QR in two dimensions

Autor(en): **Kreyszig, Erwin / Todd, John**

Objekttyp: **Article**

Zeitschrift: **Elemente der Mathematik**

Band (Jahr): **31 (1976)**

Heft 5

PDF erstellt am: **28.04.2024**

Persistenter Link: https://doi.org/10.5169/seals-31401

7. *Déterminisme et libre arbitre.* Entretiens recueillis et rédigés par H.S. Gagnebin (Griffon, Neuchâtel 1944).
8. *La géométrie et le problème de l'espace.* 1. La doctrine préalable (1945). 2. Les trois aspects de la géométrie (1946). 3. L'édification géométrique (1947). 4. La synthèse dialectique (1949). 5. Les géométries non euclidiennes (1953). 6. Le problème de l'espace (1956). (Griffon, Neuchâtel).
9. *Leitfaden der Planimetrie,* F. Gonseth und P. Marti, 8. Auflage 1971 (Orell Füssli, Zürich 1933).
10. *Aufgabensammlung der Planimetrie,* F. Gonseth und P. Marti (Orell Füssli, Zürich 1939).
11. *Eléments de géométrie,* F. Gonseth et S. Gagnebin (Payot, Lausanne 1942).
12. *Analytische Geometrie der Ebene in moderner Behandlung,* F. Gonseth und M. Rueff (Haupt, Bern 1946).
13. *Philosophie mathématique.* Chronique de l'Institut international de philosophie (Hermann, Paris 1950).
14. *Philosophie des mathématiques.* Chronique de l'Institut international de philosophie (Hermann, Paris 1954).
15. *Leitfaden der Geometrie für Sekundarschulen,* F. Gonseth und E. Moser (Orell Füssli, Zürich 1955).
16. *Philosophie néo-scolastique et philosophie ouverte.* Entretiens du centre romain de comparaison et de synthèse; échange de vues entre un groupe de thomistes et F. Gonseth pendant et après un symposium tenu à Rome en 1952 (PUF 1954).
17. *La métaphysique et l'ouverture à l'expérience.* Suite en 1955 du symposium de Rome de 1952 (PUF 1960).
18. *Elementare und nichteuklidische Geometrie in axiomatischer Darstellung und ihr Verhältnis zur Wirklichkeit* (Orell Füssli, Zürich 1956).

QR in Two Dimensions

The QR process is one of the most efficient ways of determining the characteristic values of a matrix. It is a unitary analog of the LR process of RUTISHAUSER [1]. However even the best proofs available are unfit for beginners' consumption and the later developments of the process are not yet fully understood. We present here a discussion of the two-dimensional case, in its simplest form. The formal description of the process will be given in the n-dimensional case.

Let A be a complex $n \times n$ matrix. It is well-known that A can be written in the form

$$A = QR \tag{1}$$

where Q is unitary and R upper triangular. This is essentially the result of the Gram-Schmidt orthogonalization process. Moreover, if we require the diagonal elements of R to be positive, then the representation (1) is unique.

The QR-algorithm consists in deriving sequences of matrices $\{A_n\}$, $\{Q_n\}$, $\{R_n\}$ from $A = A_1$ by repeated use of (1). Given $A_n = Q_n R_n$ we form the reversed product $A_{n+1} = R_n Q_n$ and factorize this as $A_{n+1} = Q_{n+1} R_{n+1}$. Since

$$A_{n+1} = R_n Q_n = (Q_n^* Q_n) R_n Q_n = Q_n^* (Q_n R_n) Q_n = Q_n^* A_n Q_n \tag{2}$$

the matrices $\{A_n\}$ are all (unitarily) similar to A and have the same characteristic values as A. The basic fact is that, in certain circumstances, the sequence $\{A_n\}$ converges geometrically to an upper triangular matrix, which has the characteristic values of A on the diagonal. For discussions of this result see the original papers of H. RUTISHAUSER [1], J.G.F. FRANCIS [2], V.N. KUBLANOWSKAJA [3] and more recent work of B. PARLETT [4, 5, 6], A.S. HOUSEHOLDER [7], G.W. STEWART [8] and J.H. WILKINSON [9].

