Section 4.4 Reduction to Symmetric Tridiagonal Form

Key terms

Symmetric matrix
conditioning

Tridiagonal matrix

Similarity transformation

Orthogonal matrix

Orthogonal similarity transformation properties

Householder matrix
Here we will address the issue of computing all of the eigenvalues of a (real) symmetric matrix. We take this special case because the eigenvalues of symmetric matrices are well-conditioned (sensitivity to perturbations is small) whereas the eigenvalues of non-symmetric matrices can be poorly conditioned. In addition, an \( n \times n \) symmetric matrix always possesses \( n \) linearly independent eigenvectors whereas a non-symmetric matrix may not.

To compute all of the eigenvalues of a symmetric matrix, we will \textbf{proceed in two stages}. \textbf{First}, the matrix will be transformed to symmetric tridiagonal form. The \textbf{second} stage uses an iterative method to obtain a diagonal matrix from which we determine the eigenvalues of the original symmetric matrix.

\textbf{So why do we proceed in two stages?} Why don't we just perform the iterative technique on the original matrix? Simply put, the answer is efficiency. Transforming an \( n \times n \) symmetric matrix to symmetric tridiagonal form requires on the order of \((4/3)n^3\) arithmetic operations for large \( n \). The iterative reduction of the symmetric tridiagonal matrix to diagonal form then requires \( O(n^2) \) arithmetic operations. On the other hand, applying the iterative technique directly to the original matrix requires on the order of \((4/3)n^3\) arithmetic operations \textit{per iteration}. Thus, by first transforming the matrix to a simpler form we significantly reduce the computational cost.

\textbf{This is like a pre-conditioning procedure.}
Transforming a symmetric matrix to symmetric tridiagonal form is meaningless unless we have precise knowledge of how the eigenvalues have been affected by each transformation that has been performed. Fortunately, there is a class of transformations which does not change the spectrum (the set of all the eigenvalues) of a matrix. These are known as similarity transformations.

**Definition:** Let $A$ be an $n \times n$ matrix and let $M$ be any nonsingular $n \times n$ matrix. The matrix $B = M^{-1}AM$ is said to be similar to $A$. The process of converting $A$ to $B$ is called a similarity transformation.

**Theorem:** Similar matrices have the same eigenvalues.

Proof: Show that characteristic polynomials of $A$ and $B = M^{-1}AM$ are the same.

Although any nonsingular matrix can be used to generate a similarity transformation, we would like to use matrices whose inverses are easy to compute. The class of orthogonal matrices will suit our needs nicely.

**Definition:** The $n \times n$ matrix $Q$ is called an orthogonal matrix if $Q^{-1} = Q^T$. 

Example: Let $Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$. Direct multiplication shows that

$$QQ^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2}/2 & \sqrt{2}/2 \\ 0 & -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

so Q is an orthogonal matrix.

Aside from having an inverse matrix that is easy to compute, an orthogonal matrix has several other important properties.

The most important of these properties is related to the conditioning (sensitivity to perturbations) of the eigenvalue problem. As noted previously, the eigenvalues of symmetric matrices are well conditioned. Suppose then that A is a symmetric matrix and $B = Q^{-1}AQ = Q^TAQ$ for some orthogonal matrix Q. It follows that

$$B^T = (Q^TAQ)^T = Q^TAQ = B.$$ 

Hence, a similarity transformation with an orthogonal matrix maintains symmetry and therefore preserves the conditioning of the original eigenvalue problem.

**Theorem:** If $Q$ is an $n \times n$ orthogonal matrix and is an $n$-vector, then $\|Qx\|_2 = \|x\|_2$.

So multiplication of a vector my an orthogonal matrix preserves “length”.
Reducing a Symmetric Matrix to Tridiagonal Form

There are several different algorithms available for reducing a symmetric matrix to tridiagonal form. Most work in a sequential manner, applying a succession of similarity transformations which gradually produce the desired form. These techniques differ only in the family of orthogonal matrices used to generate the similarity transformations.

Here, we will restrict our attention to a reduction algorithm based on the use of Householder matrices.

**Definition:** A **Householder matrix** is any matrix of the form $H = I - 2ww^T$ where $w$ is a column vector with $w^Tw = 1$. (Recall that $\|w\|_2 = \sqrt{w^Tw}$. So $w$ is a unit vector.)

Alston Householder (1904 – 1993): Recognized for his impact and influence on computer science in general and particularly for his contributions to the methods and techniques for obtaining numerical solutions to very large problems through the use of digital computers, and for his many publications, including books, which have provided guidance and help to workers in the field of numerical analysis, and for his contributions to professional activities and societies as committee member, paper referee, and conference organizer.

