Linear Algebra

Lecture Notes

Rostyslav Hryniv

Ukrainian Catholic University
Business Analytics and Computer Science Programmes

4^{th} term
Spring 2020

Lecture 7. Orthogonal vectors and subspaces

Outline

Outline

Distances in $\mathbb{R}^n$

- In $\mathbb{R}^n$ (and in other vector spaces) one can measure the distance $\rho(\mathbf{x}, \mathbf{y})$ between any two vectors $\mathbf{x} = (x_1, \ldots, x_n)^\top$, $\mathbf{y} = (y_1, \ldots, y_n)^\top$
- The standard choice is the Euclidean distance ρ_2:

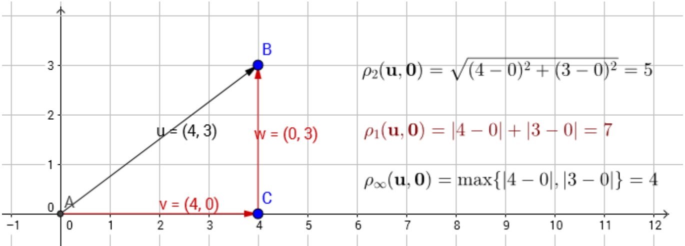
$$\rho_2(\mathbf{x}, \mathbf{y}) := \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$$

- Other possibilities: the block, or Manhattan distance:

$$\rho_1(\mathbf{x}, \mathbf{y}) = |x_1 - y_1| + |x_2 - y_2| + \cdots + |x_n - y_n|$$

or maximum-coordinate distance

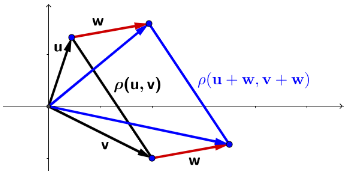
$$\rho_\infty(\mathbf{x}, \mathbf{y}) = \max\{|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n|\}$$



Distances and norms

Any distance must satisfy the following conditions:

- $\rho(\mathbf{x},\mathbf{y}) \geq 0$ and $\rho(\mathbf{x},\mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$ — positivity
- $\rho(\mathbf{x},\mathbf{y}) = \rho(\mathbf{y},\mathbf{x})$ — symmetry
- $\rho(\mathbf{x},\mathbf{y}) \leq \rho(\mathbf{x},\mathbf{z}) + \rho(\mathbf{z},\mathbf{y})$ — triangle inequality



In linear vector spaces, distance should be translation invariant, i.e.,

$$\rho(\mathbf{u},\mathbf{v}) = \rho(\mathbf{u}+\mathbf{w},\mathbf{v}+\mathbf{w})$$

and thus only depends on $\mathbf{u}-\mathbf{v}$.

Thus

$$\rho(\mathbf{u},\mathbf{v}) = \rho(\mathbf{u}-\mathbf{v},\mathbf{0}) =: \|\mathbf{u}-\mathbf{v}\|,$$

where $\|\mathbf{x}\|$ is the norm, or length of $\mathbf{x}$.

Distances and norms

Definition (Norm)

A norm $\|\cdot\|$ on a linear vector space V is a function $V \to \mathbb{R}$ with the following properties:

- $\|\mathbf{x}\| \geq 0$; moreover, $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$ — positivity
- $\|k\mathbf{x}\| = |k|\|\mathbf{x}\|$ for all $k \in \mathbb{R}$ and all $\mathbf{x} \in V$ — linearity
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ — triangle inequality

Example

- The Euclidean norm (will prove below): $\|\mathbf{x}\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$
- The block, or Manhattan norm: $\|\mathbf{x}\|_1 = |x_1| + \cdots + |x_n|$
- The maximum norm: $\|\mathbf{x}\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$

