.o.n) if its vectors are

- pairwise orthogonal: $< v_i, v_j >= 0$ if $i \neq j$
- and normalized: $\|v_i\| = 1$ for $i = 1, 2, \dots, d$.

- called **orthogonal** if pairwise orthogonal but not normalized.
- Ex: the standard basis is a b.o.n of $\mathbb{R}^n$ for Euclidean product.
- If $\mathbf{v} = \{v_1, \dots, v_n\}$ is b.o.n. of $E \Rightarrow$ the coordinates of v in basis $\mathbf{v}$ are

$$(< v, v_1 >, \dots, < v, v_n >).$$

- If $\mathbf{v}$ is b.o.n. of E and $\mathbf{u}$ basis of E then,

$$\mathbf{u} \text{ is b.o.n} \quad \Leftrightarrow \quad A^t_{\mathbf{u}\to\mathbf{v}} A_{\mathbf{u}\to\mathbf{v}} = I.$$

- For any basis $\mathbf{u}$ of E, $\exists$ scalar product such that $\mathbf{u}$ is b.o.n for it.

Orthogonal matrices

An $n \times n$ matrix that satisfies $A^t A = I$ is called an **orthogonal** matrix.

- If we call the columns $u_1, \dots, u_n$, $A = \left(u_1 \dots u_n \right)$, then,
  A is orthogonal $\Leftrightarrow \{u_1, \dots, u_n\}$ is a b.o.n. for Euclidean scalar product.
- A is orthogonal $\Leftrightarrow A^{-1} = A^t$.
- A is orthogonal $\Leftrightarrow AA^t = I$.
- A is orthogonal $\Rightarrow \det A = \pm 1$.
- If A is orthogonal, then the corresponding endomorphism preserves the Euclidean scalar product:

$$< Au, Av >=< u, v > \quad \text{for all } u, v$$

- In particular, A preserves norms, angles $\Rightarrow$ does not deform objects.

Orthogonal matrices

An $n \times n$ matrix that satisfies $A^t A = I$ is called an **orthogonal** matrix.

- If we call the columns $u_1, \dots, u_n$, $A = \begin{pmatrix} u_1 \dots u_n \end{pmatrix}$, then,
  A is orthogonal $\Leftrightarrow \{u_1, \dots, u_n\}$ is a b.o.n. for Euclidean scalar product.
- A is orthogonal $\Leftrightarrow A^{-1} = A^t$.
- A is orthogonal $\Leftrightarrow AA^t = I$.
- A is orthogonal $\Rightarrow \det A = \pm 1$.
- If A is orthogonal, then the corresponding endomorphism preserves the Euclidean scalar product:

$$< Au, Av > = < u, v > \quad \text{for all } u, v$$

- In particular, A preserves norms, angles $\Rightarrow$ does not deform objects.

Orthogonal matrices

An $n \times n$ matrix that satisfies $A^t A = I$ is called an **orthogonal** matrix.

- If we call the columns $u_1, \ldots, u_n$, $A = \left(u_1 \ldots u_n\right)$, then,
  A is orthogonal $\Leftrightarrow \{u_1, \ldots, u_n\}$ is a b.o.n. for Euclidean scalar product.
- A is orthogonal $\Leftrightarrow A^{-1} = A^t$.
- A is orthogonal $\Leftrightarrow AA^t = I$.
- A is orthogonal $\Rightarrow \det A = \pm 1$.
- If A is orthogonal, then the corresponding endomorphism preserves the Euclidean scalar product:

$$< Au, Av >=< u, v > \text{ for all } u, v$$

- In particular, A preserves norms, angles $\Rightarrow$ does not deform objects.

Orthogonal matrices

An $n \times n$ matrix that satisfies $A^t A = I$ is called an **orthogonal** matrix.

- If we call the columns $u_1, \ldots, u_n$, $A = \begin{pmatrix} u_1 \ldots u_n \end{pmatrix}$, then,
  A is orthogonal $\Leftrightarrow \{u_1, \ldots, u_n\}$ is a b.o.n. for Euclidean scalar product.
- A is orthogonal $\Leftrightarrow A^{-1} = A^t$.
- A is orthogonal $\Leftrightarrow AA^t = I$.
- A is orthogonal $\Rightarrow \det A = \pm 1$.
- If A is orthogonal, then the corresponding endomorphism preserves the Euclidean scalar product:

$$< Au, Av > = < u, v > \quad \text{for all } u, v$$

- In particular, A preserves norms, angles $\Rightarrow$ does not deform objects.

Orthogonal matrices

An $n \times n$ matrix that satisfies $A^t A = I$ is called an **orthogonal** matrix.

- If we call the columns $u_1, \dots, u_n$, $A = \left(u_1 \dots u_n\right)$, then,

  A is orthogonal $\Leftrightarrow \{u_1, \dots, u_n\}$ is a b.o.n. for Euclidean scalar product.
- A is orthogonal $\Leftrightarrow A^{-1} = A^t$.
- A is orthogonal $\Leftrightarrow AA^t = I$.
- A is orthogonal $\Rightarrow \det A = \pm 1$.
- If A is orthogonal, then the corresponding endomorphism preserves the Euclidean scalar product:

$$< Au, Av > = < u, v > \text{ for all } u, v$$

- In particular, A preserves norms, angles $\Rightarrow$ does not deform objects.

Orthogonal matrices

An $n \times n$ matrix that satisfies $A^t A = I$ is called an **orthogonal** matrix.

- If we call the columns $u_1, \dots, u_n$, $A = \begin{pmatrix} u_1 \dots u_n \end{pmatrix}$, then,
  A is orthogonal $\Leftrightarrow \{u_1, \dots, u_n\}$ is a b.o.n. for Euclidean scalar product.
- A is orthogonal $\Leftrightarrow A^{-1} = A^t$.
- A is orthogonal $\Leftrightarrow AA^t = I$.
- A is orthogonal $\Rightarrow \det A = \pm 1$.
- If A is orthogonal, then the corresponding endomorphism preserves the Euclidean scalar product:

$$< Au, Av >=< u, v > \text{ for all } u, v$$

- In particular, A preserves norms, angles $\Rightarrow$ does not deform objects.

Examples of 2×2 orthogonal matrices

The following maps $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ are linear and preserve norms:

- $f=$ **symmetry** with respect to a line l passing through the origin, $l = [v]$. E.g. $f(x, y) = (x, -y)$.
- $f=$**rotation** counterclockwise of angle α with respect to the origin; then

$$M_e(f) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

Examples of 2×2 orthogonal matrices

The following maps $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ are linear and preserve norms:

- $f =$ **symmetry** with respect to a line l passing through the origin, $l = [v]$. E.g. $f(x, y) = (x, -y)$.
- $f =$**rotation** counterclockwise of angle α with respect to the origin; then

$$M_e(f) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

Gram-Schmidt algorithm

Given a subspace F of a euclidean space E, the following algorithm produces a b.o.n. of F:

1. Take any basis of F: $u_1, \dots, u_d$ and define:
2. $v_1 := u_1$
3. $v_2 := u_2 - \frac{\langle v_1,u_2 \rangle}{\langle v_1,v_1 \rangle} v_1 \quad (\Rightarrow [v_1, v_2] = [u_1, u_2])$.
4. $v_3 := u_3 - \frac{\langle v_1,u_3 \rangle}{\langle v_1,v_1 \rangle} v_1 - \frac{\langle v_2,u_3 \rangle}{\langle v_2,v_2 \rangle} v_2$
   $(\Rightarrow [v_1, v_2, v_3] = [u_1, u_2, u_3])$.

$$\vdots$$

5. $v_d := u_d - \frac{\langle v_1,u_d \rangle}{\langle v_1,v_1 \rangle} v_1 - \dots - \frac{\langle v_{d-1},u_d \rangle}{\langle v_{d-1},v_{d-1} \rangle} v_{d-1}$.
6. Then $v_1, \dots, v_d$ are orthogonal vectors and $[v_1, \dots, v_d] = [u_1, \dots, u_d]$.
7. We normalize to obtain the desired b.o.n. $w_1, \dots, w_d$:

$$w_1 = v_1/\|v_1\|,\ w_2 = v_2/\|v_2\|,\ \dots,\ w_d = v_d/\|v_d\|.$$

Warning: Roundoff errors make vectors not orthogonal! $\rightarrow$ modified Gram-Schmidt algorithm solves this issue

Gram-Schmidt algorithm

Given a subspace F of a euclidean space E, the following algorithm produces a b.o.n. of F:

1. Take any basis of F: $u_1, \ldots, u_d$ and define:
2. $v_1 := u_1$
3. $v_2 := u_2 - \frac{<v_1,u_2>}{<v_1,v_1>} v_1 \quad (\Rightarrow [v_1, v_2] = [u_1, u_2])$.
4. $v_3 := u_3 - \frac{<v_1,u_3>}{<v_1,v_1>} v_1 - \frac{<v_2,u_3>}{<v_2,v_2>} v_2$
   $(\Rightarrow [v_1, v_2, v_3] = [u_1, u_2, u_3])$.

$$\vdots$$

5. $v_d := u_d - \frac{<v_1,u_d>}{<v_1,v_1>} v_1 - \cdots - \frac{<v_{d-1},u_d>}{<v_{d-1},v_{d-1}>} v_{d-1}$.
6. Then $v_1, \ldots, v_d$ are orthogonal vectors and $[v_1, \ldots, v_d] = [u_1, \ldots, u_d]$.
7. We normalize to obtain the desired b.o.n. $w_1, \ldots, w_d$:

$$w_1 = v_1/\|v_1\|,\ w_2 = v_2/\|v_2\|,\ \ldots,\ w_d = v_d/\|v_d\|.$$

Warning: Roundoff errors make vectors not orthogonal! $\rightarrow$ modified Gram-Schmidt algorithm solves this issue

Gram-Schmidt algorithm

Given a subspace F of a euclidean space E, the following algorithm produces a b.o.n. of F:

1. Take any basis of F: $u_1, \ldots, u_d$ and define:
2. $v_1 := u_1$
3. $v_2 := u_2 - \frac{<v_1,u_2>}{<v_1,v_1>} v_1 \quad (\Rightarrow [v_1, v_2] = [u_1, u_2])$.
4. $v_3 := u_3 - \frac{<v_1,u_3>}{<v_1,v_1>} v_1 - \frac{<v_2,u_3>}{<v_2,v_2>} v_2$
   $(\Rightarrow [v_1, v_2, v_3] = [u_1, u_2, u_3])$.

$$\vdots$$

5. $v_d := u_d - \frac{<v_1,u_d>}{<v_1,v_1>} v_1 - \cdots - \frac{<v_{d-1},u_d>}{<v_{d-1},v_{d-1}>} v_{d-1}$.
6. Then $v_1, \ldots, v_d$ are orthogonal vectors and $[v_1, \ldots, v_d] = [u_1, \ldots, u_d]$.
7. We normalize to obtain the desired b.o.n. $w_1, \ldots, w_d$:

$$w_1 = v_1/\|v_1\|,\ w_2 = v_2/\|v_2\|,\ \ldots,\ w_d = v_d/\|v_d\|.$$

Warning: Roundoff errors make vectors not orthogonal! $\rightarrow$ modified Gram-Schmidt algorithm solves this issue

Gram-Schmidt algorithm

Given a subspace F of a euclidean space E, the following algorithm produces a b.o.n. of F:

1. Take any basis of F: $u_1, \ldots, u_d$ and define:
2. $v_1 := u_1$
3. $v_2 := u_2 - \frac{<v_1,u_2>}{<v_1,v_1>} v_1 \quad (\Rightarrow [v_1, v_2] = [u_1, u_2])$.
4. $v_3 := u_3 - \frac{<v_1,u_3>}{<v_1,v_1>} v_1 - \frac{<v_2,u_3>}{<v_2,v_2>} v_2$
   $(\Rightarrow [v_1, v_2, v_3] = [u_1, u_2, u_3])$.

$$\vdots$$

5. $v_d := u_d - \frac{<v_1,u_d>}{<v_1,v_1>} v_1 - \cdots - \frac{<v_{d-1},u_d>}{<v_{d-1},v_{d-1}>} v_{d-1}$.
6. Then $v_1, \ldots, v_d$ are orthogonal vectors and $[v_1, \ldots, v_d] = [u_1, \ldots, u_d]$.
7. We normalize to obtain the desired b.o.n. $w_1, \ldots, w_d$:

$$w_1 = v_1/\|v_1\|,\ w_2 = v_2/\|v_2\|,\ \ldots,\ w_d = v_d/\|v_d\|.$$

Warning: Roundoff errors make vectors not orthogonal! $\rightarrow$ modified Gram-Schmidt algorithm solves this issue

Gram-Schmidt algorithm

Given a subspace F of a euclidean space E, the following algorithm produces a b.o.n. of F:

1. Take any basis of F: $u_1, \ldots, u_d$ and define:
2. $v_1 := u_1$
3. $v_2 := u_2 - \frac{<v_1,u_2>}{<v_1,v_1>} v_1 \quad (\Rightarrow [v_1, v_2] = [u_1, u_2])$.
4. $v_3 := u_3 - \frac{<v_1,u_3>}{<v_1,v_1>} v_1 - \frac{<v_2,u_3>}{<v_2,v_2>} v_2$ $(\Rightarrow [v_1, v_2, v_3] = [u_1, u_2, u_3])$.

$$\vdots$$

5. $v_d := u_d - \frac{<v_1,u_d>}{<v_1,v_1>} v_1 - \cdots - \frac{<v_{d-1},u_d>}{<v_{d-1},v_{d-1}>} v_{d-1}$.
6. Then $v_1, \ldots, v_d$ are orthogonal vectors and $[v_1, \ldots, v_d] = [u_1, \ldots, u_d]$.
7. We normalize to obtain the desired b.o.n. $w_1, \ldots, w_d$:

$$w_1 = v_1/\|v_1\|,\ w_2 = v_2/\|v_2\|,\ \ldots,\ w_d = v_d/\|v_d\|.$$

Warning: Roundoff errors make vectors not orthogonal! → modified Gram-Schmidt algorithm solves this issue

Gram-Schmidt algorithm

Given a subspace F of a euclidean space E, the following algorithm produces a b.o.n. of F:

1. Take any basis of F: $u_1, \ldots, u_d$ and define:
2. $v_1 := u_1$
3. $v_2 := u_2 - \frac{<v_1,u_2>}{<v_1,v_1>} v_1 \quad (\Rightarrow [v_1, v_2] = [u_1, u_2])$.
4. $v_3 := u_3 - \frac{<v_1,u_3>}{<v_1,v_1>} v_1 - \frac{<v_2,u_3>}{<v_2,v_2>} v_2$
   $(\Rightarrow [v_1, v_2, v_3] = [u_1, u_2, u_3])$.

$$\vdots$$

5. $v_d := u_d - \frac{<v_1,u_d>}{<v_1,v_1>} v_1 - \cdots - \frac{<v_{d-1},u_d>}{<v_{d-1},v_{d-1}>} v_{d-1}$.
6. Then $v_1, \ldots, v_d$ are orthogonal vectors and
   $[v_1, \ldots, v_d] = [u_1, \ldots, u_d]$.
7. We normalize to obtain the desired b.o.n. $w_1, \ldots, w_d$:

$$w_1 = v_1/\|v_1\|,\ w_2 = v_2/\|v_2\|,\ \ldots,\ w_d = v_d/\|v_d\|.$$

Warning: Roundoff errors make vectors not orthogonal! $\rightarrow$ modified Gram-Schmidt algorithm solves this issue

Gram-Schmidt algorithm

Given a subspace F of a euclidean space E, the following algorithm produces a b.o.n. of F:

1. Take any basis of F: $u_1, \ldots, u_d$ and define:
2. $v_1 := u_1$
3. $v_2 := u_2 - \frac{<v_1,u_2>}{<v_1,v_1>} v_1 \quad (\Rightarrow [v_1, v_2] = [u_1, u_2])$.
4. $v_3 := u_3 - \frac{<v_1,u_3>}{<v_1,v_1>} v_1 - \frac{<v_2,u_3>}{<v_2,v_2>} v_2$
   $(\Rightarrow [v_1, v_2, v_3] = [u_1, u_2, u_3])$.

