Jaiden Atterbury SH: 2028502 Homework 5

Component Skills Exercises

Exercise 1. (CS 3.4):

Let $u, v \in \mathbb{R}^n$. The inner-product of u and v is given by the dot product $u \cdot v = u^T v$. The outer-product of u and v is defined as uv^T. Notice that, while the inner-product takes in two vectors and produces a scalar, the outer-product takes in two vectors and produces a matrix.

(a.) Compute the outer-product of $u = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $v = \begin{bmatrix} -2 \\ 4 \\ 1 \end{bmatrix}$.

The outer-product of u and v is defined as uv^T, hence in our case this is

$$uv^T = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} -2 & 4 & 1 \end{pmatrix} = \begin{pmatrix} -2\times 1 & 4\times 1 & 1\times 1 \\ -2\times 2 & 4\times 2 & 1\times 2 \\ -2\times 3 & 4\times 3 & 1\times 3 \end{pmatrix} = \begin{pmatrix} -2 & 4 & 1 \\ -4 & 8 & 2 \\ -6 & 12 & 3 \end{pmatrix}$$

(b.) What is the flop count to compute an outer-product of two vectors in $\mathbb{R}^n$? How does it compare to the flop count of an inner-product? Which product is more expensive to compute?

If we let $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^n$, then their outer product, uv^T, is given as

$$uv^T = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & \cdots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \cdots & u_2 v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_n v_1 & u_n v_2 & \cdots & u_n v_n \end{bmatrix}$$

As can be seen above, each row of uv^T requires n multiplications, thus, since there are n rows in uv^T, there are $n \times n = n^2$ total flops to compute an outer-product.

Furthermore, as computed in lecture 4, the inner-product of two vectors in $\mathbb{R}^n$ is $2n-1$ total flops to compute.

Since the inner-product takes $\sim 2n$ flops to compute, and the outer-product takes $\sim n^2$ flops to compute, for large n (and in general), the outer-product on two vectors in $\mathbb{R}^n$ is significantly more expensive to compute.

(c.) In the special case when $u=v=q$, where q is a unit vector (with respect to the 2-norm), the outer-product qq^T becomes an example of an orthogonal projector. In general, a $n \times n$ matrix P is called an orthogonal projector if

- $P^T = P$ (i.e, P is symmetric), and
- $P^2 = P$ (i.e, P is idempotent).

Show that $P = qq^T$ is an orthogonal projector.

Let $q \in \mathbb{R}^n$ be a unit vector with respect to the 2-norm, then it follows that $\|q\|_2 = 1$. Let $P = qq^T$.

◦ STEP 1: Show that $P^T = P$:

$$
\begin{aligned}
P^T &= (qq^T)^T \\
&= (q^T)^T q^T && (\text{since } (AB)^T = B^T A^T) \\
&= qq^T && (\text{since } (A^T)^T = A) \\
&= P
\end{aligned}
$$

$\therefore P^T = P$ (P is symmetric)

◦ STEP 2: Show that $P^2 = P$:

$$
\begin{aligned}
P^2 &= PP \\
&= (qq^T)(qq^T) \\
&= q(q^T q)q^T && (\text{By associativity of matrix multiplication}) \\
&= q(q \cdot q)q^T && (\text{since } u^T v = u \cdot v) \\
&= q(1)q^T && (\text{since } \|q\|_2 = \sqrt{q \cdot q} \Rightarrow \|q\|_2^2 = q \cdot q \text{ and } \|q\|_2 = 1 \Rightarrow \|q\|_2^2 = 1) \\
&= qq^T \\
&= P
\end{aligned}
$$

$\therefore P^2 = P$ (P is idempotent)

Since $P = qq^T$ is symmetric and idempotent, $P = qq^T$ is an orthogonal projector.

(d.) Consider the orthogonal projector $P = qq^T$, where $q = (-1/\sqrt{2}, 1/\sqrt{2})^T$. Using software of your choice, plot the vectors $x = (-1, 2)^T$ and Px alongside the vector q in $\mathbb{R}^2$. Based on your plot, what does multiplication by P do geometrically to any vector $x \in \mathbb{R}^2$? (Your answer should be brief.)

In this problem, I used Python to plot the vectors $q=(-1/\sqrt{2}, 1/\sqrt{2})^T$ in red, $x=(-1, 2)^T$ in blue, and $Px=(-3/2, 3/2)^T$ in magenta to illustrate the functionality of the orthogonal projector $P=qq^T$. This plot can be found in the Homework 5 plots appendix under "Exercise 1 part d."

Geometrically, multiplication by $P=qq^T$ takes any vector $\vec{x}$ in $\mathbb{R}^2$, and projects x onto the span $\{q\}$, i.e. the line spanned by q.

(e.) Show that $C=I_n-qq^T$ is an orthogonal projector.

Let $q\in\mathbb{R}^n$ be a unit vector with respect to the 2-norm, then it follows that $\|q\|_2=1$. Let $C=I_n-qq^T$.