Euclidean norm and inner product

The Euclidean norm $\|\cdot\|_2$ is the most important one:

it comes from the scalar product of vectors,

$$\|\mathbf{x}\|_2^2 = x_1^2 + \cdots + x_n^2 = \mathbf{x}^\top \mathbf{x}$$

Recall that the inner, dot, or scalar product of vectors $\mathbf{x} = (x_1, \ldots, x_n)^\top$, $\mathbf{y} = (y_1, \ldots, y_n)^\top$ in $\mathbb{R}^n$ is

$$\boxed{\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y} = x_1 y_1 + \cdots + x_n y_n}$$

Observe that $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$

Remark

Observe that $\mathbf{x} \cdot \mathbf{y}$ is a bilinear symmetric form.
A general bilinear form is given by $\langle \mathbf{x}, \mathbf{y} \rangle_A = \mathbf{x}^\top A \mathbf{y}$ for an $n \times n$ matrix A.
The form $\langle \cdot, \cdot \rangle_A$ defines a norm $\|\cdot\|_A$ if A is positive definite
The inner products $\langle \cdot, \cdot \rangle_A$ are used in the conjugate gradient method

Inner products in linear vector spaces

An inner product in a linear vector space V is
a symmetric positive definite biliner form $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$

- $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$; moreover, $\langle \mathbf{u}, \mathbf{u} \rangle = 0 \iff \mathbf{u} = \mathbf{0}$ positivity
- $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ symmetry
- $\langle c_1\mathbf{u}_1 + c_2\mathbf{u}_2, \mathbf{v} \rangle = c_1\langle \mathbf{u}_1, \mathbf{v} \rangle + c_2\langle \mathbf{u}_2, \mathbf{v} \rangle$ linearity

Example

In the space $\mathcal{P}_\infty$, introduce the inner product via

$$\langle p, q \rangle := \int_{-1}^{1} p(t)q(t)\,dt$$

Another possibility: for $p(t) = p_0 + p_1 t + \cdots + p_n t^n$
and $q(t) = q_0 + q_1 t + \cdots + q_m t^m$, set

$$\langle p, q \rangle := \sum_{k \geq 0} p_k q_k$$

Inner products in linear vector spaces

Question

In the linear vector space $V = \mathcal{P}_n$, consider the bilinear form

$$\langle p, q\rangle := p(0)q(0) + p(1)q(1) + \cdots + p(10)q(10)$$

For what values of n is that a scalar product in $\mathcal{P}_n$?

- A *for every $n \in \mathbb{N}$*
- B *for no values of $n \in \mathbb{N}$*
- C *for $n = 10$ only*
- D *for $n = 1, 2, \ldots, 9$*
- E *none of the above*

Voting instructions:

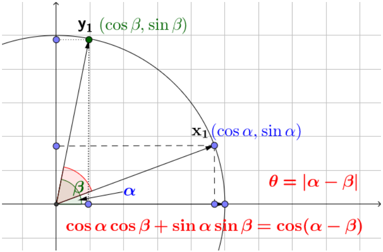
1. Go to www.socrative.com
2. Choose Student login
3. Enter the class LAUCU2020

Outline

Cosine theorem: $n = 2$

$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\| \cos\theta$, with θ the angle between $\mathbf{x}$ and $\mathbf{y}$

- Set $\mathbf{x}_1 := \mathbf{x}/\|\mathbf{x}\|$ and $\mathbf{y}_1 := \mathbf{y}/\|\mathbf{y}\| \implies$
$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|(\mathbf{x}_1 \cdot \mathbf{y}_1)$$
- $\|\mathbf{x}_1\| = \|\mathbf{y}_1\| = 1 \implies \exists \alpha, \beta$ s.t.
$$\mathbf{x}_1 = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}, \qquad \mathbf{y}_1 = \begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix}$$
- Then $\mathbf{x}_1 \cdot \mathbf{y}_1 = \cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\alpha - \beta)$
- $\theta := |\alpha - \beta|$ is the angle between $\mathbf{x}$ and $\mathbf{y}$



Cosine theorem: general case

Reduction to the case $n = 2$ (isometric transformation):

