Matrix V

REA1121

Mathematics for programming

Outline

- Roadmap
- Projections
- Non-solvable linear equations
- Exercises

ROADMAP

Roadmap

Vectors

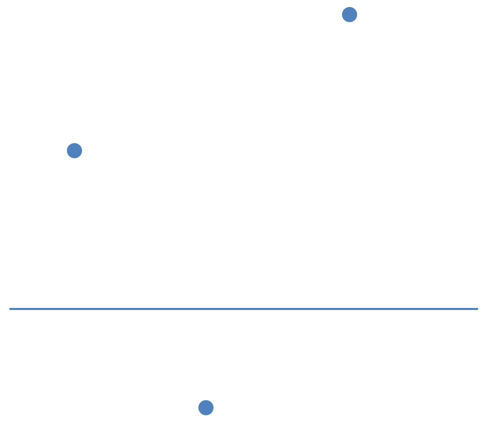
- Linear combination $c\boldsymbol{v} + d\boldsymbol{w}$
- Dot product $\boldsymbol{v} \cdot \boldsymbol{w}$
- Length
- Unit vector
- Angle

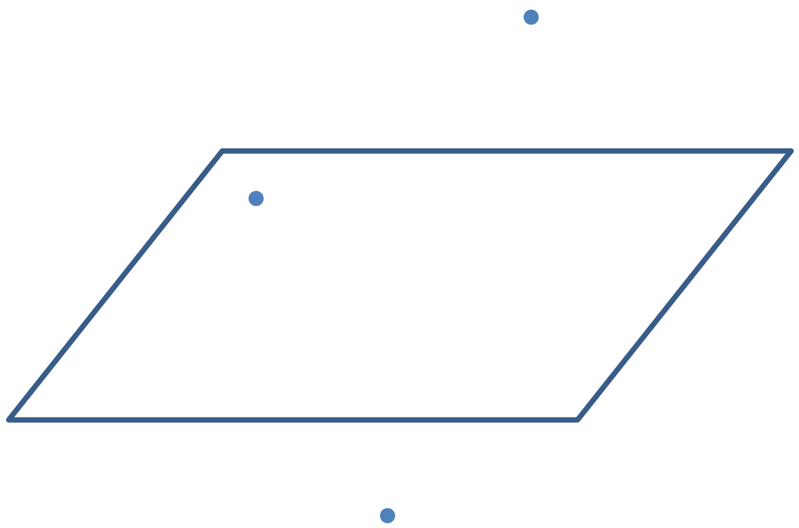
Vector space

- Subspace
- Linear independence
- Basis
- Orthonormal basis
- Dimension
- Column space
- Nullspace
- Rank

