ROTATION MATRICES

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Note that

$$A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \quad A \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

This shows the vectors $e^{(1)}$ and $e^{(2)}$ are rotated counter-clockwise thru an angle of θ radians. In particular,

$$Ae^{(2)} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} = \begin{bmatrix} \cos\left(\theta + \frac{\pi}{2}\right) \\ \sin\left(\theta + \frac{\pi}{2}\right) \end{bmatrix}$$

Thus in general, the transformation $x \longrightarrow Ax$ corresponds to a rotation of x counter-clockwise thru an angle of θ radians.

A^{-1} should correspond to a *clockwise* rotation thru $-\theta$ radians; or replacing θ by $-\theta$ in the original formula for A, we have

$$A^{-1} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

We generalize this in a very simple way to $\mathbb{R}^n$. Let $1 \leq k < l \leq n$, and define the matrix $R_{k,l}$ to the following matrix:

$$\begin{bmatrix}
1 & 0 & & & & \cdots & & & & & 0 \\
0 & \ddots & \ddots & & & & & & & & \\
 & \ddots & 1 & 0 & & & & 0 & & & \\
 & & 0 & \cos\theta & 0 & \cdots & 0 & -\sin\theta & 0 & & \\
 & & & 0 & 1 & & & 0 & & & \\
\vdots & & \vdots & \vdots & & \ddots & & \vdots & & & \vdots \\
 & & & 0 & & & 1 & 0 & & & \\
 & & 0 & \sin\theta & 0 & \cdots & 0 & \cos\theta & 0 & & \\
 & & & 0 & & & & 0 & 1 & \ddots & \\
 & & & \vdots & & & & \vdots & \ddots & \ddots & 0 \\
0 & & & & & \cdots & & & & 0 & 1
\end{bmatrix}$$

In it, we have modified the identity matrix I, changing it in the four elements in positions (k, k), (k, l) , (l, k) , and (l, l).

The matrix $R_{k,l}$ will rotate the (k, l)-plane thru an angle of θ, while leaving unchanged the remainder of $\mathbb{R}^n$, that part perpendicular to (k, l)-plane.

HOUSEHOLDER MATRICES

Let $w \in \mathbb{C}^n$ be a vector of Euclidean length 1,

$$w^* w = 1 \quad \left(w^T w = 1 \text{ for } w \in \mathbb{R}^n\right)$$

Define the matrix

$$H = I - 2ww^* \quad \left(I - 2ww^T \text{ for } w \in \mathbb{R}^n\right)$$

Then H is a Hermitian unitary matrix (orthogonal if $w \in \mathbb{R}^n$).

First,

$$\begin{aligned} H^* &= (I - 2ww^*)^* \\ &= I - 2\,(ww^*)^* \\ &= I - 2\,(w^*)^*\, w^* = H \end{aligned}$$

Also,

$$\begin{aligned} H^* H &= H^2 \\ &= (I - 2ww^*)^2 \\ &= I - 4ww^* + 4\,(ww^*)\,(ww^*) \end{aligned}$$

Note that

$$(ww^*)(ww^*) = w\underbrace{(w^*w)}_{=1}w^* = ww^*$$

Thus

$$H^*H = I - 4ww^* + 4(ww^*)(ww^*) = I$$

showing H is unitary.

What does H do in a geometric sense? First, note that

$$Hw = (I - 2ww^*)w = w - 2ww^*w = w - 2w = -w$$

Also, let v be any vector orthogonal to w. Then

$$(ww^*)v = w(w^*v) = w(0) = 0$$

$$Hv = (I - 2ww^*)v = v - 2(ww^*)v = v$$

Thus H is a reflection of space thru the $(n-1)$-dimensional hyperplane perpendicular to w.

The rotation matrices $R_{k,l}$ and the Householder matrices H are the most commonly used orthogonal (or unitary) matrices used in numerical analysis.

