Linear Least Squares

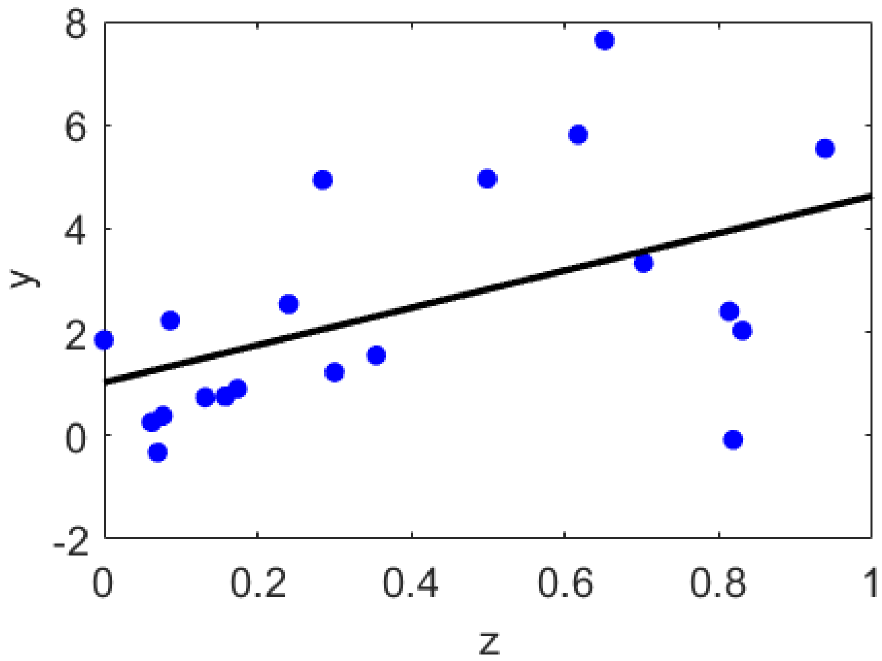
CPSC 406 – Computational Optimization

Overview

- Least-quares for data fitting
- Solution properties
- Solution methods

Fitting a line to data

- Given data points $(z_i, y_i), i = 1, \ldots, m$
- Find a line $y = c + sz$ that best fits the data



$$\min_{c,s} \sum_{i=1}^{m} (y_i - \bar{y}_i)^2 \quad \text{st} \quad \bar{y}_i = c + sz_i$$

Matrix formulation

Generic form

$$\min_x \|Ax - b\|_2^2 = \sum_{i=1}^{m} (a_i^T x - b_i)^2 \quad \text{where} \quad A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix}$$

Example

$$\min_{c,s} \sum_{i=1}^{m} (y_i - \bar{y}_i)^2 \quad \text{st} \quad \bar{y}_i = c + s z_i$$

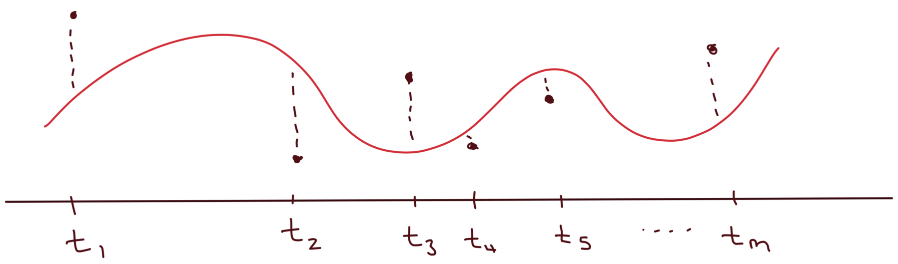
where

$$A = \begin{bmatrix} 1 & z_1 \\ \vdots & \vdots \\ 1 & z_m \end{bmatrix}, \quad b = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}, \quad x = \begin{bmatrix} c \\ s \end{bmatrix}$$

Example: Polynomial data fitting

Given m measurements y_i taken at times t_i:

$$(t_1, y_1), \dots, (t_m, y_m)$$



Polynomial model $p(t)$ of degree $(n-1)$:

