SFU
muraki

WEDNESDAY
- quiz friday
- No computing workshops FRI

LAST DAY
- the REAL (discrete) LEAST-SQUARES
- linear solve with "MORE EQUATIONS" than unknowns

$$[A]_{M\times N}\,\vec{v}_{N\times 1} = \vec{y}_{M\times 1}$$

- for $M \geq \text{rank}[A] \geq N$, the vector $\vec{v}_{LSQ}$ that satisfies

$$[A^TA]_{N\times N}\,\underbrace{\vec{v}_{LSQ}}_{N\times 1} = \underbrace{[A^T]_{N\times M}\,\vec{y}_{M\times 1}}_{N\times 1}$$

is the UNIQUE minimizer for

$$\mathcal{E}(\vec{v}) = \left|[A]\vec{v}-\vec{y}\right|^2 = \text{sum of squares error}$$

$\rightarrow$ hence, $\vec{v}_{LSQ}$ is the "least-squares" solution

- Matlab's backslash gives L-SQ solution!

UNDERSTANDING RANK

- return to linear algebra of equations

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$ has unique solution $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

equivalent to $\{(1,3),(2,4)\}$

+ LSQ line $y = 1 \cdot x + 2$

- set of equations

$$\left.\begin{aligned} &1)\quad a + b = 3 \\ &2)\quad 2a + b = 4 \\ &1)+2) \to 3)\quad 3a + 2b = 7 \end{aligned}\right\}$$ $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is still a solution

3 equations in 2 unknowns? YES.

- but ONLY 2 LINEARLY-INDEPENDENT equations in 2 unknowns

- definition: linear dependence for a set of arithmetic objects (vectors, equations, functions...) means there is some non-zero linear combination that is zero $\sum_1^N c_j \cdot (\text{thing})_j = 0$

if not, then objects are linearly independent

LECT 07W NOTES.

- so we see that the correct thing to count is the # of <u>linearly independent</u> equations

- MATRIX RANK counts something slightly different

$$[A]\vec{v} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 2 \end{bmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 7 \end{pmatrix}$$

what if this NOT 7? (3 indep eq'ns)

- define row & col rank:

  ROW RANK = # of linearly indep rows $\leq M$
  COL " = " " " " cols $\leq N$

- RANK THEOREM: for every matrix,

  row rank = column rank (p360)

- theorem useful for Normal Eq'ns (p365)

$$\text{rank of } [A]_{M \times N} = \text{rank of } [A^T A]_{N \times N}$$

- when rank $[A] = N$, the $[A^T A]$ is invertible

BACK to REAL L-SQ.

- the situation of the OVERDETERMINED linear solve

$$[A]_{M\times N}\,\vec{v} = \vec{y}$$

where $M > N$ (more equations than unknowns) is generally NOT solvable exactly, but has a L-Sq. solution $\vec{v}_{LSQ}$.

- however, the mathematically precise statement needs to be that.

$[A]_{M\times N}$ with $M \geq \operatorname{rank}[A] = N$

THEN $[A^TA]\vec{v} = [A^T]\vec{y}$ is guaranteed to generate a L-Sq solution $\vec{v}_{LSQ}$.

(and if $\vec{y}$ just happens to be very special, then it is NOT impossible that $\vec{e} = \vec{0}$)

- the lin alg thy tells us that L-Sq for an overdetermined system FAILS when. $M \geq N > \operatorname{rank}[A]$

- WHAT GOES WRONG when $M \geq N > \text{rank}\,[A]$. ?
  - there is some linear combination of columns of $[A]$ that gives zero vector
  - as in

$$[A]\vec{w} = \sum_{1}^{N} w_j \vec{a}_j = \vec{0}$$

j^{th} column of $[A]$!

homogeneous solution $\Rightarrow$ NO UNIQUE $\vec{V}_{LSQ}$

BACK TO "BEST-FIT" L-SQ

- introduce set of basis functions $\{f_1(x) \ldots f_N(x)\}$ and they must be linearly independent.

$$\{1, x, 3x-4\} + \{1, \sin^2 x, \cos^2 x\}$$

are NOT lin indep sets

- the best-fit linear combination: $y(x) = \sum_{1}^{N} c_j f_j(x)$

$$[A]\vec{c} = \begin{bmatrix} f_1(x_1) & \cdots & f_N(x_1) \\ \vdots & & \vdots \\ f_1(x_M) & \cdots & f_N(x_M) \end{bmatrix}_{M\times N} \vec{c} = \begin{pmatrix} y_1 \\ \vdots \\ y_M \end{pmatrix}$$

· finally solving Normal Equation

$$[A^T A]\vec{c} = [A^T]\vec{y}$$

$$\underbrace{\phantom{[A^TA]\vec{c}}}_{\text{gives } \vec{c} \text{ + best-fit } y(x)}$$

THINK UPSCALING...

· accuracy: it's not possible to address how well $y(x)$ "matches" $\{y_k\}$ without knowing more about the $\{y_k\}$ points

· but can address linear solve of $\vec{c}$ from $\vec{y}$

· can show condition number of $A^T A$.

$$K[A^T A] = (K[A])^2$$

· so if $K[A] \simeq 10^8$, $K[A^T A] \simeq 10^{16}$ ← all digits lost

and $K[A] \simeq 10^4$ means only 8-sig digits are reliable

· since $[A^T A]$ is determined by $\{f_j(x)\}_{j=1 \to N}$ ...

- polynomial L-SQ for large N
  - NOT surprisingly, monomial basis is NOT good choice
  - for exact interpolation, we used Lagrange. – for L-SQ, use Legendre polynomials (p523) for $-1 \le x_k \le +1$

$$P_0(x) = 1, \quad P_1(x) = x$$

$$P_2(x) = x^3 - \frac{3}{5}x, \quad P_3(x) = x^4 - \frac{6}{7}x^2 + \frac{3}{35}$$

- WHY these? $\int_{-1}^{1} P_j(x) P_k(x)\, dx = 0$ for $j \neq k$

LEGENDRE DEMO.
- From Prof Adcock, test using $f(x) = \frac{1}{1+16x^2}$

Demo

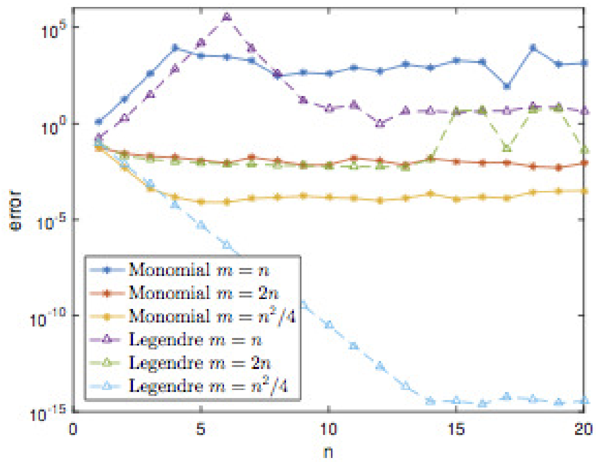
Demo: **polyLS.m** (needs **LegMat.m**)

This demo compares polynomial least-squares fitting using:

- equally-spaced nodes in $[-1, 1]$,
- the monomial and Legendre bases,
- the scalings $m = n$, $m = 2n$ and $m = n^2/4$.

The least-squares problem is solved using backslash (see later).

Demo



1) Both methods are not at all robust when $m = n$ (interpolation).

2) Both methods are slightly more robust for $m = 2n$.

3) When $m = n^2/4$ the Legendre basis is robust, achieving close to machine epsilon for large n. However, the monomial basis is still not robust in this case.