$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}_{m\times n}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1}}_{x} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}_{m\times 1}}_{b}$$

CX1104: **Linear Algebra for Computing**

Chap. No : **7.1.2**

Lecture : **Least Squares**

Topic : **Introduction**

Concept : **The Least Squares Problem**

Instructor: **A/P Chng Eng Siong**

TAs: **Zhang Su, Vishal Choudhari**

Rev: 2nd July 2020

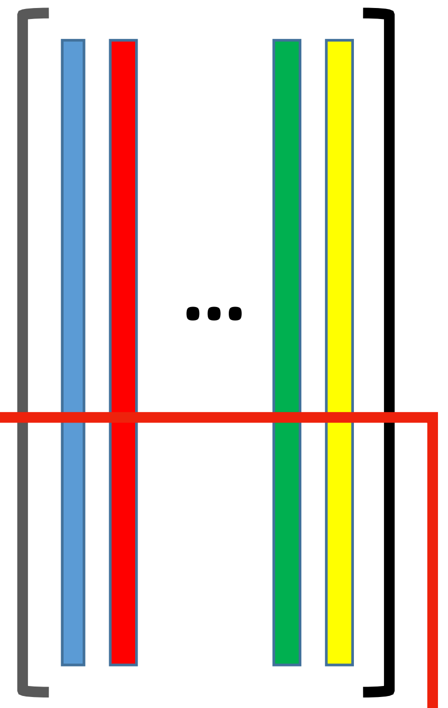
Consistency in a System of Equations

$$Ax = b$$

Consistent
(solutions exist)

- b is in column space of A, i.e,
  b is formed by linear combinations of A's columns.
- Rank (A) = Rank ($A \,|\, b$), i.e,
  rank of A is same as that of the augmented matrix.

Consider



$M \gg N$

Inconsistent
(solutions don't exist)

- b is NOT in column space of A, i.e,
  b is NOT formed by linear combinations of A's columns.
- Occurs when $M \gg N$ (**over-determined**), i.e,
  there exist more equations than unknowns.
- The rows of A are dependent but,
  their corresponding b values are not consistent.
- Rank (A) < Rank ($A \,|\, b$), i.e,
  rank of A is less than that of the augmented matrix.

Least Squares Solution for Inconsistent Equations

Consider solving the system of equations: $Ax = b$

Note:

- Matrix $A \in R^{M \times N}$, where
  - M denotes no. of rows/equations
  - N denotes no. of columns/unknowns
- $x \in R^N$
- $b \in R^M$
- When $M \gg N$,
  - the system is over-determined
  - the equations may be inconsistent
  - there may be no solution

Best we can do?

Find x such that Ax is as close to b as possible!

If A is $m \times n$ and $\mathbf{b}$ is in $\mathbb{R}^m$, a **least-squares solution** of $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

for all $\mathbf{x}$ in $\mathbb{R}^n$.

Think of $A\mathbf{x}$ as an *approximation* to $\mathbf{b}$. The smaller the distance between $\mathbf{b}$ and $A\mathbf{x}$, given by $\|\mathbf{b} - A\mathbf{x}\|$, the better the approximation. The **general least-squares problem** is to find an $\mathbf{x}$ that makes $\|\mathbf{b} - A\mathbf{x}\|$ as small as possible. The adjective "least-squares" arises from the fact that $\|\mathbf{b} - A\mathbf{x}\|$ is the square root of a sum of squares.

Lay, Linear Algebra and its Applications (4th Edition)