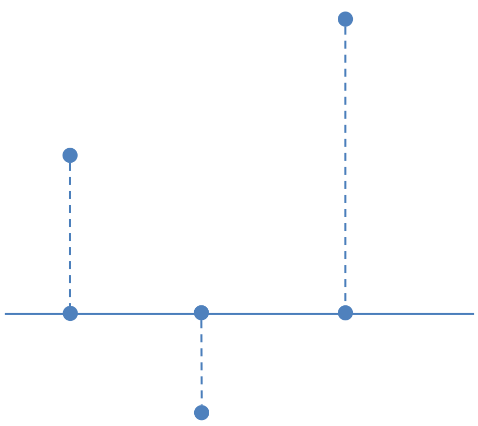
PROJECTIONS

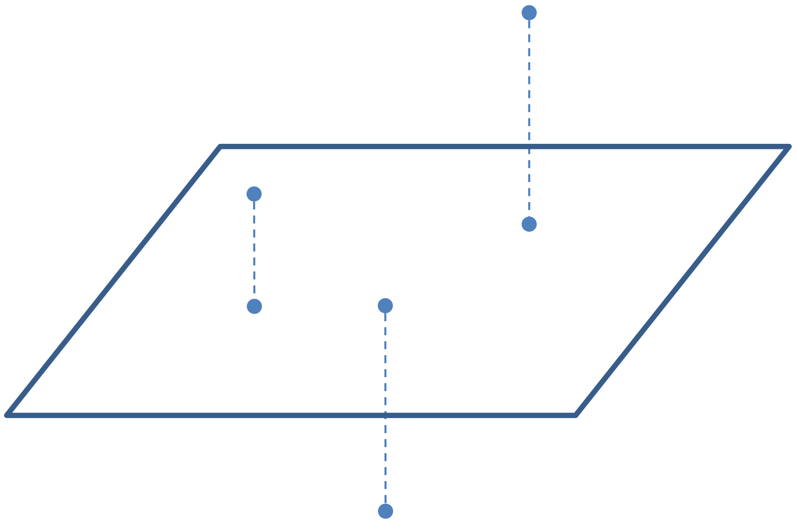
Introduction



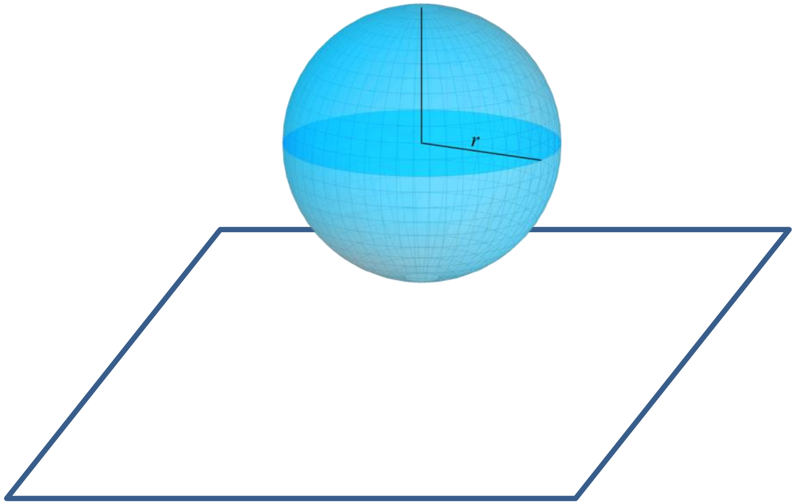


Introduction

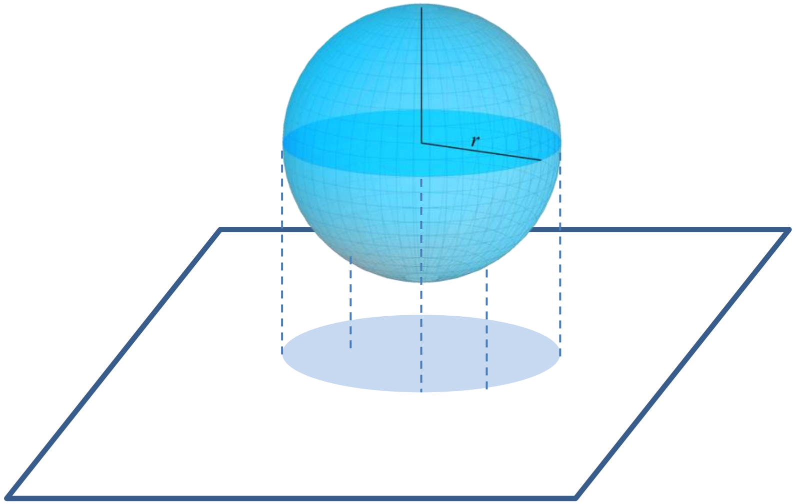


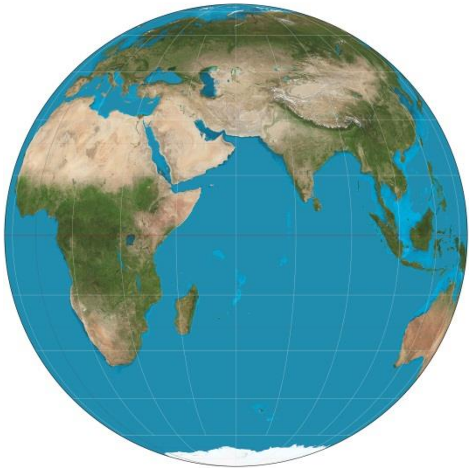


Introduction



Introduction



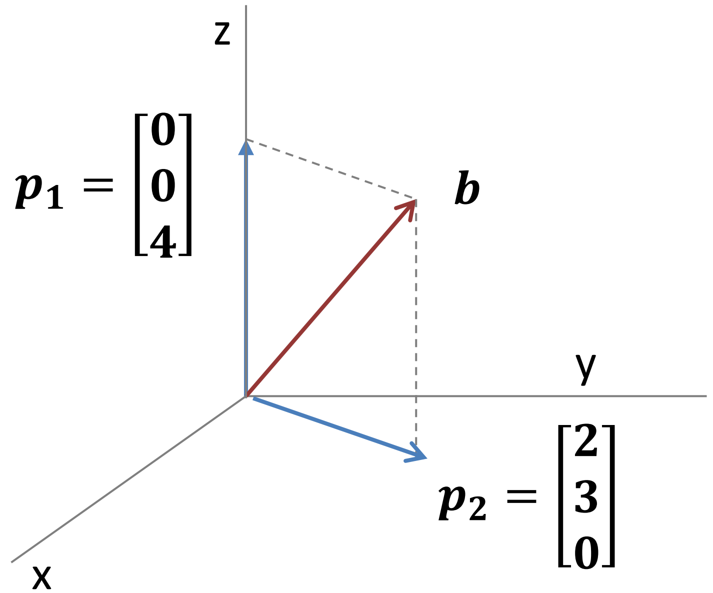


Introduction

- What are the projections of $\boldsymbol{b} = \begin{bmatrix} 2 & 3 & 4 \end{bmatrix}^T$ onto the z axis and the xy plane?

Introduction

- What are the projections of $\boldsymbol{b} = [2 \quad 3 \quad 4]^T$ onto the z axis and the xy plane?



Introduction

- What are the projections of $\boldsymbol{b} = [2 \quad 3 \quad 4]^T$ onto the z axis and the xy plane?
- What matrices produce those projections onto a line and a plane?

$$\boldsymbol{p} = P\boldsymbol{b}$$

$$P = ?$$

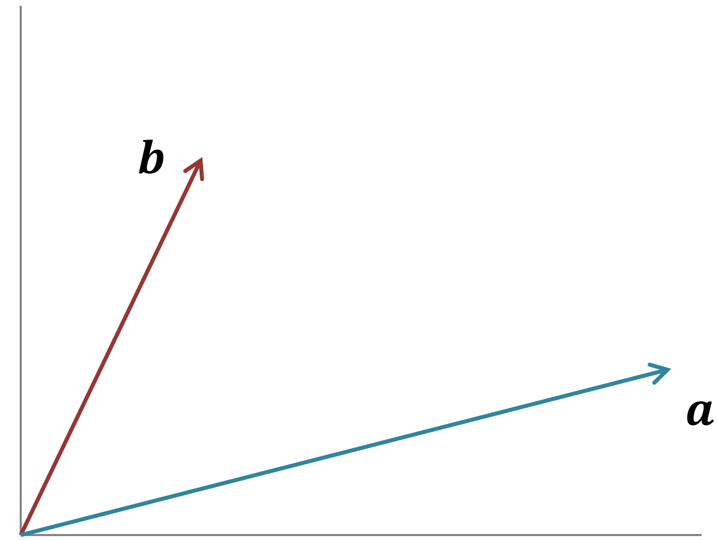
Introduction

- What are the projections of $\boldsymbol{b} = [2 \quad 3 \quad 4]^T$ onto the z axis and the xy plane?
- What matrices produce those projections onto a line and a plane?

$$P_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

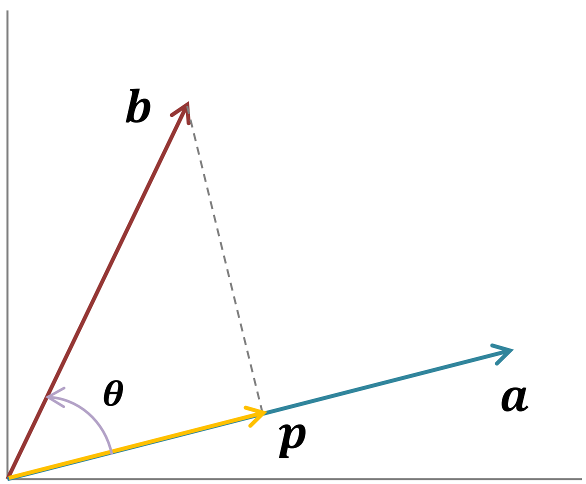
Projection onto a line

- A line goes through the origin in the direction of $\boldsymbol{a} = [a_1, a_2, \cdots, a_m]^T$.
- Along that line, we want the point $\boldsymbol{p}$ closest to $\boldsymbol{b} = [b_1, b_2, \cdots, b_m]^T$.



