Singular Value Decomposition

Rohit Budhiraja, IITK

Applied Linear Algebra for Wireless Communications

Recap and agenda for today's class

- Discussed the concept of positive definite matrices
- Discuss singular value decomposition
  - Chapter 6.6 of the book
- Discuss Hermitian and Unitary matrices
  - Chapter 9.2 of the book

Singular Value Decomposition (1)

- Singular Value Decomposition (SVD) is a highlight of linear algebra
- Consider a rectangular $m \times n$ matrix A with rank r
- We will diagonalize this A, but not by $X^{-1}AX$
- Eigenvectors in X have three big problems:
  - $A\mathbf{x} = \lambda\mathbf{x}$ requires A to be a square matrix
  - They are usually not orthogonal
  - There are not always enough eigenvectors
- Singular vectors of A solve all those problems in a perfect way
- SVD actually provides the right bases for the four subspaces
- Price we pay is to have two sets of singular vectors, $\mathbf{u}$'s and $\mathbf{v}$'s
  - $\mathbf{u}$'s are in $\mathbf{R}^m$ and the $\mathbf{v}$'s are in $\mathbf{R}^n$

Singular Value Decomposition (2)

- $\mathbf{u}$'s and $\mathbf{v}$'s give bases for the four fundamental subspaces
  - $\mathbf{u}_1, \ldots, \mathbf{u}_r$ is an orthonormal basis for the column space
  - $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m$ is an orthonormal basis for the left nullspace $N(A^T)$
  - $\mathbf{v}_1, \ldots, \mathbf{v}_r$ is an orthonormal basis for the row space
  - $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$is an orthonormal basis for the nullspace $N(A)$
- More than just orthogonality, these basis vectors diagonalize the matrix A

$$A\mathbf{v}_1 = \sigma_1 \mathbf{u}_1, A\mathbf{v}_2 = \sigma_2 \mathbf{u}_2, \ldots, A\mathbf{v}_r = \sigma_r \mathbf{u}_r$$

- Singular values σ_1 to σ_r are positive numbers as σ_i is length of $A\mathbf{v}_i$

$$\begin{aligned} ||A\mathbf{v}_i||^2 &= (A\mathbf{v}_i)^T (A\mathbf{v}_i) = (\sigma_i \mathbf{u}_i)^T (\sigma_i \mathbf{u}_i) = \sigma_i^2 \mathbf{u}_i^T \mathbf{u}_i = \sigma_i^2 \\ \Rightarrow ||A\mathbf{v}_i|| &= \sigma_i \end{aligned}$$

- These σ_i go into a diagonal matrix Σ_r

Singular Value Decomposition (3)

- Since **u**'s are orthonormal, the matrix U_r with those r columns has $U_r^T U_r = I$
- Since **v**'s are orthonormal, the matrix V_r with those r columns has $V_r^T V_r = I$
- Equations $A\mathbf{v}_i = \sigma_i \mathbf{u}_i$ tell us column by column that $AV_r = U_r \Sigma_r$

$$\begin{array}{c} (m \text{ by } n)(n \text{ by } r) \\ \boldsymbol{AV_r = U_r \Sigma_r} \\ (m \text{ by } r)(r \text{ by } r) \end{array} \qquad A \begin{bmatrix} & & \\ v_1 & \cdots & v_r \\ & & \end{bmatrix} = \begin{bmatrix} & & \\ u_1 & \cdots & u_r \\ & & \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

- This is the heart of the SVD, but there is more
- Those **v**'s and **u**'s account for the row space and column space of A
- We have $n - r$ more **v**'s and $m - r$ more **u**'s
  - They are from the nullspace $N(A)$ and the left nullspace $N(A^T)$
- They are automatically orthogonal to the first **v**'s and **u**'s
  - because the whole nullspaces are orthogonal

Singular Value Decomposition (4)

- We now include all the **v**'s and **u**'s in V and U

$$\begin{array}{c} (m \text{ by } n)(n \text{ by } n) \\ AV \;\; \text{equals} \;\; U\Sigma \\ (m \text{ by } m)(m \text{ by } n) \end{array} \quad A \begin{bmatrix} v_1 & \cdots & v_r & \cdots & v_n \end{bmatrix} = \begin{bmatrix} u_1 & \cdots & u_r & \cdots & u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix}$$

- So these matrices become square. We still have $AV = U\Sigma$
- The new Σ is $m \times n$
  - It is $r \times r$ Σ_r with $m - r$ extra zero rows and $n - r$ new zero columns
- Real change is in shapes of U and V, which are square matrices
- We also have $V^{-1} = V^T$ So $AV = U\Sigma$; becomes $A = U\Sigma V^T$
  - This is the Singular Value Decomposition
- We need to show how those amazing **u**'s and **v**'s can be constructed

