5 Orthogonal Vectors and Matrices

Orthogonal vectors and matrices are of fundamental importance in linear algebra and scientific computing. Orthogonal matrices are used in QR factorization and singular value decomposition (SVD) of a matrix. The former is applied in numerical methods for least-squares approximation and eigenvalue computations, the latter is an important tool for data reduction as well as for least-squares approximation. This lecture first considers orthogonal vectors and then defines orthogonal matrices. Applications will be discussed in subsequent lectures.

5.1 Orthogonal Vectors

A pair of vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$ is said to be *orthogonal* if

$$(\mathbf{u}, \mathbf{v}) = 0.$$

In view of formula (14) of Lecture 1, orthogonal vectors meet at a right angle. The zero-vector $\mathbf{0}$ is orthogonal to all vector, but we are more interested in nonvanishing orthogonal vectors.

The vectors in a set $\mathbb{S}_n = \{\mathbf{v}_j\}_{j=1}^n$ in $\mathbb{R}^m$ are said to be *orthonormal* if each pair of distinct vectors in $\mathbb{S}_n$ is orthogonal and all vectors in $\mathbb{S}_n$ are of unit length, i.e., if

$$(\mathbf{v}_j, \mathbf{v}_k) = \begin{cases} 0, & j \neq k, \\ 1, & j = k. \end{cases} \tag{1}$$

Here we have used that

$$(\mathbf{v}_k, \mathbf{v}_k) = \mathbf{v}_k{}^T \mathbf{v}_k = \|\mathbf{v}_k\|^2;$$

cf. equations (6) and (7) of Lecture 1. We assume throughout this section that $n \leq m$,

It is not difficult to show that orthonormal vectors are linearly independent; see Exercise 5.1 below. It follows that the m orthonormal vectors in the set $\mathbb{S}_m = \{\mathbf{v}_j\}_{j=1}^m$ form a basis for $\mathbb{R}^m$.

Example 5.1

The vectors in the subset $\mathbb{S}_3 = \{\mathbf{e}_j\}_{j=1}^3$ of $\mathbb{R}^5$ are orthonormal. Here $\mathbf{e}_j$ denotes the jth axis vector of $\mathbb{R}^m$; cf. (18) of Lecture 1. $\Box$

Example 5.2

The vectors in the set $\mathbb{S}_2 = \{\mathbf{v}_1, \mathbf{v}_2\}$ in $\mathbb{R}^2$, defined by

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}}[1, 1]^T, \qquad \mathbf{v}_2 = \frac{1}{\sqrt{2}}[-1, 1]^T,$$

[0]Version October 2, 2013

are orthonormal. Moreover, these vectors form a basis for $\mathbb{R}^2$. $\square$

An arbitrary vector $\mathbf{v} \in \mathbb{R}^m$ can be decomposed into *orthogonal components*. Consider the set of orthonormal vectors $\mathbb{S}_n = \{\mathbf{v}_j\}_{j=1}^n$ in $\mathbb{R}^m$ and regard the expression

$$\mathbf{r} = \mathbf{v} - \sum_{j=1}^{n} (\mathbf{v}_j, \mathbf{v})\mathbf{v}_j. \tag{2}$$

The vector $(\mathbf{v}_j, \mathbf{v})\mathbf{v}_j$ is referred to as the orthogonal component of $\mathbf{v}$ in the direction $\mathbf{v}_j$. Moreover, the vector $\mathbf{r}$ is orthogonal to the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. This can be seen by computing the inner products $(\mathbf{v}_k, \mathbf{r})$ for all k. We obtain

$$(\mathbf{v}_k, \mathbf{r}) = (\mathbf{v}_k, \mathbf{v} - \sum_{j=1}^{n} (\mathbf{v}_j, \mathbf{v})\mathbf{v}_j) = (\mathbf{v}_k, \mathbf{v}) - \sum_{j=1}^{n} (\mathbf{v}_j, \mathbf{v})(\mathbf{v}_k, \mathbf{v}_j).$$