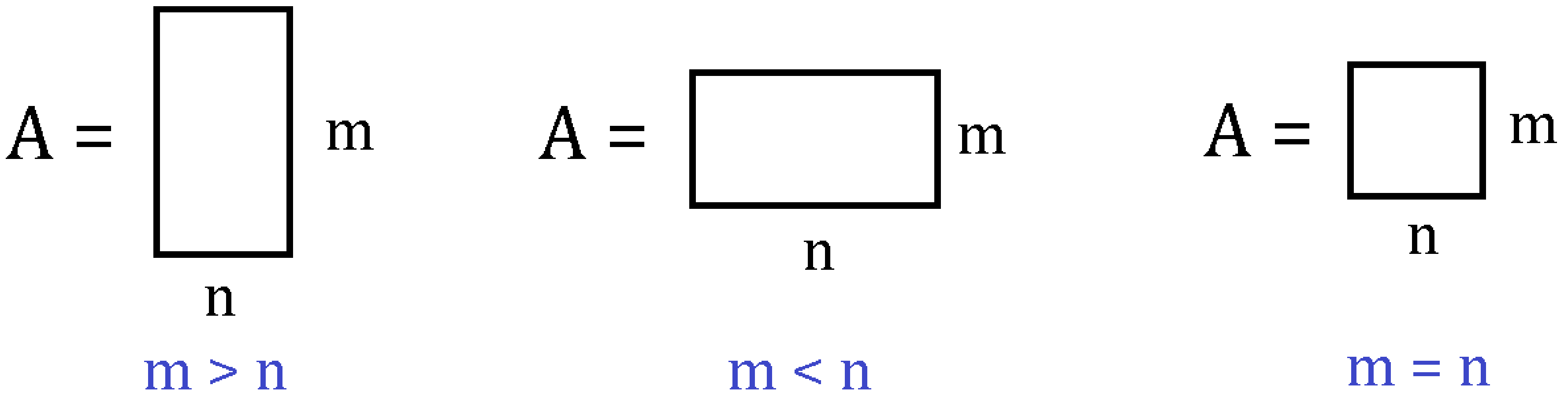
$$p(t) = x_0 + x_1 t + x_2 t^2 + \cdots + x_{n-1} t^{n-1} \quad (x_i = \text{coeff's})$$

Find coefficients $x_0, x_1, \dots, x_{n-1}$ such that

$$\begin{matrix} p(t_1) \approx y_1 \\ \vdots \\ p(t_m) \approx y_m \end{matrix} \quad \iff \quad \underbrace{\begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_m & t_m^2 & \cdots & t_m^{n-1} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_0 \\ \vdots \\ x_{n-1} \end{bmatrix}}_{x} \approx \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}_{b}$$

Solving linear systems

Find x where $Ax \approx b$



- $m > n$ (overdetermined): possibly no exact solution
- $m < n$ (underdetermined): possibly infinitely many solutions
- $m = n$ (square): possibly unique solution

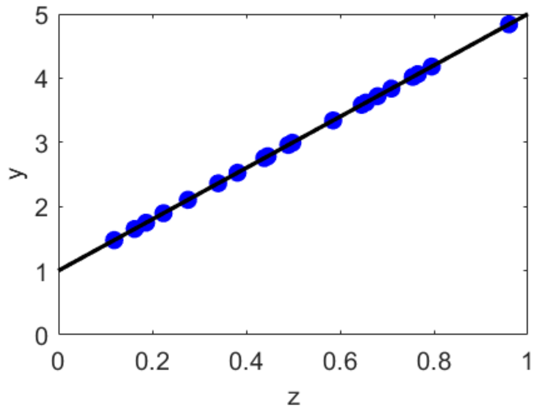
Question

Suppose that A is an $m \times n$ full-rank matrix with $m > n$. Then

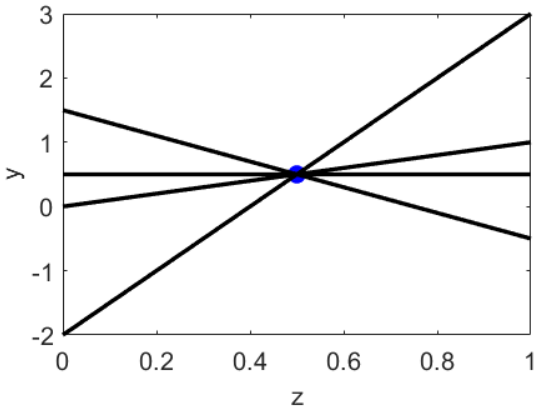
a. $Ax = b$ has a unique solution for every $b \in \mathbb{R}^m$
b. $Ax = b$ has a solution only if $b \in \mathbf{range}(A)$
c. $Ax = b$ has a solution only if $b \in \mathbf{range}(A^T)$
d. A is invertible

Over and underdetermined systems

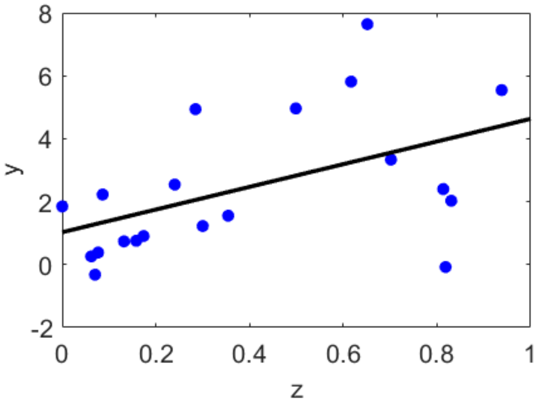
Find x where $Ax = b$



- 1 solution
- overdetermined
- feasible
- $b \in \textbf{range}(A)$



- infinitely many solutions
- underdetermined
- feasible
- $b \in \textbf{range}(A)$



- no solution
- overdetermined
- infeasible
- $b \notin \textbf{range}(A)$

Least-squares optimality

$$x^* = \operatorname*{argmin}_x f(x) := \frac{1}{2}\|Ax - b\|_2^2 = \frac{1}{2}\sum_{i=1}^{m}(a_i^T x - b_i)^2$$

