Math 22 –
Linear Algebra and its applications

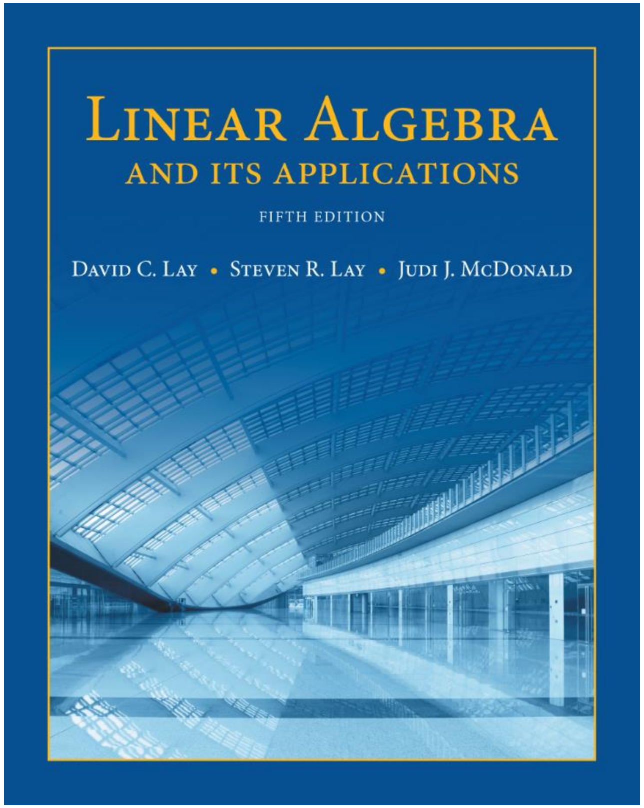
- Lecture 24 -

Instructor: Bjoern Muetzel

6 Orthogonality and Least Squares

6.5

LEAST-SQUARES PROBLEMS AND LINEAR MODELS



LEAST-SQUARES PROBLEMS

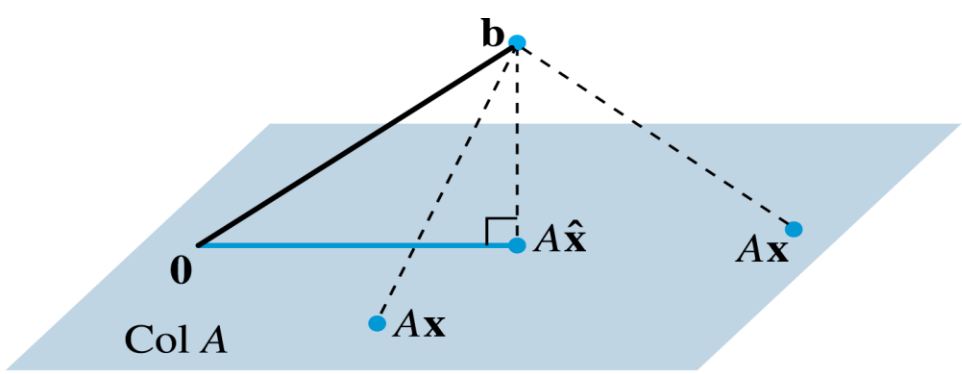
- **Definition:** If A is $m\times n$ and **b** is in $\mathbb{R}^m$, a **least-squares solution** of $A\boldsymbol{x} = \boldsymbol{b}$ is an $\hat{\boldsymbol{x}}$ in $\mathbb{R}^n$ such that

$$\boxed{\left|\left|\boldsymbol{b} - A\,\hat{\boldsymbol{x}}\right|\right| \le \left|\left|\boldsymbol{b} - A\boldsymbol{x}\right|\right|}$$ for all **x** in $\mathbb{R}^n$.

- **Note:** In this problem we assume that $A\boldsymbol{x} = \boldsymbol{b}$ has **no solution**.

Another way of seeing this, is that we find a the vector

$A\,\hat{\boldsymbol{x}} = \hat{\boldsymbol{b}}$ in Col A that is **closest** to **b** .



