Dr A. Iserles
Easter 1997

Numerical Analysis – Lecture 5

4.3 The Gram–Schmidt algorithm

Given an $m \times n$ matrix A with the columns $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$, we construct Q & R as follows. Suppose that $\boldsymbol{a}_1 \neq \boldsymbol{0}$, then we derive $\boldsymbol{q}_1$ and $R_{1,1}$ from the equation $\boldsymbol{a}_1 = R_{1,1}\boldsymbol{q}_1$. Since $\|\boldsymbol{q}_1\| = 1$, we let $\boldsymbol{q}_1 = \boldsymbol{a}_1/\|\boldsymbol{a}_1\|$, $R_{1,1} = \|\boldsymbol{a}_1\|$.
Next we form the vector $\boldsymbol{b} = \boldsymbol{a}_2 - \langle \boldsymbol{q}_1, \boldsymbol{a}_2\rangle \boldsymbol{q}_1$. It is orthogonal to $\boldsymbol{q}_1$, since

$$\langle \boldsymbol{q}_1, \boldsymbol{a}_2 - \langle \boldsymbol{q}_1, \boldsymbol{a}_2\rangle \boldsymbol{q}_1\rangle = \langle \boldsymbol{q}_1, \boldsymbol{a}_2\rangle - \langle \boldsymbol{q}_1, \boldsymbol{a}_2\rangle\langle \boldsymbol{q}_1, \boldsymbol{q}_1\rangle = 0.$$

If $\boldsymbol{b} \neq \boldsymbol{0}$, we set $\boldsymbol{q}_2 = \boldsymbol{b}/\|\boldsymbol{b}\|$, hence $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are orthonormal. Moreover,

$$\langle \boldsymbol{q}_1, \boldsymbol{a}_2\rangle \boldsymbol{q}_1 + \|\boldsymbol{b}\|\boldsymbol{q}_2 = \langle \boldsymbol{q}_1, \boldsymbol{a}_2\rangle \boldsymbol{q}_1 + \boldsymbol{b} = \boldsymbol{a}_2,$$

hence, to obey $\boldsymbol{a}_2 = R_{1,2}\boldsymbol{q}_1 + R_{2,2}\boldsymbol{q}_2$, we let $R_{1,2} = \langle \boldsymbol{q}_1, \boldsymbol{a}_2\rangle$, $R_{2,2} = \|\boldsymbol{b}\|$.

The Gram–Schmidt algorithm The above idea can be extended to all columns of A.

Step 1 Set $k := 0$, $j := 0$ (k is the number of columns of Q that have been formed and j is the number of columns of A that have been already considered);

Step 2 Increase j by 1. If $k = 0$ then set $\boldsymbol{b} := \boldsymbol{a}_j$, otherwise (i.e., when $k \geq 1$) set $R_{i,j} := \langle \boldsymbol{q}_i, \boldsymbol{a}_j\rangle$, $i = 1, 2, \ldots, k$, and $\boldsymbol{b} := \boldsymbol{a}_j - \sum_{i=1}^{k}\langle \boldsymbol{q}_i, \boldsymbol{a}_j\rangle \boldsymbol{q}_i$. *[Note: $\boldsymbol{b}$ is orthogonal to $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_k$.]*

Step 3 If $\boldsymbol{b} \neq \boldsymbol{0}$ increase k by 1. Subsequently, set $\boldsymbol{q}_k := \boldsymbol{b}/\|\boldsymbol{b}\|$, $R_{k,j} := \|\boldsymbol{b}\|$ and $R_{i,j} := 0$ for $i \geq k + 1$. *[Note: Hence, each column of Q has unit length, as required, $\boldsymbol{a}_j = \sum_{i=1}^{k} R_{i,j}\boldsymbol{q}_j$ and R is upper triangular, because $k \leq j$.]*

Step 4 Terminate if $j = n$, otherwise go to *Step 2*.

Previous lecture $\Rightarrow$ Since the columns of Q are orthonormal, there are at most m of them, i.e. the final value of k can't exceed m. If it is less then m then a previous lemma demonstrates that we can add columns so that Q becomes $m \times m$ and orthogonal.

The disadvantage of Gram–Schmidt is its *ill-conditioning*. Since we are using finite arithmetic, even small imprecisions in the calculation of inner products rapidly lead to effective loss of orthogonality. Thus, errors accumulate fast and, even for moderate values of m, we are likely to run into problems.

