INTRODUCTION TO LINEAR ALGEBRA

Sixth Edition

SOLUTIONS TO PROBLEM SETS

Gilbert Strang

Massachusetts Institute of Technology

math.mit.edu/weborder.php (orders)

math.mit.edu/linearalgebra (book website)

math.mit.edu/$\sim$gs (author website)

www.wellesleycambridge.com (all books)

Wellesley - Cambridge Press

Box 812060

Wellesley, Massachusetts 02482

Problem Set 7.1, page 295

1 $A^{\mathrm{T}}A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 64 \end{bmatrix} \qquad AA^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 64 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ give $\sigma_1 = 8$ and $\sigma_2 = 1$.

$\boldsymbol{v}_1 = (0,0,1), \quad \boldsymbol{v}_2 = (0,1,0), \quad \boldsymbol{u}_1 = (0,1,0), \quad \boldsymbol{u}_2 = (1,0,4)$. After removing row 3 of A and column 3 of A^{T}, $\begin{bmatrix} 1 & 0 \\ 0 & 64 \end{bmatrix}$ still has $\sigma_1^2 = 64$ and $\sigma_2^2 = 1$.

2 $\det(B - \lambda I) = -\lambda^3 + \frac{1}{125} = 0$ gives $\lambda = \frac{1}{5}$ times 1 and $e^{2\pi i/3}$ and $e^{4\pi i/3}$. The singular values are $\sigma = 8$ and 1 and $1/1000$. So λ changed by $1/5$ and σ only changed by $1/1000$.

3 A^{T} has the same singular values as A, and the singular vectors change from $A\boldsymbol{v} = \sigma\boldsymbol{u}$ to $A\boldsymbol{u} = \sigma\boldsymbol{v}$.

4 $\begin{bmatrix} 0 & A \\ A^{\mathrm{T}} & 0 \end{bmatrix}\begin{bmatrix} \boldsymbol{u}_k \\ \boldsymbol{v}_k \end{bmatrix} = \begin{bmatrix} A\boldsymbol{v}_k \\ A^{\mathrm{T}}\boldsymbol{u}_k \end{bmatrix} = \sigma_k \begin{bmatrix} \boldsymbol{u}_k \\ \boldsymbol{v}_k \end{bmatrix}$ and $\begin{bmatrix} 0 & A \\ A^{\mathrm{T}} & 0 \end{bmatrix}\begin{bmatrix} -\boldsymbol{u}_k \\ \boldsymbol{v}_k \end{bmatrix} = \begin{bmatrix} A\boldsymbol{v}_k \\ -A^{\mathrm{T}}\boldsymbol{u}_k \end{bmatrix} = -\sigma_k \begin{bmatrix} -\boldsymbol{u}_k \\ \boldsymbol{v}_k \end{bmatrix}$

So this one symmetric matrix S reveals the $\boldsymbol{u}$'s and $\boldsymbol{v}$'s and σ's in the SVD of A.

5 $A^{\mathrm{T}}A$ is symmetric with $\lambda_1 = 25$ and $\lambda_2 = 0$ so A has $\sigma_1 = 5$. The eigenvectors of $A^{\mathrm{T}}A$ are $\boldsymbol{v}_1 = (2,1)$ and $\boldsymbol{v}_2 = (-1,2)$: *orthogonal*. They are the $\boldsymbol{v}$'s in $A = U\Sigma V^{\mathrm{T}}$.

6 $A_1A_1^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ produces $\lambda^2 - 3\lambda + 1 = 0$ and $\lambda = \dfrac{1}{2}\left(3 \pm \sqrt{5}\right)$. The singular values are the square roots $\sigma = \dfrac{1}{2}\left(\sqrt{5} \pm 1\right)$.

$A_2A_2^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$ has $\lambda^2 - 6\lambda + 4 = 0$ and

$\lambda = \dfrac{1}{2}\left(6 \pm \sqrt{20}\right) = 3 \pm \sqrt{5}$. The singular values are the square roots $\sigma = \dfrac{\sqrt{2}}{2}\left(\sqrt{5} \pm 1\right)$.

For the singular vectors I recommend the SVD commands in MATLAB or Julia or Mathematica.

