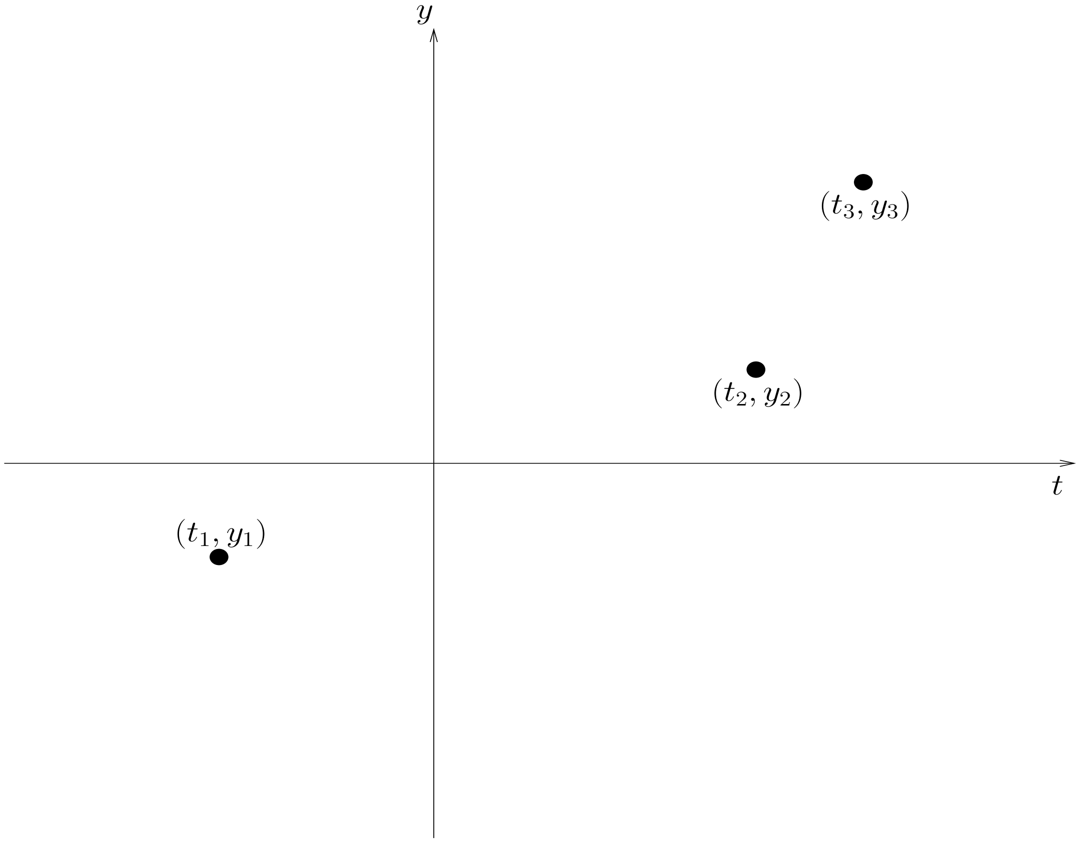
Math 304 (Spring 2010) - Lecture 15
Least Squares - Problem Definition

