- HW grading
- Examlet 1 grades
- DO NOT USE $10^{\wedge}(-3)$ in Py

  DO USE $10**(-3)$

LU: Special cases

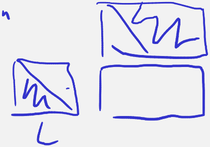
What happens if we feed a non-invertible matrix to LU?

$$\underset{\text{inv}}{P}\ \underset{\text{not}}{A} = \underset{\text{inv}}{L}\ \underset{\text{not}}{U}$$

What happens if we feed LU an $m \times n$ non-square matrices?

$$\underset{m \times n}{A} = \underset{m \times k}{L}\ \underset{k \times n}{U}$$

			L	U
- short & fat:	$m < n$		$m \times m$	$m \times n$
- tall & skinny	$n < m$		$m \times n$	$n \times n$



"TSQR"

Round-off Error in LU without Pivoting

Consider factorization of $\begin{bmatrix} \epsilon & 1 \\ 1 & 1 \end{bmatrix}$ where $\epsilon < \epsilon_{\text{mach}}$:

$$L = \begin{pmatrix} 1 & \\ 1/\varepsilon & 1 \end{pmatrix} \qquad U = \begin{pmatrix} \varepsilon & 1 \\ & 1 - 1/\varepsilon \end{pmatrix}$$

$$fl(U) = \begin{pmatrix} \varepsilon & 1 \\ & -1/\varepsilon \end{pmatrix}$$

$$A \stackrel{!}{=} L \cdot fl(U) = \begin{pmatrix} \varepsilon & 1 \\ 1 & 0 \end{pmatrix} \rightsquigarrow \text{bw error} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Round-off Error in LU with Pivoting

Permuting the rows of A in partial pivoting gives $PA = \begin{bmatrix} 1 & 1 \\ \epsilon & 1 \end{bmatrix}$

$$L = \begin{pmatrix} 1 & \\ \varepsilon & 1 \end{pmatrix} \qquad U = \begin{pmatrix} 1 & 1 \\ & 1-\varepsilon \end{pmatrix}$$

$$fl(U) = \begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}$$

$$A \stackrel{!}{=} L \cdot fl(U) = \begin{pmatrix} 1 & \\ \varepsilon & 1+\varepsilon \end{pmatrix} \rightarrow \text{bw err:} \begin{pmatrix} 0 & 0 \\ 0 & \varepsilon \end{pmatrix}$$

Changing matrices

Seen: LU cheap to re-solve if RHS changes. (Able to keep the expensive bit, the LU factorization) What if the *matrix* changes?

$PA = LU$ $\quad A\vec{x} = \vec{b}$ $\quad A x = \vec{c}$; cheap

$(A + ?)\vec{x} = \vec{b}$ $\quad \rightsquigarrow$ "if you have a 'cheap' solver for A, can you turn that into a solver for $A + ?$"?

$\hat{A} = A + \vec{u}\vec{v}^T$

$$\hat{A}^{-1} = A^{-1} - \frac{A^{-1}\vec{u}\vec{v}^T A^{-1}}{1 + \vec{v}^T A^{-1}\vec{u}}$$

$$\hat{A}^{-1}\vec{b} = \underbrace{A^{-1}b}_{n^2} - \frac{(A^{-1}u)(\vec{v}^T(A^{-1}\vec{b}))}{1 + \vec{v}^T(A^{-1}\vec{u})} \Big\} n^2$$

Demo: Sherman-Morrison [cleared]

Outline

What about non-square systems?

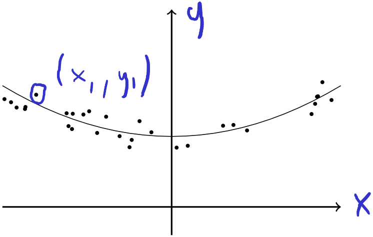
$A\vec{x} = \vec{b}$

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?

more equations than unknowns: overdet.

