Chapter 12

QR-Decomposition for Arbitrary Matrices

12.1 Orthogonal Reflections

Hyperplane reflections are represented by matrices called Householder matrices. These matrices play an important role in numerical methods, for instance for solving systems of linear equations, solving least squares problems, for computing eigenvalues, and for transforming a symmetric matrix into a tridiagonal matrix. We prove a simple geometric lemma that immediately yields the QR-decomposition of arbitrary matrices in terms of Householder matrices.

Orthogonal symmetries are a very important example of isometries. First let us review the definition of projections, introduced in Section 5.2, just after Proposition 5.5. Given a vector space E, let F and G be subspaces of E that form a direct sum $E = F \oplus G$. Since every $u \in E$ can be written uniquely as $u = v + w$, where $v \in F$ and $w \in G$, we can define the two *projections* $p_F\colon E \to F$ and $p_G\colon E \to G$ such that $p_F(u) = v$ and $p_G(u) = w$. In Section 5.2 we used the notation π_1 and π_2, but in this section it is more convenient to use p_F and p_G.

It is immediately verified that p_G and p_F are linear maps, and that

$$p_F^2 = p_F,\ p_G^2 = p_G,\ p_F \circ p_G = p_G \circ p_F = 0, \quad \text{and} \quad p_F + p_G = \mathrm{id}.$$

.

Definition 12.1. Given a vector space E, for any two subspaces F and G that form a direct sum $E = F \oplus G$, the *symmetry (or reflection) with respect to F and parallel to G* is the linear map $s\colon E \to E$ defined such that

$$s(u) = 2p_F(u) - u,$$

for every $u \in E$.

Because $p_F + p_G = \mathrm{id}$, note that we also have

$$s(u) = p_F(u) - p_G(u)$$

and

$$s(u) = u - 2p_G(u),$$

$s^2 = \mathrm{id}$, s is the identity on F, and $s = -\mathrm{id}$ on G.

We now assume that E is a Euclidean space of *finite* dimension.

Definition 12.2. Let E be a Euclidean space of finite dimension n. For any two subspaces F and G, if F and G form a direct sum $E = F \oplus G$ and F and G are orthogonal, i.e., $F = G^{\perp}$, the *orthogonal symmetry (or reflection) with respect to F and parallel to G* is the linear map $s\colon E \to E$ defined such that

$$s(u) = 2p_F(u) - u = p_F(u) - p_G(u),$$

for every $u \in E$. When F is a hyperplane, we call s a *hyperplane symmetry with respect to F (or reflection about F)*, and when G is a plane (and thus $\dim(F) = n - 2$), we call s a *flip about F*.

A reflection about a hyperplane F is shown in Figure 12.1.

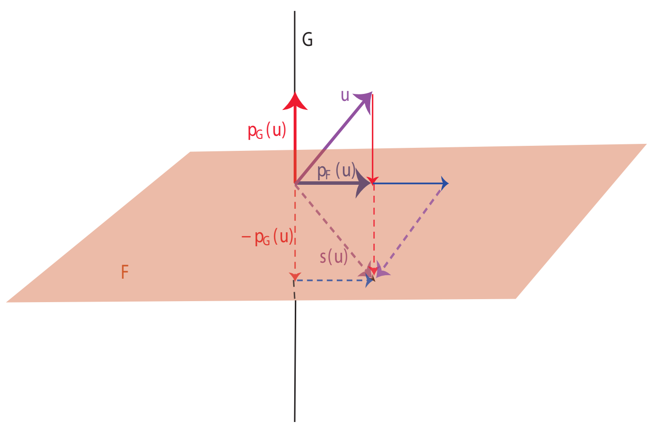

Figure 12.1: A reflection about the peach hyperplane F. Note that u is purple, $p_F(u)$ is blue and $p_G(u)$ is red.

For any two vectors $u, v \in E$, it is easily verified using the bilinearity of the inner product that

$$\|u + v\|^2 - \|u - v\|^2 = 4(u \cdot v). \tag{$*$}$$

In particular, if $u \cdot v = 0$, then $\|u + v\| = \|u - v\|$. Then since

$$u = p_F(u) + p_G(u)$$

and

$$s(u) = p_F(u) - p_G(u),$$

and since F and G are orthogonal, it follows that

$$p_F(u) \cdot p_G(v) = 0,$$

and thus by $(*)$

$$\|s(u)\| = \|p_F(u) - p_G(u)\| = \|p_F(u) + p_G(u)\| = \|u\|,$$

so that *s is an isometry.*

Using Proposition 11.10, it is possible to find an orthonormal basis $(e_1, \ldots, e_n)$ of E consisting of an orthonormal basis of F and an orthonormal basis of G. Assume that F has dimension p, so that G has dimension $n - p$. With respect to the orthonormal basis $(e_1, \ldots, e_n)$, the symmetry s has a matrix of the form

$$\begin{pmatrix} I_p & 0 \\ 0 & -I_{n-p} \end{pmatrix}.$$

Thus, $\det(s) = (-1)^{n-p}$, and s is a rotation iff $n - p$ is even. In particular, when F is a hyperplane H, we have $p = n - 1$ and $n - p = 1$, so that s is an improper orthogonal transformation. When $F = \{0\}$, we have $s = -\mathrm{id}$, which is called the *symmetry with respect to the origin*. The symmetry with respect to the origin is a rotation iff n is even, and an improper orthogonal transformation iff n is odd. When n is odd, since $s \circ s = \mathrm{id}$ and $\det(s) = (-1)^n = -1$, we observe that every improper orthogonal transformation f is the composition $f = (f \circ s) \circ s$ of the rotation $f \circ s$ with s, the symmetry with respect to the origin. When G is a plane, $p = n - 2$, and $\det(s) = (-1)^2 = 1$, so that a flip about F is a rotation. In particular, when $n = 3$, F is a line, and a flip about the line F is indeed a rotation of measure π as illustrated by Figure 12.2.