7 There are 20 singular values because a random 20 by 40 matrix almost surely has rank 20.

8 (a) The singular values of $A + I$ are square roots of eigenvalues of $(A + I)^{\mathrm{T}}(A + I)$. They are **not** eigenvalues of $A^{\mathrm{T}}A + I$.

(b) **This formula $V\Sigma^{-1}U^{\mathrm{T}}$ is the best way to compute the pseudoinverse A^+.** We could check the four Penrose conditions on A^+ from Section 4.5. For example

$$AA^+A = \left(\Sigma\, \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}}\right)\left(\Sigma\, \boldsymbol{v}_j \boldsymbol{u}_j^{\mathrm{T}}/\sigma_j\right)\left(\Sigma\, \sigma_k \boldsymbol{u}_k \boldsymbol{v}_k^{\mathrm{T}}\right) = \Sigma\, \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}}.$$

Notice also that $AA^+ = \Sigma\, \boldsymbol{u}_i \boldsymbol{v}_i^{\mathrm{T}} \boldsymbol{v}_i \boldsymbol{u}_i = \Sigma\, \boldsymbol{u}_i^{\mathrm{T}} \boldsymbol{u}_i = U^{\mathrm{T}}U =$ **projection**.

9 The singular values of Q are the positive square roots of eigenvalues of $Q^{\mathrm{T}}Q$—and all those eigenvalues are 1 because $Q^{\mathrm{T}}Q = I$ when Q is orthogonal.

10 If the λ's are in descending order, the maximum of $R(\boldsymbol{x}) = (\lambda_1 c_1^2 + \cdots + \lambda_n c_n^2)/(c_1^2 + \cdots + c_n^2)$ is λ_1 (when $\boldsymbol{x} = \boldsymbol{v}_1$). Then $c_1, c_2, \ldots, c_n$ is $1, 0, \ldots, 0$. The minimum is $R(\boldsymbol{x}) = \lambda_n$ when $\boldsymbol{x} = \boldsymbol{v}_n$ and $\boldsymbol{c} = (0, 0, \ldots, 0, 1)$.

11 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{v}_1 = 0$ means that the coefficient is $\boldsymbol{c_1 = 0}$ in $\boldsymbol{x} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n$. Then $\max \dfrac{\lambda_2 c_2^2 + \cdots + \lambda_n c_n^2}{c_2^2 + \cdots + c_n^2} = \boldsymbol{\lambda_2}$.

12 The first matrix has $A^{\mathrm{T}}A = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$ with $\lambda = 8$ and $\lambda = 2$. The eigenvectors of $A^{\mathrm{T}}A$ = right singular vectors $\boldsymbol{v}_1, \boldsymbol{v}_2$ of A are $(1, 1)/\sqrt{2}$ and $(1, -1)/\sqrt{2}$. The left singular vectors are $\boldsymbol{u} = A\boldsymbol{v}/\sigma = (4, 0)/\sqrt{2}\sqrt{8} = (1, 0)$ and $(0, 2)/\sqrt{2}\sqrt{2} = (0, 1)$.

The second matrix has $A^{\mathrm{T}}A = \begin{bmatrix} 25 & 25 \\ 25 & 25 \end{bmatrix}$ so $\lambda = 50$ and $\lambda = 0$. The right singular vectors of A are again $\boldsymbol{v}_1 = (1, 1)/\sqrt{2}$ with $\sigma_1 = \sqrt{50}$ and $\boldsymbol{v}_2 = (1, -1)/\sqrt{2}$ with no σ_2 (or you could say $\sigma_2 = 0$ but our convention is no σ_2). Then $\boldsymbol{u}_1 = A\boldsymbol{v}_1/\sqrt{50} = (3, 4)/5$.

13 This matrix has $A^{\mathrm{T}}A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ with eigenvalues $\lambda = 3, 1, 0$ and $\sigma_1 = \sqrt{3}$ and $\sigma_2 = 1$ and no σ_3. The eigenvectors of $A^{\mathrm{T}}A$ are $\boldsymbol{v}_1 = (1, 2, 1)/\sqrt{6}$ and

$\boldsymbol{v}_2 = (1, 0, -1)/\sqrt{2}$ and $\boldsymbol{v}_3 = (1, -1, 1)/\sqrt{3}$. Then $A\boldsymbol{v} = \sigma\boldsymbol{u}$ gives $\boldsymbol{u} = (1, 1)/\sqrt{2}$ and $\boldsymbol{u}_2 = (1, -1)/\sqrt{2}$.

