UNIVERSITY OF CANTERBURY

Exam

Prescription Number:	**EMTH211-17S2**

MARKS Office Use Only	
Q1	
Q2	
Q3	
Q4	
Q5	
Q6	
Q7	
Q8	
Total	

Time allowed: 180 minutes.

Write your answers in the spaces provided.

There is a *total* of 80 points.

Use black or blue ink. Do not use pencil.

Only UC approved calculators are allowed.

There is no formula sheet for this test.

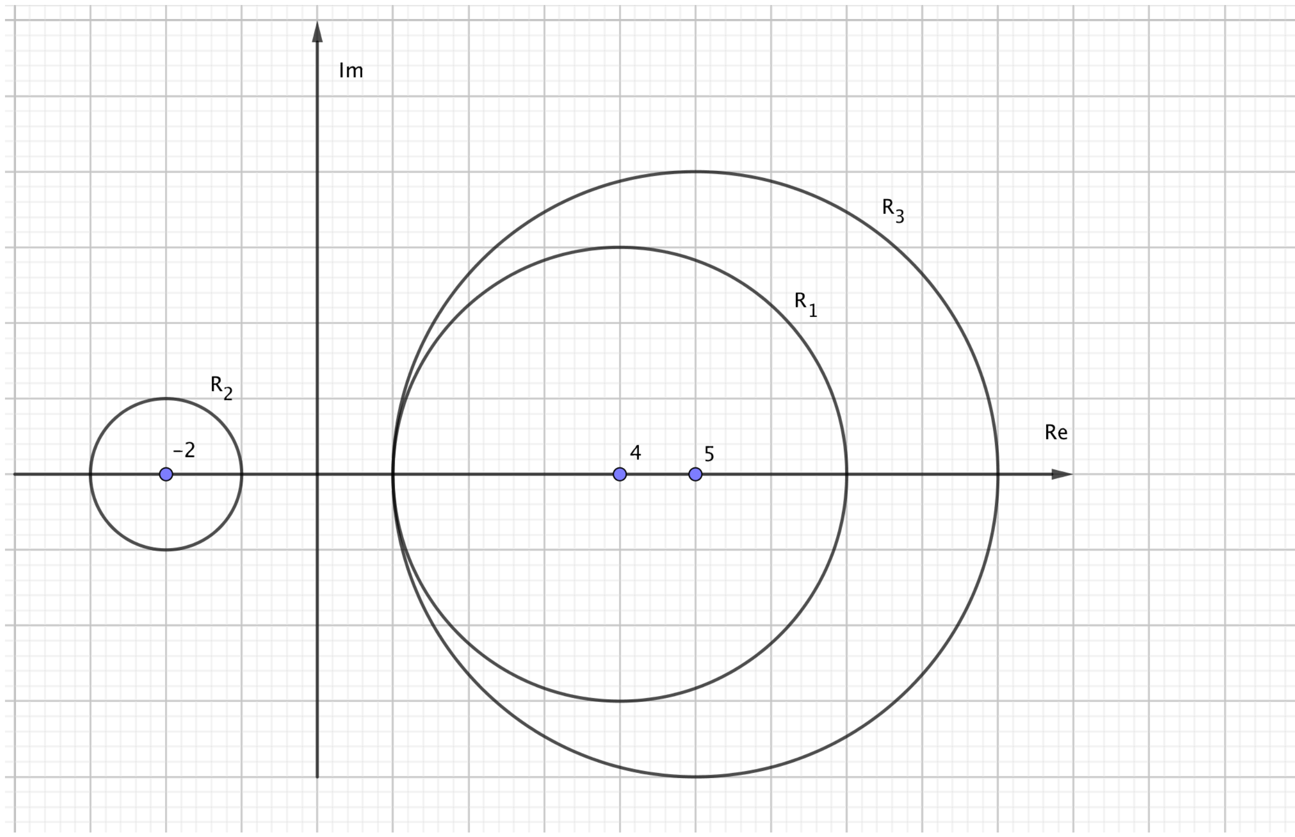
Show all working. Write neatly. Marks can be lost for poorly presented answers.

