Mathematical Tripos Part IB: Easter Term 2003

Numerical Analysis – Lecture 4

4 QR factorization (cont.)

4.1 Further properties of orthogonal matrices

The Gram-Schmidt algorithm is very unstable: round-off errors accumulate rapidly and, even for moderate values of n, the computed matrix Q is no longer orthogonal. Much better algorithms are based on constructing Q as a composition of certain elementary orthogonal mappings.

Lemma 4.1 *A matrix Q is orthogonal* $\Leftrightarrow$ $\|Qx\| = \|x\|\ \forall x \in \mathbb{R}^m$ *(i.e., the mapping Q preserves Euclidean norm $\|x\| = \sqrt{(x,x)}$)*

Proof. Set $S := Q^TQ$. Then

$$\|Qx\|^2 = (Qx, Qx) = (Q^TQx, x) =: (Sx, x).$$

If Q is orthogonal, then $S = I$, i.e. $\|Qx\|^2 = (x,x) = \|x\|^2$. Conversely, if Q is norm-preserving, then

$$(Sx, x) = (x, x) \quad \forall x \in \mathbb{R}^m \quad \text{(and } S = S^T\text{)}.$$

Taking $x = e_i$, we get $s_{ii} = (Se_i, e_i) = 1$, while the choice $x = e_i + e_j$ will provide $s_{ii} + 2s_{ij} + s_{jj} = 2$, whence $s_{ij} = 0$, i.e., $S = I$, using symmetry of S. ∎

A mapping which preserves the distance between any two points (as orthogonal mapping Q does) is called an *isometry*. It is clear that the composition of two isometrical mappings is an isometry itself. So, the following is true.

Lemma 4.2 *If P, Q are orthogonal, so is PQ.*

The simplest isometries in $\mathbb{R}^m$ are *rotations* and *reflections*. So, the idea of the next two algorithms is quite clear geometrically: given a sequence of vectors (a_i) we can always determine a sequence of elementary rotations (reflections) in $\mathbb{R}^m$ that brings the coordinates of (a_i) to the triangular form.

4.2 Givens rotations

Consider a problem: given $a, b \in \mathbb{R}$, find $c, s \in \mathbb{R}$ such that, with some r,

$$c^2 + s^2 = 1, \qquad \begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}.$$

A solution always exists and is given by

$$c = \pm\tfrac{a}{\sqrt{a^2+b^2}} =: \cos\phi, \qquad s = \pm\tfrac{b}{\sqrt{a^2+b^2}} =: \sin\phi.$$

Generally, for any vector $a_p \in \mathbb{R}^m$, we can find a matrix $\Omega^{[p,q]}$ such that

$$\Omega^{[p,q]} a_p = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & c & & s & & \\ & & & \ddots & & & \\ & & -s & & c & & \\ & & & & & \ddots & \\ & & \uparrow & & \uparrow & & 1 \\ & & p & & q & & \end{bmatrix} \begin{bmatrix} a_{1p} \\ \vdots \\ a_{pp} \\ \vdots \\ a_{qp} \\ \vdots \\ a_{np} \end{bmatrix} = \begin{bmatrix} a_{1p} \\ \vdots \\ r \\ \vdots \\ 0 \\ \vdots \\ a_{np} \end{bmatrix} \begin{matrix} \\ \\ \leftarrow p \\ \\ \leftarrow q \\ \\ \\ \end{matrix} \qquad \begin{aligned} c &= \tfrac{a_{pp}}{\sqrt{a_{pp}^2 + a_{qp}^2}}, \\ s &= \tfrac{a_{qp}}{\sqrt{a_{pp}^2 + a_{qp}^2}}. \end{aligned}$$

Such a matrix $\Omega^{[p,q]}$ is called a *Givens rotation*. The mapping $\Omega^{[p,q]}$ rotates the vectors in the two-dimensional plane spanned by e_p and e_q (clockwise by the angle ϕ), hence it is orthogonal.

