Hw3 es 2 point 1

23 November 2022 11:15

We could use a proof similar to the one done in the lecture or doing it a little bit differently

$$\|A_k\|_F^2 = \left\| \sum_{i=1}^{k} \sigma_i u_i v_i^T \right\|_F^2 = \| U \Sigma V^T \|_F^2 \Rightarrow$$

where $U = \begin{pmatrix} | & & | \\ u_i & \cdots & u_m \\ | & & | \end{pmatrix}$, $V^T = \begin{pmatrix} - v_i - \\ \vdots \\ - v_m - \end{pmatrix}$, $\Sigma = \begin{pmatrix} \sigma_1 & \cdots & & 0 \\ \vdots & \sigma_k & 0 & \vdots \\ 0 & \cdots & & 0 \end{pmatrix}$

now $\|A\|_F^2 = Tr(A^T A) \Rightarrow Tr\left((U\Sigma V^T)^T (U \Sigma V^T)\right) = Tr(V \Sigma^T U^T U \Sigma V^T) \Rightarrow$

Since U and V are orthonormal basis then $UU^T = U^T U = V^T V = VV^T = I$

$$\Rightarrow Tr(V \Sigma^T \Sigma V^T) \Rightarrow$$

Now the trace of the product is cyclic so $\Rightarrow Tr(V^T V \Sigma^T \Sigma) = Tr(\Sigma^T \Sigma) = \|\Sigma\|_F^2$

Since Σ is a matrix done in this way $\Sigma = \begin{pmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_k & & \\ & & & 0 & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$ Then the frobineaus norm is $\|\Sigma\|_F^2 = \sum_{i=1}^{k} \sigma_i^2$

$$\text{so } \|A_k\|_F^2 = \|\Sigma\|_F^2 = \sum_{i=1}^{k} \sigma_i^2$$

Hw3 es2 point 2

23 November 2022 10:34

We could use the proof done in the lecture or doing it a little bit differently

$$\| A - A_k \|_F^2 = \left\| \sum_{i=1}^{m} \sigma_i u_i v_i^T - \sum_{i=1}^{k} \sigma_i u_i v_i^T \right\|_F^2 = \left\| \sum_{i=k+1}^{m} \sigma_i u_i v_i^T \right\|_F^2 = \| U \Sigma V^T \|_F^2$$

where $U = \begin{pmatrix} | & & | \\ u_i & \cdots & u_m \\ | & & | \end{pmatrix}, V^T = \begin{pmatrix} - v_i - \\ \vdots \\ - v_m - \end{pmatrix}, \Sigma = \begin{pmatrix} 0 & & & \\ & \ddots 0 & & \\ & & \sigma_{k+1} & \\ & & & \ddots \sigma_m \end{pmatrix}$

now $\|A\|_F^2 = Tr(A^T A) \Rightarrow Tr\left((U \Sigma V^T)^T (U \Sigma V^T) \right) = Tr(V \Sigma^T U^T U \Sigma V^T) \Rightarrow$

Since U and V are orthonormal basis then $U U^T = U^T U = V^T V = V V^T = I$

$\Rightarrow Tr(V \Sigma^T \Sigma V^T) \Rightarrow$

Now the trace of the product is cyclic so $\Rightarrow Tr(V^T V \Sigma^T \Sigma) = Tr(\Sigma^T \Sigma) = \|\Sigma\|_F^2$

Since Σ is a matrix done in this way $\Sigma = \begin{pmatrix} 0 & & & \\ & \ddots 0 & & \\ & & \sigma_{k+1} & \\ & & & \ddots \sigma_m \end{pmatrix}$ Then the frobineaus norm is $\|\Sigma\|_F^2 = \sum_{i=k+1}^{m} \sigma_i^2$

$$\text{so } \|A - A_k\|_F^2 = \|\Sigma\|_F^2 = \sum_{i=k+1}^{m} \sigma_i^2$$

Hw3 es 2 point 3

23 November 2022 11:27

$\|A_k\|_2^2 =$ given the definition of 2 norm of a matrix $\max_{\|x\|_2 \neq 0} \frac{\|A_k x\|_2^2}{\|x\|_2^2} = \max_{\|x\|_2=1} \|A_k x\|_2^2 = \max_{\|x\|_2=1} \|U\Sigma V^T x\|_2^2 \Rightarrow$

where $U = \begin{pmatrix} | & & | \\ u_1 & \dots & u_m \\ | & & | \end{pmatrix}$, $V^T = \begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & v_m & - \end{pmatrix}$, $\Sigma = \begin{pmatrix} \sigma_1 & \dots & & 0 \\ \vdots & \sigma_k & & \vdots \\ & & 0 & \\ 0 & \dots & & 0 \end{pmatrix}$

