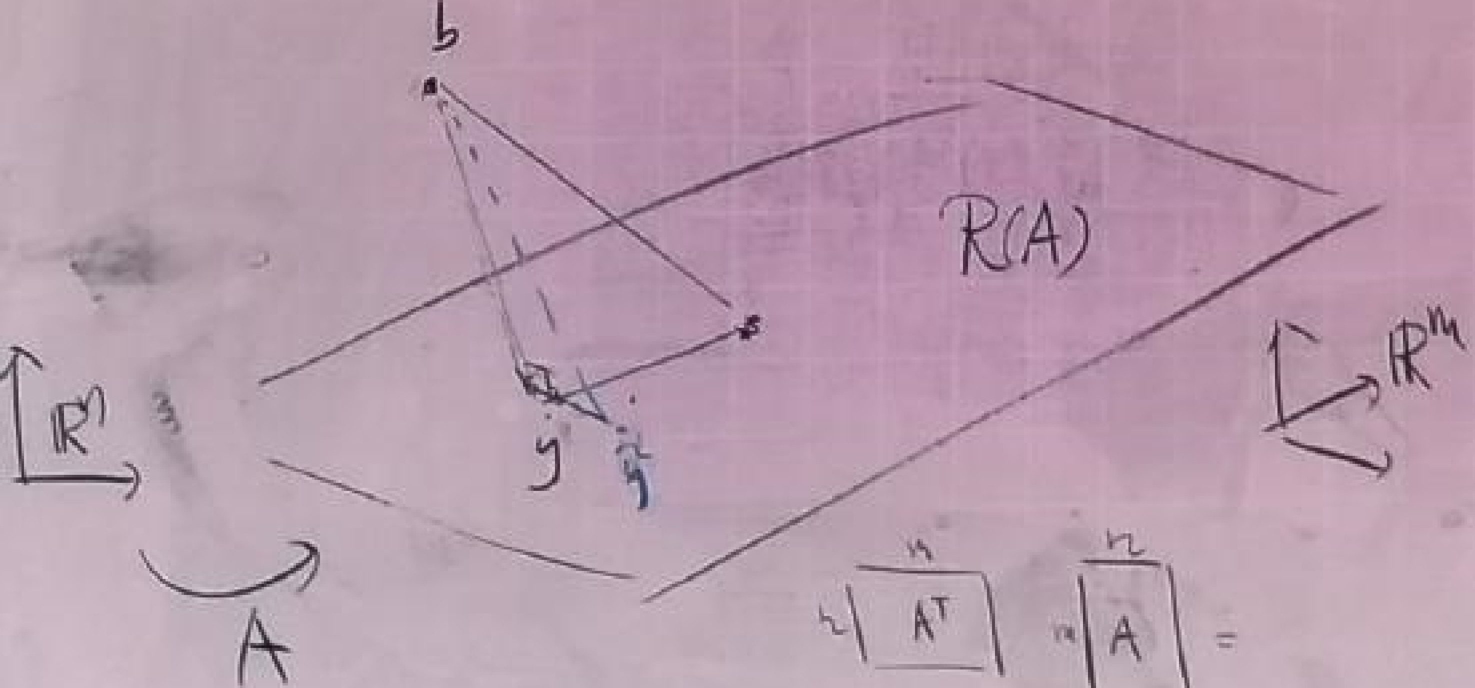
Discrete least square approximation.

$$(*) \quad Ax = b \qquad A : \mathbb{R}^n \longrightarrow \mathbb{R}^m \qquad (**)$$

$b \notin R(A)$ $(*)$ has no solution

$$(\Rightarrow \text{rank } A < m, \text{ or } R(A) \subsetneq \mathbb{R}^m)$$



$$\begin{array}{c} \\ n \end{array} \boxed{\overset{m}{A^T}} \quad \begin{array}{c} \\ m \end{array} \boxed{\overset{n}{A}} =$$

Revised problem:

(**) Find $y \in R(A)$ (and $x \in \mathbb{R}^n$, $Ax = y$)

such that $\|y - b\| = \min_{z \in R(A)} \|z - b\|$

When $\|\cdot\| = \|\cdot\|_2$

(**) is called least square app. of (*)

$\mathbb{R}^m$ Obviously, the solution y satisfies

$$(b - y) \perp R(A)$$

That is $$A^T(Ax - b) = 0_n$$

Normal equation

$$A^T A x = A^T b$$

$$A^T A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

Case 1: rank $A = n$ $(< m)$

$\ker A = \vec{0}$

$\therefore$ The solution to the normal eq is unique.

Case 2. rank $A < n$ (also rank $A < m$)

$\dim(\ker A) > 0$

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \ x^T x \quad \text{subject to} \quad A^T A x - A^T b = 0$$