Family name:	
Given names:	
Student ID:	

Question 1 [14 points]

(a) Consider the matrix $A = \begin{bmatrix} 4 & 0 & 3 \\ 0 & -2 & 1 \\ 3 & 1 & 5 \end{bmatrix}$.

(i) Draw the row based Gerschogorin disks for A.



(ii) How do you deduce from part (i) that A is invertible?

Solution: Every eigenvalue is located in one of the disks, but we see that 0 is not in one of the row based disks. This means that 0 is not an eigenvalue, which in turn means that A is invertible.

(iii) Note that the matrix A is symmetric. What extra information can we deduce about the location of the eigenvalues?

Solution: The eigenvalues are real so they are contained in the union of two intervals: $[-3,-1] \cup [1,9]$.

(b) Consider the system $\begin{cases} 2x + y = 1 \\ x + 3y = -2. \end{cases}$

(i) Using the zero vector as an initial approximation, carry out two iterations of the Jacobi method to approximate the solution of this system.

Solution: We have

$$\begin{aligned} x^{(n+1)} &= \frac{1 - y^{(n)}}{2} = \frac{1}{2} - \frac{y^{(n)}}{2} \\ y^{(n+1)} &= \frac{-2 - x^{(n)}}{3} = \frac{-2}{3} - \frac{x^{(n)}}{3} \end{aligned}$$

Now $x^{(0)} = 0$ and $y^{(0)} = 0$, which implies that $x^{(1)} = 1/2$ and $y^{(1)} = -2/3$. The second approximation then yields

$$x^{(2)} = 5/6$$

and

$$y^{(2)} = -5/6.$$

(ii) Again using the zero vector as initial approximation, carry out two iterations of the Gauss-Seidel method to approximate the solution of this system.

Solution: We have

$$x^{(n+1)} = \frac{1 - y^{(n)}}{2} = \frac{1}{2} - \frac{y^{(n)}}{2}$$
$$y^{(n+1)} = \frac{-2 - x^{(n+1)}}{3} = \frac{-2}{3} - \frac{x^{(n+1)}}{3}$$

Now $x^{(0)} = 0$ and $y^{(0)} = 0$, which implies that $x^{(1)} = 1/2$ and $y^{(1)} = -5/6$. The second approximation then yields

$$x^{(2)} = 11/12$$

and

$$y^{(2)} = -35/36.$$

(iii) Why can we be sure that the iterative processes will converge for this system?

Solution: The coefficient matrix is strictly diagonally dominant.

Question 2 [8 points]

A farmer's herd of cows can be divided into three age classes: calves (aged 0 – 1 year), juveniles (aged 1 – 2 years), and adults (aged 2 – 3 years). Each year, each juvenile produces an average of 0.8 female calves, and each adult produces an average of 0.6 female calves. Calves do not reproduce until they are juveniles. Each year 10% of calves and 20% of juveniles either die or are sold. All adult cows either die or are sold at the end of their third year.

(a) Write down the Leslie matrix L for this population.

Solution:

$$L = \begin{bmatrix} 0 & 0.8 & 0.6 \\ 0.9 & 0 & 0 \\ 0 & 0.8 & 0 \end{bmatrix}.$$

Using the command [P,D]=eig(L) in Matlab with the matrix L from part (a) gives the following output:

$$P = \begin{bmatrix} -0.68568 & 0.15547 - 0.37310i & 0.15547 + 0.37310i \\ -0.58132 & -0.47449 + 0.31631i & -0.47449 - 0.31631i \\ -0.43808 & 0.71514 & 0.71514 \end{bmatrix}$$

and

$$D = \begin{bmatrix} 1.06158 & 0 & 0 \\ 0 & -0.53079 + 0.35385i & 0 \\ 0 & 0 & -0.53079 - 0.35385i \end{bmatrix}$$

(b) Will the size of the herd of cows in the long run grow, decline, or remain stable? Explain your answer.

Solution: It will grow. The dominant eigenvalue is 1.06158 which is strictly larger than 1. The theory tells us that in this case we have population growth.

(c) What is the long term distribution of cows among the three age classes?

