Sergio Garcia Tapia
Numerical Linear Algebra, Lloyd Trefethen and David Bau III
Lecture 7: QR Factorization
June 23, 2024

Lecture 7: QR Factorization

Exercise 1. Consider again the matrices A and B of Exercise 6.4.

(a) Using any method you like, determine (on paper) a reduced QR factorization $A = \hat{Q}\hat{R}$ and a full qr factorization $A = QR$.

(b) Again, using any method you like, determine reduced and full QR factorizations $B = \hat{Q}\hat{R}$ and $B = QR$.

Solution:

(a) Matrix A is

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

Let $a_1 = (1, 0, 1)^*$ and $a_2 = (0, 1, 0)^*$. We proceed by using the Gram Schmidt Orthogonalization Procedure. Since $r_{11} = \|a_1\|_2 = \sqrt{2}$, we define $q_1 = \frac{1}{\sqrt{2}} a_1$. For q_2, we get

$$q_2 = \frac{a_2 - r_{12} q_1}{r_{22}}$$

where

$$r_{12} = q_1^* a_2 = \frac{1}{\sqrt{2}} \langle (1, 0, 1)^*, (0, 1, 0) \rangle = 0$$

Hence $r_{12} = 0$. Meanwhile, we have

$$a_2 - r_{12} q_1 = (0, 1, 0) - 0 \cdot q_1 = (0, 1, 0)$$

We see that $r_{22} = \|a_2 - r_{12} q_1\| = \|a_2\| = 1$. Thus the reduced QR factorization is

$$\hat{Q} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \\ \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \qquad \hat{R} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix}$$

As expected, $\hat{Q}$ has orthonormal columns, and $\hat{R}$ is upper-triangular. To obtain the full factorization, we add a row of 0's to $\hat{R}$ to get R. To pass from $\hat{Q}$ to Q, we append any column that is orthonormal to those in $\hat{Q}$. If $q_3 = (q_{13}, q_{23}, q_{33})$ is such a column, it must satisfy:

$$0 = q_1^* q_3 = \frac{1}{\sqrt{2}} \langle (1, 0, 1), (q_{13}, q_{23}, q_{33}) \rangle = \frac{1}{\sqrt{2}} (q_{13} + q_{33})$$
$$0 = q_2^* q_3 = \langle (0, 1, 0), (q_{13}, q_{23}, q_{33}) \rangle = q_{23}$$

Hence, $q_{23} = 0$, and $q_{13} = -q_{33}$. If we set $q_{33} = \frac{1}{\sqrt{2}}$, then $a_{13} = -\frac{1}{\sqrt{2}}$. Now

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \qquad R = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

(b) Matrix B is

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

Set $b_1 = (2, 0, 1)^*$. Then $r_{11} = \|b_1\|_2 = \sqrt{5}$, and $q_1 = \frac{1}{\sqrt{5}} b_1$. Next, we set $b_2 = (0, 1, 0)$, and compute

$$\begin{aligned} r_{12} &= q_1^* b_2 = \frac{1}{\sqrt{5}} \langle (2, 0, 1), (0, 1, 0) \rangle = 0 \\ v_2 &= b_2 - r_{12} q_2 \\ &= b_2 - 0 \cdot q_2 \\ &= b_2 \\ \|v_2\|_2 &= \|b\|_2 = 1 \\ q_2 &= \frac{v_2}{\|v_2\|} = \frac{b_2}{\|b_2\|_2} = b_2 \end{aligned}$$

Therefore the reduced QR factorization is

$$\hat{Q} = \begin{bmatrix} \frac{2}{\sqrt{5}} & 0 \\ 0 & 1 \\ \frac{1}{\sqrt{5}} & 0 \end{bmatrix} \qquad \hat{R} = \begin{bmatrix} \sqrt{5} & 0 \\ 0 & 1 \end{bmatrix}$$

If we set $q_3 = \frac{1}{\sqrt{5}}(-2, 0, 1)$, then the full QR factorization is

$$Q = \begin{bmatrix} \frac{2}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} \\ 0 & 1 & \\ \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \end{bmatrix} \qquad R = \begin{bmatrix} \sqrt{5} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Exercise 2. Let A be the matrix with the property that columns $1, 3, 5, 7, \ldots$ are orthogonal to columns $2, 4, 6, 8, \ldots$. In a reduced QR factorization $A = \hat{Q}\hat{R}$, what special structure does $\hat{R}$ possess? You may assume that A has full rank.

