- Exemlel 2
- HW 6

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

What if A doesn't have full rank?

$$S = \{ x : \|A\vec{x} - \vec{b}\|_2 = \min_{\vec{y}} \|A\vec{y} - \vec{b}\|_2 \}$$

solution $= \underset{x \in S}{\operatorname{argmin}} \|\vec{x}\|_2$

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

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

$A = U\Sigma V^T$

$$\text{min!} \leftarrow \|Ax - b\|_2$$
$$= \|U\Sigma V^T x - b\|_2$$
$$= \|\Sigma \underbrace{V^T x}_{\vec{y}} - U^T b\|_2$$
$$= \|\Sigma \vec{y} - \vec{c}\|_2$$
$$= \left\| \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} \vec{y} - \vec{c} \right\|_2$$

$\vec{c} := U^T \vec{b}$

full rank: $y_i = c_i / \sigma_i$
(all $\sigma_i \neq 0$)

rank-def.
$\exists i\ \sigma_i = 0$

$\sigma_1 \dots \sigma_k \neq 0$

$\sigma_{k+1} \dots \sigma_n = 0$

$$\left\| \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_k & \\ 0 & & & 0 \end{pmatrix} \tilde{y} - \tilde{c} \right\|_2$$

$y_{k+1} \dots y_n$ do not influence the residual norm.

$\Rightarrow$ Of all vectors $\tilde{y}$, the one w/ $y_{k+1} = \dots = y_n = 0$ minimizes $\|y\|_2$.

$\Rightarrow \|y\|_2 = \|V^T x\|_2 = \|x\|_2$.

Numerically, we'd replace the test

$\sigma_i = 0$ $\quad\quad$ $\sigma_i \geq \varepsilon$, with ε given

by the user.

$$A = U \Sigma V^T \qquad\qquad A\vec{x} \approx \vec{b}$$

$$A^+ \vec{b}$$

pseudoinverse: $A^+ = V \, \Sigma^+ U^T$

$$\Sigma^+ = \text{diag} \begin{cases} 1/\sigma_i & \text{if "}\sigma_i \neq 0\text{"} \\ 0 & \text{otherwise} \end{cases}$$

SVD: Minimum-Norm, Pseudoinverse

What is the minimum 2-norm solution to $A\boldsymbol{x} \cong \boldsymbol{b}$ and why?

✓

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

$A = U\Sigma V^T$

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

Comparing the Methods

np.lstsq

Methods to solve least squares with A an $m \times n$ matrix: (ld. order in # multipl.s)

normal eqns: A^TA : $mn^2/2$

Cholesky: $n^3/6$

Householder: $mn^2 - n^3/3$

SVD: $C(mn^2 + n^3)$

Demo: Relative cost of matrix factorizations [cleared]

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

Eigenvalue Problems: Setup/Math Recap

A is an $n \times n$ matrix.

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

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

- 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:

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

Finding Eigenvalues

How do you find eigenvalues?

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

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

Does that help algorithmically? Abel-Ruffini theorem: for $n \geqslant 5$ 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]

Multiplicity

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 $\geqslant$ GM.
If AM $>$ GM, the matrix is called *defective*.

An Example

Give characteristic polynomial, eigenvalues, eigenvectors of

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

CP: $(1-\lambda)^2$

Eigenvalues: 1, with alg. multiplicity 2

$$\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} y = y \\ x + y = x \Rightarrow y = 0 \end{cases}$$

eigenvectors: $\begin{pmatrix} x \\ 0 \end{pmatrix}$

geo. multiplicity: 1

Diagonalizability

$A \in \mathbb{R}^{n \times n}$
$\llcorner$ square!

When is a matrix called *diagonalizable*?