In practice, appropriate "shifts" are introduced and quadratic convergence can be obtained.

For simplicity we discuss the real two dimensional case. Let

$$A=\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$$

be real and unimodular, det $A=1$. We compute the QR decomposition. If, where $c=\cos\theta$, $s=\sin\theta$,

$$A=A_1=\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}=\begin{bmatrix} c & s \\ s & -c \end{bmatrix}\begin{bmatrix} \xi & \eta \\ 0 & \zeta \end{bmatrix}$$

then

$$c=\alpha\,(\alpha^2+\gamma^2)^{-1/2},\ s=\gamma\,(\alpha^2+\gamma^2)^{-1/2}.$$

Next

$$A_2=\begin{bmatrix} c & s \\ s & -c \end{bmatrix}\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}\begin{bmatrix} c & s \\ s & -c \end{bmatrix}=\begin{bmatrix} (\alpha+\delta-(\alpha/(\alpha^2+\gamma^2)) & \gamma-\beta-(\gamma/(\alpha^2+\gamma^2)) \\ -\gamma/(\alpha^2+\gamma^2) & \alpha/(\alpha^2+\gamma^2) \end{bmatrix}.$$

If we write

$$A_n=\begin{bmatrix} \alpha_n & \beta_n \\ \gamma_n & \delta_n \end{bmatrix}$$

then the recurrence relations determining α_{n+1}, γ_{n+1} are

$$\begin{cases} \alpha_{n+1}=(\alpha_1+\delta_1)-(\alpha_n/(\alpha_n^2+\gamma_n^2)), \\ \gamma_{n+1}=-\gamma_n/(\alpha_n^2+\gamma_n^2), \end{cases} \qquad n=1,2,\ldots \tag{3}$$

What we have to prove from (3) is that, in certain circumstances

$$\alpha_n\to\lambda,\ \gamma_n\to 0$$

where λ is an appropriate characteristic value of A. The solution to non-linear systems of difference equations such as (3) is not usually easy.

We assume that A has a dominant characteristic value λ. This means that A has distinct real characteristic values which are reciprocal, since A is unimodular. We may assume that λ is positive for otherwise we could deal with $-A$. Hence

$\lambda+\lambda^{-1}=\alpha+\delta=k$, say, where $k>2$.

We discuss an example first. We normalize the matrix

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

to a unimodular

$$A=A_1=\begin{bmatrix} 2/\sqrt{3} & -1/\sqrt{3} \\ -1\sqrt{3} & 2/\sqrt{3} \end{bmatrix},$$

with characteristic roots $\sqrt{3}$, $1/\sqrt{3}$.

Application of the relations (3) gives

	$n=1$	$n=2$	$n=3$	$n=4$
$\alpha_n/\sqrt{3}=$	$2/3$	$14/15$	$122/123$	$1094/1095$
$\gamma_n/\sqrt{3}=$	$-1/3$	$1/5$	$-3/41$	$9/365$

which indicates that $\alpha_n\to\sqrt{3}$, $\gamma_n\to 0$. The general form of α_n, γ_n can be conjectured from the above table and established by induction. We find

$$\frac{\alpha_n}{\sqrt{3}}=\frac{3\times 9^{n-1}+1}{3(9^{n-1}+1)},\ \frac{(-1)^n\gamma_n}{\sqrt{3}}=\frac{2\times 3^{n-1}}{3(9^{n-1}+1)},$$

so that the convergence of $\{\gamma_n\}$ is ultimately geometric with common ratio $1/3=\lambda^{-2}$ while that of $\{\alpha_n\}$ is ultimately geometric with common ratio $1/9=\lambda^{-4}$.