◦ STEP 1: show that $C^T=C$:

$$
\begin{aligned}
C^T &= (I_n-qq^T)^T \\
&= I_n^T-(qq^T)^T && (\text{since } (A+B)^T=A^T+B^T) \\
&= I_n^T-(q^T)^Tq^T && (\text{since } (AB)^T=B^TA^T) \\
&= I_n-qq^T && (\text{since } I_n^T=I_n \text{ and } (A^T)^T=A) \\
&= C
\end{aligned}
$$

$\therefore C^T=C$ (C is symmetric)

◦ STEP 2: show that $C^2=C$:

$$
\begin{aligned}
C^2 &= CC \\
&= (I_n-qq^T)(I_n-qq^T) \\
&= (I_n-qq^T)I_n-(I_n-qq^T)qq^T && (\text{since } A(B+C)=AB+AC) \\
&= (I_n-qq^T)-I_nqq^T+qq^Tqq^T && (\text{since } (A+B)C=AC+BC \text{ and } AI_n=A) \\
&= I_n-qq^T-qq^T+q(q^Tq)q^T && (\text{since } I_nA=A \text{ and associativity of matrix multiplication}) \\
&= I_n-qq^T-qq^T+q(q\cdot q)q^T && (\text{since } u^Tv=u\cdot v) \\
&= I_n-qq^T-qq^T+qq^T && (\text{since } \|q\|_2=\sqrt{q\cdot q}\Rightarrow\|q\|_2^2=q\cdot q \text{ and } \|q\|_2=1\Rightarrow\|q\|_2^2=1) \\
&= I_n-qq^T \\
&= C
\end{aligned}
$$

$\therefore C^2=C$ (C is idempotent)

Since $C=I_n-qq^T$ is symmetric and idempotent, $C=I_n-qq^T$ is an orthogonal projector.

(f.) Consider the orthogonal projector $C = I_n - qq^T$, where $q = (-1/\sqrt{2}, 1/\sqrt{2})^T$. Using software of your choice, plot the vectors $x = (-1, 2)^T$ and Cx alongside the vector q in $\mathbb{R}^2$. Based on your plot, what does multiplication by C do geometrically to any vector $x \in \mathbb{R}^2$? (Your answer should be brief.)

In this problem, I used python to plot the vectors $q = (-1/\sqrt{2}, 1/\sqrt{2})^T$ in red, $x = (-1, 2)^T$ in blue, and $Cx = (1/2, 1/2)^T$ in magenta to illustrate the functionality of the orthogonal projector $C = I_n - qq^T$. This plot can be found in the Homework 5 plots appendix under "Exercise 1 part f."

Geometrically, multiplication by $C = I_n - qq^T$ takes any vector $x \in \mathbb{R}^2$ and projects $\vec{x}$ onto the subspace orthogonal to span$\{q\}$, i.e. the line orthogonal to the span$\{q\}$, thus $q \cdot C\vec{x} = 0$.

- Exercise 2. (CS 3.4):

Suppose we have two mutually orthonormal vectors $q_1, q_2 \in \mathbb{R}^n$. Consider the nxn matrix

$$P = q_1 q_1^T + q_2 q_2^T$$

(a.) Show that P is an orthogonal projector.

Let $q_1, q_2 \in \mathbb{R}^n$ be mutually orthonormal, this implies that $\|q_1\|_2 = 1$, $\|q_2\|_2 = 1$ and $q_1 \cdot q_2 = q_2 \cdot q_1 = 0$. Let $P = q_1 q_1^T + q_2 q_2^T$.

∘ STEP 1: Show that $P^T = P$:

$$\begin{aligned}
P^T &= (q_1 q_1^T + q_2 q_2^T)^T \\
&= (q_1 q_1^T)^T + (q_2 q_2^T)^T && (\text{since } (A+B)^T = A^T + B^T) \\
&= (q_1^T)^T q_1^T + (q_2^T)^T q_2^T && (\text{since } (AB)^T = B^T A^T) \\
&= q_1 q_1^T + q_2 q_2^T && (\text{since } (A^T)^T = A) \\
&= P
\end{aligned}$$

$\therefore$ $P^T = P$ (P is symmetric)

∘ STEP 2: Show that $P^2 = P$:

$$\begin{aligned}
P^2 &= PP \\
&= (q_1 q_1^T + q_2 q_2^T)(q_1 q_1^T + q_2 q_2^T) && (\text{since } A(B+C) = AB + AC) \\
&= (q_1 q_1^T + q_2 q_2^T) q_1 q_1^T + (q_1 q_1^T + q_2 q_2^T) q_2 q_2^T && (\text{since } (A+B)C = AC + BC) \\
&= q_1 q_1^T q_1 q_1^T + q_2 q_2^T q_1 q_1^T + q_1 q_1^T q_2 q_2^T + q_2 q_2^T q_2 q_2^T && (\text{since } x^T y = x \cdot y) \\
&= q_1 (q_1 \cdot q_1) q_1^T + q_2 (q_2 \cdot q_1) q_1^T + q_1 (q_1 \cdot q_2) q_2^T + q_2 (q_2 \cdot q_2) q_2^T && (\text{since } \|q\|_2 = \sqrt{q \cdot q} \Rightarrow \|q\|_2^2 = q \cdot q) \\
&= q_1 (\|q_1\|_2^2) q_1^T + q_2 (q_2 \cdot q_1) q_1^T + q_1 (q_1 \cdot q_2) q_2^T + q_2 (\|q_2\|_2^2) q_2^T && (\text{since } \|q\|_2^2 = 1,\ q_1 \cdot q_2 = 0,\ q_2 \cdot q_1 = 0) \\
&= q_1 q_1^T + q_2 (0) q_1^T + q_1 (0) q_2^T + q_2 q_2^T \\
&= q_1 q_1^T + q_2 q_2^T \\
&= P
\end{aligned}$$

$\therefore$ $P^2 = P$ (P is idempotent.

Since $P = q_1 q_1^T + q_2 q_2^T$ is symmetric and idempotent, $P = q_1 q_1^T + q_2 q_2^T$ is an orthogonal projection.

$\boxed{\text{NOTE:}}$ In the proof of $P^2 = P$, the reasonings in parentheses to the right should all be moved one line down.