$\Rightarrow \max_{\|x\|_2=1} \|(U\Sigma V^T x)^T (U\Sigma V^T x)\|_2 = \max_{\|x\|_2=1} \|x^T V \Sigma^T U^T U \Sigma V^T x\|_2 \Rightarrow$ since U is orthogonal $\max_{\|x\|_2=1} \|x^T V \Sigma^T \Sigma V^T x\|_2$

$\Rightarrow \max_{\|x\|_2=1} \|\Sigma V^T x\|_2^2$ then we define $z = V^T x \quad \|z\| = \|x\| = 1$ because V is orthogonal $\|z\|_2^2 = \|V^T x\|_2^2 = \|x^T V V^T x\|_2 = \|x^T x\| = \|x\|_2^2 = \|x\|_2 = 1$

$\Rightarrow \max_{\|z\|=1} \|\Sigma z\|_2^2 \Rightarrow$ since Σ is diagonal $\begin{pmatrix} \sigma_1 & \dots & & 0 \\ \vdots & \sigma_k & & \vdots \\ & & 0 & \\ 0 & \dots & & 0 \end{pmatrix} \Rightarrow$ the max is σ_1^2 when z is $(1 \; 0 \dots 0)$

given a diagonal matrix the max of the l2 norm is the maximum value of its diagonal because we can set to 1 the component of the vector z that multiplies the biggest element in the matrix in our case the sigma1 and the others to 0 given that z has to have norm = 1. We want the square of this value so sigma 1 squared

so $\|A_k\|_2^2 = \max_{\|z\|=1} \|\Sigma z\|_2^2 = \sigma_1^2$

Hw3 es 2 point 4

23 November 2022 11:28

$\|A - A_k\|_2^2 =$ given the definition of 2 norm of a matrix $\max_{\|x\|_2 \neq 0} \frac{\|(A-A_k)x\|_2^2}{\|x\|_2^2} = \max_{\|x\|_2 = 1} \|(A-A_k)x\|_2^2$

$$A - A_k = \sum_{i=1}^{n} \sigma_i u_i v_i^T - \sum_{i=1}^{k} \sigma_i u_i v_i^T = \sum_{i=k+1}^{n} \sigma_i u_i v_i^T = U\Sigma V^T$$

we want to evaluate $\max_{\|x\|_2=1} \|U\Sigma V^T x\|_2^2$ where $U = \begin{pmatrix} | & & | \\ u_1 & - & u_n \\ | & & | \end{pmatrix}$, $V^T = \begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & v_n & \end{pmatrix}$, $\Sigma = \begin{pmatrix} 0 \cdots 0 & & \cdots & 0 \\ \vdots & \sigma_{k+1} & & \vdots \\ 0 \cdots & & \ddots & \sigma_n \end{pmatrix}$

$\Rightarrow \max_{\|x\|_2=1} \|(U\Sigma V^T x)^T (U\Sigma V^T x)\|_2 = \max_{\|x\|_2=1} \|x^T V \Sigma^T U^T U \Sigma V^T x\|_2 \Rightarrow$ since U is orthogonal $\max_{\|x\|_2=1} \|x^T V \Sigma^T \Sigma V^T x\|_2$

$\Rightarrow \max_{\|x\|_2=1} \|\Sigma V^T x\|_2^2$ then we define $z = V^T x$ $\|z\| = \|x\| = 1$ because V is orthogonal $\|z\|_2^2 = \|V^T x\|_2^2 = \|x^T V V^T x\|_2 = \|x^T x\| = \|x\|_2^2 = \|x\|_2 = 1$