Emre Mengi
Department of Mathematics
Koç University
emengi@ku.edu.tr

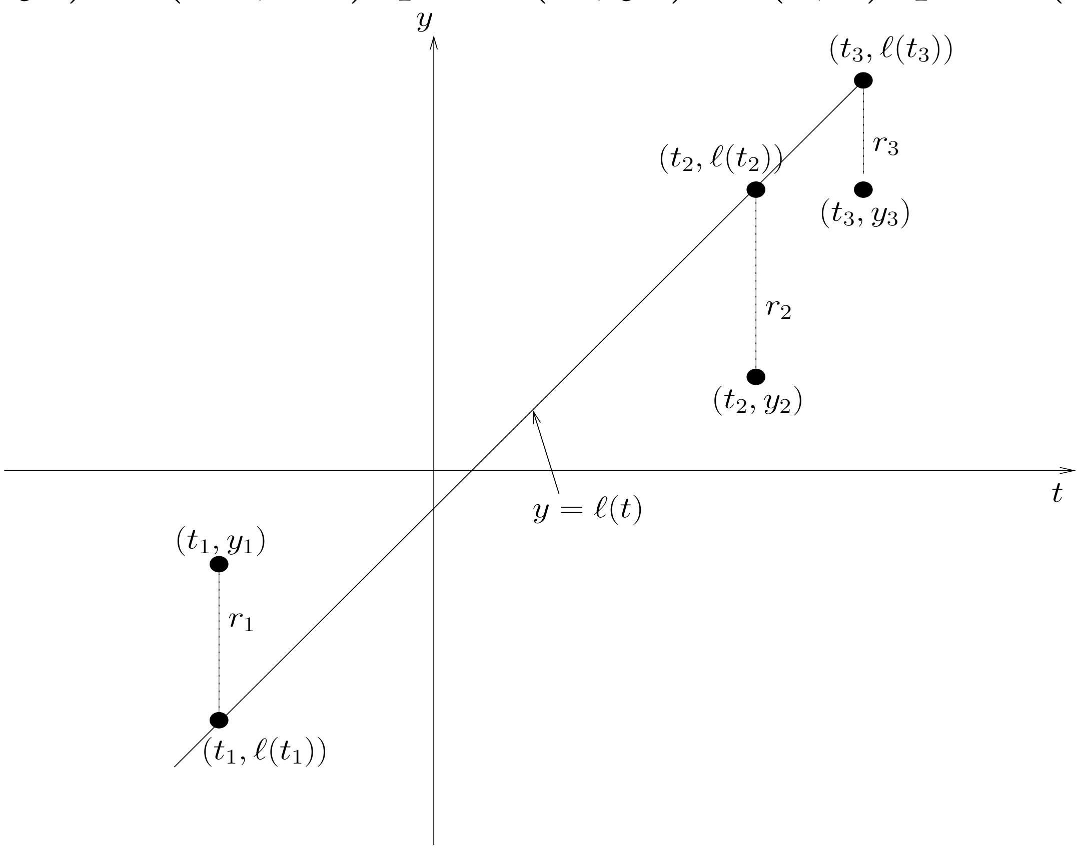
Least Squares, Problem Definition

Given $p_1 = (t_1, y_1) = (-2, -1)$, $p_2 = (t_2, y_2) = (3, 1)$, $p_3 = (t_3, y_3) = (4, 3)$.



Least Squares, Problem Definition

Given $p_1 = (t_1, y_1) = (-2, -1)$, $p_2 = (t_2, y_2) = (3, 1)$, $p_3 = (t_3, y_3) = (4, 3)$.



- Find the line $\ell(t) = x_1 t + x_0$ that best fits the points p_1, p_2, p_3. (The unknowns are x_0, x_1.)

Least Squares, Problem Definition

- Find the line $\ell(t) = x_1 t + x_0$ so that

$$\sqrt{\sum_{i=1}^{3} (\ell(t_i) - y_i)^2} = \sqrt{(-2x_1 + x_0 - (-1))^2 + (3x_1 + x_0 - 1)^2 + (4x_1 + x_0 - 3)^2}$$

is as small as possible.

Least Squares, Problem Definition

- Find the line $\ell(t) = x_1 t + x_0$ so that

  $$\sqrt{\sum_{i=1}^{3} (\ell(t_i) - y_i)^2} = \sqrt{(-2x_1 + x_0 - (-1))^2 + (3x_1 + x_0 - 1)^2 + (4x_1 + x_0 - 3)^2}$$

  is as small as possible.