- Consider the plane V through $\mathbf{x}$ and $\mathbf{y}$ and choose any orthonormal basis $E' = (\mathbf{e}'_1, \mathbf{e}'_2)$ in V
- orthonormal means $\mathbf{e}'_1 \cdot \mathbf{e}'_2 = 0$ and $\|\mathbf{e}'_1\| = \|\mathbf{e}'_2\| = 1$
- Let $(\mathbf{x})_{E'} = (x'_1, x'_2)^\top$ and $(\mathbf{y})_{E'} = (y'_1, y'_2)^\top$; then

$$\begin{aligned}\mathbf{x} &= P_{E'\to E}(\mathbf{x})_{E'} = (\mathbf{e}'_1\ \mathbf{e}'_2)(\mathbf{x})_{E'}\\ \mathbf{y} &= P_{E'\to E}(\mathbf{y})_{E'} = (\mathbf{e}'_1\ \mathbf{e}'_2)(\mathbf{y})_{E'}\end{aligned}$$

- Now

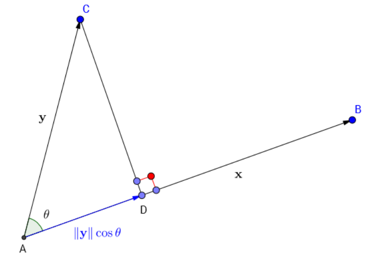
$$\begin{aligned}\mathbf{x}^\top\mathbf{y} &= (\mathbf{x})_{E'}^\top \begin{pmatrix}(\mathbf{e}'_1)^\top\\ (\mathbf{e}'_2)^\top\end{pmatrix} (\mathbf{e}'_1\ \mathbf{e}'_2)(\mathbf{y})_{E'}\\ &= (\mathbf{x})_{E'}^\top \begin{pmatrix}\mathbf{e}'_1\cdot\mathbf{e}'_1 & \mathbf{e}'_1\cdot\mathbf{e}'_2\\ \mathbf{e}'_2\cdot\mathbf{e}'_1 & \mathbf{e}'_2\cdot\mathbf{e}'_2\end{pmatrix}(\mathbf{y})_{E'} = (\mathbf{x})_{E'}^\top I_2 (\mathbf{y})_{E'}\end{aligned}$$

- Thus $\boxed{\mathbf{x}\cdot\mathbf{y} = (\mathbf{x})_{E'}\cdot(\mathbf{y})_{E'}}$

Geometric interpretation: projection

In $\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \|\mathbf{y}\| \cos\theta$,

- $\|\mathbf{y}\| \cos\theta$ is the length of projection of $\mathbf{y}$ onto $\mathbf{x}$-direction



- $|AC| \cos\theta = \pm|AD|$, so that $\mathbf{x} \cdot \mathbf{y} = \pm|AB||AD|$
- $\mathbf{x} \cdot \mathbf{y} > 0 \iff \theta < \pi/2$ (acute)
- $\mathbf{x} \cdot \mathbf{y} < 0 \iff \theta > \pi/2$ (obtuse)
- $\mathbf{x} \cdot \mathbf{y} = 0 \iff \theta = \pi/2$ (right)

Geometric interpretation: hyperplanes

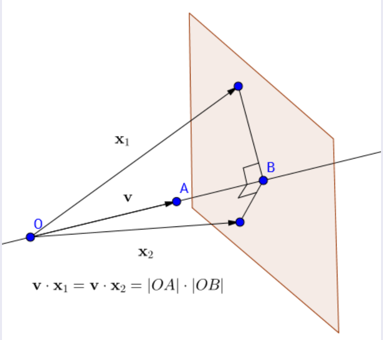
Identify a point $A(x_1, \dots, x_n)$ in $\mathbb{R}^n$
and a vector $\vec{OA} = \mathbf{x} = (x_1, \dots, x_n)^\top$