In the general real 2×2 case we can prove that

$$\left\{\begin{aligned} \alpha_{n+1}&=\frac{p(\alpha^2+\gamma^2+\alpha q)-(\alpha+q)}{(\alpha+q)^2+\gamma^2} \\ \gamma_{n+1}&=(-1)^n\frac{\gamma(pq+1)}{(\alpha+q)^2+\gamma^2} \end{aligned}\right. \tag{4}$$

where $v_n=\lambda^n-\lambda^{-n}$, $p=p_n=v_{n+1}/v_n$, $q=q_n=-v_{n-1}/v_n$. Since $p_n\sim\lambda$, $q_n\sim-\lambda^{-1}$ this gives $\alpha_n\to\lambda$, $\gamma_n\to 0$.

In order to establish (4) we write $x_n=\alpha_n$, $y_n=(-1)^n\gamma_n$ in (3) to get

$$x_{n+1}=k-(x_n/(x_n^2+y_n^2)),\ y_{n+1}=y_n/(x_n^2+y_n^2), \tag{5}$$

which we combine as $z_{n+1}=k-\bar{z}_n/(z_n\bar{z}_n)$, i.e.,

$$z_{n+1}=k-z_n^{-1},\ n=1,2,\ldots, \tag{6}$$

where $z_r=x_r+iy_r$, $\bar{z}_r=x_r-iy_r$.

We have therefore reduced our problem to that of the iteration of the fractional linear transformation

$$w=\frac{kz-1}{z}. \tag{7}$$

This is a well-known problem. An essentially geometric solution is given, e.g., by T. J. I'A. BROMWICH [10, p. 22, ex. 4]. This depends on the fact that (7) can be represented in the form

$$\frac{w-\lambda}{w-\lambda^{-1}}=\lambda^{-2}\left(\frac{z-\lambda}{z-\lambda^{-1}}\right) \tag{7'}$$

which gives

$$\frac{z_{n+1}-\lambda}{z_{n+1}-\lambda^{-1}}=\lambda^{-2n}\left(\frac{z_1-\lambda}{z_1-\lambda^{-1}}\right),$$

so that $z_n\to\lambda$, as required. We can derive (7′) using the fact that a fractional linear transformation with fixed points at $0,\infty$ is necessarily linear, or by using the cross-ratio property of fractional linear transformations. For details compare, e.g. CARATHÉODORY [11, p. 14] or KREYSZIG [12, p. 503].

A second method is simply to establish, by induction,

$$z_{n+1}=\frac{pz_1+1}{z_1+q} \tag{8}$$

where $z_1=a-iy$ and $p=p_n$, $q=q_n$ are as defined above. Taking the real and imaginary parts of (8) gives (4).

Our third method, preferable in the matrix context, follows. We begin by recalling that if

$$W=\frac{\alpha w+\beta}{\gamma w+\delta},\quad w=\frac{az+b}{cz+d}$$

then

$$W=\frac{Az+B}{Cz+D}\quad\text{where}\quad \begin{bmatrix}A & B\\ C & D\end{bmatrix}=\begin{bmatrix}\alpha & \beta\\ \gamma & \delta\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix}.$$

Thus the iteration of a fractional linear transformation is equivalent to the powering of a matrix.

We use the following result.

Lemma. If M is a 2×2 matrix with distinct characteristic values λ, μ then

$$M^n=\begin{bmatrix} ad\lambda^n-bc\mu^n & -ab(\lambda^n-\mu^n) \\ cd(\lambda^n-\mu^n) & -bc\lambda^n+ad\mu^n \end{bmatrix} \tag{9}$$

where

$$T=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is a unimodular matrix which diagonalizes M.

Proof. Either by induction or let $D=\operatorname{diag}[\lambda,\ \mu]$ so that $T^{-1}MT=D$, $M=TDT^{-1}$ and

$$\begin{aligned} M^n &= \{TDT^{-1}\}\{TDT^{-1}\}\dots\{TDT^{-1}\} \\ &= TD^nT^{-1} \\ &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \lambda^n & 0 \\ 0 & \mu^n \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \end{aligned}$$

which gives (9), on multiplying out.