We will only focus on the case $\operatorname{rank} A = n < m$

<u>Example</u> Polynomial data fitting

Given (x_i, y_i) $\quad i = 1, \dots, m$

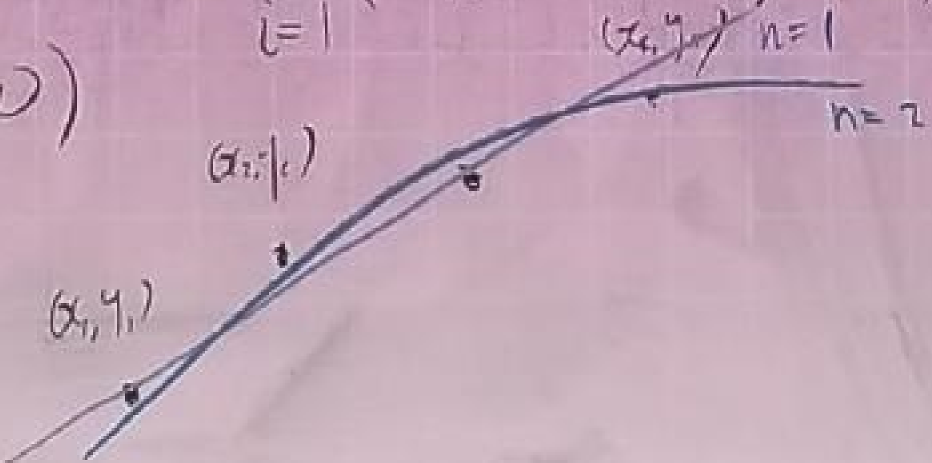
find $a_0, \dots, a_n$ $\quad n+1 < m$

such that $P_n(x) = a_n x^n + \cdots + a_0$

minimizes the quantity

$$E = \sum_{i=1}^{m} \left(y_i - P_n(x_i) \right)^2 w_i ,$$

$(w_i > 0)$



Solving the least square problem

$$E(\vec{a}) = \sum_{i=1}^{m} \left(y_i - P_n(\vec{a}, x_i)\right)^2 w_i$$

$$P_n(\vec{a}, x_i) = a_n x_i^n + a_{n-1} x_i^{n-1} + \cdots + a_0$$

$$\frac{\partial P_n}{\partial a_j} = x_i^j \qquad j = 0, \ldots, n$$

jth equation.

$$0 = \frac{\partial E}{\partial a_j} = \sum_{i=1}^{m} 2\left(y_i - P_n(\vec{a}, x_i)\right) w_i x_i^j$$

$$\sum_{k=0}^{n} a_k \left(\sum_{i=1}^{m} x_i^{j+k} w_i \right) = \sum_{i=1}^{m} w_i y_i x_i^j$$

$$\begin{array}{l} j=0 \longrightarrow \\ j=1 \rightarrow \\ \\ \end{array} \left\{ \begin{array}{l} a_0 \left(\sum_{i=1}^{m} w_i x_i^0 \right) + a_1 \left(\sum_{i=1}^{m} w_i x_i^1 \right) + \cdots + a_n \left(\sum_{i=1}^{m} w_i x_i^n \right) \\ a_0 \left(\sum_{i=1}^{m} w_i x_i^1 \right) + a_1 \left(\sum_{i=1}^{m} w_i x_i^2 \right) + \cdots + a_n \left(\sum_{i=1}^{m} w_i x_i^{n+1} \right) \\ a_0 \ldots \end{array} \right.$$

This is called the normal equation for the unknown coefficient vector $\vec{a}$

$$\begin{pmatrix} \sum_{i=0}^{m} w_i x_i^0 & \sum_{i=1}^{m} w_i x_i^1 & \cdots & \sum_{i=1}^{m} w_i x_i^n \\ \sum_{i=1}^{m} w_i x_i^1 & \sum_{i=1}^{m} w_i x_i^2 & \cdots & \sum_{i=1}^{m} w_i x_i^n \\ \vdots & & & \\ \sum_{i=1}^{m} w_i x_i^n & \sum_{i=1}^{m} w_i x_i^{n+1} & \cdots & \sum_{i=1}^{m} w_i x_i^{2n} \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix} = \text{RHS}$$

$$\ldots^{n}) = \sum_{i=1}^{m} w_i y_i x_i^0$$

$$\ldots^{+1}) = \sum_{i=1}^{m} w_i y_i x_i^1$$

Example: L^2-polynomial fitting on $[0,1]$

Given $f: [0,1] \longmapsto \mathbb{R}$

Given $n > 0$

Find $a_0, \dots, a_n \in \mathbb{R}$

$$P_n(x) = a_n x^n + \cdots + a_0$$

to $\underset{\vec{a} \in \mathbb{R}^{n+1}}{\text{minimize}}\ E = \int_0^1 (f(x) - P_n(x))^2 \, dx$

$$E(\vec{a}) = \int_0^1 f(x)^2 \, dx - 2 \int_0^1 (a_n x^n + \cdots + a_0) f(x) \, dx$$

$$+ \int_0^1 \left(\sum_{j=0}^{n} a_j x^j \right) \left(\sum_{k=0}^{n} a_k x^k \right) dx$$

(Ans: Hilbert matrix: $H_{ij} = \frac{1}{i+j-1}$)

Remark

Section 8.2

$1, x, \cdots x^n$

$\langle x^{i-1}, x^{j-1} \rangle = \frac{1}{i+j-1} = H_{ij}$

ill conditioned.

Instead, apply Gram-Schmidt Process

to $1, x, \cdots x^n$

$q_0, q_1(x) \cdots q_n(x)$

$q_j(x) \in \text{Span}\{1, x \cdots x^j\}$

$\langle q_i, q_j \rangle = \delta_{ij} \quad i \neq j$

$\min_{\vec{a} \in \mathbb{R}^n} \int (f(x) - P_n(x))^2 dx = \min_{\vec{b} \in \mathbb{R}^n} \int_a \left(f(x) - (b_0 q_0 \cdots + b_n q_n(x))\right)^2$

Ans: $b_j = \langle f, q_j \rangle$ (No linear systems, stable)