Remark: Given any two orthogonal subspaces F, G forming a direct sum $E = F \oplus G$, let f be the symmetry with respect to F and parallel to G, and let g be the symmetry with respect to G and parallel to F. We leave as an exercise to show that

$$f \circ g = g \circ f = -\mathrm{id}.$$

When $F = H$ is a hyperplane, we can give an explicit formula for $s(u)$ in terms of any nonnull vector w orthogonal to H. Indeed, from

$$u = p_H(u) + p_G(u),$$

since $p_G(u) \in G$ and G is spanned by w, which is orthogonal to H, we have

$$p_G(u) = \lambda w$$

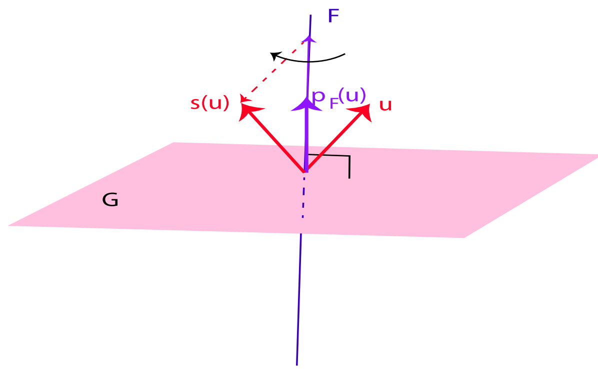

Figure 12.2: A flip in $\mathbb{R}^3$ is a rotation of π about the F axis.

for some $\lambda \in \mathbb{R}$, and we get

$$u \cdot w = \lambda \|w\|^2,$$

and thus

$$p_G(u) = \frac{(u \cdot w)}{\|w\|^2}\, w.$$

Since

$$s(u) = u - 2p_G(u),$$

we get

$$s(u) = u - 2\, \frac{(u \cdot w)}{\|w\|^2}\, w.$$

Since the above formula is important, we record it in the following proposition.

Proposition 12.1. *Let E be a finite-dimensional Euclidean space and let H be a hyperplane in E. For any nonzero vector w orthogonal to H, the hyperplane reflection s about H is given by*

$$s(u) = u - 2\, \frac{(u \cdot w)}{\|w\|^2}\, w, \quad u \in E.$$

Such reflections are represented by matrices called *Householder matrices*, which play an important role in numerical matrix analysis (see Kincaid and Cheney [39] or Ciarlet [14]).

Definition 12.3. A *Householder matrix* if a matrix of the form

$$H = I_n - 2\, \frac{WW^\top}{\|W\|^2} = I_n - 2\, \frac{WW^\top}{W^\top W},$$

where $W \in \mathbb{R}^n$ is a nonzero vector.

Householder matrices are symmetric and orthogonal. It is easily checked that over an orthonormal basis $(e_1, \ldots, e_n)$, a hyperplane reflection about a hyperplane H orthogonal to a nonzero vector w is represented by the matrix

$$H = I_n - 2\,\frac{WW^\top}{\|W\|^2},$$

where W is the column vector of the coordinates of w over the basis $(e_1, \ldots, e_n)$. Since

$$p_G(u) = \frac{(u \cdot w)}{\|w\|^2}\, w,$$

the matrix representing p_G is

$$\frac{WW^\top}{W^\top W},$$

and since $p_H + p_G = \mathrm{id}$, the matrix representing p_H is

$$I_n - \frac{WW^\top}{W^\top W}.$$

These formulae can be used to derive a formula for a rotation of $\mathbb{R}^3$, given the direction w of its axis of rotation and given the angle θ of rotation.

The following fact is the key to the proof that every isometry can be decomposed as a product of reflections.

Proposition 12.2. *Let E be any nontrivial Euclidean space. For any two vectors $u, v \in E$, if $\|u\| = \|v\|$, then there is a hyperplane H such that the reflection s about H maps u to v, and if $u \neq v$, then this reflection is unique. See Figure 12.3.*

Proof. If $u = v$, then any hyperplane containing u does the job. Otherwise, we must have $H = \{v - u\}^\perp$, and by the above formula,

$$s(u) = u - 2\,\frac{(u \cdot (v - u))}{\|(v - u)\|^2}\,(v - u) = u + \frac{2\|u\|^2 - 2u \cdot v}{\|(v - u)\|^2}\,(v - u),$$

and since

$$\|(v - u)\|^2 = \|u\|^2 + \|v\|^2 - 2u \cdot v$$

and $\|u\| = \|v\|$, we have

$$\|(v - u)\|^2 = 2\|u\|^2 - 2u \cdot v,$$

and thus, $s(u) = v$. $\square$

If E is a complex vector space and the inner product is Hermitian, Proposition 12.2 is false. The problem is that the vector $v - u$ does not work unless the inner product $u \cdot v$ is real! The proposition can be salvaged enough to yield the QR-decomposition in terms of Householder transformations; see Section 13.5.

We now show that hyperplane reflections can be used to obtain another proof of the QR-decomposition.

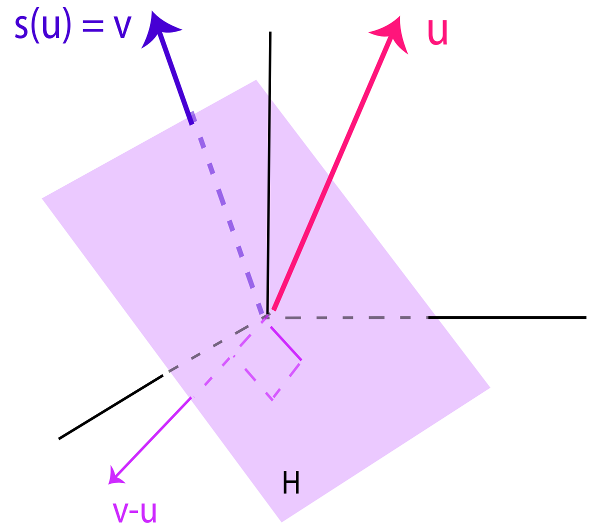

Figure 12.3: In $\mathbb{R}^3$, the (hyper)plane perpendicular to $v - u$ reflects u onto v.

