

1 QR Factorization

One of the main virtues of orthogonal matrices is that they can be easily inverted—the transpose is the inverse. This fact, combined with the factorization theorem in this section, provides a useful way to simplify many matrix calculations (for example, in [least squares approximation](#)).

Definition 1. Let A be an $m \times n$ matrix with independent columns. A *QR-factorization* of A expresses it as $A = QR$ where Q is $m \times n$ with orthonormal columns and R is an invertible and upper triangular matrix with positive diagonal entries.

The importance of the factorization lies in the fact that there are computer algorithms that accomplish it with good control over round-off error, making it particularly useful in matrix calculations. The factorization is a matrix version of the Gram-Schmidt process.

Suppose $A = \begin{bmatrix} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & & | \end{bmatrix}$ is an $m \times n$ matrix with linearly independent columns $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n$. The Gram-Schmidt algorithm can be applied to these columns to provide orthogonal columns $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ where $\mathbf{f}_1 = \mathbf{c}_1$ and

$$\mathbf{f}_k = \mathbf{c}_k - \frac{\mathbf{c}_k \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2} \mathbf{f}_1 + \frac{\mathbf{c}_k \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2} \mathbf{f}_2 - \cdots - \frac{\mathbf{c}_k \cdot \mathbf{f}_{k-1}}{\|\mathbf{f}_{k-1}\|^2} \mathbf{f}_{k-1}$$

for each $k = 2, 3, \dots, n$. Now write $\mathbf{q}_k = \frac{1}{\|\mathbf{f}_k\|} \mathbf{f}_k$ for each k . Then $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ are orthonormal columns, and the above equation becomes

$$\|\mathbf{f}_k\| \mathbf{q}_k = \mathbf{c}_k - (\mathbf{c}_k \cdot \mathbf{q}_1) \mathbf{q}_1 - (\mathbf{c}_k \cdot \mathbf{q}_2) \mathbf{q}_2 - \cdots - (\mathbf{c}_k \cdot \mathbf{q}_{k-1}) \mathbf{q}_{k-1}$$

Using these equations, express each $\mathbf{c}_k$ as a linear combination of the $\mathbf{q}_j$:

$$\begin{aligned} \mathbf{c}_1 &= \|\mathbf{f}_1\| \mathbf{q}_1 \\ \mathbf{c}_2 &= (\mathbf{c}_2 \cdot \mathbf{q}_1) \mathbf{q}_1 + \|\mathbf{f}_2\| \mathbf{q}_2 \\ \mathbf{c}_3 &= (\mathbf{c}_3 \cdot \mathbf{q}_1) \mathbf{q}_1 + (\mathbf{c}_3 \cdot \mathbf{q}_2) \mathbf{q}_2 + \|\mathbf{f}_3\| \mathbf{q}_3 \\ &\vdots \\ \mathbf{c}_n &= (\mathbf{c}_n \cdot \mathbf{q}_1) \mathbf{q}_1 + (\mathbf{c}_n \cdot \mathbf{q}_2) \mathbf{q}_2 + (\mathbf{c}_n \cdot \mathbf{q}_3) \mathbf{q}_3 + \cdots + \|\mathbf{f}_n\| \mathbf{q}_n \end{aligned}$$

These equations have a matrix form that gives the required factorization:

$$\begin{aligned} A &= \begin{bmatrix} | & | & | & \cdots & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 & \cdots & \mathbf{c}_n \\ | & | & | & & | \end{bmatrix} \\ &= \begin{bmatrix} | & | & | & \cdots & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 & \cdots & \mathbf{q}_n \\ | & | & | & & | \end{bmatrix} \begin{bmatrix} \|\mathbf{f}_1\| & \mathbf{c}_2 \cdot \mathbf{q}_1 & \mathbf{c}_3 \cdot \mathbf{q}_1 & \cdots & \mathbf{c}_n \cdot \mathbf{q}_1 \\ 0 & \|\mathbf{f}_2\| & \mathbf{c}_3 \cdot \mathbf{q}_2 & \cdots & \mathbf{c}_n \cdot \mathbf{q}_2 \\ 0 & 0 & \|\mathbf{f}_3\| & \cdots & \mathbf{c}_n \cdot \mathbf{q}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \|\mathbf{f}_n\| \end{bmatrix} \end{aligned} \quad (1)$$

