STAT 30900 : MATHEMATICAL COMPUTATIONS I

Fall 2019

•

HOMEWORK 3

•

Solutions by

JINHONG DU

12243476

1

Let $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} \neq \mathbf{0}$. A *Householder* matrix $H_\mathbf{u} \in \mathbb{R}^{n \times n}$ is defined by

$$H_\mathbf{u} = I - \frac{2\mathbf{u}\mathbf{u}^\top}{\|\mathbf{u}\|_2^2}.$$

(a) Show that $H_\mathbf{u}$ is both symmetric and orthogonal.

Proof.

$$\begin{aligned} H_\mathbf{u}^\top &= \left(I - \frac{2\mathbf{u}\mathbf{u}^\top}{\|\mathbf{u}\|_2^2}\right)^\top \\ &= I^\top - \left(\frac{2\mathbf{u}\mathbf{u}^\top}{\|\mathbf{u}\|_2^2}\right)^\top \\ &= I - \frac{2(\mathbf{u}\mathbf{u}^\top)^\top}{\|\mathbf{u}\|_2^2} \\ &= I - \frac{2\mathbf{u}\mathbf{u}^\top}{\|\mathbf{u}\|_2^2} \\ &= H_\mathbf{u}, \end{aligned}$$

i.e., $H_\mathbf{u}$ is symmetric.

$$\begin{aligned} H_\mathbf{u}^\top H_\mathbf{u} &= \left(I - \frac{2\mathbf{u}\mathbf{u}^\top}{\|\mathbf{u}\|_2^2}\right)^2 \\ &= I - \frac{4\mathbf{u}\mathbf{u}^\top}{\|\mathbf{u}\|_2^2} + \frac{2\mathbf{u}\mathbf{u}^\top}{\|\mathbf{u}\|_2^2} \cdot \frac{2\mathbf{u}\mathbf{u}^\top}{\|\mathbf{u}\|_2^2} \\ &= I - \frac{4\mathbf{u}\mathbf{u}^\top}{\|\mathbf{u}\|_2^2} + \frac{4\mathbf{u}\mathbf{u}^\top}{\|\mathbf{u}\|_2^2} \\ &= I, \end{aligned}$$

i.e., $H_\mathbf{u}$ is orthogonal. $\square$

(b) Show that for any $\alpha \in \mathbb{R}$, $\alpha \neq 0$,

$$H_{\alpha\mathbf{u}} = H_\mathbf{u}.$$

In other words, $H_\mathbf{u}$ only depends on the 'direction' of $\mathbf{u}$ and not on its 'magnitude'.

Proof.

$$\begin{aligned} H_{\alpha\mathbf{u}} &= I - \frac{2(\alpha\mathbf{u})(\alpha\mathbf{u})^\top}{\|\alpha\mathbf{u}\|_2^2} \\ &= I - \frac{2\alpha^2\mathbf{u}\mathbf{u}^\top}{\alpha^2\|\mathbf{u}\|_2^2} \\ &= H_\mathbf{u} \end{aligned}$$

$\square$

(c) In general, given a matrix $M \in \mathbb{R}^{n \times n}$ and a vector $\mathbf{x} \in \mathbb{R}^n$, computing the matrix-vector product $M\mathbf{x}$ requires n inner products — one for each row of M with $\mathbf{x}$. Show that $H_\mathbf{u}\mathbf{x}$ can be computed using only two inner products.

Proof. Since

$$
\begin{aligned}
H_{\mathbf{u}}\mathbf{x} &= \left(I - \frac{2\mathbf{u}\mathbf{u}^\top}{\|\mathbf{u}\|_2^2}\right)\mathbf{x} \\
&= \mathbf{x} - \frac{2\mathbf{u}\mathbf{u}^\top\mathbf{x}}{\|\mathbf{u}\|_2^2},
\end{aligned}
$$

it can be computed using only two inner products $\langle \mathbf{u}, \mathbf{x}\rangle$ and $\langle \mathbf{u}, \mathbf{u}\rangle$. □

(d) Given $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ where $\mathbf{a} \neq \mathbf{b}$ and $\|\mathbf{a}\|_2 = \|\mathbf{b}\|_2$. Find $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} \neq \mathbf{0}$ such that

$$H_{\mathbf{u}}\mathbf{a} = \mathbf{b}.$$

Since $\mathbf{a} \neq \mathbf{b}$ and

$$H_{\mathbf{u}}\mathbf{a} = \mathbf{a} - \frac{2\mathbf{u}\mathbf{u}^\top\mathbf{a}}{\|\mathbf{u}\|_2^2} = \mathbf{b},$$

we have

$$\frac{2\mathbf{u}^\top\mathbf{a}}{\|\mathbf{u}\|_2^2}\mathbf{u} = \mathbf{a} - \mathbf{b} \neq \mathbf{0}.$$

Notice that $\frac{2\mathbf{u}^\top\mathbf{a}}{\|\mathbf{u}\|_2^2}$ is a scalar, we must have $\mathbf{u} = c(\mathbf{a} - \mathbf{b})$ for some $c \in \mathbb{R}$ and $c \neq 0$. Also, $\frac{2\mathbf{u}^\top\mathbf{a}}{\|\mathbf{u}\|_2^2}\mathbf{u}$ is invariant for $c\mathbf{u}$, $\forall\ c \neq 0$. So $\mathbf{u} = c(\mathbf{a} - \mathbf{b})$ for all $c \neq 0$ satisfies $H_{\mathbf{u}}\mathbf{a} = \mathbf{b}$.
The condition $\|\mathbf{a}\|_2 = \|\mathbf{b}\|_2$ insures that such $\mathbf{u}$ must exist. To see this, since

$$
\begin{aligned}
(\mathbf{a} - \mathbf{b})^\top\mathbf{a} &= \|\mathbf{a}\|_2^2 - \mathbf{b}^\top\mathbf{a} \\
&= \|\mathbf{b}\|_2^2 - \mathbf{a}^\top\mathbf{b} \\
&= (\mathbf{a} - \mathbf{b})^\top(-\mathbf{b})
\end{aligned}
$$

we have

$$
\begin{aligned}
H_{c(\mathbf{a}-\mathbf{b})}\mathbf{a} &= \mathbf{a} - \frac{2(\mathbf{a} - \mathbf{b})^\top\mathbf{a}}{\|\mathbf{a} - \mathbf{b}\|_2^2}(\mathbf{a} - \mathbf{b}) \\
&= \mathbf{a} - \frac{(\mathbf{a} - \mathbf{b})^\top\mathbf{a} + (\mathbf{a} - \mathbf{b})^\top(-\mathbf{b})}{\|\mathbf{a} - \mathbf{b}\|_2^2}(\mathbf{a} - \mathbf{b}) \\
&= \mathbf{a} - (\mathbf{a} - \mathbf{b}) \\
&= \mathbf{b}
\end{aligned}
$$

