Question 1

If Q is an $n \times n$ orthogonal matrix, then $Q^\top Q = QQ^\top = I_n$ where I_n is the $n \times n$ identity matrix.

Part A

Let Q_1 and Q_2 be orthogonal matrices. Then,

$$\begin{aligned}(Q_1Q_2)^\top(Q_1Q_2) &= Q_2^\top \cancel{Q_1^\top Q_1} Q_2 = Q_2^\top Q_2 = I_n \\ &= (Q_1Q_2)(Q_1Q_2)^\top = Q_1 \cancel{Q_2 Q_2^\top} Q_1^\top = Q_1 Q_1^\top = I_n.\end{aligned}$$

Hence, Q_1Q_2 must be orthogonal.

Part B

Suppose $Q^\top$ is an orthogonal matrix. Then,

$$\begin{aligned}(Q^\top)^\top Q^\top &= QQ^\top = I_n \\ &= Q^\top (Q^\top)^\top = Q^\top Q = I_n.\end{aligned}$$

As such, $Q^\top$ must be orthogonal. Subsequently, if Q is an orthogonal matrix (where its column form an orthonormal basis), then its rows must also form an orthonormal basis.

Proof: Suppose Q is made up of column vectors $\vec{\mathbf{q}}_i$ for each column $i = 1, \ldots, n$. If Q is an $n \times n$ orthogonal matrix, then

$$Q^\top Q = \begin{pmatrix} — \vec{\mathbf{q}}_1^\top — \\ \vdots \\ — \vec{\mathbf{q}}_n^\top — \end{pmatrix} \begin{pmatrix} | & & | \\ \vec{\mathbf{q}}_1 & \cdots & \vec{\mathbf{q}}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix} = I_n.$$

Since the columns $\vec{\mathbf{q}}_i$ form an orthonormal basis of $\mathbb{R}^n$, the value of the diagonal entries (i, i) is $\langle \vec{\mathbf{q}}_i^\top, \vec{\mathbf{q}}_i \rangle = 1$ and the value of all other entries (i, j) where $i \neq j$ is $\langle \vec{\mathbf{q}}_i^\top, \vec{\mathbf{q}}_j \rangle = 0$. Thus, $||\vec{\mathbf{q}}_i|| = ||\vec{\mathbf{q}}_i^\top|| = 1$ for all rows i.

Part C

Let $Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Then, $Q^\top = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. As such,

$$\begin{aligned} QQ^\top &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \sin\theta\cos\theta \\ \sin\theta\cos\theta - \sin\theta\cos\theta & \sin^2\theta + \cos^2\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= Q^\top Q \\ &= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos^2\theta + \sin^2\theta & -\cos\theta\sin\theta + \sin\theta\cos\theta \\ -\sin\theta\cos\theta + \sin\theta\cos\theta & \sin^2\theta + \cos^2\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= I_2. \end{aligned}$$

Thus, $Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is an orthogonal matrix. This special matrix is known as the rotational matrix. Let $\vec{\mathbf{v}} = \begin{pmatrix} x \\ y \end{pmatrix}$ be a vector on the Cartesian plane. Applying Q to a vector $\vec{\mathbf{v}}$ will rotate it counterclockwise by θ.

$$Q\vec{\mathbf{v}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}$$

For example, let $\vec{\mathbf{v}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ be a vector on an xy-plane. We can flip the vector across y-axis by rotating the $\vec{\mathbf{v}}$ by $\theta = \pi$.

$$-\vec{\mathbf{v}} = \begin{pmatrix} \cos\pi & -\sin\pi \\ \sin\pi & \cos\pi \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1(\cos\pi) - 0(\sin\pi) \\ 1(\sin\pi) + 0(\cos\pi) \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Part D

By definition, the norm of a vector $\vec{\mathbf{x}} \in \mathbb{R}^n$ is given by $||\vec{\mathbf{x}}|| = \sqrt{\langle \vec{\mathbf{x}}, \vec{\mathbf{x}} \rangle} = \sqrt{\vec{\mathbf{x}}^\top \vec{\mathbf{x}}}$. If Q is an orthogonal matrix, then:

$$\begin{aligned} ||Q\vec{\mathbf{x}}|| &= \sqrt{\langle Q\vec{\mathbf{x}}, Q\vec{\mathbf{x}} \rangle} \\ &= \sqrt{(Q\vec{\mathbf{x}})^\top (Q\vec{\mathbf{x}})} \\ &= \sqrt{\vec{\mathbf{x}}^\top \cancel{Q^\top Q} \vec{\mathbf{x}}} \\ &= \sqrt{\langle \vec{\mathbf{x}}, \vec{\mathbf{x}} \rangle} \\ &= ||\vec{\mathbf{x}}||. \end{aligned}$$

Question 2

Refer to section `%% Question 2` in the `math425hw4.m` file for the relevant code.

Parts A and B

See snippets provided under `% 2(a)` and `% 2(b)`. The results are saved in variable named `x_gauss_n` and `x_qr_n` for the respective $n \times n$ Hilbert matrices and the computation method. The results are omitted here for brevity.

Part C

By subtracting the expected values for $\vec{\mathbf{x}}_n^*$ for the respective $n \times n$ Hilbert matrices, we can find the error between the computed and expected solutions $\Delta\vec{\mathbf{x}}_n^*$. For the purpose of illustration, we can use the norm to see how the difference grows, as it also disregards the alternating signs for each entries.

	$n = 5$	$n = 10$	$n = 20$
$\lvert\lvert\vec{\mathbf{x}}^* - \vec{\mathbf{x}}_{\text{Gauss},n}^*\rvert\rvert$	$1.6234\ldots \times 10^{-12}$	$1.5150\ldots \times 10^{-04}$	$12.9343\ldots$
$\lvert\lvert\vec{\mathbf{x}}^* - \vec{\mathbf{x}}_{\text{QR},n}^*\rvert\rvert$	$4.6829\ldots \times 10^{-12}$	$7.5286\ldots \times 10^{-05}$	$59.1000\ldots$

Here, we can see that the method using Gaussian elimination suffers from numerical instability greatly, especially with ill-conditioned matrices such as the Hilbert matrix. At the cost of complexity, the QR factorization approach is more stable and produces far more reliable results for larger n.

Question 3

Refer to section `%% Question 3` in the `math425hw4.m` file for the relevant code.