[http://www-history.mcs.st-and.ac.uk/Biographies/Householder.html](http://www-history.mcs.st-and.ac.uk/Biographies/Householder.html)
Properties of Householder matrices $H = I - 2ww^T$:

- $H$ is symmetric and orthogonal.

- Geometrically, multiplication of a vector $x$ by the Householder matrix $H$ results in the reflection of $x$ across the hyperplane whose normal vector is $w$. (The figure depicts this in 3-dimensional space.)

In practice, the Householder matrices are not computed explicitly, only the vector $w$ is computed. For once the vector $w$ is known the similarity transformation $HAH$ is given by

$$\begin{align*}
(I - 2ww^T)A(I - 2ww^T) &= A - 2ww^TA - 2Aww^T + 4ww^TAww^T,
\end{align*}$$

which is completely determined by $w$.

The computation of $HAH$ can be simplified tremendously if we define $u = Aw$ and $K = w^Tu = w^TAw$. Then we have that

$$\begin{align*}
HAH &= A - 2ww^TA - 2Aww^T + 4ww^TAww^T \\
&= A - 2wu^T - 2uw^T + 4Kww^T \\
&= A - 2w(u^T - Kw) - 2(u - Kw)w^T.
\end{align*}$$

If we now let $q = u - Kw$, then $HAH = A - 2qw^T - 2qw^T$. 

The algorithm to reduce a symmetric matrix to tridiagonal form using Householder matrices involves a sequence of \( n - 2 \) similarity transformations, as illustrated in the Figure for the case \( n = 5 \).

The first Householder matrix, \( H_1 \), is selected so that \( H_1 A \) will have zeros in the first \( n - 2 \) rows of the \( n \)-th column and the \( n \)-th row of \( A \) will not be affected. However, by symmetry, when \( H_1 A H_1 \) is computed to complete the transformation, the zeros in the \( n \)-th column will not be changed, but zeros will appear in the first \( n - 2 \) columns of the \( n \)-th row.

Each subsequent Householder matrix, \( H_i \) (\( i = 2, 3, 4, ..., n - 2 \)), is then selected so that

\[
H_i H_{i-1} ... H_2 H_1 A H_1 H_2 ... H_{i-1}
\]

will have zeros in the first \( n - i - 1 \) rows of the \((n - i + 1)\)-st column, but will not affect the bottom \( i \) rows. Completing the \( i \)-th transformation will place zeros in the first \( n - i - 1 \) columns of the \((n - i + 1)\)-st row.

Note \( H_1 A \) puts zero in a portion of the last column while \((H_1 A)H_1\) puts zero in the same size portion of the last row. The pattern continues. ETC.
Determining the appropriate Householder matrix for use in each step of the algorithm requires the solution of the following fundamental problem:

Given an integer \( k \) and an \( n \)-dimensional column vector \( x \), select \( w \) so that \( Hx = (I - 2ww^T)x \) has zeros in the first \( n - k - 1 \) rows, but leaves the last \( k \) elements in \( x \) unchanged.

Note that this problem specification contains only \( n - 1 \) conditions on the vector \( w \). The last condition comes from the requirement that \( w^Tw = 1 \), or equivalently, that the vector \((I - 2ww^T)x\) have the same Euclidean norm as the vector \( x \).

To solve this problem, we first note that in order for the last \( k \) elements in \( x \) to be unchanged, the last \( k \) elements in \( w \) must be zero. This guarantees that the last \( k \) rows and columns of \( H \) are identical to those in the identity matrix. Thus vector \( w \) must be of the form

\[
w = \begin{bmatrix} w_1 & w_2 & w_3 & \cdots & w_{n-k} & 0 & \cdots & 0 \end{bmatrix}^T.
\]

Let \( b = (I - 2ww^T)x \), where by construction \( b \) (the first \( n - k - 1 \) rows are zero) will have the form

\[
b = \begin{bmatrix} 0 & \cdots & 0 & \alpha & x_{n-k+1} & \cdots & x_n \end{bmatrix}^T,
\]

with \( n - k - 1 \) zeros. To be determined is \( \alpha \) and the values of \( x_{n-k+1}, \ldots, x_n \).
\[ b = (I - 2ww^T)x \]

Since multiplication by the Householder matrix must preserve the Euclidean norm, we must have \( b^T b = x^T x \), which implies that

\[ \alpha^2 = x_1^2 + x_2^2 + x_3^2 + \cdots + x_{n-k}^2. \]

To proceed further, let's rearrange the equation defining the vector \( b \) as

Pre-multiplying equation (1) by \( w^T \) yields

\[ w^T x - 2w^T ww^T x = w^T b, \]

which simplifies to

\[ -w^T x = \alpha w_{n-k} \]

upon taking into account the form of both \( w \) and \( b \) and using the fact that \( w^T w = 1 \).