(b.) Suppose $q_1 = (1\ 0\ 0)^T$ and $q_2 = (0\ \frac{1}{\sqrt{2}}\ \frac{1}{\sqrt{2}})^T$.

(i) Using software of your choice, sketch span$\{q_1, q_2\}$. Superimpose onto your sketch the vectors $x = (-1\ 1\ 2)^T$ and Px.

In this problem I used Python to plot the vectors $q_1 = (1\ 0\ 0)^T$ and $q_2 = (0\ \frac{1}{\sqrt{2}}\ \frac{1}{\sqrt{2}})^T$ in red (I used the same color for both vectors to represent that it is the span$\{q_1, q_2\}$, $x = (-1\ 1\ 2)^T$ in blue and $Px = (-1, 3/2, 3/2)^T$ in magenta to illustrate the functionality of the orthogonal projector $P = q_1 q_1^T + q_2 q_2^T$. This plot can be found in the Homework 5 plots appendix under "Exercise 2 Part b(i)."

(ii) What geometrically does multiplying x by P do to x? (Your answer should be brief.)

Geometrically, multiplication by $P = q_1 q_1^T + q_2 q_2^T$ takes any vector $x \in \mathbb{R}^n$ projects x into the span$\{q_1, q_2\}$. i.e. Px is on the (hyper)plane spanned by q_1 and q_2. In our case, x is projected onto the plane spanned by q_1 and q_2 in $\mathbb{R}^3$.

(c.) Suppose $\{q_1, q_2, \cdots, q_K\}$ forms an orthonormal basis for some K-dimensional subspace of $\mathbb{R}^n$. Propose a matrix P that orthogonally projects any vector $x \in \mathbb{R}^n$ onto this subspace.

Based on parts c and d from Exercise 1 and parts a and b from Exercise 2, I propose the matrix

$$P = q_1 q_1^T + q_2 q_2^T + \cdots + q_K q_K^T$$

orthogonally projects any vector $x \in \mathbb{R}^n$ onto the subspace spanned by the orthonormal basis $\{q_1, q_2, \cdots, q_K\}$.

(d.) Suppose $\{q_1, q_2, \cdots, q_K\}$ forms an orthonormal basis for some K-dimensional subspace of $\mathbb{R}^n$. Propose a matrix C that orthogonally projects any vector $x \in \mathbb{R}^n$ onto the subspace perpendicular to the subspace spanned by $\{q_1, q_2, \cdots, q_K\}$.

Hint: see Exercise 1(e)-(f) above.

Based on Exercise 1(e)-(f) above, I propose the matrix

$$C = I_n - q_1 q_1^T - q_2 q_2^T - \cdots - q_K q_K^T$$

orthogonally projects any vector $x \in \mathbb{R}^n$ onto the subspace perpendicular to the subspace spanned by the K-dimensional orthonormal basis $\{q_1, q_2, \cdots, q_K\}$.

- Exercise 3. (CS 3.4):

Let $q \in \mathbb{R}^n$ be a unit vector (with respect to the 2-norm). Define the $n \times n$ matrix

$$H_q = 2qq^T - I_n.$$

(a.) Is H_q an orthogonal projector? If not, which property of an orthogonal projector does H_q not satisfy?

Let $q \in \mathbb{R}^n$ be a unit vector with respect to the two-norm, then it follows that $\|q\|_2 = 1$. Let $H_q = 2qq^T - I_n$.

∘ STEP 1: show that $H_q^T = H_q$:

$$\begin{aligned} H_q^T &= (2qq^T - I_n)^T \\ &= (2qq^T)^T - (I_n)^T && (\text{since } (A+B)^T = A^T + B^T) \\ &= 2(qq^T)^T - I_n && (\text{since } (cA)^T = cA^T \text{ and } I_n^T = I_n) \\ &= 2(q^T)^T q^T - I_n && (\text{since } (AB)^T = B^T A^T) \\ &= 2qq^T - I_n && (\text{since } (A^T)^T = A) \\ &= H_q \end{aligned}$$

$\therefore H_q^T = H_q$ (H_q is symmetric)

∘ STEP 2: show that $H_q^2 = H_q$:

$$\begin{aligned} H_q^2 &= H_q H_q \\ &= (2qq^T - I_n)(2qq^T - I_n) \\ &= (2qq^T - I_n)2qq^T - (2qq^T - I_n)I_n && (\text{since } A(B+C) = AB + AC) \\ &= 2qq^T 2qq^T - I_n 2qq^T - 2qq^T + I_n && (\text{since } (A+B)C = AC + BC \text{ and } AI_n = A) \\ &= 4qq^T - 2qq^T - 2qq^T + I_n && (\text{since } c(AB) = (cA)B = A(cB) \text{ and } I_n A = A) \\ &= I_n \end{aligned}$$

$\therefore H_q^2 \neq H_q$ (H_q is not idempotent)

Since $H_q^2 \neq H_q$, H_q is not idempotent and is thus not an orthogonal projector.

(b.) As a direct consequence of a calculation that you did in (a), you should find the inverse of H_q. What is it?

Hint: Do not try to use Gauss-Jordan elimination to construct the inverse. See if you can find a matrix H_q^{-1} from your calculations in (a) such that $H_q^{-1} H_q = H_q H_q^{-1} = I_n$.

As found in my calculations for part (a), $H_q^2 = I_n$, this implies that $H_q H_q = I_n$, and hence $H_q^{-1} = H_q$. Thus the inverse of H_q is itself, i.e. $H_q^{-1} = H_q = 2qq^T - I_n$.