EXAMPLES

For the $n = 3$ case, with $w \in \mathbb{R}^3$,

$$H = \begin{bmatrix} 1-2w_1^2 & -2w_1w_2 & -2w_1w_3 \\ -2w_1w_2 & 1-2w_2^2 & -2w_2w_3 \\ -2w_1w_3 & -2w_2w_3 & 1-2w_3^2 \end{bmatrix}$$

Then $H^T = H$ and $H^2 = I$.

For $w = \left[\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right]^T$,

$$H = \begin{bmatrix} \frac{7}{9} & -\frac{4}{9} & -\frac{4}{9} \\ -\frac{4}{9} & \frac{1}{9} & -\frac{8}{9} \\ -\frac{4}{9} & -\frac{8}{9} & \frac{1}{9} \end{bmatrix}$$

For $w = \left[0, \frac{3}{5}, \frac{4}{5}\right]^T$,

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{7}{25} & -\frac{24}{25} \\ 0 & -\frac{24}{25} & -\frac{7}{25} \end{bmatrix}$$

REDUCTION OF A VECTOR

Given a vector $d \in \mathbb{R}^m$, we want to find $v \in \mathbb{R}^m$ with $\|v\|_2 = 1$ and

$$\left(I - 2vv^T\right) d = \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad (1)$$

Since $I - 2vv^T$ is orthogonal, the length of d is preserved in the tranformation of d:

$$|\alpha| = \|d\|_2 \equiv S$$

and

$$\alpha = \pm S = \pm \,\mathsf{sqrt}\left(d_1^2 + \cdots d_m^2\right)$$

with the sign to be determined later.

Introduce

$$p = v^T d$$

From (1),

$$d - 2pv = [\alpha, 0, ..., 0]^T \tag{2}$$

Multiply on the left by v^T to get

$$\begin{aligned} p - 2pv^T v &= v_1 \alpha \\ p &= -\alpha v_1 \end{aligned}$$

Substitute this back into (2). Then look at the individual components, obtaining

$$\begin{aligned} d_1 + 2\alpha v_1^2 &= \alpha \\ d_i - 2pv_i &= 0, \quad i = 2, ..., m \end{aligned} \tag{3}$$

From the first equation,

$$v_1^2 = \frac{\alpha - d_1}{2\alpha} = \frac{1}{2}\left(1 - \frac{d_1}{\alpha}\right) \tag{4}$$

Recall that we have not yet chosen the sign of α. Now choose the sign of α according to

$$\text{sign}(\alpha) = -\,\text{sign}(d_1)$$

The subtraction in (4) is now actually an addition, so as to avoid a 'loss of significance' error. We can take the square root in (4) to find v_1, and there is no obvious choice of the sign here, although most people would choose $v_1 > 0$.

Return to $p = -\alpha v_1$ to obtain p. Then return to (3) to find $v_2, ..., v_m$:

$$v_i = \frac{d_i}{2p}, \quad i = 2, ..., m$$

This completes the construction of v and therefore the Householder matrix $I - 2vv^T$ based upon it. Again, we now have

$$\left(I - 2vv^T\right) d = \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

EXAMPLE

Let

$$d = [2, 2, 1]^T$$

Then

$$\alpha = -\|d\|_2 = -3, \quad v_1 = \text{sqrt}\left(\frac{5}{6}\right), \quad p = \text{sqrt}\left(\frac{15}{2}\right)$$

$$v_2 = \frac{2}{\text{sqrt}\,(30)}, \quad v_3 = \frac{1}{\text{sqrt}\,(30)}$$

Then

$$H = \begin{bmatrix} -\frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{11}{15} & -\frac{2}{15} \\ -\frac{1}{3} & -\frac{2}{15} & \frac{14}{15} \end{bmatrix}$$

In practice, one does not need H explicitly.

QR FACTORIZATION

Let A be a matrix that is $n \times n$. We want to factor it into the form

$$A = QR$$

with Q orthogonal and R upper triangular. We do this by working on each of the columns in succession.