Projection onto a line

- The key to projection is orthogonality:
  - The line from $\boldsymbol{b}$ to $\boldsymbol{p}$ is perpendicular to the vector $\boldsymbol{a}$.
- The projection $\boldsymbol{p}$ is some multiple of $\boldsymbol{a}$. Call it $\boldsymbol{p} = \hat{x}\boldsymbol{a}$.



Projection onto a line

- Projecting $\boldsymbol{b}$ onto $\boldsymbol{a}$.
- Error $\boldsymbol{e} = \boldsymbol{b} - \hat{x}\boldsymbol{a}$.
- Orthogonality: $\boldsymbol{a} \cdot (\boldsymbol{b} - \hat{x}\boldsymbol{a}) = \boldsymbol{0}$ or $\boldsymbol{a} \cdot \boldsymbol{b} - \hat{x}\boldsymbol{a} \cdot \boldsymbol{a} = \boldsymbol{0}$
- $\hat{x} = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\boldsymbol{a} \cdot \boldsymbol{a}} = \frac{\boldsymbol{a}^T \boldsymbol{b}}{\boldsymbol{a}^T \boldsymbol{a}}$

Projection onto a line

- The projection of $\boldsymbol{b}$ onto the line through $\boldsymbol{a}$ is the vector $\boldsymbol{p} = \hat{x}\boldsymbol{a} = \dfrac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}}\boldsymbol{a}$.
- $\boldsymbol{p} = \boldsymbol{a}\hat{x} = \boldsymbol{a}\dfrac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}} = \dfrac{\boldsymbol{a}\boldsymbol{a}^T}{\boldsymbol{a}^T\boldsymbol{a}}\boldsymbol{b}, \boldsymbol{p} = P\boldsymbol{b}$

  Projection matrix $P = \dfrac{\boldsymbol{a}\boldsymbol{a}^T}{\boldsymbol{a}^T\boldsymbol{a}}$
- Two special cases:
  - If $\boldsymbol{b} = \boldsymbol{a}$, then $\hat{x} = 1, \boldsymbol{p} = \boldsymbol{a}$
  - If $\boldsymbol{b}$ is perpendicular to $\boldsymbol{a}$, then $\boldsymbol{a}^T\boldsymbol{b} = 0, \boldsymbol{p} = 0$

Example 1

Project $\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ onto $\boldsymbol{a} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ to find $\boldsymbol{p} = \hat{x}\boldsymbol{a}$

Example 1

$$\hat{x} = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\boldsymbol{a} \cdot \boldsymbol{a}} = \frac{\boldsymbol{a}^T \boldsymbol{b}}{\boldsymbol{a}^T \boldsymbol{a}} = \frac{5}{9}$$

$$\boldsymbol{p} = \hat{x}\boldsymbol{a} = \frac{5}{9}\boldsymbol{a}$$

$$\boldsymbol{p} \cdot \boldsymbol{e} = ?$$

$$\|\boldsymbol{p}\| = ?$$

Example 1

$$\hat{x} = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\boldsymbol{a} \cdot \boldsymbol{a}} = \frac{\boldsymbol{a}^T \boldsymbol{b}}{\boldsymbol{a}^T \boldsymbol{a}} = \frac{5}{9}$$

$$\boldsymbol{p} = \hat{x}\boldsymbol{a} = \frac{5}{9}\boldsymbol{a}$$

$$\boldsymbol{p} \cdot \boldsymbol{e} = \boldsymbol{0}$$

$$\|\boldsymbol{p}\| = \frac{\|\boldsymbol{a}\|\|\boldsymbol{b}\| \cos \boldsymbol{\theta}}{\|\boldsymbol{a}\|^2}\|\boldsymbol{a}\| = \|\boldsymbol{b}\|\cos \boldsymbol{\theta}$$

Example 2

Find the projection matrix $P = \dfrac{\boldsymbol{a}\boldsymbol{a}^T}{\boldsymbol{a}^T\boldsymbol{a}}$ onto the line through $\boldsymbol{a} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$.

Example 2

Projection matrix P

$$P = \frac{\boldsymbol{a}\boldsymbol{a}^T}{\boldsymbol{a}^T\boldsymbol{a}} = \frac{1}{9}\begin{bmatrix}1\\2\\2\end{bmatrix}\begin{bmatrix}1 & 2 & 2\end{bmatrix} = \frac{1}{9}\begin{bmatrix}1 & 2 & 2\\2 & 4 & 4\\2 & 4 & 4\end{bmatrix}$$

Example 2

- P: $m \times m$ matrix
- Its rank is 1
- P projects any $\boldsymbol{b}$ onto $\boldsymbol{a}$,1-dimensional subspace which is the column space of P
- $||a||$ does not influence P
- $P^2 = P$

Example 2

- P projects onto a subspace ($\boldsymbol{a}$)
- $I - P$ projects onto the perpendicular subspace (the plane perpendicular to $\boldsymbol{a}$).

Projection onto a subspace

- Find the combination $\boldsymbol{p} = \widehat{x_1}\boldsymbol{a}_1 + \cdots + \widehat{x_n}\boldsymbol{a}_n$ closest to a given vector $\boldsymbol{b}$.
- We are projecting each $\boldsymbol{b}$ in $\mathbf{R}^m$ onto the subspace spanned by the $\boldsymbol{a}$'s, to get $\boldsymbol{p}$.

Projection onto a subspace

$$\begin{bmatrix} \boldsymbol{a}_1{}^T \\ \vdots \\ \boldsymbol{a}_n{}^T \end{bmatrix} [\boldsymbol{b} - A\widehat{\boldsymbol{x}}] = \boldsymbol{0}$$

$$A^T(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = 0$$

$$A^T A\widehat{\boldsymbol{x}} = A^T\boldsymbol{b}$$

Projection onto a subspace

If the $\boldsymbol{a}$'s are independent, $A^T A$ is invertible.

$$A^T A\hat{\boldsymbol{x}} = A^T \boldsymbol{b}$$
$$\hat{\boldsymbol{x}} = (A^T A)^{-\mathbf{1}} A^T \boldsymbol{b}$$

$$\boldsymbol{p} = A\hat{\boldsymbol{x}} = A(A^T A)^{-\mathbf{1}} A^T \boldsymbol{b}$$
$$P = A(A^T A)^{-1} A^T$$

Example 3

If $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$ onto $\boldsymbol{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$, find $\hat{x}$ and $\boldsymbol{p}$.

