Today:
- Linear least sq.

Announcements:
- HW2 due
- Exam 1 next week
- Grading weights
- [E++]
What about non-square systems?

Specifically, what about linear systems with ‘tall and skinny’ matrices? (A: $m \times n$ with $m > n$) (aka overetermined linear systems)

Specifically, any hope that we will solve those exactly?
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
Rewriting Least Squares

Rewrite in matrix form.

ʻ||Ax − b||^2_2 → min!ʼ is cumbersome to write → new notation, defined to be equivalent:

\[ Ax \approx b \]
Least Squares: Nonlinearity

Q: Give an example of a nonlinear least squares problem.

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

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

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

Demo: Interactive Polynomial Fit
Properties of Least-Squares

Consider LSQ problem \( Ax \cong b \) and its associated objective function \( \varphi(x) = \|b - Ax\|_2^2 \). Does this always have a solution?

Is it always unique?

Examine the objective function, find its minimum.
Least squares: Demos

**Demo:** Polynomial fitting with the normal equations

What’s the shape of $A^T A$?

Square

**Demo:** Issues with the normal equations
Least Squares, Viewed Geometrically

Why is $\mathbf{r} \perp \text{span}(A)$ a good thing to require?
Least Squares, Viewed Geometrically (II)

Phrase the Pythagoras observation as an equation.

\[ \text{span}(A) \perp b - Ax \]

\[ A^*b - A^*Ax = 0 \]

Write that with an orthogonal projection matrix \( P \).

\[ Ax = Pb \]
About Orthogonal Projectors

What is a projector?

\[ p^2 = p \]

What is an orthogonal projector?

iff \( P \) symmetric

How do I make one projecting onto span\( \{q_1, q_2, \ldots, q_\ell\} \) for orthogonal \( q_i \)?

\[ Q \approx \begin{pmatrix} q_1 & \cdots & q_\ell \end{pmatrix} \begin{pmatrix} q_1 \cdots q_\ell \end{pmatrix} x \]
Least Squares and Orthogonal Projection

\[ \mathbf{A}^\dagger \mathbf{A} \mathbf{x} = \mathbf{A}^\dagger \mathbf{b} \]

\[ \mathbf{x} = (\mathbf{A}^\dagger \mathbf{A})^{-1} \mathbf{A}^\dagger \mathbf{b} \]

Check that \( P = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \) is an orthogonal projector onto cols\( \text{span}(\mathbf{A}) \).

\[
\begin{array}{c}
P = P \\
P \text{ symmetric} \checkmark
\end{array}
\]

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

\[
\mathbf{A}^\dagger \mathbf{A} \text{ needs to be invertible}
\]

\( \iff \mathbf{A} \text{ has full rank} \)
Pseudoinverse

What is the pseudoinverse of $A$?

$$A^+ = (A^+ A)^{-1} A^+$$

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

$$\text{cond}(A) = \|A\| \|A^+\|$$

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

$$x = A^+ b$$
In-Class Activity: Least Squares
Sensitivity and Conditioning of Least Squares

\[
\frac{\|Ax\|}{\|x\|} \leq \frac{1}{\cos \theta} \text{ cond } (A) \cdot \frac{\|b\|}{\|b\|}
\]

What values of \( \theta \) are bad?

\( \theta \) near \( 90^\circ \) is likely not good.
Sensitivity and Conditioning of Least Squares (II)

Any comments regarding dependencies?

\[ \frac{\Delta x}{\| x \|} \leq \left[ \text{cond}(A)^2 \tan(B) + \text{cond}(A) \right] \frac{\| \Delta A \|}{\| A \|}. \]

What about changes in the matrix?
Least-squares by Transformation

Want a matrix $Q$ so that

\[ QAx \approx Qb \]

has the same solution as

\[ Ax \approx b. \]

I.e. want

\[ \| Q(Ax - b) \|_2 = \| Ax - b \|_2. \]

What type of matrix does that? Any invertible one?

\[ Q \text{ orthogonal} \]
Orthogonal Matrices

What’s an orthogonal (=orthonormal) matrix?
One that satisfies $Q^TQ = I$ and $QQ^T = I$.

Are orthogonal projectors orthogonal?
Nope, not in general.

Now what about that norm property?

$$\|Qv\|_2^2 = (Qv)^T(Qv) = v^TQ^TQv = v^Tv = \|v\|_2^2.$$
Simpler Problems: Triangular

Would we win anything from transforming a least-squares system to upper triangular form?

\[
\begin{pmatrix}
\ast \\
0
\end{pmatrix} X \succeq \begin{pmatrix}
\ast \\
0
\end{pmatrix}
\]

If so, how would we minimize the residual norm?

\[A = QR\]
Computing QR

- Gram-Schmidt
- Householder Reflectors
- Givens Rotations

Latter two similar to LU:
- Successively zero out below-diagonal part
- But: using orthogonal matrices

**Demo:** Gram-Schmidt–The Movie
**Demo:** Gram-Schmidt and Modified Gram-Schmidt
**Demo:** Keeping track of coefficients in Gram-Schmidt