For the first step, let

$$P_1 = I - 2w^{(1)}w^{(1)T}, \quad \left\| w^{(1)} \right\|_2 = 1$$

Let $A = \left[A_{*,1}, ..., A_{*,n} \right]$. Then

$$P_1 A = \left[P_1 A_{*,1}, ..., P_1 A_{*,n} \right]$$

We choose P_1 so that

$$P_1 A_{*,1} = [\alpha, 0, ..., 0]^T$$

We can of course do this by the construction already described above, with $d = A_{*,1}$.

With this choice of P_1, the matrix $P_1 A$ will have zeroes in the first column below the diagonal position. Next, we wish to do the same to the second column, but without changing the first column.

Define

$$P_2 = I - 2w^{(2)}w^{(2)T}, \quad \left\| w^{(2)} \right\|_2 = 1$$

But now require

$$w^{(2)} = \begin{bmatrix} 0 \\ v \end{bmatrix}, \quad v^T v = 1, \quad v \in \mathbb{R}^{n-1}$$

Then

$$P_2 = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & I - 2vv^T \end{bmatrix}$$

Calculate $P_2 P_1 A$:

$$P_2 P_1 A = \left[P_2 P_1 A_{*,1}, ..., P_2 P_1 A_{*,n} \right]$$

We know that

$$P_1 A_{*,1} = [\alpha, 0, ..., 0]^T$$

and therefore

$$P_2 P_1 A_{*,1} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & I - 2vv^T \end{bmatrix} \begin{bmatrix} \alpha \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \alpha \\ \mathbf{0} \end{bmatrix}$$

Thus the first column of $P_2 P_1 A$ retains its zero structure below the diagonal. Now choose v so as to force the second column of $P_2 P_1 A$ to have zeroes below the diagonal position. Writing

$$P_1 A_{*,2} = \begin{bmatrix} c \\ \mathbf{d} \end{bmatrix}, \quad \mathbf{d} \in \mathbb{R}^{n-1}$$

we choose v by forcing

$$\left(I - 2vv^T\right) \mathbf{d} = [\alpha_2, 0, ..., 0]^T$$

for a suitable value of α_2.

Continue in this manner, defining

$$P_i = I - 2w^{(i)}w^{(i)T}, \quad \left\|w^{(i)}\right\|_2 = 1$$

with

$$w^{(i)} = \begin{bmatrix} 0 \\ v \end{bmatrix}, \quad v^T v = 1, \quad v \in \mathbb{R}^{n-i+1}$$

and choosing v to force $P_i \cdots P_1 A$ to have zeros below the diagonal in column i. After $n-1$ steps, the matrix

$$P_{n-1} \cdots P_2 P_1 A \equiv R$$

is upper triangular. Note that the matrix $P_{n-1} \cdots P_2 P_1$ is orthogonal,

$$\begin{aligned} &(P_{n-1} \cdots P_2 P_1)^T (P_{n-1} \cdots P_2 P_1) \\ &\quad = P_1^T \cdots P_{n-1}^T P_{n-1} \cdots P_2 P_1 \\ &\quad = I \end{aligned}$$

We define

$$Q^T = P_{n-1} \cdots P_2 P_1$$

Then

$$\begin{aligned} Q^T A &= R \\ A &= QR \end{aligned}$$

EXAMPLE

An example for the matrix

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

is given on page 613 in the text. The result is $A = QR$ with

$$R = \begin{bmatrix} -4.24264 & -2.12132 & -2.12132 \\ 0 & -3.67423 & -1.22475 \\ 0 & 0 & 3.46410 \end{bmatrix}$$

$$Q = P_1 P_2 = \begin{bmatrix} -0.94281 & 0.27217 & -0.19245 \\ -0.23570 & -0.95258 & -0.19245 \\ -0.23570 & -0.13608 & 0.96225 \end{bmatrix}$$

This factorization can also be accomplished using the Matlab instruction qr:

$$[Q \quad R] = qr(A)$$

HOUSEHOLDER MULTIPLICATIONS

Consider multiplying a Householder matrix H times another matrix:

$$\begin{aligned} HA = \left(I - 2ww^T\right) A &= A - \left(2ww^T\right) A \\ &= A - 2w\left(w^T A\right) \end{aligned}$$

The quantity $w^T A$ is a vector, and it can be produced with approximately $2n^2$ operations if w has all nonzero components. Then calculating $2w\left(w^T A\right)$ will cost a further n^2 multiplications, approximately, and $A - 2w\left(w^T A\right)$ will cost n^2 subtractions. Thus the total operations cost to produce HA will be around $4n^2$ operations, rather than the $2n^3$ one would usually expect with multiplying two $n \times n$ matrices.