Example 3

$$A^T A\hat{\boldsymbol{x}} = A^T \boldsymbol{b}$$

$$A^T A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}$$

$$A^T \boldsymbol{b} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}, \hat{\boldsymbol{x}} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

Example 3

$$\boldsymbol{p} = A\widehat{\boldsymbol{x}} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$$

$$\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

$$\boldsymbol{p} \cdot \boldsymbol{e} = ?$$

Example 3

$$P = A(A^T A)^{-\mathbf{1}} A^T$$

$$(A^T A)^{-\mathbf{1}} = \frac{1}{6}\begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix}$$

$$P = \frac{1}{6}\begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}$$

Example 3

- $A^T A$ is invertible if and only if A has linearly independent columns.
- When A has independent columns, $A^T A$ is square, symmetric, and invertible.

$$\begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 9 \end{bmatrix}$$

NON-SOLVABLE LINEAR EQUATIONS

Non-solvable equations

- It often happens that $A\boldsymbol{x} = \boldsymbol{b}$ has no solution.
- The usual reason is: too many equations.
  - *m* is greater than *n* (more equations than unknowns)
  - The *n* columns span a small part of *m*-dimensional space
  - $\boldsymbol{b}$ is outside the column space of A
  - Elimination reaches an impossible equation and stops

Example 4

Find the closest line to the points $(0,6)$, $(1,0)$, and $(2,0)$.

Example 4

- Suppose a straight line $b = C + Dt$ goes through these three points

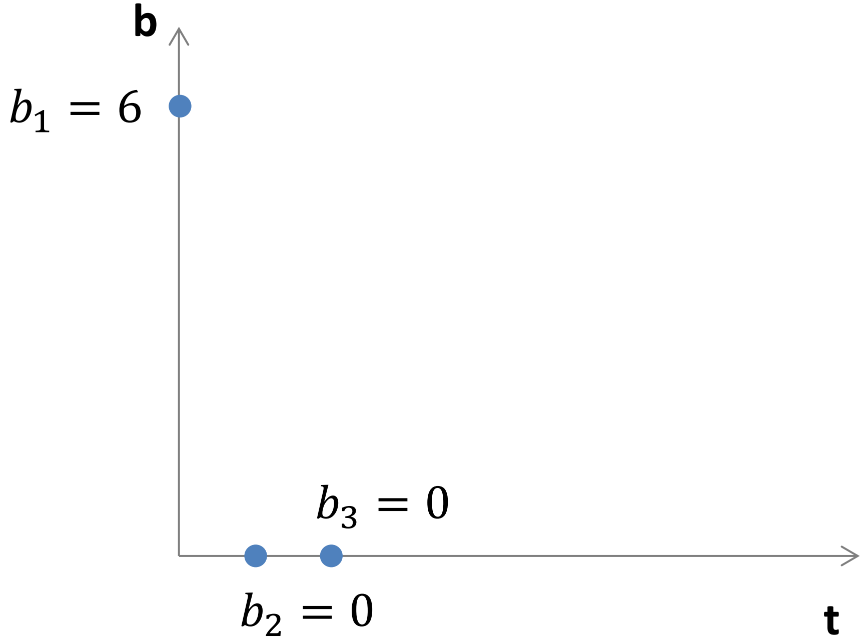
$$\begin{cases} C + D \cdot 0 = 6 \\ C + D \cdot 1 = 0 \\ C + D \cdot 2 = 0 \end{cases}$$

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, x = \begin{bmatrix} C \\ D \end{bmatrix}, b = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

Example 4

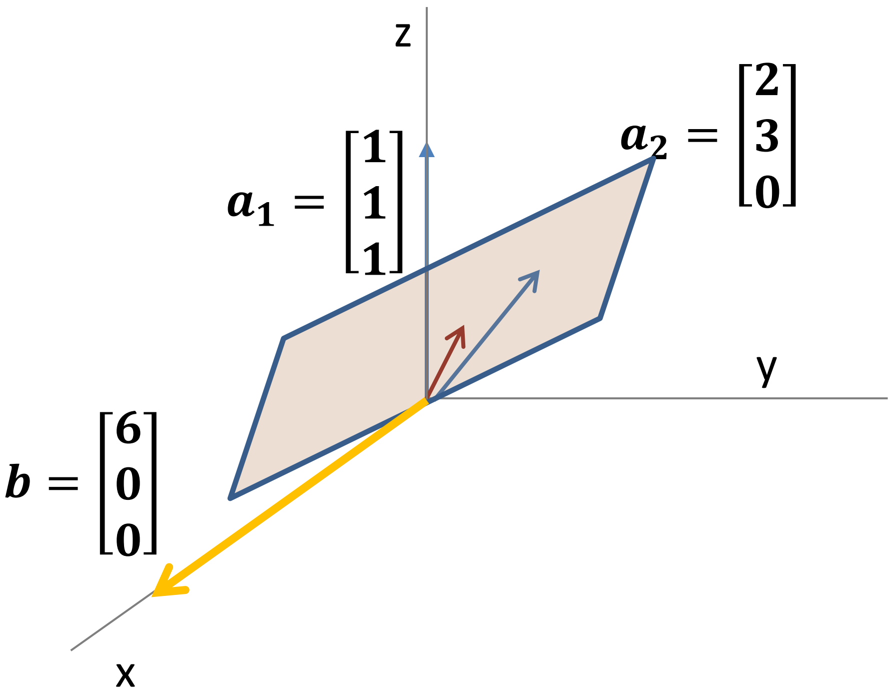
$Ax = b$ is not solvable

no straight line goes through the three points



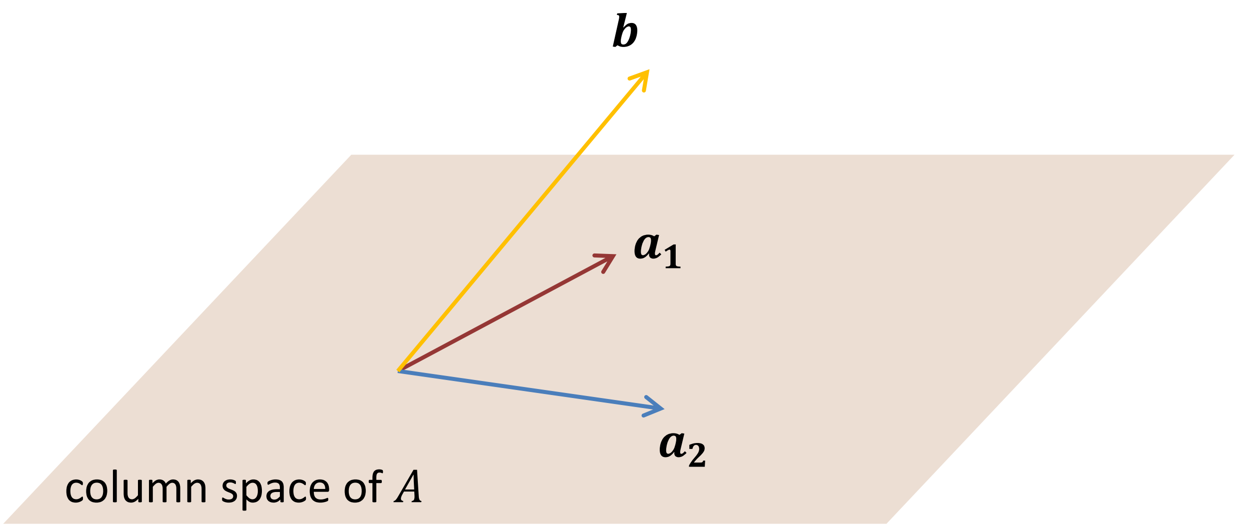
Example 4

The vector $\boldsymbol{b}$ does not lie in the column space of A.