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

- Given A and **b**, we apply the **Best Approximation Theorem** to the subspace Col A. So we take $\boxed{\hat{\boldsymbol{b}} = \ proj_{ColA}(\boldsymbol{b})}$

- Because $\hat{\boldsymbol{b}}$ is in the column space A, the equation $A\,\boldsymbol{x} = \hat{\boldsymbol{b}}$ *is* consistent, and there is an $\hat{\boldsymbol{x}}$ in $\mathbb{R}^n$ such that

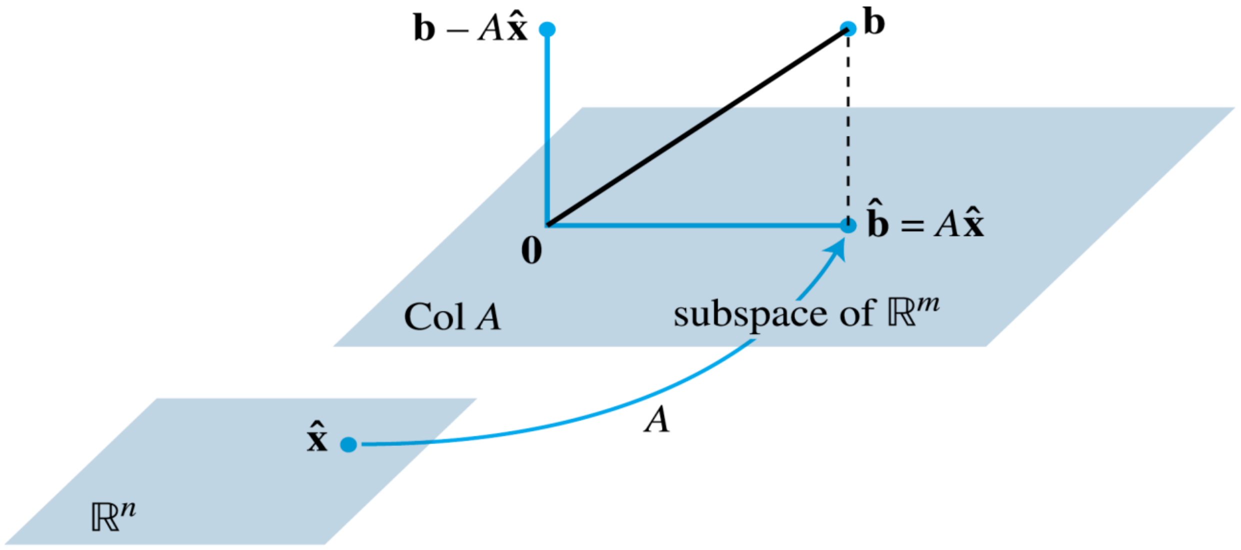
$$\boxed{A\,\hat{\boldsymbol{x}} = \hat{\boldsymbol{b}}} \qquad (1)$$

- Since $\hat{\boldsymbol{b}}$ is the closest point in Col A to **b**, a vector $\hat{\boldsymbol{x}}$ is a least-squares solution of $A\,\boldsymbol{x} = \hat{\boldsymbol{b}}$ if and only if $\hat{\boldsymbol{x}}$ satisfies (1).

How can we find $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{b}} = \ \boldsymbol{proj_{ColA}}(\boldsymbol{b})$?

Trick: ($\hat{\boldsymbol{b}}$ **- b**) in $ColA^{\perp}$

SOLUTION OF THE GENREAL LEAST-SQUARES PROBLEM



The least-squares solution $\hat{\mathbf{x}}$ is in $\mathbb{R}^n$.

SOLUTION OF THE GENREAL LEAST-SQUARES PROBLEM

- **Theorem 13:** The set **of least-squares solutions** of $A\boldsymbol{x} = \boldsymbol{b}$ coincides with the nonempty set of solutions of the **normal equation**.

$$\boxed{A^T A\mathrm{x} = A^T \mathrm{b}}$$

- **Example 1:** Find a least-squares solution of the inconsistent system for

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \mathrm{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

- Then draw a picture.