$\Leftrightarrow$ not defective $\Leftrightarrow$ $\exists$ n linearly indep. eigvec.

$$X = \begin{bmatrix} | & & | \\ \vec{x}_1 & \cdots & \vec{x}_n \\ | & & | \end{bmatrix}$$ basis of eigenvectors, X invertible

$$AX = \begin{pmatrix} | & & | \\ \lambda_1 \vec{x}_1 & \cdots & \lambda_n \vec{x}_n \\ | & & | \end{pmatrix} = XD \Leftrightarrow A = XDX^{-1}$$

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

Similar Matrices

Related definition: Two matrices A and B are called similar if there exists an invertible matrix X so that $A = XBX^{-1}$.

In that sense: "Diagonalizable" = "Similar to a diagonal matrix".

Observe: Similar A and B have same eigenvalues. (Why?)

Suppose $A\vec{v} = \lambda\vec{v}$ $\quad B = X^{-1}AX \quad \vec{w} = X^{-1}\vec{v}$

$B\vec{w} = X^{-1}AXX^{-1}v = X^{-1}A\vec{v} = X^{-1}\lambda\vec{v}$

$= \lambda\vec{w}$

Eigenvalue Transformations (I)

What do the following transformations of the eigenvalue problem $A\boldsymbol{x} = \lambda\boldsymbol{x}$ do?

Shift. $A \to A - \sigma I$

$$(A - \sigma I)\vec{x} = A\vec{x} - \sigma\vec{x} = \lambda\vec{x} - \sigma\vec{x} = (\lambda - \sigma)\vec{x}$$

Inversion. $A \to A^{-1}$

$$A^{-1}\vec{x} = \frac{1}{\lambda}\vec{x}$$

Power. $A \to A^k$

$$A^k\vec{x} = \underbrace{A \cdots A}_{k \text{ times}}\vec{x} = \lambda^k\vec{x}$$

Eigenvalue Transformations (II)

Polynomial $A \to aA^2 + bA + cI$

$$(aA^2 + bA + cI)\vec{x} = (a\lambda^2 + b\lambda + c)\vec{x}$$

Similarity $T^{-1}AT$ with T invertible

$$\vec{y} = T^{-1}\vec{x}$$

$$T^{-1}ATy = \lambda y$$

Sensitivity (I)

Assume A not defective. Suppose $X^{-1}AX = D$. Perturb $A \to A + E$. What happens to the eigenvalues?

$$X^{-1}(A+E)X = D + F$$

defines F.

$A+E$, $D+F$ have the same eigenvalues

(sim)

Suppose $\vec{v}$ is eigenvector of $D+F$

assume $\mu \notin \operatorname{diag}(D)$

$$\Leftrightarrow (D+F)\vec{v} = \mu\vec{v}$$

$$\Leftrightarrow F\vec{v} = (\mu I - D)\vec{v} \quad \Big|\ (\mu I - D)^{-1}\cdot$$

$$\Leftrightarrow (\mu I - D)^{-1} F\vec{v} = \vec{v}$$

$$\Rightarrow \quad 1 \cancel{\|v\|} \le \|(\mu I - D)^{-1}\|\,\|F\|\,\cancel{\|v\|}$$

$$\|(\mu I - D)^{-1}\|^{-1} \leq \|F\|$$

distance between μ and the eigenvalue of A closest to μ.

Bauer Fike

$\Rightarrow$ no* assumptions on μ or v: that bound holds

$\rightarrow$ For all eigenvalues of $A+E$.

$$F = X^{-1} E X$$

$$\|(\mu I - D)^{-1}\|^{-1} \leq \|X^{-1} E X\| \leq \text{cond}(X)\,\|E\|$$

max distance between perturbed and closest true eigenvalue $\leq$ cond(X) $\|E\|$

Sensitivity (II)

$X^{-1}(A+E)X = D + F$. Have $\left\|(\mu I - D)^{-1}\right\|^{-1} \le \|F\|$.

(covered, see above)

Demo: Bauer-Fike Eigenvalue Sensitivity Bound [cleared]

Power Iteration

What are the eigenvalues of A^{1000}?

Assume $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$ with eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$.
Further assume $\|\boldsymbol{x}_i\| = 1$.