(e) Show that $\mathbf{u}$ is an eigenvector of $H_{\mathbf{u}}$. What is the corresponding eigenvalue?

Proof. Since

$$
\begin{aligned}
H_{\mathbf{u}}\mathbf{u} &= \mathbf{u} - \frac{2\mathbf{u}\mathbf{u}^\top\mathbf{u}}{\|\mathbf{u}\|_2^2} \\
&= \mathbf{u} - 2\mathbf{u} \\
&= -\mathbf{u},
\end{aligned}
$$

we know that $\mathbf{u}$ is an eigenvector of $H_{\mathbf{u}}$ with respect to eigenvalue -1. □

(f) Show that every $\mathbf{v} \in \operatorname{span}\{\mathbf{u}\}^\perp$ (cf. orthogonal complement in Homework **1**) is an eigenvector of $H_\mathbf{u}$. What are the corresponding eigenvalues? What is $\dim(\operatorname{span}\{\mathbf{u}\}^\perp)$?

Proof. $\forall\ \mathbf{v} \in \operatorname{span}\{\mathbf{u}\}^\perp$, $\mathbf{u}^\top \mathbf{v} = 0$. So

$$\begin{aligned} H_\mathbf{u}\mathbf{v} &= \mathbf{v} - \frac{2\mathbf{u}\mathbf{u}^\top \mathbf{v}}{\|\mathbf{u}\|_2^2} \\ &= \mathbf{v}, \end{aligned}$$

i.e., $\mathbf{v}$ is an eigenvector of $H_\mathbf{u}$ with respect to eigenvalue 1.

$$\begin{aligned} \dim(\operatorname{span}\{\mathbf{u}\}^\perp) &= \dim(\mathbb{R}^n) - \dim(\operatorname{span}\{\mathbf{u}\}) \\ &= n - 1 \end{aligned}$$

□

(g) Find the eigenvalue decomposition of $H_\mathbf{u}$, i.e., find an orthogonal matrix Q and a diagonal matrix Λ such that

$$H_\mathbf{u} = Q\Lambda Q^\top.$$

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}\}$ be the orthonormal basis of $\operatorname{span}\{\mathbf{u}\}^\perp$. Then $\{\frac{\mathbf{u}}{\|\mathbf{u}\|_2}, \mathbf{v}_1, \ldots, \mathbf{v}_{n-1}\}$ form the orthonormal basis of $\mathbb{R}^n$. Let $Q = \begin{bmatrix} \frac{\mathbf{u}}{\|\mathbf{u}\|_2} & \mathbf{v}_1 & \cdots & \mathbf{v}_{n-1} \end{bmatrix}$, $\Lambda = \operatorname{diag}(-1, 1, \cdots, 1)$, then the columns of Q are eigenvectors of $H_\mathbf{u}$, Q is unitary and $H_\mathbf{u} = Q\Lambda Q^\top$

2

Let $A \in \mathbb{R}^{m\times n}$ and suppose its complete orthogonal decomposition is given by

$$A = Q_1 \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top,$$

where Q_1 and Q_2 are orthogonal, and L is an nonsingular lower triangular matrix. Recall that $X \in \mathbb{R}^{n\times m}$ is the unique pseudo-inverse of A if the following Moore–Penrose conditions hold:

(i) $AXA = A$,

(ii) $XAX = X$,

(iii) $(AX)^\top = AX$,

(iv) $(XA)^\top = XA$

and in which case we write $X = A^\dagger$.

(a) Let

$$A^- = Q_2 \begin{bmatrix} L^{-1} & Y \\ 0 & 0 \end{bmatrix} Q_1^\top, \qquad Y \neq 0.$$

Which of the four conditions (i)–(iv) are satisfied?

Since

$$\begin{aligned} AA^- &= Q_1 \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top Q_2 \begin{bmatrix} L^{-1} & Y \\ 0 & 0 \end{bmatrix} Q_1^\top \\ &= Q_1 \begin{bmatrix} I & LY \\ 0 & 0 \end{bmatrix} Q_1^\top \\ A^-A &= Q_2 \begin{bmatrix} L^{-1} & Y \\ 0 & 0 \end{bmatrix} Q_1^\top Q_1 \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top \\ &= Q_2 \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top, \end{aligned}$$

we have

$$\begin{aligned} AA^-A &= Q_1 \begin{bmatrix} I & LY \\ 0 & 0 \end{bmatrix} Q_1^\top Q_1 \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top \\ &= Q_1 \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top = A \\ A^-AA^- &= Q_2 \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top Q_2 \begin{bmatrix} L^{-1} & Y \\ 0 & 0 \end{bmatrix} Q_1^\top \\ &= Q_2 \begin{bmatrix} L^{-1} & Y \\ 0 & 0 \end{bmatrix} Q_1^\top = A^- \\ (AA^-)^\top &= Q_1 \begin{bmatrix} I & 0 \\ Y^\top L^\top & 0 \end{bmatrix} Q_1^\top \neq AA^- \\ (A^-A)^\top &= Q_2 \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top = A^-A, \end{aligned}$$

Solution (cont.)

i.e. A^- statisfies condition (i) (ii) and (iv).

