

24.5 The QR Algorithm

Chapter 11 §08 brought a square matrix to upper Hessenberg form with reflections and then stopped, with a warning attached: no finite procedure built from arithmetic and square roots can carry that reduction on to triangular form, because such a procedure would extract the roots of an arbitrary polynomial from its coefficients, and the theorem of Abel and Ruffini forbids it — a theorem this book quotes, there and here, and does not prove. Every eigenvalue method must therefore iterate, and nothing proved below rests on that quotation. The iteration this section develops is the one Chapter 11 named in a single sentence — factor $\mathbf{Z} = \mathbf{QR}$, then replace $\mathbf{Z}$ by $\mathbf{RQ}$ — and we now ask why it works, why it is always started from a Hessenberg matrix, and why it is always run with a shift.

24.5.1 Factor and multiply back

Section 4 iterated a matrix against a **vector**: the power method drives $\mathbf{x}^{(k)} = \mathbf{A}\mathbf{x}^{(k-1)}$ towards the dominant eigenvector, one eigenvalue at a time. The whole spectrum at once needs an iteration on **matrices**, and it must be an iteration by similarity, or the spectrum would not survive it. Chapter 11's reflections already gave us a supply of similarities that cost little: $\mathbf{A} \mapsto \mathbf{Q}^*\mathbf{A}\mathbf{Q}$ with $\mathbf{Q}$ unitary. The question is which $\mathbf{Q}$ to take.

Here is the answer, and it looks like sleight of hand. Factor $\mathbf{A} = \mathbf{QR}$ and then multiply the two factors back **in the other order**. Since $\mathbf{Q}^*\mathbf{Q} = \mathbf{I}$, the product $\mathbf{RQ}$ is $\mathbf{Q}^*(\mathbf{QR})\mathbf{Q} = \mathbf{Q}^*\mathbf{A}\mathbf{Q}$, so nothing has been lost; and the reversal shifts weight from the upper triangle into the lower one in a way that, repeated, empties the lower triangle altogether.

Factor the matrix, then put the factors back the wrong way round, and repeat.

Definition 24.5.1. The QR Iteration

Let $\mathbf{A} \in M_n(F)$, where $F = \mathbb{R}$ or $F = \mathbb{C}$. Set $\mathbf{A}^{(1)} = \mathbf{A}$, and for $k \geq 1$ define

$$\mathbf{A}^{(k)} = \mathbf{Q}^{(k)}\mathbf{R}^{(k)}, \quad \mathbf{A}^{(k+1)} = \mathbf{R}^{(k)}\mathbf{Q}^{(k)},$$

where $\mathbf{A}^{(k)} = \mathbf{Q}^{(k)}\mathbf{R}^{(k)}$ is a factorization with $\mathbf{Q}^{(k)} \in M_n(F)$ unitary and $\mathbf{R}^{(k)} \in M_n(F)$ upper triangular. The sequence $\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \dots$ is the **unshifted QR iteration** started at $\mathbf{A}$.

In words: at every step we factor the current matrix, then form the product of its factors in the reversed order, and that product is the next matrix. The superscript in parentheses counts steps, never coordinates.

Well-definedness. A factorization exists at every step, for every square matrix and with no hypothesis on its columns, by Theorem 11.8.4. It need not be unique: if $\mathbf{D}$ is diagonal with $|d_j| = 1$, then $(\mathbf{QD})(\mathbf{D}^*\mathbf{R})$ is another one. When $\mathbf{A}^{(k)}$ is **invertible** the ambiguity is exactly that, and demanding a positive real diagonal in $\mathbf{R}^{(k)}$ removes it, by Theorem 11.8.1; and if $\mathbf{A}$ is invertible then so is every $\mathbf{A}^{(k)}$, being similar to $\mathbf{A}$. We do not need the normalization for anything below, so the definition leaves the choice open.

Proposition 24.5.1. Every Step Is a Unitary Similarity

In the notation of Definition 24.5.1, $\mathbf{A}^{(k+1)} = (\mathbf{Q}^{(k)})^*\mathbf{A}^{(k)}\mathbf{Q}^{(k)}$ for every k . Consequently all the matrices $\mathbf{A}^{(k)}$ are unitarily similar to $\mathbf{A}$, and they share its characteristic polynomial, its

eigenvalues with multiplicities, its trace, and its determinant.

