Midterm Examination #2

You should fill in the blanks with answers in the simplest form. You don’t need to justify anything, but you may include your reasoning for partial credits, if applicable, in the provided boxes.

1. **Optimization (30 points).**
   
   (a) Let $A \in \mathbb{R}^{m \times n}$ be full-rank and fat, and $\hat{x}$ be the least-norm solution to $Ax = b$. Then
   
   $$\|\hat{x}\|^2 = \frac{b'(AA')^{-1}b}{c}.$$  

   **Solution:** The last norm solution $\hat{x} = A^+b = A'(AA')^{-1}b$. Thus, $\|\hat{x}\|^2 = \hat{x}'\hat{x} = b'(AA')^{-1}'AA'(AA')^{-1}b = b'(AA')^{-1}b$, since $(AA')^{-1}$ is symmetric.

   (b) Let $A \in \mathbb{R}^{m \times n}$ be full-rank and fat. Then the orthogonal projection of $c \in \mathbb{R}^n$ on to the set $\{x : Ax = b\}$ is $\frac{A'(AA')^{-1}(b - Ac) + c}{c}$.  

   **Solution:** Let $\hat{c}$ be the orthogonal projection of $c$ on to the set $\{x : Ax = b\}$. By definition of orthogonal projection, we must have $c - \hat{c} \perp x - \hat{c}$ for any $x$ in this set. It then follows by Pythagoras theorem that $\|x - c\|^2 = \|c - \hat{c}\|^2 + \|\hat{c} - x\|^2$, and therefore that $\hat{c}$ is the orthogonal projection of $c$ if and only if it is the minimizer of $\|x - c\|^2$ in the set $\{x : Ax = b\}$. Thus, this problem reduces to minimizing $\|x - c\|$ subject to $Ax = b$. Making the substitution $y = x - c$, problem can be seen to be equivalent to minimizing $\|y\|$ subject to $Ay = b - Ac$, the solution to which is $\hat{y} = A^+(b - Ac)$. We then see that the required $\hat{x} = \hat{y} + c = A^+(b - Ac) + c = A'(AA')^{-1}(b - Ac) + c$.

   **Alternative solution:** Translating everything by $c$, it is equivalent to finding the orthogonal projection of 0 to $A(\hat{x} + c) = b$, which is given by the least-norm solution of $A\hat{x} = b - Ac$, i.e., $\hat{x} = A'(AA')^{-1}(b - Ac)$. By translating back to the original coordinates, we have the desired solution $\hat{x} = \hat{x} + c = A'(AA')^{-1}(b - Ac) + c$.  

   (c) Let $A \in \mathbb{R}^{m \times n}$ be full-rank and tall with rows $\tilde{a}_1, \ldots, \tilde{a}_m$. Then $x \in \mathbb{R}^n$ that minimizes

   $$\sum_{i=1}^{m} w_i(\tilde{a}_i'x - b_i)^2$$

   for given positive weights $w_1, \ldots, w_m$ is

   $$\left(\sum_{i=1}^{m} w_i\tilde{a}_i\tilde{a}_i'\right)^{-1}\sum_{i=1}^{m} w_i b_i \tilde{a}_i.$$  

   (Your answer should be in terms of $A, b = (b_1, \ldots, b_m)$, and $w = (w_1, \ldots, w_m)$.)

   **Solution:** Note that $\sum_{i=1}^{n}(\tilde{a}_i'x - b_i)^2 = \|Ax - b\|^2$. Now define $W = \text{diag}(w) = \text{diag}(w_1, \ldots, w_m)$, and $\sqrt{W} = \text{diag}(\sqrt{w_1}, \ldots, \sqrt{w_m})$. It is then easy to see that $\sum_{i=1}^{n} w_i(\tilde{a}_i'x - b_i)^2 = \|\sqrt{W}Ax - \sqrt{W}b\|^2$. Note that if $A$ is tall and full-rank, then $\sqrt{W}A$ is also tall and full rank. Therefore, $(\sqrt{W}A)^+ = ((\sqrt{W}A)'\sqrt{W}A)^{-1}(\sqrt{W}A)' = (A'WA)^{-1}A'\sqrt{W}$, and the least-squares solution to $\|\sqrt{W}Ax - \sqrt{W}b\|^2$ is given by $\hat{x} = (\sqrt{W}A)^+\sqrt{W}b = (A'WA)^{-1}A'Wb = (\sum_{i=1}^{m} w_i\tilde{a}_i\tilde{a}_i')^{-1}\sum_{i=1}^{m} w_i b_i \tilde{a}_i$.  