(b) Prove that

$$A^\dagger = Q_2 \begin{bmatrix} L^{-1} & 0 \\ 0 & 0 \end{bmatrix} Q_1^\top$$

by letting

$$A^\dagger = Q_2 \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} Q_1^\top$$

and by completing the following steps:

- Using (i), prove that $X_{11} = L^{-1}$.
- Using the symmetry conditions (iii) and (iv), prove that $X_{12} = 0$ and $X_{21} = 0$.
- Using (ii), prove that $X_{22} = 0$.

Proof. Since $AA^\dagger A = A$, we have

$$\begin{aligned} AA^\dagger A &= Q_1 \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top Q_2 \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} Q_1^\top Q_1 \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top \\ &= Q_1 \begin{bmatrix} LX_{11} & LX_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top \\ &= Q_1 \begin{bmatrix} LX_{11}L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top \end{aligned}$$

$$Q_1 \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top.$$

Since Q_1 and Q_2 are orthonormal matrix, we have $LX_{11}L = L$. Also, L is nonsingular, so $X_{11} = L^{-1}$.
Since $(A^\dagger A)^\top = A^\dagger A$,

$$\begin{aligned} A^\dagger A &= Q_2 \begin{bmatrix} L^{-1} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} Q_1^\top Q_1 \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top \\ &= Q_2 \begin{bmatrix} I & 0 \\ X_{21}L & 0 \end{bmatrix} Q_2^\top \\ \left(A^\dagger A\right)^\top &= Q_2 \begin{bmatrix} I & L^{-1}X_{21}^\top \\ 0 & 0 \end{bmatrix} Q_2 \end{aligned}$$

we have $X_{21}L = 0$, i.e. $X_{21} = 0$.
Since $(AA^\dagger)^\top = AA^\dagger$,

$$\begin{aligned} AA^\dagger &= Q_1 \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top Q_2 \begin{bmatrix} L^{-1} & X_{12} \\ 0 & X_{22} \end{bmatrix} Q_1^\top \\ &= Q_1 \begin{bmatrix} I & LX_{12} \\ 0 & 0 \end{bmatrix} Q_1^\top \\ \left(AA^\dagger\right)^\top &= Q_2 \begin{bmatrix} I & 0 \\ X_{12}^\top L^{-1} & 0 \end{bmatrix} Q_1 \end{aligned}$$

Solution (cont.)

we have $X_{12}L = 0$, i.e. $X_{12} = 0$.
Since $A^\dagger A A^\dagger = A^\dagger$,

$$\begin{aligned} A^\dagger A A^\dagger &= Q_2 \begin{bmatrix} L^{-1} & 0 \\ 0 & X_{22} \end{bmatrix} Q_1^\top Q_1 \begin{bmatrix} L & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top Q_2 \begin{bmatrix} L^{-1} & 0 \\ 0 & X_{22} \end{bmatrix} Q_1^\top \\ &= Q_2 \begin{bmatrix} L^{-1} & 0 \\ 0 & 0 \end{bmatrix} Q_2^\top \\ A^\dagger &= Q_2 \begin{bmatrix} L^{-1} & 0 \\ 0 & X_{22} \end{bmatrix} Q_2^\top, \end{aligned}$$

we have $X_{22} = 0$.
Therefore, $A^\dagger = Q_2 \begin{bmatrix} L^{-1} & 0 \\ 0 & 0 \end{bmatrix} Q_1^\top$. $\square$

3

Let $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $\mathbf{c} \in \mathbb{R}^n$. We are interested in the least squares problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|A\mathbf{x} - \mathbf{b}\|_2^2. \tag{1}$$

(a) Show that $\mathbf{x}$ is a solution to (1) if and only if $\mathbf{x}$ is a solution to the *augmented system*

$$\begin{bmatrix} I & A \\ A^\top & 0 \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}. \tag{2}$$

Proof. Since

$$\begin{aligned} \|A\mathbf{x} - \mathbf{b}\|_2^2 &= \|A\mathbf{x} - P_{\mathrm{im}(A)}\mathbf{b} - P_{\mathrm{im}(A)^\perp}\mathbf{b}\|_2^2 \\ &= \|A\mathbf{x} - P_{\mathrm{im}(A)}\mathbf{b}\|_2^2 + \|P_{\mathrm{im}(A)^\perp}\mathbf{b}\|_2^2 \\ &\geq \|P_{\mathrm{im}(A)^\perp}\mathbf{b}\|_2^2 \\ &= \|P_{\ker(A^\top)}\mathbf{b}\|_2^2 \end{aligned}$$

and the equality holds when $A\mathbf{x} = P_{\mathrm{im}(A)}\mathbf{b}$, which is attainable since $P_{\mathrm{im}(A)}\mathbf{b} \in \mathrm{im}(A)$. So $\mathbf{x}$ is the solution if and only if

$$A^\top(\mathbf{b} - A\mathbf{x}) = \mathbf{0}$$

which is equal to

$$\begin{cases} \mathbf{r} + A\mathbf{x} = \mathbf{b} \\ A^\top \mathbf{r} = \mathbf{0} \end{cases}$$

by setting $\mathbf{r} = \mathbf{b} - A\mathbf{x}$, i.e.,

$$\begin{bmatrix} I & A \\ A^\top & 0 \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}.$$

□

(b) Show that the $(m+n)\times(m+n)$ matrix in (2) is nonsingular if and only if A has full column rank.

Proof. Let $B = \begin{bmatrix} I & A \\ A^\top & 0 \end{bmatrix}$, $\mathbf{a}_i$ be the ith columnof A.

$\Longrightarrow$

Suppose that A does not have full column rank, then there exists $c_1, \ldots, c_n$ such that at least one c_i is nonzero and $\sum_{i=1}^n c_i\mathbf{a}_i = \mathbf{0}_m$. Then $\sum_{i=1}^n c_i \begin{bmatrix} \mathbf{a}_i \\ \mathbf{0}_n \end{bmatrix} = \mathbf{0}_{m+n}$, which means that the last n columns of B are linear dependent. Since $B \in \mathbb{R}^{(m+n)\times(m+n)}$ is nonsingular, all columns of B should be linear independent. Contradiction. So A has full column rank.

