Chapter 10

QR-Decomposition for Arbitrary Matrices

10.1 Orthogonal Reflections

Orthogonal symmetries are a very important example of isometries. First let us review the definition of a (linear) *projection*.

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

$$p_F(u) = v \quad \text{and} \quad p_G(u) = w.$$

It is immediately verified that p_G and p_F are linear maps, and that $p_F^2 = p_F$, $p_G^2 = p_G$, $p_F \circ p_G = p_G \circ p_F = 0$, and $p_F + p_G = \mathrm{id}$.

Definition 10.1. Given a vector space E, for any two subspaces F and G that form a direct sum $E = F \oplus G$, the *symmetry with respect to F and parallel to G, or reflection about F* 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 10.2. 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 with respect to F and parallel to G, or orthogonal reflection about F* is the linear map $s\colon E \to E$, defined such that

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

for every $u \in E$.

When F is a hyperplane, we call s an *hyperplane symmetry with respect to F or reflection about F*, and when G is a plane, we call s a *flip about F*.

It is easy to show that s is an isometry.

Using Proposition 9.8, 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, we observe that every improper orthogonal transformation is the composition of a rotation with 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.

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.

We get

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

Such reflections are represented by matrices called *Householder matrices*, and they play an important role in numerical matrix analysis. Householder matrices are symmetric and orthogonal.

Over an orthonormal basis $(e_1, \ldots, e_n)$, a hyperplane reflection about a hyperplane H orthogonal to a nonnull vector w is represented by the matrix

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

where W is the column vector of the coordinates of w.

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

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

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

Proposition 10.1. *Let E be any nontrivial Euclidean space. For any two vectors $u, v \in E$, if $\|u\| = \|v\|$, then there is an hyperplane H such that the reflection s about H maps u to v, and if $u \neq v$, then this reflection is unique.*

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

10.2 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 10.2. *Let E be a nontrivial Euclidean space of dimension n. Given any orthonormal basis $(e_1, \ldots, e_n)$, 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.

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.

Proposition 10.2 immediately yields the QR-decomposition in terms of Householder transformations.

Theorem 10.3. *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 non-negative.

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 10.2 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

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

and

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

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 $(0,\ldots,0,u_{k+1}^{k+1},\ldots,u_n^{k+1})$ to $(0,\ldots,0,r_{k+1,k+1},0,\ldots,0)$, where

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

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

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

(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. This is very good for numerical stability.