Mathematical Tripos Part IB: Easter Term 2003

Numerical Analysis – Lecture 3

3 QR factorization

3.1 Inner product spaces

Definition 3.1 An *inner product* on a real vector space $\mathbb{X}$ is a function $(\cdot,\cdot) : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ which satisfies the following axioms

$$\begin{array}{ll} 1) & (x,x) \ge 0 \quad \forall x \in \mathbb{X}, \quad \text{with equality only if } x = 0 \\ 2) & (x,y) = (y,x) \quad \forall x,y \in \mathbb{X}, \\ 3) & (\alpha x, y) = \alpha(x,y) \quad \forall \alpha \in \mathbb{R} \quad \forall x,y \in \mathbb{X}, \\ 4) & (x+y,z) = (x,z) + (y,z) \quad \forall x,y,z \in \mathbb{X}. \end{array}$$

A vector space with an inner product is called an *inner product* space. If $(x,y) = 0$, then the vectors x, y are called *orthogonal*. A set of vectors $(x_i) \in \mathbb{X}$ is called *orthonormal* if

$$(x_i, x_j) = \delta_{ij} := \begin{cases} 0, & i = j; \\ 1, & i \neq j\,. \end{cases}$$

For $x \in \mathbb{X}$, the function $\|x\| := (x,x)^{1/2}$ is called the *norm* of x (induced by the given inner product), and we have the Cauchy–Schwarz inequality

$$(x,y) \le \|x\|\,\|y\| \quad \forall x,y \in \mathbb{X}.$$

Example 3.2 For $\mathbb{X} = \mathbb{R}^m$, the following rule defines the so-called Euclidean inner product

$$(u,v) := u^T v := \sum_{i=1}^{m} u_i v_i, \quad u,v \in \mathbb{R}^m$$

Example 3.3 Another example, which will be used later in construction of orthogonal polynomials, is the inner product on the space $C[a,b]$ of real-valued continuous functions on $[a,b]$ with respect to the fixed positive *weight function* w. It is given by the rule

$$(f,g)_w := \int_a^b [f(x)g(x)]\,w(x)\,dx, \quad f,g \in C[a,b].$$

3.2 Orthogonal matrices

Definition 3.4 An $m \times n$ matrix Q $(m \ge n)$ is called *orthogonal* if $Q^T Q = I$, e.g., $Q = \frac{1}{3}\begin{bmatrix} 2 & 2 & 1 \\ 1 & -2 & 2 \\ 2 & -1 & -2 \end{bmatrix}$.

Thus, the columns $q_1, ..., q_n \in \mathbb{R}^m$ of the orthogonal Q satisfy $q_i^T q_j = \delta_{ij}$, i.e., they are *orthonormal* with respect to the Euclidean inner product in $\mathbb{R}^m$. For a square Q we get

$$Q^{-1} = Q^T, \quad I = Q^T Q = Q Q^T = (Q^T)^T Q^T,$$

therefore, Q^T is also an orthogonal matrix, so that the rows of Q are orthonormal as well. In particular, the column and the row elements of orthogonal $Q \in \mathbb{R}^{n\times n}$ satisfy $\sum_{i=1}^n q_{ij}^2 = \sum_{j=1}^n q_{ij}^2 = 1$. It follows also that a square orthogonal Q is nonsingular. Moreover

$$1 = \det I = \det(QQ^T) = \det Q \det Q^T = (\det Q)^2 \quad \Rightarrow \quad \det Q = \pm 1\,.$$

3.3 QR factorization

Definition 3.5 The *QR factorization* of an $m \times n$ matrix A is the representation

$$A = QR,$$

where Q is an $m \times m$ *orthogonal* matrix and R is an $m \times n$ *upper triangular* matrix,

$$\underbrace{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}}_{n} = \underbrace{\begin{bmatrix} q_{11} & \cdots & q_{1n} & \cdots & q_{1m} \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ q_{n1} & & q_{nn} & & q_{nm} \\ \vdots & & \vdots & & \vdots \\ q_{m1} & \cdots & q_{mn} & \cdots & q_{mm} \end{bmatrix}}_{m \geq n} \underbrace{\left.\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & & \vdots \\ & & \ddots & \vdots \\ & & & r_{nn} \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}\right\}}_{n} {\scriptstyle m \geq n} \, .$$

Due to the bottom zero element of R, the columns $q_{n+1}, ..., q_m$ are inessential for the representation itself, hence we can safely write

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ q_{n1} & \cdots & q_{nn} \\ \vdots & & \vdots \\ q_{m1} & \cdots & q_{mn} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & & \vdots \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix} . \tag{3.1}$$

The latter formula is called the *skinny* QR factorization.

Theorem 3.6 *Every matrix A has a QR factorization. If A is non-singular, then a factorization $A = QR$ where R has a positive main diagonal is unique.*

Proof. Three algorithms of the QR factorization are given below:
1) the Gram–Schmidt orthogonalization (for the 'skinny' version),
2) the Givens rotations and
3) the Householder reflections.

Let $A = QR$ be non-singular. Then $A^T A = R^T R$ is positive definite, and there is a unique Cholesky factorization $A = \widetilde{L}\widetilde{L}^T$ with $\widetilde{L}$ having a positive main diagonal. Thus, $R^T = \widetilde{L}$ is uniquely determined. ∎

Method 3.7 (Solution to a linear system) If A is square nonsingular, we can solve $Ax = b$ by calculating the QR factorization of A and then proceeding in two steps

$$Ax = Q \underbrace{Rx}_{y} = b$$

solving first $Qy = b$ (hence $y = Q^T b$) and then $Rx = y$ (a triangular system, back-substitution).