$$\vdots$$

5. $v_d := u_d - \frac{<v_1,u_d>}{<v_1,v_1>} v_1 - \cdots - \frac{<v_{d-1},u_d>}{<v_{d-1},v_{d-1}>} v_{d-1}$.
6. Then $v_1, \ldots, v_d$ are orthogonal vectors and $[v_1, \ldots, v_d] = [u_1, \ldots, u_d]$.
7. We normalize to obtain the desired b.o.n. $w_1, \ldots, w_d$:

$$w_1 = v_1/\|v_1\|,\ w_2 = v_2/\|v_2\|,\ \ldots,\ w_d = v_d/\|v_d\|.$$

Warning: Roundoff errors make vectors not orthogonal! $\rightarrow$ modified Gram-Schmidt algorithm solves this issue

Gram-Schmidt algorithm

Given a subspace F of a euclidean space E, the following algorithm produces a b.o.n. of F:

1. Take any basis of F: $u_1, \dots, u_d$ and define:
2. $v_1 := u_1$
3. $v_2 := u_2 - \frac{<v_1,u_2>}{<v_1,v_1>} v_1 \quad (\Rightarrow [v_1, v_2] = [u_1, u_2])$.
4. $v_3 := u_3 - \frac{<v_1,u_3>}{<v_1,v_1>} v_1 - \frac{<v_2,u_3>}{<v_2,v_2>} v_2$
   $(\Rightarrow [v_1, v_2, v_3] = [u_1, u_2, u_3])$.

$$\vdots$$

5. $v_d := u_d - \frac{<v_1,u_d>}{<v_1,v_1>} v_1 - \dots - \frac{<v_{d-1},u_d>}{<v_{d-1},v_{d-1}>} v_{d-1}$.
6. Then $v_1, \dots, v_d$ are orthogonal vectors and $[v_1, \dots, v_d] = [u_1, \dots, u_d]$.
7. We normalize to obtain the desired b.o.n. $w_1, \dots, w_d$:

$$w_1 = v_1/\|v_1\|,\ w_2 = v_2/\|v_2\|,\ \dots,\ w_d = v_d/\|v_d\|.$$

Warning: Roundoff errors make vectors not orthogonal! $\rightarrow$ modified Gram-Schmidt algorithm solves this issue

Outline

Spectral theorem

Theorem (Spectral Theorem)

Let A be a symmetric $n \times n$ matrix. Then A has real eigenvalues, diagonalizes, and there exists an orthonormal basis $\{v_1, \dots, v_n\}$ of eigenvectors (in the Euclidean product); if V has columns $v_1, \dots, v_n$, and D is the diagonal matrix of eigenvalues (in the corresponding order) then A decomposes as

$$A = VDV^t.$$

The orthonormal basis of eigenvectors is not difficult to find:

- If u, v are eigenvectors of A of eigenvalues $\lambda \neq \mu$, then $u \perp v$.
- If the eigenvalues are all distinct, then normalizing the eigenvectors we obtain an orthonormal basis of eigenvectors.
- If the eigenvalues are not all distinct, we have to use Gram-Schmidt algorithm on each subspace of eigenvectors.

Spectral theorem

Theorem (Spectral Theorem)

Let A be a symmetric $n \times n$ matrix. Then A has real eigenvalues, diagonalizes, and there exists an orthonormal basis $\{v_1, \ldots, v_n\}$ of eigenvectors (in the Euclidean product); if V has columns $v_1, \ldots, v_n$, and D is the diagonal matrix of eigenvalues (in the corresponding order) then A decomposes as

$$A = VDV^t.$$

The orthonormal basis of eigenvectors is not difficult to find:

- If u, v are eigenvectors of A of eigenvalues $\lambda \neq \mu$, then $u \perp v$.
- If the eigenvalues are all distinct, then normalizing the eigenvectors we obtain an orthonormal basis of eigenvectors.
- If the eigenvalues are not all distinct, we have to use Gram-Schmidt algorithm on each subspace of eigenvectors.

Spectral theorem

Theorem (Spectral Theorem)

Let A be a symmetric $n \times n$ matrix. Then A has real eigenvalues, diagonalizes, and there exists an orthonormal basis $\{v_1, \ldots, v_n\}$ of eigenvectors (in the Euclidean product); if V has columns $v_1, \ldots, v_n$, and D is the diagonal matrix of eigenvalues (in the corresponding order) then A decomposes as

$$A = VDV^t.$$

The orthonormal basis of eigenvectors is not difficult to find:

- If u, v are eigenvectors of A of eigenvalues $\lambda \neq \mu$, then $u \perp v$.
- If the eigenvalues are all distinct, then normalizing the eigenvectors we obtain an orthonormal basis of eigenvectors.
- If the eigenvalues are not all distinct, we have to use Gram-Schmidt algorithm on each subspace of eigenvectors.

Spectral theorem

Theorem (Spectral Theorem)

Let A be a symmetric $n \times n$ matrix. Then A has real eigenvalues, diagonalizes, and there exists an orthonormal basis $\{v_1, \dots, v_n\}$ of eigenvectors (in the Euclidean product); if V has columns $v_1, \dots, v_n$, and D is the diagonal matrix of eigenvalues (in the corresponding order) then A decomposes as

$$A = VDV^t.$$

The orthonormal basis of eigenvectors is not difficult to find:

- If u, v are eigenvectors of A of eigenvalues $\lambda \neq \mu$, then $u \perp v$.
- If the eigenvalues are all distinct, then normalizing the eigenvectors we obtain an orthonormal basis of eigenvectors.
- If the eigenvalues are not all distinct, we have to use Gram-Schmidt algorithm on each subspace of eigenvectors.

Characterization of scalar products

Let A be a symmetric matrix.

- A is the matrix of a scalar product if and only if all eigenvalues of A are positive.
- Sylvester criterion: if δ_1, $\delta_2, \ldots, \delta_n$ are the principal minors of A, A is the matrix of a scalar product if and only if $\delta_i > 0 \; \forall i$.

Characterization of scalar products

Let A be a symmetric matrix.

- A is the matrix of a scalar product if and only if all eigenvalues of A are positive.
- Sylvester criterion: if δ_1, δ_2,. . . , δ_n are the principal minors of A, A is the matrix of a scalar product if and only if $\delta_i > 0 \ \forall i$.

Outline

Cross-product in $\mathbb{R}^3$

The **cross-product** between two vectors $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3)$ of $\mathbb{R}^3$ is the following vector (in standard basis)

$$
\begin{aligned}
(u_1, u_2, u_3) \times (v_1, v_2, v_3) &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \\
&= (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1).
\end{aligned}
$$

Main properties:

- bilineal
- $v \times u = -u \times v$ (anti-commutative)
- $u \times v$ is orthogonal to both u and v
- $\langle u \times v, w \rangle = \det(u, v, w)$
- $\|u \times v\| = \|u\| \cdot \|v\| \cdot |\sin(\widehat{uv})|$
- $u \times v = 0 \Leftrightarrow u, v$ linearly dependent
- If u, v are orthogonal and normalized $\Rightarrow u, v, u \times v$ is b.o.n.

Cross-product in $\mathbb{R}^3$

The **cross-product** between two vectors $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3)$ of $\mathbb{R}^3$ is the following vector (in standard basis)

$$\begin{aligned}(u_1, u_2, u_3) \times (v_1, v_2, v_3) &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \\ &= (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1).\end{aligned}$$

Main properties:

- bilineal
- $v \times u = -u \times v$ (anti-commutative)
- $u \times v$ is orthogonal to both u and v
- $\langle u \times v, w \rangle = \det(u, v, w)$
- $\|u \times v\| = \|u\| \cdot \|v\| \cdot |\sin(\widehat{uv})|$
- $u \times v = 0 \Leftrightarrow u, v$ linearly dependent
- If u, v are orthogonal and normalized $\Rightarrow u, v, u \times v$ is b.o.n.

Cross-product in $\mathbb{R}^3$

The **cross-product** between two vectors $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3)$ of $\mathbb{R}^3$ is the following vector (in standard basis)

$$\begin{aligned}(u_1, u_2, u_3) \times (v_1, v_2, v_3) &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \\ &= (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1).\end{aligned}$$

Main properties:

- bilineal
- $v \times u = -u \times v$ (anti-commutative)
- $u \times v$ is orthogonal to both u and v
- $\langle u \times v, w \rangle = \det(u, v, w)$
- $\|u \times v\| = \|u\| \cdot \|v\| \cdot |\sin(\widehat{uv})|$
- $u \times v = 0 \Leftrightarrow u, v$ linearly dependent
- If u, v are orthogonal and normalized $\Rightarrow u, v, u \times v$ is b.o.n.

Cross-product in $\mathbb{R}^3$

The **cross-product** between two vectors $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3)$ of $\mathbb{R}^3$ is the following vector (in standard basis)

$$\begin{aligned}(u_1, u_2, u_3) \times (v_1, v_2, v_3) &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \\ &= (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1).\end{aligned}$$

Main properties:

- bilineal
- $v \times u = -u \times v$ (anti-commutative)
- $u \times v$ is orthogonal to both u and v
- $\langle u \times v, w \rangle = \det(u, v, w)$
- $\|u \times v\| = \|u\| \cdot \|v\| \cdot |\sin(\widehat{uv})|$
- $u \times v = 0 \Leftrightarrow u, v$ linearly dependent
- If u, v are orthogonal and normalized $\Rightarrow u, v, u \times v$ is b.o.n.

Cross-product in $\mathbb{R}^3$

The **cross-product** between two vectors $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3)$ of $\mathbb{R}^3$ is the following vector (in standard basis)

$$(u_1, u_2, u_3) \times (v_1, v_2, v_3) = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1).$$

Main properties:

- bilineal
- $v \times u = -u \times v$ (anti-commutative)
- $u \times v$ is orthogonal to both u and v
- $\langle u \times v, w \rangle = \det(u, v, w)$
- $\|u \times v\| = \|u\| \cdot \|v\| \cdot |\sin(\widehat{uv})|$
- $u \times v = 0 \Leftrightarrow u, v$ linearly dependent
- If u, v are orthogonal and normalized $\Rightarrow u, v, u \times v$ is b.o.n.

Cross-product in $\mathbb{R}^3$

The **cross-product** between two vectors $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3)$ of $\mathbb{R}^3$ is the following vector (in standard basis)

$$(u_1, u_2, u_3) \times (v_1, v_2, v_3) = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1).$$

Main properties:

- bilineal
- $v \times u = -u \times v$ (anti-commutative)
- $u \times v$ is orthogonal to both u and v
- $\langle u \times v, w \rangle = \det(u, v, w)$
- $\|u \times v\| = \|u\| \cdot \|v\| \cdot |\sin(\widehat{uv})|$
- $u \times v = 0 \Leftrightarrow u, v$ linearly dependent
- If u, v are orthogonal and normalized $\Rightarrow u, v, u \times v$ is b.o.n.

Cross-product in $\mathbb{R}^3$

The **cross-product** between two vectors $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3)$ of $\mathbb{R}^3$ is the following vector (in standard basis)

$$\begin{aligned}(u_1, u_2, u_3) \times (v_1, v_2, v_3) &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \\ &= (u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1).\end{aligned}$$

Main properties:

- bilineal
- $v \times u = -u \times v$ (anti-commutative)
- $u \times v$ is orthogonal to both u and v
- $\langle u \times v, w \rangle = \det(u, v, w)$
- $\|u \times v\| = \|u\| \cdot \|v\| \cdot |\sin(\widehat{uv})|$
- $u \times v = 0 \Leftrightarrow u, v$ linearly dependent
- If u, v are orthogonal and normalized $\Rightarrow u, v, u \times v$ is b.o.n.

Outline

Orthogonal complement

The **orthogonal complement** to a given subspace F of a Euclidean space E is the subspace

$$F^{\perp} = \{u \in E \,|\, u \perp v \text{ for all } v \in F\}.$$

Properties when E has finite dimension:

- If $F = [v_1, \ldots, v_d] \Rightarrow F^{\perp} = \left\{ u \in E \;\middle|\; \begin{matrix} < u, v_1 >= 0 \\ \vdots \\ < u, v_d >= 0 \end{matrix} \right\}$
- $(F^{\perp})^{\perp} = F$, $\quad F \subseteq G \Leftrightarrow G^{\perp} \subseteq F^{\perp}$,
- $(F + G)^{\perp} = F^{\perp} \cap G^{\perp}$, $\quad (F \cap G)^{\perp} = F^{\perp} + G^{\perp}$.
- $F \cap F^{\perp} = \{\mathbf{0}\}$.

Orthogonal complement

The **orthogonal complement** to a given subspace F of a Euclidean space E is the subspace

$$F^\perp = \{u \in E \,|\, u \perp v \text{ for all } v \in F\}.$$

Properties when E has finite dimension:

- If $F = [v_1, \dots, v_d] \Rightarrow F^\perp = \left\{ u \in E \;\middle|\; \begin{array}{c} < u, v_1 >= 0 \\ \vdots \\ < u, v_d >= 0 \end{array} \right\}$
- $(F^\perp)^\perp = F$, $\qquad F \subseteq G \Leftrightarrow G^\perp \subseteq F^\perp$,
- $(F+G)^\perp = F^\perp \cap G^\perp$, $\qquad (F \cap G)^\perp = F^\perp + G^\perp$.
- $F \cap F^\perp = \{\mathbf{0}\}$.

Orthogonal complement

The **orthogonal complement** to a given subspace F of a Euclidean space E is the subspace

$$F^{\perp} = \{u \in E \,|\, u \perp v \text{ for all } v \in F\}.$$

Properties when E has finite dimension:

- If $F = [v_1, \ldots, v_d] \Rightarrow F^{\perp} = \left\{ u \in E \;\middle|\; \begin{array}{c} < u, v_1 >= 0 \\ \vdots \\ < u, v_d >= 0 \end{array} \right\}$
- $(F^{\perp})^{\perp} = F$, $\qquad F \subseteq G \Leftrightarrow G^{\perp} \subseteq F^{\perp}$,
- $(F + G)^{\perp} = F^{\perp} \cap G^{\perp}$, $\qquad (F \cap G)^{\perp} = F^{\perp} + G^{\perp}$.
- $F \cap F^{\perp} = \{\mathbf{0}\}$.

Orthogonal complement

The **orthogonal complement** to a given subspace F of a Euclidean space E is the subspace

$$F^{\perp} = \{u \in E \,|\, u \perp v \text{ for all } v \in F\}.$$

Properties when E has finite dimension:

- If $F = [v_1, \dots, v_d] \Rightarrow F^{\perp} = \left\{ u \in E \;\middle|\; \begin{array}{c} < u, v_1 >= 0 \\ \vdots \\ < u, v_d >= 0 \end{array} \right\}$
- $(F^{\perp})^{\perp} = F$, $\quad F \subseteq G \Leftrightarrow G^{\perp} \subseteq F^{\perp}$,
- $(F + G)^{\perp} = F^{\perp} \cap G^{\perp}$, $\quad (F \cap G)^{\perp} = F^{\perp} + G^{\perp}$.
- $F \cap F^{\perp} = \{\mathbf{0}\}$.

In $\mathbb{R}^n$ with the Euclidean scalar product,

- If F is defined by generators $\Rightarrow$ the equations of $F^\perp$ are easy to get: their coefficients are the generators coordinates.
- If F is given by equations $\Rightarrow$ the generators of $F^\perp$ are easy to get: their coordinates are the coefficients of the equations.

F	$F^\perp$
$[(1,3,2),(-2,1,8)]$	$\begin{cases} x+3y+2z=0 \\ -2x+y+8z=0 \end{cases}$
$3x-5y+\frac{11}{2}z=0$	$[(3,-5,\frac{11}{2})]$

- If A is a real matrix, then

$$\mathrm{Nuc}(A)^\perp = \mathrm{Im}(A^t).$$

In $\mathbb{R}^n$ with the Euclidean scalar product,

- If F is defined by generators $\Rightarrow$ the equations of $F^\perp$ are easy to get: their coefficients are the generators coordinates.
- If F is given by equations $\Rightarrow$ the generators of $F^\perp$ are easy to get: their coordinates are the coefficients of the equations.

F	$F^\perp$
$[(1,3,2),(-2,1,8)]$	$\begin{cases} x+3y+2z=0 \\ -2x+y+8z=0 \end{cases}$
$3x-5y+\frac{11}{2}z=0$	$[(3,-5,\frac{11}{2})]$

- If A is a real matrix, then

$$\text{Nuc}(A)^\perp = \text{Im}(A^t).$$

In $\mathbb{R}^n$ with the Euclidean scalar product,

- If F is defined by generators $\Rightarrow$ the equations of $F^\perp$ are easy to get: their coefficients are the generators coordinates.
- If F is given by equations $\Rightarrow$ the generators of $F^\perp$ are easy to get: their coordinates are the coefficients of the equations.