Solution: Suppose we applied the Gram Schmidt Orthogonalization Procedure to the columns of A. Since we are assuming that A has full rank, we know that its columns are nonzero, and hence they all have a nonzero norm. We begin by setting $q_1 = \frac{a_1}{\|a_1\|_2}$. Since a_1 is orthogonal to a_2, we find that

$$r_{12} = q_1^* a_2 = \frac{1}{\|a_1\|_2} a_1^* a_2 = 0$$

Thus,

$$q_2 = \frac{a_2 - r_{12}q_1}{r_{22}} = \frac{a_2}{r_{22}}$$

Here, $r_{22} = \|a_2\|_2$. Next, note that $r_{13} = q_1^* a_3$, and $r_{23} = q_2^* a_3 = 0$. Thus,

$$q_3 = \frac{a_3 - r_{13}q_1 - r_{23}q_2}{r_{33}} = \frac{a_3 - r_{13}q_1}{r_{33}}$$

Proceeding this way, we see that $\hat{R}$ is an upper-triangular matrix with a checkered pattern. That is, given any upper-triangular entry, its four neighboring entries are always 0. That is, the entry above, below, left, and right of $r_{ij} = 0$, where $i \leq j$.

Exercise 3. Let A be an $m \times m$ matrix, and let a_j be its jth column. Give an algebraic proof of *Hadamard's inequality*:

$$|\det A| \leq \prod_{j=1}^{m} \|a_j\|_2$$

Also, give a geometric interpretation of this result, making use of the fact that the determinant equals the volume of a parallelepiped.

Solution:

Proof. Suppose A is nonzero, and let Q and R be the components of a QR factorization of A. This means that $A = QR$, where Q has orthonormal columns, and R is upper-triangular. The fact that Q has orthonormal means that it is *unitary*, and hence $Q^* = Q^{-1}$. By Theorem 5.4, the singular values of Q are the eigenvalues of Q^*Q, and since $Q^*Q = Q^{-1}Q = I$, this implies that all of the singular values of Q are 1. By Theorem 5.6, we know that $|\det(Q)|$ is the product of its singular values, and hence, $|\det(Q)| = 1$.

Since R is upper-triangular, its determinant is the product of its diagonal entries. On the other hand, we know that

$$a_j = \sum_{i=1}^{j} r_{ij} q_i$$

and since the list $q_1, \ldots, q_j$ is orthonormal, we get

$$\|a_j\|_2 = \sqrt{\sum_{i=1}^{j} r_{ij}^2} = \sqrt{r_{1j}^2 + \cdots + r_{j-1,j}^2 + r_{jj}^2} \geq |r_{jj}|$$

Now using the fact the determinant of a product is a product of the determinants, we get

$$\begin{aligned}|\det(A)| &= |\det(QR)| \\ &= |\det(Q)\cdot\det(R)| \\ &= |\det(Q)|\cdot|\det(R)| \\ &= |\det(R)| \\ &= \left|\prod_{j=1}^{m} r_{jj}\right| \\ &= \prod_{j=1}^{m} |r_{jj}| \\ &\le \prod_{j=1}^{m} \|a_j\|_2\end{aligned}$$

This implies that the volume of the parallelpiped determined by the vectors corresponding to the columns of A does not exceed the product of the lengths of these vectors. □

Exercise 4. Let $x^{(1)}$, $y^{(1)}$, $x^{(2)}$, and $y^{(2)}$ be nonzero vectors in $\mathbb{R}^3$ with the property that $x^{(1)}$ and $y^{(1)}$ are linearly independent, and so are $x^{(2)}$ and $y^{(2)}$. Consider the two planes in $\mathbb{R}^3$,

$$P^{(1)} = \langle x^{(1)}, y^{(1)}\rangle, \quad P^{(2)} = \langle x^{(2)}, y^{(2)}\rangle$$

Suppose we wish to find a nonzero vector $v \in \mathbb{R}^3$ that lies in the intersection $P = P^{(1)} \cap P^{(2)}$. Devise a method for solving this problem by reducing it to the computation of the QR factorization of three 3×2 matrices.

