Imperial College London

Maths for Signals and Systems

Linear Algebra in Engineering

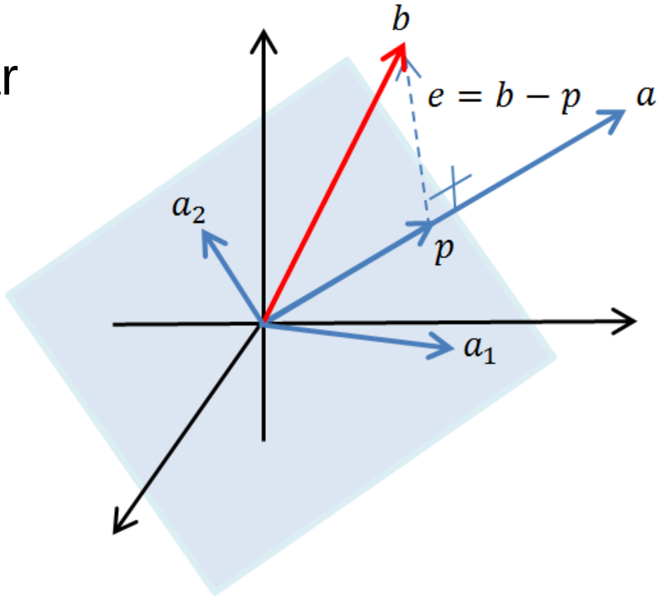
Lecture 9, Friday 1st November 2014

DR TANIA STATHAKI

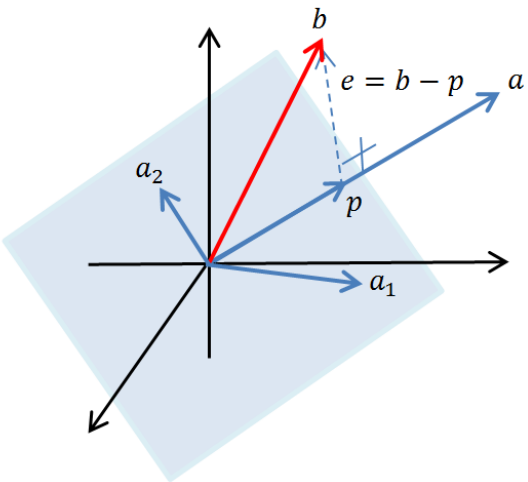
READER (ASSOCIATE PROFFESOR) IN SIGNAL PROCESSING
IMPERIAL COLLEGE LONDON

Problem formulation

- Consider again the scenario where the system $Ax = b$ doesn't have solution.
- Goal: Solve $A\hat{x} = p$ instead, where p is the projection of b onto the column space of A.
- The error $e = b - p$ is again perpendicular to the column space of A.
- This scenario is depicted in the figure on the right for R^3.
- The quantities shown are column vectors and $A = [a_1 \quad a_2]$.



Projection error, solution, projection matrix



- The error e is perpendicular to a_1 and a_2.

  $a_1{}^T(b - A\hat{x}) = 0$ and $a_2{}^T(b - A\hat{x}) = 0$

  $$\begin{bmatrix} a_1{}^T \\ a_2{}^T \end{bmatrix}(b - A\hat{x}) = 0 \Rightarrow A^T(b - A\hat{x}) = 0$$
- Consider $A^T(b - A\hat{x}) = 0$.
- **Question:** In which subspace does $(b - A\hat{x})$ belong?

  Answer: $e \in N(A^T)$. Therefore e is perpendicular to the column space of A. This is also obvious from $A^T(b - A\hat{x}) = 0$.
- Solution: $A^T A\hat{x} = A^T\, b \Longrightarrow \hat{x} = (A^T A)^{-1} A^T b$
- Projection: $p = Pb = A(A^T A)^{-1}\, A^T b$
- Projection matrix: $P = A(A^T A)^{-1} A^T$

Projection matrix

- Projection matrix: $P = A(A^T A)^{-1} A^T$

- A is not square (it is rectangular) and therefore we cannot use the property $(A^T A)^{-1} = A^{-1}(A^T)^{-1}$.

- **Question:** If A **<u>was</u>** a square and invertible matrix of size $n \times n$ what would P be?

  Answer: In that case the column space of A would be the entire R^n and therefore the projection of any vector on $C(A)$ would be the vector itself.

  This can be also verified by
  $P = A(A^T A)^{-1} A^T = AA^{-1}(A^T)^{-1} A^T = I \cdot I = I$

Projection matrix

- Projection matrix: $P = A(A^T A)^{-1} A^T$

- **Question:** If b is perpendicular to $C(A)$, what do I get when I apply projection?

  Answer: In that case b belongs to the null space of A^T and therefore $A^T b = 0$. We then obtain $Pb = A(A^T A)^{-1} A^T b = 0$

- **Question:** What is the projection that gives me e?

