Today

- rank-def. lsq
- SVD
- Eigenvalues
  - ↳ existence
  - ↳ sensitivity

Announcements

- Example / mistake hotspots
- Looking at quiz answers

$$\left|\frac{x f'(x)}{f(x)}\right|$$

Rank-Deficient Matrices and Least-Squares

What happens with Least Squares for rank-deficient matrices?

$$A\mathbf{x} \cong \mathbf{b} \quad \to \|Ax - b\|_2 \to \min$$

- QR still finds a solution with minimal residual
- By QR it's easy to see that least squares with a short-and-fat matrix is equivalent to a rank-deficient one.
- But: No longer unique. $\mathbf{x} + \mathbf{n}$ for $\mathbf{n} \in N(A)$ has the same residual.
- In other words: Have more freedom
  Or: Can demand another condition, for example:
  - Minimize $\|\mathbf{b} - A\mathbf{x}\|_2^2$ *and*
  - minimize $\|\mathbf{x}\|_2^2$, simultaneously.
    Unfortunately, QR does not help much with that $\to$ Need better tool.

Singular Value Decomposition (SVD)

What is the *Singular Value Decomposition* of an $m \times n$ matrix?

cols of V; right sing. vec

$$A = U \Sigma V^*$$

QR

cols of U; left singular vect

singular values

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$

"reduced" SVD: shrink to smallest size around nonzeros

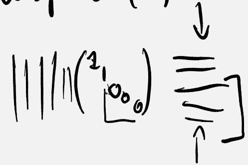
SVD: What's this thing good for? (I)

$$A = U\Sigma V^T$$

$$\|A\|_2 = \|U\Sigma V^T\|_2 = \|\Sigma\|_2 = \sigma_1$$

$$\mathrm{cond}(A) = \sigma_1/\sigma_n$$

nullspace (A)



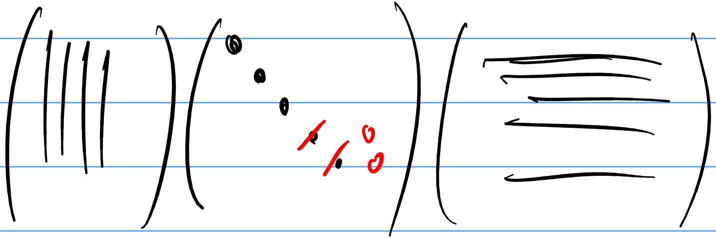
$\mathrm{rank}(A) = \mathrm{rank}(\sigma)$ (easy)

$\mathrm{numrank}(A, \varepsilon)$

$$= \#\{\sigma_i : |\sigma_i| > \varepsilon\}$$

$A = U \Sigma V^T$

To force this to rank 3:

$A =$ 

low-rank approximation

$$A = \begin{pmatrix} & & \\ & U & \\ & & \end{pmatrix} \begin{pmatrix} 3 & & & & & \\ & 2 & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & 0 & \\ & & & & & 0 \end{pmatrix} \underbrace{\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{pmatrix}}_{V^T}$$

$$\text{nullspace}(A) = \text{span}(v_4, v_5, v_6)$$

$$\subseteq$$

$$\supseteq$$

$$A v_5 = U \begin{pmatrix} \cdot & & & & & \\ & \cdot & & & & \\ & & \cdot & & & \\ & & & 0 & & \\ & & & & 0 & \\ & & & & & 0 \end{pmatrix} V^T v_5 = U \begin{pmatrix} \cdot & & & & & \\ & \cdot & & & & \\ & & \cdot & & & \\ & & & 0 & & \\ & & & & 0 & \\ & & & & & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

SVD: What's this thing good for? (II)

- *Low-rank Approximation*

Theorem (Eckart-Young-Mirsky)

If $k < \operatorname{rank}(A) = r$

columns of U (LSV)

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

columns of V (RSV)

then

$$\min_{\operatorname{rank}(B)=k} \|A - B\|_2 = \|A - A_k\|_2$$

Demo: Image compression

SVD: What's this thing good for? (III)

$A = U\Sigma V$

- The minimum norm solution to $A\mathbf{x} \cong \mathbf{b}$:

$\vec{y}$ has the smallest norm of all vectors minimizing $\|Ax - b\|_2$.

