Name: ______________________________ Student ID: ____________________

Math 2740 – Fall 2025
Sample final examination (Variant 2) – SOLUTIONS
2 hours

Instructions

- This examination has **9 exercises**.
- Show all your work. Correct answers without justification will receive little or no credit.
- You may use the back of pages if needed.
- No electronic devices (including calculators) are permitted.
- The exam is out of 130 points.

Exercise 1. [**Definitions and Core Results – 15 points**]
State the definition or theorem for each of the following. Be precise and complete.

1. [**5 pts**] State the Gram-Schmidt procedure.
2. [**4 pts**] Define a discrete-time Markov chain.
3. [**3 pts**] Define the dot product of $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$.
4. [**4 pts**] Define the *principal components* of a centered data matrix.

Solution of Exercise 1.

1. **Gram-Schmidt procedure:** Let $W \subset \mathbb{R}^n$ be a subspace and $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ a basis of W. Let

$$
\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{v}_1 \bullet \mathbf{x}_2}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{v}_1 \bullet \mathbf{x}_3}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\mathbf{v}_2 \bullet \mathbf{x}_3}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_k &= \mathbf{x}_k - \frac{\mathbf{v}_1 \bullet \mathbf{x}_k}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \cdots - \frac{\mathbf{v}_{k-1} \bullet \mathbf{x}_k}{\|\mathbf{v}_{k-1}\|^2}\mathbf{v}_{k-1}
\end{aligned}
$$

and

$$W_1 = \mathsf{span}(\mathbf{x}_1), W_2 = \mathsf{span}(\mathbf{x}_1, \mathbf{x}_2), \ldots, W_k = \mathsf{span}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$$

Then $\forall i = 1, \ldots, k$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ is an orthogonal basis for W_i.

2. **Discrete-time Markov chain:** An experiment with finite number of possible outcomes $S_1, \ldots, S_n$ is repeated. The sequence of outcomes is a **Markov chain** if there is a set of n^2 numbers $\{p_{ij}\}$ such that the conditional probability of outcome S_i on any experiment given outcome S_j on the previous experiment is p_{ij}, i.e., for $1 \leq i, j \leq n$, $t = 1, \ldots,$

$$p_{ij} = \mathbb{P}(S_i \text{ on experiment } t+1 \mid S_j \text{ on experiment } t)$$

Outcomes $S_1, \ldots, S_n$ are **states** and p_{ij} are **transition probabilities**. $P = [p_{ij}]$ is the **transition matrix**.

3. **Dot product:** Let $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n$, $\mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$. The **dot product** of $\mathbf{a}$ and $\mathbf{b}$ is the scalar

$$\mathbf{a} \bullet \mathbf{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + \cdots + a_n b_n$$

4. **Principal components:** For $i = 1, \ldots, p$, the ith principal component of a centered data matrix is

$$z_i = \mathbf{v}_i^T \mathbf{x}$$

where $\mathbf{v}_i$ is an eigenvector of the sample covariance matrix S_X associated to the ith largest eigenvalue λ_i. If $\mathbf{v}_i$ is normalized, then $\lambda_i = \mathrm{Var}(z_i)$.

Exercise 2. **[Gram–Schmidt Orthonormalization – 20 points]**

Consider the vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

1. **[4 pts]**Can you apply the Gram–Schmidt procedure to these vectors? Justify your answer.
2. **[4 pts]** Apply the Gram–Schmidt procedure to $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ to obtain an *orthogonal* set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.
3. **[4 pts]** Normalize your vectors to obtain an *orthonormal* set $\{q_1, q_2, q_3\}$.
4. **[4 pts]** Verify orthonormality by computing the inner products $\langle q_i, q_j \rangle$ for all i, j and by checking $\|q_i\| = 1$.
5. **[4 pts]** Form the matrix $Q = [q_1\ q_2\ q_3]$ and state whether Q is orthogonal (justify your answer).

Solution of Exercise 2. **Q1:** Yes, the Gram–Schmidt procedure can be applied to these vectors because they are linearly independent. We can verify this by checking that the matrix $[\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3]$ has full rank (equivalently, that $\det([\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3]) \neq 0$).