Solution: The distribution is proportional to the eigenvector associated with the dominant eigenvalue. This is the first column of D. Scaing by dividing by the sum of the entries in the vector gives us

$$\begin{bmatrix} 0.40214 \\ 0.34095 \\ 0.25694 \end{bmatrix}.$$

So in the long run, the distribution is roughly 40% calves, 34% juveniles and 26% adults.

Question 3 [10 points]

(a) Use the power method to find estimates for the dominant eigenvalue **and** an associated eigenvector for
$$A = \begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix}.$$
Use $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as the initial vector and calculate two iterations. Normalise your iteration using the ∞-norm.

Solution:

First step:
$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$
so $\mathbf{x}^{(1)} = \begin{bmatrix} 1 \\ 2/3 \end{bmatrix}$.
$$R(\mathbf{x}^{(1)}) = \frac{A\mathbf{x}^{(1)} \cdot \mathbf{x}^{(1)}}{\mathbf{x}^{(1)} \cdot \mathbf{x}^{(1)}} = \frac{\begin{bmatrix} 10/3 \\ 4/3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2/3 \end{bmatrix}}{1 + 4/9} = \frac{38}{13}$$

Second step:
$$A\mathbf{x}^{(1)} = \begin{bmatrix} 10/3 \\ 4/3 \end{bmatrix}$$
so $\mathbf{x}^{(2)} = \begin{bmatrix} 1 \\ 2/5 \end{bmatrix}$.
$$R(\mathbf{x}^{(2)}) = \frac{A\mathbf{x}^{(2)} \cdot \mathbf{x}^{(2)}}{\mathbf{x}^{(2)} \cdot \mathbf{x}^{(1)}} = \frac{\begin{bmatrix} 18/5 \\ 4/5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2/5 \end{bmatrix}}{1 + 4/25} = \frac{98}{29}$$

(b) Let A be an $m \times n$ matrix with columns $\mathbf{c}_1$, $\mathbf{c}_2$, …, $\mathbf{c}_n$. Express the product $A\mathbf{x}$ in terms of these columns.

Solution:

$$A\mathbf{x} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n \ ,$$

a linear combination of columns.

(c) (i) Use block multiplication to calculate

$$\begin{bmatrix} A^{-1} & O \\ -D^{-1}CA^{-1} & D^{-1} \end{bmatrix} \begin{bmatrix} A & O \\ C & D \end{bmatrix}.$$

Solution:

$$\begin{bmatrix} A^{-1} & O \\ -D^{-1}CA^{-1} & D^{-1} \end{bmatrix} \begin{bmatrix} A & O \\ C & D \end{bmatrix} = \begin{bmatrix} I & O \\ O & I \end{bmatrix}.$$

(ii) What does the result of part (i) tell you about $\begin{bmatrix} A & O \\ C & D \end{bmatrix}^{-1}$? Justify your answer.

Solution: If H is square and $GH = I$ then H is invertible with inverse G. Thus the block matrix

$$\begin{bmatrix} A^{-1} & O \\ -D^{-1}CA^{-1} & D^{-1} \end{bmatrix}$$

is the inverse of $\begin{bmatrix} A & O \\ C & D \end{bmatrix}$.

(iii) Using the results of the previous parts calculate

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 2 & 2 & 0 \\ 2 & 1 & 0 & 2 \end{bmatrix}^{-1}.$$

Solution:

Putting things into the form of the block matrix $\begin{bmatrix} A & O \\ C & D \end{bmatrix}$ we have $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, $C = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$.

We can find the inverse of any 2×2 matrix with the formula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

or alternatively by using row reduction. Thus we conclude

$$A^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, \quad D^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix},$$

and

$$\begin{aligned} -D^{-1}CA^{-1} &= -\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \\ &= -\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} = -\begin{bmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}. \end{aligned}$$

Thus

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 2 & 2 & 0 \\ 2 & 1 & 0 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & O \\ -D^{-1}CA^{-1} & D^{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -\frac{1}{2} & -1 & \frac{1}{2} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}.$$

This answer can be double checked by multiplying the original 4×4 matrix by the proposed inverse.