F	$F^\perp$
$[(1,3,2),(-2,1,8)]$	$\begin{cases} x+3y+2z=0 \\ -2x+y+8z=0 \end{cases}$
$3x-5y+\frac{11}{2}z=0$	$[(3,-5,\frac{11}{2})]$

- If A is a real matrix, then

$$\mathrm{Nuc}(A)^\perp = \mathrm{Im}(A^t).$$

Outline

Orthogonal projection

Let E be a Euclidean space of dimension n.

Theorem (Orthogonal Decomposition)

$E = F \oplus F^{\perp}$ for any subspace F. This is, any $v \in E$ can be written in a unique way as $v = w + w'$ where $w \in F$ and $w' \in F^{\perp}$.

- w is called the *orthogonal projection* of v on F and is denoted as $proj_F(v)$,
- w' is called the *orthogonal projection* of v on $F^{\perp}$ and is denoted as $proj_{F^{\perp}}(v)$.
- Thus, $v = proj_F(v) + proj_{F^{\perp}}(v)$ and $proj_F(v)$ is the unique vector of F such that $v - proj_F(v)$ belongs to $F^{\perp}$.
- If $F \subseteq E$ has dimension $d \Rightarrow F^{\perp}$ has dimension $n - d$.

Orthogonal projection

Let E be a Euclidean space of dimension n.

Theorem (Orthogonal Decomposition)

$E = F \oplus F^{\perp}$ for any subspace F. This is, any $v \in E$ can be written in a unique way as $v = w + w'$ where $w \in F$ and $w' \in F^{\perp}$.

- w is called the *orthogonal projection* of v on F and is denoted as $proj_F(v)$,
- w' is called the *orthogonal projection* of v on $F^{\perp}$ and is denoted as $proj_{F^{\perp}}(v)$.
- Thus, $v = proj_F(v) + proj_{F^{\perp}}(v)$ and $proj_F(v)$ is the unique vector of F such that $v - proj_F(v)$ belongs to $F^{\perp}$.
- If $F \subseteq E$ has dimension $d \Rightarrow F^{\perp}$ has dimension $n - d$.

Orthogonal projection

Let E be a Euclidean space of dimension n.

Theorem (Orthogonal Decomposition)

$E = F \oplus F^{\perp}$ for any subspace F. This is, any $v \in E$ can be written in a unique way as $v = w + w'$ where $w \in F$ and $w' \in F^{\perp}$.

- w is called the *orthogonal projection* of v on F and is denoted as $proj_F(v)$,
- w' is called the *orthogonal projection* of v on $F^{\perp}$ and is denoted as $proj_{F^{\perp}}(v)$.
- Thus, $v = proj_F(v) + proj_{F^{\perp}}(v)$ and $proj_F(v)$ is the unique vector of F such that $v - proj_F(v)$ belongs to $F^{\perp}$.
- If $F \subseteq E$ has dimension $d \Rightarrow F^{\perp}$ has dimension $n - d$.

Orthogonal projection

Let E be a Euclidean space of dimension n.

Theorem (Orthogonal Decomposition)

$E = F \oplus F^{\perp}$ for any subspace F. This is, any $v \in E$ can be written in a unique way as $v = w + w'$ where $w \in F$ and $w' \in F^{\perp}$.

- w is called the *orthogonal projection* of v on F and is denoted as $proj_F(v)$,
- w' is called the *orthogonal projection* of v on $F^{\perp}$ and is denoted as $proj_{F^{\perp}}(v)$.
- Thus, $v = proj_F(v) + proj_{F^{\perp}}(v)$ and $proj_F(v)$ is the unique vector of F such that $v - proj_F(v)$ belongs to $F^{\perp}$.
- If $F \subseteq E$ has dimension $d \Rightarrow F^{\perp}$ has dimension $n - d$.

Orthogonal projection

Let E be a Euclidean space of dimension n.

Theorem (Orthogonal Decomposition)

$E = F \oplus F^{\perp}$ for any subspace F. This is, any $v \in E$ can be written in a unique way as $v = w + w'$ where $w \in F$ and $w' \in F^{\perp}$.

- w is called the *orthogonal projection* of v on F and is denoted as $proj_F(v)$,
- w' is called the *orthogonal projection* of v on $F^{\perp}$ and is denoted as $proj_{F^{\perp}}(v)$.
- Thus, $v = proj_F(v) + proj_{F^{\perp}}(v)$ and $proj_F(v)$ is the unique vector of F such that $v - proj_F(v)$ belongs to $F^{\perp}$.
- If $F \subseteq E$ has dimension $d \Rightarrow F^{\perp}$ has dimension $n - d$.

Geometric interpretation

Proposition

The orthogonal projection of v on F is the vector of F that is closest to v; this is,

$$\|v - proj_F(v)\| = \min_{w \in F}\{\|v - w\|\}$$

(and this equals $\|proj_{F^\perp}(v)\|$). The orthogonal projection $proj_F(v)$ is the best approximation to v in F.

Computation of the orthogonal projection

Proposition

$proj_F(v)$ is the unique vector w that satisfies $w \in F$ and $v - w \in F^{\perp}$. If F has basis $u_1, \ldots, u_d$, then $proj_F(v)$ is the unique vector w such that

$$w = c_1 u_1 + \ldots c_d u_d \in F \quad and \left\{ \begin{array}{c} < u_1, w > = < u_1, v > \\ \vdots \\ < u_d, w > = < u_d, v > \end{array} \right.$$

Thus, $proj_F(v)$ is the vector $c_1 u_1 + \cdots + c_d u_d$ such that $c_1, \ldots, c_d$ are solution to the system

$$\begin{pmatrix} < u_1, u_1 > & \ldots & < u_1, u_d > \\ \vdots & \vdots & \vdots \\ < u_d, u_1 > & \ldots & < u_d, u_d > \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_d \end{pmatrix} = \begin{pmatrix} < u_1, v > \\ \vdots \\ < u_d, v > \end{pmatrix}$$

Orthogonal projection with orthogonal basis

Corollary

If $\dim F = 1$, $F = [u]$, *then* $proj_F(v) = \frac{<v,u>}{<u,u>}\ u$.

Proposition

If $u_1, \ldots, u_d$ *is an orthogonal basis of* F *and* $v \in \mathbb{R}^n$, *then*

$$proj_F(v) = \frac{< v, u_1 >}{< u_1, u_1 >} u_1 + \cdots + \frac{< v, u_d >}{< u_d, u_d >} u_d.$$

Proposition

If $u_1, \ldots, u_d$ *is an b.o.n. of* F *and* $v \in \mathbb{R}^n$, *then*

$$proj_F(v) =< v, u_1 > u_1 + \cdots + < v, u_d > u_d.$$

That is, the coordinates of $proj_F(v)$ *in the basis* $u_1, \ldots, u_d$ *are* $< v, u_1 >, \ldots, < v, u_d >$.

Orthogonal projection with orthogonal basis

Corollary

If $\dim F = 1$, $F = [u]$*, then* $proj_F(v) = \frac{<v,u>}{<u,u>}\ u$.

Proposition

If $u_1, \ldots, u_d$ *is an orthogonal basis of* F *and* $v \in \mathbb{R}^n$*, then*

$$proj_F(v) = \frac{< v, u_1 >}{< u_1, u_1 >} u_1 + \cdots + \frac{< v, u_d >}{< u_d, u_d >} u_d .$$

Proposition

If $u_1, \ldots, u_d$ *is an b.o.n. of* F *and* $v \in \mathbb{R}^n$*, then*

$$proj_F(v) = < v, u_1 > u_1 + \cdots + < v, u_d > u_d .$$

That is, the coordinates of $proj_F(v)$ *in the basis* $u_1, \ldots, u_d$ *are* $< v, u_1 >, \ldots, < v, u_d >$.

Orthogonal projection with orthogonal basis

Corollary

If $\dim F = 1$, $F = [u]$, *then* $proj_F(v) = \frac{<v,u>}{<u,u>}\ u.$

Proposition

If $u_1, \ldots, u_d$ *is an orthogonal basis of* F *and* $v \in \mathbb{R}^n$, *then*

$$proj_F(v) = \frac{< v, u_1 >}{< u_1, u_1 >} u_1 + \cdots + \frac{< v, u_d >}{< u_d, u_d >} u_d.$$

Proposition

If $u_1, \ldots, u_d$ *is an b.o.n. of* F *and* $v \in \mathbb{R}^n$, *then*

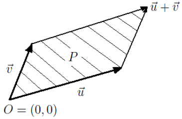
$$proj_F(v) =< v, u_1 > u_1 + \cdots + < v, u_d > u_d.$$

That is, the coordinates of $proj_F(v)$ *in the basis* $u_1, \ldots, u_d$ *are* $< v, u_1 >, \ldots, < v, u_d >$.

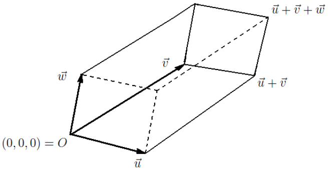
Determinants and volumes

From orthogonal projection and properties of cross product we can prove:

- In $\mathbb{R}^2$, the parallelogram determined by two vectors u, v has area equal to $|\det(u, v)|$.



- In $\mathbb{R}^3$, the parallelepiped determined by three vectors u, v, w has volume equal to $|\det(u, v, w)|$.



Outline

Singular value decomposition (SVD)

Theorem (Singular value decomposition)

Let A be a real $m \times n$ matrix. There there exists a decomposition $A = U \cdot D \cdot V^t$, where U is $m \times m$, V is $n \times n$, U, V are orthogonal and D is the following $m \times n$ matrix

$$D = \begin{pmatrix} \sigma_1 & & 0 & \cdots & 0 \\ & \ddots & & & \vdots \\ 0 & & \sigma_r & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

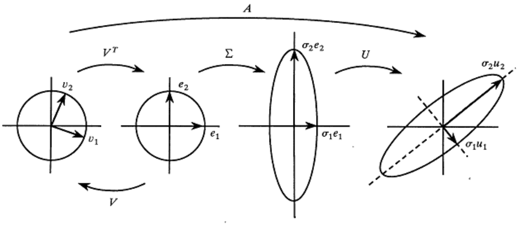
with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ and $r = \text{rank } A$.
$\sigma_1, \ldots, \sigma_r$ are called **singular values** *of A and are uniquely determined by A.*

Geometric interpretation of the SVD

If A is the standard matrix of a linear map $f : \mathbb{R}^n \to \mathbb{R}^m$, and we call $u_1, \ldots, u_m$, $v_1, \ldots, v_n$, the columns of U and V respectively, then D the matrix associated to f in orthonormal basis $v_1, \ldots, v_n$ and $u_1, \ldots, u_m$:

$$A = M_e(f) = \underbrace{U}_{A_{u\to e}} * \underbrace{D}_{M_{v,u}(f)} * \underbrace{V^t}_{A_{e\to v}}$$

(note that $V^t = V^{-1} = A_{e\to v}$).



[T. Hern, C. Long]

How to get the SVD?

The singular values are determined by the matrix A:

$$A = UDV^t \Rightarrow A^tA = VD^tU^tUDV^t = VD^tDV^t$$

but U and V are not (almost determined in most cases). How do we compute the SVD?

(1) Diagonalize the symmetric matrix $S = A^t \cdot A$

(2) If $\lambda_1 \geq \cdots \geq \lambda_r$ are the non-zero eigenvalues of $S \Rightarrow$ the **singular values** are $\sigma_1 = \sqrt{\lambda_1}, \ldots \sigma_r = \sqrt{\lambda_r}$ (fact: A^tA always has non-negative eigenvalues).

(3) The columns of V are an orthonormal basis $v_1, \ldots, v_n$ of eigenvectors of S.

(4) $u_1 = \frac{1}{\sigma_1}Av_1, \ldots, u_r = \frac{1}{\sigma_r}Av_r$ are orthonormal vectors in $\mathbb{R}^m$ (which can be completed to an orthonormal basis of $\mathbb{R}^m$ if necessary) and they form the columns of U.

How to get the SVD?

The singular values are determined by the matrix A:

$$A = UDV^t \Rightarrow A^t A = VD^t U^t UDV^t = VD^t DV^t$$

but U and V are not (almost determined in most cases). How do we compute the SVD?

(1) Diagonalize the symmetric matrix $S = A^t \cdot A$

(2) If $\lambda_1 \geq \cdots \geq \lambda_r$ are the non-zero eigenvalues of $S \Rightarrow$ the **singular values** are $\sigma_1 = \sqrt{\lambda_1}, \ldots \sigma_r = \sqrt{\lambda_r}$ (fact: $A^t A$ always has non-negative eigenvalues).

(3) The columns of V are an orthonormal basis $v_1, \ldots, v_n$ of eigenvectors of S.

(4) $u_1 = \frac{1}{\sigma_1} A v_1, \ldots, u_r = \frac{1}{\sigma_r} A v_r$ are orthonormal vectors in $\mathbb{R}^m$ (which can be completed to an orthonormal basis of $\mathbb{R}^m$ if necessary) and they form the columns of U.

How to get the SVD?

The singular values are determined by the matrix A:

$$A = UDV^t \Rightarrow A^tA = VD^tU^tUDV^t = VD^tDV^t$$

but U and V are not (almost determined in most cases). How do we compute the SVD?

(1) Diagonalize the symmetric matrix $S = A^t \cdot A$

(2) If $\lambda_1 \geq \cdots \geq \lambda_r$ are the non-zero eigenvalues of $S \Rightarrow$ the **singular values** are $\sigma_1 = \sqrt{\lambda_1}, \ldots \sigma_r = \sqrt{\lambda_r}$ (fact: A^tA always has non-negative eigenvalues).

(3) The columns of V are an orthonormal basis $v_1, \ldots, v_n$ of eigenvectors of S.

(4) $u_1 = \frac{1}{\sigma_1}Av_1, \ldots, u_r = \frac{1}{\sigma_r}Av_r$ are orthonormal vectors in $\mathbb{R}^m$ (which can be completed to an orthonormal basis of $\mathbb{R}^m$ if necessary) and they form the columns of U.

How to get the SVD?

The singular values are determined by the matrix A:

$$A = UDV^t \Rightarrow A^tA = VD^tU^tUDV^t = VD^tDV^t$$

but U and V are not (almost determined in most cases). How do we compute the SVD?

(1) Diagonalize the symmetric matrix $S = A^t \cdot A$

(2) If $\lambda_1 \geq \cdots \geq \lambda_r$ are the non-zero eigenvalues of $S \Rightarrow$ the **singular values** are $\sigma_1 = \sqrt{\lambda_1}, \ldots \sigma_r = \sqrt{\lambda_r}$ (fact: A^tA always has non-negative eigenvalues).

(3) The columns of V are an orthonormal basis $v_1, \ldots, v_n$ of eigenvectors of S.

(4) $u_1 = \frac{1}{\sigma_1}Av_1, \ldots, u_r = \frac{1}{\sigma_r}Av_r$ are orthonormal vectors in $\mathbb{R}^m$ (which can be completed to an orthonormal basis of $\mathbb{R}^m$ if necessary) and they form the columns of U.