Hyperplane in $\mathbb{R}^n$

Fix $\mathbf{v} \in \mathbb{R}^n$ and $c \in \mathbb{R}$; then endpoints of vectors

$$\{\mathbf{x} \in \mathbb{R}^n \ : \ \mathbf{x} \cdot \mathbf{v} = c\}$$

form a hyperplane $v_1 x_1 + v_2 x_2 + \cdots + v_n x_n = c$ orthogonal to $\mathbf{v}$



Further properties of the scalar product in $\mathbb{R}^n$

To summarize:

- The standard inner, dot, or scalar product in $\mathbb{R}^n$ is

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y} = \sum_{k=1}^{n} x_k y_k$$

  in a general vector space V, it is

  a symmetric positive definite bilinear form:

  - $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}$ — positivity
  - $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ — symmetry
  - $\langle c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2, \mathbf{y} \rangle = c_1 \langle \mathbf{x}_1, \mathbf{y} \rangle + c_2 \langle \mathbf{x}_2, \mathbf{y} \rangle$ — linearity

- Euclidean norm, or length of a vector $\mathbf{x}$: $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$
- Transposition: for an $m \times n$ matrix A, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{y} \in \mathbb{R}^m$

$$\langle A\mathbf{x}, \mathbf{y} \rangle = (A\mathbf{x})^\top y = (\mathbf{x}^\top A^\top)\mathbf{y} = \mathbf{x}^\top (A^\top \mathbf{y}) = \langle \mathbf{x}, A^\top \mathbf{y} \rangle$$

  i.e., for every $m \times n$ matrix A and vectors $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$

$$\boxed{(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A^\top \mathbf{y})}$$

Outline

Basic inequalities and theorems

Cauchy–Bunyakovsky–Schwarz inequality: $\boxed{|\langle \mathbf{x}, \mathbf{y} \rangle| \le \|\mathbf{x}\| \|\mathbf{y}\|}$

- use that $f(t) = \|\mathbf{x} + t\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2t\langle \mathbf{x}, \mathbf{y} \rangle + t^2\|\mathbf{y}\|^2 \ge 0$
- the discriminant nonpositive $\implies \quad |\langle \mathbf{x}, \mathbf{y} \rangle|^2 \le \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$
- alternatively: $|\langle \mathbf{x}, \mathbf{y} \rangle| = \|\mathbf{x}\| \|\mathbf{y}\| |\cos\theta| \le \|\mathbf{x}\| \|\mathbf{y}\|$

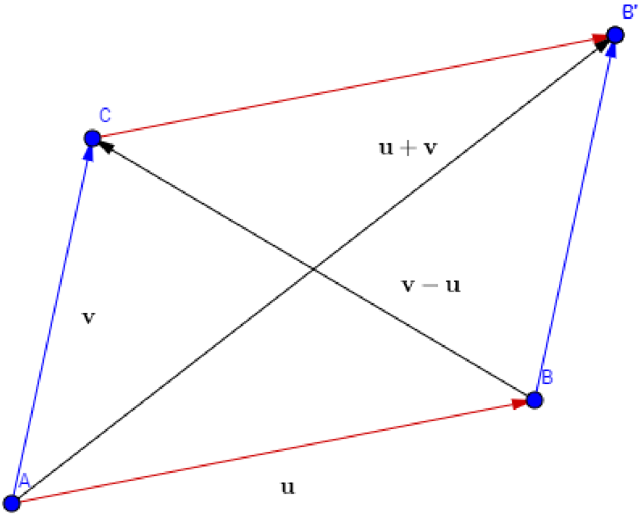
Triangle inequality: $\boxed{\|\mathbf{x} \pm \mathbf{y}\| \le \|\mathbf{x}\| + \|\mathbf{y}\|}$

This follows from bilinearity:

$$\begin{aligned} \|\mathbf{x} \pm \mathbf{y}\|^2 &= \langle \mathbf{x} \pm \mathbf{y}, \mathbf{x} \pm \mathbf{y} \rangle = \|\mathbf{x}\|^2 \pm 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 \\ &\le \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \|\mathbf{y}\| + \|\mathbf{y}\|^2 = \left(\|\mathbf{x}\| + \|\mathbf{y}\|\right)^2 \end{aligned}$$