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{matrix} /\sqrt{6} \\ /\sqrt{2} \end{matrix}$$

14 This small question is a key to everything. It is based on the associative law $(AA^{\mathrm{T}})A = A(A^{\mathrm{T}}A)$. Here we are applying both sides to an eigenvector $\boldsymbol{v}$ of $A^{\mathrm{T}}A$:

$$(AA^{\mathrm{T}})A\boldsymbol{v} = A(A^{\mathrm{T}}A)\boldsymbol{v} = A\lambda\boldsymbol{v} = \lambda A\boldsymbol{v}.$$

So $\boldsymbol{Av}$ **is an eigenvector of** $\boldsymbol{AA}^{\mathrm{T}}$ with the same eigenvalue $\lambda = \sigma^2$.

15 $$A = U\Sigma V^{\mathrm{T}} = \begin{bmatrix} \\ \boldsymbol{u}_1 & \boldsymbol{u}_2 \\ \\ \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & 0 \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{v}_1 & \boldsymbol{v}_2 \\ \\ \end{bmatrix}^{\mathrm{T}} = \frac{\begin{bmatrix} 1 & 3 \\ 3 & -1 \end{bmatrix}}{\sqrt{10}} \begin{bmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{bmatrix} \frac{\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}}{\sqrt{5}}$$

This $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ is a 2 by 2 matrix of rank 1. Its row space has basis $\boldsymbol{v}_1$, its nullspace has basis $\boldsymbol{v}_2$, its column space has basis $\boldsymbol{u}_1$, its left nullspace has basis $\boldsymbol{u}_2$:

$$\text{Row space} \quad \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{Nullspace} \quad \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$\text{Column space} \quad \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \mathbf{N}(A^{\mathrm{T}}) \quad \frac{1}{\sqrt{10}} \begin{bmatrix} 3 \\ -1 \end{bmatrix}.$$

16 (a) The main diagonal of $A^{\mathrm{T}}A$ contains the squared lengths $||\text{row } 1||^2, \cdots, ||\text{row } m||^2$. So the trace of $A^{\mathrm{T}}A$ is the sum of all a_{ij}^2.

(b) If A has rank 1, then $A^{\mathrm{T}}A$ has rank 1. So the only singular value of A is $\sigma_1 = (\text{trace } A^{\mathrm{T}}A)^{1/2}$.

17 The number $\sigma_{\max}(A^{-1})\sigma_{\max}(A)$ is the same as $\sigma_{\max}(A)/\sigma_{\min}(A)$. This is ≥ 1. It equals 1 if all σ's are equal, and $A = U\Sigma V^{\mathrm{T}}$ is a multiple of an orthogonal matrix. The ratio $\sigma_{\max}/\sigma_{\min}$ is the important **condition number** of A.

18 The smallest change in A is to set its smallest singular value σ_2 to zero.

Problem Set 7.2, page 301

1 (a) Suppose the identity matrix I is N by N, and an N by N approximating matrix A has rank $r < N$. Then $I - A$ will have $N - r$ eigenvalues equal to 1, meaning that **the error norm** $||\boldsymbol{I} - \boldsymbol{A}||$ **is at least** $\mathbf{1}$, and I is impossible to compress by a lower rank matrix.

(b) A matrix with a horizontal-vertical cross looks like A :

$$\begin{bmatrix} \text{zeros} & \text{ones} & \text{zeros} \\ \text{ones} & \text{ones} & \text{ones} \\ \text{zeros} & \text{ones} & \text{zeros} \end{bmatrix} = \begin{bmatrix} \text{zeros} & \text{ones} & \text{zeros} \\ \text{zeros} & \text{ones} & \text{zeros} \\ \text{zeros} & \text{ones} & \text{zeros} \end{bmatrix} + \begin{bmatrix} \text{zeros} & \text{zeros} & \text{zeros} \\ \text{ones} & \text{zeros} & \text{ones} \\ \text{zeros} & \text{zeros} & \text{zeros} \end{bmatrix}$$

Those are both rank one matrices (all nonzero rows equal) so $\boldsymbol{A}$ **has rank 2**.