Lemma 4.3 *Let A be an $m \times n$ matrix and, for any $1 \le p \le q \le m$, let $\widetilde{A} = \Omega^{[p,q]} A$. Then*

1) $\widetilde{a}_{qp} = 0$;
2) *the p-th and the q-th rows of $\widetilde{A}$ are linear combinations of the p-th and q-th rows of A;*
3) *all other rows of $\widetilde{A}$ remains the same as those of A.*

Proof. Follows immediately from the form of $\Omega^{[p,q]}$. ∎

From this lemma it follows that we can subsequently annihilate the subdiagonal elements, column by column, and this implies the next statement.

Theorem 4.4 *For any matrix A, there exist Givens matrices $\Omega^{[p,q]}$ such that*

$$\left(\Omega^{[m-1,m]}\right) \dots \left(\Omega^{[2,m]} \dots \Omega^{[2,3]}\right) \left(\Omega^{[1,m]} \dots \Omega^{[1,3]} \Omega^{[1,2]}\right) A = R$$

is an upper triangular matrix.

Method 4.5

$$\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \overset{\Omega^{[1,2]}}{\to} \begin{bmatrix} \bullet & \bullet & \bullet \\ 0 & \bullet & \bullet \\ * & * & * \\ * & * & * \end{bmatrix} \overset{\Omega^{[1,3]}}{\to} \begin{bmatrix} \bullet & \bullet & \bullet \\ 0 & * & * \\ 0 & \bullet & \bullet \\ * & * & * \end{bmatrix} \overset{\Omega^{[1,4]}}{\to} \begin{bmatrix} \bullet & \bullet & \bullet \\ 0 & * & * \\ 0 & * & * \\ 0 & \bullet & \bullet \end{bmatrix} \overset{\Omega^{[2,3]}}{\to} \begin{bmatrix} * & * & * \\ 0 & \bullet & \bullet \\ 0 & 0 & \bullet \\ 0 & * & * \end{bmatrix} \overset{\Omega^{[2,4]}}{\to} \begin{bmatrix} * & * & * \\ 0 & \bullet & \bullet \\ 0 & 0 & * \\ 0 & 0 & \bullet \end{bmatrix} \overset{\Omega^{[3,4]}}{\to} \begin{bmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & \bullet \\ 0 & 0 & 0 \end{bmatrix}$$

The $\bullet$-elements have changed through a single rotation while the $*$-elements remain the same.

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

If the matrix Q is required in an explicit form, say, for solving the system with many right-hand sides, set $\Omega = I$ and, for each successive rotation, replace Ω by $\Omega^{[p,q]}\Omega$. The final Ω is the product of all the rotations, in correct order, and we let $Q = \Omega^T$. The extra cost is $\mathcal{O}(m^2 n)$.

If only one vector $Q^T b$ is required, we multiply the vector by successive rotations, the cost being $\mathcal{O}(mn)$.

For $m = n$ each rotation requires four times more multiplications as the corresponding Gaussian elimination, hence the total cost is $\frac{4}{3}n^3 + \mathcal{O}(n^2)$, four times as expensive. However, the QR factorization is generally more stable than the LU one.

4.3 Householder transformations

Definition 4.6 Given any nonzero $u \in \mathbb{R}^m$, the $m \times m$ matrix

$$H = H_u = I - \tfrac{2}{\|u\|^2} u u^T$$

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

$$Hu = -u, \qquad Hv = v \quad \text{if} \quad u^T v = 0,$$

this transformation reflects any vector $x \in \mathbb{R}^m$ in the $(m-1)$-dimensional hyperplane orthogonal to u. Each such matrix H is symmetric and (since reflection is an isometry) orthogonal.