360 CHAPTER 6 Orthogonality and Least Squares

Definitions

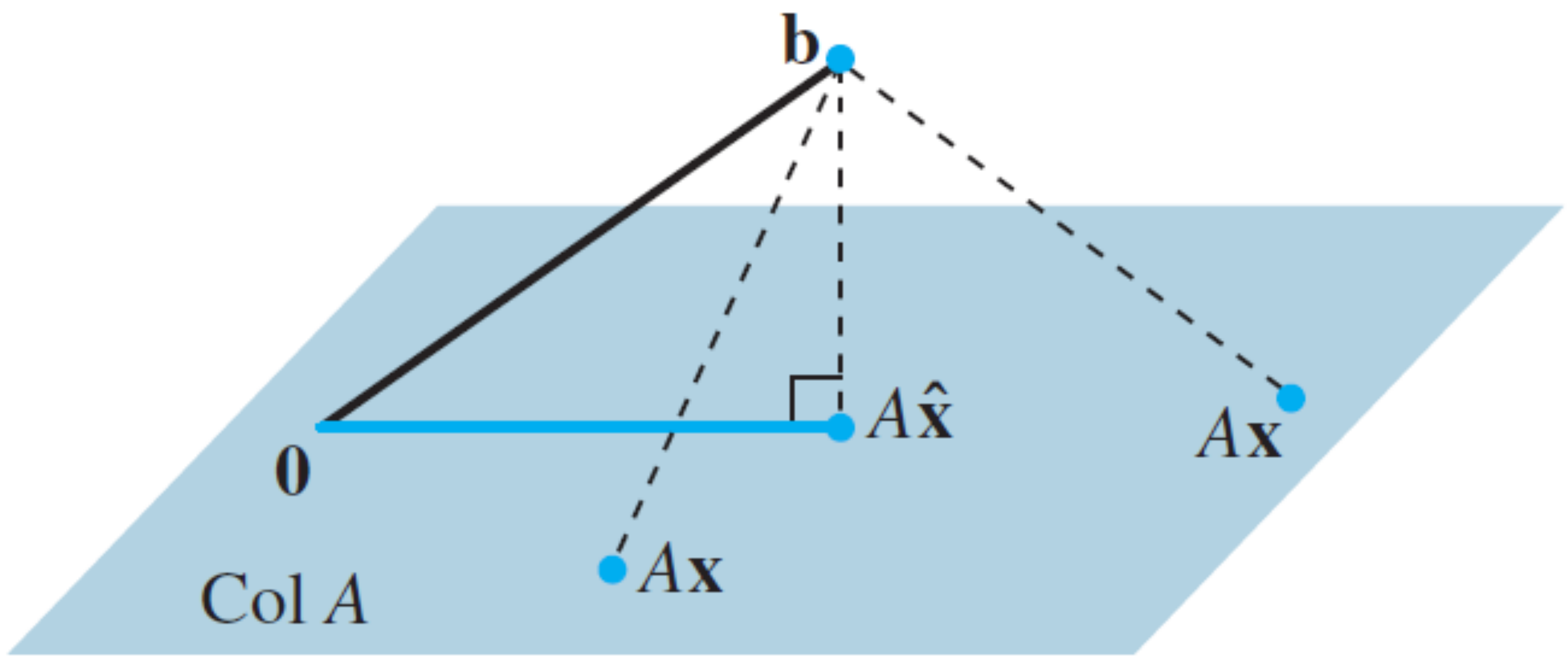


FIGURE 1 The vector $\mathbf{b}$ is closer to $A\hat{\mathbf{x}}$ than to $A\mathbf{x}$ for other $\mathbf{x}$.

The most important aspect of the least-squares problem is that no matter what $\mathbf{x}$ we select, the vector $A\mathbf{x}$ will necessarily be in the column space, Col A. So we seek an $\mathbf{x}$ that makes $A\mathbf{x}$ the closest point in Col A to $\mathbf{b}$. See Fig. 1. (Of course, if $\mathbf{b}$ happens to be in Col A, then $\mathbf{b}$ *is* $A\mathbf{x}$ for some $\mathbf{x}$, and such an $\mathbf{x}$ is a "least-squares solution.")

If a linear system is consistent, then its exact solutions are the same as its least squares solutions, in which case the least squares error is zero.

NOTE:
When the linear system $Ax = b$ is inconsistent, b does not lie in the column space of A.

Lay, Linear Algebra and its Applications (4th Edition)

360 **CHAPTER 6** Orthogonality and Least Squares

To explain the terminology in this problem, suppose that the column form of $\mathbf{b} - A\mathbf{x}$ is