Q2–Q5: Given $v_1 = (1, 1, 0)^T$, $v_2 = (1, 0, 1)^T$, $v_3 = (0, 1, 1)^T$.

$$u_1 = v_1, \quad q_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{2}}(1, 1, 0)^T.$$

$$\alpha = \langle v_2, q_1 \rangle = \frac{1}{\sqrt{2}}, \quad u_2 = v_2 - \alpha q_1 = (1 - \tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}}, 1)^T, \quad \|u_2\| = \sqrt{3 - \sqrt{2}}, \quad q_2 = \frac{u_2}{\|u_2\|}.$$

$$\beta = \langle v_3, q_1 \rangle = \tfrac{1}{\sqrt{2}}, \quad \gamma = \langle v_3, q_2 \rangle = \frac{1 - \frac{1}{\sqrt{2}}}{\sqrt{3 - \sqrt{2}}},$$

$$u_3 = v_3 - \beta q_1 - \gamma q_2 = \Big(-\tfrac{1}{2}, \tfrac{1}{2}, 1 \Big)^T - \frac{1 - \frac{1}{\sqrt{2}}}{\sqrt{3 - \sqrt{2}}} \cdot \frac{(1 - \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 1)^T}{\sqrt{3 - \sqrt{2}}},$$

which simplifies to a nonzero orthogonal vector; hence

$$q_3 = \frac{u_3}{\|u_3\|}.$$

Then $Q = [q_1\ q_2\ q_3]$ satisfies $Q^T Q = I$, so Q is orthogonal.

Exercise 3. **[Least Squares via QR – 15 points]**

Let $A \in \mathbb{R}^{m\times n}$ have full column rank and let $A = QR$ be its *reduced* QR decomposition, where $Q \in \mathbb{R}^{m\times n}$ has orthonormal columns and $R \in \mathbb{R}^{n\times n}$ is upper triangular.

1. [**8 pts**] Using an *important theorem*, prove that the least-squares solution to $A\mathbf{x} = \mathbf{b}$ is $\tilde{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$.

   Important Theorem 1 (Least Squares via QR)**.** Let $A = QR$ be a reduced QR decomposition with $Q^TQ = I$ and R upper triangular. Then the least-squares solution to $A\mathbf{x} = \mathbf{b}$ satisfies $\tilde{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$ and the residual is orthogonal to $\operatorname{col}(A)$.

2. [**7 pts**] For

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix},$$

   compute the reduced QR decomposition $A = QR$ (you may use Gram–Schmidt on the columns) and find $\tilde{\mathbf{x}}$.

Solution of Exercise 3. Least Squares via QR. Important theorem: If $A = QR$ with $Q^TQ = I$, R upper triangular, then $\tilde{x} = R^{-1}Q^Tb$. For $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{pmatrix}$, $b = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}$, the normal matrix and right-hand side are

$$A^TA = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}, \quad A^Tb = \begin{pmatrix} 3 \\ 3 \end{pmatrix}.$$

Solving $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \tilde{x} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$ gives $\tilde{x} = \left(\frac{3}{4}, \frac{3}{4}\right)^T$. Residual: $r = b - A\tilde{x} = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}, 0\right)^T$, with $\|r\| = \sqrt{\frac{3}{8}}$. Moreover $r \perp \operatorname{col}(A)$.

Exercise 4. **[Singular Value Decomposition – 15 points]**
Consider

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}.$$

1. **[8 pts]** Compute the full singular value decomposition $A = U\Sigma V^T$ of A. Show your work by computing $A^T A$, its eigenvalues and eigenvectors, then construct V, Σ, and U.
2. **[3 pts]** What is the rank of A?
3. **[4 pts]** Compute the Moore-Penrose pseudoinverse A^+ using the SVD.

Solution of Exercise 4. **Singular Value Decomposition.**

Q1: For $A = \begin{pmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}$, compute

$$A^T A = \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix}.$$

Eigenvalues are $\lambda_1 = 9, \lambda_2 = 4$. Corresponding eigenvectors: $v_1 = (1,0)^T, v_2 = (0,1)^T$, so $V = I_2$.
Singular values: $\sigma_1 = 3, \sigma_2 = 2$.

$$\Sigma = \begin{pmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}.$$

To find U, compute Av_i/σ_i:

$$u_1 = \frac{1}{3}\begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad u_2 = \frac{1}{2}\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Complete to orthonormal basis: $u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Thus

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3.$$

Q2: The rank of A is 2 (two nonzero singular values).
Q3: The Moore-Penrose pseudoinverse is

$$A^+ = V\Sigma^+U^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix}.$$

Exercise 5. [PCA on Centered Data – 10 points]

Let the centered data matrix be

$$\tilde{X} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \\ 1 & -1 \\ 0 & 0 \end{pmatrix}.$$

1. **[6 pts]** Compute the covariance matrix $S = \frac{1}{n-1}\tilde{X}^T\tilde{X}$ and its eigenvalues/eigenvectors.
2. **[4 pts]** Identify the first principal component and the variance explained by it.