2 $$A = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 4 & 2 & 2 \\ 1 & 2 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & -2 & 0 & 0 \end{bmatrix}$$

and the rank is 2.

$$B = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \end{bmatrix}$$ also has rank 2.

3 $$BB^{\mathrm{T}} = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 9 & 13 \\ 13 & 19 \end{bmatrix}$$ trace $= \mathbf{28}$ and $\det = \mathbf{2}$.

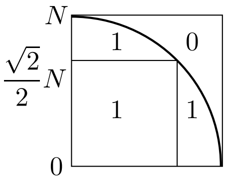
$$B^{\mathrm{T}}B = \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 5 & 5 \\ 5 & 13 & 13 \\ 5 & 13 & 13 \end{bmatrix}$$ trace $= \mathbf{28}$ and $\det = \mathbf{0}$.

The 2 nonzero eigenvalues must be the same for both matrices. They are $\sigma_1,\ \sigma_2 = 14 \pm \sqrt{14^2 - 2}$. I would call B compressible when σ_2 is so much smaller than σ_1.

4 (computer question $\mathsf{svd}(A)$).

5 The Japanese flag has a circle filled by 1's, with diameter $= 2N$ 1's. Outside the circle are zeros. The rank is approximately CN. What is the number C ? Alex Townsend contributed this key idea: The circle contains a big square matrix filled by 1's. The rank of that all-ones matrix is only 1.

So we only have to count the rows above and below that square! Multiply by 2 to include the columns to the left and right of the square.



The picture shows $\left(1-\frac{\sqrt{2}}{2}\right)N$ rows of 1's above the square—and repeated below the square. It also shows $\left(1-\frac{\sqrt{2}}{2}\right)N$ columns of 1's to the right of the square—and repeated to the left.

Combined, those $(2-\sqrt{2})N$ rows and columns (plus 1 for the big square) tell us the rank of this $2N$ by $2N$ Japanese flag containing the red circle.

6 The N by N matrix A is filled by the values $A_{ij} = F(i/N, j/N)$ of the two-variable function $F(x,y)$, by taking the points $(x,y) = (i/N, j/N)$ on a uniform square grid (x and y go from 0 to 1). Three choices of that function F:

1) $F = xy$ produces a symmetric **rank-1 matrix**. Its i,j entry is a multiple of the product i times j. All rows of F contain a multiple of the vector $(1, 2, \ldots, N)$.

2) $F_2 = x + y$ gives a sum of 2 rank-one matrices (**the rank is 2**). One matrix has constants along each row. The other has constants down each column.

3) $F_3 = (x,y) = x^2 + y^2$ will also produce a sum of constant rows (from x^2) and constant columns (from y^2). Again rank $= 2$.

7 Symmetric matrix S if $F(x,y) = F(y,x)$. Example $F = x + y$.

Antisymmetric matrix A if $F(x,y) = -F(y,x)$. Example $F = x - y$.

Matrix of rank 2 if $F(x,y) = F(x) + F(y)$ (and other possibilities too ?)

Singular matrix M from a sum of less than n rank-one matrices (please expand this part of the answer).

Problem Set 7.3, page 307

1 The row averages of A_0 are 3 and 0. Therefore

$$A = \begin{bmatrix} 2 & 1 & 0 & -1 & -2 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix} \quad \text{and} \quad S = \frac{AA^{\mathrm{T}}}{4} = \frac{1}{4}\begin{bmatrix} 10 & 0 \\ 0 & 4 \end{bmatrix}$$

The eigenvalues of S are $\lambda_1 = \dfrac{10}{4}$ and $\lambda_2 = \dfrac{4}{4} = 1$. The top eigenvector of S is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. I think this means that a **horizontal line** (the x axis) is closer to the five points $(2,-1), \ldots, (-2,-1)$ in the columns of A than any other line through the origin $(0,0)$.