Here the first factor $Q = \begin{bmatrix} | & | & | & \cdots & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 & \cdots & \mathbf{q}_n \\ | & | & | & \cdots & | \end{bmatrix}$ has orthonormal columns, and the second factor is an $n \times n$ upper triangular matrix R with positive diagonal entries (and so is invertible). We record this in the following theorem.

Theorem 1 (QR-Factorization). Every $m \times n$ matrix A with linearly independent columns has a QR-factorization $A = QR$ where Q has orthonormal columns and R is upper triangular with positive diagonal entries.

The matrices Q and R in Theorem ?? are uniquely determined by A ; we return to this below.

Example 1. Find the QR-factorization of $A = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

Explanation. Denote the columns of A as $\mathbf{c}_1, \mathbf{c}_2$, and $\mathbf{c}_3$, and observe that $\{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ is independent. If we apply the Gram-Schmidt algorithm to these columns, the result is:

$$\mathbf{f}_1 = \mathbf{c}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{f}_2 = \mathbf{c}_2 - \frac{1}{2}\mathbf{f}_1 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{f}_3 = \mathbf{c}_3 + \frac{1}{2}\mathbf{f}_1 - \mathbf{f}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Write $\mathbf{q}_j = \frac{1}{\|\mathbf{f}_j\|}\mathbf{f}_j$ for each j , so $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ is orthonormal. Then equation (??) preceding Theorem ?? gives $A = QR$ where

$$Q = \begin{bmatrix} | & | & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{2}{\sqrt{6}} & 0 \\ 0 & \frac{1}{\sqrt{6}} & 1 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} & 1 & 0 \\ -\sqrt{3} & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & \sqrt{6} \end{bmatrix}$$

$$R = \begin{bmatrix} \|\mathbf{f}_1\| & \mathbf{c}_2 \cdot \mathbf{q}_1 & \mathbf{c}_3 \cdot \mathbf{q}_1 \\ 0 & \|\mathbf{f}_2\| & \mathbf{c}_3 \cdot \mathbf{q}_2 \\ 0 & 0 & \|\mathbf{f}_3\| \end{bmatrix} = \begin{bmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{3}}{\sqrt{2}} & \frac{\sqrt{3}}{\sqrt{2}} \\ 0 & 0 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 1 & -1 \\ 0 & \sqrt{3} & \sqrt{3} \\ 0 & 0 & \sqrt{2} \end{bmatrix}$$

The reader can verify that indeed $A = QR$.

If a matrix A has independent rows and we apply QR-factorization to A^T , the result is:

Corollary 1. If A has independent rows, then A factors uniquely as $A = LP$ where P has orthonormal rows and L is an invertible lower triangular matrix with positive main diagonal entries.

Since a square matrix with orthonormal columns is orthogonal, we have

Theorem 2. Every square invertible matrix A has factorizations $A = QR$ and $A = LP$ where Q and P are orthogonal, R is upper triangular with positive diagonal entries, and L is lower triangular with positive diagonal entries.

We now take the time to prove the uniqueness of the QR-factorization.

Theorem 3. Let A be an $m \times n$ matrix with independent columns. If $A = QR$ and $A = Q_1 R_1$ are QR-factorizations of A , then $Q_1 = Q$ and $R_1 = R$.