Pythagorean theorem: $\mathbf{x} \cdot \mathbf{y} = 0 \iff \boxed{\|\mathbf{x} \pm \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2}$

Follows from the above equality for $\|\mathbf{x} \pm \mathbf{y}\|^2$



Parallelogram identity:

$$\|\mathbf{u}+\mathbf{v}\|^2+\|\mathbf{u}-\mathbf{v}\|^2=2\|\mathbf{u}\|^2+2\|\mathbf{v}\|^2$$

Orthogonal vectors and subspaces

Definitions:

- $\mathbf{x}$ and $\mathbf{y}$ are orthogonal ($\mathbf{x} \perp \mathbf{y}$) $\iff \langle \mathbf{x}, \mathbf{y} \rangle = 0$
- for a subspace M:

$$\mathbf{x} \perp M \iff \forall \mathbf{y} \in M : \langle \mathbf{x}, \mathbf{y} \rangle = 0$$

- for two subspaces L and M:

$$L \perp M \iff \forall \mathbf{x} \in L \, \forall \mathbf{y} \in M : \langle \mathbf{x}, \mathbf{y} \rangle = 0$$

- an orthogonal complement $L^{\perp}$ of a subspace L:

$$L^{\perp} = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{y} \perp L\}$$

Question:

Are the wall and the floor two orthogonal subspaces in $\mathbb{R}^3$?

Properties of orthogonal vectors/subspaces

$L^\perp$ is a subspace:

$$\mathbf{x}, \mathbf{y} \in L^\perp \implies \mathbf{x} \cdot \mathbf{z} = \mathbf{y} \cdot \mathbf{z} = 0 \text{ for all } \mathbf{z} \in L$$

Therefore, $(a\mathbf{x} + b\mathbf{y}) \cdot \mathbf{z} = 0 \implies a\mathbf{x} + b\mathbf{y} \in L^\perp$

$\mathbf{x} \perp \mathbf{v}_1, \ldots, \mathbf{v}_m \implies \mathbf{x} \perp \mathrm{ls}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$

$$\mathbf{x} \cdot (c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m) = c_1\mathbf{x} \cdot \mathbf{v}_1 + \cdots + c_m\mathbf{x} \cdot \mathbf{v}_m = 0$$

$\dim(L^\perp) = n - \dim L$

- Set $m := \dim L$ and take any basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ of L
- Form an $m \times n$ matrix A with rows $\mathbf{v}_1, \ldots, \mathbf{v}_m$
- Then $\mathcal{N}(A) = L^\perp$: $\quad \begin{array}{lcl} \mathbf{x} \in \mathcal{N}(A) & \implies & \mathbf{x} \perp L, \\ \mathbf{x} \perp L & \implies & \mathbf{x} \in \mathcal{N}(A) \end{array}$
- $\dim(\mathcal{N}(A)) = n - m$

Properties of orthogonal vectors/subspaces

- if subspaces L and M are orthogonal ($L \perp M$), can form their orthogonal sum $L \oplus M$:

$$L \oplus M = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in L, \mathbf{y} \in M\}$$

- $L \oplus M$ is a subspace
- $\dim(L \oplus M) = \dim L + \dim M$:

  indeed, the union of bases of L and M is a basis of $L \oplus M$
- in particular, since $\dim(L^\perp) = n - \dim L$,

$$\boxed{L \oplus L^\perp = \mathbb{R}^n}$$

Outline

Four orthogonal subspaces of a matrix

Every $m \times n$ matrix A generates 4 subspaces:

- column space $\mathcal{C}(A) \subset \mathbb{R}^m$ — linear span of its columns
- nullspace $\mathcal{N}(A) \subset \mathbb{R}^n$ — solutions of $A\mathbf{x} = \mathbf{0}$
- row space $\mathcal{R}(A)$ of A $=$ column space $\mathcal{C}(A^T)$ of $A^\top \subset \mathbb{R}^n$
- left nullspace of A $= \mathcal{N}(A^T) \subset \mathbb{R}^m$