$\Longleftarrow$

Since A has full column rank, $A^\top A$ also has full rank and thus invertible. By elementary row transfor-

Solution (cont.)

mation,

$$\begin{bmatrix} I & A \\ A^\top & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} I & A \\ 0 & -A^\top A \end{bmatrix} \longrightarrow \begin{bmatrix} I & 0 \\ 0 & -A^\top A \end{bmatrix}$$

rank of these matrices remains the same. While the diagonal block matrix $\begin{bmatrix} I & 0 \\ 0 & A^\top A \end{bmatrix}$ has rank $\text{rank}(I)+\text{rank}(A^\top A) = m+n$, so does B. So $\begin{bmatrix} I & A \\ A^\top & 0 \end{bmatrix}$ is nonsingular.

□

(c) Suppose A has full column rank and the QR decomposition of A is

$$A = Q\begin{bmatrix} R \\ 0 \end{bmatrix}.$$

Show that the solution to the augmented system

$$\begin{bmatrix} I & A \\ A^\top & 0 \end{bmatrix}\begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{c} \end{bmatrix}$$

can be computed from

$$\mathbf{z} = R^{-\top}\mathbf{c}, \qquad \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} = Q^\top \mathbf{b},$$

and

$$\mathbf{x} = R^{-1}(\mathbf{d}_1 - \mathbf{z}), \qquad \mathbf{y} = Q\begin{bmatrix} \mathbf{z} \\ \mathbf{d}_2 \end{bmatrix}.$$

Proof. Since

$$\begin{bmatrix} Q^\top & \\ & R^{-\top} \end{bmatrix}\begin{bmatrix} I & A \\ A^\top & 0 \end{bmatrix}\begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} Q^\top & \begin{bmatrix} R \\ 0 \end{bmatrix} \\ \begin{bmatrix} I_n & 0 \end{bmatrix} Q^\top & 0 \end{bmatrix}\begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix}$$

$$= \begin{bmatrix} Q^\top & \\ & R^{-\top} \end{bmatrix}\begin{bmatrix} \mathbf{b} \\ \mathbf{c} \end{bmatrix} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \mathbf{z} \end{bmatrix}$$

Therefore,

$$\begin{cases} Q^\top \mathbf{y} + \begin{bmatrix} R \\ 0 \end{bmatrix}\mathbf{x} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} & (1) \\ \begin{bmatrix} I_n & 0 \end{bmatrix} Q^\top \mathbf{y} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} & (2) \end{cases}$$

Multiplying $\begin{bmatrix} I_n & 0 \end{bmatrix}$ on (1), we have

$$R\mathbf{x} = \mathbf{d}_1 - \begin{bmatrix} I_n & 0 \end{bmatrix} Q^\top \mathbf{y} = \mathbf{d}_1 - \mathbf{z}.$$

Solution (cont.)

Since A has full column rank, R also has full column rank and is invertible. So

$$\mathbf{x} = R^{-1}(\mathbf{d}_1 - \mathbf{z}).$$

Then

$$Q^\top \mathbf{y} + \begin{bmatrix} \mathbf{d}_1 - \mathbf{z} \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix},$$

i.e.

$$\mathbf{y} = Q \begin{bmatrix} \mathbf{z} \\ \mathbf{d}_2 \end{bmatrix}.$$

□

(d) Hence deduce that if A has full column rank, then

$$A^\dagger = R^{-1} Q_1^\top$$

where $Q = [Q_1, Q_2]$ with $Q_1 \in \mathbb{R}^{m\times n}$ and $Q_2 \in \mathbb{R}^{m\times(m-n)}$. Check that this agrees with the general formula derived for a rank-retaining factorization $A = GH$ in the lectures.

Proof. Since $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$,

$$\begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{bmatrix} = \begin{bmatrix} Q_1^\top \\ Q_2^\top \end{bmatrix} \mathbf{b} = \begin{bmatrix} Q_1^\top \mathbf{b} \\ Q_2^\top \mathbf{b} \end{bmatrix}.$$

For $\mathbf{c} = \mathbf{0}_n$ in (c), we have that

$$\mathbf{x} = R^{-1}\mathbf{d}_1 = R^{-1} Q_1^\top \mathbf{b}$$

so $A^\dagger = R^{-1} Q_1^\top$.
Since $A = Q_1 R$ is the rank retaining decomposition of A, we have that the minimum length solution is given by

$$\begin{aligned} \mathbf{x} &= R^\top (RR^\top)^{-1} (Q_1^\top Q_1)^{-1} Q_1^\top \mathbf{b} \\ &= R^\top R^{-\top} R^{-1} Q_1^\top \mathbf{b} \\ &= R^{-1} Q_1^\top \mathbf{b}. \end{aligned}$$

□

4

Let $A \in \mathbb{R}^{m \times n}$. Suppose we apply QR with column pivoting to obtain the decomposition

$$A = Q \begin{bmatrix} R & S \\ 0 & 0 \end{bmatrix} \Pi^\top$$

where Q is orthogonal and R is upper triangular and invertible. Let $\mathbf{x}_B$ be the *basic solution*, i.e.,

$$\mathbf{x}_B = \Pi \begin{bmatrix} R^{-1} & 0 \\ 0 & 0 \end{bmatrix} Q^\top \mathbf{b},$$

and let $\hat{\mathbf{x}} = A^\dagger \mathbf{b}$. Show that

$$\frac{\|\mathbf{x}_B - \hat{\mathbf{x}}\|_2}{\|\hat{\mathbf{x}}\|_2} \leq \|R^{-1}S\|_2.$$

(*Hint*: If

$$\Pi^\top \mathbf{x} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} \quad \text{and} \quad Q^\top \mathbf{b} = \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix},$$

consider the associated linearly constrained least-squares problem

$$\min \|\mathbf{u}\|_2^2 + \|\mathbf{v}\|_2^2 \quad \text{s.t. } R\mathbf{u} + S\mathbf{v} = \mathbf{c}$$

and write down the augmented system for the constrained problem.)