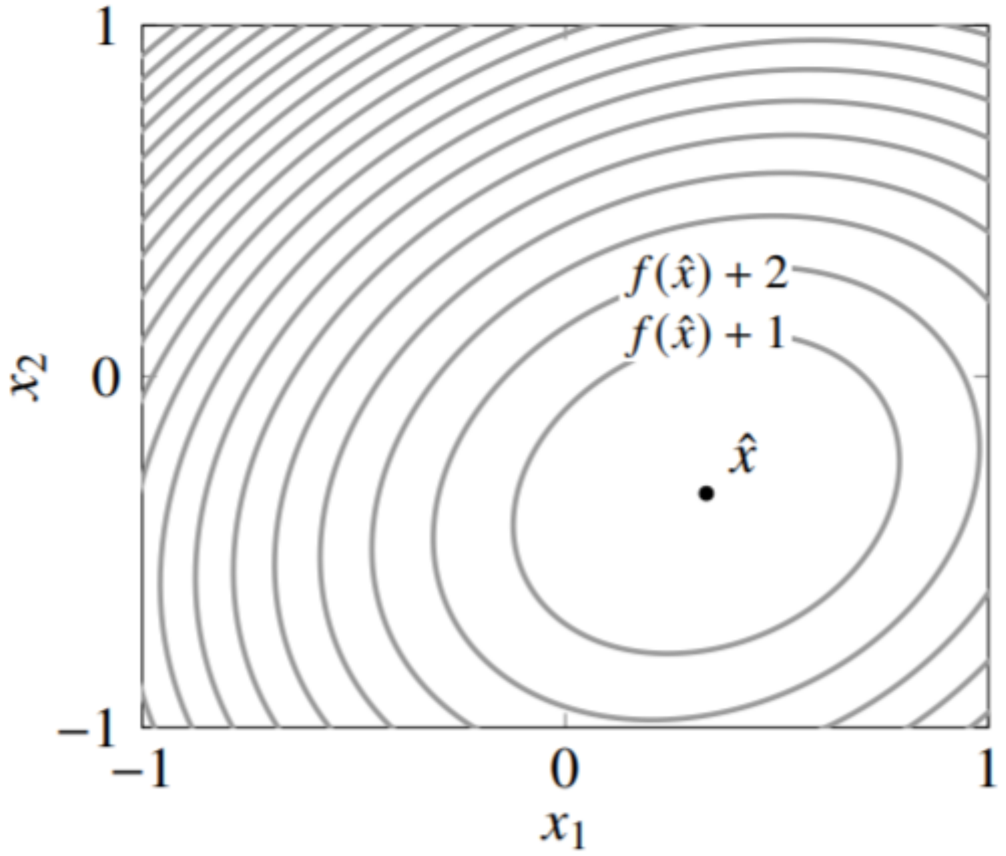
$$\mathbf{b} - A\mathbf{x} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}$$

The term "least squares solution" results from the fact that minimizing $\|\mathbf{b} - A\mathbf{x}\|$ also has the effect of minimizing $\|\mathbf{b} - A\mathbf{x}\|^2 = e_1^2 + e_2^2 + \cdots + e_m^2$.

Example

Example

$$A = \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$



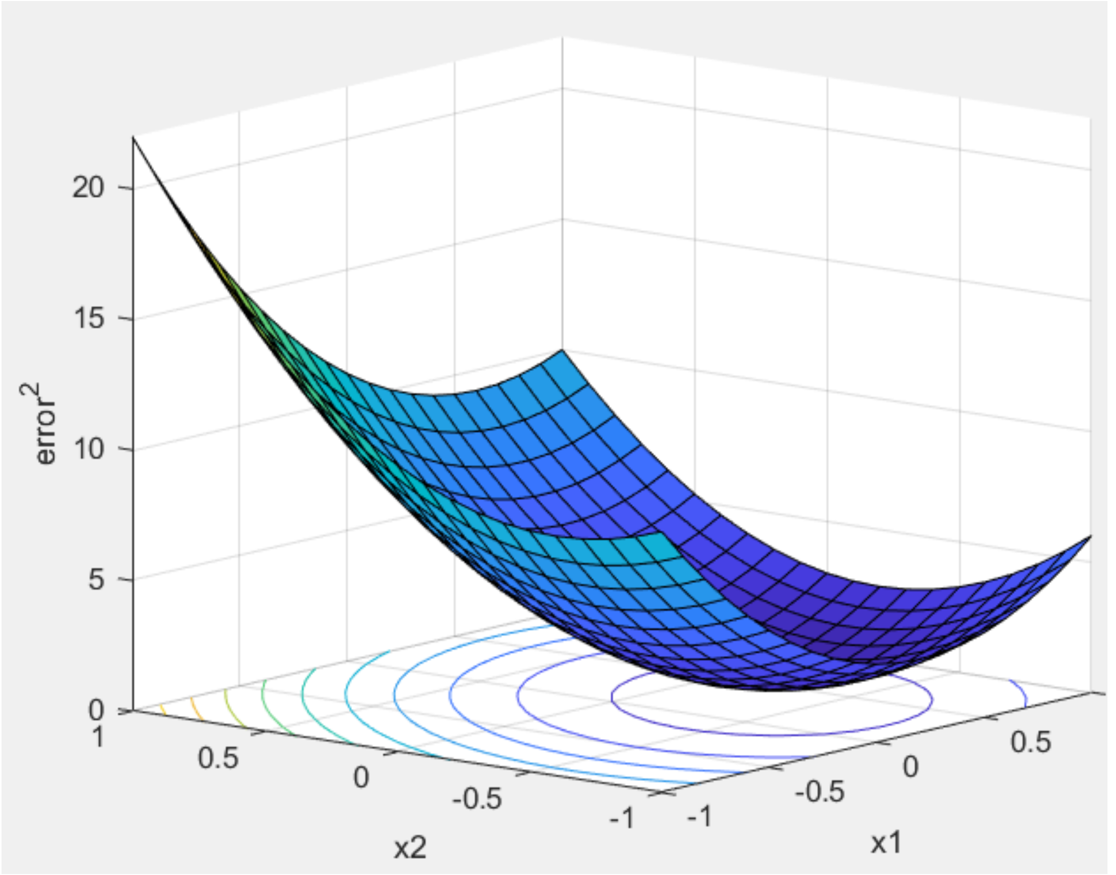
- the least squares solution $\hat{x}$ minimizes