The fundamental theorem of linear algebra

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear map and let A be its standard matrix. Then $\mathbb{R}^n = \text{Nuc}(A) \oplus \text{Im}(A^t)$ ($\text{Im}(A^t)$=row space of A), $\mathbb{R}^m = \text{Im}(A) \oplus \text{Nuc}(A^t)$, these decompositions give orthogonal complements and there exist b.o.n.'s $v_1, \dots, v_n$ (of $\mathbb{R}^n$) and $u_1, \dots, u_m$ (of $\mathbb{R}^m$) such that

1. $\text{Im}(A) = [u_1, \dots, u_r]$
2. $\text{Nuc}(A) = [v_{r+1}, \dots, v_n]$
3. $\text{Im}(A^t) = [v_1, \dots, v_r]$
4. $\text{Nuc}(A^t) = [u_{r+1}, \dots, u_m]$

Moreover, the restriction of the map f to the row space $\text{Im}(A^t)$ and onto $\text{Im}(A)$ in the bases $v_1, \dots, v_r$, $u_1, \dots, u_r$ (left and right, respectively) is the diagonal matrix of singular vaules,

$$D = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}.$$

The fundamental theorem of linear algebra

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear map and let A be its standard matrix. Then $\mathbb{R}^n = \text{Nuc}(A) \oplus \text{Im}(A^t)$ ($\text{Im}(A^t)$=row space of A), $\mathbb{R}^m = \text{Im}(A) \oplus \text{Nuc}(A^t)$, these decompositions give orthogonal complements and there exist b.o.n.'s $v_1, \dots, v_n$ (of $\mathbb{R}^n$) and $u_1, \dots, u_m$ (of $\mathbb{R}^m$) such that

1. $\text{Im}(A) = [u_1, \dots, u_r]$
2. $\text{Nuc}(A) = [v_{r+1}, \dots, v_n]$
3. $\text{Im}(A^t) = [v_1, \dots, v_r]$
4. $\text{Nuc}(A^t) = [u_{r+1}, \dots, u_m]$

Moreover, the restriction of the map f to the row space $\text{Im}(A^t)$ and onto $\text{Im}(A)$ in the bases $v_1, \dots, v_r$, $u_1, \dots, u_r$ (left and right, respectively) is the diagonal matrix of singular vaules,

$$D = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}.$$

The fundamental theorem of linear algebra

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear map and let A be its standard matrix. Then $\mathbb{R}^n = \text{Nuc}(A) \oplus \text{Im}(A^t)$ ($\text{Im}(A^t)$=row space of A), $\mathbb{R}^m = \text{Im}(A) \oplus \text{Nuc}(A^t)$, these decompositions give orthogonal complements and there exist b.o.n.'s $v_1, \ldots, v_n$ (of $\mathbb{R}^n$) and $u_1, \ldots, u_m$ (of $\mathbb{R}^m$) such that

1. $\text{Im}(A) = [u_1, \ldots, u_r]$
2. $\text{Nuc}(A) = [v_{r+1}, \ldots, v_n]$
3. $\text{Im}(A^t) = [v_1, \ldots, v_r]$
4. $\text{Nuc}(A^t) = [u_{r+1}, \ldots, u_m]$

Moreover, the restriction of the map f to the row space $\text{Im}(A^t)$ and onto $\text{Im}(A)$ in the bases $v_1, \ldots, v_r$, $u_1, \ldots, u_r$ (left and right, respectively) is the diagonal matrix of singular vaules,

$$D = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}.$$

The fundamental theorem of linear algebra

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear map and let A be its standard matrix. Then $\mathbb{R}^n = \mathrm{Nuc}(A) \oplus \mathrm{Im}(A^t)$ ($\mathrm{Im}(A^t)$=row space of A), $\mathbb{R}^m = \mathrm{Im}(A) \oplus \mathrm{Nuc}(A^t)$, these decompositions give orthogonal complements and there exist b.o.n.'s $v_1, \ldots, v_n$ (of $\mathbb{R}^n$) and $u_1, \ldots, u_m$ (of $\mathbb{R}^m$) such that

1. $\mathrm{Im}(A) = [u_1, \ldots, u_r]$
2. $\mathrm{Nuc}(A) = [v_{r+1}, \ldots, v_n]$
3. $\mathrm{Im}(A^t) = [v_1, \ldots, v_r]$
4. $\mathrm{Nuc}(A^t) = [u_{r+1}, \ldots, u_m]$

Moreover, the restriction of the map f to the row space $\mathrm{Im}(A^t)$ and onto $\mathrm{Im}(A)$ in the bases $v_1, \ldots, v_r$, $u_1, \ldots, u_r$ (left and right, respectively) is the diagonal matrix of singular vaules,

$$D = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}.$$

The fundamental theorem of linear algebra

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear map and let A be its standard matrix. Then $\mathbb{R}^n = \text{Nuc}(A) \oplus \text{Im}(A^t)$ ($\text{Im}(A^t)$=row space of A), $\mathbb{R}^m = \text{Im}(A) \oplus \text{Nuc}(A^t)$, these decompositions give orthogonal complements and there exist b.o.n.'s $v_1, \dots, v_n$ (of $\mathbb{R}^n$) and $u_1, \dots, u_m$ (of $\mathbb{R}^m$) such that

1. $\text{Im}(A) = [u_1, \dots, u_r]$
2. $\text{Nuc}(A) = [v_{r+1}, \dots, v_n]$
3. $\text{Im}(A^t) = [v_1, \dots, v_r]$
4. $\text{Nuc}(A^t) = [u_{r+1}, \dots, u_m]$

Moreover, the restriction of the map f to the row space $\text{Im}(A^t)$ and onto $\text{Im}(A)$ in the bases $v_1, \dots, v_r$, $u_1, \dots, u_r$ (left and right, respectively) is the diagonal matrix of singular vaules,

$$D = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix}.$$

2-norm of a matrix

To "measure" a linear map we measure how big the image of the unit sphere is under this map:

Definition

The **2-norm** of an $m \times n$ matrix A is

$$\|A\|_2 = \max_{\|x\|=1} \|Ax\|.$$

- This is a **matrix norm**: $\|A\|_2 \geq 0$, $\|A\|_2 = 0 \Leftrightarrow A = 0$, $\|cA\|_2 = |c|\|A\|_2$, $\|A+B\|_2 \leq \|A\|_2 + \|B\|_2$
- $\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$
- $\|Av\| \leq \|A\|_2 \|v\| \ \forall v$.
- $\|AX\|_2 = \|A\|_2$ if X is an orthogonal matrix.
- $\|YA\|_2 = \|A\|_2$ if Y is an orthogonal matrix.
- $\|AB\|_2 \leq \|A\|_2 \|B\|_2$.

2-norm of a matrix

To "measure" a linear map we measure how big the image of the unit sphere is under this map:

Definition

The **2-norm** of an $m \times n$ matrix A is

$$\|A\|_2 = \max_{\|x\|=1} \|Ax\|.$$

- This is a **matrix norm**: $\|A\|_2 \geq 0$, $\|A\|_2 = 0 \Leftrightarrow A = 0$, $\|cA\|_2 = |c|\|A\|_2$, $\|A+B\|_2 \leq \|A\|_2 + \|B\|_2$
- $\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$
- $\|Av\| \leq \|A\|_2 \|v\| \ \forall v$.
- $\|AX\|_2 = \|A\|_2$ if X is an orthogonal matrix.
- $\|YA\|_2 = \|A\|_2$ if Y is an orthogonal matrix.
- $\|AB\|_2 \leq \|A\|_2 \|B\|_2$.

2-norm of a matrix

To "measure" a linear map we measure how big the image of the unit sphere is under this map:

Definition

The **2-norm** of an $m \times n$ matrix A is

$$\|A\|_2 = \max_{\|x\|=1} \|Ax\|.$$

- This is a **matrix norm**: $\|A\|_2 \geq 0$, $\|A\|_2 = 0 \Leftrightarrow A = 0$, $\|cA\|_2 = |c|\|A\|_2$, $\|A + B\|_2 \leq \|A\|_2 + \|B\|_2$
- $\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$
- $\|Av\| \leq \|A\|_2\|v\| \ \forall v$.
- $\|AX\|_2 = \|A\|_2$ if X is an orthogonal matrix.
- $\|YA\|_2 = \|A\|_2$ if Y is an orthogonal matrix.
- $\|AB\|_2 \leq \|A\|_2\|B\|_2$.

2-norm of a matrix

To "measure" a linear map we measure how big the image of the unit sphere is under this map:

Definition

The **2-norm** of an $m \times n$ matrix A is

$$\|A\|_2 = \max_{\|x\|=1} \|Ax\|.$$

- This is a **matrix norm**: $\|A\|_2 \geq 0$, $\|A\|_2 = 0 \Leftrightarrow A = 0$, $\|cA\|_2 = |c|\|A\|_2$, $\|A + B\|_2 \leq \|A\|_2 + \|B\|_2$
- $\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$
- $\|Av\| \leq \|A\|_2 \|v\| \; \forall v$.
- $\|AX\|_2 = \|A\|_2$ if X is an orthogonal matrix.
- $\|YA\|_2 = \|A\|_2$ if Y is an orthogonal matrix.
- $\|AB\|_2 \leq \|A\|_2 \|B\|_2$.

2-norm of a matrix

To "measure" a linear map we measure how big the image of the unit sphere is under this map:

Definition

The **2-norm** of an $m \times n$ matrix A is

$$\|A\|_2 = \max_{\|x\|=1} \|Ax\|.$$

- This is a **matrix norm**: $\|A\|_2 \geq 0$, $\|A\|_2 = 0 \Leftrightarrow A = 0$, $\|cA\|_2 = |c|\|A\|_2$, $\|A + B\|_2 \leq \|A\|_2 + \|B\|_2$
- $\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$
- $\|Av\| \leq \|A\|_2\|v\| \ \forall v$.
- $\|AX\|_2 = \|A\|_2$ if X is an orthogonal matrix.
- $\|YA\|_2 = \|A\|_2$ if Y is an orthogonal matrix.
- $\|AB\|_2 \leq \|A\|_2\|B\|_2$.

2-norm of a matrix

To "measure" a linear map we measure how big the image of the unit sphere is under this map:

Definition

The **2-norm** of an $m \times n$ matrix A is

$$\|A\|_2 = \max_{\|x\|=1} \|Ax\|.$$

- This is a **matrix norm**: $\|A\|_2 \geq 0$, $\|A\|_2 = 0 \Leftrightarrow A = 0$, $\|cA\|_2 = |c|\|A\|_2$, $\|A+B\|_2 \leq \|A\|_2 + \|B\|_2$
- $\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$
- $\|Av\| \leq \|A\|_2 \|v\| \; \forall v$.
- $\|AX\|_2 = \|A\|_2$ if X is an orthogonal matrix.
- $\|YA\|_2 = \|A\|_2$ if Y is an orthogonal matrix.
- $\|AB\|_2 \leq \|A\|_2 \|B\|_2$.

2-norm of a matrix

To "measure" a linear map we measure how big the image of the unit sphere is under this map:

Definition

The **2-norm** of an $m \times n$ matrix A is

$$\|A\|_2 = \max_{\|x\|=1} \|Ax\|.$$

- This is a **matrix norm**: $\|A\|_2 \geq 0$, $\|A\|_2 = 0 \Leftrightarrow A = 0$, $\|cA\|_2 = |c|\|A\|_2$, $\|A+B\|_2 \leq \|A\|_2 + \|B\|_2$
- $\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$
- $\|Av\| \leq \|A\|_2 \|v\| \ \forall v$.
- $\|AX\|_2 = \|A\|_2$ if X is an orthogonal matrix.
- $\|YA\|_2 = \|A\|_2$ if Y is an orthogonal matrix.
- $\|AB\|_2 \leq \|A\|_2 \|B\|_2$.

Geometric consequence of the *SVD*:

Proposition

- $\|A\|_2 = \sigma_1$
- *The maximum is attained at* $\pm v_1$*:* $\max_{||x||=1} \|Ax\| = \|Av_1\|$.
- $\min_{||x||=1} \|Ax\| = \begin{cases} \sigma_n & \textit{if A has rank n, and is attained at } \pm v_n \\ 0 & \textit{if A has rank } < n \end{cases}$
- *If A is invertible,* $\|A^{-1}\|_2 = \frac{1}{\sigma_r}$.

Geometric consequence of the *SVD*:

Proposition

- $\|A\|_2 = \sigma_1$
- *The maximum is attained at* $\pm v_1$*:* $\max_{||x||=1} \|Ax\| = \|Av_1\|$.
- $\min_{||x||=1} \|Ax\| = \begin{cases} \sigma_n & \textit{if A has rank n, and is attained at } \pm v_n \\ 0 & \textit{if A has rank } < n \end{cases}$
- *If A is invertible,* $\|A^{-1}\|_2 = \frac{1}{\sigma_r}$.

Geometric consequence of the *SVD*:

Proposition

- $\|A\|_2 = \sigma_1$
- *The maximum is attained at* $\pm v_1$*:* $\max_{||x||=1} \|Ax\| = \|Av_1\|$.
- $\min_{||x||=1} \|Ax\| = \begin{cases} \sigma_n & \textit{if A has rank n, and is attained at } \pm v_n \\ 0 & \textit{if A has rank } < n \end{cases}$
- *If A is invertible,* $\|A^{-1}\|_2 = \frac{1}{\sigma_r}$.

Geometric consequence of the *SVD*:

Proposition

- $\|A\|_2 = \sigma_1$
- *The maximum is attained at* $\pm v_1$*:* $\max_{||x||=1} \|Ax\| = \|Av_1\|$.
- $\min_{||x||=1} \|Ax\| = \begin{cases} \sigma_n & \text{if } A \text{ has rank } n, \text{ and is attained at } \pm v_n \\ 0 & \text{if } A \text{ has rank } < n \end{cases}$
- *If A is invertible,* $\|A^{-1}\|_2 = \frac{1}{\sigma_r}$.

Geometric consequence of the *SVD*:

Proposition

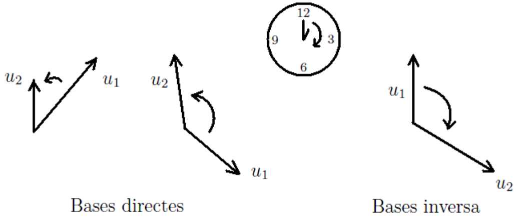
- $\|A\|_2 = \sigma_1$
- *The maximum is attained at* $\pm v_1$*:* $\max_{||x||=1} \|Ax\| = \|Av_1\|$.
- $\min_{||x||=1} \|Ax\| = \begin{cases} \sigma_n & \textit{if A has rank n, and is attained at } \pm v_n \\ 0 & \textit{if A has rank } < n \end{cases}$
- *If A is invertible,* $\|A^{-1}\|_2 = \frac{1}{\sigma_r}$.

Outline

Orientation of $\mathbb{R}^2$

A basis u_1, u_2 of $\mathbb{R}^2$ has

- direct/positive orientation if the shortest rotation from u_1 to u_2 is counter-clockwise.
- inverse/negative if the shortest rotation from u_1 to u_2 is clockwise.



Orientations

In $\mathbb{R}^n$ we say that the standard basis has direct/positive orientation. For the other bases:

Definition

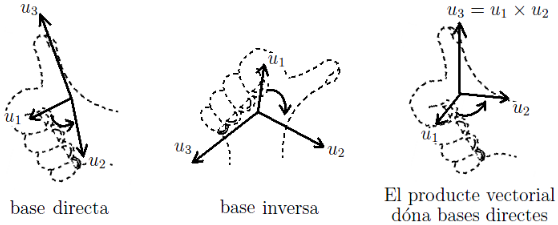
A basis $u_1, \ldots, u_n$ of $\mathbb{R}^n$ has direct/positive orientation, if

$$\det(u_1, u_2, \ldots, u_n) > 0$$

(computed in standard coordinates); otherwise, the basis is said to have inverse/negative orientation.

Geometric intuition in $\mathbb{R}^3$

In $\mathbb{R}^3$, to see if a basis u_1, u_2, u_3 has direct orientation we use the *right-hand rule*: put your thumb pointing to u_3 and if the sense of closing your hand is the same as the shortest from u_1 and u_2, then it has direct orientation.



- If $u, v \in \mathbb{R}^3$ are l.i. $\Rightarrow u, v, u \times v$ is a direct basis,

$$\det(u, v, u \times v) > 0.$$

Isometries

Definition

An endomorphism $f \in End(E)$ is an isometry if it preserves the scalar product,

$$\langle f(u), f(v)\rangle = \langle u, v\rangle \quad \forall u, v.$$

Ex: If A is an orthogonal matrix, then $x \overset{f}{\mapsto} Ax$ is an isometry.

Proposition

If $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a linear map, the following are equivalent

- *f is an isometry*
- *f maps the standard basis to a b.o.n*
- *$M_e(f)$ (or in any b.o.n) is an orthogonal matrix*

Isometries

Definition

An endomorphism $f \in End(E)$ is an isometry if it preserves the scalar product,

$$\langle f(u), f(v)\rangle = \langle u, v\rangle \quad \forall u, v.$$

Ex: If A is an orthogonal matrix, then $x \overset{f}{\mapsto} Ax$ is an isometry.