Proof. Suppose $\operatorname{rank}(A) = r$. Since $A\mathbf{x} = \mathbf{b}$ equals to $Q \begin{bmatrix} R & S \\ 0 & 0 \end{bmatrix} \Pi^\top \mathbf{x} = \mathbf{b}$, i.e., $\begin{bmatrix} R & S \\ 0 & 0 \end{bmatrix} \Pi^\top \mathbf{x} = Q^\top \mathbf{b}$. Let $\Pi^\top \mathbf{x} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$ and $Q^\top \mathbf{b} = \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix}$ where $\mathbf{u}, \mathbf{c} \in \mathbb{R}^r$ and $\mathbf{v}, \mathbf{d} \in \mathbb{R}^{n-r}$, then $\begin{bmatrix} R\mathbf{u} + S\mathbf{v} \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix}$. So the optimization problem $\min\limits_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|_2$ equals to

$$\begin{aligned} \min\ & \|\mathbf{u}\|_2^2 + \|\mathbf{v}\|_2^2 \\ \text{s.t. }\ & R\mathbf{u} + S\mathbf{v} = \mathbf{c} \end{aligned}$$

Also, the solution to these two problems is $\hat{\mathbf{x}} = \Pi \begin{bmatrix} \hat{\mathbf{u}} \\ \hat{\mathbf{v}} \end{bmatrix} = A^\dagger \mathbf{b}$.

Let $f(\mathbf{u}, \mathbf{v}, \lambda) = \|\mathbf{u}\|_2^2 + \|\mathbf{v}\|_2^2 - 2\lambda^\top (R\mathbf{u} + S\mathbf{v} - \mathbf{c})$, and set

$$\begin{aligned} \nabla_{\mathbf{u}} f &= 2\mathbf{u} - 2R^\top \lambda = 0 \\ \nabla_{\mathbf{v}} f &= 2\mathbf{v} - 2S^\top \lambda = 0, \end{aligned}$$

Solution (cont.)

we have $\lambda = R^{-\top}\hat{\mathbf{u}}$ and $\hat{\mathbf{v}} = S^\top R^{-\top}\hat{\mathbf{u}} = (R^{-1}S)^\top\hat{\mathbf{u}}$. So

$$\begin{aligned}
\|\mathbf{x}_B - \hat{\mathbf{x}}\|_2^2 &= \|\Pi^\top\mathbf{x}_B - \Pi^\top\hat{\mathbf{x}}\|_2^2 \\
&= \left\| \begin{bmatrix} R^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{u}} \\ (R^{-1}S)^\top\hat{\mathbf{u}} \end{bmatrix} \right\|_2^2 \\
&= \left\| \begin{bmatrix} R^{-1}\mathbf{c} - \hat{\mathbf{u}} \\ -(R^{-1}S)^\top\hat{\mathbf{u}} \end{bmatrix} \right\|_2^2 \\
&= \left\| \begin{bmatrix} R^{-1}(R\hat{\mathbf{u}} + S\hat{\mathbf{v}}) - \hat{\mathbf{u}} \\ -(R^{-1}S)^\top\hat{\mathbf{u}} \end{bmatrix} \right\|_2^2 \\
&= \left\| \begin{bmatrix} R^{-1}S\hat{\mathbf{v}} \\ -(R^{-1}S)^\top\hat{\mathbf{u}} \end{bmatrix} \right\|_2^2 \\
&= \|R^{-1}S\hat{\mathbf{v}}\|_2^2 + \|(R^{-1}S)^\top\hat{\mathbf{u}}\|_2^2 \\
&\le \|R^{-1}S\|_2^2\|\hat{\mathbf{v}}\|_2^2 + \|(R^{-1}S)^\top\|_2^2\|\hat{\mathbf{u}}\|_2^2 \\
&= \|R^{-1}S\|_2^2(\|\hat{\mathbf{v}}\|_2^2 + \|\hat{\mathbf{u}}\|_2^2)
\end{aligned}$$

where the last equality comes from the fact that $\|R^{-1}S\|_2 = \|(R^{-1}S)^\top\|_2$ since the maximal singular values of these two matrices are the same.
Therefore,

$$\begin{aligned}
\frac{\|\mathbf{x}_B - \hat{\mathbf{x}}\|_2}{\|\hat{\mathbf{x}}\|_2} &\le \frac{\|R^{-1}S\|_2\sqrt{\|\hat{\mathbf{v}}\|_2^2 + \|\hat{\mathbf{u}}\|_2}}{\sqrt{\|\hat{\mathbf{v}}\|_2^2 + \|\hat{\mathbf{u}}\|_2}} \\
&= \|R^{-1}S\|_2
\end{aligned}$$

$\square$

5

Given a symmetric $A \in \mathbb{R}^{n\times n}$, $\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n$, and $\mathbf{b} \in \mathbb{R}^n$. Let

$$\mathbf{r} = \mathbf{b} - A\mathbf{x}$$

Consider the QR decomposition

$$[\mathbf{x}, \mathbf{r}] = QR$$

and observe that if $E\mathbf{x} = \mathbf{r}$, then

$$(Q^\top EQ)(Q^\top \mathbf{x}) = Q^\top \mathbf{r}.$$

Show how to compute a symmetric $E \in \mathbb{R}^{n\times n}$ so that it attains

$$\min_{(A+E)\mathbf{x}=\mathbf{b}} \|E\|_F,$$

where the minimum is taken over all symmetric E (Note: The point here is that one must usually take into account that errors occurring in symmetric matrices must also be symmetric).

