Mathematical Tripos Part IB: Lent 2010

Numerical Analysis – Lecture 14[1]

Banded matrices The matrix A is a *banded matrix* if there exists an integer $r < n$ such that $A_{i,j} = 0$ for $|i-j| > r$, $i,j = 1,2,\ldots,n$. In other words, all the nonzero elements of A reside in a band of width $2r+1$ along the main diagonal. In that case, according to the statement from the end of the last lecture, $A = LU$ implies that $L_{i,j} = U_{i,j} = 0 \;\forall\; |i-j| > r$ and sparsity structure is inherited by the factorization.
In general, the expense of calculating an LU factorization of an $n \times n$ *dense* matrix A is $\mathcal{O}(n^3)$ operations and the expense of solving $A\boldsymbol{x} = \boldsymbol{b}$, provided that the factorization is known, is $\mathcal{O}(n^2)$. However, in the case of a banded A, we need just $\mathcal{O}(r^2 n)$ operations to factorize and $\mathcal{O}(rn)$ operations to solve a linear system. If $r \ll n$ this represents a very substantial saving!

General sparse matrices feature a wide range of applications, e.g. the solution of partial differential equations, and there exists a wealth of methods for their solution. One approach is efficient factorization, that minimizes *fill in.* Yet another is to use iterative methods (cf. Part II Numerical Analysis course). There also exists a substantial body of other, highly effective methods, e.g. Fast Fourier Transforms, preconditioned conjugate gradients and multigrid techniques (cf. Part II Numerical Analysis course), fast multipole techniques and much more.

Sparsity and graph theory An exceedingly powerful (and beautiful) methodology of ordering pivots to minimize fill-in of sparse matrices uses graph theory and, like many other cool applications of mathematics in numerical analysis, is alas not in the schedules `:-(`

5.2 QR factorization of matrices

Scalar products, norms and orthogonality We first recall a few definitions. $\mathbb{R}^n$ is the linear space of all real n-tuples.

- For all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ we define the *scalar product*

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle = \sum_{j=1}^{n} u_j v_j = \boldsymbol{u}^\top \boldsymbol{v} = \boldsymbol{v}^\top \boldsymbol{u}\,.$$

- If $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ then $\langle \alpha\boldsymbol{u} + \beta\boldsymbol{w}, \boldsymbol{v} \rangle = \alpha\langle \boldsymbol{u}, \boldsymbol{v} \rangle + \beta\langle \boldsymbol{w}, \boldsymbol{v} \rangle$.
- The *norm* (a.k.a. the *Euclidean length*) of $\boldsymbol{u} \in \mathbb{R}^n$ is $\|\boldsymbol{u}\| = \left(\sum_{j=1}^{n} u_j^2\right)^{1/2} = \langle \boldsymbol{u}, \boldsymbol{u} \rangle^{1/2} \geq 0$.
- For $\boldsymbol{u} \in \mathbb{R}^n$, $\|\boldsymbol{u}\| = 0$ iff $\boldsymbol{u} = \boldsymbol{0}$.
- We say that $\boldsymbol{u} \in \mathbb{R}^n$ and $\boldsymbol{v} \in \mathbb{R}^n$ are *orthogonal* to each other if $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$.
- The vectors $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_m \in \mathbb{R}^n$ are *orthonormal* if

$$\langle \boldsymbol{q}_k, \boldsymbol{q}_\ell \rangle = \begin{cases} 1, & k = \ell, \\ 0, & k \neq \ell, \end{cases} \qquad k, \ell = 1, 2, \ldots, m.$$

- An $n \times n$ real matrix Q is *orthogonal* if all its columns are orthonormal. Since $(Q^\top Q)_{k,\ell} = \langle \boldsymbol{q}_k, \boldsymbol{q}_\ell \rangle$, this implies that $Q^\top Q = I$ (I is the *unit matrix*). Hence $Q^{-1} = Q^\top$ and $QQ^\top = QQ^{-1} = I$. We conclude that the rows of an orthogonal matrix are also orthonormal, and that $Q^\top$ is an orthogonal matrix. Further, $1 = \det I = \det(QQ^\top) = \det Q \det Q^\top = (\det Q)^2$, and thus we deduce that $\det Q = \pm 1$, and that an orthogonal matrix is nonsingular.

[1]Corrections and suggestions to these notes should be emailed to `A.Iserles@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`.

Proposition If P, Q are orthogonal then so is PQ.

Proof. Since $P^\top P = Q^\top Q = I$, we have $(PQ)^\top(PQ) = (Q^\top P^\top)(PQ) = Q^\top(P^\top P)Q = Q^\top Q = I$, hence PQ is orthogonal. □

Proposition Let $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_m \in \mathbb{R}^n$ be orthonormal. Then $m \le n$.

