Minimization and Least Squares Approximation

Suppose we wish to solve a system of equations:

$$f_1(\vec{x}) = 0, \quad f_2(\vec{x}) = 0, \quad \dots, \quad f_m(\vec{x}) = 0, \qquad (1)$$

for $\vec{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$. We convert this system into a minimization problem by considering the function

$$p(\vec{x}) = [f_1(\vec{x})]^2 + \dots + [f_m(\vec{x})]^2.$$

Since $p(\vec{x}) \geq 0$, the minimum values of p are heuristically indicating where solutions of (1) are. Specifically, if $P(\vec{x}^*) = 0$ then $\vec{x}^*$ is a solution of (1).

For linear systems

$$A\vec{x} = \vec{b}$$

the function $p(\vec{x})$ can be written in a concise form

$$p(\vec{x}) = \|A\vec{x} - \vec{b}\|_2^2. \quad \longleftarrow \text{standard Euclidean 2-norm.}$$

Why consider this extra construction since we already know how to solve $A\vec{x} = \vec{b}$?

If we have more equations than unknowns we may not have an exact solution of $A\vec{x} = \vec{b}$.

This is when we turn to finding an approximate solution that minimizes $\rho(\vec{x}) = \| A\vec{x} - \vec{b} \|_2^2$.

Ex Consider the problem of finding the "best" linear function that approximates the data:

$$\{ (0,1), (1,2), (2,1) \}.$$

We look for a function $f(x) = ax + b$ such that

$$\begin{aligned} f(0) &= 1 = 0 + b \\ f(1) &= 2 = a + b \\ f(2) &= 1 = 2a + b. \end{aligned}$$

The associated minimization problem is to minimize

$$\rho(a,b) = (b-1)^2 + (a+b-2)^2 + (2a+b-1)^2.$$

Even in this case, it is highly non-trivial to find the solutions of this problem. We will develop Linear algebraic tools to solve these problems in a general way.

Minimization of Quadratic Functions

- The Problem:

Find vectors $\vec{x}^*$ (hopefully unique) such that the multivariate polynomial

$$p(\vec{x}) = \sum_{i,j=1}^{n} k_{ij} x_i x_j - 2 \sum_{i=1}^{n} f_i x_i + C$$

Obtains its minimum value at $\vec{x}^*$.

- Here k_{ij}, f_i, c are all real-valued constants.

- Claim: We may assume $k_{ij} = k_{ji}$.

Why? When $n = 2$

$$p(\vec{x}) = k_{11} x_1^2 + k_{22} x_2^2 + k_{12} x_1 x_2 + k_{21} x_2 x_1 + \ldots$$

define
$$\hat{k}_{11} = k_{11}$$
$$\hat{k}_{22} = k_{22}$$
$$\hat{k}_{12} = \frac{k_{12} + k_{21}}{2}$$
$$\hat{k}_{21} = \frac{k_{12} + k_{21}}{2}$$

- then

equal.

$$p(\vec{x}) = \hat{k}_{11} x_1^2 + \hat{k}_{22} x_2^2 + \hat{k}_{12} x_1 x_2 + \hat{k}_{21} x_2 x_1 + \ldots$$

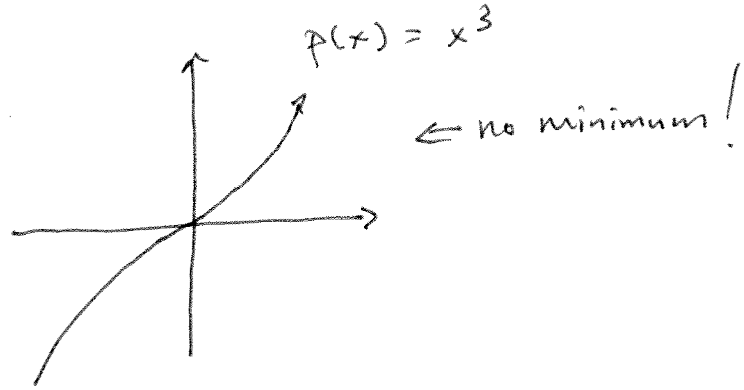
Another claim:

$$p(\vec{x}) = \vec{x}^T K \vec{x} - 2\vec{x}^T \vec{f} + c \qquad (*)$$

$$K = \begin{pmatrix} k_{11} & k_{12} & \cdots \\ k_{21} & \ddots & \\ \vdots & & \ddots \end{pmatrix}, \quad \vec{f} = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

The big questions:

When do we expect to have a minimum?