Substituting equation (2) into equation (1) produces

\[ x + 2\alpha w_{n-k} w = b \]

To determine the entries in \( b \) we look at the expressions for the entries of the left side.
In terms of the entries we have

\[ \begin{align*}
  x_i + 2\alpha w_{n-k} w_i &= 0 \\
  x_{n-k} + 2\alpha w^2_{n-k} &= \alpha.
\end{align*} \quad (i = 1, 2, 3, \ldots, n - k - 1) \]

Now use the expression for the last entry to solve for \( w_{n-k} \) and we get

\[ w_{n-k} = \sqrt{\frac{1}{2} \left( 1 - \frac{x_{n-k}}{\alpha} \right)}. \]

\[ \text{This changes as } k \text{ changes.} \]

To avoid cancellation error, we will choose \( \text{sgn}(\alpha) = -\text{sgn}(x_{n-k}) \). With \( w_{n-k} \) determined, the remaining non-zero entries in \( w \) are given by

\[ w_i = -\frac{1}{2} \frac{x_i}{\alpha w_{n-k}} \quad (i = 1, 2, 3, \ldots, n - k - 1). \]

\[ \text{Note: we need, } x_i, w_{n-k}, \text{ and } \alpha. \]

So here we can compute the \( n - k \) entries of \( w \) for the form

\[ w = \begin{bmatrix} w_1 & w_2 & w_3 & \cdots & w_{n-k} & 0 & \cdots & 0 \end{bmatrix}^T. \]

Once we program this we need not proceed step-by-step as we did in the development.
Example: Reduction to Tridiagonal Form

We want to reduce that symmetric matrix $A$ to tridiagonal form.

Since $A$ is $4 \times 4$ we will need two Householder matrices.

To start let $x = \text{col}_4(A)$; \[ x = \begin{bmatrix} 2 \\ -2 \\ 1 \\ 4 \end{bmatrix} \]

**Step 1.** Compute $\alpha$.

$\alpha^2 = (x_1)^2 + (x_2)^2 + (x_3)^2 = 4 + 4 + 1 = 9$

Since sign$(x_3) = +1$, so $\alpha = -3$.

**Step 2.** We want to produce zeros in the first two rows of the last column of $A$ and leave the last element in that column alone. Therefore, we are working with $k = 1$. So for $n = 4$ and $k = 1$ we have $n - k = 3$; \[ w_3 = \sqrt{\frac{1}{2} \left( 1 - \frac{x_3}{\alpha} \right)} = \sqrt{\frac{1}{2} \left( 1 - \frac{1}{-3} \right)} = \frac{\sqrt{6}}{3} \]

Then \[ w_2 = \frac{-1}{2} \frac{x_2}{\alpha w_3} = \frac{-1}{2} \frac{-2}{-3 \left( \frac{\sqrt{6}}{3} \right)} = \frac{-\sqrt{6}}{6} \]

and \[ w_1 = \frac{-1}{2} \frac{x_1}{\alpha w_3} = \frac{-1}{2} \frac{2}{-3 \left( \frac{\sqrt{6}}{3} \right)} = \frac{\sqrt{6}}{6} \]

Then \[ w = \frac{\sqrt{6}}{6} \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix} \]

Note: $k$ refers to the number of the Householder matrix.
Now with $w = \frac{\sqrt{6}}{6} \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}$ we can construct Householder matrix $H_1$. Using the previously developed set of computations $u = Aw$, $K = w^T u = w^T A w$, and $q = u - Kw$ we have $H_1 AH_1 = A - 2wq^T - 2qw^T$:

$$u = Aw = (\sqrt{6}/6) \begin{bmatrix} 3 & -5 & 5 & 6 \end{bmatrix}^T$$

$$K = w^T u = 3; \text{ and}$$

$$q = u - Kw = (\sqrt{6}/6) \begin{bmatrix} 0 & -2 & -1 & 6 \end{bmatrix}^T.$$
Next we want to produce a zero in the first row of the third column of $H_1A H_1$ which is the vector.

**Step 1.** Compute $\alpha$. 

$$\alpha^2 = (x_1)^2 + (x_2)^2 = \frac{16}{9} + 1 = \frac{25}{9}$$

Since $\text{sign}(x_2) = +1$, so $\alpha = -\frac{5}{3}$.