Proposition

If $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a linear map, the following are equivalent

- *f is an isometry*
- *f maps the standard basis to a b.o.n*
- *$M_e(f)$ (or in any b.o.n) is an orthogonal matrix*

Direct/inverse isometries

Properties: if f is an isometry, then

- $\|f(u)\| = \|u\|$, for all $u \in E$.
- $d(f(x), f(y)) = d(x, y)$ for all x, y.
- angle between $f(u)$ and $f(v)$ = angle between u and v

Remark: if f is an isometry of $\mathbb{R}^n \Rightarrow \det(f) = \pm 1$ and if λ is an eigenvalue of f, then $|\lambda| = 1$.

- If $\det f = +1$ we say that it is a direct isometry (preserves orientation).
- If $\det f = +1$ we say that it is an inverse isometry (changes orientation).

Examples of isometries in $\mathbb{R}^2$

The following maps $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ are isometries:

- $f=$ reflection/symmetry with along a line l passing through the origin, $l = [v]$. Then

$$f(x) = 2\frac{< v, x >}{< v, v >}v - x, \qquad M_e(f) = \frac{2}{< v, v >}vv^t - Id,$$

and taking u in $[v]$ this can be written as:

$$M_e(f) = Id - \frac{2}{< u, u >}u \cdot u^t.$$

- $f=$**rotation** counterclockwise of angle α with respect to the origin; then

$$M_e(f) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

f is a direct isometry.

Examples of isometries in $\mathbb{R}^2$

The following maps $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ are isometries:

- $f=$ reflection/symmetry with along a line l passing through the origin, $l = [v]$.Then

$$f(x) = 2\frac{< v, x >}{< v, v >}v - x, \qquad M_e(f) = \frac{2}{< v, v >}vv^t - Id,$$

  and taking u in $[v]$ this can be written as:

$$M_e(f) = Id - \frac{2}{< u, u >}u \cdot u^t.$$

- $f=$**rotation** counterclockwise of angle α with respect to the origin; then

$$M_e(f) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

  f is a direct isometry.

Classification of isometries in $\mathbb{R}^2$

Theorem

If f is an isometry of $\mathbb{R}^2$, then either

- $\det f = 1$ *and f is a counterclockwise rotation of angle α with respect to $(0,0)$ and in any direct b.o.n $\mathbf{u}$,*

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

or

- $\det f = -1$ *and f is a reflection/symmetry along a line $[v] \ni (0,0)$; if $u \in [v]^\perp \Rightarrow$*

$$M_e(f) = Id - \frac{2}{< u, u >} u \cdot u^t \qquad M_{v,u}(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Classification of isometries in $\mathbb{R}^2$

Theorem

If f is an isometry of $\mathbb{R}^2$, then either

- $\det f = 1$ *and f is a counterclockwise rotation of angle α with respect to $(0,0)$ and in any direct b.o.n* $\mathbf{u}$,

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

or

- $\det f = -1$ *and f is a reflection/symmetry along a line* $[v] \ni (0,0)$*; if* $u \in [v]^{\perp} \Rightarrow$

$$M_e(f) = Id - \frac{2}{< u, u >} u \cdot u^t \qquad M_{v,u}(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Classification of isometries in $\mathbb{R}^2$

Theorem

If f is an isometry of $\mathbb{R}^2$, then either

- $\det f = 1$ *and f is a counterclockwise rotation of angle α with respect to $(0,0)$ and in any direct b.o.n* $\mathbf{u}$,

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

or

- $\det f = -1$ *and f is a reflection/symmetry along a line* $[v] \ni (0,0)$; *if* $u \in [v]^\perp \Rightarrow$

$$M_e(f) = Id - \frac{2}{< u, u >} u \cdot u^t \qquad M_{v,u}(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Example of isometry classification

Tell if the map $f(x,y) = (\frac{x-\sqrt{3}y}{2}, \frac{\sqrt{3}x+y}{2})$ is an isomtry and describe it.

- The standard matrix of f is $M = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$.
- M is orthogonal $\Rightarrow$ f is an isometry
- $det(M) = 1 \Rightarrow f$ is a rotation (by Theorem of Classification).
- To find angle α: according to the Theorem M must be of the form

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \Rightarrow \cos\alpha = 1/2, \sin\alpha = \sqrt{3}/2 \Rightarrow \alpha = \pi/3.$$

- Or also: $\alpha = \widehat{v, f(v)}\ \forall v \Rightarrow$ take any $v \in \mathbb{R}^2$, compute $\cos\alpha = \frac{<v,f(v)>}{\|v\|\|f(v)\|}$ and if $det(v, f(v)) > 0$ (resp. $det(v, f(v)) < 0$) take $\alpha \in [0, \pi]$ (resp. $\alpha \in [\pi, 2\pi]$).

Example of isometry classification

Tell if the map $f(x,y) = (\frac{x-\sqrt{3}y}{2}, \frac{\sqrt{3}x+y}{2})$ is an isomtry and describe it.

- The standard matrix of f is $M = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$.
- M is orthogonal $\Rightarrow$ f is an isometry
- $det(M) = 1 \Rightarrow f$ is a rotation (by Theorem of Classification).
- To find angle α: according to the Theorem M must be of the form

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \Rightarrow \cos\alpha = 1/2, \sin\alpha = \sqrt{3}/2 \Rightarrow \alpha = \pi/3.$$

- Or also: $\alpha = \widehat{v, f(v)}\ \forall v \Rightarrow$ take any $v \in \mathbb{R}^2$, compute $\cos\alpha = \frac{<v,f(v)>}{\|v\|\|f(v)\|}$ and if $det(v, f(v)) > 0$ (resp. $det(v, f(v)) < 0$) take $\alpha \in [0, \pi]$ (resp. $\alpha \in [\pi, 2\pi]$).

Example of isometry classification

Tell if the map $f(x,y) = (\frac{x-\sqrt{3}y}{2}, \frac{\sqrt{3}x+y}{2})$ is an isomtry and describe it.

- The standard matrix of f is $M = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$.
- M is orthogonal $\Rightarrow$ f is an isometry
- $det(M) = 1 \Rightarrow f$ is a rotation (by Theorem of Classification).
- To find angle α: according to the Theorem M must be of the form

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \Rightarrow \cos\alpha = 1/2, \sin\alpha = \sqrt{3}/2 \Rightarrow \alpha = \pi/3.$$

- Or also: $\alpha = \widehat{v, f(v)}$ $\forall v \Rightarrow$ take any $v \in \mathbb{R}^2$, compute $\cos\alpha = \frac{<v,f(v)>}{|||v|||||f(v)|||}$ and if $det(v, f(v)) > 0$ (resp. $det(v, f(v)) < 0$) take $\alpha \in [0, \pi]$ (resp. $\alpha \in [\pi, 2\pi]$).

Example of isometry classification

Tell if the map $f(x,y) = (\frac{x-\sqrt{3}y}{2}, \frac{\sqrt{3}x+y}{2})$ is an isomtry and describe it.

- The standard matrix of f is $M = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$.
- M is orthogonal $\Rightarrow$ f is an isometry
- $det(M) = 1 \Rightarrow f$ is a rotation (by Theorem of Classification).
- To find angle α: according to the Theorem M must be of the form

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \Rightarrow \cos\alpha = 1/2, \sin\alpha = \sqrt{3}/2 \Rightarrow \alpha = \pi/3.$$

- Or also: $\alpha = \widehat{v, f(v)}\ \forall v \Rightarrow$ take any $v \in \mathbb{R}^2$, compute $\cos\alpha = \frac{<v,f(v)>}{|||v|||||f(v)|||}$ and if $det(v, f(v)) > 0$ (resp. $det(v, f(v)) < 0$) take $\alpha \in [0, \pi]$ (resp. $\alpha \in [\pi, 2\pi]$).

Example of isometry classification

Tell if the map $f(x,y) = (\frac{x-\sqrt{3}y}{2}, \frac{\sqrt{3}x+y}{2})$ is an isomtry and describe it.

- The standard matrix of f is $M = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$.
- M is orthogonal $\Rightarrow$ f is an isometry
- $det(M) = 1 \Rightarrow f$ is a rotation (by Theorem of Classification).
- To find angle α: according to the Theorem M must be of the form

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \Rightarrow \cos\alpha = 1/2, \sin\alpha = \sqrt{3}/2 \Rightarrow \alpha = \pi/3.$$

- Or also: $\alpha = \widehat{v, f(v)}\ \forall v \Rightarrow$ take any $v \in \mathbb{R}^2$, compute $\cos\alpha = \frac{<v,f(v)>}{|||v|||||f(v)|||}$ and if $det(v, f(v)) > 0$ (resp. $det(v, f(v)) < 0$) take $\alpha \in [0, \pi]$ (resp. $\alpha \in [\pi, 2\pi]$).

Examples of isometries in $\mathbb{R}^3$

$f =$ *rotation* of a certain angle with respect to a line $r \ni O = (0,0,0)$ (r is called *rotation axis*).

- To distinguish between angle θ and $-\theta$ ($= 2\pi - \theta$) need to orient r.
- Choose a vector u of r, orient $r = [u]$ by u using the right hand rule: pointing your thumb in the sense of u, positive rotation is in the sense of cloing hand (looking from the arrow of u, counterclockwise rotation on the plane $r^\perp$). Then θ is in $[0,\pi]$ or $[\pi, 2\pi]$ according to: for any $v \notin r$

$$\begin{aligned} \det(v, f(v), u) \geq 0 \quad &\Leftrightarrow \quad \theta \in [0,\pi] \\ \det(v, f(v), u) \leq 0 \quad &\Leftrightarrow \quad \theta \in [\pi, 2\pi] \end{aligned} \qquad (1)$$

- Preserves orientation of bases, so it's a direct isometry ($\det(f) = 1$).

Examples of isometries in $\mathbb{R}^3$

$f =$ *rotation* of a certain angle with respect to a line $r \ni O = (0,0,0)$ (r is called *rotation axis*).

- To distinguish between angle θ and $-\theta$ $(= 2\pi - \theta)$ need to orient r.
- Choose a vector u of r, orient $r = [u]$ by u using the right hand rule: pointing your thumb in the sense of u, positive rotation is in the sense of cloing hand (looking from the arrow of u, counterclockwise rotation on the plane $r^{\perp}$). Then θ is in $[0,\pi]$ or $[\pi, 2\pi]$ according to: for any $v \notin r$

$$\begin{array}{lcl} \det(v, f(v), u) \geq 0 & \Leftrightarrow & \theta \in [0,\pi] \\ \det(v, f(v), u) \leq 0 & \Leftrightarrow & \theta \in [\pi, 2\pi] \end{array} \quad (1)$$

- Preserves orientation of bases, so it's a direct isometry ($\det(f) = 1$).

Examples of isometries in $\mathbb{R}^3$

f = *rotation* of a certain angle with respect to a line $r \ni O = (0,0,0)$ (r is called *rotation axis*).

- To distinguish between angle θ and $-\theta$ $(= 2\pi - \theta)$ need to orient r.
- Choose a vector u of r, orient $r = [u]$ by u using the right hand rule: pointing your thumb in the sense of u, positive rotation is in the sense of cloing hand (looking from the arrow of u, counterclockwise rotation on the plane $r^{\perp}$). Then θ is in $[0, \pi]$ or $[\pi, 2\pi]$ according to: for any $v \notin r$

$$\begin{aligned} \det(v, f(v), u) \geq 0 \quad &\Leftrightarrow \quad \theta \in [0, \pi] \\ \det(v, f(v), u) \leq 0 \quad &\Leftrightarrow \quad \theta \in [\pi, 2\pi] \end{aligned} \tag{1}$$

- Preserves orientation of bases, so it's a direct isometry ($\det(f) = 1$).

Example of rotation:

f = rotation with $r = [e_3]$, oriented by e_3, and angle $\pi/3$. Then

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Matrix of a rotation

f = rotation with respect to $r = [u]$ (oriented by u) and angle θ. Take a positive b.o.n. $\mathbf{u} = u_1, u_2, u_3$ with $u_3 = \frac{u}{\|u\|}$ ("adapted b.o.n."), then

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Example: Axial symmetry

f = axial symmetry with respect to a line $r \ni O$.

- f = rotation of angle π with axis r (so $\det(f) = 1$).
- As $\pi = -\pi$, orientation not needed.
- Ex: find $M_e(f)$ for f = axial symmetry respect to z axis.

Example: Axial symmetry

f = axial symmetry with respect to a line $r \ni O$.

- f = rotation of angle π with axis r (so $\det(f) = 1$).
- As $\pi = -\pi$, orientation not needed.
- Ex: find $M_e(f)$ for f = axial symmetry respect to z axis.

Example: Axial symmetry

f = axial symmetry with respect to a line $r \ni O$.

- f = rotation of angle π with axis r (so $\det(f) = 1$).
- As $\pi = -\pi$, orientation not needed.
- Ex: find $M_e(f)$ for f = axial symmetry respect to z axis.

Example: Specular reflection

f = *specular reflexion/symmetry* along a plane $H \ni O$.

- It changes basis orientation $\Rightarrow \det(f) = -1$.
- Example: if $H = \{z = 0\}$, then $f(x, y, z) = (x, y, -z)$
- If $u \in H^\perp$, then

$$M_e(f) = Id - \frac{2}{< u, u >} u \cdot u^t.$$

Example: Specular reflection

f = *specular reflexion/symmetry* along a plane $H \ni O$.

- It changes basis orientation $\Rightarrow \det(f) = -1$.
- Example: if $H = \{z = 0\}$, then $f(x, y, z) = (x, y, -z)$
- If $u \in H^\perp$, then

$$M_e(f) = Id - \frac{2}{< u, u >} u \cdot u^t.$$

Example: Specular reflection

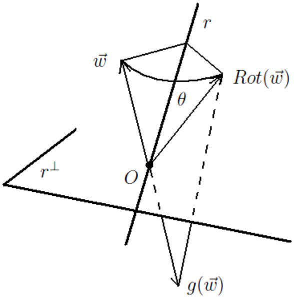
$f =$ *specular reflexion/symmetry* along a plane $H \ni O$.

- It changes basis orientation $\Rightarrow \det(f) = -1$.
- Example: if $H = \{z = 0\}$, then $f(x, y, z) = (x, y, -z)$
- If $u \in H^{\perp}$, then

$$M_e(f) = Id - \frac{2}{< u, u >} u \cdot u^t.$$

Example: Rotation followed by specular reflection

g = rotation R with axis $r = [u]$ and angle θ followed by a specular reflection along the orthogonal plane to r, $[u]^{\perp}$.



Classification of isometries in $\mathbb{R}^3$

Theorem

If $f : \mathbb{R}^3 \to \mathbb{R}^3$ is an isometry, then either:

- **Case 1.** $\boxed{\det f = +1}$ *(direct isometry): f is a rotation with axis $r = [u] =$ VEPs of VAP 1.*
- **Case 2.** $\boxed{\det f = -1}$ *(indirect isometry): $f =$ rotation R of angle θ and axis $r = [u]$ (=VEPs of VAP -1), followed by a specular reflection S along plane $[u]^\perp$, $f = S \circ R$.*

Any isometry in $\mathbb{R}^3$ can be written as one of these 2 (case 1 if $\det f = 1$, case 2 if $\det f = -1$). Important: in case 2, the plane of reflection is orthogonal to rotation axis.

Classification of isometries in $\mathbb{R}^3$

Theorem

If $f : \mathbb{R}^3 \to \mathbb{R}^3$ is an isometry, then either:

- **Case 1.** $\boxed{\det f = +1}$ *(direct isometry): f is a rotation with axis $r = [u]$ = VEPs of VAP 1.*
- **Case 2.** $\boxed{\det f = -1}$ *(indirect isometry): f = rotation R of angle θ and axis $r = [u]$ (=VEPs of VAP -1), followed by a specular reflection S along plane $[u]^\perp$, $f = S \circ R$.*

Any isometry in $\mathbb{R}^3$ can be written as one of these 2 (case 1 if $\det f = 1$, case 2 if $\det f = -1$). Important: in case 2, the plane of reflection is orthogonal to rotation axis.

Classification of isometries in $\mathbb{R}^3$

Theorem

If $f : \mathbb{R}^3 \to \mathbb{R}^3$ is an isometry, then either:

- **Case 1.** $\boxed{\det f = +1}$ *(direct isometry): f is a rotation with axis $r = [u]$ = VEPs of VAP 1.*
- **Case 2.** $\boxed{\det f = -1}$ *(indirect isometry): f = rotation R of angle θ and axis $r = [u]$ (=VEPs of VAP -1), followed by a specular reflection S along plane $[u]^\perp$, $f = S \circ R$.*

Any isometry in $\mathbb{R}^3$ can be written as one of these 2 (case 1 if $\det f = 1$, case 2 if $\det f = -1$). Important: in case 2, the plane of reflection is orthogonal to rotation axis.

