Module 9

Applications of SVD
Solutions

Problem 1: 10 points

Consider the overdetermined system in the single variable x:

$$a_1x = b_1, \ldots, a_mx = b_m,$$

where $a_i \neq 0$ for some $i \in \{1, \ldots, m\}$. Prove that the least squares solution of smallest norm is given by

$$x^+ = \frac{a_1b_1 + \cdots + a_mb_m}{a_1^2 + \cdots + a_m^2}.$$

Solution. Let A and b be $m \times 1$ matrices

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

In matrix form the system

$$a_1x = b_1, \ldots, a_mx = b_m$$

is expressed as $Ax = b$. We know from the lesson on least squares that the least squares solution of smallest norm of the system $Ax = b$ is a solution of the normal equations

$$A^\top Ax = A^\top b.$$

However

$$A^\top A = a_1^2 + \cdots + a_m^2 \neq 0$$

since by hypothesis $a_i \neq 0$ for some $i \in \{1, \ldots, m\}$ and $A^\top b = a_1b_1 + \cdots + a_mb_m$, so the normal equations have the unique solution

$$x^+ = (A^\top A)^{-1}A^\top b = \frac{a_1b_1 + \cdots + a_mb_m}{a_1^2 + \cdots + a_m^2},$$

which is least squares solution of smallest norm of the system $Ax = b$.

Problem 2: 5 points

Use Matlab to find the pseudo-inverse of the 8×6 matrix

$$A = \begin{pmatrix} 64 & 2 & 3 & 61 & 60 & 6 \\ 9 & 55 & 54 & 12 & 13 & 51 \\ 17 & 47 & 46 & 20 & 21 & 43 \\ 40 & 26 & 27 & 37 & 36 & 30 \\ 32 & 34 & 35 & 29 & 28 & 38 \\ 41 & 23 & 22 & 44 & 45 & 19 \\ 49 & 15 & 14 & 52 & 53 & 11 \\ 8 & 58 & 59 & 5 & 4 & 62 \end{pmatrix}.$$

Observe that the sums of the columns are all equal to to 260. Let b be the vector of dimension 8 whose coordinates are all equal to 256. Find the solution x^+ of the system $Ax = b$.

Solution. The above 8×6 matrix is obtained in Matlab by finding an 8×8 magic square and then dropping its last two columns using the commands:

```
A = magic(8); A = A(:, 1:6).
```

Then use the following code to find $x^+ = A^+ b$:

```
b = 256*one(8,1); x = pinv(A)*b
```

and you get

$$x^+ = \begin{pmatrix} 1.1361 \\ 1.4391 \\ 1.3633 \\ 1.3633 \\ 1.4391 \\ 1.1361 \end{pmatrix}.$$

Problem 3: 35 points

(1) (20 points) Prove (without using Matlab) that the closest rank 1 approximation (in $\| \ \|_2$) of the matrix

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

is

$$A_1 = \frac{3}{2} \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix}.$$

1) (15 points) Show that the Eckart–Young theorem fails for the operator norm $\| \ \|_\infty$ by finding a rank 1 matrix B such that $\|A - B\|_\infty < \|A - A_1\|_\infty$.

Solution. First we need to find an SVD $A = V\Sigma U^\top$ for A. For this we first compute the singular values σ_1, σ_2 of A, which are the nonnegative square roots of the eigenvalues of $A^\top A$. We have

$$A^\top A = \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} 25 & 20 \\ 20 & 25 \end{pmatrix}.$$

The characteristic polynomial of $A^\top A$ is

$$\det(\lambda I - A^\top A) = \begin{vmatrix} \lambda - 25 & -20 \\ -20 & \lambda - 25 \end{vmatrix} = \lambda^2 - 50\lambda + 225.$$

The zeros of the equation

$$\lambda^2 - 50\lambda + 225 = 0$$

are

$$\lambda = \frac{50 \pm \sqrt{50^2 - 4 \times 225}}{2} = \frac{50 \pm \sqrt{2500 - 900}}{2} = \frac{50 \pm 40}{2} = 45, 5.$$

The singular values of A are $\sigma_1 = 3\sqrt{5}$ and $\sigma_2 = \sqrt{5}$. Next we need to diagonalize $A^\top A$. To find an eigenvector associated with the eigenvalue 45 we solve the system $A^\top A u = 45u$ with $u = \begin{pmatrix} x \\ y \end{pmatrix}$, but since this system is singular it suffices to solve the first equation

$$25x + 20y = 45x,$$

that is, $20y = 20x$, so we can pick the unit vector

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Similarly, to find an eigenvector associated with the eigenvalue 5 we solve the equation

$$25x + 20y = 5x,$$

that is, $20y = -20x$, so we can pick the unit vector

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Consequently we obtain the orthogonal matrix

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

and we have

$$A^\top A = \begin{pmatrix} 25 & 20 \\ 20 & 25 \end{pmatrix} = U\Sigma^2 U^\top = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 45 & 0 \\ 0 & 5 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

as we easily verify. To find V we compute AU, that is,

$$AU = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 3 & 3 \\ 9 & -1 \end{pmatrix}.$$

To obtain V, we normalize the column of AU and we get

$$V = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix}.$$

Thus we have the SVD

$$A = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix} = V\Sigma U^\top = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 3\sqrt{5} & 0 \\ 0 & \sqrt{5} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

as we easily verify.