Using (1), the sum in the right-hand side simplifies to

$$\sum_{j=1}^{n} (\mathbf{v}_j, \mathbf{v})(\mathbf{v}_k, \mathbf{v}_j) = (\mathbf{v}_k, \mathbf{v}),$$

which shows that

$$(\mathbf{v}_k, \mathbf{r}) = 0$$

Thus, $\mathbf{v}$ can be expressed as a sum of the orthogonal vectors $\mathbf{r}, \mathbf{v}_1, \ldots, \mathbf{v}_n$,

$$\mathbf{v} = \mathbf{r} + \sum_{j=1}^{n} (\mathbf{v}_j, \mathbf{v})\mathbf{v}_j.$$

Example 5.3

Let $\mathbf{v} = [1, 2, 3, 4, 5]^T$ and let the set $\mathbb{S}$ be the same as in Example 5.1. Then $\mathbf{r} = [0, 0, 0, 4, 5]^T$. Clearly, $\mathbf{r}$ is orthogonal to the axis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. $\square$

Exercise 5.1

Let $\mathbb{S}_n = \{\mathbf{v}_j\}_{j=1}^n$ be a set of orthonormal vectors in $\mathbb{R}^m$. Show that the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent. Hint: Assume this is not the case. For instance, assume that $\mathbf{v}_1$ is a linear combination of the vectors $\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n$, and apply (1). $\square$

Let the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ in $\mathbb{R}^m$, $m \geq n$, be linearly independent. The Gram-Schmidt procedure repeatedly uses the decomposition (2) to determine a set of orthonormal vectors $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ such that

$$\text{span}\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\} = \text{span}\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}.$$

In other words, the Gram-Schmidt procedure determines an orthonormal basis for $\text{span}\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$.

Gram-Schmidt procedure

for $j = 1, 2, \ldots, n$
  $\mathbf{v} := \mathbf{a}_j$
  for $i = 1, 2, \ldots, j-1$
    $r_{i,j} := (\mathbf{q}_i, \mathbf{a}_j)$
  end
  for $i = 1, 2, \ldots, j-1$
    $\mathbf{v} := \mathbf{v} - r_{i,j}\mathbf{q}_i$
  end
  $r_{j,j} := \|\mathbf{v}\|$
  $\mathbf{q}_j := \mathbf{v}/r_{j,j}$
end □

A matrix interpretation of these recursion formulas shows that they determine a factorization of the matrix $A = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n]$. Note that the coefficients $r_{i,j}$ only are defined for $i \leq j$. They determine the upper triangular matrix

$$R = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,n} \\ & r_{2,2} & \cdots & r_{2,n} \\ & & \ddots & \vdots \\ & & & r_{n,n} \end{bmatrix} \tag{3}$$

Define the matrix

$$Q = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n] \in \mathbb{R}^{m \times n}$$

with orthonormal columns. Then the recursion formulas of the Gram-Schmidt procedure show that

$$A = QR. \tag{4}$$

This factorization is known as a *QR factorization*. We will use it for solving least-squares problems in Lecture 6. QR factorization is a useful analogue of the LU factorization of Lecture 3 for matrices A with more rows than columns.

Exercise 5.2

Let $A = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3] \in \mathbb{R}^{5 \times 3}$. Determine the orthonormal columns $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ and scalars $r_{i,j}$ for $1 \leq i \leq j \leq 3$ by the Gram-Schmidt procedure. Show that the matrix $Q = [\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3] \in \mathbb{R}^{5 \times 3}$ and the upper triangular matrix $R = [r_{i,j}]$ determine a QR factorization (4) of A. Hint: Start a matrix A with one column and then add columns, one at a time. □

Exercise 5.3

Is the QR factorization (4) unique? That is, are there other factorizations of an $m \times n$ matrix A with $m \geq n$ of the form (4) with a matrix Q with orthonormal columns and an upper triangular matrix R? $\square$

The Gram-Schmidt procedure described above is sensitive to round-off errors introduced during the computation of the columns $\mathbf{q}_j$, with the result that the computed columns may be far from orthogonal when n (and therefore also m) are large. A rearrangement of the order of the computations results in a numerically better behaved method. The following algorithm describes the latter.