Proof. Since $\mathbf{Q}^{(k)}$ is unitary, $(\mathbf{Q}^{(k)})^* \mathbf{Q}^{(k)} = \mathbf{I}_n$ (Definition 11.7.2), so

$$(\mathbf{Q}^{(k)})^* \mathbf{A}^{(k)} \mathbf{Q}^{(k)} = (\mathbf{Q}^{(k)})^* \mathbf{Q}^{(k)} \mathbf{R}^{(k)} \mathbf{Q}^{(k)} = \mathbf{R}^{(k)} \mathbf{Q}^{(k)} = \mathbf{A}^{(k+1)}.$$

A unitary matrix is invertible with $\mathbf{Q}^{-1} = \mathbf{Q}^*$, so this exhibits $\mathbf{A}^{(k+1)}$ as similar to $\mathbf{A}^{(k)}$ (Definition 4.2.2); composing these similarities for $1, \dots, k-1$ makes $\mathbf{A}^{(k)}$ similar to $\mathbf{A}^{(1)} = \mathbf{A}$. Similar matrices have the same characteristic polynomial (Theorem 7.10.3), hence the same eigenvalues with multiplicities, and the same trace and determinant. This proves the proposition. ■

Two examples fix what the iteration does and does not do.

Example. Two matrices the iteration cannot move

Determine the QR iterates of $\mathbf{A} = \text{diag}(3, 1)$ and of the rotation $\mathbf{B} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Solution. For $\mathbf{A}$: the factorization $\mathbf{A} = \mathbf{I}_2 \mathbf{A}$ has $\mathbf{I}_2$ unitary and $\mathbf{A}$ upper triangular with positive diagonal, so by the uniqueness in Theorem 11.8.1 it is *the* factorization. Then $\mathbf{A}^{(2)} = \mathbf{A} \mathbf{I}_2 = \mathbf{A}$, and the iteration stands still. It has nothing to do: $\mathbf{A}$ is already triangular and its diagonal already lists the eigenvalues (Theorem 9.7.4).

For $\mathbf{B}$: here $\mathbf{B}$ is itself unitary, since its columns $(0, 1)$ and $(-1, 0)$ are orthonormal, so $\mathbf{B} = \mathbf{B} \mathbf{I}_2$ is a factorization with $\mathbf{I}_2$ upper triangular with positive diagonal — again the unique one. So $\mathbf{B}^{(2)} = \mathbf{I}_2 \mathbf{B} = \mathbf{B}$, and the iteration stands still here too. This time it has everything to do: the eigenvalues of $\mathbf{B}$ are $\pm i$, and $\mathbf{B}$ is not triangular. The iteration is stuck.

The second matrix is the section's standing negative example, and the reason it is stuck is visible already: $|i| = |-i|$. We return to it after the convergence theorem, and repair it with a shift.

Warning. Reversing the factors is not a similarity for just any factorization. The step works because $\mathbf{Q}$ is unitary, so that $\mathbf{Q}^{-1} = \mathbf{Q}^*$ and $\mathbf{RQ} = \mathbf{Q}^* \mathbf{A} \mathbf{Q}$. Do the same with an $\mathbf{LU}$ factorization and you get $\mathbf{UL} = \mathbf{L}^{-1} \mathbf{A} \mathbf{L}$, which is still a similarity — but with $\mathbf{L}^{-1}$ in place of $\mathbf{L}^*$, and $\mathbf{L}^{-1}$ can be enormous. That is the older $\mathbf{LR}$ iteration, and its instability under rounding is exactly the instability of elimination without pivoting met in Section 2.

24.5.2 The iteration is the power method in disguise

Nothing said so far suggests that the iterates go anywhere. The identity that explains them is the following, and it is the heart of the section: *the accumulated factors of k steps are a QR factorization of the k -th power of $\mathbf{A}$* . Powers of $\mathbf{A}$ are what the power method of Section 4 runs on, so the identity says the two methods are the same method, one applied to a vector and the other to a whole basis at once.

Write

$$\mathbf{U}_k = \mathbf{Q}^{(1)}\mathbf{Q}^{(2)} \cdots \mathbf{Q}^{(k)}, \quad \mathbf{T}_k = \mathbf{R}^{(k)} \cdots \mathbf{R}^{(2)}\mathbf{R}^{(1)}.$$

The first is unitary, as a product of unitary matrices (Proposition 11.7.4), and the second is upper triangular, as a product of upper triangular matrices.

Theorem 24.5.2. QR Iterates Factor the Powers

Let $\mathbf{A} \in M_n(F)$ and let $\mathbf{A}^{(k)}$, $\mathbf{Q}^{(k)}$, $\mathbf{R}^{(k)}$ be as in Definition 24.5.1. Then for every $k \geq 1$,

$$\mathbf{A}^{(k+1)} = \mathbf{U}_k^* \mathbf{A} \mathbf{U}_k \quad \text{and} \quad \mathbf{A}^k = \mathbf{U}_k \mathbf{T}_k.$$

In particular $\mathbf{U}_k \mathbf{T}_k$ is a QR factorization of $\mathbf{A}^k$.