Proof. We argue by contradiction. Suppose that $m \ge n+1$ and let Q be the orthogonal matrix whose columns are $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_n$. Since Q is nonsingular and $\boldsymbol{q}_m \neq \boldsymbol{0}$, there exists a nonzero solution to the linear system $Q\boldsymbol{a} = \boldsymbol{q}_m$, hence $\boldsymbol{q}_m = \sum_{j=1}^n a_j \boldsymbol{q}_j$. But

$$0 = \langle \boldsymbol{q}_\ell, \boldsymbol{q}_m \rangle = \left\langle \boldsymbol{q}_\ell, \sum_{j=1}^n a_j \boldsymbol{q}_j \right\rangle = \sum_{j=1}^n a_j \langle \boldsymbol{q}_\ell, \boldsymbol{q}_j \rangle = a_\ell, \qquad \ell = 1, 2, \ldots, n,$$

hence $\boldsymbol{a} = \boldsymbol{0}$, a contradiction. We deduce that $m \le n$. □

Lemma Let $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_m \in \mathbb{R}^n$ be orthonormal and $m \le n-1$. Then there exists $\boldsymbol{q}_{m+1} \in \mathbb{R}^n$ such that $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_{m+1}$ are orthonormal.

Proof. We construct $\boldsymbol{q}_{m+1}$. Let Q be the $n \times m$ matrix whose columns are $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_m$. Since

$$\sum_{k=1}^n \sum_{j=1}^m Q_{k,j}^2 = \sum_{j=1}^m \|\boldsymbol{q}_j\|^2 = m < n,$$

it follows that $\exists\, \ell \in \{1, 2, \ldots, n\}$ such that $\sum_{j=1}^m Q_{\ell,j}^2 < 1$. We let $\boldsymbol{w} = \boldsymbol{e}_\ell - \sum_{j=1}^m \langle \boldsymbol{q}_j, \boldsymbol{e}_\ell \rangle \boldsymbol{q}_j$. Then for $i = 1, 2, \ldots, m$

$$\langle \boldsymbol{q}_i, \boldsymbol{w} \rangle = \langle \boldsymbol{q}_i, \boldsymbol{e}_\ell \rangle - \sum_{j=1}^m \langle \boldsymbol{q}_j, \boldsymbol{e}_\ell \rangle \langle \boldsymbol{q}_i, \boldsymbol{q}_j \rangle = 0,$$

i.e. by design $\boldsymbol{w}$ is orthogonal to $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_m$. Further, since $Q_{\ell,j} = \langle \boldsymbol{q}_j, \boldsymbol{e}_\ell \rangle$, we have

$$\|\boldsymbol{w}\|^2 = \langle \boldsymbol{w}, \boldsymbol{w} \rangle = \langle \boldsymbol{e}_\ell, \boldsymbol{e}_\ell \rangle - 2\sum_{j=1}^m \langle \boldsymbol{q}_j, \boldsymbol{e}_\ell \rangle \langle \boldsymbol{e}_\ell, \boldsymbol{q}_j \rangle + \sum_{j=1}^m \langle \boldsymbol{q}_j, \boldsymbol{e}_\ell \rangle \sum_{k=1}^m \langle \boldsymbol{q}_k, \boldsymbol{e}_\ell \rangle \langle \boldsymbol{q}_j, \boldsymbol{q}_k \rangle = 1 - \sum_{j=1}^m Q_{\ell,j}^2 > 0.$$

Thus we define $\boldsymbol{q}_{m+1} = \boldsymbol{w}/\|\boldsymbol{w}\|$. □

The QR factorization The QR factorization of an $m \times n$ matrix A has the form $A = QR$, where Q is an $m \times m$ *orthogonal* matrix and R is an $m \times n$ *upper triangular* matrix (i.e., $R_{i,j} = 0$ for $i > j$). We will demonstrate in the sequel that every matrix has a (non-unique) QR factorization. We say that R is in a *standard form* if, given that R_{k,j_k} is the first nonzero entry in the kth row, the j_ks form a strictly monotone sequence. (Such R is also allowed entire rows of zeros, but only at the bottom.)

An application Let $m = n$ and A be nonsingular. We can solve $A\boldsymbol{x} = \boldsymbol{b}$ by calculating the QR factorization of A and solving first $Q\boldsymbol{y} = \boldsymbol{b}$ (hence $\boldsymbol{y} = Q^\top \boldsymbol{b}$) and then $R\boldsymbol{x} = \boldsymbol{y}$ (a triangular system!).

Interpretation of the QR factorization Let $m \ge n$ and denote the columns of A and Q by $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ and $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_m$ respectively. Since

$$\begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots & \boldsymbol{q}_m \end{bmatrix} \begin{bmatrix} R_{1,1} & R_{1,2} & \cdots & R_{1,n} \\ 0 & R_{2,2} & & \vdots \\ \vdots & \ddots & \ddots & \\ & & 0 & R_{n,n} \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix},$$

we have $\boldsymbol{a}_k = \sum_{j=1}^k R_{j,k} \boldsymbol{q}_j$, $k = 1, 2, \ldots, n$. In other words, Q has the property that each kth column of A can be expressed as a linear combination of the first k columns of Q.