12.2 QR-Decomposition Using Householder Matrices

First we state the result geometrically. When translated in terms of Householder matrices, we obtain the fact advertised earlier that every matrix (not necessarily invertible) has a QR-decomposition.

Proposition 12.3. *Let E be a nontrivial Euclidean space of dimension n. For any orthonormal basis $(e_1, \ldots, e_n)$ and for any n-tuple of vectors $(v_1, \ldots, v_n)$, there is a sequence of n isometries $h_1, \ldots, h_n$ such that h_i is a hyperplane reflection or the identity, and if $(r_1, \ldots, r_n)$ are the vectors given by*

$$r_j = h_n \circ \cdots \circ h_2 \circ h_1(v_j),$$

then every r_j is a linear combination of the vectors $(e_1, \ldots, e_j)$, $1 \leq j \leq n$. Equivalently, the matrix R whose columns are the components of the r_j over the basis $(e_1, \ldots, e_n)$ is an upper triangular matrix. Furthermore, the h_i can be chosen so that the diagonal entries of R are nonnegative.

Proof. We proceed by induction on n. For $n = 1$, we have $v_1 = \lambda e_1$ for some $\lambda \in \mathbb{R}$. If $\lambda \geq 0$, we let $h_1 = \mathrm{id}$, else if $\lambda < 0$, we let $h_1 = -\mathrm{id}$, the reflection about the origin.

For $n \geq 2$, we first have to find h_1. Let

$$r_{1,1} = \|v_1\|.$$

If $v_1 = r_{1,1}e_1$, we let $h_1 = \mathrm{id}$. Otherwise, there is a unique hyperplane reflection h_1 such that

$$h_1(v_1) = r_{1,1}\, e_1,$$

defined such that

$$h_1(u) = u - 2\,\frac{(u \cdot w_1)}{\|w_1\|^2}\, w_1$$

for all $u \in E$, where

$$w_1 = r_{1,1}\, e_1 - v_1.$$

The map h_1 is the reflection about the hyperplane H_1 orthogonal to the vector $w_1 = r_{1,1}\, e_1 - v_1$. See Figure 12.4. Letting

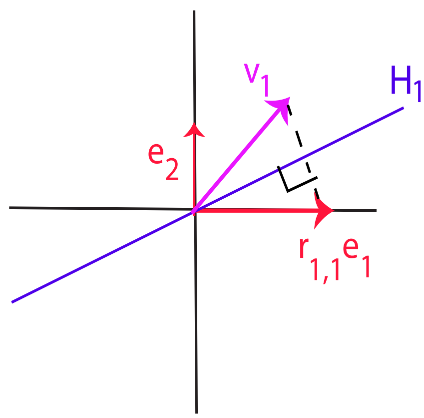

Figure 12.4: The construction of h_1 in Proposition 12.3.

$$r_1 = h_1(v_1) = r_{1,1}\, e_1,$$

it is obvious that r_1 belongs to the subspace spanned by e_1, and $r_{1,1} = \|v_1\|$ is nonnegative.

Next assume that we have found k linear maps $h_1, \ldots, h_k$, hyperplane reflections or the identity, where $1 \leq k \leq n-1$, such that if $(r_1, \ldots, r_k)$ are the vectors given by

$$r_j = h_k \circ \cdots \circ h_2 \circ h_1(v_j),$$

then every r_j is a linear combination of the vectors $(e_1, \ldots, e_j)$, $1 \leq j \leq k$. See Figure 12.5. The vectors $(e_1, \ldots, e_k)$ form a basis for the subspace denoted by U'_k, the vectors $(e_{k+1}, \ldots, e_n)$ form a basis for the subspace denoted by U''_k, the subspaces U'_k and U''_k are orthogonal, and $E = U'_k \oplus U''_k$. Let

$$u_{k+1} = h_k \circ \cdots \circ h_2 \circ h_1(v_{k+1}).$$

We can write

$$u_{k+1} = u'_{k+1} + u''_{k+1},$$

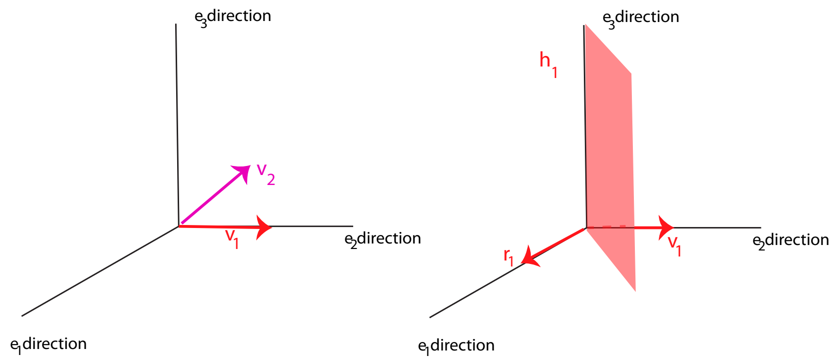

Figure 12.5: The construction of $r_1 = h_1(v_1)$ in Proposition 12.3.