- Define

  $$r = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} \ell(t_1) - y_1 \\ \ell(t_2) - y_2 \\ \ell(t_3) - y_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_0 \\ x_1 \end{bmatrix}}_{x} - \underbrace{\begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}}_{b}$$

Least Squares, Problem Definition

- Find the line $\ell(t) = x_1 t + x_0$ so that

  $$\sqrt{\sum_{i=1}^{3}(\ell(t_i) - y_i)^2} = \sqrt{(-2x_1 + x_0 - (-1))^2 + (3x_1 + x_0 - 1)^2 + (4x_1 + x_0 - 3)^2}$$

  is as small as possible.

- Define

  $$r = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} \ell(t_1) - y_1 \\ \ell(t_2) - y_2 \\ \ell(t_3) - y_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_0 \\ x_1 \end{bmatrix}}_{x} - \underbrace{\begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}}_{b}$$

- The problem can be posed as

  find $x = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$ such that $\|r\| = \|Ax - b\|$ is as small as possible.

Least Squares, Problem Definition

- More generally given m points in $\mathbf{R}^2$

$$p_i = (t_i, y_i), \quad i = 1, \ldots, m$$

Least Squares, Problem Definition

- More generally given m points in $\mathbf{R}^2$

$$p_i = (t_i, y_i), \quad i = 1, \ldots, m$$

- Suppose you want to find the polynomial of degree $n-1$ $(n < m)$ in the form

$$P(t) = x_{n-1}t^{n-1} + x_{n-2}t^{n-2} + \cdots + x_1 t + x_0$$

Least Squares, Problem Definition

- More generally given m points in $\mathbf{R}^2$

$$p_i = (t_i, y_i), \quad i = 1, \ldots, m$$

- Suppose you want to find the polynomial of degree $n-1$ $(n < m)$ in the form

$$P(t) = x_{n-1}t^{n-1} + x_{n-2}t^{n-2} + \cdots + x_1 t + x_0$$

minimizing

$$\sqrt{\sum_{i=1}^{m} (P(t_i) - y_i)^2}.$$

Least Squares, Problem Definition

- Define

$$\underbrace{\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}}_{r} = \begin{bmatrix} P(t_1) - y_1 \\ P(t_2) - y_2 \\ \vdots \\ P(t_m) - y_m \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \cdots & t_1^{n-2} & t_2^{n-1} \\ 1 & \cdots & t_2^{n-2} & t_2^{n-1} \\ & & \vdots & \vdots \\ 1 & \cdots & t_m^{n-2} & t_m^{n-1} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix}}_{x} - \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{b}$$

Least Squares, Problem Definition

- Define

$$\underbrace{\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}}_{r} = \begin{bmatrix} P(t_1) - y_1 \\ P(t_2) - y_2 \\ \vdots \\ P(t_m) - y_m \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \cdots & t_1^{n-2} & t_2^{n-1} \\ 1 & \cdots & t_2^{n-2} & t_2^{n-1} \\ & & \vdots & \vdots \\ 1 & \cdots & t_m^{n-2} & t_m^{n-1} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix}}_{x} - \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{b}$$

Remark: The matrix A above is called the *Vandermonde* matrix.

Least Squares, Problem Definition

- Define

$$\underbrace{\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}}_{r} = \begin{bmatrix} P(t_1) - y_1 \\ P(t_2) - y_2 \\ \vdots \\ P(t_m) - y_m \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \cdots & t_1^{n-2} & t_2^{n-1} \\ 1 & \cdots & t_2^{n-2} & t_2^{n-1} \\ & & \vdots & \vdots \\ 1 & \cdots & t_m^{n-2} & t_m^{n-1} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix}}_{x} - \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{b}$$

Remark: The matrix A above is called the *Vandermonde* matrix.

- We want to find $x = \begin{bmatrix} x_0 & x_1 & \cdots & x_{n-1} \end{bmatrix}^T$ minimizing

$$\|r\| = \|Ax - b\|.$$

Least Squares, Problem Definition

Definition: An $m \times n$ system $Ax = b$ is called *overdetermined* if $m > n$.