4.4 Orthogonal rotations

Given real, $m \times n$ matrix $A_0 = A$, we seek a sequence $\Omega_1, \Omega_2, \ldots, \Omega_k$ of $m \times m$ orthogonal matrices such that the matrix $A_i := \Omega_i A_{i-1}$ has more zero elements below the main diagonal than A_{i-1} for $i = 1, 2, \ldots, k$ and so that the manner of insertion of such zeros is such that A_k is upper triangular. We then let $R = A_k$, therefore

$$\Omega_k \Omega_{k-1} \cdots \Omega_2 \Omega_1 A = R$$

and $Q = (\Omega_k \Omega_{k-1} \cdots \Omega_1)^{-1} = \Omega_1^{\mathrm{T}} \Omega_2^{\mathrm{T}} \cdots \Omega_k^{\mathrm{T}}$. Hence $A = QR$, where Q is orthogonal and R upper triangular.

4.5 Givens rotations

We say that an $m \times m$ orthogonal matrix Ω_j is a *Givens rotation* if it coincides with the unit matrix, except for four elements. Specifically, we use the notation $\Omega^{[p,q]}$, where $1 \le p < q \le m$ for a matrix such that

$$\Omega^{[p,q]}_{p,p} = \Omega^{[p,q]}_{q,q} = \cos\theta, \qquad \Omega^{[p,q]}_{p,q} = \sin\theta, \qquad \Omega^{[p,q]}_{q,p} = -\sin\theta$$

for some $\theta \in [-\pi, \pi]$. The remaining elements of $\Omega^{[p,q]}$ are those of a unit matrix. For example,

$$m = 4 \quad \Longrightarrow \quad \Omega^{[1,2]} = \begin{bmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \Omega^{[2,4]} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\theta & 0 & \cos\theta \end{bmatrix}.$$

Geometrically, such matrices correspond to the underlying coordinate system being rigidly rotated on a two-dimensional plane. Trivially, they are orthogonal.

Theorem Let A be an $m \times n$ matrix. Then, for every $1 \le p < q \le m$, $i \in \{p, q\}$ and $1 \le j \le n$, there exists $\theta \in [-\pi, \pi]$ such that $(\Omega^{[p,q]}A)_{i,j} = 0$. Moreover, all the rows of $\Omega^{[p,q]}A$, except for the pth and the qth, are the same as the corresponding rows of A, whereas the pth and the qth rows are linear combinations of the 'old' pth and qth rows.

Proof. Let $i = q$. If $A_{p,j} = A_{q,j} = 0$ then any θ will do, otherwise we let

$$\cos\theta := A_{p,j}/\sqrt{A^2_{p,j} + A^2_{q,j}}, \qquad \sin\theta := A_{q,j}/\sqrt{A^2_{p,j} + A^2_{q,j}}.$$

Hence

$$(\Omega^{[p,q]}A)_{q,k} = -(\sin\theta)A_{p,k} + (\cos\theta)A_{q,k}, \quad k = 1, 2, \ldots, n \qquad \Rightarrow \qquad (\Omega^{[p,q]})_{q,j} = 0.$$

Likewise, when $i = p$ we let $\cos\theta := A_{q,j}/\sqrt{A^2_{p,j} + A^2_{q,j}}$, $\sin\theta := -A_{p,j}/\sqrt{A^2_{p,j} + A^2_{q,j}}$.
The last two statements of the theorem are an immediate consequence of the structure of $\Omega^{[p,q]}$. $\square$

An example: a 3×3 matrix Suppose that A is 3×3. We can force zeros underneath the main diagonal as follows.

1 First pick $\Omega^{[1,2]}$ so that $(\Omega^{[1,2]}A)_{1,2} = 0$.

2 Next pick $\Omega^{[1,3]}$ so that $(\Omega^{[1,3]}\Omega^{[1,2]}A)_{1,3} = 0$. Note that multiplication by $\Omega^{[1,3]}$ doesn't alter the second row, hence $(\Omega^{[1,3]}\Omega^{[1,2]}A)_{1,2}$ remains zero.

3 Finally, pick $\Omega^{[2,3]}$ so that $(\Omega^{[2,3]}\Omega^{[1,3]}\Omega^{[1,2]}A)_{2,3} = 0$. Since both second and third row of $\Omega^{[1,3]}\Omega^{[1,2]}A$ have a leading zero, their linear combination preserves these zeros, hence

$$(\Omega^{[2,3]}\Omega^{[1,3]}\Omega^{[1,2]}A)_{1,j} = 0, \qquad j = 2, 3.$$

It follows that $\Omega^{[2,3]}\Omega^{[1,3]}\Omega^{[1,2]}A$ is upper triangular.