Question 4 [12 points]

(a) Let

$$\mathbf{x}_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix}3\\1\\1\\3\end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix}5\\0\\2\\3\end{bmatrix}.$$

Use the Gram–Schmidt process to find an orthonormal basis for $\operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$.

Solution:

$$\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}.$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix}3\\1\\1\\3\end{bmatrix} - 2\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}.$$

$$\begin{aligned}\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\ &= \begin{bmatrix}5\\0\\2\\3\end{bmatrix} - \frac{10}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} - \frac{6}{4}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix} \\ &= \begin{bmatrix}5\\0\\2\\3\end{bmatrix} - \frac{5}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} - \frac{3}{2}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix} = \begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}.\end{aligned}$$

Normalizing

$$\mathbf{q}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{2}\begin{bmatrix}1\\1\\1\\1\end{bmatrix},$$

$$\mathbf{q}_2 = \frac{1}{|v_2\|}\mathbf{v}_2 = \frac{1}{2}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}, \quad \text{and} \quad \mathbf{q}_3 = \frac{1}{2}\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix},$$

are an orthonormal basis.

(b) The matrix $B = \begin{bmatrix} 2 & -1 & 0 \\ 2 & 3 & 0 \\ 2 & -1 & -2 \\ 2 & 3 & 2 \end{bmatrix}$ has an economy QR decomposition with $Q = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$.

(i) Find R.

Solution:

$$QR = B \implies Q^T QR = Q^T B \implies R = Q^T B.$$

Therefore,

$$\begin{aligned} R &= \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & -1 & 0 \\ 2 & 3 & 0 \\ 2 & -1 & -2 \\ 2 & 3 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 2 & 0 \\ 0 & 4 & 2 \\ 0 & 0 & 2 \end{bmatrix}. \end{aligned}$$

(ii) Use this economy QR decomposition to solve the system $B\mathbf{x} \approx \begin{bmatrix} 8 \\ 4 \\ 4 \\ 4 \end{bmatrix}$ by least squares.

Solution: To solve $B\mathbf{x} \approx \mathbf{b}$ by least squares when $B = QR$ we solve $R\mathbf{x} = Q^T\mathbf{b}$ by back substitution.

$$Q^T\mathbf{b} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} 8 \\ 4 \\ 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ -2 \\ 2 \end{bmatrix}.$$

Therefore,

$$\begin{aligned} & R\mathbf{x} = Q^T\mathbf{b} \\ \Longrightarrow \quad & \begin{bmatrix} 4 & 2 & 0 \\ 0 & 4 & 2 \\ 0 & 0 & 2 \end{bmatrix}\mathbf{x} = \begin{bmatrix} 10 \\ -2 \\ 2 \end{bmatrix} \end{aligned}$$

Solving by back substitution, $x_3 = 1$,

$$4x_2 + 2x_3 = -2 \implies 4x_2 = -2 - 2 \implies x_2 = -1.$$

and

$$4x_1 + 2x_2 = 10 \implies 4x_1 = 10 + 2 \implies x_1 = 3.$$

Therefore,

$$\mathbf{x} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}.$$

Question 5 [12 points]

(a) Consider $A = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 1 \end{bmatrix}$. Given that $A^T A$ has eigenvalues $\mu_1 = 3$ and $\mu_2 = 1$ with corresponding eigenvectors $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Find a singular value decomposition of A.

Solution:

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

We are given that $A^T A$ has eigenvalues $\mu_1 = 3$ and $\mu_2 = 1$ with eigenvectors $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, respectively. Therefore $\sigma_1 = \sqrt{3}$ and $\sigma_2 = \sqrt{1} = 1$, and

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix},$$

$$\mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v}_2 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

We need to find $\mathbf{u}_3$ to complete an orthonormal basis for $\mathbb{R}^3$. Start with $\mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and use GS to construct a vector orthogonal to $\mathbf{u}_1$ and $\mathbf{u}_2$.

$$\begin{aligned}
\mathbf{x}_3 &= \mathbf{e}_3 - (\mathbf{e}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{e}_3\mathbf{u}_2)\mathbf{u}_2 \\
&= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad - \quad \frac{1}{\sqrt{6}}. - 1.\frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} \quad - \quad \frac{1}{\sqrt{2}}.1.\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\
&= \frac{1}{6}\left(\begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix} + \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} - 3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right) \\
&= \frac{1}{6} \begin{bmatrix} -2 \\ -2 \\ 2 \end{bmatrix}.
\end{aligned}$$

Normalizing $\mathbf{u}_3 = \dfrac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$.