Lemma 4.7 *For any two vectors a, b of equal length, the Householder transformation H_u with $u = a - b$ reflects a onto b. In particular, for any $a \in \mathbb{R}^m$, the choice $b = \gamma e_1$ implies*

$$H_u a = \gamma e_1, \qquad u = a - \gamma e_1, \qquad \gamma = \pm \|a\|.$$

Proof. Draw a picture.

Theorem 4.8 *For any matrix A, there exist Housholder matrices H_k such that*

$$H_{n-1} \cdots H_2 H_1 A =: R$$

is an upper triangular matrix.

Proof. We take $H_1 = H_u$ with the vector $u = a - \gamma e_1$, and a being the first column of A, or explicitly

$$u_1 = a_1 \pm \left(\sum_{i=1}^m a_i^2\right)^{1/2}, \qquad u_i = a_i, \quad i > 1.$$

Then $H_1 A$ has γe_1 as its first column. Suppose that the first $k-1$ columns of $\widetilde{A} = H_{k-1} \cdots H_1 A$ have an upper triangular form, and let $\widetilde{a}$ be the k-th column of $\widetilde{A}$. So, we define the next $H_k = H_u$ taking u as

$$u_i = 0, \quad i < k, \qquad u_k = \widetilde{a}_k \pm \left(\sum_{i=k}^m \widetilde{a}_i^2\right)^{1/2}, \qquad u_i = \widetilde{a}_i, \quad i > k.$$

Due to its zero components, such a u is orthogonal to the previous $k-1$ columns of $\widetilde{A}$, hence they remain invariant under reflection $H_k \widetilde{A}$ (as well as the first $k-1$ rows of $\widetilde{A}$). The bottom $m-k$ component of $H_k a$ will vanish due to Lemma 4.7. ∎

Example 4.9

$$A = \left[\begin{array}{ccc} 2 & 4 & 7 \\ 0 & 3 & -1 \\ \hline 0 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & -2 \end{array}\right] \quad \Rightarrow \quad \gamma = -\|\widetilde{a}\| = -3, \quad u = \begin{bmatrix} 0 \\ 0 \\ 5 \\ 1 \\ -2 \end{bmatrix} \quad \Rightarrow \quad \left(I - 2\frac{uu^{\mathrm{T}}}{\|u\|^2}\right) A = \begin{bmatrix} 2 & 4 & 7 \\ 0 & 3 & -1 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Calculation of Q and $Q^T b$. If the matrix Q is required in an explicit form, set $\Omega = I$ initially and, for each successive reflection, replace Ω by $H_u \Omega$. As in the case of Givens rotations, by the end of the computation, $Q = \Omega^T$. The same algorithm is being used to calculate the vector $c = Q^T b$.

Remark 4.10 For practical computations notice the following. For a matrix B, respectively a vector x, one should compute the products $H_u B$ and $H_u x$ as

$$H_u B = B - \frac{2}{\|u\|^2} u (u^T B), \qquad H_u x = x - 2\frac{u^T x}{\|u\|^2} u.$$

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!

4.4 Examples

Example 4.11 (Givens rotations) For

$$A = \begin{bmatrix} 2 & 4 & 5 \\ 1 & -1 & 1 \\ 2 & 1 & -1 \end{bmatrix},$$

we have the following steps of the Givens rotation algorithm:

$$\begin{aligned}
R \leftarrow A &= \begin{bmatrix} \boxed{2} & 4 & 5 \\ \boxed{1} & -1 & 1 \\ 2 & 1 & -1 \end{bmatrix} \\
R \leftarrow \Omega^{[1,2]} R &= \underbrace{\tfrac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & \sqrt{5} \end{bmatrix}}_{\Omega^{[1,2]}} \begin{bmatrix} 2 & 4 & 5 \\ 1 & -1 & 1 \\ 2 & 1 & -1 \end{bmatrix} = \tfrac{1}{\sqrt{5}} \begin{bmatrix} \boxed{5} & 7 & 11 \\ 0 & -6 & -3 \\ \boxed{2\sqrt{5}} & \sqrt{5} & -\sqrt{5} \end{bmatrix} \\
R \leftarrow \Omega^{[1,3]} R &= \underbrace{\tfrac{1}{3\sqrt{5}} \begin{bmatrix} 5 & 0 & 2\sqrt{5} \\ 0 & 3\sqrt{5} & 0 \\ -2\sqrt{5} & 0 & 5 \end{bmatrix}}_{\Omega^{[1,3]}} \tfrac{1}{\sqrt{5}} \begin{bmatrix} 5 & 7 & 11 \\ 0 & -6 & -3 \\ 2\sqrt{5} & \sqrt{5} & -\sqrt{5} \end{bmatrix} = \tfrac{3}{\sqrt{5}} \begin{bmatrix} \sqrt{5} & \sqrt{5} & \sqrt{5} \\ 0 & \boxed{-2} & -1 \\ 0 & \boxed{-1} & -3 \end{bmatrix}, \\
R \leftarrow \Omega^{[2,3]} R &= \underbrace{\tfrac{1}{\sqrt{5}} \begin{bmatrix} \sqrt{5} & 0 & 0 \\ 0 & -2 & -1 \\ 0 & 1 & -2 \end{bmatrix}}_{\Omega^{[2,3]}} \tfrac{3}{\sqrt{5}} \begin{bmatrix} \sqrt{5} & \sqrt{5} & \sqrt{5} \\ 0 & -2 & -1 \\ 0 & -1 & -3 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 3 \\ 0 & 3 & 3 \\ 0 & 0 & 3 \end{bmatrix}.
\end{aligned}$$

Finally,

$$\begin{aligned}
Q^T = \Omega^{[2,3]}\Omega^{[1,3]}\Omega^{[1,2]} &= \tfrac{1}{\sqrt{5}} \begin{bmatrix} \sqrt{5} & 0 & 0 \\ 0 & -2 & -1 \\ 0 & 1 & -2 \end{bmatrix} \tfrac{1}{3} \begin{bmatrix} \sqrt{5} & 0 & 2 \\ 0 & 3 & 0 \\ -2 & 0 & \sqrt{5} \end{bmatrix} \qquad \times \Omega^{[1,2]} \\
&= \tfrac{1}{3\sqrt{5}} \begin{bmatrix} 5 & 0 & 2\sqrt{5} \\ 2 & -6 & -\sqrt{5} \\ 4 & 3 & -2\sqrt{5} \end{bmatrix} \tfrac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & \sqrt{5} \end{bmatrix} = \tfrac{1}{3} \begin{bmatrix} 2 & 1 & 2 \\ 2 & -2 & -1 \\ 1 & 2 & -2 \end{bmatrix}.
\end{aligned}$$

so that

$$A = \begin{bmatrix} 2 & 4 & 5 \\ 1 & -1 & 1 \\ 2 & 1 & -1 \end{bmatrix} = \underbrace{\frac{1}{3} \begin{bmatrix} 2 & 2 & 1 \\ 1 & -2 & 2 \\ 2 & -1 & -2 \end{bmatrix}}_{Q} \cdot \underbrace{\begin{bmatrix} 3 & 3 & 3 \\ & 3 & 3 \\ & & 3 \end{bmatrix}}_{R}.$$

Example 4.12 (Householder reflection) For

$$A = \begin{bmatrix} 2 & 4 & 5 \\ 1 & -1 & 1 \\ 2 & 1 & -1 \end{bmatrix},$$

we have the following steps of the Housholder algorithm:

$$A = \begin{bmatrix} \boxed{\begin{matrix} 2 \\ 1 \\ 2 \end{matrix}} & 4 & 5 \\ & -1 & 1 \\ & 1 & -1 \end{bmatrix},$$

