Mathematical Tripos Part IB: Easter 2001

Numerical Analysis – Lecture 6[1]

The Givens algorithm. Given an $m \times n$ matrix A, one uses the Givens rotations,

$$\Omega^{[1,2]}, \ldots, \Omega^{[1,m]}, \qquad \Omega^{[2,3]}, \ldots, \Omega^{[2,m]}, \qquad \Omega^{[3,4]}, \ldots, \Omega^{[3,m]}, \qquad \ldots$$

and so on, to annihilate the subdiagonal elements of the resulting matrices in each column successively. The final matrix is upper triangular, i.e., 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 $\mathcal{O}\big(mn^2\big)$.

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 $\mathcal{O}\big(m^2 n\big)$. 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 transformations

Let $\boldsymbol{u} \in \mathbb{R}^m \setminus \{\boldsymbol{0}\}$. The $m \times m$ matrix

$$H = I - 2\frac{\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}}{\|\boldsymbol{u}\|^2}$$

is called a *Householder transformation* or a *Householder reflection*. Since

$$H\boldsymbol{u} = -\boldsymbol{u}, \qquad H\boldsymbol{v} = \boldsymbol{v} \quad \text{if} \quad \boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = 0,$$

this transformation reflects any vector $\boldsymbol{x} \in \mathbb{R}^n$ in the $(n-1)$-dimensional hyperplane spanned by the vectors orthogonal to $\boldsymbol{u}$.

Each such matrix H is symmetric and orthogonal. The latter follows because reflection leaves the Eucleadian distance invariant, or by direct calculation

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

Householder transformations offer an alternative to Given rotations in the calculation of a QR factorization.

Deriving the first column of $\boldsymbol{R}$. Our goal is to multiply an $m \times n$ matrix A by a sequence of Householder transformations so that each product induces zeros under the diagonal in an entire successive column. To start with, we seek a reflection 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 $\boldsymbol{u} \in \mathbb{R}^m$ s.t. the bottom $m-1$ entries of

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

[1]Corrections and suggestions to these notes should be emailed to a.shadrin@damtp.cam.ac.uk. All handouts are available on the WWW at the URL http://www.damtp.cam.ac.uk/user/na/PartIB/Handouts.html.

vanish and, in addition, we normalise $\boldsymbol{u}$ 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 normalisation implies that

$$2u_1a_1 + 2\sum_{i=2}^{m} a_i^2 = u_1^2 + \sum_{i=2}^{m} a_i^2 \quad \Rightarrow \quad u_1^2 - 2u_1a_1 + a_1^2 - \sum_{i=1}^{m} a_i^2 = 0 \quad \Rightarrow \quad u_1 = a_1 \pm \|\boldsymbol{a}\|.$$

It is usual to let the sign be the same as the sign of a_1, since $\|\boldsymbol{u}\| \ll 1$ might lead to a division by a tiny number, hence to numerical difficulties.

For large m we do not execute explicit matrix multiplication. Instead, to calculate

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

we first evaluate $\boldsymbol{w}^{\mathrm{T}} := \boldsymbol{u}^{\mathrm{T}}A$, subsequently forming $A - \frac{2}{\|\boldsymbol{u}\|^2}\boldsymbol{u}\boldsymbol{w}^{\mathrm{T}}$.

Subsequent columns of $\boldsymbol{R}$. Suppose that the Householder transformation has been applied $k-1$ times, so that the first $k-1$ columns of the resulting matrix $\tilde{A}$ have an upper triangular form. Let $\boldsymbol{a}$ be the k-th column of $\tilde{A}$. By the next reflection we want annihilate the $k+1, \ldots, m$ components $\boldsymbol{a}$, however, leaving the previous columns as they are. So, we define the next $\boldsymbol{u}$ as

$$u_i = 0, \quad i < k, \qquad u_k = a_k \pm \left(\textstyle\sum_{i=k}^{m} a_i^2\right)^{1/2}, \qquad u_i = a_i, \quad i > k.$$

Due to its zero components, such a $\boldsymbol{u}$ is orthogonal to the previous $k-1$ columns of $\tilde{A}$, hence they remain invariant under reflection (as well as the first $k-1$ rows of $\tilde{A}$). The bottom $m-k$ component of $H\boldsymbol{a}$ will vanish due to the computations given above

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.

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