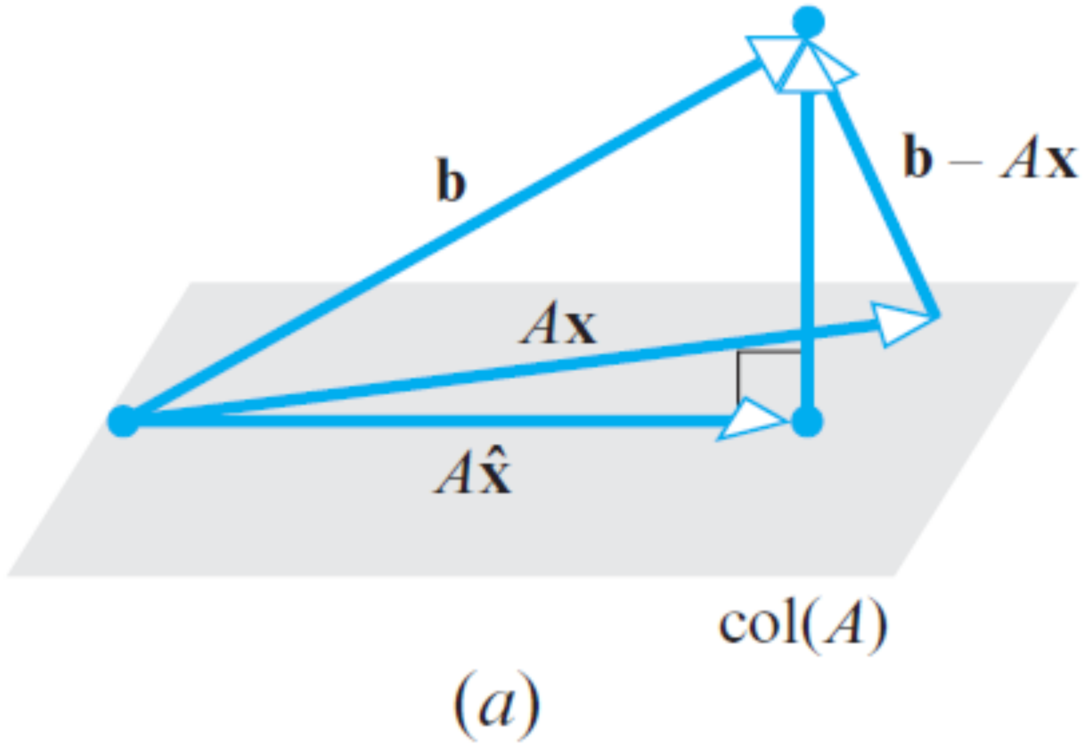
$$f(x) = \|Ax - b\|^2 = (2x_1 - 1)^2 + (-x_1 + x_2)^2 + (2x_2 + 1)^2$$

- to find $\hat{x}$, set derivatives with respect to x_1 and x_2 equal to zero:

$$10x_1 - 2x_2 - 4 = 0, \qquad -2x_1 + 10x_2 + 4 = 0$$

solution is $(\hat{x}_1, \hat{x}_2) = (1/3, -1/3)$






```
opStr =

    'x1=0.30,   x2=-0.30,   err^2=0.680 '
```

```
%ch6_4_Ex1.m
%Chng Eng Siong, plotting the error wrt x
close all; clear all;
A = [2 0; -1 1; 0 2];
b = [1 0 -1]';
[x1,x2] = meshgrid(-1:0.1:1, -1:0.1:1);
[m,n] = size(x1);
z = zeros(m,n);
for i=1:m
    for j=1:n
        z(i,j) = norm(b - (x1(i,j)*A(:,1)+x2(i,j)*A(:,2))).^2;
    end
end
surfc(x1,x2,z)
xlabel('x1'); ylabel('x2'); zlabel('error^2');

% Lets print the min value and the x vector
minIdx = find(z == min(z(:)));
x1(minIdx), x2(minIdx), z(minIdx)
opStr = sprintf('x1=%0.2f,  x2=%0.2f,  err^2=%.3f ',x1(minIdx),x2(minIdx),z(minIdx))
```

Boyd – Lecture 6

Ref: https://www.youtube.com/watch?v=ZKhZclqfR5E&t=1850s