Final Exam

You may not use any books, notes, or computer programs (e.g., Julia). Throughout this exam (with one exception noted below) we use standard mathematical notation; and in particular, we do not use (and you may not use) notation from any computer language, or from any strange or non-standard mathematical dialect (e.g., physics). (One problem, however, specifically involves the Julia language.)

The first three questions are multiple choice. For these problems simply circle the appropriate response. You do not need to give any justification for your answers to these questions. We will give partial credit for multiple choice problems left with no answer. If we can’t tell which response you are selecting, we will give zero credit.

For the other problems you are asked for a free-form answer, which must be written in the space between the lines below the problem statement. You should do your scratch work on scratch paper.

All problems have equal weight. Some are easy. Others, not so much. We recommend doing the easy ones first.

Name: ____________________________________________

SUID: ____________________________________________

(For EE103 staff only)

<table>
<thead>
<tr>
<th>Problem</th>
<th>Score</th>
<th>Problem</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>/10</td>
<td>8</td>
<td>/10</td>
</tr>
<tr>
<td>2</td>
<td>/10</td>
<td>9</td>
<td>/10</td>
</tr>
<tr>
<td>3</td>
<td>/10</td>
<td>10</td>
<td>/10</td>
</tr>
<tr>
<td>4</td>
<td>/10</td>
<td>11</td>
<td>/10</td>
</tr>
<tr>
<td>5</td>
<td>/10</td>
<td>12</td>
<td>/10</td>
</tr>
<tr>
<td>6</td>
<td>/10</td>
<td>13</td>
<td>/10</td>
</tr>
<tr>
<td>7</td>
<td>/10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>/130</td>
</tr>
</tbody>
</table>
1. **Stacked matrix.** Let $A$ be an $m \times n$ matrix, and consider the stacked matrix $S$ defined by

$$S = \begin{bmatrix} A \\ I \end{bmatrix}.$$ 

Circle the (one) correct choice in each subproblem below.

(a) $S$ has linearly independent columns
   - never
   - always
   - only when $A$ has independent columns
   - only when $A$ has independent rows

(b) $S$ has linearly independent rows
   - never
   - always
   - only when $A$ has independent columns
   - only when $A$ has independent rows
2. *Appropriate response.* For each of the situations (a), (b), and (c) described below, circle the most appropriate response from among the five choices. You can circle only one for each situation.

(a) An intern working for you develops several different models to predict the daily demand for a product. How should you choose which model is the best one, the one to put into production?

<table>
<thead>
<tr>
<th>Laplacian</th>
<th>Magic</th>
<th>Regularization</th>
<th>Validation</th>
<th>k-means</th>
</tr>
</thead>
</table>

(b) As an intern you develop an auto-regressive model to predict tomorrow’s sales volume. It works very well, making predictions that are typically within 5% of the actual sales volume. Your boss, who has an MBA and is not particularly interested in mathematical details, asks how your predictor works. How do you respond?

<table>
<thead>
<tr>
<th>Laplacian</th>
<th>Magic</th>
<th>Regularization</th>
<th>Validation</th>
<th>k-means</th>
</tr>
</thead>
</table>

(c) A colleague needs a quantitative measure of how rough an image is, *i.e.*, how much adjacent pixel values differ. What do you suggest?

<table>
<thead>
<tr>
<th>Laplacian</th>
<th>Magic</th>
<th>Regularization</th>
<th>Validation</th>
<th>k-means</th>
</tr>
</thead>
</table>
3. *Regularized least squares in Julia.* You are asked to write some Julia code to compute the \( \hat{x} \) that minimizes \( \|Ax - b\|^2 + \lambda \|x\|^2 \), where \( A \) is an \( m \times n \) matrix, \( b \) is an \( m \)-vector, and \( \lambda \) is a positive scalar. These are given as the Julia quantities \( A \), \( b \), and \( \text{lambda} \), respectively, and the dimensions \( m \) and \( n \) are given as \( m \) and \( n \). You are to put the value of \( \hat{x} \) in \( \text{xhat} \).

Which of the following Julia snippets will carry this out correctly? Circle the correct response for each code snippet below. You do not need to justify your responses.