Proof Write $Q = \begin{bmatrix} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & & | \end{bmatrix}$ and $Q_1 = \begin{bmatrix} | & | & & | \\ \mathbf{d}_1 & \mathbf{d}_2 & \cdots & \mathbf{d}_n \\ | & | & & | \end{bmatrix}$ in terms of their columns, and observe

first that $Q^T Q = I_n = Q_1^T Q_1$ because Q and Q_1 have orthonormal columns. Hence it suffices to show that $Q_1 = Q$ (then $R_1 = Q_1^T A = Q^T A = R$). Since $Q_1^T Q_1 = I_n$, the equation $QR = Q_1 R_1$ gives $Q_1^T Q = R_1 R^{-1}$; for convenience we write this matrix as

$$Q_1^T Q = R_1 R^{-1} = \begin{bmatrix} t_{ij} \end{bmatrix}$$

This matrix is upper triangular with positive diagonal elements (since this is true for R and R_1), so $t_{ii} > 0$ for each i and $t_{ij} = 0$ if $i > j$. On the other hand, the (i, j) -entry of $Q_1^T Q$ is $\mathbf{d}_i^T \mathbf{c}_j = \mathbf{d}_i \cdot \mathbf{c}_j$, so we have $\mathbf{d}_i \cdot \mathbf{c}_j = t_{ij}$ for all i and j . But each $\mathbf{c}_j$ is in $\text{span}\{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_n\}$ because $Q = Q_1(R_1 R^{-1})$. We know how to write a vector as a linear combination of an orthonormal basis (using Corollary ?? from [Orthogonality and Projections](#)):

$$\mathbf{c}_j = (\mathbf{d}_1 \cdot \mathbf{c}_j)\mathbf{d}_1 + (\mathbf{d}_2 \cdot \mathbf{c}_j)\mathbf{d}_2 + \cdots + (\mathbf{d}_n \cdot \mathbf{c}_j)\mathbf{d}_n = t_{1j}\mathbf{d}_1 + t_{2j}\mathbf{d}_2 + \cdots + t_{nj}\mathbf{d}_n,$$

because $\mathbf{d}_i \cdot \mathbf{c}_j = t_{ij} = 0$ if $i > j$. The first few equations here are

$$\begin{aligned} \mathbf{c}_1 &= t_{11}\mathbf{d}_1 \\ \mathbf{c}_2 &= t_{12}\mathbf{d}_1 + t_{22}\mathbf{d}_2 \\ \mathbf{c}_3 &= t_{13}\mathbf{d}_1 + t_{23}\mathbf{d}_2 + t_{33}\mathbf{d}_3 \\ \mathbf{c}_4 &= t_{14}\mathbf{d}_1 + t_{24}\mathbf{d}_2 + t_{34}\mathbf{d}_3 + t_{44}\mathbf{d}_4 \\ &\vdots \end{aligned}$$

The first of these equations gives $1 = \|\mathbf{c}_1\| = \|t_{11}\mathbf{d}_1\| = |t_{11}| \|\mathbf{d}_1\| = t_{11}$, whence $\mathbf{c}_1 = \mathbf{d}_1$. But then we have $t_{12} = \mathbf{d}_1 \cdot \mathbf{c}_2 = \mathbf{c}_1 \cdot \mathbf{c}_2 = 0$, so the second equation becomes $\mathbf{c}_2 = t_{22}\mathbf{d}_2$. Now a similar argument gives $\mathbf{c}_2 = \mathbf{d}_2$, and then $t_{13} = 0$ and $t_{23} = 0$ follows in the same way. Hence $\mathbf{c}_3 = t_{33}\mathbf{d}_3$ and $\mathbf{c}_3 = \mathbf{d}_3$. Continue in this way to get $\mathbf{c}_i = \mathbf{d}_i$ for all i . This proves that $Q_1 = Q$. ■

QR-Algorithm for approximating eigenvalues

In The Power Method and the Dominant Eigenvalue, we learned about an iterative method for computing eigenvalues. We also mentioned that a better method for approximating the eigenvalues of an invertible matrix A depends on the QR-factorization of A . While it is beyond the scope of this book to pursue a detailed discussion of this method, we give an example and conclude with some remarks on the QR-algorithm. The interested reader is referred to J. M. Wilkinson, *The Algebraic Eigenvalue Problem* (Oxford, England: Oxford University Press, 1965) or G. W. Stewart, *Introduction to Matrix Computations* (New York: Academic Press, 1973).