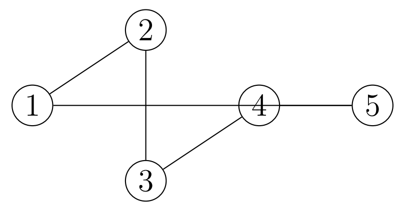
Solution of Exercise 5. **PCA.** With $\tilde{X} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \\ 1 & -1 \\ 0 & 0 \end{pmatrix}$ and $n = 4$,

$$X^T X = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad S = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}.$$

Eigenvalues: $\lambda_1 = 1$, $\lambda_2 = \frac{1}{3}$. Corresponding unit eigenvectors: for λ_1, $u_1 = \frac{1}{\sqrt{2}}(1,-1)^T$; for λ_2, $u_2 = \frac{1}{\sqrt{2}}(1,1)^T$. First principal component: u_1; variance explained: $\lambda_1 = 1$.

Exercise 6. **[Graph Measures I – 12 points]**
Consider the simple undirected graph G on vertices $V = \{1, 2, 3, 4, 5\}$ with edge set

$$E = \{\{1, 2\}, \{1, 5\}, \{2, 3\}, \{3, 4\}, \{4, 5\}\}.$$



1. **[4 pts]** Compute the degree $\deg(i)$ of each vertex and give the degree sequence in nonincreasing order.
2. **[4 pts]** Compute the density of G, defined as $\delta(G) = \dfrac{2|E|}{|V|(|V| - 1)}$.
3. **[4 pts]** Compute the local clustering coefficient C_i for each vertex with $\deg(i) \geq 2$ and state the average clustering coefficient.

Solution of Exercise 6. **Graph Measures I.** Degrees: $\deg(1) = 2, \deg(2) = 2, \deg(3) = 2, \deg(4) = 2, \deg(5) = 2$; degree sequence $(2, 2, 2, 2, 2)$. Density: $\delta(G) = \frac{2|E|}{|V|(|V|-1)} = \frac{2 \cdot 5}{5 \cdot 4} = \frac{1}{2}$. Clustering: For each vertex, neighbors are non-adjacent, so $e_i = 0$ and $C_i = 0$ whenever $\deg(i) \geq 2$. Average clustering $= 0$.

Exercise 7. **[Graph Measures II – 13 points]**

For the same graph G as in Exercise 6:

1. **[5 pts]** Compute the graph diameter and the average shortest-path length $\ell(G)$.
2. **[4 pts]** Compute the (normalized) degree centrality of each vertex, $C_D(i) = \deg(i)/(n-1)$ where $n = |V|$.
3. **[4 pts]** Compute the closeness centrality of each vertex, $C_C(i) = \dfrac{n-1}{\sum_{j \neq i} d(i,j)}$.

Solution of Exercise 7. **Graph Measures II.** Shortest paths yield diameter $= 3$ (e.g., $2 \to 5$ via $2 \to 1 \to 5$ or $2 \to 3 \to 4 \to 5$). Average shortest-path length: sum over unordered pairs is 15, giving $\ell(G) = 15/10 = 1.5$. Degree centrality: all equal $C_D(i) = \deg(i)/(5-1) = 2/4 = 0.5$. Closeness: for each vertex, $\sum_{j \neq i} d(i,j) = 6$, so $C_C(i) = \frac{4}{6} = \frac{2}{3}$ for all i.

Exercise 8. [Absorbing Markov Chains – 20 points]

Consider a Markov chain with four states $\{1, 2, 3, 4\}$ and column-stochastic transition matrix:

$$P = \begin{pmatrix} 1 & 0.3 & 0.2 & 0 \\ 0 & 0.4 & 0.1 & 0 \\ 0 & 0.2 & 0.5 & 0.5 \\ 0 & 0.1 & 0.2 & 0.5 \end{pmatrix},$$

where P_{ij} is the probability of moving from state j to state i.

1. **[4 pts]** Draw the directed graph representation of this Markov chain, showing all states and transition probabilities on the edges.
2. **[3 pts]** Identify which states are absorbing and which are transient. Justify your answer.
3. **[3 pts]** Explain why this Markov chain is classified as an absorbing Markov chain.
4. **[5 pts]** Reorder the states (if necessary) to write P in canonical form

$$P = \begin{pmatrix} I & 0 \\ R & Q \end{pmatrix}$$

and identify the matrices I, R, and Q.