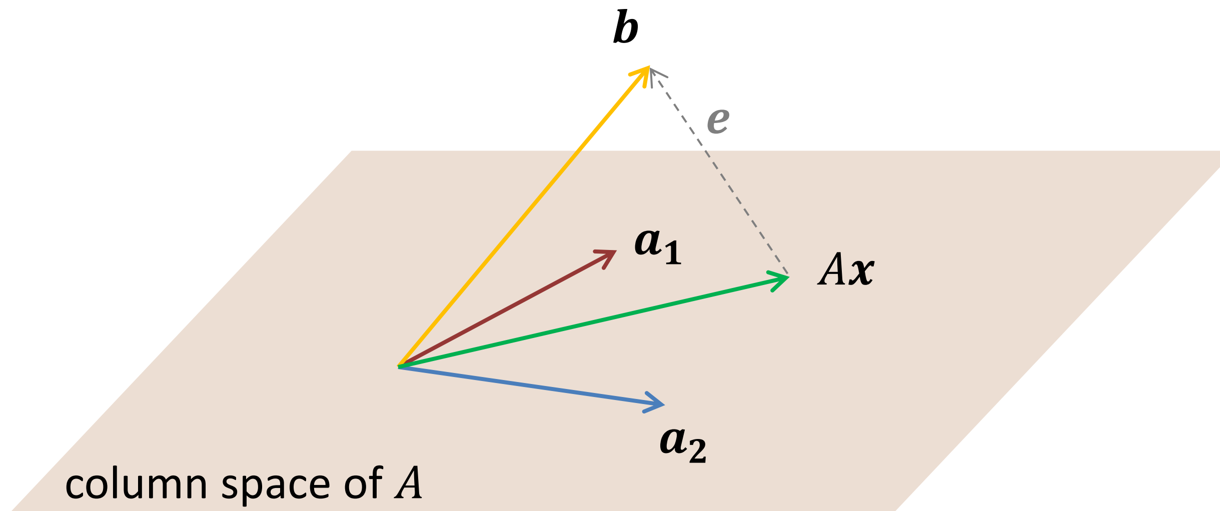
Example 4

The vector $\boldsymbol{b}$ does not lie in the column space of A.



Example 4

How do we find $\hat{\boldsymbol{x}}$ that minimizes $\boldsymbol{e} = \boldsymbol{b} - A\boldsymbol{x}$?

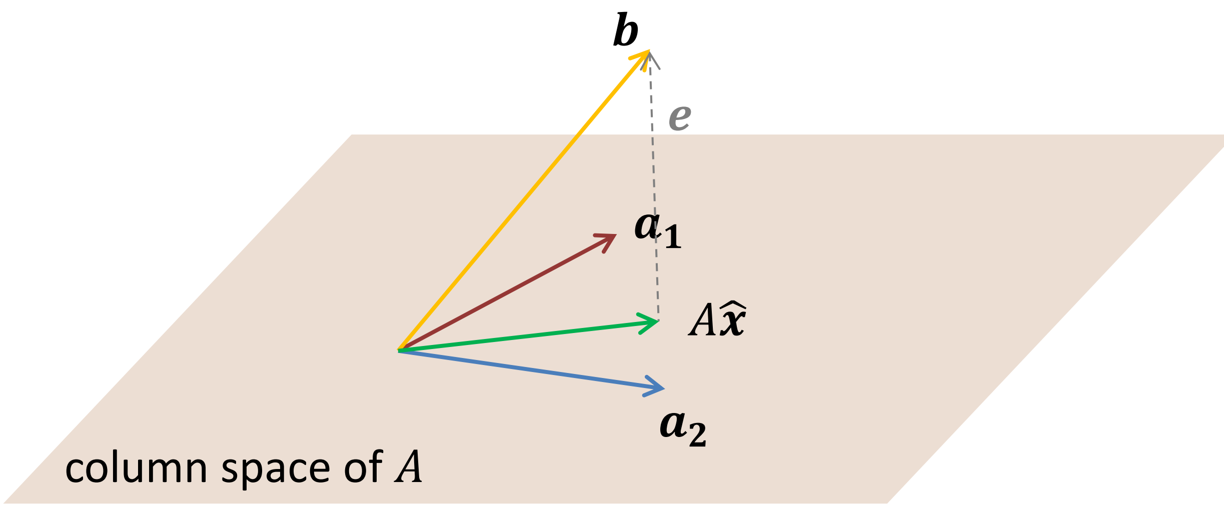


Example 4

By geometry:

