Sample Solutions for Assignment 3.

Reading: Lectures 6-8 in the text.


$I - 2P$ is unitary since $(I - 2P)^*(I - 2P) = I - 2P - 2P^* + 4P^*P = I$, where the last equality follows because $P = P^*$ and $P = P^2$. Geometrically, $(I - P)v$ is a vector orthogonal to the subspace onto which $P$ projects, and $(I - 2P)v$ is the reflection of $v$ through the subspace orthogonal to the range of $P$.


(a) 

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}.$$  

The columns of $A$ are orthogonal and if we normalize the first column by dividing by $\sqrt{2}$, then they are orthonormal. Thus the orthogonal projector $P$ onto the range of $A$ is 

$$P = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1 \\ 1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix}.$$  

The image of the vector $(1, 2, 3)^*$ is 

$$\begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}.$$  

(b) 

$$B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}.$$  

An orthonormal basis for the range of $B$ is $(1/\sqrt{2}, 0, 1/\sqrt{2})^T$, $(1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3})^T$. Hence the orthogonal projector $P$ onto the range of $B$ is 

$$P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} \\ 0 & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3} \end{bmatrix} = \begin{bmatrix} 5/6 & 1/3 & 1/6 \\ 1/3 & 1/3 & -1/3 \\ 1/6 & -1/3 & 5/6 \end{bmatrix}.$$
The image of the vector \((1, 2, 3)^*\) is
\[
\begin{bmatrix}
2 \\
0 \\
2
\end{bmatrix}.
\]


(a) Since the columns of \(A\) are already orthogonal, all we have to do is normalize to get the reduced QR factorization:
\[
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
1 & 0
\end{bmatrix} = \begin{bmatrix}
\sqrt{2}/2 & 0 \\
0 & 1 \\
\sqrt{2}/2 & 0
\end{bmatrix} \begin{bmatrix}
\sqrt{2} \\
0 \\
1
\end{bmatrix} \equiv \hat{Q}\hat{R}.
\]

To get the full QR factorization, append a vector to \(\hat{Q}\) that is orthogonal to both columns of \(\hat{Q}\) and has norm 1, and append a row of zeros to \(\hat{R}\):
\[
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
1 & 0
\end{bmatrix} = \begin{bmatrix}
\sqrt{2}/2 & 0 & \sqrt{2}/2 \\
0 & 1 & 0 \\
\sqrt{2}/2 & 0 & -\sqrt{2}/2
\end{bmatrix} \begin{bmatrix}
\sqrt{2} \\
0 \\
1
\end{bmatrix} \equiv QR.
\]

(b) Using Gram-Schmidt orthogonalization, start by normalizing the first column:
\[
q_1 = \begin{bmatrix}
\sqrt{2}/2 \\
0 \\
\sqrt{2}/2
\end{bmatrix}, \quad r_{11} = \sqrt{2}.
\]

Now orthogonalize the second column against \(q_1\):
\[
\tilde{q}_2 = \begin{bmatrix}
2 \\
1 \\
0
\end{bmatrix} - \sqrt{2} \begin{bmatrix}
\sqrt{2}/2 \\
0 \\
\sqrt{2}/2
\end{bmatrix} = \begin{bmatrix}
1 \\
1 \\
-1
\end{bmatrix}, \quad r_{12} = \sqrt{2}.
\]

Finally, normalize to get \(q_2\):
\[
q_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|} = \begin{bmatrix}
\sqrt{3}/3 \\
\sqrt{3}/3 \\
-\sqrt{3}/3
\end{bmatrix}, \quad r_{22} = \sqrt{3}.
\]

Thus the reduced QR factorization is:
\[
\begin{bmatrix}
1 & 2 \\
0 & 1 \\
1 & 0
\end{bmatrix} = \begin{bmatrix}
\sqrt{2}/2 & \sqrt{3}/3 & -\sqrt{6}/6 \\
0 & \sqrt{3}/3 & \sqrt{6}/3 \\
\sqrt{2}/2 & -\sqrt{3}/3 & \sqrt{6}/6
\end{bmatrix} \begin{bmatrix}
\sqrt{2} \\
\sqrt{2} \\
\sqrt{3}
\end{bmatrix} \equiv \hat{Q}\hat{R}.
\]