Modified Gram-Schmidt procedure

for $i = 1, 2, \ldots, n$
  $\mathbf{v}_i := \mathbf{a}_i$
for $i = 1, 2, \ldots, n$
  $r_{i,i} := \|\mathbf{v}_i\|$
  $\mathbf{q}_i := \mathbf{v}_i / r_{i,i}$
  for $j = i+1, i+2, \ldots, n$
    $r_{i,j} := (\mathbf{q}_i, \mathbf{v}_j)$
    $\mathbf{v}_j := \mathbf{v}_j - r_{i,j}\mathbf{q}_i$
  end
end $\square$

In exact arithmetic the Gram-Schmidt and Modified Gram-Schmidt procedures yield the same output. Thus, the coefficients $r_{i,j}$ and columns $\mathbf{q}_i$ computed by the Modified Gram-Schmidt procedure also determine a QR factorization (4).

Exercise 5.4

Compute a QR factorization of the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix}$$

using the modified Gram-Schmidt procedure. $\square$

Exercise 5.5

Let the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ in the modified Gram-Schmidt procedure be linearly indepenedent and in $\mathbb{R}^m$ with $m \geq n$. How many arithmetic floating point operations does the procedure require? Only determine the dominant term of the form $cm^\alpha n^\beta$ in the flop count, i.e., determine the coefficient c and the powers α and β. $\square$

5.2 Orthogonal Matrices

A square matrix $Q = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m] \in \mathbb{R}^{m\times m}$ is said to be *orthogonal* if its columns $\{\mathbf{q}_j\}_{j=1}^m$ form an orthonormal basis for $\mathbb{R}^m$. Since the columns $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m$ are linearly independent, cf. Exercise 5.1, the matrix Q is nonsingular. Thus, Q has an inverse, which we denote by Q^{-1}. It follows from the orthonormality of the columns of Q that

$$Q^T Q = \begin{bmatrix} (\mathbf{q}_1, \mathbf{q}_1) & (\mathbf{q}_1, \mathbf{q}_2) & \cdots & (\mathbf{q}_1, \mathbf{q}_m) \\ (\mathbf{q}_2, \mathbf{q}_1) & (\mathbf{q}_2, \mathbf{q}_2) & \cdots & (\mathbf{q}_2, \mathbf{q}_m) \\ \vdots & \vdots & & \vdots \\ (\mathbf{q}_m, \mathbf{q}_1) & (\mathbf{q}_m, \mathbf{q}_2) & \cdots & (\mathbf{q}_m, \mathbf{q}_m) \end{bmatrix} = I,$$

where I denotes the identity matrix. Multiplying the above expression by the inverse Q^{-1} from the right-hand side shows that

$$Q^T = Q^{-1}.$$

Thus, the transpose of an orthogonal matrix is the inverse.

Example 5.4

The identity matrix I is orthogonal. $\square$

Example 5.5

The matrix

$$Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

is orthogonal. Its inverse is its transpose,

$$Q^{-1} = Q^T = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

$\square$

Geometrically, multiplying a vector by an orthogonal matrix reflects the vector in some plane and/or rotates it. Therefore, multiplying a vector by an orthogonal matrices does not change its length. That is, the Euclidean norm of a vector $\mathbf{u}$ is invariant under multiplication by an orthogonal matrix Q:

$$\|Q\mathbf{u}\| = \|\mathbf{u}\|. \tag{5}$$

This can be shown by using the properties (17) and (19) of Lecture 1. We have

$$\|Q\mathbf{u}\|^2 = (Q\mathbf{u})^T(Q\mathbf{u}) = \mathbf{u}^T Q^T(Q\mathbf{u}) = \mathbf{u}^T(Q^TQ)\mathbf{u} = \mathbf{u}^T\mathbf{u} = \|\mathbf{u}\|^2.$$

Taking square-roots of the right-hand side and left-hand side, and using that $\|\cdot\|$ is nonnegative gives (5).