Solution: Let $A \in \mathbb{R}^{3\times 2}$ be the matrix whose columns are $x^{(1)}$ and $y^{(1)}$, and let $B \in \mathbb{R}^{3\times 2}$ be the matrix whose columns are $x^{(2)}$ and $y^{(2)}$. Then

$$P^{(1)} = \text{range}(A) \quad P^{(2)} = \text{range}(B) \quad P = \text{range}(A) \cap \text{range}(B)$$

Let $A = \hat{Q}_A\hat{R}_A$ be the reduced QR factorization of A and $B = \hat{Q}_B\hat{R}_B$ be the QR factorization of B, where $q_{A,1}$, $q_{A,2}$ are the orthonormal columns of Q_A and $q_{B,1}$, $q_{B,2}$ are the orthonormal columns of Q_B, with $\langle x^{(1)}, y^{(1)}\rangle = \langle q_{A,1}, q_{A,2}\rangle$ and $\langle x^{(2)}, y^{(2)}\rangle = \langle q_{B,1}, q_{B,2}\rangle$. We can extend an orthonormal basis of $\mathbb{R}^3$ by appending a vector $q_{A,3}$ to the list consisting of the columns of $\hat{Q}_A$. Similarly, we can append $q_{B,3}$ to the list consisting of the columns of $\hat{Q}_B$.

Note that $q_{A,3}$ is orthogonal to $\langle q_{A,1}, q_{A,2}\rangle$ and $q_{B,3}$ is orthogonal to $\langle q_{B,1}, q_{B,2}\rangle$. Let $C \in \mathbb{R}^{3\times 2}$ be the matrix whose columns are $q_{A,3}$ and $q_{B,3}$, and let $C = \hat{Q}_C\hat{R}_C$ be the QR factorization of C.

Let $C = \hat{Q}_C\hat{R}_C$ be the reduced QR factorization of C:

- If C has linearly dependent columns, the $\hat{Q}_C$ will have a single column, otherwise it will have two. In either case, extend the list of orthonormal columns of $\hat{Q}_C$ to a basis of $\mathbb{R}^3$, and take the third vector, calling it v. Then v has the property that it is orthogonal to the first two columns. Suppose for a moment that the columns of C were linearly independent. Since $q_{A,3}$ is a basis for range$(A)^\perp$ and $q_{B,3}$ is a basis for range$(B)^\perp$, and v is orthogonal to both, we see that $v \in (\text{range}(A)^\perp)^\perp = \text{range}(A)$, and $v \in (\text{range}(B)^\perp)^\perp = \text{range}(B)$. Hence, $v \in \text{range}(A) \cap \text{range}(B)$, as we set out to show.

- If the columns were linearly dependent, then this implies that $q_{A,3}$ and $q_{B,3}$ are linearly dependent, meaning they are scalar multiples of one another. In this case, either $q_{A,3}$ or $q_{B,3}$ is the first column of $\hat{Q}_C$, and nevertheless, v is perpendicular to it by construction.

Hence, the three 3×2 matrices to which we apply QR factorization are A, after which we save the third column $q_{A,3}$ of Q_A, then B, after which we save the third column $q_{B,3}$ of Q_B, and C, the matrix consisting of $q_{A,3}$ and $q_{B,3}$. Then v is the third column of Q_C in the full QR factorization of C.

Exercise 5. Let A be an $m \times n$ matrix ($m \geq n$), and let $A = \hat{Q}\hat{R}$ be a reduced QR factorization.

(a) Show that A has rank n if and only if all the diagonal entries of $\hat{R}$ are nonzero.

(b) Suppose $\hat{R}$ has k nonzero diagonal entries for some k with $) \leq k < n$. What does this imply about the rank of A? Exactly k? At least k? At most k? Give a precise answer, and prove it.

Solution:

(a) *Proof.* By Equation (7.8), the diagonal entries of $\hat{R}$ satisfy

$$|r_{jj}| = \left\| a_j - \sum_{i=1}^{j} r_{ij} q_i \right\|_2$$

Note that the span of $q_1, \dots, q_{j-1}$ is equivalent to the span of $a_1, \dots, a_{j-1}$ by construction. Therefore, the equation above implies that the vector whose norm is $|r_{jj}|$ belongs to the span of $a_1, \dots, a_{j-1}, a_j$, meaning that it is a linear combination of the. The matrix A has full rank n if and only if the columns $a_1, \dots, a_n$ is linearly independent, which is if and only if $a_1, \dots, a_j$ is linearly independent for every $j \in \{1, \dots, n\}$. This happens if and only if no linear combination of them yields the 0 vector, which in turn implies that the vector whose 2-norm we are computing above can never be zero, and hence $|r_{jj}| \neq 0$.