where $u'_{k+1} \in U'_k$ and $u''_{k+1} \in U''_k$. See Figure 12.6. Let

$$r_{k+1,k+1} = \|u''_{k+1}\|.$$

If $u''_{k+1} = r_{k+1,k+1}\, e_{k+1}$, we let $h_{k+1} = \mathrm{id}$. Otherwise, there is a unique hyperplane reflection h_{k+1} such that

$$h_{k+1}(u''_{k+1}) = r_{k+1,k+1}\, e_{k+1},$$

defined such that

$$h_{k+1}(u) = u - 2\,\frac{(u \cdot w_{k+1})}{\|w_{k+1}\|^2}\, w_{k+1}$$

for all $u \in E$, where

$$w_{k+1} = r_{k+1,k+1}\, e_{k+1} - u''_{k+1}.$$

The map h_{k+1} is the reflection about the hyperplane H_{k+1} orthogonal to the vector $w_{k+1} = r_{k+1,k+1}\, e_{k+1} - u''_{k+1}$. However, since $u''_{k+1}, e_{k+1} \in U''_k$ and U'_k is orthogonal to U''_k, the subspace U'_k is contained in H_{k+1}, and thus, the vectors $(r_1, \ldots, r_k)$ and u'_{k+1}, which belong to U'_k, are invariant under h_{k+1}. This proves that

$$h_{k+1}(u_{k+1}) = h_{k+1}(u'_{k+1}) + h_{k+1}(u''_{k+1}) = u'_{k+1} + r_{k+1,k+1}\, e_{k+1}$$

is a linear combination of $(e_1, \ldots, e_{k+1})$. Letting

$$r_{k+1} = h_{k+1}(u_{k+1}) = u'_{k+1} + r_{k+1,k+1}\, e_{k+1},$$

since $u_{k+1} = h_k \circ \cdots \circ h_2 \circ h_1(v_{k+1})$, the vector

$$r_{k+1} = h_{k+1} \circ \cdots \circ h_2 \circ h_1(v_{k+1})$$

is a linear combination of $(e_1, \ldots, e_{k+1})$. See Figure 12.7. The coefficient of r_{k+1} over e_{k+1} is $r_{k+1,k+1} = \|u''_{k+1}\|$, which is nonnegative. This concludes the induction step, and thus the proof. $\square$

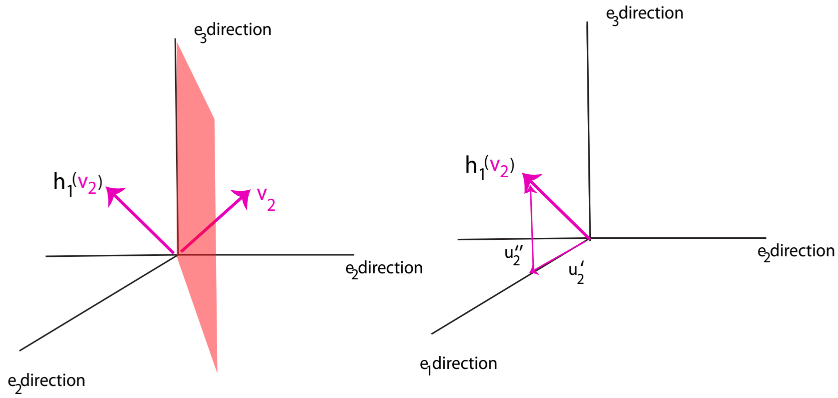

Figure 12.6: The construction of $u_2 = h_1(v_2)$ and its decomposition as $u_2 = u_2' + u_2''$.

Remarks:

(1) Since every h_i is a hyperplane reflection or the identity,

$$\rho = h_n \circ \cdots \circ h_2 \circ h_1$$

is an isometry.

(2) If we allow negative diagonal entries in R, the last isometry h_n may be omitted.

(3) Instead of picking $r_{k,k} = \|u_k''\|$, which means that

$$w_k = r_{k,k}\, e_k - u_k'',$$

where $1 \leq k \leq n$, it might be preferable to pick $r_{k,k} = -\|u_k''\|$ if this makes $\|w_k\|^2$ larger, in which case

$$w_k = r_{k,k}\, e_k + u_k''.$$

Indeed, since the definition of h_k involves division by $\|w_k\|^2$, it is desirable to avoid division by very small numbers.

(4) The method also applies to any m-tuple of vectors $(v_1, \ldots, v_m)$, with $m \leq n$. Then R is an upper triangular $m \times m$ matrix and Q is an $n \times m$ matrix with orthogonal columns ($Q^\top Q = I_m$). We leave the minor adjustments to the method as an exercise to the reader

Proposition 12.3 directly yields the QR-decomposition in terms of Householder transformations (see Strang [63, 64], Golub and Van Loan [30], Trefethen and Bau [68], Kincaid and Cheney [39], or Ciarlet [14]).

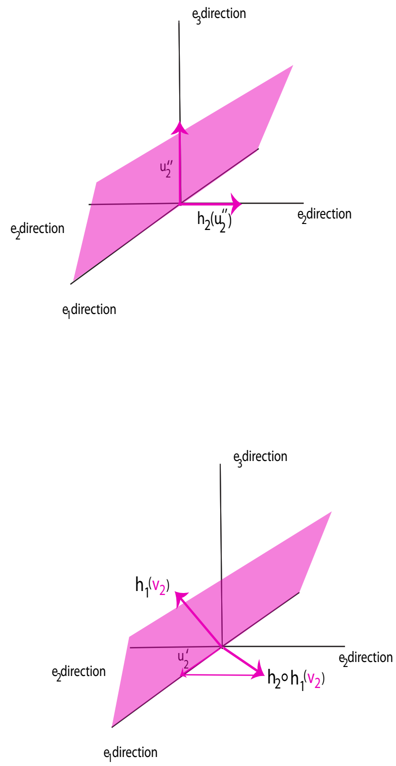

Figure 12.7: The construction of h_2 and $r_2 = h_2 \circ h_1(v_2)$ in Proposition 12.3.