This also shows why we generally do not produce $Q = P_1 \cdots P_{n-1}$ explicitly, as it is cheaper to carry out matrix multiplications when Q is in factored form.

REDUCTION OF SYMMETRIC MATRICES

Let A be a symmetric matrix. We want to use a similarity transformation to reduce it to symmetric tridiagonal form.

Assume A is $n \times n$. We will perform a sequence of similarity transformations, using Householder matrices, to reduce A to tridiagonal form. We begin by using a similarity transformation to put zeros into the first column of the matrix, in the positions below the (2,1) position.

$$P_1^T A P_1 = A_2 = \begin{bmatrix} a_{1,1} & \widehat{a}_{2,1} & 0 & \cdots & 0 \\ \widehat{a}_{2,1} & \widehat{a}_{2,2} & & \cdots & \widehat{a}_{n,2} \\ 0 & & & & \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & \widehat{a}_{n,2} & & \cdots & \widehat{a}_{n,n} \end{bmatrix}$$

We use the Householder matrix

$$P_1 = I - 2w^{(2)}w^{(2)T}, \quad \left\| w^{(2)} \right\|_2 = 1$$

with

$$w^{(2)} = [0, v_1, ..., v_m]^T, \qquad m = n - 1$$

Then

$$P_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1-2v_1^2 & \cdots & -2v_1v_m \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -2v_1v_m & \cdots & 1-2v_m^2 \end{bmatrix}$$

Let $A = \left[A_{*,1}, ..., A_{*,n}\right]$. Then

$$P_1A = \left[P_1A_{*,1}, ..., P_1A_{*,n}\right]$$

The first column looks like

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1-2v_1^2 & \cdots & -2v_1v_m \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -2v_1v_m & \cdots & 1-2v_m^2 \end{bmatrix} \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{n,1} \end{bmatrix} = \begin{bmatrix} a_{1,1} \\ \widehat{a}_{2,1} \\ 0 \\ \vdots \end{bmatrix}$$

With reference to our earlier work on reducing a vector d to a simpler form with only a single nonzero component, we use that algorithm with

$$d = \left[a_{2,1}, ..., a_{n,1}\right]^T$$

With that algorithm, we can find $v \in \mathbb{R}^m$ with

$$\left(I - 2vv^T\right)d = \left[\widehat{a}_{2,1}, 0, ..., 0\right]^T$$

In fact,

$$\widehat{a}_{2,1} = -\,\mathsf{sign}\left(a_{2,1}\right)\mathsf{sqrt}\left(a_{2,1}^2 + \cdots + a_{n,1}^2\right)$$

Now look at $P_1 A$ and $P_1 A P_1$:

$$P_1 A = \begin{bmatrix} a_{1,1} & * & \cdots & * \\ \widehat{a}_{2,1} & * & \cdots & * \\ 0 & * & & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

$P_1 A P_1$ is given by

$$\begin{bmatrix} a_{1,1} & * & \cdots & * \\ \widehat{a}_{2,1} & * & \cdots & * \\ 0 & * & & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 - 2v_1^2 & \cdots & -2v_1 v_m \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -2v_1 v_m & \cdots & 1 - 2v_m^2 \end{bmatrix}$$

$$= \begin{bmatrix} a_{1,1} & * & \cdots & * \\ \widehat{a}_{2,1} & * & \cdots & * \\ 0 & * & & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