**Step 2.** We want to produce a zero in the first row of the third column of $H_1A H_1$ and leave the last two elements in that column alone. Therefore, we are working with $k = 2$. So for $n = 4$ and $k = 2$ we have $n - k = 2$;

$$w_2 = \sqrt{\frac{1}{2} \left( 1 - \frac{x_2}{\alpha} \right)} = \sqrt{\frac{1}{2} \left( 1 - \frac{1}{-5/3} \right)} = \frac{2\sqrt{5}}{5}$$

$$w_1 = \frac{-1}{2 \alpha w_2} \frac{x_1}{\alpha} = \frac{-1}{2} \frac{4/3}{(-5/3)(2\sqrt{5}/5)} = \frac{\sqrt{5}}{5}$$

Hence

$$w = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

$$w_i = -\frac{1}{2 \alpha w_n-k} x_i$$
Step 3. Using $w$ we construct Householder matrix $H_2$ computed from $H_1 A H_1 - 2w q^T - 2q w^T$.

$$u = A w = (\sqrt{5}/5) \begin{bmatrix} -11/3 & 2 & 10/3 & 0 \end{bmatrix}^T;$$

$$K = w^T u = 1/15; \text{ and}$$

$$q = u - K w = (\sqrt{5}/5) \begin{bmatrix} -56/15 & 28/15 & 10/3 & 0 \end{bmatrix}^T$$

Thus we have reduced $A$ to this tridiagonal matrix.

$$H_2 H_1 A H_1 H_2 = \begin{bmatrix}
149/75 & 68/75 & 0 & 0 \\
68/75 & -33/25 & -5/3 & 0 \\
0 & -5/3 & 10/3 & -3 \\
0 & 0 & -3 & 4 \\
\end{bmatrix}$$

For comparison:

$$\gg (\text{eig}(A_{\text{new}}) - \text{eig}(A))^T$$

ans = $1.0e-14 \times [0.3109, -0.1110, -0.0444, -0.0888]$
Obtaining the Eigenvectors of a Symmetric Matrix

We need a bit of background material here. We know that similar matrices have the same eigenvalues, but they (generally) do not have the eigenvectors. Let \((\lambda, \mathbf{x})\) be an eigenpair of matrix \(\mathbf{A}\), then \(\mathbf{Ax} = \lambda \mathbf{x}\). Suppose \(\mathbf{B} = \mathbf{PAP}^{-1}\) then \(\mathbf{A} = \mathbf{P}^{-1}\mathbf{BP}\) and \(\lambda \mathbf{x} = \mathbf{Ax} = \mathbf{P}^{-1}\mathbf{BPx}\). Multiplying both side by \(\mathbf{P}\) we get \(\lambda \mathbf{Px} = \mathbf{BPx}\), thus \((\lambda, \mathbf{Px})\) is an eigenpair of \(\mathbf{B}\).

Suppose that \(\mathbf{A}\) is symmetric and \(H_n H_{n-1} ... H_2 H_1 A H_1 H_2 ... H_{n-1} H_n\) is symmetric tridiagonal. In addition suppose we use the second stage to find all the eigenvalues of the symmetric tridiagonal matrix and matrix \(\mathbf{M}\) is the product of the matrices needed for the similarity transformation of the symmetric tridiagonal to diagonal matrix \(\mathbf{D}\). Then we have \(\mathbf{M}^{-1}H_n H_{n-1} ... H_2 H_1 A H_1 H_2 ... H_{n-1} H_n \mathbf{M} = \mathbf{D}\). Since we are using similarity transformation the diagonal entries of \(\mathbf{D}\) are eigenvalues of the symmetric tridiagonal matrix and of the original symmetric matrix \(\mathbf{A}\). Let the diagonal entries of \(\mathbf{D}\) be denoted, in order, by \(\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n\).

It follows that the eigenvectors of \(\mathbf{A}\) can computed from the eigen vectors of \(\mathbf{D}\), which are just the columns of the identity matrix. Thus the eigenvector of \(\mathbf{A}\) corresponding to eigenvalue \(\lambda_i\) is \(\mathbf{v}_i = H_1 H_2 ... H_{n-1} H_n \mathbf{M} \mathbf{e}_i\) where \(\mathbf{e}_i\) is column \(i\) of the identity matrix.

So while computing all of the eigenvalues of a symmetric matrix, it is possible to simultaneously compute the corresponding eigenvectors.
Let $X$ be a 20-by-20 random symmetric matrix, with colors representing different values we can see the process of the Householder matrices being applied one by one to obtain a tridiagonal matrix.

Go to the next slide.

http://fa.bianp.net/blog/2013/householder-matrices/
Click on picture to animate.

\[ P_0 X P_0 \]
Final state: Symmetric tridiagonal matrix.