Theorem 12.4. *For every real $n \times n$ matrix A, there is a sequence $H_1, \ldots, H_n$ of matrices, where each H_i is either a Householder matrix or the identity, and an upper triangular matrix R such that*

$$R = H_n \cdots H_2 H_1 A.$$

As a corollary, there is a pair of matrices Q, R, where Q is orthogonal and R is upper triangular, such that $A = QR$ (a QR-decomposition of A). Furthermore, R can be chosen so that its diagonal entries are nonnegative.

Proof. The jth column of A can be viewed as a vector v_j over the canonical basis $(e_1, \ldots, e_n)$ of $\mathbb{E}^n$ (where $(e_j)_i = 1$ if $i = j$, and 0 otherwise, $1 \leq i, j \leq n$). Applying Proposition 12.3 to $(v_1, \ldots, v_n)$, there is a sequence of n isometries $h_1, \ldots, h_n$ such that h_i is a hyperplane reflection or the identity, and if $(r_1, \ldots, r_n)$ are the vectors given by

$$r_j = h_n \circ \cdots \circ h_2 \circ h_1(v_j),$$

then every r_j is a linear combination of the vectors $(e_1, \ldots, e_j)$, $1 \leq j \leq n$. Letting R be the matrix whose columns are the vectors r_j, and H_i the matrix associated with h_i, it is clear that

$$R = H_n \cdots H_2 H_1 A,$$

where R is upper triangular and every H_i is either a Householder matrix or the identity. However, $h_i \circ h_i = \mathrm{id}$ for all i, $1 \leq i \leq n$, and so

$$v_j = h_1 \circ h_2 \circ \cdots \circ h_n(r_j)$$

for all j, $1 \leq j \leq n$. But $\rho = h_1 \circ h_2 \circ \cdots \circ h_n$ is an isometry represented by the orthogonal matrix $Q = H_1 H_2 \cdots H_n$. It is clear that $A = QR$, where R is upper triangular. As we noted in Proposition 12.3, the diagonal entries of R can be chosen to be nonnegative. $\square$

Remarks:

(1) Letting

$$A_{k+1} = H_k \cdots H_2 H_1 A,$$

with $A_1 = A$, $1 \leq k \leq n$, the proof of Proposition 12.3 can be interpreted in terms of the computation of the sequence of matrices $A_1, \ldots, A_{n+1} = R$. The matrix A_{k+1} has the shape

$$A_{k+1} = \begin{pmatrix} \times & \times & \times & u_1^{k+1} & \times & \times & \times & \times \\ 0 & \times & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \times & u_k^{k+1} & \times & \times & \times & \times \\ 0 & 0 & 0 & u_{k+1}^{k+1} & \times & \times & \times & \times \\ 0 & 0 & 0 & u_{k+2}^{k+1} & \times & \times & \times & \times \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & u_{n-1}^{k+1} & \times & \times & \times & \times \\ 0 & 0 & 0 & u_n^{k+1} & \times & \times & \times & \times \end{pmatrix},$$

where the $(k+1)$th column of the matrix is the vector

$$u_{k+1} = h_k \circ \cdots \circ h_2 \circ h_1(v_{k+1}),$$

and thus

$$u'_{k+1} = \left(u_1^{k+1}, \ldots, u_k^{k+1}\right)$$

and

$$u''_{k+1} = \left(u_{k+1}^{k+1}, u_{k+2}^{k+1}, \ldots, u_n^{k+1}\right).$$

If the last $n-k-1$ entries in column $k+1$ are all zero, there is nothing to do, and we let $H_{k+1} = I$. Otherwise, we kill these $n-k-1$ entries by multiplying A_{k+1} on the left by the Householder matrix H_{k+1} sending

$$\left(0, \ldots, 0, u_{k+1}^{k+1}, \ldots, u_n^{k+1}\right) \quad \text{to} \quad (0, \ldots, 0, r_{k+1,k+1}, 0, \ldots, 0),$$

where $r_{k+1,k+1} = \|(u_{k+1}^{k+1}, \ldots, u_n^{k+1})\|$.

(2) If A is invertible and the diagonal entries of R are positive, it can be shown that Q and R are unique.

(3) If we allow negative diagonal entries in R, the matrix H_n may be omitted ($H_n = I$).

(4) The method allows the computation of the determinant of A. We have

$$\det(A) = (-1)^m r_{1,1} \cdots r_{n,n},$$

where m is the number of Householder matrices (not the identity) among the H_i.

(5) The "condition number" of the matrix A is preserved (see Strang [64], Golub and Van Loan [30], Trefethen and Bau [68], Kincaid and Cheney [39], or Ciarlet [14]). This is very good for numerical stability.

(6) The method also applies to a rectangular $m \times n$ matrix. If $m \geq n$, then R is an $n \times n$ upper triangular matrix and Q is an $m \times n$ matrix such that $Q^\top Q = I_n$.

The following `Matlab` functions implement the QR-factorization method of a real square (possibly singular) matrix A using Householder reflections

The main function `houseqr` computes the upper triangular matrix R obtained by applying Householder reflections to A. It makes use of the function `house`, which computes a unit vector u such that given a vector $x \in \mathbb{R}^p$, the Householder transformation $P = I - 2uu^\top$ sets to zero all entries in x but the first entry x_1. It only applies if $\|x(2:p)\|_1 = |x_2| + \cdots + |x_p| > 0$. Since computations are done in floating point, we use a tolerance factor tol, and if $\|x(2:p)\|_1 \leq tol$, then we return $u = 0$, which indicates that the corresponding Householder transformation is the identity. To make sure that $\|Px\|$ is as large as possible, we pick $uu = x + \operatorname{sign}(x_1)\,\|x\|_2\, e_1$, where $\operatorname{sign}(z) = 1$ if $z \geq 0$ and $\operatorname{sign}(z) = -1$ if $z < 0$. Note that as a result, diagonal entries in R may be negative. We will take care of this issue later.

```
function s = signe(x)
%  if x >= 0, then signe(x) = 1
%  else if x < 0 then signe(x) = -1
%

if x < 0
   s = -1;
else
    s = 1;
end
end

function [uu, u] = house(x)
% This constructs the unnormalized  vector uu
% defining the Householder reflection that
% zeros all but the first entries in x.
% u is the normalized vector uu/||uu||
```

```
%

tol = 2*10^(-15);   % tolerance
uu = x;
p = size(x,1);
% computes l^1-norm of x(2:p,1)
n1 = sum(abs(x(2:p,1)));
if n1 <= tol
   u = zeros(p,1);  uu = u;
else
   l = sqrt(x'*x);  % l^2 norm of x
   uu(1) = x(1) + signe(x(1))*l;
   u = uu/sqrt(uu'*uu);
end
end
```