Least Squares, Problem Definition

Definition: An $m \times n$ system $Ax = b$ is called *overdetermined* if $m > n$.

- Overdetermined systems are usually inconsistent. (*e.g.* It is unlikely that three lines in $\mathbf{R}^2$ intersect each other at a common point.)

Least Squares, Problem Definition

Definition: An $m \times n$ system $Ax = b$ is called *overdetermined* if $m > n$.

- Overdetermined systems are usually inconsistent. (*e.g.* It is unlikely that three lines in $\mathbf{R}^2$ intersect each other at a common point.)

  Example:

$$[\, A \mid b \,] = \begin{bmatrix} 4 & 1 & 3 \\ 3 & 1 & 1 \\ -2 & 1 & -1 \end{bmatrix} \rightsquigarrow \underbrace{\begin{bmatrix} 4 & 1 & 3 \\ 0 & 1/4 & -5/4 \\ 0 & 0 & 8 \end{bmatrix}}_{inconsistent}$$

Least Squares, Problem Definition

<u>Justification:</u>

$\text{Col}(A) = \text{span}\{a_1, a_2, \ldots, a_n\}$ is at most an n-dimen subspace in $\mathbf{R}^m$

Least Squares, Problem Definition

Justification:

$\mathrm{Col}(A) = \mathrm{span}\{a_1, a_2, \ldots, a_n\}$ is at most an n-dimen subspace in $\mathbf{R}^m$

$$\Longrightarrow$$

Most $b \in \mathbf{R}^m$ are not in $\mathrm{Col}(A)$

Least Squares, Problem Definition

Justification:

$\mathrm{Col}(A) = \mathrm{span}\{a_1, a_2, \ldots, a_n\}$ is at most an n-dimen subspace in $\mathbf{R}^m$

$$\Longrightarrow$$

Most $b \in \mathbf{R}^m$ are not in $\mathrm{Col}(A)$

$$\Longrightarrow$$

$Ax = b$ is inconsistent for most $b \in \mathbf{R}^m$

Least Squares, Problem Definition

<u>Justification:</u>

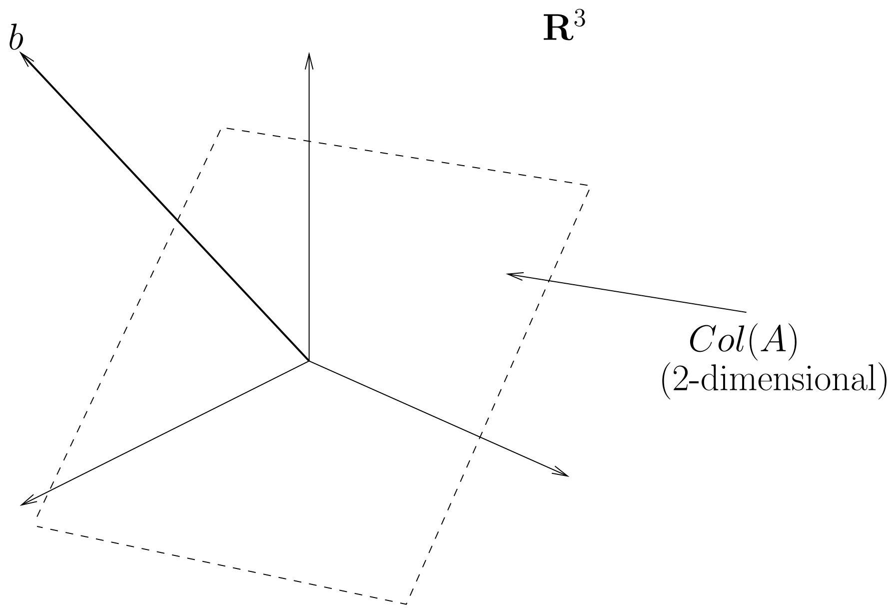
$\mathrm{Col}(A) = \mathrm{span}\{a_1, a_2, \ldots, a_n\}$ is at most an n-dimen subspace in $\mathbf{R}^m$

