Rank-Deficient Matrices and QR

What happens with QR for rank-deficient matrices?

\[ A = QR \]

- small diagonal entries
- "rank-revealing" so QR → orthogonally column-pivoted
Rank-Deficient Matrices and Least-Squares
What happens with Least Squares for rank-deficient matrices?

\[ Ax \approx b \]

\[ \| A(x + u) - b \|_2 = \| Ax + b \|_2 \]

\[ \min _{u \in N(A)} \| A(x + u) - b \|_2 \]

"Total least squares"
SVD: What’s this thing good for? (I)

\[ A = \begin{bmatrix} \sigma_1 \psi_1 \cdots \psi_n \end{bmatrix} \begin{bmatrix} \sigma_1 & & \cdots & \sigma_n \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \ddots & \sigma_n \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{\sigma_1}} v_1 \\ \vdots \\ \frac{1}{\sqrt{\sigma_n}} v_n \end{bmatrix} \]

\[ \|Ax\|_2 = \sigma_1 \]

\[ \text{cond}_2(A) = \frac{\sigma_1}{\sigma_n} \]

\[ N(A) = \text{span}(v_{k+1}, \ldots, v_n) \]

\[ N(\Sigma) = \text{span}\left(\begin{bmatrix} 0 & \cdots & 0 & 0 \sigma_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \sigma_n \\ 0 & \cdots & \cdots & 0 \end{bmatrix}^{\top} \right) \]

\[ \text{rank}(A) = k \]

\[ \text{num-rank}(A, \varepsilon) = \# \{ \sigma_i \geq \varepsilon \} \]

\[ \geq 0 \text{  and  } |\sigma_i| \geq |\sigma_k| \]
SVD: What’s this thing good for? (II)

- Low-rank Approximation

Theorem (Eckart-Young-Mirsky)

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

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

then

$$\min_{\text{rank}(B)=k} \|A - B\|_2 = \|A - A_k\|_2 = \sigma_{k+1},$$

$$\min_{\text{rank}(B)=k} \|A - B\|_F = \|A - A_k\|_F = \sqrt{\sum_{j=k}^{n} \sigma_j^2}.$$
SVD: What’s this thing good for? (III)

- The minimum norm solution to $Ax \approx b$:

$$A = U \Sigma V^T = \sum_{i=1}^{n} \sigma_i u_i v_i^T$$

$$Ax \approx b \iff \Sigma v_i x \approx u_i b \quad y = V^T x$$

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix}$$

$$\Sigma^+ = \begin{bmatrix} \sigma_1^{-1} & 0 & \cdots & 0 \\ 0 & \sigma_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$\|\Sigma y - c\|_2^2 = \sum_{i=1}^{r} (\sigma_i y_i - c_i)^2 + \sum_{i=r+1}^{n} c_i^2$$

$$\|y\|_2^2 = \sum_{i=1}^{r} y_i^2 + \sum_{i=r+1}^{n} y_i^2$$
What is the minimum 2-norm solution to $Ax \cong b$ and why?

$y = \Sigma^+ U^T b$ is the minimum norm solution.

$\Sigma y \cong U^T b$. $x = V y = V \Sigma^+ U^T b$

Generalize the pseudoinverse to the case of a rank-deficient matrix.

When $A$ is full rank, $V \Sigma^+ U^T = (A^T A)^{-1} A$
Comparing the Methods

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

1. **normal eqns.**
   \[ A^T A \text{ solve. } n^2 m^2/2 + n^3/6 \]

2. **QR with Householder dr.**
   \[ mn^2 - n^3/3 \]

3. **SVD:** \( mn^2 + n^3 \)

**Demo:** Relative cost of matrix factorizations [cleared]
In-Class Activity: Householder, Givens, SVD

**In-class activity:** Householder, Givens, SVD
Outline

Introduction to Scientific Computing

Systems of Linear Equations

Linear Least Squares

**Eigenvalue Problems**
- Properties and Transformations
- Sensitivity
- 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

Additional Topics
$A$ is an $n \times n$ matrix.

- $x \neq 0$ is called an eigenvector of $A$ if there exists a $\lambda$ so that $Ax = \lambda x$.

- In that case, $\lambda$ 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: 
  \[ \rho(A) = \max \{|\lambda| : \lambda(A)\} \]
Finding Eigenvalues

How do you find eigenvalues?

\[ A\mathbf{x} = \lambda \mathbf{x} \iff (A - \lambda I)\mathbf{x} = 0 \]

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

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

Does that help algorithmically? Abel-Ruffini theorem: for \( n \geq 5 \) is no general formula for roots of polynomial. IOW: no.

- For LU and QR, we obtain exact answers (except rounding).
- For eigenvalue problems: not possible—must approximate.

**Demo:** Rounding in characteristic polynomial using SymPy [cleared]
What is the multiplicity of an eigenvalue?

Actually, there are two notions called multiplicity:

- **Algebraic Multiplicity**: multiplicity of the root of the characteristic polynomial
- **Geometric Multiplicity**: \# of lin. indep. eigenvectors

In general: \( AM \geq GM \).

If \( AM > GM \), the matrix is called **defective**.
An Example

Give characteristic polynomial, eigenvalues, eigenvectors of

\[
\begin{pmatrix}
1 & 1 \\
1 & 1
\end{pmatrix}.
\]

\[|\lambda I - A| = (\lambda - 1)^2 \]

AM = 2

GM = |< AM |

\[\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow y = 0 \]

\[\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix} \]

\[1D\]