Write your name below. This exam is worth 100 points. On each problem, you must show all your work to receive credit on that problem. You are allowed to use two pages of notes (one piece of paper, front and back). You are not allowed to use a calculator or any computational software.

Name: Section:
(Chi/9:05/001; Grooms/11:15/002; Grooms/1:25/003)

1. (28 points: 4 each) If the statement is **always true** mark "TRUE"; if it is possible for the statement to be false then mark "FALSE". **No justification is necessary**.

_____(a) A matrix whose columns form an orthogonal basis of $\mathbb{R}^n$ is an orthogonal matrix.

Solution: False. The columns need to be *orthonormal*.

_____(b) An orthogonal matrix has determinant 1.

Solution: False. An orthogonal matrix has determinant ± 1.

_____(c) When obtaining a QR decomposition for a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ with the Householder method, the resulting $\mathbf{Q}$ matrix will be rectangular in the case that $m > n$.

Solution: False. The $\mathbf{Q}$ matrix is the product of the orthogonal Householder matrices so it is also square.

_____(d) Suppose that $\mathbf{A}$ is symmetric and has an LU factorization without pivoting. True or False: $\mathbf{A}$ is positive definite.

Solution: False. The additional condition needed to ensure that $\mathbf{A}$ is positive definite is that the diagonal elements of the $\mathbf{U}$ factor (i.e. the REF) are positive.

_____(e) If the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are mutually orthogonal and nonzero, then they are linearly independent. (Mutually orthogonal means $\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = 0$ whenever $i \neq j$.)

Solution: True.

_____(f) Let $\boldsymbol{u}$ be a unit vector. True or False: $(\mathbf{I} - \boldsymbol{u}\boldsymbol{u}^T)\boldsymbol{x}$ is the orthogonal projection of $\boldsymbol{x}$ onto span$\{\boldsymbol{u}\}^\perp$.

Solution: True. The projection of $\boldsymbol{x}$ onto the span of $\boldsymbol{u}$ is $\boldsymbol{u}\boldsymbol{u}^T\boldsymbol{x}$, so $\boldsymbol{x} - \boldsymbol{u}\boldsymbol{u}^T\boldsymbol{x}$ is the projection onto the orthogonal complement.

_____(g) Let $p(\boldsymbol{x}) = \boldsymbol{x}^T\mathbf{K}\boldsymbol{x} - 2\boldsymbol{x}^T\boldsymbol{f} + c$, and let $\mathbf{K}$ be symmetric and non-negative definite. True or False: If $\boldsymbol{f}$ is orthogonal to the kernel of $\mathbf{K}$ then the function p has a minimum value.

Solution: True. Since $\mathbf{K}$ is symmetric its kernel and its cokernel are equal. If $\boldsymbol{f}$ is orthogonal to the kernel then it is in the range, so a solution of $\mathbf{K}\boldsymbol{x} = \boldsymbol{f}$ exists. This solution corresponds to a minimizer of the function p. If there are multiple solutions, they all produce the same minimum value of the function.

2. (30 points, 6 each) Let $\mathcal{W} \subset \mathbb{R}^3$ be the plane spanned by the vectors

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

and let $\mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. We want to find the $\mathbf{w} \in \text{span}(\mathbf{w}_1, \mathbf{w}_2)$ that is closest to $\mathbf{b}$. We do this by finding the $\mathbf{w}$ that minimizes the distance between $\mathbf{b}$ and all the vectors in $\text{span}(\mathbf{w}_1, \mathbf{w}_2)$. We denote the solution as $\mathbf{w}^\star$.

(a) Write this problem in terms of minimizing a quadratic function $p(\mathbf{x}) = \mathbf{x}^T\mathbf{K}\mathbf{x} - 2\mathbf{x}^T\mathbf{f} + c$. Assume that we employ the standard dot product so that we are minimizing the Euclidean distance between $\mathbf{b}$ and $\mathbf{w}$. What are $\mathbf{K}$, $\mathbf{f}$, and c?

(b) What are the coordinates of $\mathbf{w}^\star$ in terms of $\mathbf{w}_1$ and $\mathbf{w}_2$?

