Exercises for 8.3

Exercise 8.3.1 Find the Cholesky decomposition of each of the following matrices.

a. $\left[\begin{array}{ll} 4 & 3 \\ 3 & 5 \end{array}\right]$

b. $\left[\begin{array}{rr} 2 & -1 \\ -1 & 1 \end{array}\right]$

c. $\left[\begin{array}{rrr} 12 & 4 & 3 \\ 4 & 2 & -1 \\ 3 & -1 & 7 \end{array}\right]$

d. $\left[\begin{array}{lll} 20 & 4 & 5 \\ 4 & 2 & 3 \\ 5 & 3 & 5 \end{array}\right]$

Exercise 8.3.2

a. If A is positive definite, show that A^k is positive definite for all $k \geq 1$.

b. Prove the converse to (a) when k is odd.

c. Find a symmetric matrix A such that A^2 is positive definite but A is not.

Exercise 8.3.3 Let $A = \left[\begin{array}{ll} 1 & a \\ a & b \end{array}\right]$. If $a^2 < b$, show that A is positive definite and find the Cholesky factorization.

Exercise 8.3.4 If A and B are positive definite and $r > 0$, show that $A+B$ and rA are both positive definite.

Exercise 8.3.5 If A and B are positive definite, show that $\left[\begin{array}{ll} A & 0 \\ 0 & B \end{array}\right]$ is positive definite.

Exercise 8.3.6 If A is an $n \times n$ positive definite matrix and U is an $n \times m$ matrix of rank m, show that U^TAU is positive definite.

Exercise 8.3.7 If A is positive definite, show that each diagonal entry is positive.

Exercise 8.3.8 Let A_0 be formed from A by deleting rows 2 and 4 and deleting columns 2 and 4. If A is positive definite, show that A_0 is positive definite.

Exercise 8.3.9 If A is positive definite, show that $A = CC^T$ where C has orthogonal columns.

Exercise 8.3.10 If A is positive definite, show that $A = C^2$ where C is positive definite.

Exercise 8.3.11 Let A be a positive definite matrix. If a is a real number, show that aA is positive definite if and only if $a > 0$.

Exercise 8.3.12

a. Suppose an invertible matrix A can be factored in $\mathbf{M}_{nn}$ as $A = LDU$ where L is lower triangular with 1s on the diagonal, U is upper triangular with 1s on the diagonal, and D is diagonal with positive diagonal entries. Show that the factorization is unique: If $A = L_1D_1U_1$ is another such factorization, show that $L_1 = L$, $D_1 = D$, and $U_1 = U$.

b. Show that a matrix A is positive definite if and only if A is symmetric and admits a factorization $A = LDU$ as in (a).

Exercise 8.3.13 Let A be positive definite and write $d_r = \det {}^{(r)}A$ for each $r = 1, 2, \dots, n$. If U is the upper triangular matrix obtained in step 1 of the algorithm, show that the diagonal elements $u_{11}, u_{22}, \dots, u_{nn}$ of U are given by $u_{11} = d_1$, $u_{jj} = d_j/d_{j-1}$ if $j > 1$. [*Hint*: If $LA = U$ where L is lower triangular with 1s on the diagonal, use block multiplication to show that $\det {}^{(r)}A = \det {}^{(r)}U$ for each r.]

8.4 QR-Factorization[7]

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).

[7]This section is not used elsewhere in the book

Definition 8.6 QR-factorization

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 $A = \left[\begin{array}{cccc} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \end{array} \right]$ 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 $\mathbf{q}_i$:

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

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

$$\begin{aligned}
A &= \left[\begin{array}{ccccc} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 & \cdots & \mathbf{c}_n \end{array} \right] \\
&= \left[\begin{array}{ccccc} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 & \cdots & \mathbf{q}_n \end{array} \right] \left[\begin{array}{ccccc}
\|\mathbf{f}_1\| & \mathbf{c}_2 \cdot \mathbf{q}_1 & \mathbf{c}_3 \cdot \mathbf{q}_1 & \cdots & \mathbf{c}_n \cdot \mathbf{q}_1 \\
0 & \|\mathbf{f}_2\| & \mathbf{c}_3 \cdot \mathbf{q}_2 & \cdots & \mathbf{c}_n \cdot \mathbf{q}_2 \\
0 & 0 & \|\mathbf{f}_3\| & \cdots & \mathbf{c}_n \cdot \mathbf{q}_3 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \|\mathbf{f}_n\|
\end{array} \right]
\end{aligned} \tag{8.5}$$

Here the first factor $Q = \left[\begin{array}{ccccc} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 & \cdots & \mathbf{q}_n \end{array} \right]$ 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 8.4.1: 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 8.4.1 are uniquely determined by A; we return to this below.