We remark thar the matrix Q in the QR factorization (4) is an orthogonal matrix only if $m = n$.

Exercise 5.6

What is the determinant of a real orthogonal matrix? Hint: Let A and B be square matrices of the same size. Recall that $\det(AB) = \det(A)\det(B)$ and $\det(A) = \det(A^T)$. □

5.3 Householder Matrices

Matrices of the form

$$H = I - \rho\,\mathbf{u}\mathbf{u}^T \in \mathbb{R}^{m\times m}, \qquad \mathbf{u} \in \mathbb{R}^m\backslash\{\mathbf{0}\}, \qquad \rho = \frac{2}{\mathbf{u}^T\mathbf{u}}, \tag{6}$$

are known as *Householder matrices*. They are used in numerical methods for least-squares approximation and eigenvalue computations. We will discuss the former application in the next section.

Householder matrices are symmetric, i.e., $H = H^T$, and orthogonal. The latter property follows from

$$\begin{aligned} H^TH &= H^2 = (I - \rho\,\mathbf{u}\mathbf{u}^T)(I - \rho\,\mathbf{u}\mathbf{u}^T) = I - \rho\,\mathbf{u}\mathbf{u}^T - \rho\,\mathbf{u}\mathbf{u}^T + (\rho\,\mathbf{u}\mathbf{u}^T)(\rho\,\mathbf{u}\mathbf{u}^T) \\ &= I - \rho\,\mathbf{u}\mathbf{u}^T - \rho\,\mathbf{u}\mathbf{u}^T + \rho\,\mathbf{u}(\rho\,\mathbf{u}^T\mathbf{u})\mathbf{u}^T = I - \rho\,\mathbf{u}\mathbf{u}^T - \rho\,\mathbf{u}\mathbf{u}^T + 2\rho\,\mathbf{u}\mathbf{u}^T = I, \end{aligned}$$

where we have used the fact that $\rho\,\mathbf{u}^T\mathbf{u} = 2$; cf. (6).

Our interest in Householder matrices stems from the facts that they are orthogonal and the vector $\mathbf{u}$ in their definition can be chosen so that an arbitrary (but fixed) vector $\mathbf{w} \in \mathbb{R}^m\backslash\{\mathbf{0}\}$ is mapped by H onto a multiple of the axis vector $\mathbf{e}_1$. We will now show how this can be done. Let $\mathbf{w} \neq \mathbf{0}$ be given. We would like

$$H\mathbf{w} = \sigma\,\mathbf{e}_1 \tag{7}$$

for some scalar σ. It follows from (5) that

$$\|\mathbf{w}\| = \|H\mathbf{w}\| = \|\sigma\,\mathbf{e}_1\| = |\sigma|\|\mathbf{e}_1\| = |\sigma|.$$

Therefore,

$$\sigma = \pm\|\mathbf{w}\|. \tag{8}$$

Moreover, using the definition (6) of H, we obtain

$$\sigma\,\mathbf{e}_1 = H\mathbf{w} = (I - \rho\,\mathbf{u}\mathbf{u}^T)\mathbf{w} = \mathbf{w} - \tau\,\mathbf{u}, \qquad \tau = \rho\,\mathbf{u}^T\mathbf{w},$$

from which it follows that

$$\tau\,\mathbf{u} = \mathbf{w} - \sigma\,\mathbf{e}_1.$$

The matrix H is independent of the scaling factor τ in the sense that the entries of the matrix H do not change if we replace $\tau\,\mathbf{u}$ by $\mathbf{u}$. This is a consequence of the definition (6). We therefore may choose

$$\mathbf{u} = \mathbf{w} - \sigma\,\mathbf{e}_1. \tag{9}$$

This choice of $\mathbf{u}$ and either one of the choices (8) of σ give a Householder matrix that satisfies (7). Nevertheless, in finite precision arithmetic, the choice of sign in (8) may be important. To see this, let $\mathbf{w} = [w_1, w_2, \ldots, w_m]^T$ and write the vector (9) in the form

$$\mathbf{u} = \begin{bmatrix} w_1 - \sigma \\ w_2 \\ w_3 \\ \vdots \\ w_m \end{bmatrix}.$$