Idea. Induction on k , with the two statements proved together, because each feeds the other. The similarity statement is Proposition 24.5.1 telescoped. For the factorization, the move to find is the one that turns $\mathbf{A} \mathbf{U}_k$ into something with a $\mathbf{Q}$ at the front: the similarity statement, read backwards, says $\mathbf{A} \mathbf{U}_k = \mathbf{U}_k \mathbf{A}^{(k+1)}$, and $\mathbf{A}^{(k+1)}$ is by definition $\mathbf{Q}^{(k+1)} \mathbf{R}^{(k+1)}$. One line of bookkeeping then converts $\mathbf{A} \cdot \mathbf{A}^k$ into $\mathbf{U}_{k+1} \mathbf{T}_{k+1}$.

Proof. Both statements by induction on k .

For $k = 1$: $\mathbf{U}_1 = \mathbf{Q}^{(1)}$ and $\mathbf{T}_1 = \mathbf{R}^{(1)}$, so $\mathbf{U}_1 \mathbf{T}_1 = \mathbf{Q}^{(1)} \mathbf{R}^{(1)} = \mathbf{A}^{(1)} = \mathbf{A}$, and $\mathbf{U}_1^* \mathbf{A} \mathbf{U}_1 = \mathbf{A}^{(2)}$ by Proposition 24.5.1.

Let $k \geq 1$ and suppose both statements hold for k . By Proposition 24.5.1 applied at step $k + 1$ and the inductive hypothesis,

$$\mathbf{A}^{(k+2)} = (\mathbf{Q}^{(k+1)})^* \mathbf{A}^{(k+1)} \mathbf{Q}^{(k+1)} = (\mathbf{Q}^{(k+1)})^* \mathbf{U}_k^* \mathbf{A} \mathbf{U}_k \mathbf{Q}^{(k+1)} = \mathbf{U}_{k+1}^* \mathbf{A} \mathbf{U}_{k+1},$$

where the last equality uses $\mathbf{U}_{k+1} = \mathbf{U}_k \mathbf{Q}^{(k+1)}$ and $(\mathbf{XY})^* = \mathbf{Y}^* \mathbf{X}^*$. That is the first statement for $k + 1$.

For the second, multiply the inductive similarity $\mathbf{A}^{(k+1)} = \mathbf{U}_k^* \mathbf{A} \mathbf{U}_k$ on the left by $\mathbf{U}_k$, which is legitimate because $\mathbf{U}_k \mathbf{U}_k^* = \mathbf{I}_n$:

$$\mathbf{A} \mathbf{U}_k = \mathbf{U}_k \mathbf{A}^{(k+1)} = \mathbf{U}_k \mathbf{Q}^{(k+1)} \mathbf{R}^{(k+1)} = \mathbf{U}_{k+1} \mathbf{R}^{(k+1)}.$$

Hence, using the inductive factorization $\mathbf{A}^k = \mathbf{U}_k \mathbf{T}_k$,

$$\mathbf{A}^{k+1} = \mathbf{A} (\mathbf{U}_k \mathbf{T}_k) = (\mathbf{A} \mathbf{U}_k) \mathbf{T}_k = \mathbf{U}_{k+1} \mathbf{R}^{(k+1)} \mathbf{T}_k = \mathbf{U}_{k+1} \mathbf{T}_{k+1},$$

since $\mathbf{T}_{k+1} = \mathbf{R}^{(k+1)}\mathbf{T}_k$ by definition. As $\mathbf{U}_{k+1}$ is unitary and $\mathbf{T}_{k+1}$ upper triangular, this is a QR factorization of $\mathbf{A}^{k+1}$, and the induction is complete. ■

The consequence to keep is about the **first column**. Comparing first columns in $\mathbf{A}^k = \mathbf{U}_k\mathbf{T}_k$ and writing $\tau_k = (\mathbf{T}_k)_{11}$, we get

$$\mathbf{A}^k \mathbf{e}_1 = \mathbf{U}_k \mathbf{T}_k \mathbf{e}_1 = \tau_k \mathbf{U}_k \mathbf{e}_1, \quad (24.5.1)$$

because the first column of an upper triangular matrix is $\tau_k \mathbf{e}_1$. So whenever $\tau_k \neq 0$, which the Quick check just below records for invertible $\mathbf{A}$, the first column of $\mathbf{U}_k$ is the normalized vector $\mathbf{A}^k \mathbf{e}_1 / \|\mathbf{A}^k \mathbf{e}_1\|$, up to a scalar of modulus one — which is precisely the k -th iterate of the power method (Theorem 24.4.1) started at $\mathbf{e}_1$. The QR algorithm runs the power method in its first column, and, as the next theorem shows, that alone forces the first column of $\mathbf{A}^{(k+1)}$ to converge.

Quick check. Why is $\tau_k \neq 0$ when $\mathbf{A}$ is invertible?

Solution. Each $\mathbf{A}^{(j)}$ is similar to $\mathbf{A}$ (Proposition 24.5.1), hence invertible, and $\mathbf{R}^{(j)} = (\mathbf{Q}^{(j)})^* \mathbf{A}^{(j)}$ is then invertible as a product of invertible matrices. So $\mathbf{T}_k$ is invertible, being a product of invertible matrices; an upper triangular matrix is invertible exactly when its diagonal entries are all nonzero (Lemma 3.6.2), so $\tau_k = (\mathbf{T}_k)_{11} \neq 0$.