Theorem (Second fundamental theorem of LA)

(i) *The column space $\mathcal{C}(A)$ of A is the orthogonal complement of its left nullspace $\mathcal{N}(A^T)$*

(ii) *The nullspace $\mathcal{N}(A)$ of A is the orthogonal complement of its row space $\mathcal{R}(A) = \mathcal{C}(A^T)$*

Proof: Orthogonality of $\mathcal{C}(A)$ and $\mathcal{N}(A^\top)$.

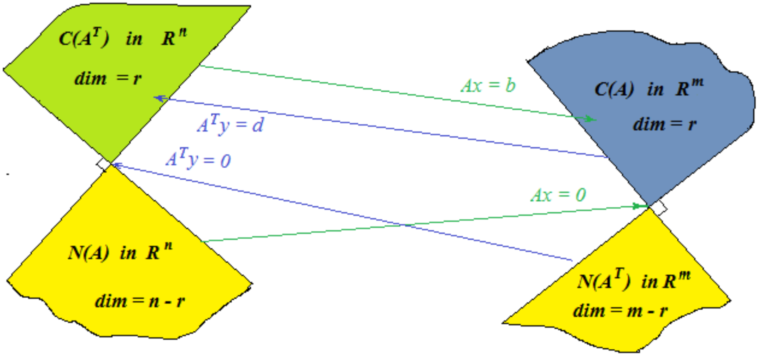
$$\mathcal{C}(A) = \{\mathbf{z} = A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}, \quad \mathcal{N}(A^T) = \{\mathbf{y} \in \mathbb{R}^m \mid \mathbf{y}^T A = \mathbf{0}\} \implies$$

$$\mathbf{y}^T\mathbf{z} = \mathbf{y}^T(A\mathbf{x}) = (\mathbf{y}^T A)\mathbf{x} = \mathbf{0}^T\mathbf{x} = 0$$

$$\dim(\mathcal{C}(A)) = r, \qquad \dim(\mathcal{N}(A^T)) = m - r \implies$$

$$\mathcal{C}(A) \oplus \mathcal{N}(A^T) = \mathbb{R}^m$$

Four orthogonal subspaces of a matrix



Example

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \implies \begin{array}{ll} \mathcal{C}(A) = \mathbb{R}^2, & \mathcal{N}(A) = \mathrm{ls}\{(1,-1,1)\}, \\ \mathcal{C}(A^T) = \mathrm{ls}\{(1,0,-1),(0,1,1)\}, & \mathcal{N}(A^T) = \{\mathbf{0}\} \end{array}$$

Further examples

Matrices with given column and null-spaces

Assume $\mathcal{C}(A) \; = \mathrm{ls}\{\mathbf{u}_1 = (1,1,0)^\top, \mathbf{u}_2 = (0,1,2)^\top)\}$ and
$\mathcal{N}(A) = \mathrm{ls}\{\mathbf{v}_1 = (1,1,0,1)^\top, \mathbf{v}_2 = (0,1,-1,0)^\top)\}$.
How many such matrices A are there?

Idea:

- fix any linearly independent vectors $\mathbf{v}_3$ and $\mathbf{v}_4$ in $\mathcal{R}(A) = \big(\mathcal{N}(A)\big)^\perp$
- find a unique matrix that sends $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$ into $\mathbf{0}, \mathbf{0}, \mathbf{u}_1, \mathbf{u}_2$
- all A are obtained this way