If the components w_j for $j \geq 2$ are of small magnitude compared to w_1, then $\|\mathbf{w}\| \approx |w_1|$ and, therefore, the first component of $\mathbf{u}$ satisfies

$$u_1 = w_1 - \sigma = w_1 \pm \|\mathbf{w}\| \approx w_1 \pm |w_1|. \tag{10}$$

We would like to avoid the situation that $|w_1|$ is large and $|u_1|$ is small, because then u_1 is determined with low relative accuracy; see Exercise 5.8 below. We therefore let

$$\sigma = -\mathrm{sign}(w_1)\|\mathbf{w}\|, \tag{11}$$

which yields

$$u_1 = w_1 + \mathrm{sign}(w_1)\|\mathbf{w}\|. \tag{12}$$

Then u_1 is computed by adding numbers of the same sign and cancellation of significant digits is avoided.

Example 5.6

Let $\mathbf{w} = [1, 1, 1, 1]^T$. We are interested in determining the Householder matrix H that maps w onto a multiple of $\mathbf{e}_1$. The parameter σ in (7) is chosen so that $|\sigma| = \|\mathbf{w}\| = 2$, i.e., σ is 2 or -2. The vector $\mathbf{u}$ in (6) is given by

$$u = w \pm \sigma\,\mathbf{e}_1 = \begin{bmatrix} 1 \pm \sigma \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

The choice $\sigma = 2$ yields $u = [3 \;\; 1 \;\; 1 \;\; 1]^T$. We let σ be positive, because the first entry of $\mathbf{w}$ is positive. Then we obtain $\rho = 2/\|\mathbf{u}\|^2 = 1/6$ and

$$H = I - \rho\, uu^T = \begin{bmatrix} -1/2 & -1/2 & -1/2 & -1/2 \\ -1/2 & 5/6 & -1/6 & -1/6 \\ -1/2 & -1/6 & 5/6 & -1/6 \\ -1/2 & -1/6 & -1/6 & 5/6 \end{bmatrix}.$$

It is easy to verify that H is orthogonal and maps $\mathbf{w}$ to $-2\mathbf{e}_1$. □

In most applications it is not necessary to explicitly form the Householder matrix H. For instance, the products of a Householder matrix H with a vector $\mathbf{v}$ can be evaluated efficiently using the definition (6) of H, i.e.,

$$H\mathbf{v} = (I - \rho\, \mathbf{u}\mathbf{u}^T)\mathbf{v} = \mathbf{v} - (\rho\, \mathbf{u}^T\mathbf{v})\mathbf{u}.$$

The left-hand side is computed by first evaluating the scalar $\tau = \rho\, \mathbf{u}^T\mathbf{v}$ and then computing the vector scaling and addition $\mathbf{v} - \tau\, \mathbf{u}$. This way of evaluating $H\mathbf{w}$ requires fewer arithmetic floating-point operations than straightforward computation of the matrix-vector product using the entries of H; see Exercise 5.11. Moreover, the entries of H do not have to be stored, only the vector $\mathbf{u}$ and scalar ρ. The savings in arithmetic operations and storage is important for large-scale problems.

