Mathematical Tripos Part IB: Lent 2010

Numerical Analysis – Lecture 15[1]

The Gram–Schmidt algorithm Given an $m \times n$ matrix $A \neq O$ with the columns $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$, we construct Q & R where Q is orthogonal, R upper-triangular and $A = QR$: in other words,

$$\sum_{k=1}^{\ell} R_{k,\ell} \boldsymbol{q}_k = \boldsymbol{a}_\ell, \quad \ell = 1, 2, \ldots, n, \qquad \text{where} \qquad A = [\, \boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \cdots \quad \boldsymbol{a}_n \,]. \tag{5.2}$$

Assuming $\boldsymbol{a}_1 \neq \boldsymbol{0}$, we derive $\boldsymbol{q}_1$ and $R_{1,1}$ from the equation (5.2) for $k = 1$. 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 (5.2) for $k = 2$, we let $R_{1,2} = \langle \boldsymbol{q}_1, \boldsymbol{a}_2 \rangle$, $R_{2,2} = \|\boldsymbol{b}\|$.

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 already formed and j is the number of columns of A that have been already considered, clearly $k \leq j$);

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 ⇒ 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*: using finite arithmetic, small imprecisions in the calculation of inner products spread rapidly, leading to effective loss of orthogonality. Errors accumulate fast and the computed off-diagonal elements of $Q^\top Q$ may become large.

Orthogonality conditions are preserved well when one generates a new orthogonal matrix by computing the product of two given orthogonal matrices. Therefore algorithms that express Q as a product of simple orthogonal matrices are highly useful. This suggests an alternative way forward.

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

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, and $\det \Omega_j = 1$. Specifically, we use the notation $\Omega^{[p,q]}$, where $1 \leq p < q \leq m$ 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,

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

[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`.

Geometrically, such matrices correspond to the underlying coordinate system being rigidly rotated along a two-dimensional plane (in mechanics this is called an *Euler rotation*). It is trivial to confirm that 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

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

Hence

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

Likewise, when $i = p$ we let $\cos\theta := A_{q,j}/\sqrt{A_{p,j}^2 + A_{q,j}^2}$, $\sin\theta := -A_{p,j}/\sqrt{A_{p,j}^2 + A_{q,j}^2}$.

The last two statements of the theorem are an immediate consequence of the construction of $\Omega^{[p,q]}$. □

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

1 First pick $\Omega^{[1,2]}$ so that $(\Omega^{[1,2]}A)_{2,1} = 0 \quad \Rightarrow \quad \Omega^{[1,2]}A = \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ \times & \times & \times \end{bmatrix}$.

2 Next pick $\Omega^{[1,3]}$ so that $(\Omega^{[1,3]}\Omega^{[1,2]}A)_{3,1} = 0$. Multiplication by $\Omega^{[1,3]}$ doesn't alter the second row, hence $(\Omega^{[1,3]}\Omega^{[1,2]}A)_{2,1}$ remains zero $\quad \Rightarrow \quad \Omega^{[1,3]}\Omega^{[1,2]}A = \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{bmatrix}$.

3 Finally, pick $\Omega^{[2,3]}$ so that $(\Omega^{[2,3]}\Omega^{[1,3]}\Omega^{[1,2]}A)_{3,2} = 0$. Since both second and third row of $\Omega^{[1,3]}\Omega^{[1,2]}A$ have a leading zero, $(\Omega^{[2,3]}\Omega^{[1,3]}\Omega^{[1,2]}A)_{2,1} = (\Omega^{[2,3]}\Omega^{[1,3]}\Omega^{[1,2]}A)_{3,1} = 0$. It follows that $\Omega^{[2,3]}\Omega^{[1,3]}\Omega^{[1,2]}A$ is upper triangular. Therefore

$$R = \Omega^{[2,3]}\Omega^{[1,3]}\Omega^{[1,2]}A = \begin{bmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{bmatrix}, \qquad Q = (\Omega^{[2,3]}\Omega^{[1,3]}\Omega^{[1,2]})^\top.$$

The Givens algorithm Given $m \times n$ matrix A, let ℓ_i be the number of leading zeros in the ith row of A, $i = 1, 2, \ldots, m$.

Step 1 Stop if the (integer) sequence $\{\ell_1, \ell_2, \ldots, \ell_m\}$ increases monotonically, the increase being strictly monotone for $\ell_i \le n$.

Step 2 Pick any two integers $1 \le p < q \le m$ such that either $\ell_p > \ell_q$ or $\ell_p = \ell_q < n$.

Step 3 Replace A by $\Omega^{[p,q]}A$, using the Givens rotation that annihilates the $(q, \ell_q + 1)$ element.

Update the values of ℓ_p and ℓ_q and go to *Step 1.*

The final matrix A is upper triangular and also has the property that the number of leading zeros in each row increases *strictly monotonically* until all the rows of A are zero – a matrix of this form is said to be in *standard form*. This end result, as we recall, is 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 $\mathcal{O}(mn^2)$.

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 $Q = \Omega^\top$. The extra cost is $\mathcal{O}(m^2 n)$. However, in most applications we don't need Q but, instead, just the action of $Q^\top$ on a given vector (recall: solution of linear systems!). This can be accomplished by multiplying the vector by successive rotations, the cost being $\mathcal{O}(mn)$.