24.5.3 What converges, and how fast

Theorem 24.5.3. Convergence of the First Column

Let $\mathbf{A} \in M_n(\mathbb{C})$ be invertible and diagonalizable, say $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}$ with $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$ and with $\mathbf{v}_1, \dots, \mathbf{v}_n$ the columns of $\mathbf{X}$. Assume a **single dominant eigenvalue**,

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n| > 0,$$

and write $\mathbf{e}_1 = c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n$, assuming $c_1 \neq 0$. Put $t = |\lambda_2/\lambda_1| < 1$. Then there is a constant C , depending on $\mathbf{A}$ and $\mathbf{X}$ but not on k , such that the unshifted QR iterates satisfy

$$\|\mathbf{A}^{(k+1)} \mathbf{e}_1 - \lambda_1 \mathbf{e}_1\|_2 \leq C t^k \quad (k \geq 1).$$

In particular $a_{11}^{(k)} \rightarrow \lambda_1$ and $a_{i1}^{(k)} \rightarrow 0$ for every $i \geq 2$, each at the rate t^k . For $n = 2$ this says that the iterates converge to upper triangular form, with $a_{11}^{(k)} \rightarrow \lambda_1$ and $a_{22}^{(k)} \rightarrow \lambda_2$.

Idea. By Eq. 24.5.1 the first column $\mathbf{u}_1$ of $\mathbf{U}_k$ is the power-method iterate, so it is close to the dominant eigenvector, and the vector $\mathbf{A}\mathbf{u}_1 - \lambda_1 \mathbf{u}_1$ is small. Now read $\mathbf{A}^{(k+1)} = \mathbf{U}_k^* \mathbf{A} \mathbf{U}_k$ column by column: its first column is $\mathbf{U}_k^* \mathbf{A} \mathbf{u}_1$, and subtracting $\lambda_1 \mathbf{e}_1 = \mathbf{U}_k^* (\lambda_1 \mathbf{u}_1)$ leaves $\mathbf{U}_k^* (\mathbf{A}\mathbf{u}_1 - \lambda_1 \mathbf{u}_1)$. A unitary matrix does not change lengths, so the whole error in the first

column *equals* the length of that small vector. Nothing else is needed.

Proof. Write $\mathbf{u}_1 = \mathbf{U}_k \mathbf{e}_1$ for the first column of $\mathbf{U}_k$, suppressing k from the notation.

Step 1: the error in the first column is a residual. Since $\mathbf{U}_k$ is unitary and $\mathbf{A}^{(k+1)} = \mathbf{U}_k^* \mathbf{A} \mathbf{U}_k$ by Theorem 24.5.2,

$$\mathbf{A}^{(k+1)} \mathbf{e}_1 - \lambda_1 \mathbf{e}_1 = \mathbf{U}_k^* \mathbf{A} \mathbf{u}_1 - \lambda_1 \mathbf{U}_k^* \mathbf{u}_1 = \mathbf{U}_k^* (\mathbf{A} \mathbf{u}_1 - \lambda_1 \mathbf{u}_1),$$

using $\mathbf{u}_1 = \mathbf{U}_k \mathbf{e}_1$ twice. A unitary matrix preserves the Euclidean norm (Theorem 11.7.3), so

$$\|\mathbf{A}^{(k+1)} \mathbf{e}_1 - \lambda_1 \mathbf{e}_1\|_2 = \|\mathbf{A} \mathbf{u}_1 - \lambda_1 \mathbf{u}_1\|_2. \quad (*)$$

Step 2: $\mathbf{u}_1$ is a power-method iterate. By Eq. 24.5.1, $\mathbf{A}^k \mathbf{e}_1 = \tau_k \mathbf{u}_1$ with $\tau_k \neq 0$, as the Quick check above records. Since $\mathbf{A} \mathbf{v}_i = \lambda_i \mathbf{v}_i$,

$$\mathbf{A}^k \mathbf{e}_1 = \sum_{i=1}^n c_i \lambda_i^k \mathbf{v}_i = c_1 \lambda_1^k (\mathbf{v}_1 + \mathbf{z}_k), \quad \mathbf{z}_k = \sum_{i \geq 2} \frac{c_i}{c_1} \left(\frac{\lambda_i}{\lambda_1} \right)^k \mathbf{v}_i,$$

which is legitimate because $c_1 \neq 0$ and $\lambda_1 \neq 0$. Every ratio obeys $|\lambda_i/\lambda_1| \leq t$, so

$$\|\mathbf{z}_k\|_2 \leq M t^k, \quad M = \sum_{i \geq 2} \frac{|c_i|}{|c_1|} \|\mathbf{v}_i\|_2,$$

by the triangle inequality (Corollary 11.1.6). Put $\mathbf{y}_k = \mathbf{v}_1 + \mathbf{z}_k$. Then $\mathbf{u}_1 = \mathbf{A}^k \mathbf{e}_1 / \tau_k = \omega_k \mathbf{y}_k / \|\mathbf{y}_k\|_2$ for the scalar $\omega_k = c_1 \lambda_1^k \|\mathbf{y}_k\|_2 / \tau_k$, and $|\omega_k| = 1$ because $\|\mathbf{u}_1\|_2 = 1$.