As an alternative proof, Suppose that $\hat{U}\hat{\Sigma}V^*$ is a reduced SVD of A, where $\hat{U}$ is $m \times n$, $\hat{\Sigma}$ is $n \times n$, and V^* is $n \times n$. Then

$$\begin{aligned} A &= \hat{U}\hat{\Sigma}V^* \\ \hat{Q}\hat{R} &= \hat{U}\hat{\Sigma}V^* \\ \hat{R} &= (\hat{Q}^*\hat{U})\hat{\Sigma}\hat{V}^* \end{aligned}$$

Since $\hat{Q}^*$ and $\hat{U}$ are unitary, their product is also unitary. Therefore, the equation above is an SVD of $\hat{R}$. This implies that the singular values of A, which are the diagonal entries in $\hat{\Sigma}$, are precisely the singular values of $\hat{R}$. In particular, since A the rank of A is n, all of its singular values are nonzero, and hence all of the singular values of R are nonzero. This occurs if and only if all the eigenvalues of R are nonzero, which are the diagonal entries of R because R is upper-triangular. □

(b) *Proof.* I will prove that the rank of A is at least k, and show an example to conclude that it may not necessarily be equal to k.

Suppose that entries $r_{i_1 i_1}, \ldots, r_{i_k, i_k}$ on the diagonal of $\hat{R}$ are nonzero, where $i_1, \ldots, i_k \in \{1, \ldots, n\}$, with $i_1 < \cdots < i_k$ and such that no diagonal entry r_{jj} is zero for $j < i_k$. Let a_j denote the jth column of A, and q_j denotes the jth column of $\hat{Q}$. Then $a_{i_1}, \ldots, a_{i_k}$ is a linearly independent list. Since the span of the a_j's equals the span of the q_j's, this means that a linear combination of the a_j's is a linear combination of the q_j's. Then the coefficient of a_{ik} is precisely the coefficient of q_{ik} in this combination because a_{ik} is the only vector in the list $a_{i_1}, \ldots, a_{i_k}$ depending on q_{i_k}. Suppose that such a linear combination resulted in 0. If the coefficient of q_{i_k} were nonzero, then q_{ik} would be a linear combination of $q_{i_1}, \ldots, q_{i_{k-1}}$, which is impossible because the latter list is orthonormal. Hence the coefficient of q_{i_k} and hence a_{i_k} must be 0. In a similar fashion, we conclude that the rest of the coefficients are also 0, implying that $a_{i_1}, \ldots, a_{i_k}$ is a linearly independent list, and hence that $\operatorname{rank}(A) \geq k$.

To show that it need not be exactly k, its enough to provide an example. Suppose $A \in \mathbb{R}^{m \times n}$ and the columns of A are $e_1, \ldots, e_{n-2}, 0, e_{n-1}$. Then in the QR factorization, we have $q_j = e_j$ for $1 \leq j \leq n-2$. When we reach step $j-1$, get $r_{j-1,j-1} = 0$, so we must arbitrarily pick an orthonormal vector. Suppose we choose $q_{j-1} = e_{j-1}$. When we go on to step j, we find that e_{j-1} is already in the span of $q_1, \ldots, q_{j-1}$. Thus we get

$$
\begin{aligned}
r_{ij} = q_i^* a_j = q_i^* e_{j-1} = q_i q_{j-1} &= \begin{cases} 0 & i < j-1 \\ 1 & i = j-1 \end{cases} \\
|r_{jj}| &= \left\| a_j - \sum_{i=1}^{j-1} r_{ij} q_i \right\|_2 \\
&= \|e_{j-1} - q_{j-1}\| = 0
\end{aligned}
$$

Hence $r_{jj} = 0$. In other words, the matrix $\hat{R}$ has $n-2$ nonzero entries, but the rank of A is $n-1$. Hence, we can only conclude that $\operatorname{rank}(A) \geq k$, where k is the number of nonzero entries in the diagonal of $\hat{R}$. □