The Householder transformations are recorded in an array u of $n-1$ vectors. There are more efficient implementations, but for the sake of clarity we present the following version.

```
function [R,  u] = houseqr(A)
%  This function computes the upper triangular R in the QR
%  factorization of A using Householder reflections, and an
%  implicit representation of Q as  a sequence of n - 1
%  vectors u_i representing Householder reflections

n = size(A, 1);
R = A;
u = zeros(n,n-1);
for i = 1:n-1
    [~, u(i:n,i)] = house(R(i:n,i));
    if u(i:n,i) == zeros(n - i + 1,1)
       R(i+1:n,i) = zeros(n - i,1);
    else
       R(i:n,i:n) = R(i:n,i:n)
                    - 2*u(i:n,i)*(u(i:n,i)'*R(i:n,i:n));
    end
end
end
```

If only R is desired, then `houseqr` does the job. In order to obtain R, we need to compose the Householder transformations. We present a simple method which is not the most efficient (there is a way to avoid multiplying explicity the Householder matrices).

The function `buildhouse` creates a Householder reflection from a vector v.

```
function P = buildhouse(v,i)
% This function builds a Householder reflection
%   [I 0 ]
%   [0 PP]
%   from a Householder reflection
%   PP = I - 2uu*uu'
%   where uu = v(i:n)
%   If uu = 0 then P - I
%

n = size(v,1);
if v(i:n) == zeros(n - i + 1,1)
   P = eye(n);
else
   PP = eye(n - i + 1) - 2*v(i:n)*v(i:n)';
   P = [eye(i-1) zeros(i-1, n - i + 1);
   zeros(n - i + 1, i - 1) PP];
end
end
```

The function `buildQ` builds the matrix Q in the QR-decomposition of A.

```
function Q = buildQ(u)
% Builds the matrix Q in the QR decomposition
% of an nxn  matrix A using Householder matrices,
% where u is a representation of the n - 1
% Householder reflection by a list u of vectors produced by
% houseqr

n = size(u,1);
Q = buildhouse(u(:,1),1);
for i = 2:n-1
  Q = Q*buildhouse(u(:,i),i);
end
end
```

The function `buildhouseQR` computes a QR-factorization of A. At the end, if some entries on the diagonal of R are negative, it creates a diagonal orthogonal matrix P such that PR has nonnegative diagonal entries, so that $A = (QP)(PR)$ is the desired QR-factorization of A.

```
function [Q,R] = buildhouseQR(A)
%
%   Computes the QR decomposition of a square
```

```
%   matrix A (possibly singular) using Householder reflections

n = size(A,1);
[R,u] = houseqr(A);
Q = buildQ(u);
% Produces a matrix R whose diagonal entries are
% nonnegative
P = eye(n);
for i = 1:n
   if R(i,i) < 0
      P(i,i) = -1;
   end
end
Q = Q*P; R = P*R;
end
```

Example 12.1. Consider the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{pmatrix}.$$

Running the function `buildhouseQR`, we get

$$Q = \begin{pmatrix} 0.1826 & 0.8165 & 0.4001 & 0.3741 \\ 0.3651 & 0.4082 & -0.2546 & -0.7970 \\ 0.5477 & -0.0000 & -0.6910 & 0.4717 \\ 0.7303 & -0.4082 & 0.5455 & -0.0488 \end{pmatrix}$$

and

$$R = \begin{pmatrix} 5.4772 & 7.3030 & 9.1287 & 10.9545 \\ 0 & 0.8165 & 1.6330 & 2.4495 \\ 0 & -0.0000 & 0.0000 & 0.0000 \\ 0 & -0.0000 & 0 & 0.0000 \end{pmatrix}.$$

Observe that A has rank 2. The reader should check that $A = QR$.

Remark: Curiously, running `Matlab` built-in function `qr`, the same R is obtained (up to column signs) but a different Q is obtained (the last two columns are different).

12.3 Summary

The main concepts and results of this chapter are listed below:

- *Symmetry (or reflection) with respect to F and parallel to G.*
- *Orthogonal symmetry (or reflection) with respect to F and parallel to G; reflections, flips.*
- Hyperplane reflections and *Householder matrices*.
- A key fact about reflections (Proposition 12.2).
- *QR-decomposition in terms of Householder transformations* (Theorem 12.4).

12.4 Problems

Problem 12.1. (1) Given a unit vector $(-\sin\theta, \cos\theta)$, prove that the Householder matrix determined by the vector $(-\sin\theta, \cos\theta)$ is

$$\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}.$$

Give a geometric interpretation (i.e., why the choice $(-\sin\theta, \cos\theta)$?).

(2) Given any matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

Prove that there is a Householder matrix H such that AH is lower triangular, i.e.,

$$AH = \begin{pmatrix} a' & 0 \\ c' & d' \end{pmatrix}$$

for some $a', c', d' \in \mathbb{R}$.

Problem 12.2. Given a Euclidean space E of dimension n, if h is a reflection about some hyperplane orthogonal to a nonzero vector u and f is any isometry, prove that $f \circ h \circ f^{-1}$ is the reflection about the hyperplane orthogonal to $f(u)$.