  Answer: $P'b = e \Rightarrow P'b = b - p \Rightarrow P'b = b - Pb$

  $\Rightarrow P'b = Ib - Pb = (I - P)b \Rightarrow P' = I - P$

Projection matrix: Properties

- Properties of P

➢Symmetric

$A^T A$ symmetric and therefore $(A^T A)^{-1}$ is symmetric

[to prove this we use the property $(A^{-1})^T = (A^T)^{-1}$]

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

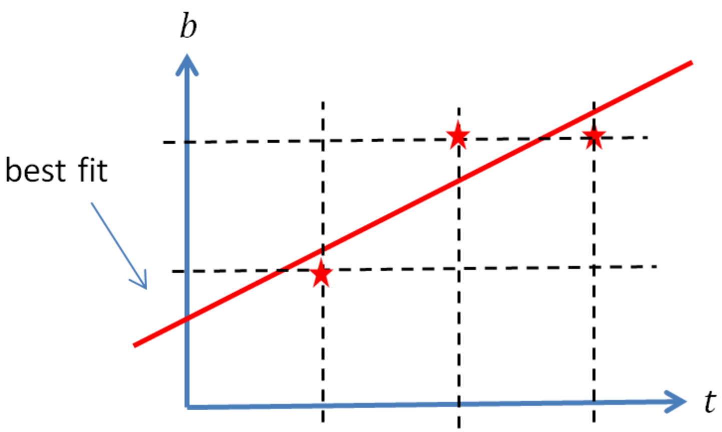
➢$P^2 = P$

$P^2 = A\ (A^T A)^{-1} A^T A\ (A^T A)^{-1} A^T = A\ (A^T A)^{-1} [A^T A\ (A^T A)^{-1}] A^T =$

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

Application: Least squares method. Fitting by a line.

- Problem: I am given the three points shown with stars in the figure on the right. I want to fit them on the "best" possible straight line.
- The points are (1,1), (2,2), (3,2)
- The required line is $b = C + Dt$ with C and D unknowns.



- The three points must satisfy the line equation:

$$C + D = 1$$
$$C + 2D = 2$$
$$C + 3D = 2$$

Application: Least squares method. Fitting by a line.

- Problem: Solve the system $Ax = b$

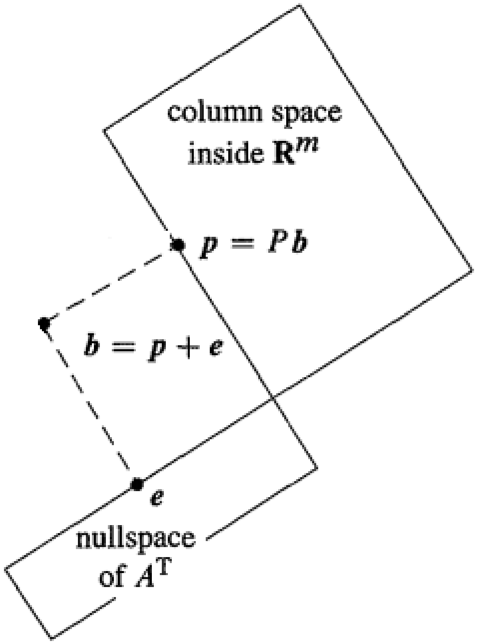
$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

- The system is not solvable because $b \notin C(A)$ (show that).
- Solution: Solve $A\hat{x} = p$ instead, where p is the projection of b onto $C(A)$.

  A random b is written as:

  $b - p = e \Rightarrow b = p + e$

  p is in column space and e is perpendicular to the column space.
- Projection kills e and keeps p.



Solution of the specific example

- The proposed approach $A\hat{x} = p$ yields the solution which can be also obtained if we look for an $\hat{x}$ that minimizes the function
$$\|A\hat{x} - b\|^2 = \|e\|^2$$
- The above function is the magnitude of the total error.
- This method is called **linear regression**.
- Let's solve this specific problem at the end!

$A^T A = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$ $A^T b = \begin{bmatrix} 5 \\ 11 \end{bmatrix}$

- We can use the inverse $(A^T A)^{-1} = \begin{bmatrix} 7/3 & -1 \\ -1 & 1/2 \end{bmatrix}$
- Or we can solve directly the equations $3C + 6D = 5, 6C + 14D = 11$.
- Final solution is $D = \frac{1}{2}, C = \frac{2}{3}$. The "best" line is $b = \frac{2}{3} + \frac{1}{2}t$ (red line).

Solution of the specific example

- As mentioned, an alternative approach is to find the unknowns that minimize the error function:

$$(C+D-1)^2+(C+2D-2)^2+(C+3D-2)^2$$

- We must take the partial derivatives with respect to the two unknowns and set them to zero.
- Verify that, by implementing this method, you get the same solution as previously.
- The vector p is obtained by:

$$p_1 = C + D = \frac{1}{2} + \frac{2}{3} = \frac{7}{6}$$

$$p_2 = C + 2D = \frac{2}{3} + 1 = \frac{5}{3}$$

$$p_3 = C + 3D = \frac{2}{3} + \frac{3}{2} = \frac{13}{6}$$

Solution of the specific example

- Furthermore:

$$e_1 = \frac{-1}{6}$$
$$e_2 = \frac{2}{6}$$
$$e_3 = \frac{-1}{6}$$

- Verify that:
  - $b = p + e$
  - p and e are perpendicular to each other

Problem

- **Problem:** Show that if A has independent columns then $A^T A$ is invertible. I will need this in the next lecture.
- I must prove that $A^T A x = 0$ implies $x = 0$. I assume then that $A^T A x = 0$.
- The above implies $x^T A^T A x = 0 \Rightarrow (Ax)^T A x = 0$
- I define $Ax = y$ and therefore $y^T y = 0 \Rightarrow ||y||^2 = 0 \Rightarrow y = 0 \Rightarrow Ax = 0 \Rightarrow x = 0$ if A has independent columns!
- There is one case in which the columns of A are for sure independent and this is when they are perpendicular unit vectors!