$\Leftrightarrow\ U\Sigma V^T x \cong b$

$\Leftrightarrow\ \Sigma \underbrace{V^T x}_{y} \cong U^T b$

$\Leftrightarrow\ \Sigma y \cong U^T b$

$k\left[\begin{pmatrix} \ddots & & \\ & 0 & \\ & & 0 \; 0 \end{pmatrix}\right] y \cong U^T b$

$y = V^T x$

$x = Vy$

$\|x\|_2 = \|y\|_2$

because V orthogonal.

$$\begin{cases} y_i = (U^T b)_i / \sigma_i & (i = 1..k) \\ y_i = 0 & (i = k+1..n) \end{cases}$$

$$\Sigma^{+} = \begin{pmatrix} 1/\sigma_1 & & & & \\ & \ddots & & & \\ & & \frac{1}{\sigma_n} & & \\ & & & 0 & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

write down the solution to total LSQ

$Ax \cong b$ with $A = U\Sigma V^{T}$

$$x = V\Sigma^{+}U^{T}b$$

"la. pinv" Moore-Penrose pseudoinv. of A

SVD: Minimum-Norm, Pseudoinverse

$\mathbf{y} = \Sigma^+ U^T \mathbf{b}$ is the minimum norm-solution to $\Sigma \mathbf{y} \cong U^T \mathbf{b}$.
Observe $\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2$.

$$\mathbf{x} = V \Sigma^+ U^T \mathbf{b}$$

solves the minimum-norm least-squares problem.

Define $A^+ = V \Sigma^+ U^T$ and call it the pseudoinverse of A.
Coincides with prior definition in case of full rank.

In-Class Activity: Householder, Givens, SVD

In-class activity: Householder, Givens, SVD

Comparing the Methods

Methods to solve least squares with A an $m \times n$ matrix:

- Form: $A^T A$: $n^2 m/2$
  Solve with $A^T A$: $n^3/6$
- Solve with Householder: $mn^2 - n^3/3$
- If $m \approx n$, about the same
- If $m \gg n$: Householder QR requires about twice as much work as normal equations
- SVD: $mn^2 + n^3$ (with a large constant)

Demo: Relative cost of matrix factorizations

Outline

Eigenvalue Problems: Setup/Math Recap

R-SVD

la.svd(A)
eigenv.(A^T A)

A is an $n \times n$ matrix.

- $\mathbf{x} \neq \mathbf{0}$ is called an *eigenvector* of A if there exists a λ so that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

- In that case, λ is called an *eigenvalue*.
- The set of all eigenvalues $\lambda(A)$ is called the *spectrum*.
- The *spectral radius* is the magnitude of the biggest eigenvalue:

$$\rho(A) = \max\{|\lambda| : \lambda(A)\}$$

$x^6 + ax^5 + bx^4 + \cdots + = 0$

Finding Eigenvalues

How do you find eigenvalues?

$$A\mathbf{x} = \lambda \mathbf{x} \Leftrightarrow (A - \lambda I)\mathbf{x} = 0$$
$$\Leftrightarrow A - \lambda I \text{ singular} \Leftrightarrow \det(A - \lambda I) = 0$$

$\det(A - \lambda I)$ is called the *characteristic polynomial*, which has degree n, and therefore n (potentially complex) roots.

Does that help algorithmically? Abel showed that for $n \geqslant 5$ there is no general formula for the roots of the polynomial. (i.e. no analog to the quadratic formula for $n = 5$) IOW: no.

Algorithmically, that means we will need to *approximate*. So far (e.g. for LU and QR), if it had not been for FP error, we would have obtained *exact* answers. For eigenvalue problems, that is no longer true–we can only hope for an *approximate* answer.