Problem 12.3. (1) Given a matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

prove that there are Householder matrices G, H such that

$$GAH = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} = D,$$

where D is a diagonal matrix, iff the following equations hold:

$$(b + c)\cos(\theta + \varphi) = (a - d)\sin(\theta + \varphi),$$
$$(c - b)\cos(\theta - \varphi) = (a + d)\sin(\theta - \varphi).$$

(2) Discuss the solvability of the system. Consider the following cases:
Case 1: $a - d = a + d = 0$.
Case 2a: $a - d = b + c = 0$, $a + d \neq 0$.
Case 2b: $a - d = 0$, $b + c \neq 0$, $a + d \neq 0$.
Case 3a: $a + d = c - b = 0$, $a - d \neq 0$.
Case 3b: $a + d = 0$, $c - b \neq 0$, $a - d \neq 0$.
Case 4: $a + d \neq 0$, $a - d \neq 0$. Show that the solution in this case is

$$\begin{aligned} \theta &= \frac{1}{2}\left[\arctan\left(\frac{b+c}{a-d}\right) + \arctan\left(\frac{c-b}{a+d}\right)\right], \\ \varphi &= \frac{1}{2}\left[\arctan\left(\frac{b+c}{a-d}\right) - \arctan\left(\frac{c-b}{a+d}\right)\right]. \end{aligned}$$

If $b = 0$, show that the discussion is simpler: basically, consider $c = 0$ or $c \neq 0$.

(3) Expressing everything in terms of $u = \cot\theta$ and $v = \cot\varphi$, show that the equations in (2) become

$$\begin{aligned} (b+c)(uv-1) &= (u+v)(a-d), \\ (c-b)(uv+1) &= (-u+v)(a+d). \end{aligned}$$

Problem 12.4. Let A be an $n \times n$ real invertible matrix.

(1) Prove that $A^\top A$ is symmetric positive definite.

(2) Use the Cholesky factorization $A^\top A = R^\top R$ with R upper triangular with positive diagonal entries to prove that $Q = AR^{-1}$ is orthogonal, so that $A = QR$ is the QR-factorization of A.

Problem 12.5. Modify the function `houseqr` so that it applies to an $m \times n$ matrix with $m \geq n$, to produce an $m \times n$ upper-triangular matrix whose last $m - n$ rows are zeros.

Problem 12.6. The purpose of this problem is to prove that given any self-adjoint linear map $f\colon E \to E$ (i.e., such that $f^* = f$), where E is a Euclidean space of dimension $n \geq 3$, given an orthonormal basis $(e_1, \ldots, e_n)$, there are $n - 2$ isometries h_i, hyperplane reflections or the identity, such that the matrix of

$$h_{n-2} \circ \cdots \circ h_1 \circ f \circ h_1 \circ \cdots \circ h_{n-2}$$

is a symmetric tridiagonal matrix.

(1) Prove that for any isometry $f\colon E \to E$, we have $f = f^* = f^{-1}$ iff $f \circ f = \mathrm{id}$.

Prove that if f and h are self-adjoint linear maps ($f^* = f$ and $h^* = h$), then $h \circ f \circ h$ is a self-adjoint linear map.

(2) Let V_k be the subspace spanned by $(e_{k+1}, \ldots, e_n)$. Proceed by induction. For the base case, proceed as follows.

Let

$$f(e_1) = a_1^0 e_1 + \cdots + a_n^0 e_n,$$

and let

$$r_{1,\,2} = \left\| a_2^0 e_2 + \cdots + a_n^0 e_n \right\|.$$

Find an isometry h_1 (reflection or id) such that

$$h_1(f(e_1) - a_1^0 e_1) = r_{1,\,2}\, e_2.$$

Observe that

$$w_1 = r_{1,\,2}\, e_2 + a_1^0 e_1 - f(e_1) \in V_1,$$

and prove that $h_1(e_1) = e_1$, so that

$$h_1 \circ f \circ h_1(e_1) = a_1^0 e_1 + r_{1,\,2}\, e_2.$$

Let $f_1 = h_1 \circ f \circ h_1$.

Assuming by induction that

$$f_k = h_k \circ \cdots \circ h_1 \circ f \circ h_1 \circ \cdots \circ h_k$$

has a tridiagonal matrix up to the kth row and column, $1 \leq k \leq n-3$, let

$$f_k(e_{k+1}) = a_k^k e_k + a_{k+1}^k e_{k+1} + \cdots + a_n^k e_n,$$

and let

$$r_{k+1,\,k+2} = \left\| a_{k+2}^k e_{k+2} + \cdots + a_n^k e_n \right\|.$$

Find an isometry h_{k+1} (reflection or id) such that

$$h_{k+1}(f_k(e_{k+1}) - a_k^k e_k - a_{k+1}^k e_{k+1}) = r_{k+1,\,k+2}\, e_{k+2}.$$

Observe that

$$w_{k+1} = r_{k+1,\,k+2}\, e_{k+2} + a_k^k e_k + a_{k+1}^k e_{k+1} - f_k(e_{k+1}) \in V_{k+1},$$

and prove that $h_{k+1}(e_k) = e_k$ and $h_{k+1}(e_{k+1}) = e_{k+1}$, so that

$$h_{k+1} \circ f_k \circ h_{k+1}(e_{k+1}) = a_k^k e_k + a_{k+1}^k e_{k+1} + r_{k+1,\,k+2}\, e_{k+2}.$$

Let $f_{k+1} = h_{k+1} \circ f_k \circ h_{k+1}$, and finish the proof.

(3) Prove that given any symmetric $n \times n$-matrix A, there are $n-2$ matrices $H_1, \ldots, H_{n-2}$, Householder matrices or the identity, such that

$$B = H_{n-2} \cdots H_1 A H_1 \cdots H_{n-2}$$

is a symmetric tridiagonal matrix.

(4) Write a computer program implementing the above method.