Step 3: the residual is small. Since $\mathbf{A} \mathbf{v}_1 = \lambda_1 \mathbf{v}_1$, the vector $\mathbf{v}_1$ drops out of $(\mathbf{A} - \lambda_1 \mathbf{I}) \mathbf{y}_k$, leaving

$$\begin{aligned} \mathbf{A} \mathbf{u}_1 - \lambda_1 \mathbf{u}_1 &= \omega_k \frac{(\mathbf{A} - \lambda_1 \mathbf{I}) \mathbf{z}_k}{\|\mathbf{y}_k\|_2}, \\ \text{so } \|\mathbf{A} \mathbf{u}_1 - \lambda_1 \mathbf{u}_1\|_2 &\leq \frac{(\|\mathbf{A}\|_2 + |\lambda_1|) M t^k}{\|\mathbf{y}_k\|_2}. \end{aligned}$$

Choose k_0 with $M t^{k_0} \leq \frac{1}{2} \|\mathbf{v}_1\|_2$, possible because $t < 1$. For $k \geq k_0$ the triangle inequality gives $\|\mathbf{y}_k\|_2 \geq \|\mathbf{v}_1\|_2 - M t^k \geq \frac{1}{2} \|\mathbf{v}_1\|_2 > 0$, whence

$$\|\mathbf{A} \mathbf{u}_1 - \lambda_1 \mathbf{u}_1\|_2 \leq C_1 t^k, \quad C_1 = \frac{2(\|\mathbf{A}\|_2 + |\lambda_1|) M}{\|\mathbf{v}_1\|_2}.$$

For the finitely many $k < k_0$ the crude bound $\|\mathbf{A} \mathbf{u}_1 - \lambda_1 \mathbf{u}_1\|_2 \leq \|\mathbf{A}\|_2 + |\lambda_1|$ holds, since $\|\mathbf{u}_1\|_2 = 1$. So the displayed estimate holds for all $k \geq 1$ with

$$C = \max \left\{ C_1, (\|\mathbf{A}\|_2 + |\lambda_1|) t^{-k_0} \right\},$$

and combining with (*) proves the main estimate. The entrywise statements follow because each coordinate of a vector is at most its Euclidean norm.

For $n = 2$, the estimate forces $a_{21}^{(k)} \rightarrow 0$ and $a_{11}^{(k)} \rightarrow \lambda_1$; and $a_{22}^{(k)} = \operatorname{tr} \mathbf{A}^{(k)} - a_{11}^{(k)} = \lambda_1 + \lambda_2 - a_{11}^{(k)} \rightarrow \lambda_2$, the trace being preserved by Proposition 24.5.1. This proves the theorem. ■

What is proved here, and what is not. The theorem above settles the first column, for every n , and with it the whole 2×2 case. The general statement is this: if $\mathbf{A}$ is diagonalizable with $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0$, **and** if $\mathbf{X}^{-1}$ admits an $\mathbf{LU}$ factorization without row swaps (Theorem 3.6.4), then the whole of $\mathbf{A}^{(k)}$ below the diagonal tends to zero, with

$$a_{i+1,i}^{(k)} = O(|\lambda_{i+1}/\lambda_i|^k) \quad \text{and} \quad a_{ii}^{(k)} \rightarrow \lambda_i \quad (1 \leq i \leq n).$$

This is the classical convergence theorem of the algorithm, due to Rutishauser, Francis, and Kublanovskaya; **it is not proved here.** The proof runs through the factorization $\mathbf{A}^k = \mathbf{X} \mathbf{\Lambda}^k \mathbf{LU}$ of Theorem 24.5.2's identity, the observation that $\mathbf{\Lambda}^k \mathbf{L} \mathbf{\Lambda}^{-k} \rightarrow \mathbf{I}$, and a continuity argument for the QR factorization that this book has not developed. Nothing in this section or later depends on it: the Hessenberg and shifting results below are proved outright, and the worked example at the end exhibits the rates rather than assuming them. The extra hypothesis on $\mathbf{X}^{-1}$ is not cosmetic, and the strict ordering of the moduli is not either — the next example shows what happens without it.

Example. The unshifted algorithm on a rotation

Let $\mathbf{B} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Explain why no unshifted QR iteration can bring $\mathbf{B}$ closer to triangular form, and say which hypothesis of Theorem 24.5.3 fails.