Thus the SVD is

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

where

$$U = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ -2/\sqrt{6} & 0 & -1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad \text{and} \quad V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

(b) The pixel values of a greyscale image of a certain sports team's lineout calls are stored in an $m \times n$ matrix A. An SVD is then performed on the matrix giving

$$A = \sum_{j=1}^{r} \sigma_j \mathbf{u}_j \mathbf{v}_j^T$$

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are the nonzero singular values, and $\{\mathbf{u}_j\}$ and $\{\mathbf{v}_j\}$ are the left and right singular vectors. A spy wishes to transmit the image to a galaxy far away, using only a very limited amount of data. Can you suggest how he/she might do so?

Solution: The spy can achieve compression by dropping the terms involving small singular values and only transmitting

$$\sigma_1, \mathbf{u}_1, \mathbf{v}_1; \ \sigma_2, \mathbf{u}_2, \mathbf{v}_2; \ \ldots; \sigma_k, \mathbf{u}_k, \mathbf{v}_k;$$

with a suitable choice of k. This is much less data to transmit than the whole $m \times n$ matrix. The sequence of cat images in the lecture notes was an example of using SVD for compression.

Question 6 [7 points]

Suppose we have two data vectors $\mathbf{x}$ and $\mathbf{y}$ of length n.
Note: In this exam the notation $\mathbf{x}$ and $\mathbf{y}$ will be used for $\overrightarrow{x}$ and $\overrightarrow{y}$.

(a) Give the definition of covariance and correlation for the data vectors.

Solution

$$\begin{aligned}
cov(\mathbf{x},\mathbf{y}) &= \frac{1}{n-1}\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}} \\
\text{alternatively} \qquad &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}) \\
cor(\mathbf{x},\mathbf{y}) &= \frac{\tilde{\mathbf{x}}\cdot\tilde{\mathbf{y}}}{\|\tilde{\mathbf{x}}\|\|\tilde{\mathbf{y}}\|} \\
\text{alternatively} \qquad &= \frac{cov(\mathbf{x},\mathbf{y})}{\sqrt{var(\mathbf{x})var(\mathbf{y})}}
\end{aligned}$$

(b) Assume that $\mathbf{x}$ is temperature in Fahrenheit. How will the covariance $cov(\mathbf{x},\mathbf{y})$ and the correlation $cor(\mathbf{x},\mathbf{y})$ change, when $\mathbf{x}$ is transformed into degree Celsius? The transformation is given by

$$\frac{5(\mathbf{x}-32)}{9}.$$

Solution

$$\begin{aligned}
cov\left(\frac{5(\mathbf{x}-32)}{9},\mathbf{y}\right) &= \frac{5}{9}cov(\mathbf{x},\mathbf{y}) \\
cor\left(\frac{5(\mathbf{x}-32)}{9},\mathbf{y}\right) &= cor(\mathbf{x},\mathbf{y})
\end{aligned}$$

(c) Let $\mathbf{y}$ be the amount of carbon dioxide dissolved in water and $\mathbf{x}$ the water temperature. The warmer the water gets the smaller is the amount of dissolved carbon dioxide. What is the sign of $cov(\mathbf{x}, \mathbf{y})$ and $cor(\mathbf{x}, \mathbf{y})$?

Since the values of $\mathbf{x}$ and $\mathbf{y}$ tend to go into opposite directions, the signs of $cov(\mathbf{x}, \mathbf{y})$ and $cor(\mathbf{x}, \mathbf{y})$ will be negative.

(d) Assume we run a simple linear regression

$$\mathbf{y} = b_0 + b_1\mathbf{x} + \mathbf{e}.$$

Which sign do you expect for b_1?

Solution Since the values of $\mathbf{x}$ and $\mathbf{y}$ tend to go into opposite directions, the sign of b_1 will be negative.