(a) \( \text{xhat} = [A \sqrt{\text{lambda}} \cdot \text{eye}(n)] \backslash b \) **Works.** ** Doesn’t work.**
(b) \( \text{xhat} = [A \sqrt{\text{lambda}} \cdot \text{eye}(n)] \{b; \text{zeros}(m)\} \) **Works.** ** Doesn’t work.**
(c) \( \text{xhat} = \text{inv}(A' \cdot A + \lambda \cdot \text{eye}(n)) \cdot (A' \cdot b) \) **Works.** ** Doesn’t work.**
(d) \( \text{xhat} = [A; \sqrt{\text{lambda}} \cdot \text{eye}(n)] \backslash b \) **Works.** ** Doesn’t work.**
(e) \( \text{xhat} = [A; \sqrt{\text{lambda}} \cdot \text{eye}(n)] \{b; \text{zeros}(n)\} \) **Works.** ** Doesn’t work.**
4. *Converting from purchase quantity matrix to purchase dollar matrix*. An $n \times N$ matrix $Q$ gives the purchase history of a set of $n$ products by $N$ customers, over some period, with $Q_{ij}$ being the quantity of product $i$ bought by customer $j$. The $n$-vector $p$ gives the product prices. A data analyst needs the $n \times N$ matrix $D$, where $D_{ij}$ is the total dollar value that customer $j$ spent on product $i$. Express $D$ in terms of $Q$ and $p$, using compact matrix/vector notation. You can use any notation or ideas we have encountered, *e.g.*, stacking, slicing, block matrices, transpose, matrix-vector product, matrix-matrix product, inner product, norm, correlation, $\text{diag}()$, and so on.
5. **A matrix identity.** Suppose $A$ is a square matrix that satisfies $A^3 = 0$. A student guesses that $(I - A)^{-1} = I + A + A^2$, based on the infinite series $1/(1-a) = 1 + a + a^2 + \cdots$, which holds for numbers $a$ that satisfy $|a| < 1$.

Is the student right or wrong? If right, show that her assertion holds with no further assumptions about $A$. If she is wrong, give a counterexample, i.e., a matrix $A$ that satisfies $A^3 = 0$, but $I + A + A^2$ is not the inverse of $I - A$.

Circle one and give explanation or counterexample below: *She is right. She is wrong.*

---

6
6. *Polynomial differentiation.* Suppose \( p \) is a polynomial of degree \( n - 1 \) or less, given by 
\[
p(t) = c_1 + c_2 t + \cdots + c_n t^{n-1}.
\]
Its derivative (with respect to \( t \)) \( p'(t) \) is a polynomial of degree \( n - 2 \) or less, given by 
\[
p'(t) = d_1 + d_2 t + \cdots + d_{n-1} t^{n-2}.
\]
Find a matrix \( D \) for which \( d = Dc \). (Give the entries of \( D \), and be sure to specify its dimensions.)
7. **Minimum trading to achieve target sector exposures.** A hedge fund currently holds a portfolio in a universe of \( n \) assets given by the current weight \( n \)-vector \( w^{\text{curr}} \). Its entries add up to one, but need not be positive. We are to choose a new weight \( n \)-vector \( w \), whose entries also add up to one. The quantity \( \|w - w^{\text{curr}}\| \) is a measure of how much trading will be required to move to the new portfolio vector \( w \). We would like this to be as small as possible, subject to some requirements on sector exposure discussed below.

A given \( k \times n \) sector exposure matrix \( S \) is used to analyze a portfolio's exposure to \( k \) sectors (such as energy, pharmaceuticals, consumer electronics, financial services). The \( k \)-vector \( s = Sw \) gives the exposure of the portfolio with weight vector \( w \) to the \( k \) sectors. (In the simplest case, the entries of \( S \) are 0 or 1, depending on whether the asset is in the sector. In more sophisticated sector models the entries of \( S \) have fractional entries. This will not matter in this problem.)

Suppose that a target exposure vector \( s^{\text{tar}} \) is given. How would you choose \( w \) to minimize the trading measure \( \|w - w^{\text{curr}}\| \) while achieving the target exposures (and of course, the sum of the entries of \( w \) is one)?