APPLICATIONS TO LINEAR MODELS

- **Least-Square Lines:** We can fit a line to a set of data points.

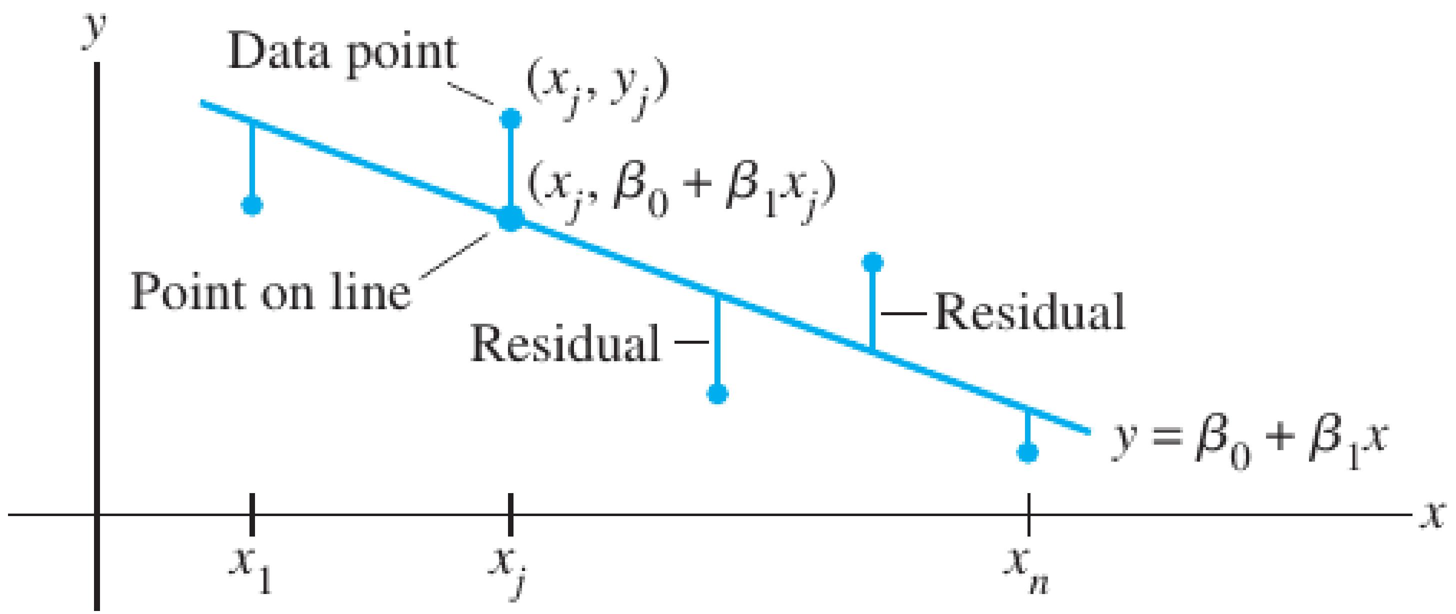


FIGURE 1 Fitting a line to experimental data.

- **How** to set up the **system of linear equations** can be seen from the following **table**:

Predicted y-value		Observed y-value
$\beta_0 + \beta_1 x_1$	=	y_1
$\beta_0 + \beta_1 x_2$	=	y_2
$\vdots$		$\vdots$
$\beta_0 + \beta_1 x_n$	=	y_n

We obtain an **equation**

$$X\boldsymbol{\beta} = \mathbf{y}, \quad \text{where } X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Looking at the table we want to **minimize**:

APPLICATIONS TO LINEAR MODELS

Hence we have to solve the **Least Square Problem:**

- **Example:** Find the **Least-Square Line** that best fits the data points (2,1), (5,2), (7,3) and (8,3). Draw a picture.