Proof. Since Q is unitary, $(A+E)\mathbf{x} = \mathbf{b}$ equals to $(Q^\top EQ)(Q^\top \mathbf{x}) = Q^\top \mathbf{r}$ and $Q^\top EQ$ is also symmetric. Also,

$$Q^\top \begin{bmatrix} \mathbf{x} & \mathbf{r} \end{bmatrix} = \begin{bmatrix} Q^\top \mathbf{x} & Q^\top \mathbf{r} \end{bmatrix} = R = \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}$$

where $r_{11} \neq 0$ since $\mathbf{x} \neq \mathbf{0}$. Let $F = Q^\top EQ$, then the optimization problem is equal to $\min \|F\|_F$, s.t. $(A+E)\mathbf{x} = \mathbf{b}$ and $F^\top F$. Since

$$\begin{bmatrix} f_{11} & \cdots & f_{1n} \\ \vdots & \ddots & \vdots \\ f_{1n} & \cdots & f_{nn} \end{bmatrix} \begin{bmatrix} r_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} r_{12} \\ r_{22} \\ \vdots \\ 0 \end{bmatrix}$$

Then

$$\begin{cases} f_{11} r_{11} = r_{12} \\ f_{12} r_{11} = r_{22} \end{cases} \implies \begin{cases} f_{11} &= \frac{r_{12}}{r_{11}} \\ f_{12} &= \frac{r_{22}}{r_{11}} \end{cases}$$

Let $F_0 = \begin{bmatrix} \frac{r_{12}}{r_{11}} & \frac{r_{22}}{r_{11}} & \cdots & 0 \\ \frac{r_{22}}{r_{11}} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$. So $\|E\|_F = \|F\|_F \geq \|F_0\|_F = \sqrt{\frac{r_{12}^2 + 2r_{22}^2}{r_{11}^2}}$. We conclude that the solution to $\min_E \|E\|_F$ s.t. $(A+E)\mathbf{x} = \mathbf{b}$ and $E^\top = E$, is $\sqrt{\frac{r_{12}^2 + 2r_{22}^2}{r_{11}^2}}$ and the minimum is achieved at $E_0 = QF_0Q^\top$. $\square$

6

In this exercise, we will implement and compare Gram–Schmidt and Householder QR. Your implementation should be tailored to the program you are using for efficiency (e.g. vectorize your code in Matlab/Octave/Scilab). Assume in the following that the input is a matrix $A \in \mathbb{R}^{m\times n}$ with $\operatorname{rank}(A) = n \le m$ and we want to find its full QR decomposition $A = QR$ where $Q \in \mathrm{O}(m)$ and $R \in \mathbb{R}^{m\times n}$ is upper-triangular.

(a) Implement the (classical) Gram–Schmidt algorithm to obtain Q and R.

For solving $A = QR$ by the Gram–Schmidt algorithm, suppose we have get the first $k-1$ columns of Q and R, denoted by $Q_{k-1} = \begin{bmatrix}\mathbf{q}_1 & \cdots & \mathbf{q}_{k-1}\end{bmatrix}$ and $R_{k-1} = \begin{bmatrix}\mathbf{r}_1 & \cdots & \mathbf{r}_{k-1}\end{bmatrix}$ where $\mathbf{r}_{(i+1):m}$, the subvector of $\mathbf{r}_i$, are $\mathbf{0}$. For $\mathbf{q}_k$ and $\mathbf{r}_k$, notice that

$$\begin{aligned} Q_{k-1}^\top \mathbf{q}_k &= \mathbf{0} \\ \mathbf{a}_k &= Q_k \mathbf{r}_{k,1:k} \\ &= Q_{k-1}\mathbf{r}_{k,1:(k-1)} + \mathbf{q}_k r_{kk} \end{aligned}$$

we have

$$Q_{k-1}^\top \mathbf{a}_k = Q_{k-1}^\top Q_{k-1}\mathbf{r}_{k,1:(k-1)} + Q_{k-1}^\top \mathbf{q}_k \mathbf{r}_k = \mathbf{r}_{k,1:(k-1)}$$

which can use to solve $\mathbf{r}_{k,1:(k-1)}$. Then $\mathbf{q}_k r_{kk} = \mathbf{a}_k - Q_{k-1}\mathbf{r}_{k,1:(k-1)}$. Since $\mathbf{q}_k$ is a unit vector, we have

$$\begin{aligned} r_{kk} &= \|\mathbf{a}_k - Q_{k-1}\mathbf{r}_{k,1:(k-1)}\|_2 \\ \mathbf{q}_k &= \frac{\mathbf{a}_k - Q_{k-1}\mathbf{r}_{k,1:(k-1)}}{r_{kk}}. \end{aligned}$$

```
function [Q, R] = GramSchmidt_QR(A)
    % Input: A, m by n matrix with full column rank
    [m, n] = size(A);
    Q = zeros(m,m);
    R = zeros(m,n);

    % Gram-Schmidt
    R(1,1) = norm(A(:,1),'fro');
    Q(:,1) = A(:,1)/R(1,1);

    for i=2:n
        R(1:i-1,i) = Q(:,1:(i-1))'*A(:,i);
        Q(:,i) = A(:,i) - Q(:,1:(i-1))*R(1:(i-1),i);
        R(i,i) = norm(Q(:,i),'fro');
        Q(:,i) = Q(:,i)/R(i,i);
    end
```

Solution (cont.)

```
    % fill bank columns of Q if m>n
    for i=(n+1):m
        for j=i-n:m
            ej = zeros(m,1);
            ej(j) = 1;
            Q(:,i) = ej - Q(:,1:i)*(ej'*Q(:,1:i))';
            if sum(abs(Q(:,i)))==0
                continue
            end
        end
    end
end
```

(b) Implement the Householder QR algorithm to obtain Q and R. You should (i) store Q implicitly, taking advantage of the fact that it can be uniquely specified by a sequence of vectors of decreasing dimensions; (ii) choose α in your Householder matrices to have the opposite sign of x_1 to avoid cancellation in v_1 (cf. notations in lecture notes).