Your answer can involve any of the concepts we have covered, including inverses and pseudo-inverses, block matrices, and so on. Any matrices or vectors that appear in your answer must be fully defined. If your method requires an assumption (such as linear independence of some set of vectors), state it clearly.
8. \textit{Left and right inverses of a vector.} Suppose that $x$ is a nonzero $n$-vector with $n > 1$.

(a) Does $x$ have a left inverse? If so, give one. If not say so.

(b) Does $x$ have a right inverse? If so, give one. If not say so.

If the left or right inverse does not exist, it is enough to say that it doesn’t exist. If a left or right inverse exists, then give one. You do not need to justify your answer.
9. *Immigration.* The population dynamics of a country is given by $x_{t+1} = Ax_t + u$, $t = 1, \ldots, T$, where the 100-vector $x_t$ gives the population age distribution in year $t$, and $u$ gives the immigration age distribution (with negative entries meaning emigration), which we assume is constant (i.e., does not vary with $t$). You are given $A$, $x_1$, and $x^{\text{des}}$, a 100-vector that represents a desired population distribution in year 4. We seek $u$ that achieves $x_4 = x^{\text{des}}$.

Give a matrix formula for $u$. If your formula only makes sense when some conditions hold (for example invertibility of one or more matrices), say so.
10. *Products, materials, and locations.* $P$ different products each require some amounts of $M$ different materials, and are manufactured in $L$ different locations, which have different material costs. We let $C_{lm}$ denote the cost of material $m$ in location $l$, for $l = 1, \ldots, L$ and $m = 1, \ldots, M$. We let $Q_{mp}$ denote the amount of material $m$ required to manufacture one unit of product $p$, for $m = 1, \ldots, M$ and $p = 1, \ldots, P$. Let $T_{pl}$ denote the total cost to manufacture product $p$ in location $l$, for $p = 1, \ldots, P$ and $l = 1, \ldots, L$.

Give an expression for the matrix $T$. 

---

11
11. Gram matrix. Let $a_1, \ldots, a_n$ be the columns of the $m \times n$ matrix $A$. Suppose that the columns all have norm one, and for $i \neq j$, $\angle(a_i, a_j) = 60^\circ$. What can you say about the Gram matrix $G = A^T A$? Be as specific as you can be.
12. *Transit system tomography.* An urban transit system (say, a subway) consists of \( n \) links between pairs of stations. Each passenger enters the system at an origin station, traverses over a sequence of the links, and then exits the system at their destination station. The fare collection system keeps track of the following information for passenger \( i \): \( s_i \), the trip starting time (when she enters the origin station), measured in minutes after 6AM, \( f_i \), the trip finishing time (when she leaves the destination station), and the sequence of links over which she traveled. For example, \( s_i = 128 \), \( f_i = 144 \), and link list \((3, 7, 8, 10, 4)\) means the passenger entered the origin station at 8:08AM, left the destination station at 8:24AM, and traversed links 3, 7, 8, 10, 4, in that order. The total trip time is \( f_i - s_i \).

We model the trip time as the sum of the delays over the links the passenger traverses. We let \( d_i \) denote the delay associated with link \( i \), for \( i = 1, \ldots, n \). We do not know the link delays, but wish to estimate them, based on the passenger information described above, for a very large number \( m \) of passengers passing through the system.

We will do this using least-squares. We choose our estimate \( \hat{d} \) of the \( n \)-vector of link delays as the minimizer of the norm squared of the residual of our trip time model. In other words, we choose \( \hat{d} \) to minimize \( \| Rd - c \|^2 \), for some \( m \times n \) matrix \( R \) and some \( m \)-vector \( c \).

Say what \( R \) and \( c \) are. (That is, give all their entries.)
13. Suppose $A$ is a $5 \times 10$ matrix, $B$ is a $20 \times 10$ matrix, and $C$ is a $10 \times 10$ matrix. Determine whether each of the following expressions make sense. If the expression makes sense, give its dimensions.

(a) $A^T A + C$.
(b) $BC^3$.
(c) $I + BC^T$.
(d) $B^T - [C \ I]$.
(e) $B \begin{bmatrix} A \\ A \end{bmatrix} C$. 