Further examples

Problem

The last two columns of a 3×3 *matrix* A *are* $\mathbf{u}_2 = (1, 1, 0)^\top$ *and* $\mathbf{u}_3 = (0, 1, 2)^\top$ *respectively, while its RREF is*

$$R = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

What is the first column of A*?*

Solution

- $\mathcal{R}(A) = \mathcal{R}(R) = \mathrm{ls}\{(1, 0, 2)^\top, (0, 1, 1)^\top\} = (\mathcal{N}(A))^\perp$
- *therefore,* $\mathcal{N}(A) = \mathcal{N}(R) = \mathrm{ls}\{(2, 1, -1)^\top\}$
- *thus* $A\,(2, 1, -1)^\top = 2\mathbf{u}_1 + \mathbf{u}_2 - \mathbf{u}_3 = \mathbf{0}$, *i.e.,*
  $$\mathbf{u}_1 = (\mathbf{u}_3 - \mathbf{u}_2)/2 = (-1/2, 0, 1)^\top$$

Further examples

Solvability of the normal equation

- The least square solution of an inconsistent system $A\mathbf{x} = \mathbf{b}$ is a solution to the normal equation $A^\top A\mathbf{x} = A^\top \mathbf{b}$
- the latter is always solvable since $A^\top \mathbf{b} \in \mathcal{C}(A^\top A)$
- in fact, the equality $\mathcal{N}(A) = \mathcal{N}(A^\top A)$ and the theorem yield

$$\mathcal{C}(A^\top) = \mathcal{C}(A^\top A)$$

so that $A^\top \mathbf{b} \in \mathcal{C}(A^\top) = \mathcal{C}(A^\top A)$

Outline

Pythagorean theorem and shortest distance

Pythagorean theorem

$\mathbf{u}$ and $\mathbf{v}$ are orthogonal $\iff$ $\boxed{\|\mathbf{u} \pm \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2}$

Reason:

$$\begin{aligned}\|\mathbf{u} \pm \mathbf{v}\|^2 = \langle \mathbf{u} \pm \mathbf{v}, \mathbf{u} \pm \mathbf{v}\rangle &= \langle \mathbf{u}, \mathbf{u}\rangle \pm 2\langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle \\ &= \|\mathbf{u}\|^2 \pm 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2\end{aligned}$$

Shortest distance to a line

ℓ is a line in $\mathbb{R}^n$ in direction of $\mathbf{a}$; P is a point in $\mathbb{R}^n$ outside ℓ

Problem: Find a point Q on ℓ that is closest to P.

Solution: Set $\mathbf{p} := \overrightarrow{OP}$, $\mathbf{q} := \overrightarrow{OQ} = k\mathbf{a}$. The optimal k minimizes

$$|PQ|^2 = \|k\mathbf{a} - \mathbf{p}\|^2 = k^2\|\mathbf{a}\|^2 - 2k\langle \mathbf{p}, \mathbf{a}\rangle + \|\mathbf{p}\|^2 \implies k = \langle \mathbf{p}, \mathbf{a}\rangle / \|\mathbf{a}\|^2$$

Observe that $\overrightarrow{PQ} = \mathbf{q} - \mathbf{p}$ and $\mathbf{a}$ are then orthogonal:

$$\langle k\mathbf{a} - \mathbf{p}, \mathbf{a}\rangle = k\langle \mathbf{a}, \mathbf{a}\rangle - \langle \mathbf{p}, \mathbf{a}\rangle = 0$$

Shortest distance and orthogonality

Orthogonal PQ is the shortest one!

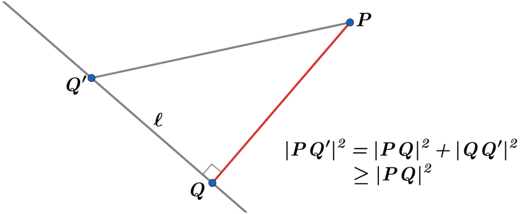
ℓ is a line in $\mathbb{R}^n$ in direction of $\mathbf{a}$; P is a point in $\mathbb{R}^n$ outside ℓ

Claim: If Q on ℓ is s.t. $\mathbf{u} := \overrightarrow{PQ} \perp \mathbf{a}$, then $|PQ|$ is the smallest one

