Today
- rank-def. LSq
- SVD
- Eigenvalues
  - existence
  - sensitivity

Announcements
- Exam #1 mistakes hotspots
- Looking at quiz answers

\[ \frac{\mathbf{x} \cdot \partial f(x)}{\partial f(x)} \]
Rank-Deficient Matrices and Least-Squares

What happens with Least Squares for rank-deficient matrices?

\[ Ax \cong b \quad \Rightarrow \quad \| Ax - b \|_1 \Rightarrow m \]

- 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. \( \hat{x} + n \) for \( n \in N(A) \) has the same residual.
- In other words: Have more freedom

Or: Can demand another condition, for example:
- Minimize \( \| b - Ax \|_2^2 \) and
- minimize \( \| x \|_2^2 \), simultaneously.

Unfortunately, QR does not help much with that \( \Rightarrow \) Need better tool.
Singular Value Decomposition (SVD)

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

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

- $A$: Original matrix
- $U$: Left singular vectors
- $\Sigma$: Diagonal matrix of singular values
- $V^T$: Right singular vectors

- $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{\min(m,n)} > 0$

"Reduced" SVD: shrink to smallest size around non-zeros.
SVD: What’s this thing good for? (I)

\[ A = U \Sigma V^T \]

\[ \|A\|_2 = \| U \Sigma V^T \|_2 = \| \Sigma \|_2 = \sigma, \]

\[ \text{cond}(A) = \sigma_1 / \sigma_c \]

nullspace \((A) \)

\[ \text{rank}(A) = \text{rank}(\sigma) \] (easy)

\[ \text{numrank}(A, \varepsilon) = \# \{ \sigma_i : |\sigma_i| > \varepsilon \} \]
$A = U \Sigma V^T$

To force this to rank 3:

$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \vdots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \vdots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$

Low-rank approximation
\[ A = \begin{pmatrix} U \\ \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{pmatrix} \end{pmatrix} \begin{pmatrix} 3 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} \]

\[ \text{null span } (A) = \text{span } (v_4, v_5, v_6) \]

\[ A v_5 = U \begin{pmatrix} \cdots \\ 0 \\ 0 \end{pmatrix} \]
SVD: What’s this thing good for? (II)

- **Low-rank Approximation**

Theorem (Eckart–Young–Mirsky)

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

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

then

\[
\min_{\text{rank}(B)=k} \|A - B\|_2 = \|A - A_k\|_2
\]

**Demo:** Image compression
SVD: What’s this thing good for? (III)

- The minimum norm solution to $Ax \approx b$: $A = U \Sigma V^T$

\[ y \text{ has the smallest norm of all vectors } \text{ minimizing } \|Ax - y\|. \]

\[ \begin{align*}
\text{\( \exists \) } U \Sigma V^T x & \approx b \\
\begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} & \approx U^T b \\
\end{align*} \]

\[ y = V^T x \\
x = Vy \\
\|x\|_2 = \|y\|_2 \text{ because } V \text{ orthogonal.} \]

\[ \begin{align*}
\begin{bmatrix} y_1 \\ \vdots \\ y_n \\
\end{bmatrix} & = (U^T b)_i / \sigma_i \\
(i = 1, \ldots, n) \\
y_i & = 0 \\
(i = n+1, \ldots, m) \\
\end{align*} \]
\[ E^+ = \begin{pmatrix}
\frac{1}{\sigma_1} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & 0
\end{pmatrix} \]

Write down the solution to total CSE

\[ \mathbf{A} \mathbf{x} = \mathbf{b} \quad \text{will } \mathbf{A} = \text{MNE}\mathbf{V}^+ \]

\[ \mathbf{x} = \text{MNE}\mathbf{V}^+ \mathbf{b} \]

\text{"Carrying" Moore-Penrose pseudoinverse of } \mathbf{A}
SVD: Minimum-Norm, Pseudoinverse

\[ y = \Sigma^+ U^T b \] is the **minimum norm-solution** to \( \Sigma y \approx U^T b \).

Observe \( \| x \|_2 = \| y \|_2 \).

\[ x = V \Sigma^+ U^T b \]

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

Introduction to Scientific Computing

Systems of Linear Equations

Linear Least Squares

Eigenvalue Problems
  - Sensitivity
  - Properties and Transformations
  - Computing Eigenvalues
  - Krylov Space Methods

Nonlinear Equations

Optimization

Interpolation

Numerical Integration and Differentiation

Initial Value Problems for ODEs

Boundary Value Problems for ODEs

Partial Differential Equations and Sparse Linear Algebra

Fast Fourier Transform
Eigenvalue Problems: Setup/Math Recap

\[ \text{R-SVD} \]

\[ \text{left svd (} A \text{)} \]

\[ \text{eigenv. (} A^T A \text{)} \]

\[ A \text{ is an } n \times n \text{ matrix.} \]

\[ \Rightarrow \text{ If } x \neq 0 \text{ is called an } \text{eigenvector} \text{ of } A \text{ if there exists a } \lambda \text{ so that} \]

\[ Ax = \lambda x. \]

\[ \Rightarrow \text{ In that case, } \lambda \text{ is called an } \text{eigenvalue}. \]

\[ \Rightarrow \text{ The set of all eigenvalues } \lambda(A) \text{ is called the } \text{spectrum}. \]

\[ \Rightarrow \text{ The } \text{spectral radius} \text{ is the magnitude of the biggest eigenvalue:} \]

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

\[ x^6 + ax^5 + bx^4 + \cdots + c = 0 \]
Finding Eigenvalues

How do you find eigenvalues?

\[
Ax = \lambda x \iff (A - \lambda I)x = 0 \\
\iff A - \lambda I \text{ singular} \iff \det(A - \lambda I) = 0
\]

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

Does that help algorithmically? Abel showed that for \(n \geq 5\) 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.