(c) What is this closest point $\mathbf{w}^\star$?

(d) What is the *squared* Euclidean distance between $\mathbf{w}^\star$ and $\mathbf{b}$? (Hint: No need to take square roots!)

(e) Suppose that we employ instead the weighted norm defined by $\|\mathbf{v}\|^2 = 3v_1^2 + v_2^2 + \frac{1}{3}v_3^2$. What are $\mathbf{K}$ and $\mathbf{f}$ in this case?

Solution:

(a) Let $\mathbf{w} = c_1\mathbf{w}_1 + c_2\mathbf{w}_2$. Expanding $\|\mathbf{b} - \mathbf{w}\|^2$, we have

$$\begin{aligned} \|\mathbf{b} - \mathbf{w}\|^2 &= \langle \mathbf{b} - \mathbf{w}, \mathbf{b} - \mathbf{w} \rangle \\ &= \langle \mathbf{b}, \mathbf{b} \rangle - 2\langle \mathbf{b}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle \\ &= \underbrace{\|\mathbf{b}\|^2}_{c} - 2\underbrace{\sum_i c_i \langle \mathbf{b}, \mathbf{w}_i \rangle}_{\mathbf{x}^T\mathbf{f}} + \underbrace{\sum_i \sum_j c_i c_j \langle \mathbf{w}_i, \mathbf{w}_j \rangle}_{\mathbf{x}^T\mathbf{K}\mathbf{x}}, \end{aligned}$$

where $\mathbf{K}_{ij} = \langle \mathbf{w}_i, \mathbf{w}_j \rangle$. Stacking the $\mathbf{w}_i$ into a matrix, we have $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 0 & -1 \end{pmatrix}$ so that

$$\mathbf{K} = \mathbf{A}^T\mathbf{A} = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$$

and

$$\mathbf{f} = \mathbf{A}^T\mathbf{b} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

(b) We notice in $\mathbf{K}$ that $a = 5 > 0$ and $ac - b^2 = 6 > 0$ so $\mathbf{K}$ is positive definite so the solution to $\mathbf{K}\mathbf{x} = \mathbf{f}$ is unique and given by

$$\mathbf{x}^\star = \mathbf{K}^{-1}\mathbf{f} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{5}{6} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{3} \end{pmatrix}.$$

The coordinates of $\mathbf{w}^\star$ in terms of the $\mathbf{w}_1$ and $\mathbf{w}_2$ are then $c_1 = \frac{1}{3}$ and $c_2 = -\frac{1}{3}$.

(c) The closest point $\mathbf{w}^\star$ is

$$\mathbf{w}^\star \quad = \quad \mathbf{A}\mathbf{x}^\star = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}.$$

(d) The squared Euclidean distance is then

$$\|\mathbf{b} - \mathbf{w}^\star\|^2 \quad = \quad \left\| \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} \right\|^2 = \left\| \begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \end{pmatrix} \right\|^2 = \frac{4}{9} + \frac{1}{9} + \frac{1}{9} = \frac{2}{3}.$$

(e) In this case, $\mathbf{C} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$ so that

$$\mathbf{K} \quad = \quad \mathbf{A}^T\mathbf{C}\mathbf{A} = \begin{pmatrix} 7 & 2 \\ 2 & \frac{4}{3} \end{pmatrix}$$

and

$$\mathbf{f} \quad = \quad \mathbf{A}^T\mathbf{C}\mathbf{b} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}.$$

3. (30 points, 10 each) Consider the function

$$p(x,y,z) = x^2 - 4xy + 6xz + 8x + 8y^2 - 20yz - 24y + 14z^2 + 34z + 5.$$

(a) This function can be written as $p(x,y,z) = \boldsymbol{x}^T\mathbf{K}\boldsymbol{x} - 2\boldsymbol{x}^T\boldsymbol{f} + c$, where $\boldsymbol{x} = (x,y,z)^T$. What are $\mathbf{K}$, $\boldsymbol{f}$, and c?

(b) Find all critical points of this function.