Problem 12.7. Recall from Problem 5.6 that an $n \times n$ matrix H is *upper Hessenberg* if $h_{jk} = 0$ for all (j, k) such that $j - k \geq 0$. Adapt the proof of Problem 12.6 to prove that given any $n \times n$-matrix A, there are $n - 2 \geq 1$ matrices $H_1, \ldots, H_{n-2}$, Householder matrices or the identity, such that

$$B = H_{n-2} \cdots H_1 A H_1 \cdots H_{n-2}$$

is upper Hessenberg.

Problem 12.8. The purpose of this problem is to prove that given any linear map $f \colon E \to E$, where E is a Euclidean space of dimension $n \geq 2$, given an orthonormal basis $(e_1, \ldots, e_n)$, there are isometries g_i, h_i, hyperplane reflections or the identity, such that the matrix of

$$g_n \circ \cdots \circ g_1 \circ f \circ h_1 \circ \cdots \circ h_n$$

is a lower bidiagonal matrix, which means that the nonzero entries (if any) are on the main descending diagonal and on the diagonal below it.

(1) Let U'_k be the subspace spanned by $(e_1, \ldots, e_k)$ and U''_k be the subspace spanned by $(e_{k+1}, \ldots, e_n)$, $1 \leq k \leq n - 1$. Proceed by induction For the base case, proceed as follows.

Let $v_1 = f^*(e_1)$ and $r_{1,1} = \|v_1\|$. Find an isometry h_1 (reflection or id) such that

$$h_1(f^*(e_1)) = r_{1,1} e_1.$$

Observe that $h_1(f^*(e_1)) \in U'_1$, so that

$$\langle h_1(f^*(e_1)), e_j \rangle = 0$$

for all $j, 2 \leq j \leq n$, and conclude that

$$\langle e_1, f \circ h_1(e_j) \rangle = 0$$

for all $j, 2 \leq j \leq n$.

Next let

$$u_1 = f \circ h_1(e_1) = u'_1 + u''_1,$$

where $u'_1 \in U'_1$ and $u''_1 \in U''_1$, and let $r_{2,1} = \|u''_1\|$. Find an isometry g_1 (reflection or id) such that

$$g_1(u''_1) = r_{2,1} e_2.$$

Show that $g_1(e_1) = e_1$,

$$g_1 \circ f \circ h_1(e_1) = u'_1 + r_{2,1} e_2,$$

and that

$$\langle e_1, g_1 \circ f \circ h_1(e_j) \rangle = 0$$

for all $j, 2 \leq j \leq n$. At the end of this stage, show that $g_1 \circ f \circ h_1$ has a matrix such that all entries on its first row except perhaps the first are zero, and that all entries on the first column, except perhaps the first two, are zero.

Assume by induction that some isometries $g_1, \ldots, g_k$ and $h_1, \ldots, h_k$ have been found, either reflections or the identity, and such that

$$f_k = g_k \circ \cdots \circ g_1 \circ f \circ h_1 \cdots \circ h_k$$

has a matrix which is lower bidiagonal up to and including row and column k, where $1 \leq k \leq n-2$.

Let

$$v_{k+1} = f_k^*(e_{k+1}) = v'_{k+1} + v''_{k+1},$$

where $v'_{k+1} \in U'_k$ and $v''_{k+1} \in U''_k$, and let $r_{k+1,\,k+1} = \left\| v''_{k+1} \right\|$. Find an isometry h_{k+1} (reflection or id) such that

$$h_{k+1}(v''_{k+1}) = r_{k+1,\,k+1} e_{k+1}.$$

Show that if h_{k+1} is a reflection, then $U'_k \subseteq H_{k+1}$, where H_{k+1} is the hyperplane defining the reflection h_{k+1}. Deduce that $h_{k+1}(v'_{k+1}) = v'_{k+1}$, and that

$$h_{k+1}(f_k^*(e_{k+1})) = v'_{k+1} + r_{k+1,\,k+1} e_{k+1}.$$

Observe that $h_{k+1}(f_k^*(e_{k+1})) \in U'_{k+1}$, so that

$$\langle h_{k+1}(f_k^*(e_{k+1})), e_j \rangle = 0$$

for all j, $k+2 \leq j \leq n$, and thus,

$$\langle e_{k+1}, f_k \circ h_{k+1}(e_j) \rangle = 0$$

for all j, $k+2 \leq j \leq n$.

Next let

$$u_{k+1} = f_k \circ h_{k+1}(e_{k+1}) = u'_{k+1} + u''_{k+1},$$

where $u'_{k+1} \in U'_{k+1}$ and $u''_{k+1} \in U''_{k+1}$, and let $r_{k+2,\,k+1} = \left\| u''_{k+1} \right\|$. Find an isometry g_{k+1} (reflection or id) such that

$$g_{k+1}(u''_{k+1}) = r_{k+2,\,k+1} e_{k+2}.$$

Show that if g_{k+1} is a reflection, then $U'_{k+1} \subseteq G_{k+1}$, where G_{k+1} is the hyperplane defining the reflection g_{k+1}. Deduce that $g_{k+1}(e_i) = e_i$ for all i, $1 \leq i \leq k+1$, and that

$$g_{k+1} \circ f_k \circ h_{k+1}(e_{k+1}) = u'_{k+1} + r_{k+2,\,k+1} e_{k+2}.$$

Since by induction hypothesis,

$$\langle e_i, f_k \circ h_{k+1}(e_j) \rangle = 0$$

for all i, j, $1 \leq i \leq k+1$, $k+2 \leq j \leq n$, and since $g_{k+1}(e_i) = e_i$ for all i, $1 \leq i \leq k+1$, conclude that

$$\langle e_i, g_{k+1} \circ f_k \circ h_{k+1}(e_j) \rangle = 0$$

for all i, j, $1 \leq i \leq k+1$, $k+2 \leq j \leq n$. Finish the proof.