The *QR-algorithm* uses QR-factorization repeatedly to create a sequence of matrices $A_1 = A, A_2, A_3, \dots$, as follows:

Algorithm 1. Let A be an invertible matrix.

Step 1: Define $A_1 = A$ and factor it as $A_1 = Q_1 R_1$.

Step 2: Define $A_2 = R_1 Q_1$ and factor it as $A_2 = Q_2 R_2$.

Step 3: Define $A_3 = R_2 Q_2$ and factor it as $A_3 = Q_3 R_3$.

$\vdots$

Step $k + 1$: Define $A_{k+1} = R_k Q_k$ (where $A_k = Q_k R_k$).

Note that A_{k+1} is similar to A_k (in fact, $A_{k+1} = R_k Q_k = (Q_k^{-1} A_k) Q_k$), and hence each A_k has the same eigenvalues as A . If the eigenvalues of A are real and have distinct absolute values, the sequence of matrices $A_1, A_2, A_3, \dots$ converges to an upper triangular matrix with these eigenvalues on the main diagonal. We will discuss the complex case eigenvalue case below.

Example 2. If $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$ as in Example ?? of The Power Method and the Dominant Eigenvalue, use the QR-algorithm to approximate the eigenvalues.

Explanation. The matrices A_1, A_2 , and A_3 are as follows:

$$A_1 = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} = Q_1 R_1 \quad \text{where } Q_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \text{ and } R_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 5 & 1 \\ 0 & 2 \end{bmatrix}$$

$$A_2 = \frac{1}{5} \begin{bmatrix} 7 & 9 \\ 4 & -2 \end{bmatrix} = \begin{bmatrix} 1.4 & -1.8 \\ -0.8 & -0.4 \end{bmatrix} = Q_2 R_2$$

$$\text{where } Q_2 = \frac{1}{\sqrt{65}} \begin{bmatrix} 7 & 4 \\ 4 & -7 \end{bmatrix} \text{ and } R_2 = \frac{1}{\sqrt{65}} \begin{bmatrix} 13 & 11 \\ 0 & 10 \end{bmatrix}$$

$$A_3 = \frac{1}{13} \begin{bmatrix} 27 & -5 \\ 8 & -14 \end{bmatrix} = \begin{bmatrix} 2.08 & -0.38 \\ 0.62 & -1.08 \end{bmatrix}$$

This is converging to $\begin{bmatrix} 2 & * \\ 0 & -1 \end{bmatrix}$ and so is approximating the eigenvalues 2 and -1 on the main diagonal.

Shifting.

We first learned about the concept of *shifting* in The Power Method and the Dominant Eigenvalue. Convergence is accelerated if, at stage k of the algorithm, a number τ_k is chosen and $A_k - \tau_k I$ is factored in the form $Q_k R_k$ rather than A_k itself. Then

$$Q_k^{-1} A_k Q_k = Q_k^{-1} (Q_k R_k + \tau_k I) Q_k = R_k Q_k + \tau_k I$$

so we take $A_{k+1} = R_k Q_k + \tau_k I$. If the shifts τ_k are carefully chosen, convergence can be greatly improved.

Preliminary Preparation.

A matrix such as

$$\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * \end{bmatrix}$$

is said to be in *upper Hessenberg* form, and the QR-factorizations of such matrices are greatly simplified. Given an $n \times n$ matrix A , a series of orthogonal matrices $H_1, H_2, \dots, H_m$ (called *Householder matrices*) can be easily constructed such that

$$B = H_m^T \cdots H_1^T A H_1 \cdots H_m$$

is in upper Hessenberg form. Then the QR-algorithm can be efficiently applied to B and, because B is similar to A , it produces the eigenvalues of A .

Complex Eigenvalues.