(c) ~~Is Hq an orthogonal matrix~~ (i,e., is $H_q^T H_q = I_n$

As seen in part (a), $H_q^T = H_q$, and $H_q^2 = I_n$, thus we can see that

$$H_q^T H_q = H_q H_q = H_q^2 = I_n.$$

Hence we can see that H_q is an orthogonal matrix.

(d.) suppose $q = (\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})^T$.

(i) Using software of your choice, sketch the line spanned by q. Superimpose on your sketch the vectors $x = (3 \;\; 1)^T$ and $H_q x$.

In this problem, I used Python to plot the vectors $q = (\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}})^T$ in red, $x = (3, 1)^T$ in blue, and $H_q x = (-1, 3)^T$ in magenta to illustrate the functionality of the householder reflector $H_q = 2qq^T - I_n$. This plot can be found in the Homework 5 plots appendix under "Exercise 3 part d."

(ii) What geometriony does multiplying x by H_q do to x? (Your answer should be brief.)

Geometrically, multiplication by $H_q = 2qq^T - I_n$ takes any vector $x \in \mathbb{R}^n$ reflects x about the span$\{q\}$. In our case H_q reflects x about the span$\{q\}$ in 2-dimensional space.

• Exercise 4. (CS 3.7):

For the linear system, $Ax=b$, where

$$A=\begin{pmatrix}1 & 2\\ 2 & -1\\ 1 & 2\end{pmatrix}, \text{ and } b=\begin{pmatrix}1\\3\\7\end{pmatrix},$$

please compute the solution of this least squares problem using both the normal equation directly and the QR equation. verify that your answers are the same regardless of the approach. use the QR decomposition:

$$A=\tilde{Q}\tilde{R}=\begin{pmatrix}1/\sqrt{6} & 1/\sqrt{3}\\ 2/\sqrt{6} & -1/\sqrt{3}\\ 1/\sqrt{6} & 1/\sqrt{3}\end{pmatrix}\begin{pmatrix}\sqrt{6} & 2/\sqrt{6}\\ 0 & 5/\sqrt{3}\end{pmatrix}$$

∘ STEP 1: compute the least squares solution using the normal equation: $A^TAx=A^Tb$

we see that

$$A=\begin{pmatrix}1 & 2\\ 2 & -1\\ 1 & 2\end{pmatrix}, \quad A^T=\begin{pmatrix}1 & 2 & 1\\ 2 & -1 & 2\end{pmatrix}, \quad b=\begin{pmatrix}1\\3\\7\end{pmatrix}$$

hence using the normal equation, $A^TAx=A^Tb$, we obtain

$$\begin{pmatrix}1 & 2 & 1\\ 2 & -1 & 2\end{pmatrix}\begin{pmatrix}1 & 2\\ 2 & -1\\ 1 & 2\end{pmatrix}\vec{x}=\begin{pmatrix}1 & 2 & 1\\ 2 & -1 & 2\end{pmatrix}\begin{pmatrix}1\\3\\7\end{pmatrix}$$

$$\begin{pmatrix}6 & 2\\ 2 & 9\end{pmatrix}\vec{x}=\begin{pmatrix}14\\13\end{pmatrix}$$

forming the augmented matrix and using gaussian elimination we see that

$$\left[\begin{array}{cc|c}6 & 2 & 14\\ 2 & 9 & 13\end{array}\right]\xrightarrow{-\frac{1}{3}R_1+R_2}\left[\begin{array}{cc|c}6 & 2 & 14\\ 0 & \frac{25}{3} & \frac{25}{3}\end{array}\right]$$

using back substitution we obtain

$$\frac{25}{3}x_2=\frac{25}{3} \Rightarrow x_2=1$$

$$6x_1+2x_2=14 \Rightarrow 6x_1+2=14 \Rightarrow 6x_1=12 \Rightarrow x_1=2$$

thus the least squares solution using the normal equation is

$$\vec{x}=(2,\ 1)^T.$$

○ STEP 2: compute the least squares solution using the QR equation: $\tilde{R}x = \tilde{Q}^T b$

we see that

$\tilde{Q} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{3} \\ 2/\sqrt{6} & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$, $\tilde{Q}^T = \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$, $\tilde{R} = \begin{bmatrix} \sqrt{6} & 2/\sqrt{6} \\ 0 & 5/\sqrt{3} \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 3 \\ 7 \end{bmatrix}$

hence using the QR equation, $\tilde{R}x = \tilde{Q}^T b$, we obtain

$$\begin{bmatrix} \sqrt{6} & 2/\sqrt{6} \\ 0 & 5/\sqrt{3} \end{bmatrix} \vec{x} = \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 7 \end{bmatrix}$$

$$\begin{bmatrix} \sqrt{6} & 2/\sqrt{6} \\ 0 & 5/\sqrt{3} \end{bmatrix} \vec{x} = \begin{bmatrix} 14/\sqrt{6} \\ 5/\sqrt{3} \end{bmatrix}$$

Using back substitution we obtain

$$5/\sqrt{3}\, x_2 = 5\sqrt{3} \Rightarrow x_2 = 1$$

$$\sqrt{6}x_1 + 2/\sqrt{6}\, x_2 = 14/\sqrt{6} \Rightarrow \sqrt{6}x_1 = 12/\sqrt{6} \Rightarrow x_1 = 2$$

thus the least squares solution using the QR equation is

$\vec{x} = (2, 1)^T$

Thus we have verified that the least squares solution is the same regardless of the approach.