We look for the point closest to $\boldsymbol{b}$

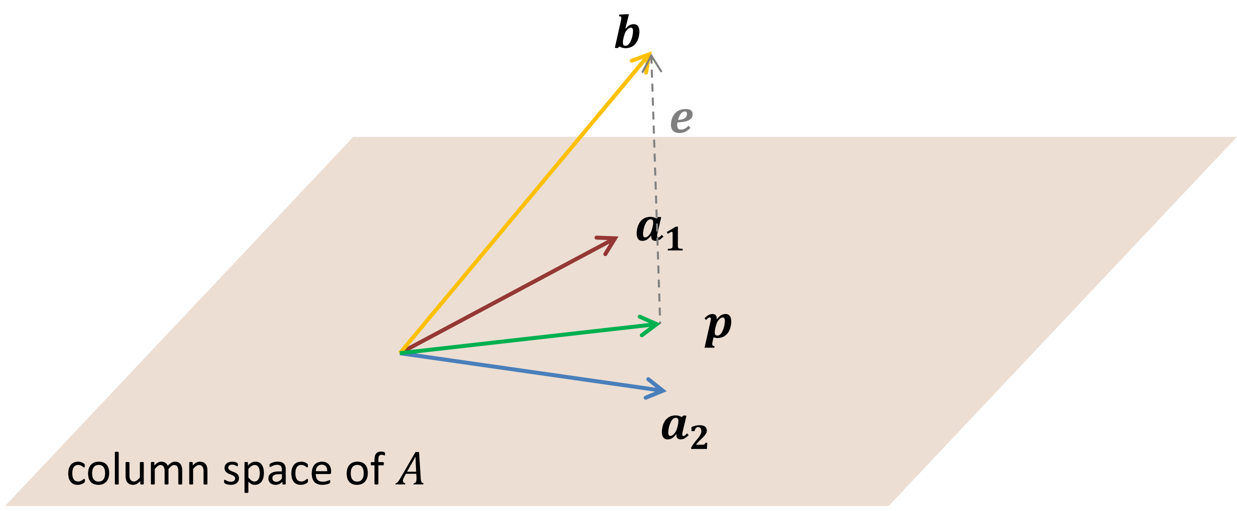
The nearest point is the projection $\boldsymbol{p}$



Example 4

By algebra

Every vector $\boldsymbol{b}$ splits into two parts. The part in the column space is $\boldsymbol{p}$. The perpendicular part in the nullspace of A^T is $\boldsymbol{e}$. $\boldsymbol{b} = \boldsymbol{p} + \boldsymbol{e}$

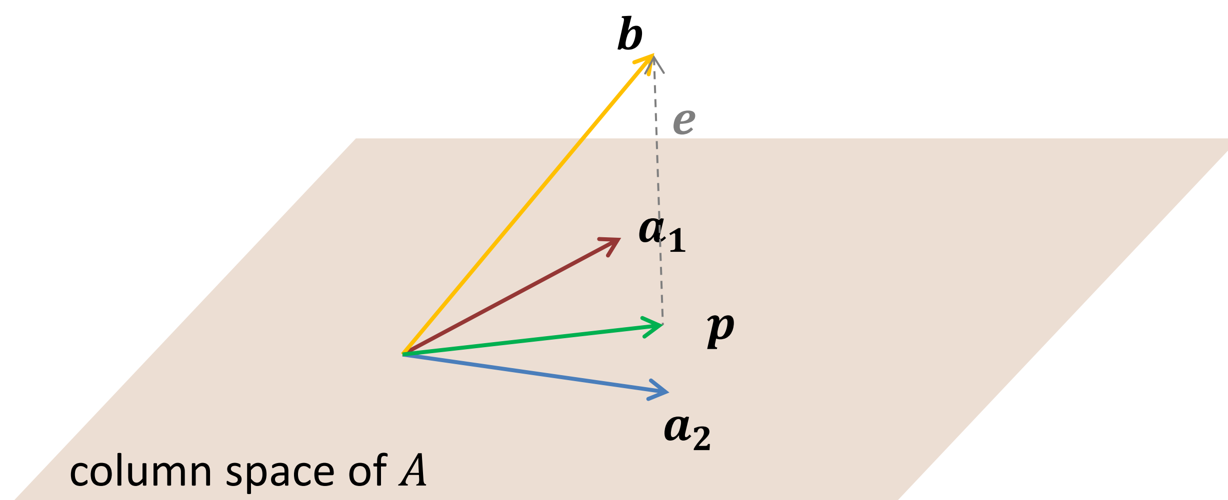


Example 4

By algebra

$A\boldsymbol{x} = \boldsymbol{b}$ is not solvable.

$A\boldsymbol{x} = \boldsymbol{p}$ is solvable.



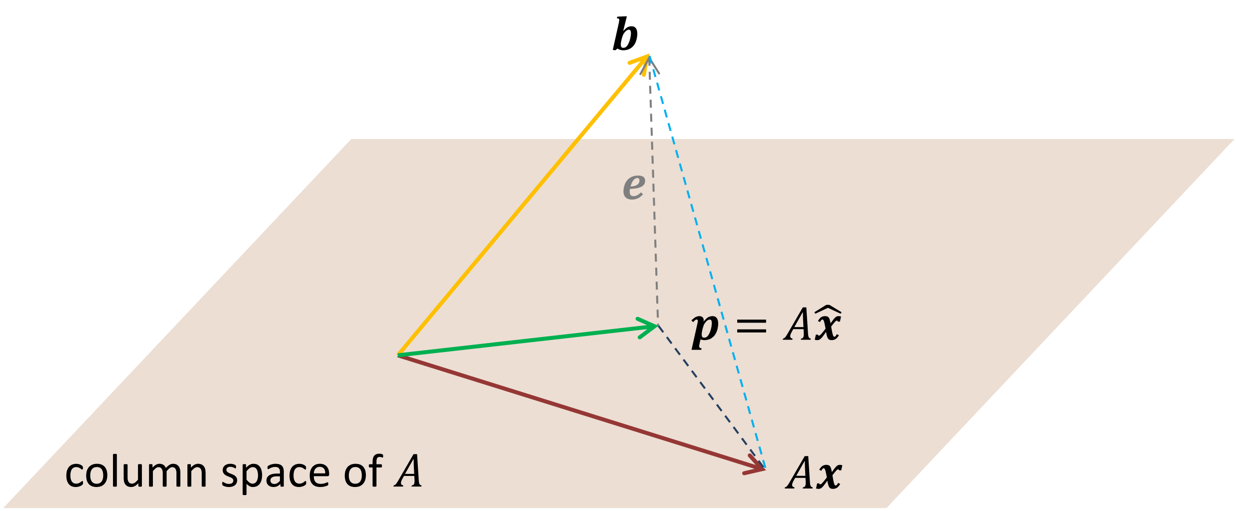
Example 4

By algebra

For any $\boldsymbol{x}$, the squared length

$$||A\boldsymbol{x} - \boldsymbol{b}||^2 = ||A\boldsymbol{x} - \boldsymbol{p}||^2 + ||\boldsymbol{e}||^2$$

