Mathematical Tripos Part IB: Lent 2020

Numerical Analysis – Lecture 15[1]

The Gram–Schmidt algorithm Assume that $m \geq n$ and that the columns of $A \in \mathbb{R}^{m\times n}$ are linearly independent. We will see how to construct a reduced QR factorization of A, i.e., $Q \in \mathbb{R}^{m\times n}$ having orthonormal columns, $R \in \mathbb{R}^{n\times n}$ upper-triangular and $A = QR$: in other words,

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

Equation (5.2) for $\ell = 1$ tells us that we must have $\boldsymbol{q}_1 = \boldsymbol{a}_1/\|\boldsymbol{a}_1\|$ and $R_{1,1} = \|\boldsymbol{a}_1\|$. 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 $\langle \boldsymbol{q}_1, \boldsymbol{a}_2 - \langle \boldsymbol{q}_1, \boldsymbol{a}_2\rangle\boldsymbol{q}_1\rangle = \langle \boldsymbol{q}_1, \boldsymbol{a}_2\rangle - \langle \boldsymbol{q}_1, \boldsymbol{a}_2\rangle\langle \boldsymbol{q}_1, \boldsymbol{q}_1\rangle = 0$. Since the columns of A are assumed linearly independent, $\boldsymbol{b} \neq \boldsymbol{0}$ and 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 $\ell = 2$, we let $R_{1,2} = \langle \boldsymbol{q}_1, \boldsymbol{a}_2\rangle$, $R_{2,2} = \|\boldsymbol{b}\|$.

More generally we get the following classical Gram-Schmidt algorithm to compute a QR factorization: Set $\boldsymbol{q}_1 = \boldsymbol{a}_1/\|\boldsymbol{a}_1\|$ and $R_{11} = \|\boldsymbol{a}_1\|$. For $j = 2, \ldots, n$: Set $R_{ij} = \langle \boldsymbol{q}_i, \boldsymbol{a}_j\rangle$ for $i \leq j-1$, and $\boldsymbol{b}_j = \boldsymbol{a}_j - \sum_{i=1}^{j-1} R_{ij}\boldsymbol{q}_i$. Set $\boldsymbol{q}_j = \boldsymbol{b}_j/\|\boldsymbol{b}_j\|$ and $R_{jj} = \|\boldsymbol{b}_j\|$.

The total cost of the classical Gram–Schmidt algorithm is $\mathcal{O}(n^2 m)$, since at each iteration j a total of $\mathcal{O}(mj)$ operations are performed.

The disadvantage of the classical 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.

The Gram-Schmidt algorithm operates by performing "triangular orthogonalization" on A: triangular operations are applied to A to produce the orthonormal system $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$. We are now going to see two algorithms for QR factorization that are based on "orthogonal triangularization": we will repeatedly apply orthogonal transformations to A to put it into triangular form.

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 Recall that the matrix associated to clockwise rotation in $\mathbb{R}^2$ by angle θ is $\left[\begin{smallmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{smallmatrix}\right]$. An $m \times m$ *Givens rotation* matrix Ω is an orthogonal matrix specified by two integers $1 \leq p < q \leq m$ and an angle $\theta \in [-\pi, \pi]$, which coincides with the identity matrix except for the 2×2 submatrix associated to rows/columns $\{p, q\}$ which correspond to a 2×2 rotation matrix. 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 an identity 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 a rotation in the two-dimensional coordinate subspace spanned by $\{\boldsymbol{e}_p, \boldsymbol{e}_q\}$, where $\boldsymbol{e}_i$ is the vector with zeros everywhere except for a 1 in position i.

[1] Corrections and suggestions to these notes should be emailed to `h.fawzi@damtp.cam.ac.uk`.

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 of $\Omega^{[p,q]}A$ are linear combinations of the pth and qth rows of A.

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 $\Omega^{[p,q]}$. $\square$

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: For each j from 1 to n and i from $j+1$ to m, replace A by $\Omega^{[j,i]}A$ where $\Omega^{[j,i]}$ is chosen to annihilate the (i, j) entry.

This algorithm transforms A into an upper triangular matrix by a sequence of orthogonal transformations. The final orthogonal matrix Q however is not computed explicitly in this algorithm. If we want to compute Q explicitly, we commence by letting Ω be the $m \times m$ identity matrix and, each time A is premultiplied by $\Omega^{[j,i]}$, 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 $Q = \Omega^\top$. Note 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 *given vector*, e.g., the right-hand side $\boldsymbol{b}$ if we are solving a linear system, by successive rotations.

The cost For each $j < i$, the cost of computing $\Omega^{[j,i]}A$ is $\mathcal{O}(n)$ since we just have to replace the j'th and i'th rows of A by their appropriate linear combinations. This has to be done less than mn times (the number of pairs (j, i)) and so the total cost is $\mathcal{O}(mn^2)$.