Lecture 6

Announcements

HW2 due Wednesday

HW3 posted (start early!)

Readings

Strang I. 8

Outline

- Singular Value Decomposition
  - Definition
  - Intuition
  - Existence
  - Geometric interpretation

Review

Real symmetric matrices have real eigenvalues and orthogonal eigenvectors.

Moreover, the column space and row space are identical.

Real symmetric matrices are diagonalizable so

$$S = Q \Lambda Q^T \qquad \text{so } Q^T \text{ is the matrix of left eigenvectors}$$

$$SQ = Q\Lambda$$

What about for non-square matrices?

Consider $A \in \mathbb{R}^{m \times n}$ and suppose $A\vec{x} = \lambda \vec{x}$ $\quad \vec{x} \in \mathbb{R}^n$

$A\vec{x} \in \mathbb{R}^m$

Singular Value Decomposition (SVD)

For a real $m \times n$ matrix, instead of constructing a set of orthogonal eigenvectors we will construct two sets of orthogonal singular vectors,

- n right singular vectors $\vec{v}_1, \ldots, \vec{v}_n$
- m left singular vectors $\vec{u}_1, \ldots, \vec{u}_m$

These will form bases for the row and column spaces of A.

For eigenvalues/eigenvectors $A\vec{x} = \lambda \vec{x}$

For singular vectors $\boxed{A\vec{v}_i = \sigma_i \vec{u}_i}$

row space

column space

In particular, $A\vec{v}_1 = \sigma_1 \vec{u}_1, \ldots, A\vec{v}_r = \sigma_r \vec{u}_r$ and $A\vec{v}_{r+1} = \ldots = A\vec{v}_n = \vec{0}$

where $r = \text{rank}(A) = \dim(C(A)) = \dim(C(A^T))$

So the last $n-r$ $\vec{v}$'s are in the null space of A

the last $m-r$ $\vec{u}$'s are in the null space of A^T

$$A \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \vec{u}_1 & \cdots & \vec{u}_m \\ | & & | \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right]$$

$m \times n \quad n \times n \quad m \times m \quad m \times n$

$$A \quad V \quad = \quad U \quad \Sigma$$

(compare w/ $SQ = Q\Lambda$)

$$\boxed{A = U\Sigma V^T}$$ (since $V^{-1} = V^T$)

- V is the matrix of right singular vectors $(n \times n)$ orthogonal
- U is the matrix of left singular vectors $(m \times m)$ orthogonal.
- Σ is the matrix of singular values $(m \times n)$

ex $A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$

$$\underset{A}{\begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}} \underset{V}{\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}} = \underset{U}{\begin{bmatrix} \frac{1}{\sqrt{10}} & \frac{-3}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{bmatrix}} \underset{\Sigma}{\begin{bmatrix} 3\sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix}}$$

- A is not symmetric (if it were $V = U$)
- $\det(\Sigma) = 15 = \det(A)$
- V and U are orthogonal matrices
- $A = \sigma_1 \vec{u}_1 v_1^T + \sigma_2 \vec{u}_2 \vec{v}_2^T$ (VERIFY)

$A = U\Sigma V^T$

In general if A is rank r,

$$\boxed{A = \sigma_1 \vec{u}_1 \vec{v}_1^T + \cdots + \sigma_r \vec{u}_r \vec{v}_r^T}$$

We can write A in a "reduced" SVD form

$$A V_r = U_r \Sigma_r$$

$$\underset{m \times n}{A} \underset{n \times r}{\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_r \\ | & & | \end{bmatrix}} = \underset{m \times r}{\begin{bmatrix} | & & | \\ \vec{u}_1 & \cdots & \vec{u}_r \\ | & & | \end{bmatrix}} \underset{r \times r}{\begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix}}$$

note $V_r \in \mathbb{R}^{n \times r}$ and $V_r^T V_r = I_r$ but $V_r V_r^T \neq I_n$

Why does this matter?

$A_k = \sigma_1 \vec{u}_1 \vec{v}_1^T + \dots + \sigma_k \vec{u}_k \vec{v}_k^T$ is the "best" rank-k approximation to A.

(Eckart-Young — we will discuss next class)

"Proof" of existence of SVD

$A = U \Sigma V^T$ What is A^TA? $A^TA = (V\Sigma^T U^T)(U\Sigma V^T) = V\Sigma^2 V^T$

What is AA^T? $AA^T = U\Sigma^2 U^T$

Note that A^TA and AA^T are symmetric and thus diagonalizable.

$A^TA = V\Lambda V^T$ (where $\Lambda = \Sigma^2$)

$AA^T = U\Lambda U^T$

- V contains orthonormal eigenvectors of A^TA
- U contains orthonormal eigenvectors of AA^T
- σ_1^2 to σ_r^2 are the nonzero eigenvalues of A^TA and AA^T

Construction

1. Choose orthonormal eigenvectors of A^TA $\vec{v}_1, \dots, \vec{v}_n$
2. Choose $\sigma_k = \sqrt{\lambda_k}$ for all k. $\boxed{A\vec{v} = \sigma\vec{u}}$
3. Set $\vec{u}_k = \dfrac{A\vec{v}_k}{\sigma_k}$ for all $k = 1, \dots, r$
4. and choose $\vec{u}_k$ to be orthogonal for $k = r+1, \dots, m$. to complete U
   and choose v_k to orthogonal for $k = r+1, \dots, n$ to complete V

ex $A = \begin{bmatrix} 3 & 0 \\ 4 & 5 \end{bmatrix}$ $A^TA = \begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix}$ $AA^T = \begin{bmatrix} 9 & 12 \\ 12 & 41 \end{bmatrix}$

① eigenvalues of $A^TA \longrightarrow$ solve $\det(A^TA - \lambda I) = 0$

② $\lambda_1 = 45 = \sigma_1^2$

$\lambda_2 = 5 = \sigma_2^2$

solve for eigenvectors

$$\begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 45 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Longrightarrow \vec{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\begin{bmatrix} 25 & 20 \\ 20 & 25 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 5 \begin{bmatrix} -1 \\ 1 \end{bmatrix} \Longrightarrow \vec{v}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

③ $$A\vec{v}_1 = \begin{bmatrix} 3/\sqrt{2} \\ 9/\sqrt{2} \end{bmatrix} = \sqrt{45} \begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix} = \sigma_1 \vec{u}_1$$

$$A\vec{v}_2 = \begin{bmatrix} -3/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \sqrt{5} \begin{bmatrix} -3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix} = \sigma_2 \vec{u}_2$$

Remarks

① If S is symmetric pos. def. then the SVD is $Q\Lambda Q^T = U\Sigma V^T$ and all singular values are positive.

② If S has a negative eigenvalue then $S\vec{x} = -\alpha\vec{x}$ for some $\alpha \geq 0$. The corresponding singular value will $\sigma = +\alpha$ and one of the corresponding singular vectors will be $-\vec{x}$.

③ If $A = Q$ is orthogonal, what is its SVD?

All singular values are 1 and $A^TA = I$ and $\Sigma = I$.