CS475 / CS675
Lecture 19: July 5, 2016

Singular value decomposition

Reading: [TB] Chapter 31

Singular Value Decomposition

- Geometric view:

- Let S be the unit sphere in $\Re^n$. The image AS is an ellipse in $\Re^m$.

Interpretation

- The n singular values of A are the lengths of the n principal semi-axes of AS: $\sigma_1, \sigma_2, \dots, \sigma_n$
  - Convention: $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$
- The n left singular vectors of A are the unit vectors $\{u_1, \dots, u_n\}$ in the direction of the principal semi-axes.
- The n right singular vectors of A are the unit vectors $\{v_1, \dots, v_n\} \in S$ such that $Av_j = \sigma_j u_j$

Reduced SVD

- Decomposition: $A = \widehat{U}\hat{\Sigma}V^T$
  - Picture:
  - Here $\hat{\Sigma} = diag, \widehat{U}$ and V have orthonormal columns.
- Equivalently: $AV = \widehat{U}\hat{\Sigma}$
  - Hence $Av_j = \sigma_j u_j \quad \forall j = 1,2, \dots, n$
  - Picture:

Full SVD

- Extend $\widehat{U} \rightarrow U$ = orthogonal
- Accordingly, $\hat{\Sigma} \rightarrow \Sigma = \begin{bmatrix} \hat{\Sigma} \\ 0 \end{bmatrix} \begin{matrix} \}n \\ \}m-n \end{matrix}$
- Then $A = U\Sigma V^T$ where $\Sigma = diag, U, V = orth$
  - Picture:

SVD vs Eigendecomposition

- They both diagonalize a matrix A. SVD uses 2 bases (left and right singular vectors). Eigendecomposition uses 1 basis (eigenvectors)

- SVD uses orthonormal vectors where as eigenvectors are not orthonormal in general

- Not all matrices have an eigendecomposition. But all matrices have a singular value decomposition

Matrix properties of SVD

- Let $A \in \Re^{m \times n}, p = \min(m, n)$,
  $r =$ # of nonzero singular values of A.

- <u>Theorem:</u> $rank(A) = r$
- <u>Proof:</u> The rank of a diagonal matrix = # of nonzero diagonal entries. Since $A = U\Sigma V^T$, then $U, V \text{ orth} \Longrightarrow rank(A) = rank(\Sigma)$

Matrix properties of SVD

- Theorem: $range(A) = span\{U_1, \dots, U_r\}$ and $null(A) = span\{V_{r+1}, \dots, V_n\}$

- Theorem: $||A||_2 = \sigma_1$ and $||A||_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$

  Note: $||A||_2^2 = \lambda_{\max}(A^T A)$, $||A||_F^2 = \sum_{ij} a_{ij}^2$

Matrix properties of SVD

- Proof: $||\mathrm{A}||_2^2 =$

$||A||_F^2 =$

Matrix properties of SVD

- Theorem: The nonzero singular values of A are the square roots of the nonzero eigenval. of $A^T A$ or AA^T.
- Proof: $A^T A$ and AA^T are similar to Σ^2

- Theorem: If $A = A^T$, then $\sigma(A) = \{|\lambda|: \lambda \in \Lambda(A)\}$. In particular, if A is SPD, then $\sigma(A) = \Lambda(A)$.

Matrix properties of SVD

- Theorem: the condition number of $A \in \Re^{n \times n}$ is $\frac{\sigma_1}{\sigma_n}$
- Proof:

Computing the SVD

- Recall:
  $A = U\Sigma V^T$
  $A^T A = (V\Sigma^T U^T)(U\Sigma V^T)$
  $= V\Sigma^T \Sigma V^T$

  $\therefore$ eigenvalues of $A^T A$ are $\{\sigma_i^2\}$

An SVD algorithm

(1) Form $A^T A$

(2) Compute the eigendecomposition $A^T A = V \Lambda V^T$

(3) Compute

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}, \sigma_i = \sqrt{\lambda_i}, \quad \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

(4) Solve the equation

$$U\Sigma = AV$$

for orthogonal U (by QR factorization)

An SVD algorithm

- Unstable algorithm
  - Suppose $\lambda_k(A^T A)$ is computed stably, i.e.,
  $$\left|\tilde{\lambda}_k - \lambda_k\right| = O\left(\epsilon\left||A^T A|\right|\right) = O(\epsilon\left||A|\right|^2)$$
  - Take square root to get σ_k:
  $|\tilde{\sigma}_k - \sigma_k|$
  $$= O\left(\frac{\left|\tilde{\lambda}_k - \lambda_k\right|}{\sqrt{\lambda_k}}\right) = O\left(\frac{\epsilon\left||A|\right|^2}{\sigma_k}\right) = O\left(\epsilon\left||A|\right|\frac{\left||A|\right|}{\sigma_k}\right)$$
  - If $\sigma_k \ll ||A||$, (e.g., σ_n), then
  $$|\tilde{\sigma}_k - \sigma_k| \approx O\left(\epsilon\left||A|\right|\frac{\sigma_1}{\sigma_n}\right) = O\left(\epsilon\left||A|\right|\kappa(A)\right)$$
  $\Longrightarrow$ loss of accuracy $\approx O(\kappa(A))$

Example

- Find the SVD of $A = \begin{bmatrix} 0 & -1/2 \\ 3 & 0 \\ 0 & 0 \end{bmatrix}$
- Method 1:

Example (continued)

- Method 2:
- Method 3:

Example (continued)

- Method 3 (continued):