$$\boldsymbol{u}_1 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \qquad \boldsymbol{u}_1\boldsymbol{u}_1^T = \begin{bmatrix} 1 & -1 & -2 \\ -1 & 1 & 2 \\ -2 & 2 & 4 \end{bmatrix}, \qquad \Omega_1 = I - 2\frac{\boldsymbol{u}_1\boldsymbol{u}_1^T}{\|\boldsymbol{u}_1\|^2} = \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & -2 \\ 2 & -2 & -1 \end{bmatrix},$$

$$R_1 = \Omega_1 A = \begin{bmatrix} 3 & 3 & 3 \\ 0 & \boxed{0} & 3 \\ 0 & \boxed{3} & 3 \end{bmatrix},$$

$$\boldsymbol{u}_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \qquad \boldsymbol{u}_2\boldsymbol{u}_2^T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}, \qquad \Omega_2 = I - 2\frac{\boldsymbol{u}_2\boldsymbol{u}_2^T}{\|\boldsymbol{u}_2\|^2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

$$R = \Omega_2 R_1 = \begin{bmatrix} 3 & 3 & 3 \\ 0 & 3 & 3 \\ 0 & 0 & 3 \end{bmatrix}.$$

Finally,

$$Q = (\Omega_2\Omega_1)^T = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & -2 \\ 2 & -2 & -1 \end{bmatrix}\right)^T = \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ 2 & -2 & -1 \\ 1 & 2 & -2 \end{bmatrix}^T = \frac{1}{3}\begin{bmatrix} 2 & 2 & 1 \\ 1 & -2 & 2 \\ 2 & -1 & -2 \end{bmatrix},$$

so that

$$A = \begin{bmatrix} 2 & 4 & 5 \\ 1 & -1 & 1 \\ 2 & 1 & -1 \end{bmatrix} = \underbrace{\frac{1}{3}\begin{bmatrix} 2 & 2 & 1 \\ 1 & -2 & 2 \\ 2 & -1 & -2 \end{bmatrix}}_{Q} \cdot \underbrace{\begin{bmatrix} 3 & 3 & 3 \\ & 3 & 3 \\ & & 3 \end{bmatrix}}_{R}.$$

4.5 Exercises

Exercises with a $*$-mark are *not* subject to supervision! They are just for fun!

4.1. Calculate the QR factorization of the matrix of Exercise 3.1 by using three Givens rotations. Explain why the initial rotation can be any one of the three types $\Omega^{(1,2)}$, $\Omega^{(1,3)}$ and $\Omega^{(2,3)}$. Prove that the final factorization is independent of this initial choice in exact arithmetic, provided that we satisfy the condition that in each row of R the leading nonzero element is positive.

4.2. Let A be an $n \times n$ matrix, and for $i = 1, 2, \ldots, n$ let $k(i)$ be the number of zero elements in the i-th row of A that come before all nonzero elements in this row and before the diagonal element a_{ii}. Show that the QR factorization of A can be calculated by using at most $\frac{1}{2}n(n-1) - \sum k(i)$ Givens rotations. Hence show that, if A is an upper triangular matrix except that there are nonzero elements in its first column, i.e. $a_{ij} = 0$ when $2 \leq j < i \leq n$, then its QR factorization can be calculated by using only $2n - 3$ Givens rotations. [*Hint*: Your should find the order of the first $(n-2)$ rotations that brings your matrix to the form considered above.]

4.3. Calculate the QR factorization of the matrix of Exercise 8 by using two Householder reflections. Show that, if this technique is used to generate the QR factorization of a general $n \times n$ matrix A, then the computation can be organised so that the total number of additions and multiplications is bounded above by a constant multiple of n^3.

4.4.* Denote by $O(n)$ the group of orthogonal matrices, and by $SO(n)$ its subgroup of orthogonal matrices with determinant equal to one. Prove:

a) every $Q \in O(n)$ is a product of $\leq n-1$ Housholder reflections,

b) every $Q \in SO(n)$ is a product of $\leq \frac{1}{2}n(n-1)$ Givens rotations.