By Lesson 3 of Week 9 (also Proposition 21.9 of Vol I), the closest rank 1 approximation A_1 to A is obtained by setting the smaller singular value σ_1 to 0, so we get

$$A_1 = V\Sigma U^\top = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 3\sqrt{5} & 0 \\ 0 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{3}{2} \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix}.$$

According to Lesson 3 from Week 4, for any matrix C,

$$\|C\|_\infty = \max_i \sum_{j=1}^n |c_{ij}|.$$

Applying the above to $C = A - A_1$, since

$$A - A_1 = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix} - \frac{3}{2} \begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix} = \begin{pmatrix} 3/2 & -3/2 \\ -1/2 & 1/2 \end{pmatrix},$$

we find that

$$\|A - A_1\|_\infty = \max\left\{\frac{3}{2} + \frac{3}{2},\ \frac{1}{2} + \frac{1}{2}\right\} = \max\{3,\ 1\} = 3.$$

We are now looking for a rank 1 matrix B such that $\|A - B\|_\infty < 3$. If we pick

$$B = \begin{pmatrix} 3/2 & 4/3 \\ 9/2 & 4 \end{pmatrix},$$

we have

$$A - B = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix} - \begin{pmatrix} 3/2 & 4/3 \\ 9/2 & 4 \end{pmatrix} = \begin{pmatrix} 3/2 & -4/3 \\ -1/2 & 1 \end{pmatrix},$$

and we have

$$\|A - B\|_\infty = \max\left\{\frac{3}{2} + \frac{4}{3},\ \frac{1}{2} + 1\right\} = \max\left\{\frac{17}{6},\ \frac{3}{2}\right\} = \frac{17}{6} < 3,$$

as desired.

A more systematic method is to look for a matrix B whose first column is equal to the first column of A, say

$$B = \begin{pmatrix} 3 & 3a \\ 4 & 4a \end{pmatrix}$$

for some $a \in \mathbb{R}$. We have

$$A - B = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix} - \begin{pmatrix} 3 & 3a \\ 4 & 4a \end{pmatrix} = \begin{pmatrix} 0 & -3a \\ 0 & 5-4a \end{pmatrix},$$

so

$$\|A - B\|_\infty = \max\{|-3a|,\ |5-4a|\}.$$

In order to have $\|A - B\|_\infty < 3$, we must have $a > 0$, and then $3a < 3$ and $5-4a < 3$, namely $a < 1$ and $2 < 4a$, that is, $1/2 < a < 1$. For all such matrices, we have $\|A - B\|_\infty < 3$. In particular, this holds for $a = 3/4$.

These counterexamples shows that the Eckart–Young theorem fails for the operator norm $\| \, \|_\infty$

Problem 4: 10 points

Let S be a real symmetric positive definite matrix and let $S = U\Sigma U^\top$ be a diagonalization of S. Prove that the closest rank 1 matrix (in the 2-norm) to S is $u_1\sigma_1 u_1^\top$, where u_1 is the first column of U.

Solution. Recall that the singular values of a matrix A are the nonnegative square roots of the eigenvalues of $A^\top A$. If $A = S$ is symmetric, then $S^\top S = S^2$. Since S is also positive definite, its eigenvalues are strictly positive, so if $S = U\Sigma U^\top$ is a diagonalization of S with $\sigma_1 \geq \cdots \geq \sigma_n > 0$, then $S^2 = U\Sigma^2 U^\top$ is diagonalization of S^2. Since Σ consists of positive entries $\sigma_1, \ldots, \sigma_n$, these are the positive square roots of $\sigma_1^2, \ldots, \sigma_n^2$, the eigenvalues of S^2, so $\sigma_1, \ldots, \sigma_n$ are also the singular values of S. It follows that $S = U\Sigma U^\top$ is also an SVD of S. Then we know by Lesson 3 of Week 9 (also Proposition 21.9 of Vol I) that a closest rank 1 approximation S_1 of S (in the 2-norm) is obtained by setting $\sigma_2 = \cdots = \sigma_n = 0$, and since the SVD of S can also be expressed in terms of the columns $u_1, \ldots, u_n$ of U as

$$S = u_1\sigma_1 u_1^\top + u_2\sigma_2 u_2^\top + \cdots + u_n\sigma_n u_n^\top,$$

we obtain

$$S_1 = u_1\sigma_1 u_1^\top.$$

Total: 60 points