$$\Longrightarrow$$

Most $b \in \mathbf{R}^m$ are not in $\mathrm{Col}(A)$

$$\Longrightarrow$$

$Ax = b$ is inconsistent for most $b \in \mathbf{R}^m$



e.g. $n = 2$, $m = 3$

$$A = \begin{bmatrix} x & x \\ x & x \\ x & x \end{bmatrix}$$

Least Squares, Problem Definition

Least Squares Problem: Given an overdetermined system $Ax = b$.

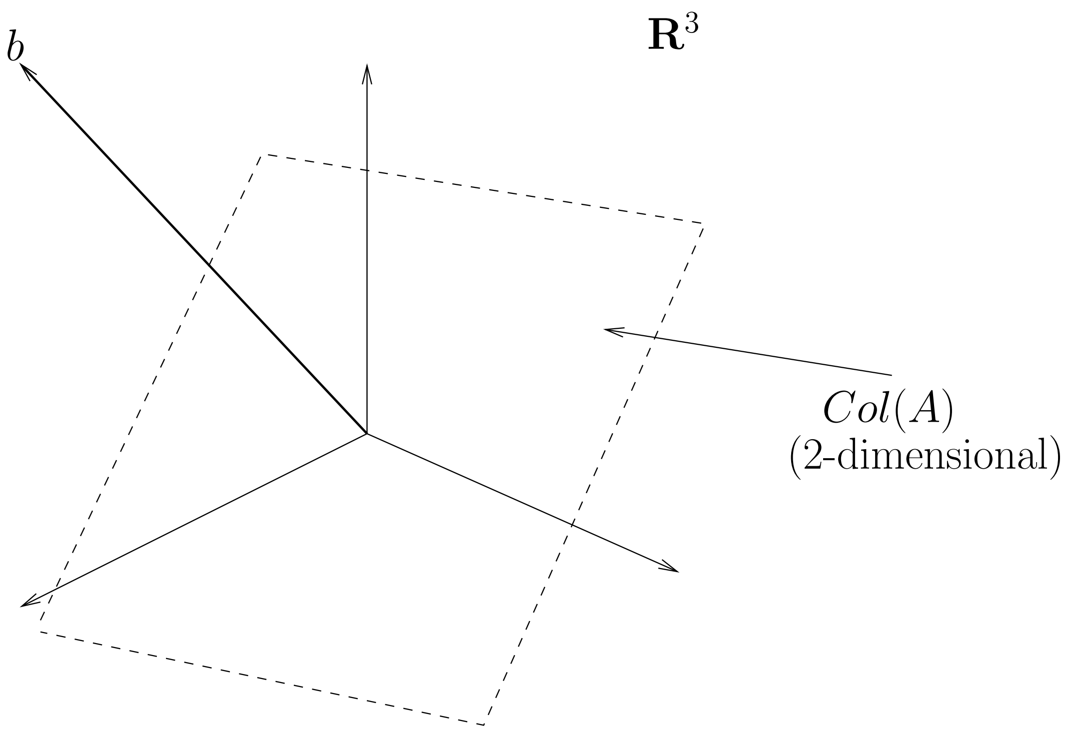
Find $x \in \mathbf{R}^n$ such that $\|Ax - b\|$ is as small as possible.

Least Squares, Problem Definition

Least Squares Problem: Given an overdetermined system $Ax = b$.

Find $x \in \mathbf{R}^n$ such that $\|Ax - b\|$ is as small as possible.

- Geometric interpretation: Find the point on the hyperplane $\mathrm{Col}(A)$ that is closest to b.



Next Lecture

- Solution of the least squares problem by exploiting the QR factorization
- Interpolation (Fausett, Chapter 8)