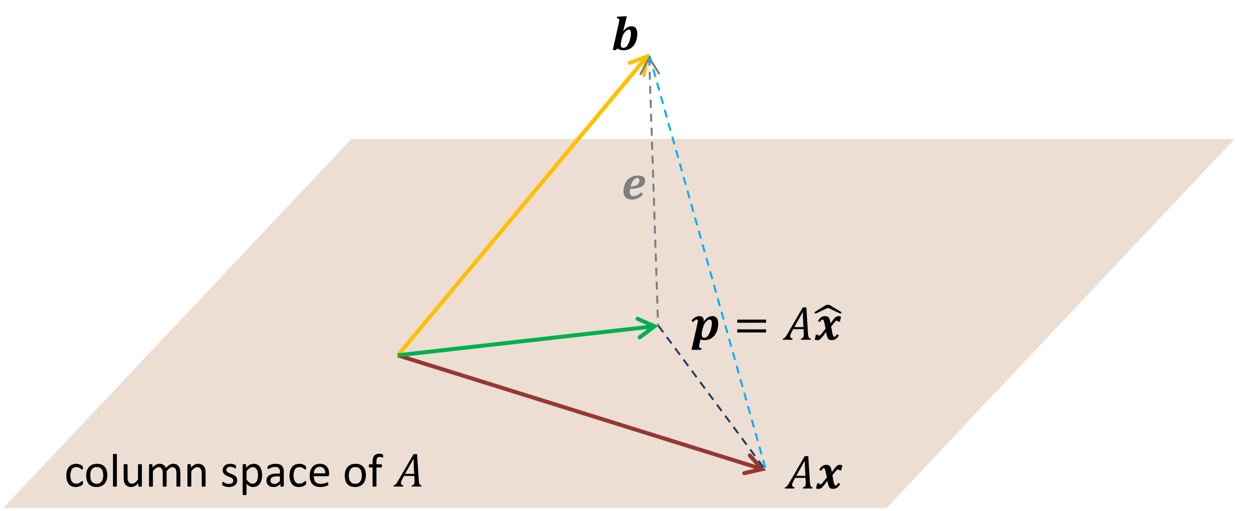
$\boldsymbol{x} = \widehat{\boldsymbol{x}}$ leads to the smallest possible error $\boldsymbol{e}$



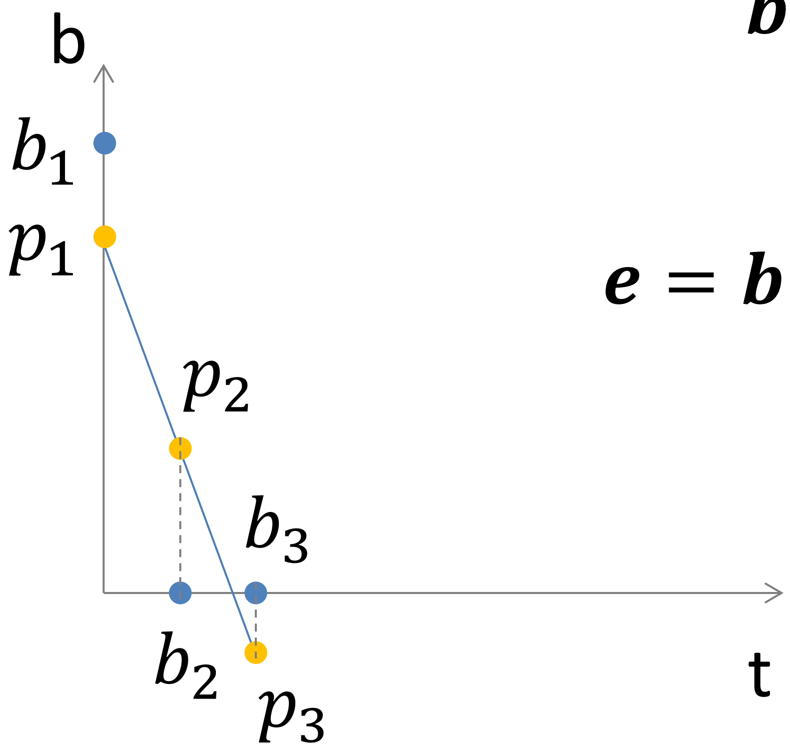
Example 4

The ***least squares*** solution $\widehat{\boldsymbol{x}}$

makes $E = ||A\boldsymbol{x} - \boldsymbol{b}||^2$ as small as possible.