Proof of SVD

- The $\mathbf{v}$'s are orthonormal eigenvectors of A^TA

$$A^TA = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V\Sigma^T\Sigma V^T$$

- V: Eigenvector matrix for symmetric positive (semi) definite matrix A^TA
- And $\Sigma^T\Sigma$ must be the eigenvalue matrix of (A^TA): each σ_i^2 is $\lambda(A^TA)$!
- Now $A\mathbf{v}_i = \sigma_i\mathbf{u}_i$ tells us about unit vectors $\mathbf{u}_1$ to $\mathbf{u}_r$
- Essential point that SVD succeeds is that $\mathbf{u}_1$ to $\mathbf{u}_r$ are orthogonal

Key step $i \neq j$

$$u_i^{\mathrm{T}}u_j = \left(\frac{Av_i}{\sigma_i}\right)^{\mathrm{T}}\left(\frac{Av_j}{\sigma_j}\right) = \frac{v_i^{\mathrm{T}}A^{\mathrm{T}}Av_j}{\sigma_i\sigma_j} = \frac{\sigma_j^2}{\sigma_i\sigma_j}v_i^{\mathrm{T}}v_j = \textbf{zero}$$

- $\mathbf{u}'s$ are eigenvectors of AA^T
- Complete $\mathbf{v}$'s and $\mathbf{u}$'s to n $\mathbf{v}$'s and m $\mathbf{u}$'s with orthogonal basis from $N(A)$ and $N(A^T)$
- We have found U, Σ, and V in $A = U\Sigma V^T$

Complex vectors and matrices

- Consider complex vector $\mathbf{z}$ and matrix A
- Main message of this section can be presented in one sentence
  - While transposing a complex vector $\mathbf{z}$ or matrix A, take complex conjugate
- Conjugate transpose

$$\bar{\mathbf{z}}^T = [\bar{z}_1, \cdots, \bar{z}_n] = z^H$$

- Here is one reason to go to $\bar{\mathbf{z}}$- length squared of a real vector is $x_1^2 + \cdots + x_r^2$,
- The length squared of a complex vector is not $z_1^2 + \cdots + z_r^2$
- With that wrong definition, the length of $(1, i)$ would be $1^2 + i^2 = 0$
- A non-zero vector would have zero length which is not good
- Instead of $(a + bi)^2$ we want $a^2 + b^2$, the absolute value squared
  - This is $(a + bi)$ times $(a - bi)$

Length of a complex vector

- For each component we want z_j times $\bar{z}_j$, which is $|z_j|^2 = a_j^2 + b_j^2$
- That comes when the components of $\mathbf{z}$ multiply the components of $\bar{\mathbf{z}}$

$$\begin{aligned} [\bar{z}_1, \cdots, \bar{z}_n] \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} &= |z_1|^2 + \cdots + |z_n|^2 \\ \bar{\mathbf{z}}^T \mathbf{z} &= ||\mathbf{z}||^2 \\ \mathbf{z}^H \mathbf{z} &= ||\mathbf{z}||^2 \end{aligned}$$

- $\mathbf{z}^H$ is the $\mathbf{z}$ Hermitian and length $||\mathbf{z}||$ is the square root of $\mathbf{z}^H \mathbf{z}$
- Similarly we have A^H. If $A = \begin{bmatrix} 1 & i \\ 0 & 1+i \end{bmatrix}$ then $A^H = \begin{bmatrix} 1 & 0 \\ -i & 1-i \end{bmatrix}$

Operation on complex vectors and matrices

- Inner product of real or complex vectors $\mathbf{u}$ and $\mathbf{v}$ is $\mathbf{u}^H\mathbf{v}$
- With complex vectors $\mathbf{u}$ and $\mathbf{v}$, $\mathbf{u}^H\mathbf{v}$ is different from $\mathbf{v}^H\mathbf{u}$
- Conjugate transpose of $(AB)^H = B^H A^H$
- Hermitian matrix: If $S = S^H$
- If $S = S^H$ and $\mathbf{z}$ is a real /complex column vector, the number $\mathbf{z}^H S\mathbf{z}$ is real
  - Proof: $(\mathbf{z}^H S\mathbf{z})^H = \mathbf{z}^H S\mathbf{z}$
  - If conjugate transpose of number is the same number, then number is real
- Every eigenvalue of a Hermitian matrix is real.
  - Proof: $S\mathbf{z} = \lambda\mathbf{z} \Rightarrow \mathbf{z}^H S\mathbf{z} = \lambda\mathbf{z}^H\mathbf{z} \Rightarrow \mathbf{z}^H S\mathbf{z} = \lambda||\mathbf{z}||^2$
  - Since λ is obtained by dividing two real numbers, it is real
- Eigenvectors of a Hermitian matrix are orthogonal (when they correspond to different eigenvalues)
- Unitary matrix: square matrix Q such that $Q^H Q = I$