- Exercise 5. (cs3.7):

Note: this exercise involves some programming. For this exercise you do not need to attach your MATLAB code. You only need to provide the results of your code

In MATLAB create the following 20x10 matrix

$$A = \begin{pmatrix} 1_{10\times 10} \\ 10^{-8} I_{10} \end{pmatrix}$$

Where $1_{10\times10}$ is a 10x10 matrix whose entries are all ones and I_{10} is the 10x10 identity matrix. This particular matrix has a very large condition number since the columns of A are nearly linearly dependent.

Consider the least-squares problem

$$Ax = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

(a.) Solve this least-squares problem in MATLAB using the normal equation. That is, form the matrix A^TA and solve $A^TAx = A^Tb$. You may use MATLAB's backslash command to solve this system, or you may compute the inverse of A^TA to solve for x directly. Are you able to get a solution? What (if any) warnings does MATLAB give you.

Using the backslash command to solve for x, I get no solution, as x is returned as a 10x1 vector of NaN values. The warning that MATLAB gives me is "Warning: Matrix is singular to working precision."

(b.) Solve this least-squares problem in MATLAB by the QR equation. That is, obtain the reduced QR factorization of A and solve $\tilde{R}x = \tilde{Q}^Tb$ by MATLAB's backslash command or by computing the inverse of $\tilde{R}$.

Note: Please use MATLAB's built-in QR factorization function as follows: [Q,R] = qr(A). Remember that MATLAB returns a full QR factorization. To get the reduced QR factorization, take only the first ten columns of Q and the first ten rows of R since A is a 20 x 10 matrix.

Rounded to the second decimal place, what is the solution x to this least-squares problem? Does this solution seem reasonable, i.e., does $Ax \approx b$ for the least-squares solution x?

Using the backslash command to solve for x in the QR equation I obtained

$$\vec{x} = (0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)^T$$

for the least-squares solution to $Ax = 1_{10\times1}$.

Computing $A\vec{x}$ with the above $\vec{x}$ I obtained

$$A\vec{x} = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 10^{-9}, 10^{-9}, 10^{-9}, 10^{-9}, 10^{-9}, 10^{-9}, 10^{-9}, 10^{-9}, 10^{-9}, 10^{-9})^T$$

this seems reasonable as the first 10 entries of $A\vec{x}$ are 1, furthermore, since the last 10 rows of A only have entries that are 0 or close to 0 (i.e., 10^{-8}), $A\vec{x}$ will have very small last 10 entries for reasonable numbers in $\vec{x}$. In order to get the last 10 entries of $A\vec{x}$ to be close to 1 would take the first 10 entries of $A\vec{x}$ much farther away from 1 and thus increases $\|A\vec{x} - \vec{b}\|_2^2$.

Exercise 6. (CS 3.7):

Let $A = (1\ 1\ 1)$ and $b = (4)$. Because A is short and fat, the columns of A are not full rank, so we cannot apply/uniquely solve the least-squares problem for $Ax = b$. Instead, we search for x such that

$$\|Ax - b\|_2^2 + \|x\|_2^2$$

is as small as possible. [That is, we determine x that solves the least-squares problem and has the smallest 2-norm.] Find this x.

Note/Hint: see the Lecture 14 course notes. After you have constructed a least-squares problem that is equivalent to the minimization problem above, you may use either the normal equation or QR equation to solve it.

In the lecture notes, if $k > n$, minimizing $\|Ax - b\|_2^2 + \|x\|_2^2$ leads to the least-squares problem:

$$\underset{n+k \times k}{\begin{pmatrix} A_{n\times k} \\ I_{k\times k} \end{pmatrix}} \vec{x} = \underset{n+k \times 1}{\begin{pmatrix} \vec{b}_{n\times 1} \\ \vec{0}_{k\times 1} \end{pmatrix}}$$

In our case:

$$\begin{pmatrix} A_{1\times 3} \\ I_{3\times 3} \end{pmatrix} \vec{x} = \begin{pmatrix} \vec{b}_{1\times 1} \\ \vec{0}_{3\times 1} \end{pmatrix}$$

which equals

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \vec{x} = \begin{pmatrix} 4 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

We will compute the least squares solution using the normal equation $A^TAx = A^Tb$. We see that

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 4 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

hence using the normal equation $A^TAx = A^Tb$ we obtain

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \vec{x} = \begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix}$$

Forming the augmented matrix and using Gaussian elimination we see that

$$\left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 1 & 2 & 1 & 4 \\ 1 & 1 & 2 & 4 \end{array}\right] \underset{\sim}{-\tfrac{1}{2}R_1+R_2} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 0 & 3/2 & 1/2 & 2 \\ 1 & 1 & 2 & 4 \end{array}\right] \underset{\sim}{-\tfrac{1}{2}R_1+R_3} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 0 & 3/2 & 1/2 & 2 \\ 0 & 1/2 & 3/2 & 2 \end{array}\right] \underset{\sim}{-\tfrac{1}{3}R_2+R_3} \left[\begin{array}{ccc|c} 2 & 1 & 1 & 4 \\ 0 & 3/2 & 1/2 & 2 \\ 0 & 0 & 4/3 & 4/3 \end{array}\right]$$

Using back substitution we obtain

$$\tfrac{4}{3}x_3 = \tfrac{4}{3} \Rightarrow x_3 = 1$$

$$\tfrac{3}{2}x_2 + \tfrac{1}{2}x_3 = 2 \Rightarrow \tfrac{3}{2}x_2 + \tfrac{1}{2} = 2 \Rightarrow \tfrac{3}{2}x_2 = \tfrac{3}{2} \Rightarrow x_2 = 1$$

$$2x_1 + x_2 + x_3 = 4 \Rightarrow 2x_1 + 2 = 4 \Rightarrow 2x_1 = 2 \Rightarrow x_1 = 1$$

thus the least-squares solution $\vec{x}$ that minimizes $\|Ax-b\|_2^2 + \|x\|_2^2$ is

$$\vec{x} = (1, 1, 1)^T.$$

Multi-Step Problem

(a.) See the Homework 5 plots appendix for the scatter plot of the data called "Multi-Step Problem part a."