Case 1: $\det f = +1$, f = rotation of axis $r = [u]$.

Orient the axis by u and let θ be the angle of rotation. Take u of norm 1. Then if $\mathbf{u} = \{u_1, u_2, u_3 = u\}$ is a direct b.o.n (called basis adapted to f),

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The matrix of f in any other basis can be obtained by change of basis.

- $[u]$ = VEPs of VAP 1 = vectors fixed by f.
- $\theta = \widehat{v, f(v)}$ if v is orthogonal to axis $[u]$.
- $\cos(\theta) = \frac{tr(f)-1}{2}$ and θ is in $[0, \pi]$ or $[\pi, 2\pi]$ according to (for any $v \notin [u]$):

$$\begin{aligned} \det(v, f(v), u) \geq 0 \quad &\Leftrightarrow \quad \theta \in [0, \pi] \\ \det(v, f(v), u) \leq 0 \quad &\Leftrightarrow \quad \theta \in [\pi, 2\pi] \end{aligned} \qquad (2)$$

Case 1: $\det f = +1$, f = rotation of axis $r = [u]$.

Orient the axis by u and let θ be the angle of rotation. Take u of norm 1. Then if $\mathbf{u} = \{u_1, u_2, u_3 = u\}$ is a direct b.o.n (called basis adapted to f),

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The matrix of f in any other basis can be obtained by change of basis.

- $[u]$ = VEPs of VAP 1 = vectors fixed by f.
- $\theta = \widehat{v, f(v)}$ if v is orthogonal to axis $[u]$.
- $\cos(\theta) = \frac{tr(f)-1}{2}$ and θ is in $[0,\pi]$ or $[\pi, 2\pi]$ according to (for any $v \notin [u]$):

$$\begin{aligned} \det(v, f(v), u) \geq 0 \quad &\Leftrightarrow \quad \theta \in [0,\pi] \\ \det(v, f(v), u) \leq 0 \quad &\Leftrightarrow \quad \theta \in [\pi, 2\pi] \end{aligned} \tag{2}$$

Case 1: $\det f = +1$, $f =$ rotation of axis $r = [u]$.

Orient the axis by u and let θ be the angle of rotation. Take u of norm 1. Then if $\mathbf{u} = \{u_1, u_2, u_3 = u\}$ is a direct b.o.n (called basis adapted to f),

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The matrix of f in any other basis can be obtained by change of basis.

- $[u] =$ VEPs of VAP 1 $=$ vectors fixed by f.
- $\theta = \widehat{v, f(v)}$ if v is orthogonal to axis $[u]$.
- $\cos(\theta) = \frac{tr(f)-1}{2}$ and θ is in $[0, \pi]$ or $[\pi, 2\pi]$ according to (for any $v \notin [u]$):

$$\begin{array}{lcl} \det(v, f(v), u) \geq 0 & \Leftrightarrow & \theta \in [0, \pi] \\ \det(v, f(v), u) \leq 0 & \Leftrightarrow & \theta \in [\pi, 2\pi] \end{array} \qquad (2)$$

Case 2: $\det f = -1$, $f = S \circ R$

$f = S \circ R$, R= rotation angle θ of axis $r = [u] \ni O$ (oriented by u), S specular reflection with respect to plane $H = [u]^{\perp}$.

- Case 2.a $\theta = 0$, $f = S =$ specular reflection along a plane $H = [u]^{\perp}$. We have:
  - $[u]$: VEPs of VAP -1
  - Image of a vector v:

    $$\boxed{S(v) = v - 2\frac{\langle u,v\rangle}{\langle u,u\rangle}u}$$

  - Matrix of f:

    $$M(S) = \mathrm{Id} - \frac{2}{u^t u} uu^t$$

  - H: vectors fixed by f, plane of VEPs of VAP 1.
  - $H = [v - f(v)]^{\perp}$ for any $v \notin H$.

Case 2: $\det f = -1$, $f = S \circ R$

$f = S \circ R$, $R=$ rotation angle θ of axis $r = [u] \ni O$ (oriented by u), S specular reflection with respect to plane $H = [u]^\perp$.

- Case 2.a $\theta = 0$, $f = S =$ specular reflection along a plane $H = [u]^\perp$. We have:
  - $[u]$: VEPs of VAP -1
  - Image of a vector v:

$$\boxed{S(v) = v - 2\frac{\langle u,v\rangle}{\langle u,u\rangle}u}$$

  - Matrix of f:

$$M(S) = \mathrm{Id} - \frac{2}{u^t u} uu^t$$

  - H: vectors fixed by f, plane of VEPs of VAP 1.
  - $H = [v - f(v)]^\perp$ for any $v \notin H$.

Case 2: $\det f = -1$, $f = S \circ R$

$f = S \circ R$, R= rotation angle θ of axis $r = [u] \ni O$ (oriented by u), S specular reflection with respect to plane $H = [u]^\perp$.

- Case 2.a $\theta = 0$, $f = S =$ specular reflection along a plane $H = [u]^\perp$. We have:
  - $[u]$: VEPs of VAP -1
  - Image of a vector v:

$$\boxed{S(v) = v - 2\frac{\langle u,v\rangle}{\langle u,u\rangle}u}$$

  - Matrix of f:

$$M(S) = \mathrm{Id} - \frac{2}{u^t u} uu^t$$

  - H: vectors fixed by f, plane of VEPs of VAP 1.
  - $H = [v - f(v)]^\perp$ for any $v \notin H$.

Case 2: $\det f = -1$, $f = S \circ R$

$f = S \circ R$, R= rotation angle θ of axis $r = [u] \ni O$ (oriented by u), S specular reflection with respect to plane $H = [u]^\perp$.

- Case 2.a $\theta = 0$, $f = S =$ specular reflection along a plane $H = [u]^\perp$. We have:
  - $[u]$: VEPs of VAP -1
  - Image of a vector v:

$$\boxed{S(v) = v - 2\frac{\langle u,v\rangle}{\langle u,u\rangle}u}$$

  - Matrix of f:

$$M(S) = \mathrm{Id} - \frac{2}{u^t u} uu^t$$

  - H: vectors fixed by f, plane of VEPs of VAP 1.
  - $H = [v - f(v)]^\perp$ for any $v \notin H$.

Case 2: $\det f = -1$, $f = S \circ R$

$f = S \circ R$, R= rotation angle θ of axis $r = [u] \ni O$ (oriented by u), S specular reflection with respect to plane $H = [u]^\perp$.

- Case 2.a $\theta = 0$, $f = S =$ specular reflection along a plane $H = [u]^\perp$. We have:
  - $[u]$: VEPs of VAP -1
  - Image of a vector v:

$$\boxed{S(v) = v - 2\frac{\langle u,v\rangle}{\langle u,u\rangle}u}$$

  - Matrix of f:

$$M(S) = \mathrm{Id} - \frac{2}{u^t u}uu^t$$

  - H: vectors fixed by f, plane of VEPs of VAP 1.
  - $H = [v - f(v)]^\perp$ for any $v \notin H$.

Case 2: $\det f = -1$, $f = S \circ R$

$f = S \circ R$, R= rotation angle θ of axis $r = [u] \ni O$ (oriented by u), S specular reflection with respect to plane $H = [u]^{\perp}$.

- Case 2.a $\theta = 0$, $f = S =$ specular reflection along a plane $H = [u]^{\perp}$. We have:
  - $[u]$: VEPs of VAP -1
  - Image of a vector v:

$$\boxed{S(v) = v - 2\frac{\langle u,v\rangle}{\langle u,u\rangle}u}$$

  - Matrix of f:

$$M(S) = \mathrm{Id} - \frac{2}{u^t u}uu^t$$

  - H: vectors fixed by f, plane of VEPs of VAP 1.
  - $H = [v - f(v)]^{\perp}$ for any $v \notin H$.

- Case 2.b: $\theta \neq 0$, $f = S \circ R$. Orient rotation axis by u. Then if $\mathbf{u} = \{u_1, u_2, u_3 = \frac{u}{\|u\|}\}$ is a direct b.o.n (basis adapted to f),

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The matrix of f in any other basis can be obtained by change of basis.

- $[u]$ = VEPs of VAP -1.
- H invariant subspace of dimension 2.
- $\theta = \widehat{v, f(v)}$ if v orthogonal to axis $[u]$.
- $\cos(\theta) = \frac{tr(f)+1}{2}$
- We have θ in $[0, \pi]$ or $[\pi, 2\pi]$ according to (for any $v \notin [u]$):

$$\begin{cases} \det(v, f(v), u) \geq 0 & \Leftrightarrow \quad \theta \in [0, \pi] \\ \det(v, f(v), u) \leq 0 & \Leftrightarrow \quad \theta \in [\pi, 2\pi] \end{cases} \tag{3}$$

- Case 2.b: $\theta \neq 0$, $f = S \circ R$. Orient rotation axis by u. Then if $\mathbf{u} = \{u_1, u_2, u_3 = \frac{u}{\|u\|}\}$ is a direct b.o.n (basis adapted to f),

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The matrix of f in any other basis can be obtained by change of basis.

  - $[u] =$ VEPs of VAP -1.
  - H invariant subspace of dimension 2.
  - $\theta = \widehat{v, f(v)}$ if v orthogonal to axis $[u]$.
  - $\cos(\theta) = \frac{tr(f)+1}{2}$
  - We have θ in $[0, \pi]$ or $[\pi, 2\pi]$ according to (for any $v \notin [u]$):

$$\begin{cases} \det(v, f(v), u) \geq 0 & \Leftrightarrow \quad \theta \in [0, \pi] \\ \det(v, f(v), u) \leq 0 & \Leftrightarrow \quad \theta \in [\pi, 2\pi] \end{cases} \tag{3}$$

- Case 2.b: $\theta \neq 0$, $f = S \circ R$. Orient rotation axis by u. Then if $\mathbf{u} = \{u_1, u_2, u_3 = \frac{u}{\|u\|}\}$ is a direct b.o.n (basis adapted to f),

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The matrix of f in any other basis can be obtained by change of basis.

  - $[u]$ = VEPs of VAP -1.
  - H invariant subspace of dimension 2.
  - $\theta = \widehat{v, f(v)}$ if v orthogonal to axis $[u]$.
  - $\cos(\theta) = \frac{tr(f)+1}{2}$
  - We have θ in $[0,\pi]$ or $[\pi, 2\pi]$ according to (for any $v \notin [u]$):

$$\begin{cases} \det(v, f(v), u) \geq 0 & \Leftrightarrow \quad \theta \in [0,\pi] \\ \det(v, f(v), u) \leq 0 & \Leftrightarrow \quad \theta \in [\pi, 2\pi] \end{cases} \quad (3)$$

- Case 2.b: $\theta \neq 0$, $f = S \circ R$. Orient rotation axis by u. Then if $\mathbf{u} = \{u_1, u_2, u_3 = \frac{u}{\|u\|}\}$ is a direct b.o.n (basis adapted to f),

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The matrix of f in any other basis can be obtained by change of basis.

  - $[u]$ = VEPs of VAP -1.
  - H invariant subspace of dimension 2.
  - $\theta = \widehat{v, f(v)}$ if v orthogonal to axis $[u]$.
  - $\cos(\theta) = \frac{tr(f)+1}{2}$
  - We have θ in $[0, \pi]$ or $[\pi, 2\pi]$ according to (for any $v \notin [u]$):

$$\begin{cases} \det(v, f(v), u) \geq 0 & \Leftrightarrow \quad \theta \in [0, \pi] \\ \det(v, f(v), u) \leq 0 & \Leftrightarrow \quad \theta \in [\pi, 2\pi] \end{cases} \tag{3}$$

- Case 2.b: $\theta \neq 0$, $f = S \circ R$. Orient rotation axis by u. Then if $\mathbf{u} = \{u_1, u_2, u_3 = \frac{u}{\|u\|}\}$ is a direct b.o.n (basis adapted to f),

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The matrix of f in any other basis can be obtained by change of basis.

  - $[u]$ = VEPs of VAP -1.
  - H invariant subspace of dimension 2.
  - $\theta = \widehat{v, f(v)}$ if v orthogonal to axis $[u]$.
  - $\cos(\theta) = \frac{tr(f)+1}{2}$
  - We have θ in $[0, \pi]$ or $[\pi, 2\pi]$ according to (for any $v \notin [u]$):

$$\begin{cases} \det(v, f(v), u) \geq 0 & \Leftrightarrow \quad \theta \in [0, \pi] \\ \det(v, f(v), u) \leq 0 & \Leftrightarrow \quad \theta \in [\pi, 2\pi] \end{cases} \tag{3}$$

- Case 2.b: $\theta \neq 0$, $f = S \circ R$. Orient rotation axis by u. Then if $\mathbf{u} = \{u_1, u_2, u_3 = \frac{u}{\|u\|}\}$ is a direct b.o.n (basis adapted to f),

$$M_{\mathbf{u}}(f) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The matrix of f in any other basis can be obtained by change of basis.

- $[u]$ = VEPs of VAP -1.
- H invariant subspace of dimension 2.
- $\theta = \widehat{v, f(v)}$ if v orthogonal to axis $[u]$.
- $\cos(\theta) = \frac{tr(f)+1}{2}$
- We have θ in $[0, \pi]$ or $[\pi, 2\pi]$ according to (for any $v \notin [u]$):

$$\begin{cases} \det(v, f(v), u) \geq 0 & \Leftrightarrow \quad \theta \in [0, \pi] \\ \det(v, f(v), u) \leq 0 & \Leftrightarrow \quad \theta \in [\pi, 2\pi] \end{cases} \tag{3}$$

Example

$$M_e(f) = \begin{pmatrix} 2/3 & 2/3 & 1/3 \\ -2/3 & 1/3 & 2/3 \\ 1/3 & -2/3 & 2/3 \end{pmatrix}$$

- $\det(f) = 1 \Rightarrow f =$ rotation of axis $[u]$ and angle θ.
- Axis: VEPs of eigenvalue 1, $r = [u = (1,0,1)]$.
- Angle: $1 + 2\cos\theta = tr(f) = 5/3 \Rightarrow \cos\theta = 1/3$
- Orient r by u, take $v = (1,0,0) \notin r$,
  $\det(v, f(v), u) = -2/3 < 0 \Rightarrow \theta \in [\pi, 2\pi]$, $\theta = 2\pi - 1.23$.

Example

$$M_e(f) = \begin{pmatrix} 2/3 & 2/3 & 1/3 \\ -2/3 & 1/3 & 2/3 \\ 1/3 & -2/3 & 2/3 \end{pmatrix}$$

- $\det(f) = 1 \Rightarrow f =$ rotation of axis $[u]$ and angle θ.
- Axis: VEPs of eigenvalue 1, $r = [u = (1, 0, 1)]$.
- Angle: $1 + 2\cos\theta = tr(f) = 5/3 \Rightarrow \cos\theta = 1/3$
- Orient r by u, take $v = (1, 0, 0) \notin r$,
  $\det(v, f(v), u) = -2/3 < 0 \Rightarrow \theta \in [\pi, 2\pi]$, $\theta = 2\pi - 1.23$.

Example

$$M_e(f) = \begin{pmatrix} 2/3 & 2/3 & 1/3 \\ -2/3 & 1/3 & 2/3 \\ 1/3 & -2/3 & 2/3 \end{pmatrix}$$

- $\det(f) = 1 \Rightarrow f =$ rotation of axis $[u]$ and angle θ.
- Axis: VEPs of eigenvalue 1, $r = [u = (1, 0, 1)]$.
- Angle: $1 + 2\cos\theta = tr(f) = 5/3 \Rightarrow \cos\theta = 1/3$
- Orient r by u, take $v = (1, 0, 0) \notin r$,
  $\det(v, f(v), u) = -2/3 < 0 \Rightarrow \theta \in [\pi, 2\pi]$, $\theta = 2\pi - 1.23$.