Solution. By Example (Two matrices the iteration cannot move), $\mathbf{B}^{(k)} = \mathbf{B}$ for every k : the iteration is stationary, and $\mathbf{B}$ is as far from triangular as a 2×2 matrix can be, with $b_{21} = 1$. The eigenvalues are $\pm i$, so $|\lambda_1| = |\lambda_2| = 1$ and the hypothesis $|\lambda_1| > |\lambda_2|$ fails. The failure is not an artifact of the proof. Over $\mathbb{R}$ the matrix is orthogonally similar to no real triangular matrix at all, since a real triangular matrix has real eigenvalues; and over $\mathbb{C}$ the power method has nothing to converge to, because $\mathbf{B}^k \mathbf{e}_1$ simply rotates around the plane forever (Theorem 24.4.1).

Warning. Convergence is to triangular form, not to diagonal form. Even when the iterates do converge, the strictly upper triangular part generally does not: only the entries **below** the diagonal die out. What is more, in the convergence theorem quoted above — the one this section does not prove — the order in which the eigenvalues appear on the diagonal is decided by the moduli, largest first, so that a run it covers reveals λ_1 in the top-left corner and λ_n in the bottom-right. With shifts, which is how the algorithm is actually run, even that ordering goes away.

24.5.4 Hessenberg first

As written, one step costs a QR factorization of a full $n \times n$ matrix and a matrix product: that is $O(n^3)$ flops, in the counting convention of Proposition 24.2.3, and the count really does grow

like n^3 , since the product alone touches n^2 entries with n terms each. The iteration needs many steps, and no amount of cleverness about the steps rescues a method that spends n^3 operations on each of them.

The remedy is Chapter 11's. Reduce $\mathbf{A}$ once to upper Hessenberg form by Theorem 11.8.5, at a cost of $O(n^3)$ paid a single time, and then iterate. This pays off only if the pattern survives the step, which is the content of the next proposition — and it is the sentence Chapter 11 §08 closed on: “Chapter 24 develops that iteration and explains why every implementation starts with the reduction proved here.”

Proposition 24.5.4. A QR Step Preserves Hessenberg Form

Let $\mathbf{Z} \in M_n(\mathbb{R})$ be upper Hessenberg, $n \geq 2$. Then $\mathbf{Z}$ has a factorization $\mathbf{Z} = \mathbf{QR}$ with $\mathbf{Q}$ orthogonal and $\mathbf{R}$ upper triangular in which $\mathbf{Q}$ is a product of $n - 1$ Givens rotations and is **itself upper Hessenberg**; and for this factorization $\mathbf{RQ}$ is again upper Hessenberg. Producing $\mathbf{R}$ and then $\mathbf{RQ}$ costs about $6n^2$ flops.

Idea. A Hessenberg matrix has only $n - 1$ entries in the way of triangularity, one per column, and Definition 11.8.2 removes one entry per rotation. So the natural factorization to use is the rotation one, not the reflection one — this is the case Chapter 11 §08 flagged, where “the count reverses”. Then two bookkeeping facts finish it: the accumulated product of the rotations is Hessenberg, and upper triangular times upper Hessenberg is upper Hessenberg.

Proof. For $1 \leq i \leq n - 1$ let $\mathbf{G}_i = \mathbf{G}(i, i + 1; c_i, s_i)$ be a Givens rotation in the coordinates $i, i + 1$ (Definition 11.8.2). Working left to right, choose $\mathbf{G}_i$ to annihilate the entry in position $(i + 1, i)$ of the matrix produced so far: if that matrix has entries a in position (i, i) and b in position $(i + 1, i)$ with $(a, b) \neq (0, 0)$, take $c_i = a/r$ and $s_i = b/r$ with $r = \sqrt{a^2 + b^2}$, as in Chapter 11 §08; if $(a, b) = (0, 0)$, take $\mathbf{G}_i = \mathbf{I}_n$. This is the only entry below the diagonal in column i : the entries in rows $i + 2, \dots, n$ of that column were zero in $\mathbf{Z}$, because $\mathbf{Z}$ is upper Hessenberg, and the earlier rotations $\mathbf{G}_1, \dots, \mathbf{G}_{i-1}$ touch only rows $1, \dots, i$. Left multiplication by $\mathbf{G}_i$ recombines rows i and $i + 1$ only, so it cannot disturb the zeros already created: in a column $j \leq i - 1$ both of the rows being combined are already zero, since $i \geq j + 1$. Hence

$$\mathbf{R} = \mathbf{G}_{n-1} \cdots \mathbf{G}_1 \mathbf{Z}$$

is upper triangular, and $\mathbf{Z} = \mathbf{QR}$ with $\mathbf{Q} = \mathbf{G}_1^\top \mathbf{G}_2^\top \cdots \mathbf{G}_{n-1}^\top$, orthogonal as a product of orthogonal matrices.