(b.) The form of our quadratic is $z(t) = at^2 + bt + c$, thus the coefficients of the quadratic that best fits the data above computed using the normal equation are: $a = -4.9$, $b = 297.7$, and $c = 127.2$. Thus our least-squares polynomial is: $z(t) = -4.9t^2 + 297.7t + 127.1879 \Rightarrow z(t) = -4.9t^2 + 297.7t + 127.2$

(c.) According to the least-squares solution in (b), the initial velocity of the rocket is $V_0 = 297.7$ meters per second, the inital altitude ~~above sea level~~ at which the rocket was launched is $z_0 = 127.2$ meters, and lastly the magnitude of the acceleration due to Earth's gravity is $-\frac{1}{2}g = -4.9 \Rightarrow g = 9.8$ meters per second squared.

(d.) See the Homework 5 plots appendix for the superimposed scatter plot of the data with the least-squares polynomial overlaid called "Multi-Step problem part d."

(e.) To find the time at which the rocket reaches its pinnacle of trajectory according to the least-squares quadratic we will:

- STEP 1: Find 1st derivative:

$$z(t) = -4.9t^2 + 297.7t + 127.2 \Rightarrow z'(t) = -9.8t + 297.2$$

- STEP 2: Find critical point(s) t^*:

$$z'(t) = 0 \Rightarrow -9.8t + 297.7 = 0$$
$$-9.8t = -297.7$$
$$t^* = 30.377551 \approx 30.4$$

- STEP 3: Find 2nd derivative:

$$z'(t) = -9.8t + 297.2 \Rightarrow z''(t) = -9.8$$

°STEP 4: 2nd derivative test:

since $z''(t) = -9.8 < 0 \ \forall t$, it follows that $z''(t^*) = z''(30.4) = -9.8 < 0$.

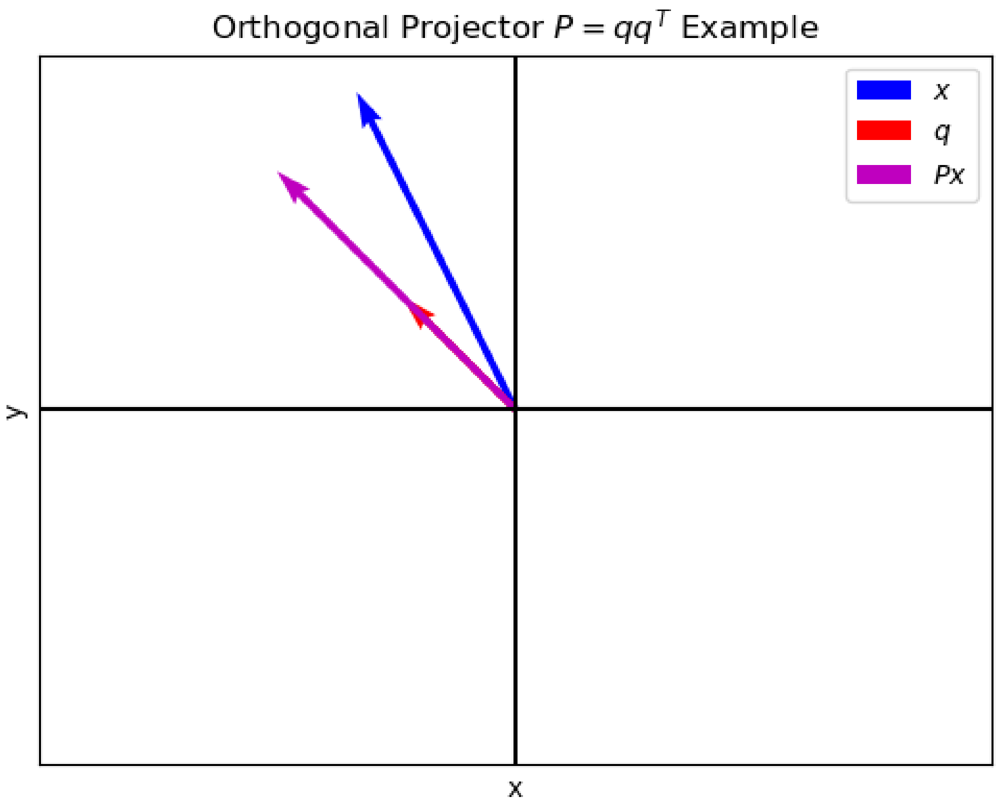
Hence $t^* = 30.4$ is a maximum.

Therefore, the rocket reaches the pinnacle of its trajectory at a time $t = 30.4$ seconds after its launch (according to the least squares quadratic fit to the data.)