Reason: for any other point P' on ℓ, we have

$$|PQ'|^2 = \|\overrightarrow{PQ'}\|^2 = \|\overrightarrow{PQ} + \overrightarrow{QQ'}\|^2 = \|\overrightarrow{PQ}\|^2 + \|\overrightarrow{QQ'}\|^2 \geq \|\overrightarrow{PQ}\|^2$$

and the inequality is strict unless $Q = Q'$



Conclusion:

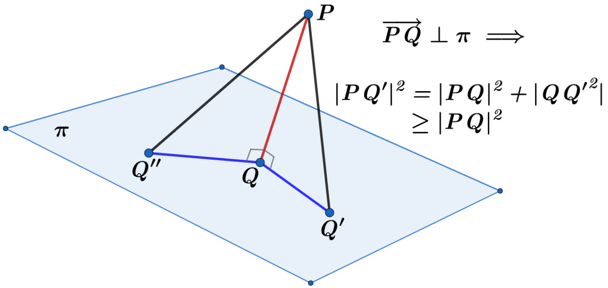
$$\boxed{Q \in \ell \text{ minimizes } |PQ| \iff \overrightarrow{PQ} \perp \ell}$$

Shortest distance to a plane

Remark

The same arguments work if instead of a line ℓ we take a plane π:

$$\boxed{Q \in \pi \text{ minimizes } |PQ| \iff \overrightarrow{PQ} \perp \pi}$$



Remark

In fact, instead of line ℓ or plane π we can take any subspace W in $\mathbb{R}^n$

Outline

Shortest distance in $\mathcal{P}_\infty$

Consider $\mathcal{P}_\infty$ with the inner product

$$\langle p, q \rangle = \int_{-1}^{1} p(t)q(t)\,dt.$$

The distance from the polynomial $p(t) = t^3$ to the subspace $\mathcal{P}_2$ can be obtained this way:

- let $p_0 = a + bt + ct^2 \in \mathcal{P}_2$ be the closest polynomial;
- then $q := p - p_0$ is orthogonal to $\mathcal{P}_2 = \operatorname{ls}\{1, t, t^2\}$
- therefore, we get three equations:

$$\begin{aligned} \langle q, 1 \rangle &= -a - c/3 = 0, \\ \langle q, t \rangle &= 1/5 - b/3 = 0, \\ \langle q, t^2 \rangle &= -a/3 - c/5 = 0 \end{aligned}$$

and find that $a = c = 0$ and $b = 3/5$, so that $p_0(t) = 3t/5$

Example: modelling COVID-19 exponential spread

COVID-19 is believed to spread exponentially

- Assume we have the numbers

$$x_1 = x_2 = 1, x_3 = x_4 = 2, x_5 = 4, x_{10} = 11$$

  of confirmed infections in some region on the day $n = 1, 2, 3, 4, 5, 10$
- It is believed that the number of infected people grows exponentially, i.e., that $x_n = a \cdot e^{bn}$
- Based on these data, what is the best estimate for a and b?

Solution:

Will be discussed next Wed, so stay healthy and tuned!

Summary:

- In inner product vector spaces
  - the norm $\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}$ of **u** and
  - the distance $\|\mathbf{u} - \mathbf{v}\|$ between **u** and **v**

  can be introduced;
- the inner product $\langle \cdot, \cdot \rangle$ and the related norm $\| \cdot \|$ satisfy
  - the Cauchy–Bunyakovski–Schwarz inequality
  - the triangle inequality
- every $m \times n$ matrix A generates two pairs of orthogonal subspaces (column space=range and nullspace of A and its transposed A^T)
  - $\mathbb{R}^m = \mathcal{C}(A) \oplus \mathcal{N}(A^T)$
  - $\mathbb{R}^n = \mathcal{C}(A^T) \oplus \mathcal{N}(A)$
- orthogonal vectors **u** and **v** satisfy the Pythagorean theorem:

$$\boxed{\|\mathbf{u} \pm \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2}$$

- a point Q in a subspace W is the closest to $P \iff \overrightarrow{PQ} \perp W$