2 Now the row averages of A_0 are $\frac{1}{2}$ and 2. Therefore

$$A = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -1 & 0 & 1 & 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad S = \frac{AA^{\mathrm{T}}}{5} = \frac{1}{5}\begin{bmatrix} \frac{3}{2} & 0 \\ 0 & 4 \end{bmatrix}.$$

Again the rows of A are accidentally orthogonal (because of the special patterns of those rows). This time the top eigenvector of S is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. So a **horizontal line** is closer to the six points $\left(\frac{1}{2},-1\right), \ldots, \left(-\frac{1}{2},-1\right)$ from the columns of A than any other line through the center point $(0,0)$.

3 $A_0 = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 2 & 2 \end{bmatrix}$ has row averages 2 and 3 so $A = \begin{bmatrix} -1 & 0 & 1 \\ 2 & -1 & -1 \end{bmatrix}$.

Then $S = \dfrac{1}{2}AA^{\mathrm{T}} = \dfrac{1}{2}\begin{bmatrix} 2 & -3 \\ -3 & 6 \end{bmatrix}$.

Then trace $(S) = \frac{1}{2}(8)$ and $\det(S) = \left(\frac{1}{2}\right)^2(3)$. The eigenvalues $\lambda(S)$ are $\frac{1}{2}$ times the roots of $\lambda^2 - 8\lambda + 3 = 0$. Those roots are $4 \pm \sqrt{16-3}$. Then the σ's are $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$.

4 This matrix A with orthogonal rows has $S = \dfrac{AA^{\mathrm{T}}}{n-1} = \dfrac{1}{3}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

With λ's in descending order $\lambda_1 > \lambda_2 > \lambda_3$, the eigenvectors are $(0, 1, 0)$ and $(0, 0, 1)$ and $(1, 0, 0)$. The first eigenvector shows the $\boldsymbol{u}_1$ direction $= y$ axis. Combined with the second eigenvector $\boldsymbol{u}_2$ in the z direction, the best plane is the yz plane.

These problems are examples where the sample **correlation matrix** (rescaling S so all its diagonal entries are 1) would be the identity matrix. If we think the original scaling is not meaningful and the rows should have the same length, then there is no reason to choose $\boldsymbol{u}_1 = (0, 1, 0)$ from the 8 in row 2.

5 Recall that least squares measures vertical errors (squared distances up or down from data points to the closest line) while PCA measures perpendicular distances to the line. They are different problems. Ordinary least squares is different from PCA = perpendicular least squares.

$A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}b$ is $\begin{bmatrix} 3 & 0 \\ 0 & 14 \end{bmatrix} \widehat{\boldsymbol{x}} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$ leads to $\widehat{\boldsymbol{x}} = \begin{bmatrix} 0 \\ 5/14 \end{bmatrix}$. Best line is $y = \dfrac{5}{14}t$.

PCA finds the line through $(0, 0)$ whose perpendicular distances to the points $(-3, -1)$, $(1, 0)$, $(2, 1)$ is smallest. The computation finds the top eigenvector of $A^{\mathrm{T}}A$, where A is now the 2 by 3 matrix of data points:

$$AA^{\mathrm{T}} = \begin{bmatrix} -3 & 1 & 2 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -3 & -1 \\ 1 & 0 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 5 \\ 5 & 2 \end{bmatrix} \text{ has } \lambda^2 - 16\lambda + 3 = 0.$$

Then $\lambda = 8 \pm \sqrt{61}$ and the top eigenvector of AA^{T} is in the direction of $(5, \sqrt{61} - 6)$ $\approx \mathbf{(5, 1.8)}$. That is the (approximate) direction of the line $y = \dfrac{1.8}{5}t$.

6 See **eigenfaces** on Wikipedia.

7 The closest matrix A_3 of rank 3 has the 3 top singular values $5, 4, 3$. Then $A - A_3$ has singular values **2** and **1**.

8 If A has $\sigma_1 = 9$ and B has $\sigma_1 = 4$, then $A + B$ has $\sigma_1 \leq 13$ because $||A + B|| \leq ||A|| + ||B||$. Also $\sigma_1 \geq 5$ for $A + B$ because $||A + B|| + || - B|| \geq ||A||$.