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

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

$$\vdots$$

$$\alpha + \beta x_n + \gamma x_n^2 = y_n$$

$$\left\| \begin{matrix} \alpha \cdot 1 + \beta x_1 + \gamma x_1^2 - y_1 \\ \vdots \\ \alpha \cdot 1 + \beta x_n + \gamma x_n^2 - y_n \end{matrix} \right\|_2 \to \min!$$

Rewriting Data Fitting

Rewrite in matrix form.

$$\|A\vec{x} - \vec{b}\|_2 \to \min$$

$$A = \underbrace{\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & & \\ 1 & x_n & x_n^2 \end{pmatrix}}_{\text{Vandermonde}} \qquad \vec{x} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \qquad \vec{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

x_i^0

Least Squares: The Problem In Matrix Form

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

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.

Data Fitting: Nonlinearity

Give an example of a nonlinear data fitting problem.

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

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

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

Demo: Interactive Polynomial Fit [cleared]

Properties of Least-Squares

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

$\varphi \geq 0$, $\varphi \to \infty$ as $\|x\| \to \infty$, φ continuous

$\to \varphi$ has a minimum

Is it always unique?

Not if A has a null space

If it doesn't, then yes.

Least-Squares: Finding a Solution by Minimization

$$\|\vec{v}\|_2^2 = v_1^2 + \cdots v_n^2 = \vec{v}^T\vec{v}$$

Examine the objective function, find its minimum.

$$\varphi(\vec{x}) = \|Ax - b\|_2^2 = (\vec{b} - A\vec{x})^T(\vec{b} - A\vec{x})$$

$$= \vec{b}^T\vec{b} - 2\vec{x}^T A^T\vec{b} + \vec{x}^T A^T A\vec{x}$$

$$0 \stackrel{!}{=} \nabla_x \varphi(\vec{x}) = -2A^T b - 2A^T A\vec{x}$$

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

↑ normal eqns.

Least squares: Demos

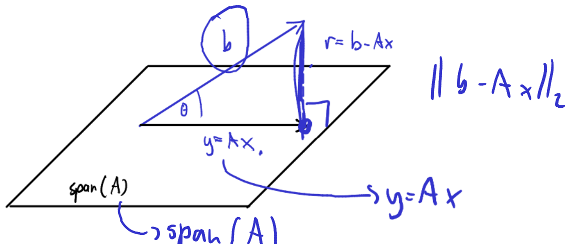
Demo: Polynomial fitting with the normal equations [cleared]

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

square

Demo: Issues with the normal equations [cleared]

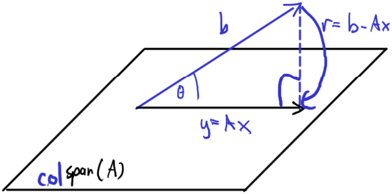
Least Squares, Viewed Geometrically



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

Pythagoras: any other y would increase $\|Ax-b\|_2$

Least Squares, Viewed Geometrically (II)



Phrase the Pythagoras observation as an equation.

$$\mathrm{span}(A) \perp b - Ax$$

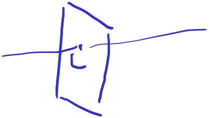
$$\vec{0} = A^T(b - Ax) = A^Tb - A^TAx$$

normal eqns again!

Write that with an orthogonal projection matrix P.

$$A\vec{x} = P\vec{b}$$ where P is the orth. projector onto colspan(A)

About Orthogonal Projectors



LA recap

What is a *projector*?

$$P^2 = P$$

What is an *orthogonal projector*?

$\Leftrightarrow$ P symmetric (try proving)

How do I make one projecting onto span$\{\boldsymbol{q}_1, \boldsymbol{q}_2, \dots, \boldsymbol{q}_\ell\}$ for ortho~~gonal~~normal $\boldsymbol{q}_i$?

$$Q = \begin{bmatrix} q_1 & q_2 & q_3 & \cdots & q_\ell \end{bmatrix}$$

$$\|q_i\|_2 = 1$$

$$QQ^T$$

Least Squares and Orthogonal Projection

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

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

$x = (A^T A)^{-1} A^T b$

$y = Ax = Pb$

$P^2 = P$; ✓

P symmetric ✓ $P = P^T$

$P\vec{b}$

$A\underbrace{(\ldots)\vec{b}}$ $\in$ colspan(A)

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

$A^T A$ has full rank ($\Leftrightarrow$ invertible)

Pseudoinverse

What is the pseudoinverse of A?

What can we say about the condition number in the case of a tall-and-skinny, full-rank matrix?

What does all this have to do with solving least squares problems?