Dr A. Iserles
Easter 1997

Numerical Analysis – Lecture 6

The Givens algorithm Given $m \times n$ matrix A, let ℓ_i be the number of leading zeros in the ith row of A, $i = 1, 2, \ldots, m$.

Step 1 Stop if the (integer) sequence $\{\ell_1, \ell_2, \ldots, \ell_m\}$ increases monotonically, the increase being strictly monotone for $\ell_i \le n$.

Step 2 Pick any two integers $1 \le p < q \le m$ such that either $\ell_p > \ell_q$ or $\ell_p = \ell_q < n$.

Step 3 Replace A by $\Omega^{[p,q]}$, a Givens rotation that annihilates the $(q, \ell_q + 1)$ element.

Update the values of ℓ_p and ℓ_q and go to *Step 1*.

The final matrix A is upper triangular and also has the property that the number of leading zeros in each row increases *strictly monotonically* until all the rows of A are zero – a matrix of this form is said to be in *standard form*. This end result, as we recall, is the required matrix R.

The cost There are less than mn rotations and each rotation replaces two rows by their linear combinations, hence the total cost is $\mathcal{O}(mn^2)$.

If we wish to obtain explicitly an orthogonal Q s.t. $A = QR$ then we commence by letting Ω be the $m \times m$ unit matrix and, each time A is premultiplied by $\Omega^{[p,q]}$, we also premultiply Ω by the same rotation. Hence the final Ω is the product of all the rotations, in correct order, and we let $Q = \Omega^{\mathrm{T}}$. The extra cost is $\mathcal{O}(m^2 n)$. However, in most applications we don't need Q but, instead, just the action of Q^{T} on a given vector (recall: solution of linear systems). This can be accomplished by multiplying the vector by successive rotations, the cost being $\mathcal{O}(mn)$.

4.6 Householder rotations

Let $\boldsymbol{u} \in \mathbb{R}^m \setminus \{\boldsymbol{0}\}$. The $m \times m$ matrix $I - 2\frac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|^2}$ is called a *Householder rotation*. Each such matrix is symmetric (that's trivial) and orthogonal, since

$$\left(I - 2\frac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|^2}\right)^{\mathrm{T}} \left(I - 2\frac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|^2}\right) = \left(I - 2\frac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|^2}\right)^2 = I - 4\frac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|^2} + 4\frac{\boldsymbol{u}(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u})\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|^4} = I.$$

Deriving the first column of R Our goal is to multiply an $m \times n$ matrix A by a sequence of Householder rotations so that each product 'peels off' requisite nonzero elements in a whole successive column. To start with, we seek a rotation that transforms the first nonzero column of A to a multiple of $\boldsymbol{e}_1$.

Let $\boldsymbol{a} \in \mathbb{R}^m$ be the first nonzero column of A. We wish to choose $\boldsymbol{u} \in \mathbb{R}^n$ s.t. the last $m-1$ entries of $\left(I - 2\frac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|^2}\right)\boldsymbol{a} = \boldsymbol{a} - 2\frac{\boldsymbol{u}^{\mathrm{T}}\boldsymbol{a}}{\|\boldsymbol{u}\|^2}\boldsymbol{u}$ vanish and, in addition, we normalize $\boldsymbol{u}$ so that $2\boldsymbol{u}^{\mathrm{T}}\boldsymbol{a} = \|\boldsymbol{u}\|^2$ (recall that $\boldsymbol{a} \neq \boldsymbol{0}$). Therefore $u_i = a_i$, $i = 2, \ldots, m$ and the normalization implies that

$$2u_1 a_1 + 2\sum_{i=2}^{m} a_i^2 = u_1^2 + \sum_{i=2}^{m} a_i^2 \quad \Rightarrow \quad u_1^2 - 2u_1 a_1 + a_1^2 - \sum_{i=1}^{m} a_i^2 = 0 \quad \Rightarrow \quad u_1 = a_1 \pm \|\boldsymbol{a}\|.$$

It is usual to let the sign be the same as the sign of a_1.

For large m we do not execute explicit matrix multiplication. Instead, to calculate $\left(I - 2\frac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|^2}\right)A = A - 2\frac{\boldsymbol{u}(\boldsymbol{u}^{\mathrm{T}}A)}{\|\boldsymbol{u}\|^2}$, first evaluate $\boldsymbol{w}^{\mathrm{T}} := \boldsymbol{u}^{\mathrm{T}}A$, subsequently forming $A - \frac{2}{\|\boldsymbol{u}\|^2}\boldsymbol{u}\boldsymbol{w}^{\mathrm{T}}$.

Next columns of R Suppose that $\boldsymbol{a}$ is the first column of A that isn't compatible with standard form (previous columns have been, presumably, already dealt with by Householder rotations) and that the standard form requires to bring the $k+1,\ldots,m$ components to zero. Hence, nonzero elements in previous columns must be confined to the first $k-1$ rows and we want them to be unamended by the rotation. Thus, we let the first $k-1$ components of $\boldsymbol{u}$ be zero and choose $u_k = a_k \pm \left(\sum_{i=k}^m a_i^2\right)^{1/2}$ and $u_i = a_i$, $i = k+1,\ldots,m$.

The Householder method We process columns of A in sequence, in each stage premultiplying a current A by the requisite Householder transformation. The end result is an upper triangular matrix R in its standard form.

Example

$$A = \begin{bmatrix} 2 & 4 & 7 \\ 0 & 3 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & -2 \end{bmatrix} \quad \Rightarrow \quad \boldsymbol{u} = \begin{bmatrix} 0 \\ 0 \\ 5 \\ 1 \\ -2 \end{bmatrix} \quad \Rightarrow \quad \left(I - 2\frac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|^2}\right) A = \begin{bmatrix} 2 & 4 & 7 \\ 0 & 3 & -1 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Calculation of Q If the matrix Q is required in an explicit form, set $\Omega = I$ initially and, for each successive rotation, replace Ω by

$$\left(I - 2\frac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|^2}\right)\Omega = \Omega - \frac{2}{\|\boldsymbol{u}\|^2}\boldsymbol{u}(\boldsymbol{u}^{\mathrm{T}}\Omega).$$

As in the case of Givens rotations, by the end of the computation, $Q = \Omega^{\mathrm{T}}$. However, if we require just the vector $\boldsymbol{c} = Q^{\mathrm{T}}\boldsymbol{b}$, say, rather than the matrix Q, then we set initially $\boldsymbol{c} = \boldsymbol{b}$ and in each stage replace $\boldsymbol{c}$ by

$$\left(I - 2\frac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|^2}\right)\boldsymbol{c} = \boldsymbol{c} - 2\frac{\boldsymbol{u}^{\mathrm{T}}\boldsymbol{c}}{\|\boldsymbol{u}\|^2}\boldsymbol{u}.$$

Deciding between Givens and Householder rotations If A is dense, it is in general more convenient to use Householder rotations. Givens rotations come into their own, however, when A has many leading zeros in its rows. In an exterme case, if an $n \times n$ matrix A consists of zeros underneath the first subdiagonal, they can be 'rotated away' in $n-1$ Givens rotations, at the cost of just $\mathcal{O}(n^2)$ operations!