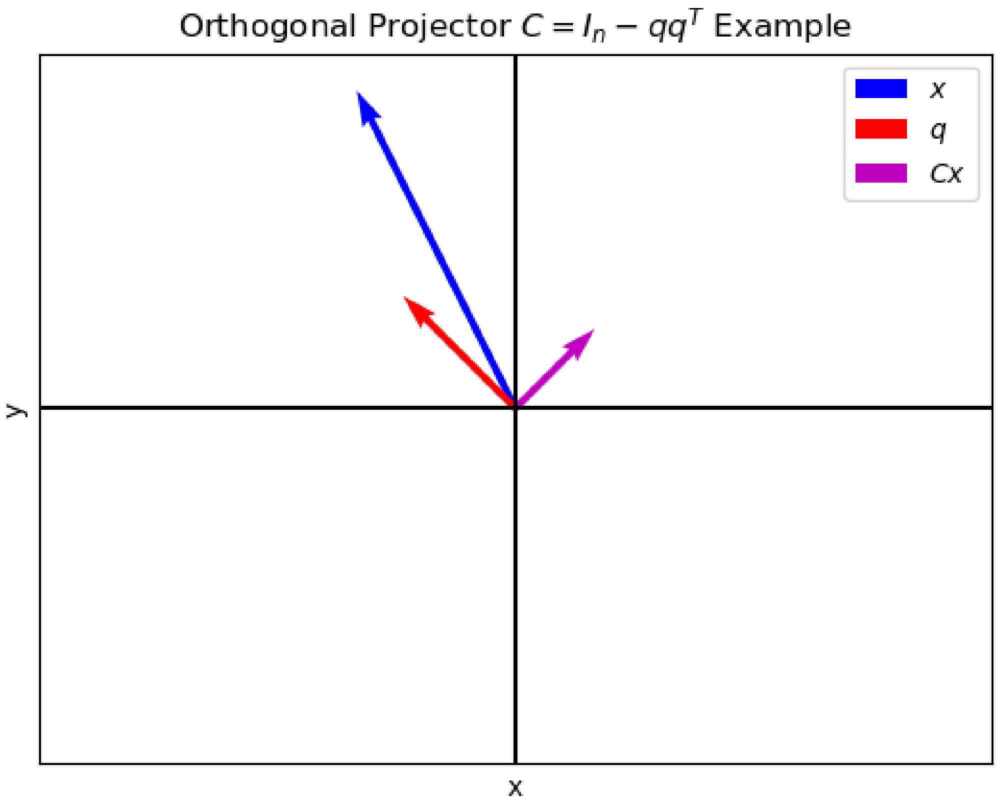
Homework 5 Plots

Note: Axis markers are removed due to the confusing way that matplotlib in Python scales the axes. Refer the the write-up to see what each vector equals.

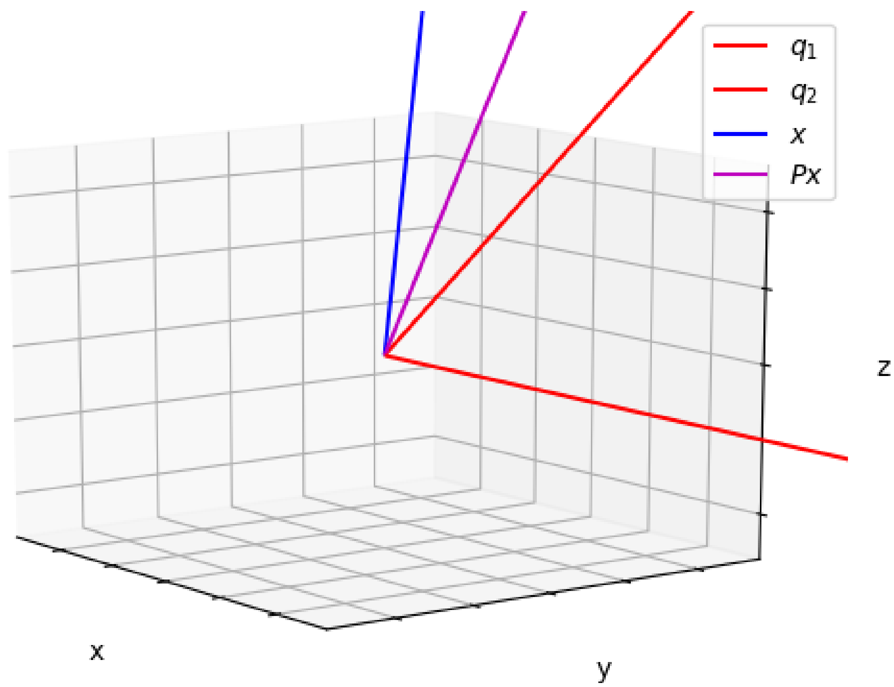
Exercise 1 Part d:



Exercise 1 Part f:

