September 26, 2024

Announcements

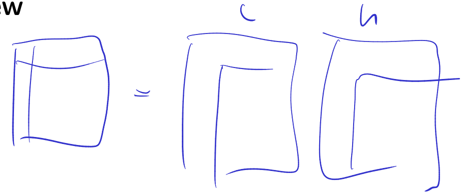
Exam 2: make res!

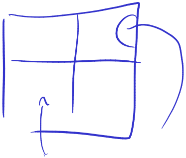
HW4

Goals

LSQ

Review



$A =$ 

$B = {}_{n}[\;]_{k}\;{}_{k}[\;]^{n} = UV^T$

low rank, $n \times n$

↳ k

What about non-square systems?

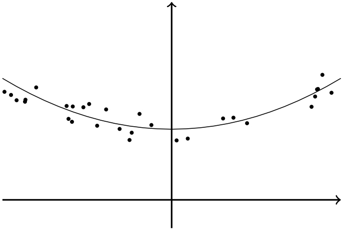
Specifically, what about linear systems with 'tall and skinny' matrices? (A: $m \times n$ with $m > n$) (aka *overdetermined* linear systems)

Specifically, any hope that we will solve those exactly?

only "as good as possible" $\|b - Ax\|_2$

$$A \, x = b$$

Example: Data Fitting



Have data: (x_i, y_i) and model:

$$y(x) = \alpha + \beta x + \gamma x^2$$

Find data that (best) fit model!

Data Fitting Continued

$$(x_1, y_1) \ldots$$

$$\alpha + \beta x_1 + \gamma x_1^2 \approx y_1$$

$$\vdots$$

$$|\alpha + \beta x_1 + \gamma x_1^2 - y_1|^2$$

$$+ |\alpha + \beta x_2 + \gamma x_2^2 - y_2|^2 \to \min!$$

Rewriting Data Fitting

Rewrite in matrix form.

Vandermonde matrix

$$\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ & \vdots & \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \end{pmatrix}$$

Least Squares: The Problem In Matrix Form

$$\|A\boldsymbol{x} - \boldsymbol{b}\|_2^2 \to \text{min!}$$

is cumbersome to write.
Invent new notation, defined to be equivalent:

$$A\boldsymbol{x} \cong \boldsymbol{b}$$

NOTE:

- Data Fitting is *one example* where LSQ problems arise.
- Many other application lead to $A\boldsymbol{x} \cong \boldsymbol{b}$, with different matrices.

Linear LSQ

Data Fitting: Nonlinearity

Give an example of a nonlinear data fitting problem.

$$\begin{aligned}&\left|\exp(\alpha) + \beta x_1 + \gamma x_1^2 - y_1\right|^2 \\ &+ \cdots + \\ &\left|\exp(\alpha) + \beta x_n + \gamma x_n^2 - y_n\right|^2 \quad \to \quad \text{min!}\end{aligned}$$

But that would be easy to remedy: Do linear least squares with $\exp(\alpha)$ as the unknown. More difficult:

$$\begin{aligned}&\left|\alpha + \exp(\beta x_1 + \gamma x_1^2) - y_1\right|^2 \\ &+ \cdots + \\ &\left|\alpha + \exp(\beta x_n + \gamma x_n^2) - y_n\right|^2 \quad \to \quad \text{min!}\end{aligned}$$

NLSQ

Demo: Interactive Polynomial Fit [cleared]

Properties of Least-Squares

$Ax = b$

Consider LSQ problem $\boldsymbol{Ax} \cong \boldsymbol{b}$ and its associated *objective function* $\varphi(x) = \|\boldsymbol{b} - \boldsymbol{Ax}\|_2^2$. Assume A has full rank. Does this always have a solution?

calc

$\varphi \to \infty$ as $\|x\| \to \infty$ $\Rightarrow$ φ has a minimum

Is it always unique?

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

Yes.

What happens if A does not have full rank?

$x + \alpha z$

Suppos $\|\vec{b} - A\vec{x}\| = \min$ Suppose $\vec{z} \in N(A) \setminus \{0\}$ $A\vec{z} = \vec{0}$

$\|b - A(x + \alpha z)\|_2 = \|b - A\vec{x}\|_2$

Least-Squares: Finding a Solution by Minimization

n

B spd $\iff$ $\vec{x}^T B \vec{x} > 0$ if $\vec{x} \neq 0$

Examine the objective function, find its minimum.

SPD

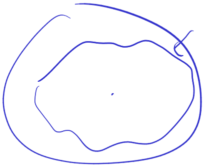
$$\varphi(\vec{x}) = (\vec{b} - A\vec{x})^T(\vec{b} - A\vec{x}) = b^T b - b^T A\vec{x} - x^T A^T b + x^T A^T A x$$

$$\nabla\varphi(\vec{x}) = -2A^T b + 2A^T A\vec{x}$$

$2b^T A\vec{x}$

To find crit. points: $\nabla\varphi(\vec{x}) = 0 \iff A^T A\vec{x} = A^T\vec{b}$

$\{\varphi \le c\}$



Hessian spd $\Rightarrow$ min.

normal equations

Least squares: Demos

Demo: Polynomial fitting with the normal equations [cleared]

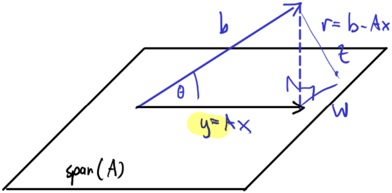
What's the shape of $A^T A$?

$k \times k$

Demo: Issues with the normal equations [cleared]

Least Squares, Viewed Geometrically

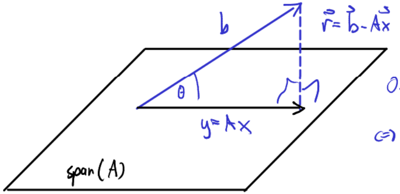
$\|r\|_2 = \|b - Ax\|_2$



Why is $\boldsymbol{r} \perp \operatorname{span}(A)$ a good thing to require?

$$\|z\|_2^2 = \|r\|_2^2 + \underbrace{\|w\|_2^2}_{>0} > \|r\|_2^2$$

Least Squares, Viewed Geometrically (II)



$0 \le \|b - Ax\|_2 \leftarrow \min$

$\Leftrightarrow \|b - Ax\|_2^2 \leftarrow \min$

Phrase the Pythagoras observation as an equation.

$$\operatorname{span}(A) \perp b - A\vec{x}$$

$$A^T(\vec{b} - A\vec{x}) = 0 \Leftrightarrow A^T\vec{b} = A^TA\vec{x}$$

Write that with an orthogonal projection matrix P. (onto span(A))

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

About Orthogonal Projectors

What is a *projector*?

$P^2 = P$

What is an *orthogonal projector*?

P symmetric

How do I make one projecting onto span$\{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_\ell\}$ for orthonormal $\boldsymbol{q}_i$?

$Q = \begin{pmatrix} \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_\ell \end{pmatrix}$

$P = QQ^T$

$(Q^T x)$ basis coeffs

$Q(Q^T x)$

Least Squares and Orthogonal Projection

Check that $P = A(A^T A)^{-1} A^T$ is an orthogonal projector onto colspan(A).

What assumptions do we need to define the P from the last question?