Today:

- Linear least sq.

Announcements:

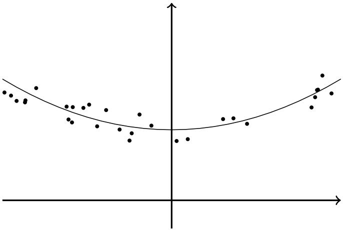
- HW2 due
- Exam 1 next week
- Grading weights
- IE7

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?

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.

'$\|A\mathbf{x} - \mathbf{b}\|_2^2 \to \min!$' is cumbersome to write $\to$ new notation, defined to be equivalent:

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

Least Squares: Nonlinearity

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

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

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

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

Demo: Interactive Polynomial Fit

Properties of Least-Squares

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

Is it always unique?

Examine the objective function, find its minimum.

Least squares: Demos

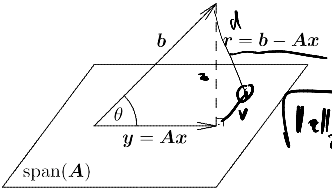
Demo: Polynomial fitting with the normal equations

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

Square

Demo: Issues with the normal equations

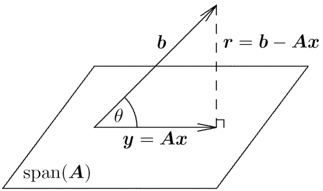
Least Squares, Viewed Geometrically



$\sqrt{\|z\|_2^2 + \|v\|_2} = \|d\|_2^2$

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

Least Squares, Viewed Geometrically (II)



Phrase the Pythagoras observation as an equation.

$$\mathrm{span}(A) \perp b - Ax$$
$$A^T b - A^T A x = 0$$

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$\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_\ell\}$ for orthogonal $\mathbf{q}_i$?

$$Q = (q_1 \cdots q_\ell) \qquad \underline{Q}\, \underline{Q^T x}$$

Least Squares and Orthogonal Projection

$A^TAx = A^Tb$

$x = (A^TA)^{-1}A^Tb$

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

$P^2 = P$ $\quad$ P symmetric $\quad$ ✓

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

A^TA needs to be invertible

$\Leftrightarrow$ A has full rank

Pseudoinverse

$Ax \cong b$ $\rightsquigarrow$ A^+ should $x = A^+ b$

What is the *pseudoinverse* of A?

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

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

square : $\|A\| \|A^{-1}\|$

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

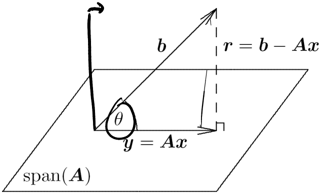
What does all this have to do with solving least squares problems?

$$x = A^+ b$$

In-Class Activity: Least Squares

In-class activity: Least Squares

Sensitivity and Conditioning of Least Squares



$$\frac{\|\Delta x\|}{\|x\|} \le \frac{1}{\cos\theta} \operatorname{cond}(A) \cdot \frac{\|\Delta b\|}{\|b\|}$$

What values of θ are bad?

θ near 90° is likely not good

Sensitivity and Conditioning of Least Squares (II)

Any comments regarding dependencies?

have to have a dep. on b

What about changes in the matrix?

$$\frac{\|\Delta x\|}{\|x\|} \le \left[\underline{\operatorname{cond}(A)^2 \tan(\theta)} + \operatorname{cond}(A) \right] \cdot \frac{\|\Delta A\|}{\|A\|}$$

Least-squares by Transformation

Want a matrix Q so that

$$QA\mathbf{x} \cong Q\mathbf{b}$$

has the same solution as

$$A\mathbf{x} \cong \mathbf{b}.$$

I.e. want

$$\|Q(A\mathbf{x}-\mathbf{b})\|_2 = \|A\mathbf{x}-\mathbf{b}\|_2 \,.$$

What type of matrix does that? Any invertible one?

Q orthogonal

Orthogonal Matrices

What's an *orthogonal (=orthonormal) matrix*?

One that satisfies $Q^T Q = I$ and $QQ^T = I$.

Are orthogonal projectors orthogonal?

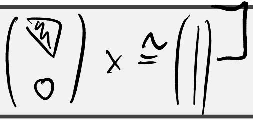
Nope, not in general.

Now what about that norm property?

$$\|Q\mathbf{v}\|_2^2 = (Q\mathbf{v})^T (Q\mathbf{v}) = \mathbf{v}^T Q^T Q\mathbf{v} = \mathbf{v}^T \mathbf{v} = \|\mathbf{v}\|_2^2 .$$

Simpler Problems: Triangular

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



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