Example

$$M_e(f) = \begin{pmatrix} 2/3 & 2/3 & 1/3 \\ -2/3 & 1/3 & 2/3 \\ 1/3 & -2/3 & 2/3 \end{pmatrix}$$

- $\det(f) = 1 \Rightarrow f =$ rotation of axis $[u]$ and angle θ.
- Axis: VEPs of eigenvalue 1, $r = [u = (1, 0, 1)]$.
- Angle: $1 + 2\cos\theta = tr(f) = 5/3 \Rightarrow \cos\theta = 1/3$
- Orient r by u, take $v = (1, 0, 0) \notin r$,
  $\det(v, f(v), u) = -2/3 < 0 \Rightarrow \theta \in [\pi, 2\pi]$, $\theta = 2\pi - 1.23$.

Outline

Outline

SVD and rank approximation

Theorem (Eckhart-Young)

Let A be any matrix. If $A = UDV^t$ and the singular values of A are $\sigma_1, \ldots, \sigma_r$ then for any $k \leq r$,

$$M = U \begin{pmatrix} \sigma_1 & & 0 & \cdots & 0 \\ & \ddots & & & \vdots \\ 0 & & \sigma_k & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix} V^t$$

is the matrix of rank k closest to A (in the sense that $\|A - M\|_2$ is minimal among matrices M of rank k). Note that $||A - M||_2 = \sigma_{k+1}$.

This is used in image compression, for example. Note that $A = \sigma_1 u_1 v_1^t + \sigma_2 u_2 v_2^t + \ldots + \sigma_r u_r v_r^t$.

Outline

Linear least squares approximation

Problem: $Ax = b$ might be incompatible due to measure errors in b, but we would still like to have an approximated solution:

$$\begin{array}{c}\text{Incompatible}\\ Ax = b \\ \text{system}\end{array} \quad \Leftrightarrow \quad b \notin \operatorname{Im}(A)$$

Want: $\tilde{x}$ such that $A\tilde{x}$ is as close to b as possible.

Definition

A **least squares solution of $Ax = b$** is a vector $\tilde{x}$ that minimizes $\|Ax - b\|$, that is

$$\|A\tilde{x} - b\| \leq \|Ax - b\| \text{ for all } x$$

Solution to the least squares problem

Solution given by Gauss (1801)

- Change b by the vector of $\mathrm{Im}(A)$ that is closest to b: the *orthogonal projection* of b in $\mathrm{Im}(A)$, $proj_{\mathrm{Im}(A)}(b)$.
- Then $\tilde{x}$ is a least squares solution $\Leftrightarrow$ $\tilde{x}$ is a solution of $Ax = proj_{\mathrm{Im}(A)}(b)$.
- If x is a least squares solution then it does not satisfy $Ax - b = \vec{0}$, but minimizes the norm $\|Ax - b\|$
- The **residual** measures how far $\tilde{x}$ is from a solution to the system:

  $$\text{residual} = A\tilde{x} - b \text{ (which is } = proj_{\mathrm{Im}(A)}(b) - b).$$

  norm of the residual: $\|A\tilde{x} - b\|$
- Important point: we do not need to compute $proj_{\mathrm{Im}(A)}(b)$ (see next slide).

Solution to the least squares problem

Solution given by Gauss (1801)

- Change b by the vector of $\text{Im}(A)$ that is closest to b: the *orthogonal projection* of b in $\text{Im}(A)$, $proj_{\text{Im}(A)}(b)$.
- Then $\tilde{x}$ is a least squares solution $\Leftrightarrow$ $\tilde{x}$ is a solution of $Ax = proj_{\text{Im}(A)}(b)$.
- If x is a least squares solution then it does not satisfy $Ax - b = \vec{0}$, but minimizes the norm $\|Ax - b\|$
- The **residual** measures how far $\tilde{x}$ is from a solution to the system:

  residual $= A\tilde{x} - b$ (which is $= proj_{\text{Im}(A)}(b) - b$).

  norm of the residual: $\|A\tilde{x} - b\|$
- Important point: we do not need to compute $proj_{\text{Im}(A)}(b)$ (see next slide).

Solution to the least squares problem

Solution given by Gauss (1801)

- Change b by the vector of $\mathrm{Im}(A)$ that is closest to b: the *orthogonal projection* of b in $\mathrm{Im}(A)$, $proj_{\mathrm{Im}(A)}(b)$.
- Then $\tilde{x}$ is a least squares solution $\Leftrightarrow$ $\tilde{x}$ is a solution of $Ax = proj_{\mathrm{Im}(A)}(b)$.
- If x is a least squares solution then it does not satisfy $Ax - b = \vec{0}$, but minimizes the norm $\|Ax - b\|$
- The **residual** measures how far $\tilde{x}$ is from a solution to the system:

  $$\text{residual} = A\tilde{x} - b \text{ (which is } = proj_{\mathrm{Im}(A)}(b) - b).$$

  norm of the residual: $\|A\tilde{x} - b\|$
- Important point: we do not need to compute $proj_{\mathrm{Im}(A)}(b)$ (see next slide).

Solution to the least squares problem

Solution given by Gauss (1801)

- Change b by the vector of $\mathrm{Im}(A)$ that is closest to b: the *orthogonal projection* of b in $\mathrm{Im}(A)$, $proj_{\mathrm{Im}(A)}(b)$.
- Then $\tilde{x}$ is a least squares solution $\Leftrightarrow$ $\tilde{x}$ is a solution of $Ax = proj_{\mathrm{Im}(A)}(b)$.
- If x is a least squares solution then it does not satisfy $Ax - b = \vec{0}$, but minimizes the norm $\|Ax - b\|$
- The **residual** measures how far $\tilde{x}$ is from a solution to the system:

  residual $= A\tilde{x} - b$ (which is $= proj_{\mathrm{Im}(A)}(b) - b$).

  norm of the residual: $\|A\tilde{x} - b\|$
- Important point: we do not need to compute $proj_{\mathrm{Im}(A)}(b)$ (see next slide).

Solution to the least squares problem

Solution given by Gauss (1801)

- Change b by the vector of $\text{Im}(A)$ that is closest to b: the *orthogonal projection* of b in $\text{Im}(A)$, $proj_{\text{Im}(A)}(b)$.
- Then $\tilde{x}$ is a least squares solution $\Leftrightarrow$ $\tilde{x}$ is a solution of $Ax = proj_{\text{Im}(A)}(b)$.
- If x is a least squares solution then it does not satisfy $Ax - b = \vec{0}$, but minimizes the norm $\|Ax - b\|$
- The **residual** measures how far $\tilde{x}$ is from a solution to the system:

$$\text{residual} = A\tilde{x} - b \text{ (which is } = proj_{\text{Im}(A)}(b) - b).$$

  norm of the residual: $\|A\tilde{x} - b\|$

- Important point: we do not need to compute $proj_{\text{Im}(A)}(b)$ (see next slide).

Theorem

- *$\tilde{x}$ is a least squares solution of $Ax = b$ if and only if it is a solution of the* **normal equations***:*

$$A^t Ax = A^t b.$$

- *If the rank of A equals the number of columns, then $A^t A$ is invertible and the least squares solution is unique and given by*

$$\tilde{x} = (A^t A)^{-1} A^t b$$

(although computing the inverse is not efficient)

- *If the original system is compatible, $\tilde{x}$ is a solution to the original system as well.*

Theorem

- *$\tilde{x}$ is a least squares solution of $Ax = b$ if and only if it is a solution of the* **normal equations***:*

$$A^t Ax = A^t b.$$

- *If the rank of A equals the number of columns, then $A^t A$ is invertible and the least squares solution is unique and given by*

$$\tilde{x} = (A^t A)^{-1} A^t b$$

(although computing the inverse is not efficient)

- *If the original system is compatible, $\tilde{x}$ is a solution to the original system as well.*

Theorem

- *$\tilde{x}$ is a least squares solution of $Ax = b$ if and only if it is a solution of the* **normal equations***:*

$$A^t Ax = A^t b.$$

- *If the rank of A equals the number of columns, then $A^t A$ is invertible and the least squares solution is unique and given by*

$$\tilde{x} = (A^t A)^{-1} A^t b$$

  (although computing the inverse is not efficient)

- *If the original system is compatible, $\tilde{x}$ is a solution to the original system as well.*

Theorem

- *$\tilde{x}$ is a least squares solution of $Ax = b$ if and only if it is a solution of the* **normal equations***:*

$$A^t A x = A^t b.$$

- *If the rank of A equals the number of columns, then $A^t A$ is invertible and the least squares solution is unique and given by*

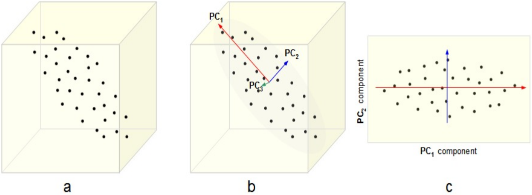
$$\tilde{x} = (A^t A)^{-1} A^t b$$

  (although computing the inverse is not efficient)
- *If the original system is compatible, $\tilde{x}$ is a solution to the original system as well.*

Outline

Principal component analysis

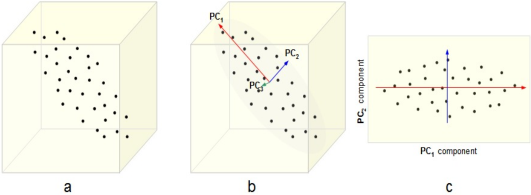
Goal: Given N data points in $\mathbb{R}^3$, $p_i = (x_i, y_i, z_i)$, $i = 1, \dots, N$ highly correlated, one wants to find $v_1 = (a, b, c)$ of norm 1 such that the set $\{t_i = ax_i + by_i + cz_i\}_i$ has maximum variance:



- Note that $proj_{[v_1]}(p_i) = t_i v_1$
- $v_1 = (a, b, c)$ is called the first principal component.
- Then one can look for $v_2 \in [v_1]^{\perp}$ (2nd principal component) maximizing variance of $proj_{[v_1]^{\perp}}(p_i)$.
- Keep going or project down to the first components in order to reduce the dimension of the problem.

Principal component analysis

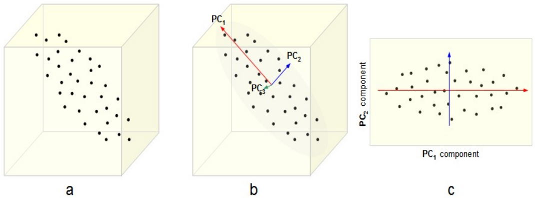
Goal: Given N data points in $\mathbb{R}^3$, $p_i = (x_i, y_i, z_i)$, $i = 1, \ldots, N$ highly correlated, one wants to find $v_1 = (a, b, c)$ of norm 1 such that the set $\{t_i = ax_i + by_i + cz_i\}_i$ has maximum variance:



- Note that $proj_{[v_1]}(p_i) = t_i v_1$
- $v_1 = (a, b, c)$ is called the first principal component.
- Then one can look for $v_2 \in [v_1]^\perp$ (2nd principal component) maximizing variance of $proj_{[v_1]^\perp}(p_i)$.
- Keep going or project down to the first components in order to reduce the dimension of the problem.

Principal component analysis

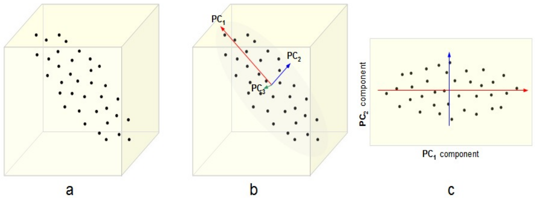
Goal: Given N data points in $\mathbb{R}^3$, $p_i = (x_i, y_i, z_i)$, $i = 1, \dots, N$ highly correlated, one wants to find $v_1 = (a, b, c)$ of norm 1 such that the set $\{t_i = ax_i + by_i + cz_i\}_i$ has maximum variance:



- Note that $proj_{[v_1]}(p_i) = t_i v_1$
- $v_1 = (a, b, c)$ is called the first principal component.
- Then one can look for $v_2 \in [v_1]^\perp$ (2nd principal component) maximizing variance of $proj_{[v_1]^\perp}(p_i)$.
- Keep going or project down to the first components in order to reduce the dimension of the problem.

Principal component analysis

Goal: Given N data points in $\mathbb{R}^3$, $p_i = (x_i, y_i, z_i)$, $i = 1, \dots, N$ highly correlated, one wants to find $v_1 = (a, b, c)$ of norm 1 such that the set $\{t_i = ax_i + by_i + cz_i\}_i$ has maximum variance:



- Note that $proj_{[v_1]}(p_i) = t_i v_1$
- $v_1 = (a, b, c)$ is called the first principal component.
- Then one can look for $v_2 \in [v_1]^\perp$ (2nd principal component) maximizing variance of $proj_{[v_1]^\perp}(p_i)$.
- Keep going or project down to the first components in order to reduce the dimension of the problem.

Procedure

Assume that set $\{p_i\}$ is centered at the origin. Let

$$M = \begin{pmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{pmatrix} \text{ so that } \sum_i x_i = \sum_i y_i = \sum_i z_i = 0.$$

- Want $v_1 = (a, b, c)$ of norm 1 such that $\sum_i t_i^2 = \sum_i (ax_i + by_i + cz_i)^2 = ||Mv_1||$ is maximum.
- v_1 is the first column vector of V in the SVD: $M = UDV^t$.
- Then the matrix $M_2 = M - Mv_1 v_1^t$ has $proj_{[v_1]^\perp}(p_i)$ in its rows.
- $M_2 = \sigma_2 u_2 v_2^t + \ldots + \sigma_r u_r v_r^t$.
- The direction which maximizes the variance is v_2 (2nd vector in V).
- Keep going.

Procedure

Assume that set $\{p_i\}$ is centered at the origin. Let

$$M = \begin{pmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{pmatrix}$$ so that $\sum_i x_i = \sum_i y_i = \sum_i z_i = 0$.

- Want $v_1 = (a, b, c)$ of norm 1 such that $\sum_i t_i^2 = \sum_i (ax_i + by_i + cz_i)^2 = ||Mv_1||$ is maximum.
- v_1 is the first column vector of V in the SVD: $M = UDV^t$.
- Then the matrix $M_2 = M - Mv_1v_1^t$ has $proj_{[v_1]^\perp}(p_i)$ in its rows.
- $M_2 = \sigma_2 u_2 v_2^t + \ldots + \sigma_r u_r v_r^t$.
- The direction which maximizes the variance is v_2 (2nd vector in V).
- Keep going.

Procedure

Assume that set $\{p_i\}$ is centered at the origin. Let

$$M = \begin{pmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{pmatrix} \text{ so that } \sum_i x_i = \sum_i y_i = \sum_i z_i = 0.$$

- Want $v_1 = (a, b, c)$ of norm 1 such that $\sum_i t_i^2 = \sum_i (ax_i + by_i + cz_i)^2 = ||Mv_1||$ is maximum.
- v_1 is the first column vector of V in the SVD: $M = UDV^t$.
- Then the matrix $M_2 = M - Mv_1v_1^t$ has $proj_{[v_1]^\perp}(p_i)$ in its rows.
- $M_2 = \sigma_2 u_2 v_2^t + \ldots + \sigma_r u_r v_r^t$.
- The direction which maximizes the variance is v_2 (2nd vector in V).
- Keep going.

Procedure

Assume that set $\{p_i\}$ is centered at the origin. Let

$$M = \begin{pmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{pmatrix}$$ so that $\sum_i x_i = \sum_i y_i = \sum_i z_i = 0.$

- Want $v_1 = (a, b, c)$ of norm 1 such that $\sum_i t_i^2 = \sum_i (ax_i + by_i + cz_i)^2 = ||Mv_1||$ is maximum.
- v_1 is the first column vector of V in the SVD: $M = UDV^t$.
- Then the matrix $M_2 = M - Mv_1 v_1^t$ has $proj_{[v_1]^\perp}(p_i)$ in its rows.
- $M_2 = \sigma_2 u_2 v_2^t + \ldots + \sigma_r u_r v_r^t$.
- The direction which maximizes the variance is v_2 (2nd vector in V).
- Keep going.

Procedure

Assume that set $\{p_i\}$ is centered at the origin. Let

$M = \begin{pmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{pmatrix}$ so that $\sum_i x_i = \sum_i y_i = \sum_i z_i = 0$.

- Want $v_1 = (a, b, c)$ of norm 1 such that $\sum_i t_i^2 = \sum_i (ax_i + by_i + cz_i)^2 = ||Mv_1||$ is maximum.
- v_1 is the first column vector of V in the SVD: $M = UDV^t$.
- Then the matrix $M_2 = M - Mv_1v_1^t$ has $proj_{[v_1]^\perp}(p_i)$ in its rows.
- $M_2 = \sigma_2 u_2 v_2^t + \ldots + \sigma_r u_r v_r^t$.
- The direction which maximizes the variance is v_2 (2nd vector in V).
- Keep going.