- quadratic objective

$$\|r\|_2^2 = r^T r \quad \Longrightarrow \quad f(x) = \frac{1}{2}(Ax - b)^T(Ax - b) = \frac{1}{2}x^T A^T A x - b^T A x + \frac{1}{2}b^T b$$

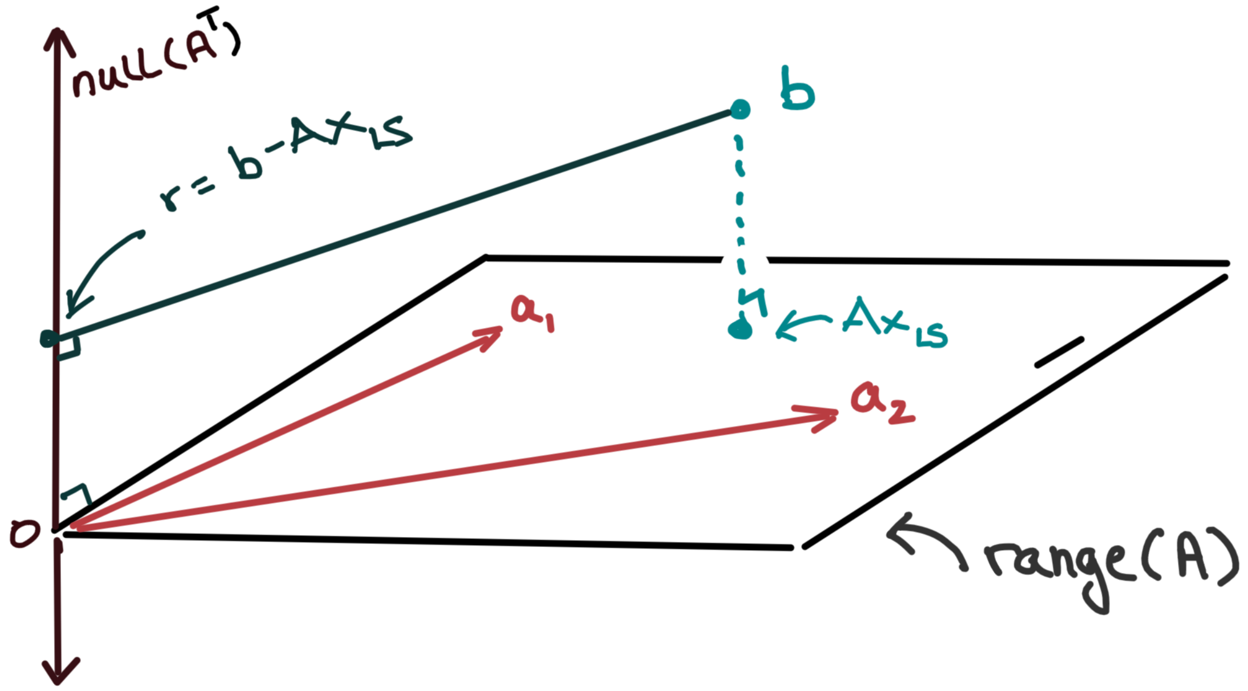
- gradient: $\nabla f(x) = A^T A x - A^T b$
- the solution of LS must be a stationary point of f:

$$\nabla f(x^*) = 0 \quad \Longleftrightarrow \quad A^T A x^* - A^T b = 0 \quad \Longleftrightarrow \quad \underbrace{A^T A x^* = A^T b}_{\text{normal equations}}$$

- If A has full column rank $\quad \Longrightarrow \quad x^* = (A^T A)^{-1} A^T b \quad \text{(unique)}$

Geometric view

$$A = [a_1 \; a_2 \; \cdots \; a_n] \quad \text{where} \quad a_i \in \mathbb{R}^m$$



$$\textbf{range}(A) = \{y \mid y = Ax \text{ for some } \quad x \in \mathbb{R}^n\}$$
$$\textbf{null}(A^T) = \{z \mid A^T z = 0\}$$
$$\textbf{range}(A)^\perp = \textbf{null}(A^T)$$

Orthogonal projection

- orthogonality of residual $r = b - Ax$ and columns of A

$$\left.\begin{array}{c} a_1^T r = 0 \\ a_2^T r = 0 \\ \vdots \\ a_n^T r = 0 \end{array}\right\} \iff \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{bmatrix} r \iff A^T r = 0 \iff r \in \mathbf{null}(A^T)$$

- the following conditions are equivalent
  1. $A^T r = 0$
  2. $r \in \mathbf{null}(A^T)$
  3. $A^T A x = A^T b$
- projection is unique:

$$y^* = Ax^* = \mathbf{proj}_{\mathbf{range}(A)}(b)$$

Question

Consider the single-variable least-squares problem with data

$$A = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad (n = 1)$$

If $m = 3$ and $b = (1, 3, 5)$, what is the least-squares solution x^*?

a. $x^* = 1$

b. $x^* = 3$

c. $x^* = 5$

d. $x^* = 9$

In-class exercise

Run the greenhouse gas example in the least-squares notebook:

- Compute the sum-of-squares residual in the trendline.