Lecture 7

Announcements

- Last day to switch classes soon
- HW 2 due today
- HW 3 posted

Reading

- Strang I.9

Outline

- Geometric interpretation of SVD
- Vector and matrix norms
- Eckart - Young Theorem
- Rayleigh Quotients

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

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

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

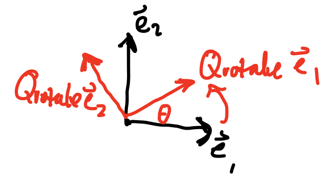
Geometric interpretation of SVD

Review An $n \times n$ matrix is orthogonal if its columns are orthonormal
- unit vectors
- orthogonal columns

For $n=2$, orthogonal matrices are rotation or reflections

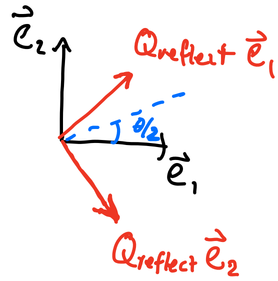
Rotation

$$Q_{rotate} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$ (rotation by angle θ)



Reflection

$$Q_{reflect} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$ (reflection across the $\frac{\theta}{2}$ line)



Consider a general 2×2 orthogonal matrix

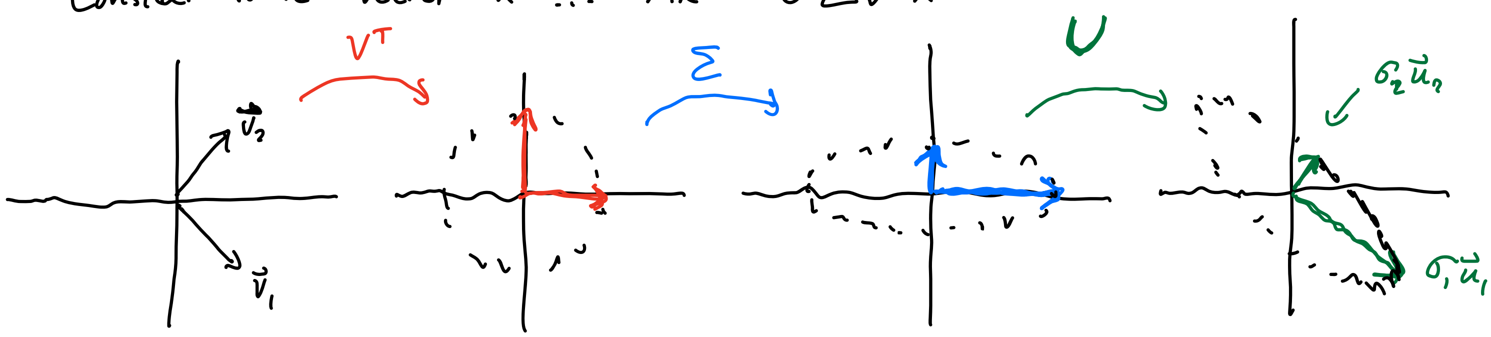
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad -1 \le a \le 1 \qquad \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ OR } \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

$\cos^2\theta + \sin^2\theta = 1$

For SVD we write $A = U\Sigma V^T$

$$A = \underbrace{U}_{\text{(rotation/reflection)}} \times \underbrace{\Sigma}_{\text{(stretching)}} \times \underbrace{V^T}_{\text{(rotation/reflection)}}$$

Consider some vector $\vec{x}$ ... $A\vec{x} = U\Sigma V^T\vec{x}$



$$A\vec{v}_1 = \sigma_1 \vec{u}_1 \qquad A\vec{v}_2 = \sigma_2 \vec{u}_2$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}$$

How does this change for symmetric matrices?

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

Review Vector norms

- $\|\vec{x}\|_1 = \sum_{i=1}^{n} |x_i|$ (1-norm)
- $\|\vec{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$ (2-norm)
- $\|\vec{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$ (p-norm)

Matrix norms

- For any vector norm, we can define an operator norm

def Let $\|\cdot\|$ be any vector norm. The corresponding operator norm is $\|A\| = \sup_{\vec{x}\neq\vec{0}} \frac{\|A\vec{x}\|}{\|\vec{x}\|}$

$$\|Q\| = \sup_{\vec{x}\neq\vec{0}} \frac{\|Q\vec{x}\|}{\|\vec{x}\|} = 1$$

note
- $\|A\| \geq 0$ for all A
- $\|A\| = 0$ if and only if $A = 0$ ✓
- $\|\alpha A\| = |\alpha| \, \|A\|$ for all real numbers α
- $\|A+B\| \leq \|A\| + \|B\|$ (triangle inequality)
- distance can be quantified $\|A-B\|$ for $m\times n$ matrices A and B

Common matrix norms

- $\|A\|_2 = \left[\max_{1\le i\le n} \lambda_i (A^TA)\right]^{1/2}$ (sq. root of the largest eigenvalue of A^TA)

  (note this is hard to compute!)