1
2. Eigenvalues and eigenvectors (20 points).

(a) The eigenvalues of \( A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \) (including multiplicity) are \( 1, 1, -1 \).

Solution: The characteristic equation of \( A \) is given by \( \chi(\lambda) = \det(\lambda I - A) = (\lambda - 1)^2(\lambda + 1) \). Thus, the eigenvalues (counting multiplicity) of the matrix \( A \) are 1, 1, and -1.

(b) Let \( B = xx' \) for some unit vector \( x \in \mathbb{R}^n \) (i.e., \( \|x\| = 1 \)). The eigenvalues of \( B \) (including multiplicity) are \( 0, \ldots, 0, 1 \).

Solution: We have \( \det(\lambda I - xx') = \lambda^n \det(I - xx'/\lambda) = \lambda^n \det(1 - x'x/\lambda) = \lambda^{n-1}(\lambda - 1) \), where the second equality follows by property of the determinant. Therefore, the eigenvalues of this matrix are given by \( (1, 0, \ldots, 0) \).

(c) Suppose that the eigenvalues of \( C \in \mathbb{R}^{n \times n} \) are \( \lambda_1, \ldots, \lambda_n \). The eigenvalues of \( I - C^2 \) (in terms of \( \lambda_1, \ldots, \lambda_n \)) are \( 1 - \lambda_1^2, \ldots, 1 - \lambda_n^2 \).

Solution: Let the Jordan canonical form of \( C = T J T^{-1} \). Then, \( C^2 = T J^2 T^{-1} \), and subsequently \( I - C^2 = T(I - J^2)^{-1} \). Since eigenvalues are unchanged by similarity transform, the eigenvalues of \( I - C^2 \) are the same as those of \( I - J^2 \). Since \( I - J^2 \) is an upper triangular matrix with diagonal elements given by \( 1 - \lambda_1^2, \ldots, 1 - \lambda_n^2 \), we see that the eigenvalues of \( I - J^2 \) are given by \( 1 - \lambda_1^2, \ldots, 1 - \lambda_n^2 \) (since the eigenvalues of a triangular matrix are simply the diagonal elements).

Alternative solution. Note that we have \( \lambda I - C^2 = (\sqrt{\lambda} I + C)(\sqrt{\lambda} I - C) \). Hence, he characteristic polynomial of \( C^2 \) can be written as

\[
\det(\lambda I - C^2) = \det(\sqrt{\lambda} I + C) \det(\sqrt{\lambda} I - C) = (\sqrt{\lambda} + \lambda_1) \cdots (\sqrt{\lambda} + \lambda_n)(\sqrt{\lambda} - \lambda_1) \cdots (\sqrt{\lambda} - \lambda_n) = (\lambda - \lambda_1^2) \cdots (\lambda - \lambda_n^2),
\]

which implies that the eigenvalues of \( C^2 \) are \( \lambda_1^2, \ldots, \lambda_n^2 \).

(d) Let \( Dx = \lambda x \) and \( y'D = \mu y' \). If \( \lambda \neq \mu \), then \( y'x = \mathbf{0} \).

Solution: Observe that \( y'Dx = \mu y'x \). But, since \( Dx = \lambda x \), we also have \( y'Dx = \lambda y'x \), which implies that \( (\lambda - \mu)y'x = 0 \). If \( \lambda \neq \mu \), it follows that \( y'x = 0 \).