We now observe that when $A=\begin{bmatrix} k & -1 \\ 1 & 0 \end{bmatrix}$ then

$$T=(\lambda-\lambda^{-1})^{-1/2}\begin{bmatrix} \lambda & \lambda^{-1} \\ 1 & 1 \end{bmatrix},$$

where λ is as already defined. The rest of the discussion is a matter of elementary algebra.

Writing as before, $v_n=\lambda^n-\lambda^{-n}$, we find from (9)

$$z_{n+1}\equiv x_{n+1}+iy_{n+1}=\frac{v_{n+1}(x_1+iy_1)-v_n}{v_n(x_1+iy_1)-v_{n-1}}.$$

Multiplying above and below on the right by $v_n(x_1-iy_1)-v_{n-1}$ and equating real and imaginary parts shows that

$$\begin{cases} x_{n+1}=\{v_nv_{n+1}(x_1^2+y_1^2)+v_nv_{n-1}-x_1(v_n^2+v_{n+1}v_{n-1})\}/D, \\ y_{n+1}=(v_n^2-v_{n+1}v_{n-1})y_1/D, \end{cases}$$

where

$$D = v_n^2(x_1^2 + y_1^2) - 2v_{n-1}v_n x_1 + v_{n-1}^2 .$$

This is another form of (4). We find that, as $n \to \infty$,

$$\begin{cases} x_{n+1} - \lambda = [2(\lambda - \lambda^{-1})\{a^2 + \gamma^2 + 1 - ka\} + O(\lambda^{-2n})]/D, \\ y_{n+1} = (\lambda - \lambda^{-1})^2 \gamma / D, \end{cases} \tag{10}$$

where

$$D = [(a - \lambda^{-1})^2 + \gamma^2]\lambda^{2n} + O(1) = O(\lambda^{2n}).$$

The relations (10) establish the convergence of the *QR*-process.

Note that when the matrix A is symmetric, as well as unimodular, we have $a\delta - \gamma^2 = 1$, i.e., $a(k-a) - \gamma^2 = 1$, i.e. $a^2 + \gamma^2 + 1 = ka$ so that (10) gives $x_n - \lambda = O(\lambda^{-4n})$, $y_n = O(\lambda^{-2n})$, in agreement with the numerical results in the special case.

Erwin Kreyszig (University of Windsor, Canada) and
John Todd (California Institute of Technology, Pasadena, USA)

REFERENCES

[1] H. RUTISHAUSER, *Solution of Eigenvalue Problems with the LR-Transformation,* in: *Further contributions to the solution of simultaneous linear equations and the determination of eigenvalues,* National Bureau of Standards, Applied Math. Series *49*, 47-81 (1958).
[2] J.G.F. FRANCIS, *The QR-Transformation,* Parts 1, 2, Computer J. *4*, 265-271, 332-345 (1961/2).
[3] V.N. KUBLANOWSKAJA, *On Some Algorithms for the Solution of the Complete Problem of Proper Values,* Z. Vyčisl. Mat. i Mat. Fiz. *1*, 555-570 (1961); U.S.S.R. Computational Math. and Math. Phys. *1*, 637-657 (1961).
[4] B. PARLETT, *Convergence of the QR-Algorithm,* Numer. Math. *7*, 187-193 (1965).
[5] B. PARLETT, *The Development and Use of Methods of LR Type,* SIAM Review *6*, 275-295 (1964).
[6] B. PARLETT, *Présentation géométrique des méthodes de calcul des valeurs propres,* Numer. Math. *21*, 223-233 (1973).
[7] A.S. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis* (Blaisdell, 1964).
[8] G.W. STEWART, *Introduction to Matrix Computations* (Academic Press, New York 1973).
[9] J.H. WILKINSON, *The Algebraic Eigenvalue Problem* (Clarendon Press, Oxford 1965).
[10] T.J.I'A. BROMWICH, *An Introduction to the Theory of Infinite Series,* 2nd ed. (Macmillan, London 1926).
[11] C. CARATHÉODORY, *Conformal Representation* (University Press, Cambridge 1937).
[12] E. KREYSZIG, *Advanced Engineering Mathematics,* 3rd ed. (Wiley, New York 1972).