Question 7 [10 points]

Let $\mathbf{x} = (1, 2, 4, 5)^\top$ and $\mathbf{y} = (1, 2, 3, 10)^\top$. We consider the regression

$$\mathbf{y} = b_0 + b_1\mathbf{x} + \mathbf{e}.$$

(a) Compute the slope $\widehat{b}_1$ of the least squares regression line.

Solution

$$\widehat{b}_1 = \frac{\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}}}{\|\tilde{\mathbf{x}}\|^2}$$

$\overline{\mathbf{x}} = 3$ and $\overline{\mathbf{y}} = 4$

$\tilde{\mathbf{x}} = (-2, -1, 1, 2)$ and $\tilde{\mathbf{y}} = (-3, -2, -1, 6)$.

$\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}} = 6 + 2 - 1 + 12 = 19$

$\|\tilde{\mathbf{x}}\|^2 = 4 + 1 + 1 + 4 = 10$

Hence, $\widehat{b}_1 = 19/10 = 1.9$.

(b) Compute the intercept $\widehat{b}_0$ of the regression line.

Solution $\widehat{b}_0 = \overline{\mathbf{y}} - \widehat{b}_1\overline{\mathbf{x}} = 4 - 1.9 \times 3 = -1.7.$

(c) Use the value $t_2(0.975) = 4.30$ to give a 95% confidence interval for b_1 from $\widehat{b}_1$.

Solution

$$\begin{aligned} SSR &= (1 - 1.9 + 1.7)^2 + (2 - 1.9 \times 2 + 1.7)^2 + (3 - 1.9 \times 4 + 1.7)^2 + (10 - 1.9 \times 5 + 1.7)^2 \\ &= 0.64 + 0.01 + 8.41 + 4.84 \\ &= 13.9 \end{aligned}$$

$$se(\widehat{b}_1) = \sqrt{\frac{1}{n-2}\frac{SSR}{\|\tilde{\mathbf{x}}\|^2}} = \sqrt{\frac{1}{2}\frac{13.9}{10}} \approx 0.834$$

Hence, $b_1 = \widehat{b}_1 \pm t_2(0.975)se(\widehat{b}_1) \approx 1.9 \pm 3.585$

Or equivalently, $b_1 \in [-1.685, 5.485]$.

EXTRA PAGE FOR WORKING

Question 8 [7 points]

For a multiple linear regression model

$$\mathbf{y} = b_0 + b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + \cdots + b_p\mathbf{x}_p + \mathbf{e}$$

we can consider the vector of estimators $\widehat{\mathbf{b}} = (\widehat{b}_0, \widehat{b}_1, \widehat{b}_2, \ldots, \widehat{b}_p)^\top$ and the matrix

$$X = [\mathbf{1}, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p]\,.$$

(a) Give the normal equation for the vector of estimators $\widehat{\mathbf{b}}$ using this notation.

Solution $X^\top X\widehat{\mathbf{b}} = X^\top \mathbf{y}$

(b) A researcher regresses carbon dioxide dissolved in water y on oxygen dissolved in water $\mathbf{x}_1$, temperature in Celsius $\mathbf{x}_2$, and temperature in Fahrenheit $\mathbf{x}_3$. But the researcher gets results that seem to be odd.

(i) What might have gone wrong in this regression?

Solution The two regressors $\mathbf{x}_2$ and $\mathbf{x}_3$ are linearly dependent. Hence, b_2 and b_3 are not uniquely determined. *Alternatively you can say:* The matrix $X^\top X$ is singular and the normal equation does not determine $\widehat{\mathbf{b}}$.

(ii) How can you check mathematically that this was the reason for the unusual results?

Solution *One of the following possibilities can be mentioned.*

- Check if $det(X^\top X)$ equals 0 or is close to 0.
- Haitovsky test
- Check if the R^2 for the regressing $\mathbf{x}_2$ on the other regressors or for regressing $\mathbf{x}_3$ on the other regressors is close to 1.

(iii) How can this be fixed?

Solution We have to drop either $\mathbf{x}_2$ or $\mathbf{x}_3$ from the regression.

END OF PAPER