Exercise 5.7

Let $\mathbf{w} = [1, 2, 3]^T$. Determine the Householder matrix that maps `w` to a multiple of $\mathbf{e}_1$. Only the vector $\mathbf{u}$ in (6) has to be computed. □

Exercise 5.8

This exercise illustrates the importance of the choice of the sign of σ in (8). Let $\mathbf{w} = [1, 0.5 \cdot 10^{-8}]^T$ and let $\mathbf{u}$ be the vector in the definition of Householder matrices (6), chosen so that $H\mathbf{w} = \sigma\mathbf{e}_1$. MATLAB yields

```
>> w=[1; 5e-9]

w =

    1.0000
    0.0000

>> sigma=norm(w)

sigma =
```

```
     1

>> u1=w(1)+sigma

u1 =

     2

>> u1=w(1)-sigma

u1 =

     0
```

where u1 denotes the computed approximations of the first component, u_1, of the vectors $\mathbf{u}$. How large are the absolute and relative errors in the computed approximations u1 of the component u_1 of the vectors $\mathbf{u}$? □

Exercise 5.9

Show that the product U_1U_2 of two orthogonal matrices U_1 and U_2 is an orthogonal matrix. Is the product of $k > 2$ orthogonal matrices an orthogonal matrix? □

Exercise 5.10

Let Q be an orthogonal matrix, i.e., $Q^TQ = I$. Show that $QQ^T = I$. Hint: First show that Q is nonsingular. □

Exercise 5.11

What is the count of arithmetic floating point operations for evaluating a matrix-vector product with an $n \times n$ Householder matrix H when the representation (6) of H is used? Only $\mathbf{u}$ and ρ are stored, not the entries of H. What is the count of arithmetic floating point operations for evaluating a matrix-vector product with H when the entries of H (but not $\mathbf{u}$ and ρ) are available? Correct orders of magnitude of the arithmetic work when n is large suffices. □

Exercise 5.12

Let $\mathbf{w}_1$ and $\mathbf{w}_2$ be nonvanishing m-vectors. Determine an orthogonal matrix $\widetilde{H}$, such that $\mathbf{w}_2 = \widetilde{H}\mathbf{w}_1$. Hint: Use two Householder matrices. □

Exercise 5.13

Let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}.$$

Determine an orthonormal basis of $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Which method should be used? □

Exercise 5.14

The 2×2 orthogonal matrix Q in Example 5.5 is an example of a Givens rotation. General 2×2 Givens rotations are matrices of the form

$$G = \begin{bmatrix} c & s \\ -s & c \end{bmatrix},$$

where $s = \sin(\theta)$ and $c = \cos(\theta)$ for some "angle" $\theta \in \mathbb{R}$. The matrix Q in Example 5.5 corresponds to $\theta = \pi/4$. Show that G is orthogonal for an arbitrary $\theta \in \mathbb{R}$. The matrix G is referred to as a Givens rotation, because the matrix-vector product $G\mathbf{v}$ can be interpreted as a rotaion of the vector $\mathbf{v}$ by the angle θ. □

Exercise 5.15

Givens rotations are convenient to use for computing the QR factorization of an almost upper triangular matrix. Consider the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ & 5 & 6 \end{bmatrix}.$$

Multiply A by the matrix

$$G_1 = \begin{bmatrix} c & s & \\ -s & c & \\ & & 1 \end{bmatrix}$$

from the left with θ chosen so that the $(2,1)$-entry of $G_1 A$ vanishes. Then multiply the matrix obtained from the left by

$$G_2 = \begin{bmatrix} 1 & & \\ & c & s \\ & -s & c \end{bmatrix}$$

with θ chosen so that the $(3,2)$-entry of G_2G_1A vanishes. Then $R = G_2G_1A$ is upper triangular. Multiplying R from the left by G_2^T and G_1^T yields

$$A = G_1^T G_2^T R. \tag{13}$$

The matrices G_1^T and G_2^T are orthogonal. Therefore, by Exercise 5.9, the matrix $Q = G_1^T G_2^T$ is orthogonal and it follows that the right-hand side of (13) is a QR factorization. Determine the matrices G_1, G_2, and R. $\Box$