Lecture 36: Quick review from previous lecture

- Full SVD for a matrix:
  Let $A$ be an $m \times n$ matrix of rank $r$ with the positive singular values
  $$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r,$$
  and let $\Sigma$ be the $m \times n$ matrix defined by

  $$\Sigma = \begin{bmatrix}
    \sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\
    0 & \sigma_2 & \cdots & 0 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & 0 & \cdots & 0 \\
    0 & 0 & 0 & \sigma_r & \cdots & 0 \\
    \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
    0 & \cdots & \cdots & 0 & \cdots & 0
  \end{bmatrix}_{m \times n}$$

  Then there exist an $m \times m$ orthogonal matrix $U$ and an $n \times n$ orthogonal matrix $V$ such that

  $$A = U \Sigma V^T. \quad (\text{Full SVD})$$

- Suppose that $A \in M_{m \times n}$, $b \in \mathbb{R}^m$. The least squares problem is to find $x \in \mathbb{R}^n$ for which that $\|Ax - b\|$ is minimized.

  A vector $x$ that minimizes $\|Ax - b\|$ is called the least squares solution.

Today we will discuss Singular Value Decomposition.

- Lecture will be recorded -

- Information for Final Exam and Course Grade has been posted on Canvas, see "Announcements".
Let’s recap. Even if $Ax = b$ has no exact solution, we can always find a least squares solution $x^* = A^+b$ that minimizes the error $\|Ax - b\|^2$.

Recall that the pseudoinverse of a nonzero $m \times n$ matrix $A$ with SVD $A = U\Sigma V^T$ is the $n \times m$ matrix $A^+ = V\Sigma^+ U^T$.

Recalling: the least squares solution must satisfy the equation $Ax = b_r$.

Check: $x^* = A^+ b$ satisfies $Ax = b_r$.

[To see this] $x^* = A^+ b = V \Sigma^+ U^T b = V \begin{bmatrix} \frac{1}{\sigma_1} & \cdots & \frac{1}{\sigma_r} & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} u_1^T \\ \vdots \\ u_r^T \\ 0 \\ \cdots \\ 0 \end{bmatrix} b$

$= V \begin{bmatrix} \frac{1}{\sigma_1} & \cdots & \frac{1}{\sigma_r} & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \langle b, u_1 \rangle \\ \vdots \\ \langle b, u_r \rangle \\ 0 \\ \cdots \\ 0 \end{bmatrix} = U \begin{bmatrix} \langle b, u_1 \rangle \\ \vdots \\ \langle b, u_r \rangle \end{bmatrix}$.
Remark: The least squares solution $x^*$ to the system $Ax = b$ satisfies the normal equation $A^TAx - A^Tb = 0$.

Fact: Suppose that $A \in M_{m \times n}$, $b \in \mathbb{R}^m$. Let $x^* = A^+ b$. Then $x^*$ is the least squares solution to the linear system $Ax = b$. If $A$ has $n$ linearly independent columns ($\ker A = \{0\}$), then

$$x^* = A^+ b = (A^T A)^{-1} A^T b,$$

where

$$A^+ = (A^T A)^{-1} A^T.$$

Example. Consider the linear system

$$\begin{cases} x + y - z = 1, \\ x + y - z = 1. \end{cases}$$

Find the best approximation to a solution having minimum norm.

We write $A \hat{x} = b$, where $A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \end{pmatrix}$, $b = (1, 1)$. In lecture 35, we have found SVD of $A$

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}^T.$$

So, pseudoinverse of $A$ is

$$A^+ = U \Sigma^+ U^T = \frac{1}{6} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

So, $x^* = A^+ b = A^+ (1, 1) = \frac{1}{6} (1, 1)$. 

\[ MATH 4242-Week 15-1 \]
§ Low rank approximations to a matrix.
Suppose we want to approximate a matrix $A = A_{m \times n}$ with rank $r$ by a matrix $B = B_{m \times n}$ with rank $k < r$. That is, we want such $B$ to minimize $\|A - B\|$.

Recall a matrix $A$ with rank $r$ has full SVD as follows:

$$A = U \Sigma V^T = [u_1, \ldots, u_m] \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_r \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_r^T \end{bmatrix}, \quad \sigma_1 \geq \cdots \geq \sigma_r > 0.$$

If we take $A_k = [u_1, \ldots, u_m] \begin{bmatrix} \sigma_1 & \cdots & \sigma_k & 0 \\ & \sigma_{k+1} & \cdots & \sigma_r \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_k^T \end{bmatrix}$, then the matrix $A_k$ has rank $k$.

The best rank $k$ approximation to $A$ is

$$B = U \Sigma_k V^T.$$

$$= U \begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} V^T, \quad k < n.$$

**Fact:** This matrix $B$ minimizes the distance to $A$ as measured by Frobenius norm and operator norm:

$$\|A - B\|_F, \quad \|A - B\|_2.$$

Recall: $\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$; $\|A\|_2 = \sigma_1$ (largest singular value of $A$).

$$A - B = U \Sigma V^T - U \Sigma_k V^T = U \left( \Sigma - \Sigma_k \right) V^T$$

$$= U \left( \begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \begin{bmatrix} \sigma_1 \\ \vdots \\ \sigma_k \end{bmatrix} \right) V^T$$

So, the largest singular value of $A - B$ is $\sigma_{k+1}$, thus $\|A - B\|_2 = \sigma_{k+1}$. Moreover, $\|A - B\|_F = \sqrt{\sigma_{k+1}^2 + \cdots + \sigma_r^2}$. \#
Example. Find the best rank 1 approximation to $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Find SVD of $A$. $A^T A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. $0 = \det(A^T A - \lambda I)$

$\lambda = 3$: $A^T A - 3 I$. $v_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. $U = [ v_1, v_2 ]$.

$\lambda = 1$: $A^T A - I$. $v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

$u_1 = \frac{Av_1}{\sqrt{3}} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$; $u_2 = \frac{Av_2}{\sqrt{1}} = \frac{1}{\sqrt{1}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$

Finding $u_3$ that is orthogonal to $u_1$, $u_2$, so $u_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$.

Thus, full SVD of $A$ is

$A = U \Sigma V^T = [u_1, u_2, u_3] \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 0 \end{pmatrix} [v_1, v_2]^T$.

Reduced SVD of $A$ is

$A_{3 \times 2} = \hat{U} \hat{\Sigma} \hat{V}^T = [u_1, u_2] \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 0 \end{pmatrix} [v_1, v_2]^T$.

So, $B = \hat{U} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 0 \end{pmatrix} \hat{V} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}$. 

Exercise: $\| A - B \|_F = $