If some of the eigenvalues of a real matrix A are not real, the QR-algorithm converges to a block upper triangular matrix where the diagonal blocks are either 1×1 (the real eigenvalues) or 2×2 (each providing a pair of conjugate complex eigenvalues of A).

Practice Problems

Problem 1 In each case find the QR-factorization of A .

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

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

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

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

Click the arrow to see answer.

$$1. Q = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}, R = \frac{1}{\sqrt{5}} \begin{bmatrix} 5 & 3 \\ 0 & 1 \end{bmatrix}$$

$$2. Q = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix},$$

$$R = \frac{1}{\sqrt{3}} \begin{bmatrix} 3 & 0 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Problem 2 If R is upper triangular and invertible, show that there exists a diagonal matrix D with diagonal entries ± 1 such that $R_1 = DR$ is invertible, upper triangular, and has positive diagonal entries.

Problem 3 If A has independent columns, let

$A = QR$ where Q has orthonormal columns and R is invertible and upper triangular. (Some authors do not require a QR-factorization to have positive diagonal entries.) Show that there is a diagonal matrix D with diagonal entries ± 1 such that $A = (QD)(DR)$ is the QR-factorization of A .

Problem 4 In each case, find the exact eigenvalues and then approximate them using the QR-algorithm.

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

$$2. A = \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix}$$

Click the arrow to see answer.

$$1. \text{ Eigenvalues } \lambda_1 = \frac{1}{2}(3 + \sqrt{13}) = 3.302776, \lambda_2 = \frac{1}{2}(3 - \sqrt{13}) = -0.302776$$

$$A_1 = \begin{bmatrix} 3 & 1 \\ 1 & 0 \end{bmatrix}, Q_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 & -1 \\ 1 & 3 \end{bmatrix}, R_1 = \frac{1}{\sqrt{10}} \begin{bmatrix} 10 & 3 \\ 0 & -1 \end{bmatrix}$$

$$A_2 = \frac{1}{10} \begin{bmatrix} 33 & -1 \\ -1 & -3 \end{bmatrix},$$

$$Q_2 = \frac{1}{\sqrt{1090}} \begin{bmatrix} 33 & 1 \\ -1 & 33 \end{bmatrix},$$

$$R_2 = \frac{1}{\sqrt{1090}} \begin{bmatrix} 109 & -3 \\ 0 & -10 \end{bmatrix}$$

$$A_3 = \frac{1}{109} \begin{bmatrix} 360 & 1 \\ 1 & -33 \end{bmatrix}$$

$$= \begin{bmatrix} 3.302775 & 0.009174 \\ 0.009174 & -0.302775 \end{bmatrix}$$

Problem 5 If A is symmetric, show that each matrix A_k in the QR-algorithm is also symmetric. Deduce that they converge to a diagonal matrix.

Click the arrow to see answer. Use induction on k . If $k = 1$, $A_1 = A$. In general $A_{k+1} = Q_k^{-1} A_k Q_k = Q_k^T A_k Q_k$, so the fact that $A_k^T = A_k$ implies $A_{k+1}^T = A_{k+1}$. The eigenvalues of A are all real (Theorem ??), so the A_k converge to an upper triangular matrix T . But T must also be symmetric (it is the limit of symmetric matrices), so it is diagonal.

Problem 6 Apply the QR-algorithm to

$$A = \begin{bmatrix} 2 & -3 \\ 1 & -2 \end{bmatrix}. \text{ Explain.}$$

Problem 7 Given a matrix A , let A_k , Q_k , and R_k , $k \geq 1$, be the matrices constructed in the QR-algorithm. Show that $A_k = (Q_1 Q_2 \cdots Q_k)(R_k \cdots R_2 R_1)$ for each $k \geq 1$ and hence that this is a QR-factorization of A_k .

Text Source

This section was adapted from Sections 8.4 and 8.5 of Keith Nicholson's *Linear Algebra with Applications*. (CC-BY-NC-SA)

W. Keith Nicholson, *Linear Algebra with Applications*, Lyryx 2018, Open Edition, pp. 437–445.