Procedure

Assume that set $\{p_i\}$ is centered at the origin. Let

$$M = \begin{pmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & z_N \end{pmatrix} \text{ so that } \sum_i x_i = \sum_i y_i = \sum_i z_i = 0.$$

- Want $v_1 = (a, b, c)$ of norm 1 such that $\sum_i t_i^2 = \sum_i (ax_i + by_i + cz_i)^2 = ||Mv_1||$ is maximum.
- v_1 is the first column vector of V in the SVD: $M = UDV^t$.
- Then the matrix $M_2 = M - Mv_1 v_1^t$ has $proj_{[v_1]^\perp}(p_i)$ in its rows.
- $M_2 = \sigma_2 u_2 v_2^t + \ldots + \sigma_r u_r v_r^t$.
- The direction which maximizes the variance is v_2 (2nd vector in V).
- Keep going.

Remarks:

- If the set $\{p_i\}$ is not centered at the origin we center it: let $(\bar{x}, \bar{y}, \bar{z}) = \sum_i (x_i, y_i, z_i)/N$, and consider

$$M = \begin{pmatrix} x_1 - \bar{x} & y_1 - \bar{y} & z_1 - \bar{z} \\ \vdots & \vdots & \vdots \\ x_N - \bar{x} & y_N - \bar{y} & z_N - \bar{z} \end{pmatrix}.$$

  Proceed as before with this M and then sum $(\bar{x}, \bar{y}, \bar{z})$ to the final result.
- The matrix $M^t M$ is the *empirical covariance* matrix and the principal component v_1 is the dominant eigenvector of this matrix.
- The same can be done for clouds of points in $\mathbb{R}^n$.

Remarks:

- If the set $\{p_i\}$ is not centered at the origin we center it: let $(\bar{x}, \bar{y}, \bar{z}) = \sum_i (x_i, y_i, z_i)/N$, and consider

$$M = \begin{pmatrix} x_1 - \bar{x} & y_1 - \bar{y} & z_1 - \bar{z} \\ \vdots & \vdots & \vdots \\ x_N - \bar{x} & y_N - \bar{y} & z_N - \bar{z} \end{pmatrix}.$$

  Proceed as before with this M and then sum $(\bar{x}, \bar{y}, \bar{z})$ to the final result.

- The matrix $M^t M$ is the *empirical covariance* matrix and the principal component v_1 is the dominant eigenvector of this matrix.

- The same can be done for clouds of points in $\mathbb{R}^n$.

Remarks:

- If the set $\{p_i\}$ is not centered at the origin we center it: let $(\bar{x}, \bar{y}, \bar{z}) = \sum_i (x_i, y_i, z_i)/N$, and consider

$$M = \begin{pmatrix} x_1 - \bar{x} & y_1 - \bar{y} & z_1 - \bar{z} \\ \vdots & \vdots & \vdots \\ x_N - \bar{x} & y_N - \bar{y} & z_N - \bar{z} \end{pmatrix}.$$

  Proceed as before with this M and then sum $(\bar{x}, \bar{y}, \bar{z})$ to the final result.
- The matrix $M^t M$ is the *empirical covariance* matrix and the principal component v_1 is the dominant eigenvector of this matrix.
- The same can be done for clouds of points in $\mathbb{R}^n$.

Outline

Scalar product in $\mathbb{C}^n$

Definition

In $\mathbb{C}^n$ the analogous to the dot product is the **standard hermitian product** $< u, v >$ of two vectors $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, v = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{C}^n$ is

$$< u, v > := u^t \overline{v} = x_1 \overline{y_1} + x_2 \overline{y_2} + \ldots + x_n \overline{y_n}.$$

Example:

$$u = \begin{pmatrix} 1 \\ i \\ 1-2i \end{pmatrix}, v = \begin{pmatrix} i \\ 0 \\ 3 \end{pmatrix} \Rightarrow < u, v > = (1 \; i \; 1-2i) \begin{pmatrix} -i \\ 0 \\ 3 \end{pmatrix} = 3 - 7i.$$

Scalar product in $\mathbb{C}^n$

Definition

In $\mathbb{C}^n$ the analogous to the dot product is the **standard hermitian product** $< u, v >$ of two vectors $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, v = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{C}^n$ is

$$< u, v > := u^t \overline{v} = x_1 \overline{y_1} + x_2 \overline{y_2} + \ldots + x_n \overline{y_n}.$$

Example:

$$u = \begin{pmatrix} 1 \\ i \\ 1-2i \end{pmatrix}, v = \begin{pmatrix} i \\ 0 \\ 3 \end{pmatrix} \Rightarrow < u, v > = (1 \ i \ 1-2i) \begin{pmatrix} -i \\ 0 \\ 3 \end{pmatrix} = 3 - 7i.$$

Properties:

1. $< u, u > \geq 0 \; \forall u$ (positive)
2. $< u, u >= 0 \Leftrightarrow u = 0$ (non-degenerate).
3. $< u, v >= \overline{< v, u >}$ ("hermitian").
4. "sesquilineal":
   - $< a_1 u_1 + a_2 u_2, v >= a_1 < u_1, v > + a_2 < u_2, v >$
   - $< u, a_1 v_1 + a_2 v_2 >= \overline{a_1} < u, v_1 > + \overline{a_2} < u, v_2 >$
5. In this case, the norm of a vector $u \in \mathbb{C}^n$ is $\|u\| = \sqrt{u^t \overline{u}} = \sqrt{|u_1|^2 + \ldots + |u_n|^2} \in \mathbb{R}$.
6. Orthonormal basis for this scalar product: same definition as before.
7. If we write the columns of a matrix $A = (v_1 \ldots v_d)$ then,

$$A^t \overline{A} = Id \quad \text{if and ony if } \{v_1, \ldots, v_d\} \text{ is an orthonormal basis.}$$

Properties:

1. $< u, u > \geq 0 \ \forall u$ (positive)
2. $< u, u >= 0 \Leftrightarrow u = 0$ (non-degenerate).
3. $< u, v >= \overline{< v, u >}$ ("hermitian").
4. "sesquilineal":
   - $< a_1 u_1 + a_2 u_2, v >= a_1 < u_1, v > + a_2 < u_2, v >$;
   - $< u, a_1 v_1 + a_2 v_2 >= \overline{a_1} < u, v_1 > + \overline{a_2} < u, v_2 >$.
5. In this case, the norm of a vector $u \in \mathbb{C}^n$ is $\|u\| = \sqrt{u^t \overline{u}} = \sqrt{|u_1|^2 + \ldots + |u_n|^2} \in \mathbb{R}$.
6. Orthonormal basis for this scalar product: same definition as before.
7. If we write the columns of a matrix $A = (v_1 \ldots v_d)$ then,

$$A^t \overline{A} = Id \quad \text{if and ony if } \{v_1, \ldots, v_d\} \text{ is an orthonormal basis.}$$

Properties:

1. $< u, u > \geq 0 \ \forall u$ (positive)
2. $< u, u >= 0 \Leftrightarrow u = 0$ (non-degenerate).
3. $< u, v >= \overline{< v, u >}$ ("hermitian").
4. "sesquilineal":
   - $< a_1 u_1 + a_2 u_2, v >= a_1 < u_1, v > + a_2 < u_2, v >$;
   - $< u, a_1 v_1 + a_2 v_2 >= \overline{a_1} < u, v_1 > + \overline{a_2} < u, v_2 >$.
5. In this case, the norm of a vector $u \in \mathbb{C}^n$ is $\|u\| = \sqrt{u^t \overline{u}} = \sqrt{|u_1|^2 + \ldots + |u_n|^2} \in \mathbb{R}$.
6. Orthonormal basis for this scalar product: same definition as before.
7. If we write the columns of a matrix $A = (v_1 \ldots v_d)$ then,

$$A^t \overline{A} = Id \quad \text{if and ony if } \{v_1, \ldots, v_d\} \text{ is an orthonormal basis.}$$

Properties:

1. $< u, u > \geq 0 \ \forall u$ (positive)
2. $< u, u >= 0 \Leftrightarrow u = 0$ (non-degenerate).
3. $< u, v >= \overline{< v, u >}$ ("hermitian").
4. "sesquilineal":
   - $< a_1 u_1 + a_2 u_2, v >= a_1 < u_1, v > + a_2 < u_2, v >$;
   - $< u, a_1 v_1 + a_2 v_2 >= \overline{a_1} < u, v_1 > + \overline{a_2} < u, v_2 >$.
5. In this case, the norm of a vector $u \in \mathbb{C}^n$ is $\|u\| = \sqrt{u^t \overline{u}} = \sqrt{|u_1|^2 + \ldots + |u_n|^2} \in \mathbb{R}$.
6. Orthonormal basis for this scalar product: same definition as before.
7. If we write the columns of a matrix $A = (v_1 \ldots v_d)$ then,

$$A^t \overline{A} = Id \quad \text{if and ony if } \{v_1, \ldots, v_d\} \text{ is an orthonormal basis.}$$

Properties:

1. $< u, u > \geq 0 \ \forall u$ (positive)
2. $< u, u >= 0 \Leftrightarrow u = 0$ (non-degenerate).
3. $< u, v >= \overline{< v, u >}$ ("hermitian").
4. "sesquilineal":
   - $< a_1 u_1 + a_2 u_2, v >= a_1 < u_1, v > + a_2 < u_2, v >$;
   - $< u, a_1 v_1 + a_2 v_2 >= \overline{a_1} < u, v_1 > + \overline{a_2} < u, v_2 >$.
5. In this case, the norm of a vector $u \in \mathbb{C}^n$ is $\|u\| = \sqrt{u^t \overline{u}} = \sqrt{|u_1|^2 + \ldots + |u_n|^2} \in \mathbb{R}$.
6. Orthonormal basis for this scalar product: same definition as before.
7. If we write the columns of a matrix $A = (v_1 \ldots v_d)$ then,

$$A^t \overline{A} = Id \quad \text{if and ony if } \{v_1, \ldots, v_d\} \text{ is an orthonormal basis.}$$

Properties:

1. $< u, u > \geq 0 \ \forall u$ (positive)
2. $< u, u >= 0 \Leftrightarrow u = 0$ (non-degenerate).
3. $< u, v >= \overline{< v, u >}$ ("hermitian").
4. "sesquilineal":
   - $< a_1 u_1 + a_2 u_2, v >= a_1 < u_1, v > + a_2 < u_2, v >$;
   - $< u, a_1 v_1 + a_2 v_2 >= \overline{a_1} < u, v_1 > + \overline{a_2} < u, v_2 >$.
5. In this case, the norm of a vector $u \in \mathbb{C}^n$ is $\|u\| = \sqrt{u^t \overline{u}} = \sqrt{|u_1|^2 + \ldots + |u_n|^2} \in \mathbb{R}$.
6. Orthonormal basis for this scalar product: same definition as before.
7. If we write the columns of a matrix $A = (v_1 \ldots v_d)$ then,

$$A^t \overline{A} = Id \quad \text{if and ony if } \{v_1, \ldots, v_d\} \text{ is an orthonormal basis.}$$

Properties:

1. $< u, u > \geq 0 \ \forall u$ (positive)
2. $< u, u >= 0 \Leftrightarrow u = 0$ (non-degenerate).
3. $< u, v >= \overline{< v, u >}$ ("hermitian").
4. "sesquilineal":
   - $< a_1 u_1 + a_2 u_2, v >= a_1 < u_1, v > + a_2 < u_2, v >$;
   - $< u, a_1 v_1 + a_2 v_2 >= \overline{a_1} < u, v_1 > + \overline{a_2} < u, v_2 >$.
5. In this case, the norm of a vector $u \in \mathbb{C}^n$ is $\|u\| = \sqrt{u^t \overline{u}} = \sqrt{|u_1|^2 + \ldots + |u_n|^2} \in \mathbb{R}$.
6. Orthonormal basis for this scalar product: same definition as before.
7. If we write the columns of a matrix $A = (v_1 \ldots v_d)$ then,

   $$A^t \overline{A} = Id \quad \text{if and ony if } \{v_1, \ldots, v_d\} \text{ is an orthonormal basis.}$$

Unitary matrices

An $n \times n$ matrix that satisfies $\overline{A}^t A = Id$ is called a **unitary** matrix.

- If we call the columns $v_1, \dots, v_n$, $A = (v_1 \dots v_n)$, then,

$$\overline{A}^t A = Id \quad \text{if and ony if } \{v_1, \dots, v_n\} \text{ is an orthonormal basis.}$$

- A is unitary if and only if $A^{-1} = \overline{A}^t$.
- A is unitary $\Rightarrow |\det A| = 1$.
- If A is unitary, then the corresponding endomorphism preserves norms (preserves the measure of vectors):

$$||Ax|| = ||x|| \text{ for all } x$$

- A also preserves dot products and angles (and hence preserves orthogonality) and so it is a transformation that does not deform objects.

Unitary matrices

An $n \times n$ matrix that satisfies $\overline{A}^t A = Id$ is called a **unitary** matrix.

- If we call the columns $v_1, \ldots, v_n$, $A = (v_1 \ldots v_n)$, then,

$$\overline{A}^t A = Id \quad \text{if and ony if } \{v_1, \ldots, v_n\} \text{ is an orthonormal basis.}$$

- A is unitary if and only if $A^{-1} = \overline{A}^t$.
- A is unitary $\Rightarrow |\det A| = 1$.
- If A is unitary, then the corresponding endomorphism **preserves norms** (preserves the measure of vectors):

$$\|Ax\| = \|x\| \text{ for all } x$$

- A also preserves dot products and angles (and hence preserves orthogonality) and so it is a transformation that does not deform objects.

Unitary matrices

An $n \times n$ matrix that satisfies $\overline{A}^t A = Id$ is called a **unitary** matrix.

- If we call the columns $v_1, \ldots, v_n$, $A = (v_1 \ldots v_n)$, then,

$$\overline{A}^t A = Id \quad \text{if and ony if } \{v_1, \ldots, v_n\} \text{ is an orthonormal basis.}$$

- A is unitary if and only if $A^{-1} = \overline{A}^t$.
- A is unitary $\Rightarrow |\det A| = 1$.
- If A is unitary, then the corresponding endomorphism preserves norms (preserves the measure of vectors):

$$||Ax|| = ||x|| \text{ for all } x$$

- A also preserves dot products and angles (and hence preserves orthogonality) and so it is a transformation that does not deform objects.

Unitary matrices

An $n \times n$ matrix that satisfies $\overline{A}^t A = Id$ is called a **unitary** matrix.

- If we call the columns $v_1, \ldots, v_n$, $A = (v_1 \ldots v_n)$, then,

$$\overline{A}^t A = Id \quad \text{if and ony if } \{v_1, \ldots, v_n\} \text{ is an orthonormal basis.}$$

- A is unitary if and only if $A^{-1} = \overline{A}^t$.
- A is unitary $\Rightarrow |\det A| = 1$.
- If A is unitary, then the corresponding endomorphism preserves norms (preserves the measure of vectors):

$$\|Ax\| = \|x\| \text{ for all } x$$

- A also preserves dot products and angles (and hence preserves orthogonality) and so it is a transformation that does not deform objects.

Unitary matrices

An $n \times n$ matrix that satisfies $\overline{A}^t A = Id$ is called a **unitary** matrix.

- If we call the columns $v_1, \ldots, v_n$, $A = (v_1 \ldots v_n)$, then,

$$\overline{A}^t A = Id \quad \text{if and ony if } \{v_1, \ldots, v_n\} \text{ is an orthonormal basis.}$$

- A is unitary if and only if $A^{-1} = \overline{A}^t$.
- A is unitary $\Rightarrow |\det A| = 1$.
- If A is unitary, then the corresponding endomorphism preserves norms (preserves the measure of vectors):

$$\|Ax\| = \|x\| \text{ for all } x$$

- A also preserves dot products and angles (and hence preserves orthogonality) and so it is a transformation that does not deform objects.

Outline

Bibliography

Basic:

- D. Poole, Linear Algebra, A modern introduction (3rd edition), Brooks/Cole, 2011. Chapter 6.

Additional

- Hernández Rodríguez, E.; Vàzquez Gallo, M.J.; Zurro Moro, M.A. Álgebra lineal y geometría [en línia]