Suppose there is $H_1 = I - 2\mathbf{u}_1\mathbf{u}_1^\top$ such that $H_1\mathbf{a}_1 = \alpha\mathbf{e}_1$. Since H_1 is a reflection matrix, $H_1^\top = H_1$ and $H_1^\top H_1 = I$. From the relations $\|H_1\mathbf{a}_1\|_2 = \|\mathbf{a}_1\|_2$ and $\|\alpha\mathbf{e}_1\|_2 = |\alpha|\|\mathbf{e}_1\| = |\alpha|$, we obtain $\alpha = \pm\|\mathbf{a}_1\|_2$. Since $\mathbf{x} = P^{-1}(\alpha\mathbf{e}_1) = \alpha P\mathbf{e}_1 = \alpha(\mathbf{e}_1 - 2\mathbf{u}_1\mathbf{u}_1^\top\mathbf{e}_1) = \alpha(\mathbf{e}_1 - 2\mathbf{u}_1 u_{11})$, we obtain $u_{11} = \pm\sqrt{\frac{1}{2}\left(1 - \frac{a_{11}}{\alpha}\right)} = \pm\sqrt{\frac{1}{2\alpha}(\alpha - a_{11})}$ and $u_{i1} = -\frac{a_{i1}}{2\alpha u_{11}}$ $(i = 2, \ldots, m)$. If we choose $u_{11} > 0$, $\alpha = -\operatorname{sign}(a_{11})\|\mathbf{a}_1\|_2$, and $\tau = \frac{-\operatorname{sign}(a_{11})(\alpha - a_{11})}{\|\mathbf{a}\|_2} > 0$, then $H_1 = I - \tau\mathbf{v}\mathbf{v}^\top$ where $v_{11} = 1$, $v_{i1} = \frac{a_{i1}}{\alpha - a_{11}}$ $(i = 2, \ldots, m)$ and $\sqrt{\tau}\mathbf{v} = \sqrt{2}\mathbf{u}$.

```
function [A, tau, varargout] = Householder_QR(A)
    [m, n] = size(A);
    tau = zeros(n,1);
    if nargout==4
        Q = eye(m);
    end

    % Householder
    for j=1:n
        normx = norm(A(j:end,j));
        s = -sign(A(j,j));
        u1 = A(j,j) - s*normx;
        w = A(j:end,j)/u1;
        w(1) = 1;
```

Solution (cont.)

```
        A(j+1:end,j) = w(2:end);
        A(j,j) = s*normx;
        tau(j) = -s*u1/normx;

        A(j:end,j+1:end) = A(j:end,j+1:end)-(tau(j)*w)*(w'*A(j:end,j+1:end));

        if nargout==4
            Q(:,j:end) = Q(:,j:end)-(Q(:,j:end)*w) * (tau(j)*w)';
        end
    end

    if nargout==4
        R = triu(A);
        varargout1 = Q;
        varargout2 = R;
    end
end
```

(c) Implement an algorithm for forming the product $Q\mathbf{x}$ and another for forming the product $Q^\top\mathbf{y}$ when Q is stored implicitly as in (b).

```
function z = Qx(QR, tau, x)
    [~, n] = size(QR);
    QR(logical(eye(n))) = 1;

    z = x;
    for j=n:-1:1
        z(j:end,:) = z(j:end,:) - tau(j) * QR(j:end,j)*(QR(j:end,j)' * z(j:end,:));
    end
end

function z = QTy(QR, tau, y)
    [~, n] = size(QR);
    QR(logical(eye(n))) = 1;

    z = y;
    for j=1:n
        z(j:end,:) = z(j:end,:) - tau(j) * QR(j:end,j)*(QR(j:end,j)' * z(j:end,:));
    end
end
```

(d) For increasing values of n, generate an upper triangular $R \in \mathbb{R}^{n\times n}$ and a $B \in \mathbb{R}^{n\times n}$, both with random standard normal entries. Use your program's built-in function for QR factorization to obtain a random[1] $Q \in \mathrm{O}(n)$ from the QR factorization of B. Now form $A = QR$ and apply your algorithms in (a) and (b) to find the QR factors of A — let these be $\widehat{Q}$ and $\widehat{R}$. Tabulate (using graphs with appropriate scales) the relative errors

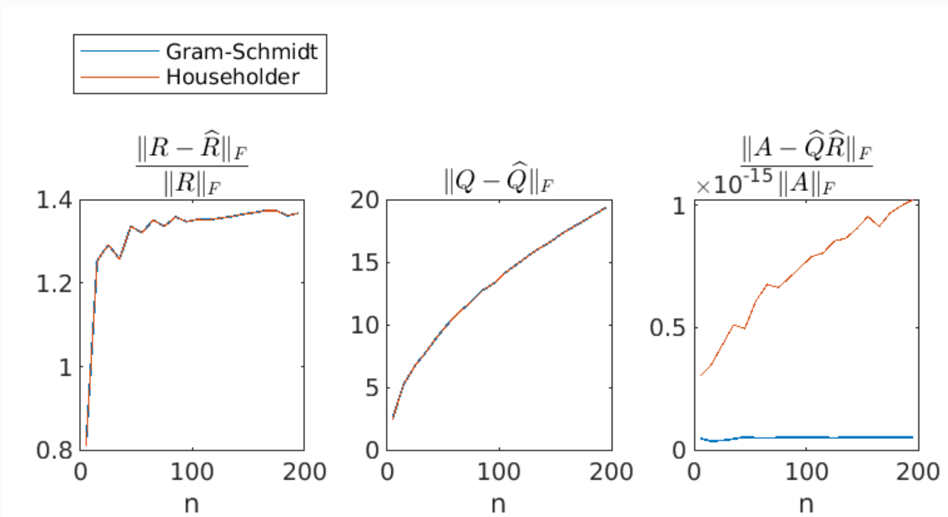
$$\frac{\|R-\widehat{R}\|_F}{\|R\|_F}, \quad \|Q-\widehat{Q}\|_F, \quad \frac{\|A-\widehat{Q}\widehat{R}\|_F}{\|A\|_F},$$

for various values of n and for each method. Scale $Q, R, \widehat{Q}, \widehat{R}$ appropriately so that R and $\widehat{R}$ have positive diagonal elements.

(i) Comment on the relative errors in $\widehat{Q}$ and $\widehat{R}$ (these are called forward errors) versus the relative error in $\widehat{Q}\widehat{R}$ (this is called backward error).

(ii) Comment on the relative error in $\widehat{Q}\widehat{R}$ computed with Gram–Schmidt versus that computed with Householder QR.