- $\|A\|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}|$ (max 1-norm of the rows of A)

- $\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{m} |a_{ij}|$ (max 1-norm of the columns of A)

- $\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2}$ (Frobenius norm) note $\|A\|_F^2 = tr(A^TA)$

Eckart-Young Theorem

If we compute $A = U\Sigma V^T = \sigma_1 \vec{u}_1 \vec{v}_1^T + \ldots + \sigma_r \vec{u}_r \vec{v}_r^T$

then $A_k = \sigma_1 \vec{u}_1 v_1^T + \ldots + \sigma_k \vec{u}_k \vec{v}_k^T$ will be the

"best" rank k approximation to A.

More precisely if B has rank k then $\|A - A_k\| \le \|A - B\|$

- This is true for $\|A\|_F$ (Frobenius) and $\|A\|_2$ (spectral norm or 2-norm)

- Note $\|A\|_F = \sqrt{\sigma_1^2 + \ldots + \sigma_r^2}$ $\quad \|I_n\|_F = \sqrt{n}$

  $\|A\|_2 = \sigma_1$ $\quad \|I_n\|_2 = 1$

- Eckart-Young applies to all norms that are computable using the singular value matrix Σ (Minsky 1955)

ex What is the rank-2 matrix closest to

$$A = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ is } A_2 = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Check

$\|A - A_2\|_2 = 2$

$\|A - A_2\|_F = \sqrt{5}$

note The set of rank 2 matrices is not convex.

(The average of two rank two matrices can have rank 4)

Proof in 2-norm/spectral norm.

- If B is $m \times n$ matrix with $\mathrm{rank}(B) \le k < r$, we want to show that $\|A - B\|_2 \geq \sigma_{k+1}$.

  Since $\|A - A_k\|_2 = \sigma_{k+1}$

- First choose $\vec{x} \neq \vec{0}$ so $B\vec{x} = \vec{0}$ and $\boxed{\vec{x} = \sum_{i=1}^{k+1} c_i \vec{v}_i}$ where $\vec{v}_1, \dots, \vec{v}_{k+1}$ are the first $k+1$ right singular vectors)

  (this is possible since $\dim(N(B)) \geq n - k$
  $\dim(\mathrm{span}(\vec{v}_1, \dots, \vec{v}_{k+1})) = k+1$

- Next $\|A - B\|_2 = \sup_{\vec{x} \neq \vec{0}} \dfrac{\|(A-B)\vec{x}\|_2}{\|\vec{x}\|_2}$ and

$$\|(A-B)\vec{x}\|_2 = \|A\vec{x} - B\vec{x}\|_2 = \|A\vec{x}\|_2 = \left\|A\left(\sum_{i=1}^{k+1} c_i \vec{v}_i\right)\right\|_2$$

$$= \left\|\sum_{i=1}^{k+1} c_i (A\vec{v}_i)\right\|_2 = \left\|\sum_{i=1}^{k+1} c_i \sigma_i \vec{u}_i\right\|_2 = \sqrt{\sum_{i=1}^{k+1} c_i^2 \sigma_i^2 \|\vec{u}_i\|_2} = \sqrt{\sum_{i=1}^{k+1} c_i^2 \sigma_i^2}$$

$$\geq \sqrt{\sum_{i=1}^{k+1} c_i^2 \sigma_{k+1}^2} = \sigma_{k+1} \sqrt{\sum_{i=1}^{k+1} c_i^2} = \sigma_{k+1} \|\vec{x}\|_2$$

So $\|(A-B)\vec{x}\|_2 \geq \sigma_{k+1} \|\vec{x}\|_2$

Therefore $\|A - B\|_2 = \sup_{\vec{x} \neq \vec{0}} \dfrac{\|(A-B)\vec{x}\|_2}{\|\vec{x}\|_2} \geq \dfrac{\sigma_{k+1} \|\vec{x}\|_2}{\|\vec{x}\|_2} = \sigma_{k+1}$