Claim. $\mathbf{P}_m = \mathbf{G}_1^\top \cdots \mathbf{G}_m^\top$ is upper Hessenberg for $1 \leq m \leq n - 1$, and its columns $m + 2, \dots, n$ are $\mathbf{e}_{m+2}, \dots, \mathbf{e}_n$.

Proof. Induction on m . Each $\mathbf{G}_i^\top$ fixes $\mathbf{e}_j$ for $j \notin \{i, i + 1\}$, so $\mathbf{P}_m \mathbf{e}_j = \mathbf{e}_j$ whenever $j > m + 1$, which is the second assertion. For $m = 1$, $\mathbf{P}_1 = \mathbf{G}_1^\top$ differs from $\mathbf{I}_n$ only in rows and columns 1, 2, so its only nonzero entry below the diagonal is in position $(2, 1)$, and it is upper Hessenberg. Let $m \geq 2$ and suppose the claim holds for $m - 1$. Right multiplication by $\mathbf{G}_m^\top$ changes only columns m and $m + 1$, replacing them by the two combinations

$$c_m \mathbf{P}_{m-1} \mathbf{e}_m + s_m \mathbf{P}_{m-1} \mathbf{e}_{m+1} \quad \text{and} \quad -s_m \mathbf{P}_{m-1} \mathbf{e}_m + c_m \mathbf{P}_{m-1} \mathbf{e}_{m+1}.$$

By the inductive hypothesis $\mathbf{P}_{m-1} \mathbf{e}_m$ has zero entries in rows $m + 2, \dots, n$, being column m of an upper Hessenberg matrix, and $\mathbf{P}_{m-1} \mathbf{e}_{m+1} = \mathbf{e}_{m+1}$ since $m + 1 > (m - 1) + 1$. So both

combinations vanish in rows $m + 2, \dots, n$. Column m of $\mathbf{P}_m$ therefore vanishes below row $m + 1$, as an upper Hessenberg matrix requires; and column $m + 1$ vanishes below row $m + 1$, which is more than required. All other columns are unchanged, so $\mathbf{P}_m$ is upper Hessenberg. ■

By the Claim, $\mathbf{Q} = \mathbf{P}_{n-1}$ is upper Hessenberg. Finally,

$$(\mathbf{RQ})_{pq} = \sum_k r_{pk} q_{kq}$$

has a nonzero term only when $k \geq p$, because $\mathbf{R}$ is upper triangular, and $k \leq q + 1$, because $\mathbf{Q}$ is upper Hessenberg; together these force $p \leq q + 1$. So $\mathbf{RQ}$ is upper Hessenberg.

Cost. Rotation $\mathbf{G}_i$ alters rows $i, i + 1$ in columns $i, \dots, n$ only, since in the earlier columns both of those rows are already zero: that is $n - i + 1$ pairs of entries, each pair replaced by two combinations of two terms, so $6(n - i + 1)$ flops. Summing over $i = 1, \dots, n - 1$ gives $3n^2 + O(n)$. Forming $\mathbf{RQ} = \mathbf{R} \mathbf{G}_1^\top \cdots \mathbf{G}_{n-1}^\top$ applies $\mathbf{G}_i^\top$ on the right to columns $i, i + 1$ of the matrix built so far, and the same induction as in the Claim shows that those two columns vanish below row $i + 1$; so $6(i + 1)$ flops suffice, and summing gives $3n^2 + O(n)$ again. The total is $6n^2 + O(n)$, which proves the proposition. ■

So the reduction is done once, for $O(n^3)$, and every subsequent step costs $O(n^2)$ — a saving of a factor of n on each of the many steps. That is the whole reason every implementation begins with Theorem 11.8.5.

Remark. The proposition is stated over $\mathbb{R}$ because Definition 11.8.2 is. Over $\mathbb{C}$ one uses the complex rotations described in Chapter 11 §08 and the argument is unchanged. The conclusion also does not depend on which QR factorization is used, as long as $\mathbf{Z}$ is invertible: two factorizations then differ by $\mathbf{Q} \mapsto \mathbf{QD}$ and $\mathbf{R} \mapsto \mathbf{D}^* \mathbf{R}$ with $\mathbf{D}$ unitary diagonal, and $(\mathbf{D}^* \mathbf{R})(\mathbf{QD}) = \mathbf{D}^* (\mathbf{RQ}) \mathbf{D}$ has exactly the same pattern of zeros as $\mathbf{RQ}$. Invertibility is doing work there: for $\mathbf{Z} = \mathbf{e}_1 \mathbf{e}_3^\top \in M_3(\mathbb{R})$, which is upper Hessenberg and singular, the reversal permutation $\mathbf{Q}$ (with $\mathbf{Q} \mathbf{e}_j = \mathbf{e}_{4-j}$) and $\mathbf{R} = \mathbf{e}_3 \mathbf{e}_3^\top$ give an orthogonal $\mathbf{Q}$, an upper triangular $\mathbf{R}$, and $\mathbf{QR} = \mathbf{Z}$, while $\mathbf{RQ} = \mathbf{e}_3 \mathbf{e}_1^\top$ has a nonzero $(3, 1)$ entry and is not upper Hessenberg.