To get the full QR factorization, append a vector to \(\hat{Q}\) that is orthogonal to both columns of \(\hat{Q}\) and has norm 1, and append a row of zeros to \(\hat{R}\):
\[
\begin{bmatrix}
1 & 2 \\
0 & 1 \\
1 & 0
\end{bmatrix} = \begin{bmatrix}
\sqrt{2}/2 & \sqrt{3}/3 & -\sqrt{6}/6 \\
0 & \sqrt{3}/3 & \sqrt{6}/3 \\
\sqrt{2}/2 & -\sqrt{3}/3 & \sqrt{6}/6
\end{bmatrix} \begin{bmatrix}
\sqrt{2} \\
\sqrt{2} \\
0
\end{bmatrix} \equiv QR.
(a) If all of the diagonal entries of \( \hat{R} \) are nonzero, then \( \hat{R} \) is nonsingular and \( A\hat{R}^{-1} = \hat{Q} \). Thus, each of the \( n \) columns of \( \hat{Q} \) lies in the range of \( A \), and since these columns are linearly independent, the rank of \( A \) is at least \( n \). The rank of \( A \) cannot be greater than \( n \) because for any vector \( x \), \( Ax = \hat{Q}\hat{R}x \), which is a linear combination of the columns of \( \hat{Q} \). Thus, when all diagonal entries of \( \hat{R} \) are nonzero, the range of \( A \) is the span of the columns of \( \hat{Q} \) and the rank of \( A \) is \( n \).

If some diagonal entry of \( \hat{R} \), say \( \hat{r}_{jj} \) is 0 (and is the first 0 on the diagonal of \( \hat{R} \)), then column \( j \) of \( A \) is a linear combination of columns 1 through \( j - 1 \). To see this, note that if \( e_j \) is the \( j \)th unit vector, then \( Ae_j = \hat{Q}\hat{R}e_j \) is a linear combination of columns 1 through \( j - 1 \) of \( \hat{Q} \); i.e., a linear combination of columns 1 through \( j - 1 \) of \( A \). Since the columns of \( A \) are not linearly independent, it must have rank less than \( n \).

(b) Suppose \( \hat{R} \) has \( k \) nonzero diagonal entries. The columns of \( A \) corresponding to the nonzero diagonal entries of \( \hat{R} \) are linearly independent. If, say, \( \hat{r}_{ss} \) and \( \hat{r}_{tt}, t > s \), are two such nonzero entries, then \( Ae_s = \hat{Q}\hat{R}e_s \) is a linear combination of columns 1 through \( s \) of \( \hat{Q} \), while \( Ae_t = \hat{Q}\hat{R}e_t \) is a linear combination of columns 1 through \( t \) of \( \hat{Q} \), with the coefficient of \( q_t \) being nonzero. Thus, it is independent of any linear combination of \( q_1, \ldots, q_s \). Therefore the rank of \( A \) is at least \( k \). The rank of \( A \) could be greater than \( k \), however. Suppose column \( j \) of \( A \) is the first column that is a linear combination of previous columns; i.e., of \( q_1, \ldots, q_{j-1} \). Thus \( r_{jj} = 0 \). We introduce an arbitrary vector \( q_j \) that is orthogonal to \( q_1, \ldots, q_{j-1} \). Now suppose column \( j + 1 \) of \( A \) is not a linear combination of \( q_1, \ldots, q_{j-1} \) (so that the rank of \( A \) is at least \( j \)) but is a linear combination of \( q_1, \ldots, q_j \). Then the diagonal entry \( r_{j+1,j+1} \) will be 0. Thus we have two zero diagonal entries, \( j - 1 \) nonzero diagonal entries, but the rank of the first \( j + 1 \) columns is \( j \). The argument can be repeated for other columns. Column \( j + 2 \) might be independent of \( q_1, \ldots, q_j \), but if it is a linear combination of \( q_1, \ldots, q_j \) and the arbitrarily chosen new column \( q_{j+1} \), then the diagonal entry \( r_{j+2,j+2} \) will be zero as well, even though the rank of the first \( j + 2 \) columns is \( j + 1 \).

As a simple example, take

\[
A = \begin{bmatrix}
1 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 \\
\vdots & \vdots & \ddots & \ddots \\
0 & 0 & \cdots & 0 & 1 \\
0 & 0 & \cdots & 0 & 0
\end{bmatrix}
\]

\( A \) is already upper triangular, so its reduced QR factorization is \( A = I \cdot A \). If \( A \) is \( m \times n \) with \( n \leq m \), then it has rank \( n - 1 \) but only one nonzero
The inner product $q_i^* v_j$ in the innermost loop requires $m$ multiplications and
$m - 1$ additions. The next line, $v_j = v_j - r_{ij} q_i$, requires $m$ multiplications
and $m$ subtractions. Thus, $4m - 1$ operations are performed in each of the
$(n - i)$ passes through the innermost loop. Outside that loop, but inside
the loop over $i$, we compute the norm of a vector, which requires an inner
product (of the vector with itself) and a square root, or $(2m - 1) + 1 = 2m$
operations. We also do $m$ divisions, giving a total of $3m$ operations. Thus,
the total number of operations for Algorithm 8.1 is

$$\sum_{i=1}^{n} \left[ 3m + \sum_{j=i+1}^{n} (4m - 1) \right] = \sum_{i=1}^{n} [3m + (4m - 1)(n - i)] = 3mn + (4m - 1) \sum_{i=1}^{n} (n - i)$$

$$= 3mn + (4m - 1) \frac{n(n - 1)}{2} = 2mn^2 + mn - \frac{1}{2} n^2 + \frac{1}{2} n.$$