Example 8.4.1

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

Solution. 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:

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

Write $\mathbf{q}_j = \frac{1}{\|\mathbf{f}_j\|^2}\mathbf{f}_j$ for each j, so $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ is orthonormal. Then equation (8.5) preceding Theorem 8.4.1 gives $A = QR$ where

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

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

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 8.4.1

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 8.4.2

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.

Remark

In Section 5.6 we found how to find a best approximation $\mathbf{z}$ to a solution of a (possibly inconsistent) system $A\mathbf{x} = \mathbf{b}$ of linear equations: take $\mathbf{z}$ to be any solution of the "normal" equations $(A^TA)\mathbf{z} = A^T\mathbf{b}$. If A has independent columns this $\mathbf{z}$ is unique (A^TA is invertible by Theorem 5.4.3), so it is often desirable to compute $(A^TA)^{-1}$. This is particularly useful in least squares approximation (Section 5.6). This is simplified if we have a QR-factorization of A (and is one of the main reasons for the importance of Theorem 8.4.1). For if $A = QR$ is such a factorization, then $Q^TQ = I_n$ because Q has orthonormal columns (verify), so we obtain

$$A^TA = R^TQ^TQR = R^TR$$

Hence computing $(A^TA)^{-1}$ amounts to finding R^{-1}, and this is a routine matter because R is upper triangular. Thus the difficulty in computing $(A^TA)^{-1}$ lies in obtaining the QR-factorization of A.

We conclude by proving the uniqueness of the QR-factorization.

Theorem 8.4.3

Let A be an $m \times n$ matrix with independent columns. If $A = QR$ and $A = Q_1R_1$ 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 $Q^TQ = I_n = Q_1^TQ_1$ because Q and Q_1 have orthonormal columns. Hence it suffices to show that $Q_1 = Q$ (then $R_1 = Q_1^TA = Q^TA = R$). Since $Q_1^TQ_1 = I_n$, the equation $QR = Q_1R_1$ gives $Q_1^TQ = R_1R^{-1}$; for convenience we write this matrix as

$$Q_1^TQ = R_1R^{-1} = \begin{bmatrix} t_{ij} \end{bmatrix}$$

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^TQ 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 $Q = Q_1(R_1R^{-1})$. Hence the expansion theorem gives

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

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

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

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 means that $Q_1 = Q$, which is what we wanted. $\square$

Exercises for 8.4

Exercise 8.4.1 In each case find the QR-factorization of A.

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

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

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

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

Exercise 8.4.2 Let A and B denote matrices.

a. If A and B have independent columns, show that AB has independent columns. [*Hint*: Theorem 5.4.3.]

b. Show that A has a QR-factorization if and only if A has independent columns.

c. If AB has a QR-factorization, show that the same is true of B but not necessarily A.

[*Hint*: Consider AA^T where $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$.]

Exercise 8.4.3 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.

Exercise 8.4.4 If A has independent columns, let $A = QR$ where Q has orthonormal columns and R is invertible and upper triangular. [Some authors call *this* a QR-factorization of A.] Show that there is a diagonal matrix D with diagonal entries ± 1 such that $A = (QD)(DR)$ is the QR-factorization of A. [*Hint*: Preceding exercise.]

8.5 Computing Eigenvalues

In practice, the problem of finding eigenvalues of a matrix is virtually never solved by finding the roots of the characteristic polynomial. This is difficult for large matrices and iterative methods are much better. Two such methods are described briefly in this section.

The Power Method

In Chapter 3 our initial rationale for diagonalizing matrices was to be able to compute the powers of a square matrix, and the eigenvalues were needed to do this. In this section, we are interested in efficiently computing eigenvalues, and it may come as no surprise that the first method we discuss uses the powers of a matrix.

Recall that an eigenvalue λ of an $n \times n$ matrix A is called a **dominant eigenvalue** if λ has multiplicity 1, and

$$|\lambda| > |\mu| \quad \text{for all eigenvalues } \mu \neq \lambda$$

Any corresponding eigenvector is called a **dominant eigenvector** of A. When such an eigenvalue exists, one technique for finding it is as follows: Let $\mathbf{x}_0$ in $\mathbb{R}^n$ be a first approximation to a dominant eigenvector λ, and compute successive approximations $\mathbf{x}_1$, $\mathbf{x}_2$, $\dots$ as follows:

$$\mathbf{x}_1 = A\mathbf{x}_0 \quad \mathbf{x}_2 = A\mathbf{x}_1 \quad \mathbf{x}_3 = A\mathbf{x}_2 \quad \cdots$$