What is this minimum?

Theorem: If K is a symmetric, positive definite matrix then $(*)$ has a unique minimizer $\vec{x}^*$:

$$\vec{x}^* = K^{-1}\vec{f}$$

and

$$p(\vec{x}^*) = c - f^T K^{-1} f.$$

Proof: First, since K is positive definite, it is non-singular so that $\vec{x}^* = K^{-1} f$ exists.

Rewrite the polynomial:

$$p(\vec{x}) = \vec{x}^T K \vec{x} - 2\vec{x}^T \underset{\substack{\uparrow \\ K\vec{x}^*}}{f} + c$$

$$= \vec{x}^T K \vec{x} - 2\vec{x}^T K \vec{x}^* + c$$

$\vec{x}^T K \vec{x}^*$ is a scalar so that

$$\vec{x}^T K \vec{x}^* = \left(\vec{x}^T K \vec{x}^*\right)^T = \vec{x}^{*T} K^T \vec{x} = \vec{x}^{*T} K \vec{x}.$$

Thus

$$p(\vec{x}) = \vec{x}^T K \vec{x} - \vec{x}^T K \vec{x}^* - \vec{x}^{*T} K \vec{x} + c$$

$$= \vec{x}^T K(\vec{x} - \vec{x}^*) - \vec{x}^{*T} K(\vec{x} - \vec{x}^*) - \vec{x}^{*T} K \vec{x}^* + c$$

add and subtract $\vec{x}^{*T} K \vec{x}^*$

$$p(\vec{x}) = (\vec{x} - \vec{x}^*)^T K (\vec{x} - x^*) - \vec{x}^{*T} K x^* + c.$$

K is positive definite: $\vec{y}^T K \vec{y} > 0$ if $y \neq 0$.

Minimize p:

set $\vec{x} = \vec{x}^*$

Remark: From the point-of-view of multi-variable calculus, this problem can also be solved. This involves finding critical points (gradient being zero), and analyzing the Hessian matrix (matrix of second-order partials.)

Example:

Minimize the function:

$$p(x,y,z) = x^2 + 2xy + xz + 2y^2 + yz + 2z^2$$
$$+6y - 7z + 5.$$

Claim:

$$p(x,y,z) = (x,y,z)\begin{pmatrix} 1 & 1 & 1/2 \\ 1 & 2 & 1/2 \\ 1/2 & 1/2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} - 2(x,y,z)\begin{pmatrix} 0 \\ -3 \\ 7/2 \end{pmatrix} + 5$$

We need to solve $K\vec{x}^* = f$ ... if K is pos def.

We can check that along the way.

$$\left(\begin{array}{ccc|c} 1 & 1 & 1/2 & 0 \\ 1 & 2 & 1/2 & -3 \\ 1/2 & 1/2 & 2 & 7/2 \end{array}\right)$$

$$\rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1/2 & 0 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 7/4 & 7/2 \end{array}\right)$$

→ no row swapping
positive pivots
symmetric
K is pos def.

Back substitution gives the unique minimizer:

$$\vec{x}^* = \begin{pmatrix} 2 \\ -3 \\ 2 \end{pmatrix}.$$

<u>The general Theorem:</u>

- If K is positive definite then the quadratic

$$p(\vec{x}) = \vec{x}^T K \vec{x} - 2\vec{x}^T \vec{f} + c$$

has a unique minimizer satisfying

$$K\vec{x}^* = F.$$

$\vec{y}^T K \vec{y} \geq 0$ for all $\vec{y}$.

- If K is just positive semi-definite and $\vec{f}$ is in the range of K then any solution of

$$K\vec{x} = f$$ is a minimizer of $p(\vec{x})$.

Least Squares Solutions

Definition: A least squares solution to a linear system of equations

$$A\vec{x} = \vec{b}, \quad (A \text{ is } m \times n)$$

is a vector $\vec{x}^* \in \mathbb{R}^n$ that minimizes the Euclidean norm.

Remarks: If the system has a solution then it is automatically a least squares solution.

Claim: If A has rank n then $A^T A$ is invertible. (non-singular).

