CS475 / CM375
Lecture 13: Oct 25, 2011

Singular Value Decomposition

Conditioning

Reading: [TB] Chapters 4, 12

Singular Value Decomposition

- A third method to solve least square problems:
  - Singular Value Decomposition (SVD)
- Idea: compute $A = \hat{U}\hat{\Sigma}V^T$
  - Picture:
  - Where $\hat{U}, V$ have orthonormal cols and $\hat{\Sigma} = diag$

Singular Value Decomposition

- Geometry:

Singular Value Decomposition

- $Ax = b \implies \hat{U}\hat{\Sigma}V^T x = b$

$$\hat{\Sigma}V^T x = \hat{U}^T b \qquad (\hat{U}^T\hat{U} = I)$$

$$V^T x = \hat{\Sigma}^{-1}\hat{U}^T b$$

$$x = V\hat{\Sigma}^{-1}\hat{U}^T b \qquad (VV^T = I)$$

- Pseudoinverse: $A^{\dagger} = V\hat{\Sigma}^{-1}\hat{U}^T$

Singular Value Decomposition

- Normal equations view:

$$A^T Ax = A^T b$$

$\Leftrightarrow$

$\Leftrightarrow$

$\Leftrightarrow$

$\Leftrightarrow$

$\Leftrightarrow$

$\Leftrightarrow$

Conditioning

- Def: Conditioning refers to the perturbation behavior of a mathematical problem
- Consider a problem $f: X \rightarrow Y$
  - Well-conditioned: small changes in x → small changes in y
  - Ill-conditioned: small changes in x → large changes in y
- Picture:

Condition number

- Let δx denote a small perturbation of x
- Let $\delta f = f(x + \delta x) - f(x)$

- Absolute condition number: $\hat{\kappa} = \sup_{\delta x} \frac{||\delta f||}{||\delta x||}$
- Relative condition number: $\kappa = \sup_{\delta x} \left(\frac{||\delta f||}{||f(x)||} / \frac{||\delta x||}{||x||} \right)$

- Well-conditioned: small κ
- Ill-conditioned: large κ

Conditioning of Matrix-Vector Multiplication

- Let $f(x) = Ax$
- Then $\kappa = \sup_{\delta x} \left(\frac{||A(x+\delta x) - Ax||}{||Ax||} \; / \; \frac{||\delta x||}{||x||} \right)$

$$= \sup_{\delta x} \frac{||A\delta x||}{||\delta x||} \; / \; \frac{||Ax||}{||x||}$$

Recall the matrix norm: $||A|| = \sup_{x} \frac{||Ax||}{||x||}$

$$= ||A|| \frac{||x||}{||Ax||}$$

Condition number of a matrix

- Let $\kappa(A)$ be the condition number of matrix A
  - Def: largest condition number achieved by multiplying some vector x by A
- Hence $\kappa(A) = \sup_x ||A|| \frac{||x||}{||Ax||}$

$$= \sup_x ||A|| \frac{||A^{-1}x||}{||x||}$$

$$= ||A|| \, ||A^{-1}||$$

Condition number of a matrix

- For Euclidean norm:

$$\kappa(A) = ||A||_2 ||A^{-1}||_2$$

$$= ||A||_2 \, ||A^{\dagger}||_2 \quad \text{(when } A \text{ is rectangular)}$$

$$= \frac{\sigma_1}{\sigma_m}$$

where σ_1 = largest singular value
and σ_m = smallest singular value

Conditioning of LS problems

- Theorem: Suppose $A \in \Re^{m \times n}$ has full rank and that x minimizes $||Ax - b||_2$. Let $r = b - Ax$.
  Let $\tilde{x}$ minimizes $||(A + \delta A)\tilde{x} - (b + \delta b)||_2$.
  Assume $\epsilon = \max\left(\frac{||\delta A||}{||A||}, \frac{||\delta b||}{||b||}\right) < \frac{1}{\kappa(A)}$
  Then $\frac{||\tilde{x} - x||}{||x||} \leq \epsilon \left[\frac{2\kappa(A)}{\cos\theta} + \tan\theta\, \kappa^2(A)\right] + O(\epsilon^2)$
  $\equiv \epsilon \kappa_{LS} + O(\epsilon^2)$
  where $\theta = \measuredangle(b, Ax)$, $\kappa_{LS} = \text{condition number of LS}$

Conditioning of LS problems

- Recall $\frac{||\tilde{x} - x||}{||x||} \leq \epsilon \left[\frac{2\kappa(A)}{\cos\theta} + \tan\theta\, \kappa^2(A)\right] + O(\epsilon^2)$
  $\equiv \epsilon \kappa_{LS} + O(\epsilon^2)$

- Notes
  - If $\theta \approx 0$, then $\kappa_{LS} \approx 2\kappa(A)$
  - If $0 < \theta < \frac{\pi}{2}$, then κ_{LS} is much larger due to $\kappa^2(A)$
  - If $\theta \approx \frac{\pi}{2}$, then $\kappa_{LS} = \infty$ even if $\kappa(A)$ is small

Stability of LS algorithms

- Recall
  - Normal equations: $A^T A x = A^T b \quad \Longrightarrow \kappa(A^T A)$
  - QR factorization: $Ax = QRx = b \quad \Longrightarrow \kappa(A)$
  - SVD: $Ax = U\Sigma V x = b \quad \Longrightarrow \kappa(A)$
- Notes
  1. Normal equations: $\kappa(A^T A) = \kappa(A)^2$

     $$\Longrightarrow \frac{||\tilde{x} - x||}{||x||} = O(\epsilon \kappa(A)^2)$$
  2. If $\theta \ll \frac{\pi}{2}$, then $\kappa(A) \le \kappa_{LS} \le \kappa(A)^2$
  3. SVD is most stable and most expensive

Eigenvalue Problems

- <u>Def:</u> Let $A \in \Re^{n \times n}$. A nonzero vector $x \in \Re^n$ is an eigenvector and $\lambda \in \mathbb{C}$ is its corresponding eigenvalue if

  $$Ax = \lambda x$$

- If x is an eigenvector, then αx (s.t. $\alpha \ne 0$) is also an eigenvector

Eigenvalue Problems

- Def: The set $\Lambda(A) = \{\lambda: \lambda \text{ is an eigenvalue of } A\}$ is the spectrum of A.
- An eigen decomposition of A is: $A = X\Lambda X^{-1}$

  where

Characteristic Polynomial

- Def: The characteristic polynomial of A, $p_A(x)$, is the degree n polynomial defined by

$$p_A(z) = \det(zI - A)$$

- Theorem: λ is an eigenvalue of A iff $p_A(\lambda) = 0$
- Proof: λ is an eigenvalue

  $\Leftrightarrow \lambda x - Ax = 0$ for some $x \neq 0$

  $\Leftrightarrow \lambda I - A$ is singular

  $\Leftrightarrow \det(\lambda I - A) = 0$

Characteristic Polynomial

1. By the fundamental theorem of algebra, $p_A(z)$ has n (complex) roots. So A has n (complex) eigenvalues
2. Given a monic polynomial of degree n,

$$p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$$

Consider $A =$

Then $\Lambda(A) = \{\text{roots of } p(z)\}$

Characteristic Polynomial

3. No analytic formula for roots of polynomial of degree ≥ 5

→ Numerical approximation: eigen decomposition techniques