3.4 The Gram-Schmidt orthogonalization

Let $A = QR$ be sought. If we denote the columns of A and Q by (a_j) and (q_j) respectively, then it follows from (3.1) that, given $(a_k) \in \mathbb{R}^m$, we want to find an orthonormal set $(q_j) \in \mathbb{R}^m$ that satisfies

$$a_k = \textstyle\sum_{j=1}^{k} r_{jk} q_j, \quad k = 1...n.$$

The latter problem can be solved for a general inner product space $\mathbb{X}$.

Theorem 3.8 *Let $\mathbb{X}$ be an inner product space, and let elements $a_1, ..., a_n \in \mathbb{X}$ be linearly independent. Then there exist elements $q_1, ..., q_n \in \mathbb{X}$ such that*

$$\textstyle\sum_{j=1}^{k} r_{jk} q_j = a_k, \quad k = 1...n, \tag{3.2}$$

$$(q_i, q_j) = \delta_{ij}, \quad i, j = 1...n. \tag{3.3}$$

Proof. From the first equation of (3.2), since $a_1 \neq 0$,

$$r_{11}q_1 = a_1, \quad \|q_1\| = 1 \quad \Rightarrow \quad r_{11}^2 = (a_1, a_1) \quad \Rightarrow \quad r_{11} = \|a_1\|, \quad q_1 = a_1/\|a_1\|.$$

Suppose that the elements q_j which satisfy (3.2)-(3.3) are constructed for all $j < k$. The next element q_k should be the part of the orthonormal set $q_1, ..., q_{k-1}, q_k$, and it should satisfy

$$r_{kk}q_k + \sum_{j=1}^{k-1} r_{jk}q_j = a_k. \tag{3.4}$$

Multiplying both sides of (3.4) by q_j (in the scalar product sense) and using orthonormality, we find the first $(k-1)$ coefficients

$$r_{jk} = (q_j, a_k), \quad j = 1...k-1.$$

Substituting them back to (3.4) we obtain

$$r_{kk}q_k = a_k - \sum_{j=1}^{k-1}(q_j, a_k)q_j =: b_k \quad \Rightarrow \quad r_{kk} = \|b_k\|, \quad q_k = b_k/\|b_k\|.$$

By construction, each q_j is a linear combination of $(a_i)_{i \leq j}$, hence $b_k := a_k - \sum_{j=1}^{k-1} c_j a_j$ is non-zero since (a_i) are linearly independent, i.e., q_k is well-defined. ∎

Algorithm 3.9 We may put the previous construction in the following algorithm

```
for k = 1 to n do
    b[k] := a[k];
    for j = 1 to k-1 do     #--  void if k=1
        r(j,k) := <q[j],a[k]>;
        b[k] := b[k] - r(j,k)*q[j]
    end
    r(k,k) := sqrt(<b[k],b[k]>);
    q[k] := b[k]/r(k,k)
end
```

Example 3.10 QR factorization by Gram-Schmidt of $A = \begin{bmatrix} 2 & 4 & 5 \\ 1 & -1 & 1 \\ 2 & 1 & -1 \end{bmatrix}$.

$$r_{11} = \|a_1\| = 3, \quad q_1 = a_1/r_{11} = \tfrac{1}{3}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix};$$

$$r_{12} = (q_1, a_2) = 3, \qquad b_2 = a_2 - r_{12}q_1 = \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix} - 3 \cdot \tfrac{1}{3}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix},$$

$$r_{22} = \|b_2\| = 3, \quad q_2 = b_2/r_{22} = \tfrac{1}{3}\begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix};$$

$$r_{13} = (q_1, a_3) = 3, \quad r_{23} = (q_2, a_3) = 3, \qquad b_3 = a_3 - r_{13}q_1 - r_{23}q_2 = \begin{bmatrix} 5 \\ 1 \\ -1 \end{bmatrix} - 3 \cdot \tfrac{1}{3}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} - 3 \cdot \tfrac{1}{3}\begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix},$$

$$r_{33} = \|b_3\| = 3, \quad q_3 = b_3/r_{33} = \tfrac{1}{3}\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}.$$

So,

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

3.5 Exercises

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

3.1. Let $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ denote the columns of the matrix

$$A = \begin{bmatrix} 6 & 6 & 1 \\ 3 & 6 & 1 \\ 2 & 1 & 1 \end{bmatrix}.$$

Apply the Gram–Schmidt procedure to A, which generates orthonormal vectors $\boldsymbol{q}_1, \boldsymbol{q}_2$ and $\boldsymbol{q}_3$. Note that this calculation provides real numbers r_{jk} such that $\boldsymbol{a}_k = \sum_{j=1}^{k} r_{jk}\boldsymbol{q}_j$, $k = 1, 2, 3$. Hence express A as the product $A = QR$, where Q and R are orthogonal and upper-triangular matrices respectively.

3.2.* For an inner product space $\mathbb{X}$ prove the Cauchy-Schwarz inequality

$$(x, y) \leq \|x\| \, \|y\| \quad \forall x, y \in \mathbb{X}.$$

[*Hint*: Given $x, y \in \mathbb{X}$, the quadratic polynomial $p(\lambda) := (\lambda x + y, \lambda x + y)$ has no real roots.]

3.3.* Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and let

$$A^T A = LL^T$$

be Cholesky factorization of $A^T A$. Prove that the QR factorization of A is given by the formula

$$A = QR, \quad Q = A(L^T)^{-1}, \quad R = L^T.$$