$\Rightarrow \max_{\|z\|=1} \|\Sigma z\|_2^2 \Rightarrow$ since Σ is diagonal $\begin{pmatrix} 0 \cdots 0 & & \cdots & 0 \\ & \sigma_{k+1} & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \sigma_n \end{pmatrix} \Rightarrow$ the max is σ_{k+1}^2 when z is $(1\ 0 \cdots 0)$

given a diagonal matrix the max of the l2 norm is the maximum value of its diagonal because we can set to 1 the component of the vector z that multiplies the biggest element in the matrix in our case the sigma1 and the others to 0 given that z has to have norm = 1. We want the square of this value so sigma 1 squared

$$\text{so } \|A_k\|_2^2 = \max_{\|z\|=1} \|\Sigma z\|_2^2 = \sigma_{k+1}^2$$

hw3 es3 point 1

25 November 2022 10:16

we have to prove that $\sigma_k \le \frac{\|A\|_F}{\sqrt{k}} \Rightarrow \|A\|_F^2 \ge k\,\sigma_k^2$

we know from the es 2 point 1 that $\|A\|_F^2 = \sum_{i=1}^{m} \sigma_i^2$ where σ_i are the singular values $\Rightarrow \sum_{i=1}^{m} \sigma_i^2 \ge k\,\sigma_k^2$

knowing that the singular values are ordered from the biggest to the lowest we can write $\sum_{i=1}^{m} \sigma_i^2 = \sum_{i=1}^{k} \sigma_i^2 + \sum_{i=k+1}^{m} \sigma_i^2 \ge k\,\sigma_k^2$

we know that $\sum_{i=1}^{k} \sigma_i^2 \ge k\,\sigma_k^2$ because the $\sigma_i \ge \sigma_k$ when $i \le k$ because of the SVD decomposition $A = U \Sigma V^T$

$$\Sigma = \begin{pmatrix} \sigma_1 & \cdots & & 0 \\ \vdots & \sigma_k & & \vdots \\ 0 & \cdots & & \sigma_m \end{pmatrix} \quad \sigma_{i+1} \ge \sigma_i \ \forall i \text{ in } 0,\dots,m-1 \text{ and } \sigma_i \ge 0 \ \forall i \text{ in } 0\dots m$$

so $\sum_{i=1}^{k} \sigma_i^2 \ge k \cdot \sigma_k^2 \Rightarrow \sum_{i=1}^{k} \sigma_i^2 + \sum_{i=k+1}^{m} \sigma_i^2 \ge k\,\sigma_k^2 \Rightarrow \|A\|_F^2 \ge k\,\sigma_k^2 \Rightarrow$

$$\|A\|_F \ge \sqrt{k}\,\sigma_k \Rightarrow \sigma_k \le \frac{\|A\|_F}{\sqrt{k}}$$

hw3 es 3 point 2

25 November 2022 10:16

we have to prove that there exists a matrix B of rank at most k such that $\|A-B\|_2 \le \frac{\|A\|_F}{\sqrt{k}} \Rightarrow$

$\|A\|_F^2 \ge k\,\|A-B\|_2^2$ from point 2.1 we know that we can rewrite $\|A\|_F^2$ as $\sum_{i=1}^{m} \sigma_i^2$

and if we take $B = A_k$ we can write $\sum_{i=1}^{n} \sigma_i^2 \ge k\,\|A-A_k\|_2^2 \Rightarrow \|A-A_k\|_2^2$ from point 2.4 is $\sigma_{k+1}^2 \Rightarrow$

$\sum_{i=1}^{m} \sigma_i^2 \ge k\,\sigma_{k+1}^2$ and this relation is verified because the eigenvalues are ordered and greater than 0. Note also that B is a matrix with rank k for the construction using SVD.

$\Downarrow$

$$\sum_{i=1}^{k} \sigma_i^2 + \sum_{i=k+1}^{m} \sigma_i^2 \ge k\,\sigma_{k+1}^2 \Rightarrow \sum_{i=1}^{k} \sigma_i \ge k\,\sigma_{k+1}^2 \text{ so } \sum_{i=1}^{k} \sigma_i^2 + \sum_{i=k+1}^{m} \sigma_i^2 \ge k\,\sigma_{k+1}^2$$

$$\Rightarrow \sum_{i=1}^{m} \sigma_i^2 \ge k\,\sigma_{k+1}^2 \Rightarrow \|A\|_F^2 \ge k\,\|A-A_k\|_2^2 \Rightarrow \|A-A_k\|_2 \le \frac{\|A\|_F}{\sqrt{k}}$$

so there exist a matrix B= Ak of rank k such that the relationship is verified