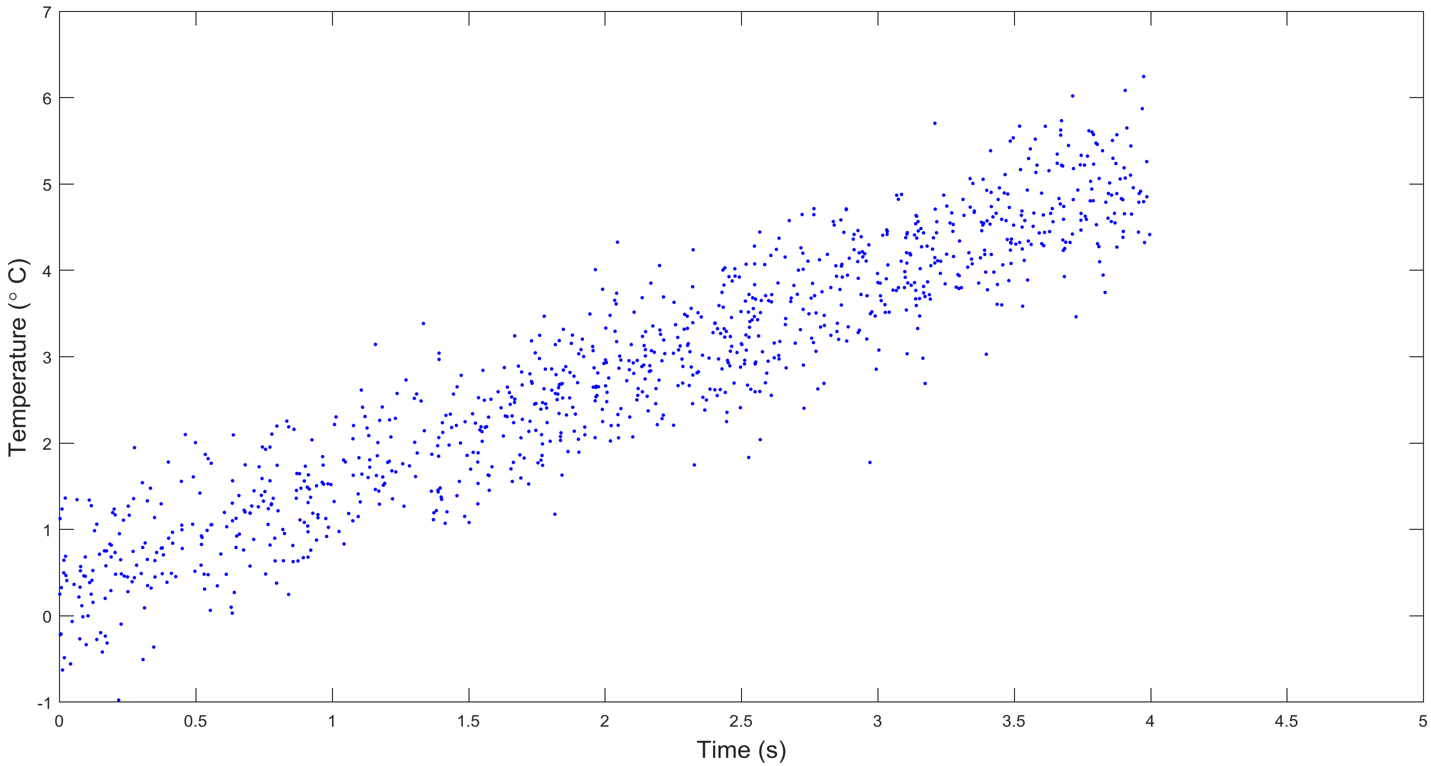
Example 4



$$\boldsymbol{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}, \qquad \boldsymbol{p} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$$

$$\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Fitting a straight line



Fitting a straight line

$$\begin{cases} C + Dt_1 = b_1 \\ C + Dt_2 = b_2 \\ \vdots \\ C + Dt_m = b_m \end{cases}$$

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} C \\ D \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Fitting a straight line

$Ax = b$ is not solvable, however we can solve $A\hat{\boldsymbol{x}} = \boldsymbol{b}$.

$$A^T A\hat{\boldsymbol{x}} = A^T \boldsymbol{b}$$

Fitting a straight line

$$A^T A = \begin{bmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_m \end{bmatrix} \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} = \begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum {t_i}^2 \end{bmatrix}$$

$$A^T \boldsymbol{b} = \begin{bmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_m \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix}$$

Fitting a straight line

The line C+Dt minimizes

$$e_1{}^2 + \cdots + e_m{}^2 = \|A\boldsymbol{x} - \boldsymbol{b}\|^2$$

when $A^T A\widehat{\boldsymbol{x}} = A^T\boldsymbol{b}$

$$\begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum t_i{}^2 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix}$$

Pseudoinverse

- For a linear system $A\boldsymbol{x} = \boldsymbol{b}$
- $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ provides the unique solution if A is invertible
- We can find the pseudoinverse A^{+}, which is a generalization of the inverse matrix.
- The most widely known type of matrix pseudoinverse is the ***Moore–Penrose pseudoinverse***.
- A^{+} exists for any matrix.

Pseudoinverse

- In particular, when A has linearly independent columns, namely, $A^T A$ is invertible,
- $A^+ = (A^T A)^{-\mathbf{1}} A^T$
- $A^+ A = I$
- A^+ provides least squares solutions $\boldsymbol{x} = A^+ \boldsymbol{b} = (A^T A)^{-\mathbf{1}} A^T \boldsymbol{b}$ to the system.
- In this particular case, $\boldsymbol{b}$ is projected onto A.

EXERCISES

Problem 1

Project the vector $\boldsymbol{b}$ onto the line through $\boldsymbol{a}$. Check that $\boldsymbol{e}$ is perpendicular to $\boldsymbol{a}$:

1) $\boldsymbol{b} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \boldsymbol{a} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

2) $\boldsymbol{b} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \boldsymbol{a} = \begin{bmatrix} -1 \\ -3 \\ -1 \end{bmatrix}$

Problem 2

- In Problem 1, find the projection matrix P onto the line through each vector $\boldsymbol{a}$.
- Verify in both cases that $P^2 = P$. Compute the projection $\boldsymbol{p}$ by $\boldsymbol{p} = P\boldsymbol{b}$.

Problem 3

1) If $P^2 = P$ show that $(I - P)^2 = I - P$.
2) When P projects onto the column space of A, $I - P$ projects onto the ______?
3) If P is the 2 by 2 projection matrix onto the line through $[1 \quad 1]^T$, then $I - P$ is the projection matrix onto _______?
4) If P is the 3 by 3 projection matrix onto the line through $[1 \quad 1 \quad 1]^T$, then $I - P$ is the projection matrix onto _______?

Problem 4

- With b = 0,8,8,20 at t = 0,1,3,4, set up and solve the normal equations $A^T A\hat{\boldsymbol{x}} = A^T \boldsymbol{b}$.
- For the best straight line, find its four heights p_i and four errors e_i.
- What is the minimum value $E = {e_1}^2 + {e_2}^2 + {e_3}^2 + {e_4}^2$?

Problem 5

- Write down three equations for the line $b = C + Dt$ to go through $b = 7$ at $t = -1$, $b = 7$ at $t = 1$, and $b = 21$ at $t = 2$.
- Find the least squares solution $\hat{x} = (C, D)$ and draw the closest line.

Solution 1

1) $\hat{x} = \frac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}} = \frac{5}{3},\ p = \frac{5}{3}a,\ e = \frac{1}{3}[-2 \quad 1 \quad 1]^T$

2) $\hat{x} = \frac{\boldsymbol{a}^T\boldsymbol{b}}{\boldsymbol{a}^T\boldsymbol{a}} = -1,\ p = -a,\ e = 0$

Solution 2

$$1)\ P = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, Pb = \frac{1}{3}\begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix}$$

$$2)\ P = \frac{1}{11}\begin{bmatrix} 1 & 3 & 1 \\ 3 & 9 & 3 \\ 1 & 3 & 1 \end{bmatrix}, Pb = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$$

Solution 3

1) If $P^2 = P$ then $(I - P)^2 = (I - P)(I - P) = I - PI - IP + P^2$.
2) When P projects onto the column space, $(I - P)$ projects onto the left nullspace, or the nullspace of A^T.
3) $I - P$ is the projection matrix onto $[1 \quad -1]^T$ in the perpendicular direction to $[1 \quad 1]^T$
4) $I - P$ projects onto the plane $x + y + z = 0$ perpendicular to $[1 \quad 1 \quad 1]^T$.

Solution 4

- $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, b = \begin{bmatrix} 0 \\ 8 \\ 8 \\ 20 \end{bmatrix}, p = \begin{bmatrix} 1 \\ 5 \\ 13 \\ 17 \end{bmatrix},\ e = \begin{bmatrix} -1 \\ 3 \\ -5 \\ 3 \end{bmatrix}$
- $E = \|e\|^2 = 44$

Solution 5

- $\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 7 \\ 7 \\ 21 \end{bmatrix}$
- $\hat{x} = \begin{bmatrix} 9 \\ 4 \end{bmatrix}$