Proof: Consider

$$\vec{x}^T A^T A \vec{x} = A\vec{x} \underset{\text{dot product.}}{\cdot} A\vec{x} = \|A\vec{x}\|_2^2$$

The fundamental theorem of linear algebra says that

$$\dim \ker A = n - \operatorname{rank} A = 0.$$

Thus

$$\vec{x}^T A^T A \vec{x} = \|A\vec{x}\|_2^2 > 0 \quad \text{if } x \neq 0$$

We conclude that $A^T A$ is positive definite,

All positive definite matrices are invertible. $\blacksquare$

We want to find a vector $\vec{x}$ that minimizes

$$\| A\vec{x} - \vec{b} \|_2 .$$

It is equivalent to minimize

$$p(\vec{x}) = \| A\vec{x} - \vec{b} \|_2^2 .$$

We expand this using the dot product:

$$p(\vec{x}) = (A\vec{x} - \vec{b})^T \cdot (A\vec{x} - \vec{b})$$

$$= \vec{x} A^T A \vec{x} - \vec{x}^T A^T b - \vec{b}^T A \vec{x} + \vec{b}^T b$$

$\vec{x}^T A^T \cdot \vec{b} = \vec{b}^T A \vec{x}$

$$= \vec{x} \underbrace{A^T A}_{K} \vec{x} - 2 \vec{x}^T \underbrace{A^T b}_{\vec{f}} + \underbrace{\vec{b}^T \vec{b}}_{c}$$

We know that minimizers of $p(\vec{x})$ are solutions of

$$K \vec{x}^* = \vec{f} \iff A^T A \vec{x}^* = A^T \vec{b} .$$

normal equations

Now, if A has rank n. (full rank) then

$$\vec{x}^* = (A^T A)^{-1} A^T \vec{b} .$$

is the unique minimizer.

Ex

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 5 \end{pmatrix}$$

Find all least squares solutions to

$$A\vec{x} = \vec{b}.$$

• Step 1: Check the rank of A.

One step of row reduction:

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & -7 & 1 \\ 0 & 4 & 1 \\ 0 & -3 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

At this point we can see there are 3 pivots.

Rank of $A = 3$, A is 5×3, $\Rightarrow$ A is fullrank

or.

rank $A = \min\{5, 3\}$.

This implies the least squares problem will have a unique solution.

- **Step 2**: Compute $A^T A$ and $A^T \vec{b}$

- $A^T = \begin{pmatrix} 1 & 3 & -1 & 1 & 0 \\ 2 & -1 & 2 & -1 & 0 \\ 0 & 1 & 1 & -2 & 1 \end{pmatrix}$, $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & -1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$

$$A^T A = \begin{pmatrix} 12 & -4 & 0 \\ -4 & 10 & 3 \\ 0 & 3 & 7 \end{pmatrix}$$

- $A^T \vec{b} = \begin{pmatrix} 0 \\ 3 \\ 7 \end{pmatrix}$

Step 3: Solve $A^T A \vec{x} = A^T b$.

$$\vec{x}^* = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Let's check the error

$$\| A\vec{x}^* - \vec{b} \|^2 = 20.$$

This relatively large. It indicates that the system is far from being compatible.

A geometric interpretation of least-squares error.

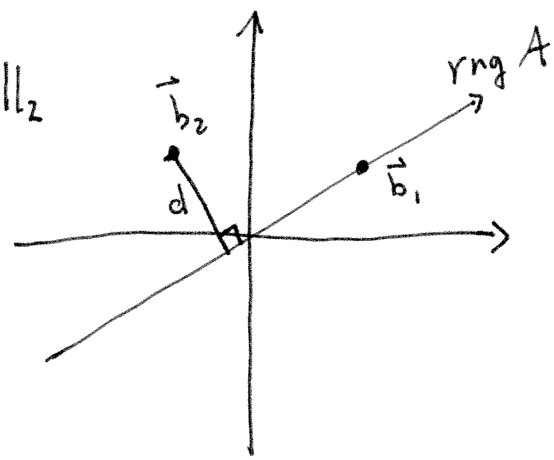
- Assume A is a 2×2 matrix with a one-dimensional kernel.

FTA says

$$\dim \operatorname{rng} A = 2 - \dim \ker A. = 1.$$

the rng A is a subspace of $\mathbb{R}^2$:
must be a line through the origin

- $d = \min_{\vec{x}} \| A\vec{x} - \vec{b}_2 \|_2$



Since $\vec{b}_1 \in \operatorname{rng} A$, $A\vec{x} = \vec{b}_1$ has solutions, How many?

Since $\vec{b}_2 \notin \operatorname{rng} A$, $A\vec{x} = \vec{b}_2$ has only least squares solutions.

It turns out that the geometric interpretation of the
- minimum of $\| A\vec{x} - \vec{b}_2 \|_2$ is the distance between

$\vec{b}_2$ and the line that defines the rng A.