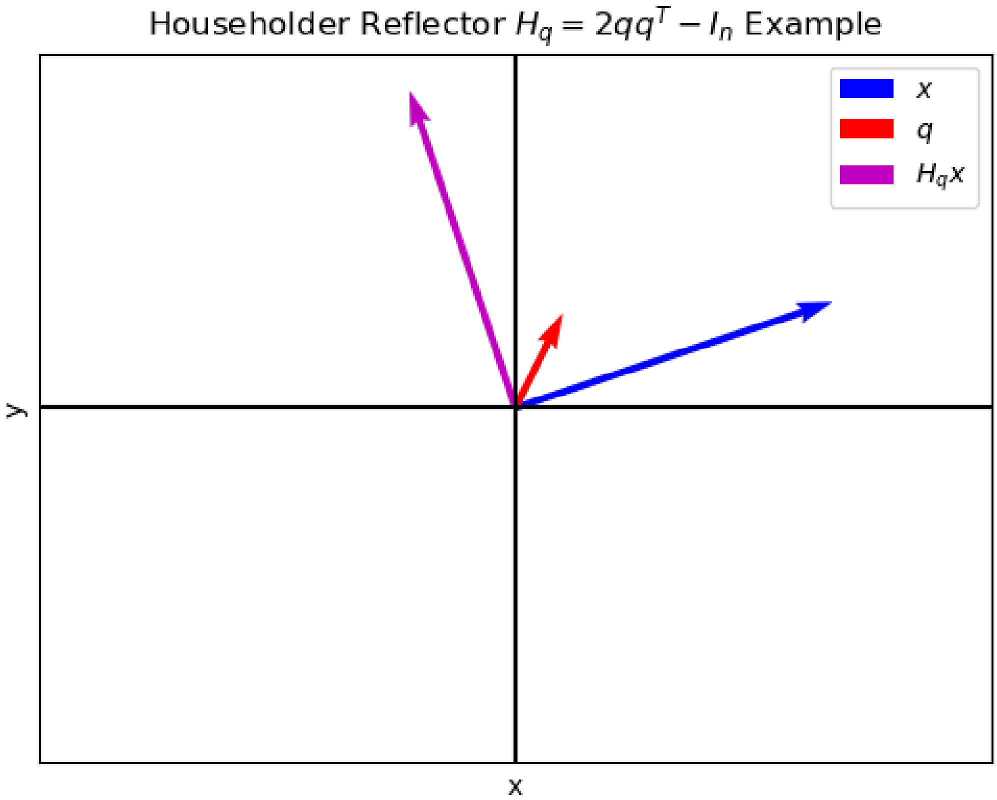


Exercise 2 Part bii:



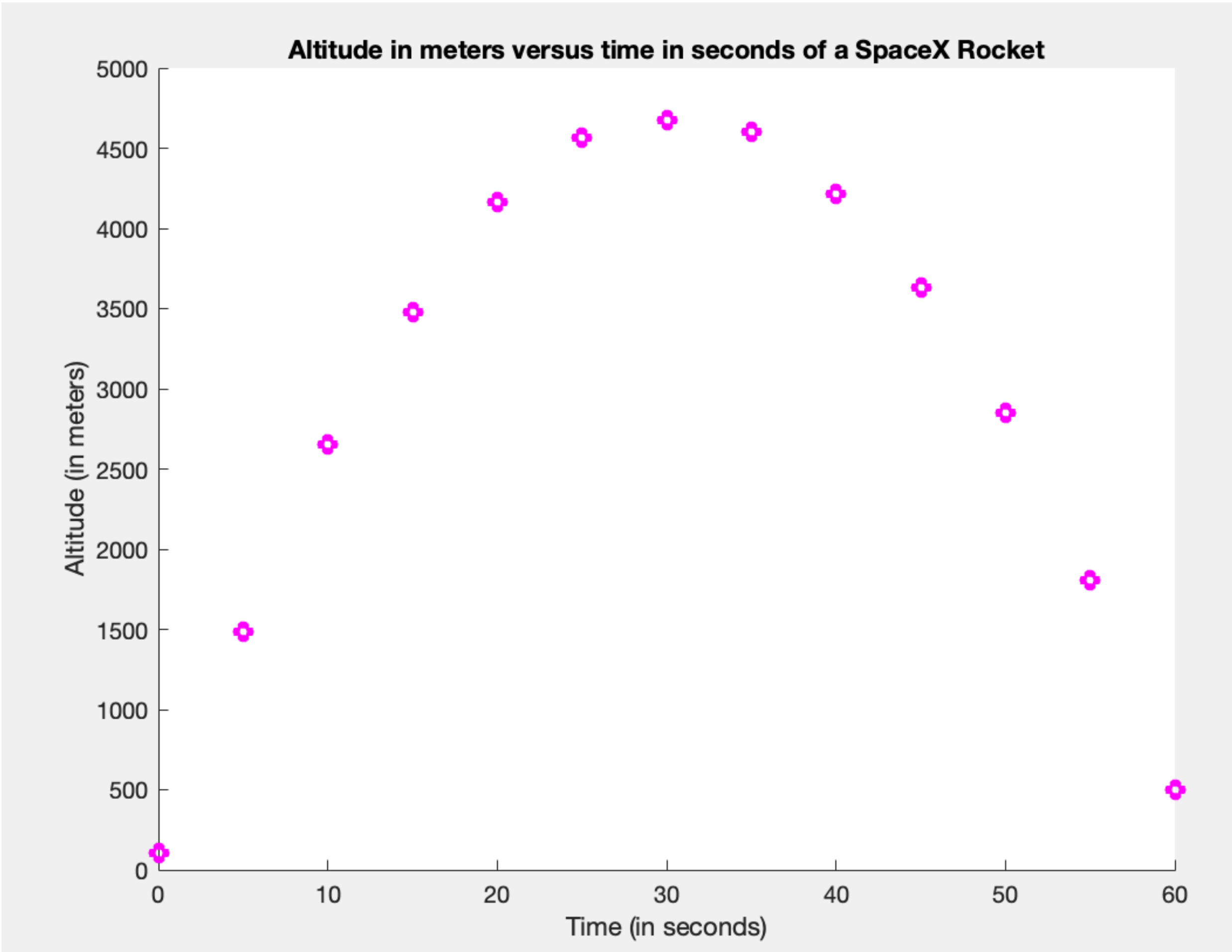
The span$\{\vec{q_1}, \vec{q_2}\}$ forms a plane in $\mathbb{R}^3$, however quiverplots in matplotlib have a hard time plotting vectors and planes in the same figure. For this figure, take the red vectors as the basis for the plane they span.

Exercise 3 Part d:



The span$\{\vec{q}\}$ forms a line in $\mathbb{R}^2$, however quiverplots in matplotlib have a hard time plotting vectors and lines in the same figure. For this figure, take the red vector as the basis for the line it spans.

Multi-Step Problem Part a:



Multi-Step Problem Part d: