The Eigenvalue Problem: The Hessenberg and Real Schur Forms

The Unsymmetric Eigenvalue Problem

Let $A$ be a real $n \times n$ matrix. It is possible that $A$ has complex eigenvalues, which must occur in complex-conjugate pairs, meaning that if $a + ib$ is an eigenvalue, where $a$ and $b$ are real, then so is $a - ib$. On the one hand, it is preferable that complex arithmetic be avoided as much as possible when using $QR$ iteration to obtain the Schur Decomposition of $A$. On the other hand, in the algorithm for $QR$ iteration, if the matrix $Q_0$ used to compute $T_0 = Q_0^H A Q_0$ is real, then every matrix $T_k$ generated by the iteration will also be real, so it will not be possible to obtain the Schur Decomposition.

We compromise by instead seeking to compute the Real Schur Decomposition $A = QTQ^T$ where $Q$ is a real, orthogonal matrix and $T$ is a real, quasi-upper-triangular matrix that has a block upper-triangular structure

$$T = \begin{bmatrix}
T_{11} & T_{12} & \cdots & T_{1p} \\
0 & T_{22} & \ddots & \vdots \\
0 & \ddots & \ddots & \vdots \\
0 & 0 & 0 & T_{pp}
\end{bmatrix},$$

where each diagonal block $T_{ii}$ is $1 \times 1$, corresponding to a real eigenvalue, or a $2 \times 2$ block, corresponding to a pair of complex eigenvalues that are conjugates of one another.

If $QR$ iteration is applied to such a matrix, then the sequence $\{T_k\}$ will not converge, but a block upper-triangular structure will be obtained, which can then be used to compute all of the eigenvalues. Therefore, the iteration can be terminated when appropriate entries below the diagonal have been made sufficiently small.

However, one significant drawback to the $QR$ iteration is that each iteration is too expensive, as it requires $O(n^3)$ operations to compute the $QR$ factorization, and to multiply the factors in reverse order. Therefore, it is desirable to first use a similarity transformation $H = U^T A U$ to reduce $A$ to a form for which the $QR$ factorization and matrix multiplication can be performed more efficiently.

Suppose that $U^T$ includes a Householder reflection, or a product of Givens rotations, that transforms the first column of $A$ to a multiple of $e_1$, as in algorithms to compute the $QR$ factorization. Then $U$ operates on all rows of $A$, so when $U$ is applied to the columns of $A$, to complete the
similarity transformation, it affects all columns. Therefore, the work of zeroing the elements of the first column of $A$ is undone.

Now, suppose that instead, $U^T$ is designed to zero all elements of the first column except the first two. Then, $U^T$ affects all rows except the first, meaning that when $U^TA$ is multiplied by $U$ on the right, the first column is unaffected. Continuing this reasoning with subsequent columns of $A$, we see that a sequence of orthogonal transformations can be used to reduce $A$ to an upper Hessenberg matrix $H$, in which $h_{ij} = 0$ whenever $i > j + 1$. That is, all entries below the subdiagonal are equal to zero.

It is particularly efficient to compute the $QR$ factorization of an upper Hessenberg, or simply Hessenberg, matrix, because it is only necessary to zero one element in each column. Therefore, it can be accomplished with a sequence of $O(n^2)$ operations. Then, these same Givens rotations can be applied, in the same order, to the columns in order to complete the similarity transformation, or, equivalently, accomplish the task of multiplying the factors of the $QR$ factorization.

Specifically, given a Hessenberg matrix $H$, we apply Givens row rotations $G_1^T, G_2^T, \ldots, G_{n-1}^T$ to $H$, where $G_i^T$ rotates rows $i$ and $i + 1$, to obtain

$$G_{n-1}^T \cdots G_2^T G_1^T H = (G_1 G_2 \cdots G_{n-1})^T H = Q^T H = R,$$

where $R$ is upper-triangular. Then, we compute

$$\tilde{H} = Q^T HQ = RQ = RG_1 G_2 \cdots G_{n-1}$$

by applying column rotations to $R$, to obtain a new matrix $\tilde{H}$.

By considering which rows or columns the Givens rotations affect, it can be shown that $Q$ is Hessenberg, and therefore $\tilde{H}$ is Hessenberg as well. The process of applying an orthogonal similarity transformation to a Hessenberg matrix to obtain a new Hessenberg matrix with the same eigenvalues that, hopefully, is closer to quasi-upper-triangular form is called a Hessenberg $QR$ step. The following algorithm overwrites $H$ with $\tilde{H} = RQ = Q^T HQ$, and also computes $Q$ as a product of Givens column rotations, which is only necessary if the full Schur Decomposition of $A$ is required, as opposed to only the eigenvalues.

```latex
\begin{verbatim}
for j = 1, 2, \ldots, n - 1 do
    [c, s] = givens(h_{jj}, h_{j+1,j})
    G_j = \begin{bmatrix}
        c & -s \\
        s & c 
    \end{bmatrix}
    H(j : j + 1, j : n) = G_j^T H(j : j + 1, :)
end
Q = I
for j = 1, 2, \ldots, n - 1 do
    H(1 : j + 1, j : j + 1) = H(1 : j + 1, j : j + 1)G_j
\end{verbatim}
```
\[ Q(1 : j + 1, j + j + 1) = Q(1 : j + 1, j : j + 1)G_j \]

end

The function \texttt{givens}(a, b) returns \( c \) and \( s \) such that
\[
\begin{bmatrix}
  c & -s \\
  s & c
\end{bmatrix}^T
\begin{bmatrix}
  a \\
  b
\end{bmatrix} =
\begin{bmatrix}
  r \\
  0
\end{bmatrix},
\quad r = \sqrt{a^2 + b^2}.
\]

Note that when performing row rotations, it is only necessary to update certain columns, and when performing column rotations, it is only necessary to update certain rows, because of the structure of the matrix at the time the rotation is performed; for example, after the first loop, \( H \) is upper-triangular.

Before a Hessenberg QR step can be performed, it is necessary to actually reduce the original matrix \( A \) to Hessenberg form \( H = U^T A U \). This can be accomplished by performing a sequence of Householder reflections \( U = P_1 P_2 \cdots P_{n-2} \) on the columns of \( A \), as in the following algorithm.

\[
U = I
\]

for \( j = 1, 2, \ldots, n - 2 \) do

\[
\begin{align*}
  v &= \texttt{house}(A(j + 1 : n, j)),
  c = 2/v^T v \\
  A(j + 1 : n, j : n) &= A(j + 1 : n, j : n) - cvv^T A(j + 1 : n, j : n) \\
  A(1 : n, j + 1 : n) &= A(1 : n, j + 1 : n) - cA(1 : n, j + 1 : n)vv^T
\end{align*}
\]

end

The function \texttt{house}(x) computes a vector \( v \) such that \( P x = I - cvv^T x = \alpha e_1 \), where \( c = 2/v^T v \) and \( \alpha = \pm \|x\|_2 \). The algorithm for the Hessenberg reduction requires \( O(n^3) \) operations, but it is performed only once, before the QR Iteration begins, so it still leads to a substantial reduction in the total number of operations that must be performed to compute the Schur Decomposition.

If a subdiagonal entry \( h_{j+1,j} \) of a Hessenberg matrix \( H \) is equal to zero, then the problem of computing the eigenvalues of \( H \) decouples into two smaller problems of computing the eigenvalues of \( H_{11} \) and \( H_{22} \), where
\[
H = \begin{bmatrix}
  H_{11} & H_{12} \\
  0 & H_{22}
\end{bmatrix}
\]
and \( H_{11} \) is \( j \times j \). Therefore, an efficient implementation of the QR Iteration on a Hessenberg matrix \( H \) focuses on a submatrix of \( H \) that is un\textit{reduced}, meaning that all of its subdiagonal entries are nonzero. It is also important to monitor the subdiagonal entries after each iteration, to determine if any of them have become nearly zero, thus allowing further decoupling. Once no further decoupling is possible, \( H \) has been reduced to quasi-upper-triangular form and the QR Iteration can terminate.

It is essential to choose an \textit{maximal} un\textit{reduced} diagonal block of \( H \) for applying a Hessenberg QR step. That is, the step must be applied to a submatrix \( H_{22} \) such that \( H \) has the structure
\[
H = \begin{bmatrix}
  H_{11} & H_{12} & H_{13} \\
  0 & H_{22} & H_{23} \\
  0 & 0 & H_{33}
\end{bmatrix}
\]
where $H_{22}$ is unreduced. This condition ensures that the eigenvalues of $H_{22}$ are also eigenvalues of $H$, as $\lambda(H) = \lambda(H_{11}) \cup \lambda(H_{22}) \cup \lambda(H_{33})$ when $H$ is structured as above. Note that the size of either $H_{11}$ or $H_{33}$ may be $0 \times 0$.

The following property of unreduced Hessenberg matrices is useful for improving the efficiency of a Hessenberg QR step.

**Theorem (Implicit Q Theorem)** Let $A$ be an $n \times n$ matrix, and let $Q$ and $P$ be $n \times n$ orthogonal matrices such that $Q^T A Q = H$ and $V^T A V = G$ are both upper Hessenberg, and $H$ is unreduced. If $Q = [\mathbf{q}_1 \cdots \mathbf{q}_n]$ and $V = [\mathbf{v}_1 \cdots \mathbf{v}_n]$, and if $\mathbf{q}_1 = \mathbf{v}_1$, then $\mathbf{q}_i = \pm \mathbf{v}_i$ for $i = 2, \ldots, n$, and $|h_{ij}| = |g_{ij}|$ for $i, j = 1, 2, \ldots, n$.

That is, if two orthogonal similarity transformations that reduce $A$ to Hessenberg form have the same first column, then they are “essentially equal”, as are the Hessenberg matrices.

Another important property of an unreduced Hessenberg matrix is that all of its eigenvalues have a geometric multiplicity of one. To see this, consider the matrix $H - \lambda I$, where $H$ is an $n \times n$ unreduced Hessenberg matrix and $\lambda$ is an arbitrary scalar. If $\lambda$ is not an eigenvalue of $H$, then $H$ is nonsingular and rank($H$) = $n$. Otherwise, because $H$ is unreduced, from the structure of $H$ it can be seen that the first $n - 1$ columns of $H - \lambda I$ must be linearly independent. We conclude that rank($H - \lambda I$) = $n - 1$, and therefore at most one vector $x$ (up to a scalar multiple) satisfies the equation $Hx = \lambda x$. That is, there can only be one linearly independent eigenvector. It follows that if any eigenvalue of $H$ repeats, then it is defective.

**The Symmetric Eigenvalue Problem**

A symmetric Hessenberg matrix is tridiagonal. Therefore, the same kind of Householder reflections that can be used to reduce a general matrix to Hessenberg form can be used to reduce a symmetric matrix $A$ to a tridiagonal matrix $T$. However, the symmetry of $A$ can be exploited to reduce the number of operations needed to apply each Householder reflection on the left and right of $A$.

It can be verified by examining the structure of the matrices involved, and the rows and columns influenced by Givens rotations, that if $T$ is a symmetric tridiagonal matrix, and $T = QR$ is its QR factorization, then $Q$ is upper Hessenberg, and $R$ is upper-bidiagonal (meaning that it is upper-triangular, with upper bandwidth 1, so that all entries below the main diagonal and above the superdiagonal are zero). Furthermore, $\tilde{T} = RQ$ is also tridiagonal.

Because each Givens rotation only affects $O(1)$ nonzero elements of a tridiagonal matrix $T$, it follows that it only takes $O(n)$ operations to compute the QR factorization of a tridiagonal matrix, and to multiply the factors in reverse order. However, to compute the eigenvectors of $A$ as well as the eigenvalues, it is necessary to compute the product of all of the Givens rotations, which still takes $O(n^2)$ operations.

The Implicit Q Theorem applies to symmetric matrices as well, meaning that if two orthogonal similarity transformations reduce a matrix $A$ to unreduced tridiagonal form, and they have the same first column, then they are essentially equal, as are the tridiagonal matrices that they produce.