5. **[5 pts]** Compute the fundamental matrix $N = (I - Q)^{-1}$. Interpret what the entries N_{ij} represent.

Solution of Exercise 8. Absorbing Markov Chains.

Q1: The directed graph has states $\{1, 2, 3, 4\}$. State 1 has a self-loop with probability 1. From state 2: to 1 with prob 0.3, to 2 with prob 0.4, to 3 with prob 0.2, to 4 with prob 0.1. From state 3: to 1 with prob 0.2, to 2 with prob 0.1, to 3 with prob 0.5, to 4 with prob 0.2. From state 4: to 3 with prob 0.5, to 4 with prob 0.5.

Q2: State 1 is absorbing (probability 1 of staying in state 1). States 2, 3, 4 are transient (they can reach the absorbing state 1 and cannot return).

Q3: This is an absorbing Markov chain because: (i) there is at least one absorbing state (state 1), and (ii) it is possible to reach the absorbing state from every transient state.

Q4: The matrix is already in canonical form with states ordered as $(1, 2, 3, 4)$:

$$P = \begin{pmatrix} 1 & 0.3 & 0.2 & 0 \\ 0 & 0.4 & 0.1 & 0 \\ 0 & 0.2 & 0.5 & 0.5 \\ 0 & 0.1 & 0.2 & 0.5 \end{pmatrix} = \begin{pmatrix} I & 0 \\ R & Q \end{pmatrix}$$

where $I = (1)$ is 1×1,

$$R = \begin{pmatrix} 0.3 \\ 0.2 \\ 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 0.4 & 0.1 & 0 \\ 0.2 & 0.5 & 0.5 \\ 0.1 & 0.2 & 0.5 \end{pmatrix}.$$

Q5: Compute $I - Q$:

$$I - Q = \begin{pmatrix} 0.6 & -0.1 & 0 \\ -0.2 & 0.5 & -0.5 \\ -0.1 & -0.2 & 0.5 \end{pmatrix}.$$

The fundamental matrix is

$$N = (I - Q)^{-1} = \begin{pmatrix} 1.724 & 0.345 & 0.345 \\ 1.034 & 2.414 & 2.414 \\ 0.517 & 1.207 & 3.207 \end{pmatrix}$$

(approximately).

The entry N_{ij} represents the expected number of times the chain visits transient state i before absorption, given that it starts in transient state j.

Exercise 9. **[Reading R Code – 10 points]**

What does the following function do? Explain your answer by describing the algorithm and its purpose. You do not need to carry out a numerical run, but you should identify what mathematical operation is being performed.

```
mystery_function <- function(A, tol = 1e-10) {
  if (!is.matrix(A)) stop("A␣must␣be␣a␣matrix")

  m <- nrow(A)
  n <- ncol(A)

  M1 <- matrix(0, nrow = m, ncol = n)
  M2 <- matrix(0, nrow = n, ncol = n)

  for (j in 1:n) {
    v <- A[, j]

    if (j > 1) {
      for (i in 1:(j-1)) {
        M2[i, j] <- sum(M1[, i] * A[, j])
        v <- v - M2[i, j] * M1[, i]
      }
    }

    M2[j, j] <- sqrt(sum(v^2))

    if (M2[j, j] < tol) {
      stop(sprintf("Column␣%d␣is␣linearly␣dependent␣on␣previous␣columns", j))
    }

    M1[, j] <- v / M2[j, j]
  }

  return(list(M1 = M1, M2 = M2))
}
```

Solution of Exercise 9. Reading R Code.

The function `mystery_function` computes the **QR decomposition** of a matrix A. Specifically:

- `M1` stores the matrix Q with orthonormal columns
- `M2` stores the upper triangular matrix R

The algorithm uses the Gram–Schmidt process:

1. For each column j of A, it starts with $v = A[, j]$
2. It subtracts projections onto all previous orthonormal columns: $v \leftarrow v - \langle v, q_i \rangle q_i$ for $i = 1, \ldots, j-1$
3. The coefficients $\langle v, q_i \rangle$ are stored in $R[i, j]$ (the upper triangular part)
4. It normalizes v by its length $R[j, j] = \|v\|$ to get $q_j = v/\|v\|$
5. The normalized vector is stored in `M1[,j]`

The function checks for linear dependence (if a column becomes zero after projection, it's linearly dependent on previous columns) and returns the factorization $A = QR$ where Q has orthonormal columns and R is upper triangular.

END OF EXAMINATION