(i) As we can see from the following plots, the backward errors are much smaller than the forward errors. As n increases, both forward errors and backward errors tend to increase.
(ii) The forward errors of these two methods are similar, while the backward error of Householder algorithm is much larger than that of Gram-Schmidt algorithm.




```
n_list = 5:10:200;
result = zeros(2,length(n_list),3);
for i = 1:length(n_list)
    n = n_list(i);
    pd = makedist('Normal');
    R = random(pd,[n,n]);
    B = random(pd,[n,n]);
    [Q, ~] = qr(B);
    Q(:,diag(R)<0) = -Q(:,diag(R)<0);
    R(diag(R)<0,:) = -R(diag(R)<0,:);
    A = Q * R;
```

[1]This is usually how one would generate a random orthogonal matrix.

Solution (cont.)

```
    [Q_G, R_G] = GramSchmidt_QR(A);
    Q_G(:,diag(R_G)<0) = -Q_G(:,diag(R_G)<0);
    R_G(diag(R_G)<0,:) = -R_G(diag(R_G)<0,:);

    [~, ~, Q_H, R_H] = Householder_QR(A);
    Q_H(:,diag(R_H)<0) = -Q_H(:,diag(R_H)<0);
    R_H(diag(R_H)<0,:) = -R_H(diag(R_H)<0,:);

    result(1,i,1) = norm(R-R_G,"fro")/norm(R,"fro");
    result(1,i,2) = norm(Q-Q_G,"fro");
    result(1,i,3) = norm(A-Q_G*R_G,"fro")/norm(A,"fro");
    result(2,i,1) = norm(R-R_H,"fro")/norm(R,"fro");
    result(2,i,2) = norm(Q-Q_H,"fro");
    result(2,i,3) = norm(A-Q_H*R_H,"fro")/norm(A,"fro");
end

title_list = {'$$\frac{\| R-_widehat{R}\|_F}{\| R\|_F}$$',
    '$$\| Q-_widehat{Q}\|_F\$$',
    '$$\frac{\| A-\widehat{Q}\widehat{R}\|_F}{\| A\|_F}$$'};
figure();
for i=1:3
    subplot(2,3,i);
    h1 = plot(n_list, result(1,:,i), 'LineWidth', 0.8);
    hold on
    h2 = plot(n_list, result(2,:,i));
    title(title_listi,'interpreter','latex')  ;
    xlabel('n');
end
% Construct a Legend with the data from the sub-plots
hL = legend([h1, h2],'Gram-Schmidt','Householder');
% Programatically move the Legend
newPosition = [0.2 0.4 0.12 0.06];
newUnits = 'normalized';
set(hL,'Position', newPosition,'Units', newUnits);
saveas(gcf,'result1.png')
```

(e) Generate a *Vandermonde matrix* and a vector,

$$A = \begin{bmatrix} 1 & \alpha_0 & \alpha_0^2 & \dots & \alpha_0^{n-1} \\ 1 & \alpha_1 & \alpha_1^2 & \dots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \dots & \alpha_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_{m-1} & \alpha_{m-1}^2 & \dots & \alpha_{m-1}^{n-1} \end{bmatrix} \in \mathbb{R}^{m\times n}, \quad \mathbf{b} = \begin{bmatrix} \exp(\sin 4\alpha_0) \\ \exp(\sin 4\alpha_1) \\ \exp(\sin 4\alpha_2) \\ \vdots \\ \exp(\sin 4\alpha_{m-1}) \end{bmatrix} \in \mathbb{R}^m,$$

where $\alpha_i = i/(m-1)$, $i = 0, 1, \ldots, m-1$. This arises when we try to do polynomial fitting

$$e^{\sin 4x} \approx c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1} x^{n-1}$$

over the interval $[0, 1]$ at discrete points $x = 0, \frac{1}{m-1}, \frac{2}{m-1}, \ldots, \frac{m-2}{m-1}, 1$. For $n = 15$ and $m = 100$, solve the least squares problem $\min \|A\mathbf{x} - \mathbf{b}\|_2$ and state your value of c_{14} using each of the following methods:

(i) Applying QR factorization to A.

(ii) Applying QR factorization to the augmented matrix $[A, \mathbf{b}] \in \mathbb{R}^{m \times (n+1)}$.

(iii) Solving the normal equations $A^\top A\mathbf{x} = A^\top \mathbf{b}$.

For (i) and (ii), your code should show how the respective QR factors are used in obtaining a solution of the least squares problem. You are free to use your program's built-in functions (e.g. `A\b` in Matlab/Octave/Scilab) for solving linear systems but for other things, use what you have implemented in (a), (b), (c). The true value of c_{14} is $2006.787453080206\ldots$. Comment on the accuracy of each method and algorithm.

For Householder algorithm, (i) and (ii) give best solutions. While Gram-Schmidt algorithm fails in (i) and (ii), which causes very large errors. Also, solving the normal equation yields largest error.

```
m=100; n=15;
alpha = (0:(m-1))/(m-1);
A = fliplr(vander(alpha));
A = A(:,1:n);
b = exp(sin(4*alpha))';

x=A\b;
x(15)
```

```
ans = 2.0068e+03
```

```
% Applying QR factorization to A
[Q, R] = GramSchmidt_QR(A);
x = R\(Q'*b);
x(15)
```

```
ans = 1.1833
```

```
[~, ~, Q, R] = Householder_QR(A);
x = R\(Q'*b);
x(15)
```

```
ans = 2.0068e+03
```

```
% Applying QR factorization to [A, b]
[~, Ra] = GramSchmidt_QR([A, b]);
R = Ra(1:m,1:n);
QTb = Ra(1:m,n+1);
x = R\QTb;
x(15)
```

Solution (cont.)

```
ans = 1.1833
```

```
[~, ~, ~, Ra] = Householder_QR([A, b]);
R = Ra(1:m,1:n);
QTb = Ra(1:m,n+1);
x = R\QTb;
x(15)
```

```
ans = 2.0068e+03
```

```
% Solving the normal equations
x = A'*A\(A'*b);
x(15)
```

```
ans = -310.2727
```