In addition, the matrix P_1AP_1 is symmetric:

$$(P_1AP_1)^T = P_1^T A^T P_1^T = P_1AP_1$$

Therefore, P_1AP_1 must look like

$$A_2 = P_1AP_1 = \begin{bmatrix} a_{1,1} & \widehat{a}_{2,1} & 0 & \cdots & 0 \\ \widehat{a}_{2,1} & * & & \cdots & * \\ 0 & * & & & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & * & & \cdots & * \end{bmatrix}$$

We continue this process by now reducing the second column by a similar operation. We use the Householder matrix

$$P_2 = I - 2w^{(3)}w^{(3)T}, \quad \left\| w^{(3)} \right\|_2 = 1$$

$$w^{(3)} = [0, 0, v_1, ..., v_m]^T, \qquad m = n - 2$$

Then

$$P_2 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 - 2v_1^2 & \cdots & -2v_1v_m \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & -2v_1v_m & \cdots & 1 - 2v_m^2 \end{bmatrix}$$

We choose $v \in \mathbb{R}^{n-2}$ such that

$$\left(I - 2vv^T\right) d = \left[\widehat{a}_{3,2}, 0, ..., 0\right]^T$$

with d the elements in positions 3 thru n of column 2 of the matrix A_2. Note that the form of P_2 is such that $P_2 A_2 P_2$ will have the same first column and row as in A_2. For example, consider calculating first $P_2 A_2$:

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 - 2v_1^2 & \cdots & -2v_1 v_m \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & -2v_1 v_m & \cdots & 1 - 2v_m^2 \end{bmatrix} \begin{bmatrix} a_{1,1} & \widehat{a}_{2,1} & 0 & \cdots & 0 \\ \widehat{a}_{2,1} & * & & \cdots & * \\ 0 & * & & & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & * & & \cdots & * \end{bmatrix}$$

We continue in this way, obtaining finally

$$P_{n-2} \cdots P_1 A P_1 \cdots P_{n-2} = T$$

with T a symmetric tridiagonal matrix. Define

$$Q = P_1 \cdots P_{n-2}$$

It is orthogonal, and

$$Q^T A Q = T$$

Therefore the eigenvalues of A and T are the same. For the eigenvectors, let

$$Tx = \lambda x, \quad x \neq 0$$

Then

$$\begin{aligned} Q^T A Q x &= \lambda x \\ A\,(Qx) &= \lambda\,(Qx) \end{aligned}$$

Thus Qx is the eigenvector of A corresponding to the eigenvector x for T.

<u>EXAMPLE</u>: (From page 617)

$$A = \begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 2 \\ 4 & 2 & 1 \end{bmatrix}, \quad w^{(2)} = \left[0, \frac{2}{\text{sqrt}(5)}, \frac{1}{\text{sqrt}(5)}\right]$$

$$P_1 = I - 2w^{(2)}w^{(2)T} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{3}{5} & -\frac{4}{5} \\ 0 & -\frac{4}{5} & \frac{3}{5} \end{bmatrix}$$

$$T = A_2 = \begin{bmatrix} 1 & -5 & 0 \\ -5 & \frac{73}{25} & -\frac{14}{25} \\ 0 & -\frac{14}{25} & -\frac{23}{25} \end{bmatrix}$$

When the rounding errors inherent in producing T from A are taken into account, how do the eigenvalues of T and A compare. This is given in the text as Theorem 9.4 (page 617). It is assume that the arithmetic being used is t-digit binary arithmetic with rounding; and moreover, it is assumed that all inner products

$$\sum_{j=1}^{m} a_j b_j$$

are accumulated in a higher precision and then rounded back to t digits at the completion of the summation process. With this, we obtain for the eigenvalues $\left\{\lambda_j\right\}$ and $\left\{\tau_j\right\}$ of A and T respectively, that

$$\left[\frac{\sum_{j=1}^{n} \left(\tau_j - \lambda_j\right)^2}{\sum_{j=1}^{n} \lambda_j^2}\right]^{\frac{1}{2}} \leq c_n 2^{-t}$$

$$c_n = 25\,(n-1)\left[1 + (12.36)\,2^{-t}\right]^{2n-4} \doteq 25\,(n-1)$$