(c) For each critical point, say whether it corresponds to a (local) minimum, maximum, or saddle point. Explain your reasoning.

Solution:

(a) The matrix $\mathbf{K}$ can be obtained by computing the Hessian and dividing by two, or by inspection. It is

$$\mathbf{K} = \begin{pmatrix} 1 & -2 & 3 \\ -2 & 8 & -10 \\ 3 & -10 & 14 \end{pmatrix}.$$

The $\boldsymbol{f}$ vector can similarly be obtained by inspection, or by computing the constant part of the gradient and dividing by -2. It is

$$\boldsymbol{f} = \begin{pmatrix} -4 \\ 12 \\ -17 \end{pmatrix}.$$

Finally, $c = 5$.

(b) Critical points are solutions of

$$\mathbf{K}\boldsymbol{x} = \boldsymbol{f}.$$

The row-reduced augmented matrix

$$\left(\begin{array}{ccc|c} 1 & -2 & 3 & -4 \\ -2 & 8 & -10 & 12 \\ 3 & -10 & 14 & -17 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -2 & 3 & -4 \\ 0 & 4 & -4 & 4 \\ 0 & -4 & 5 & -5 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & -2 & 3 & -4 \\ 0 & 4 & -4 & 4 \\ 0 & 0 & 1 & -1 \end{array}\right)$$

Back-solving yields the solution $x = -1, y = 0, z = -1$. This is the only critical point.

(c) The type of critical point depends on the Hessian, which is $2\mathbf{K}$, where $\mathbf{K}$ is given above. To check if the Hessian is positive definite we row reduce. Note that the row reduction follows the same steps as in part (b), but everything is multiplied by 2.

$$\begin{pmatrix} 2 & -4 & 6 \\ -4 & 16 & -20 \\ 6 & -20 & 28 \end{pmatrix} \to \begin{pmatrix} 2 & -4 & 6 \\ 0 & 8 & -8 \\ 0 & -8 & \text{‘}0 \end{pmatrix} \to \begin{pmatrix} 2 & -4 & 6 \\ 0 & 8 & -8 \\ 0 & 0 & 2 \end{pmatrix}$$

Since we were able to row reduce without pivoting, and since the diagonal elements of the row echelon form are positive, the Hessian is positive definite, which implies that the critical point occurs at a minimum of the function.

4. (12 points) Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and let $\mathbf{Q}$ be an $n \times n$ orthogonal matrix. Prove that the (Euclidean) angle between $\boldsymbol{x}$ and $\boldsymbol{y}$ is the same as the (Euclidean) angle between $\mathbf{Q}\boldsymbol{x}$ and $\mathbf{Q}\boldsymbol{y}$.

   Solution: The Euclidean angle θ between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is a solution of

$$\cos(\theta) = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\sqrt{\boldsymbol{x}^T\boldsymbol{x}\boldsymbol{y}^T\boldsymbol{y}}}.$$

   Using this same formula but replacing $\boldsymbol{x}$ with $\mathbf{Q}\boldsymbol{x}$ and replacing $\boldsymbol{y}$ with $\mathbf{Q}|bmy$ yields

$$\frac{\boldsymbol{x}^T\mathbf{Q}^T\mathbf{Q}\boldsymbol{y}}{\sqrt{\boldsymbol{x}^T\mathbf{Q}^T\mathbf{Q}\boldsymbol{x}\boldsymbol{y}\mathbf{Q}^T\mathbf{Q}^T\boldsymbol{y}}} = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\sqrt{\boldsymbol{x}^T\boldsymbol{x}\boldsymbol{y}^T\boldsymbol{y}}}$$

   since $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$.

   Equivalently, you can state that orthogonal matrices preserve the norm, and then show that $\boldsymbol{x}^T\boldsymbol{y} = \boldsymbol{x}^T\mathbf{Q}^T\mathbf{Q}\boldsymbol{y}$, i.e. orthogonal matrices preserve the dot product. Together these imply that orthogonal matrices preserve angles.

5. **Bonus**

   (i) (6 points, 2 each) Prove that the Householder reflection matrix $\mathbf{H} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^T$, where $\mathbf{u}$ is a unit vector, is (a) its own inverse, (b) symmetric, and (c) orthogonal.

   (ii) (2 points) Prove that orthogonal matrices preserve length. (You can use Euclidean length.)

   (iii) (2 points) Prove that the inverse of an orthogonal matrix is also orthogonal.

   Solution:

   (i) a) To see that $\mathbf{H} = \mathbf{H}^{-1}$,

$$\begin{aligned}
\mathbf{H}\mathbf{H} &= (\mathbf{I} - 2\mathbf{u}\mathbf{u}^T)(\mathbf{I} - 2\mathbf{u}\mathbf{u}^T) \\
&= \mathbf{I} - 2\mathbf{u}\mathbf{u}^T - 2\mathbf{u}\mathbf{u}^T + 4\mathbf{u}\underbrace{\mathbf{u}^T\mathbf{u}}_{1}\mathbf{u}^T \\
&= \mathbf{I} - 4\mathbf{u}\mathbf{u}^T + 4\mathbf{u}\mathbf{u}^T = \mathbf{I}.
\end{aligned}$$

   Since $\mathbf{H}$ is both a left and right inverse of $\mathbf{H}$, we know that $\mathbf{H} = \mathbf{H}^{-1}$.

   b) To see that $\mathbf{H}$ is symmetric,

$$\begin{aligned}
\mathbf{H}^T &= (\mathbf{I} - 2\mathbf{u}\mathbf{u}^T)^T \\
&= \mathbf{I}^T - 2(\mathbf{u}\mathbf{u}^T)^T \\
&= \mathbf{I} - 2\mathbf{u}\mathbf{u}^T = \mathbf{H}.
\end{aligned}$$

   c) To see that $\mathbf{H}$ is orthogonal, we combine the facts that it is its own inverse and it is symmetric to get

$$\mathbf{H}\mathbf{H}^T = \mathbf{H}\mathbf{H} = \mathbf{I}$$

   and

$$\mathbf{H}^T\mathbf{H} = \mathbf{H}\mathbf{H} = \mathbf{I}.$$

(ii) Let $\mathbf{Q} \in \mathbb{R}^{n\times n}$ be orthogonal and let $\mathbf{x} \in \mathbb{R}^n$. Then

$$\begin{aligned} \|\mathbf{Qx}\|^2 &= \mathbf{x}^T \underbrace{\mathbf{Q}^T\mathbf{Q}}_{\mathbf{I}}\mathbf{x} \\ &= \mathbf{x}^T\mathbf{x} \\ &= \|\mathbf{x}\|^2. \end{aligned}$$

Taking square roots gives the result.

(iii) Let $\mathbf{Q}$ be orthogonal and define $\mathbf{B} = \mathbf{Q}^{-1}$. Then

$$\begin{aligned} \mathbf{B}^T\mathbf{B} &= (\mathbf{Q}^{-1})^T(\mathbf{Q}^{-1}) \\ &= (\mathbf{Q}^T)^{-1}\mathbf{Q}^{-1} \\ &= (\mathbf{Q}^{-1})^{-1}\mathbf{Q}^{-1} \text{ by } \mathbf{Q}^T = \mathbf{Q}^{-1} \\ &= \mathbf{Q}\mathbf{Q}^{-1} = \mathbf{I} \text{ by definition of the inverse.} \end{aligned}$$

Also, we have

$$\begin{aligned} \mathbf{B}\mathbf{B}^T &= \mathbf{Q}^{-1}(\mathbf{Q}^{-1})^T \\ &= \mathbf{Q}^{-1}(\mathbf{Q}^T)^{-1} \\ &= \mathbf{Q}^{-1}(\mathbf{Q}^{-1})^{-1} \\ &= \mathbf{Q}^{-1}\mathbf{Q} \\ &= \mathbf{Q}^{-1}\mathbf{Q} = \mathbf{I} \text{ by definition of the inverse.} \end{aligned}$$

So $\mathbf{B} = \mathbf{Q}^{-1}$ is also orthogonal.