Quick check. A symmetric matrix is upper Hessenberg exactly when it is tridiagonal. What does Proposition 24.5.4 then say about a QR step applied to a symmetric tridiagonal matrix?

Solution. That the step returns a symmetric tridiagonal matrix. Symmetry survives because $\mathbf{RQ} = \mathbf{Q}^\top \mathbf{ZQ}$ is a congruence by an orthogonal matrix as well as a similarity, so $(\mathbf{Q}^\top \mathbf{ZQ})^\top = \mathbf{Q}^\top \mathbf{Z}^\top \mathbf{Q} = \mathbf{Q}^\top \mathbf{ZQ}$; and Hessenberg form survives by the proposition. A symmetric Hessenberg matrix has $z_{ij} = 0$ whenever $i > j + 1$, hence also whenever $j > i + 1$, which is tridiagonality. The cost then drops further, to $O(n)$ per step.

24.5.5 Shifts

The rate $|\lambda_{i+1}/\lambda_i|$ is a property of $\mathbf{A}$, and we cannot change $\mathbf{A}$. We can change the matrix we iterate on. For any $\mu \in F$, the matrix $\mathbf{A} - \mu\mathbf{I}$ has eigenvalues $\lambda_i - \mu$, so its ratios of moduli are $|\lambda_{i+1} - \mu|/|\lambda_i - \mu|$ — and a μ close to λ_n makes the last of these tiny. This is the same observation that turned the power method into inverse iteration in Section 4, where the shift μ was used to make $(\mathbf{A} - \mu\mathbf{I})^{-1}$ have a huge dominant eigenvalue (Theorem 24.4.3).

The **shifted QR step** is

$$\mathbf{A}^{(k)} - \mu_k \mathbf{I} = \mathbf{Q}^{(k)} \mathbf{R}^{(k)}, \quad \mathbf{A}^{(k+1)} = \mathbf{R}^{(k)} \mathbf{Q}^{(k)} + \mu_k \mathbf{I},$$

and it is still a unitary similarity: $\mathbf{R}^{(k)} \mathbf{Q}^{(k)} = (\mathbf{Q}^{(k)})^* (\mathbf{A}^{(k)} - \mu_k \mathbf{I}) \mathbf{Q}^{(k)}$, so adding $\mu_k \mathbf{I}$ back gives $\mathbf{A}^{(k+1)} = (\mathbf{Q}^{(k)})^* \mathbf{A}^{(k)} \mathbf{Q}^{(k)}$ exactly as in Proposition 24.5.1. The shift may be changed at every step, which is why the ratio can be driven down as the information improves. And since $\mathbf{A}^{(k)} - \mu_k \mathbf{I}$ is Hessenberg whenever $\mathbf{A}^{(k)}$ is, Proposition 24.5.4 applies to the shifted step: as it stands for a real μ_k , and in the complex form recorded in the Remark after it for a μ_k off the real line, which the second of the two shifts below produces.

Two shifts are standard. The **Rayleigh shift** takes $\mu_k = a_{nn}^{(k)}$, the trailing diagonal entry; when $\mathbf{A}$ is Hermitian this is the Rayleigh quotient $R_{\mathbf{A}^{(k)}}(\mathbf{e}_n)$ of Definition 17.1.1, hence a good eigenvalue estimate as soon as $\mathbf{e}_n$ is nearly an eigenvector. The **Wilkinson shift** takes μ_k to be the eigenvalue of the trailing 2×2 block of $\mathbf{A}^{(k)}$ nearer to $a_{nn}^{(k)}$; it costs one quadratic formula and, unlike the Rayleigh shift, it can leave the real line, which is what lets it break the deadlock of a rotation.

The extreme case says exactly why shifting works.

Proposition 24.5.5. An Exact Shift Deflates in One Step

Let $\mathbf{Z} \in M_n(F)$ be upper Hessenberg, $n \geq 2$, with all its subdiagonal entries $z_{i+1,i}$ **nonzero** — such a matrix is called **unreduced**, the word Lemma 17.4.5 used for the tridiagonal case. Let $\mu \in F$ be an eigenvalue of $\mathbf{Z}$. Let $\mathbf{Z} - \mu\mathbf{I} = \mathbf{Q}\mathbf{R}$ with $\mathbf{Q}$ unitary and $\mathbf{R}$ upper triangular, and put $\mathbf{Z}' = \mathbf{R}\mathbf{Q} + \mu\mathbf{I}$. Then the last row of $\mathbf{R}$ is zero, and the last row of $\mathbf{Z}'$ is $\mu \mathbf{e}_n^\top$. In particular $z'_{n,n-1} = 0$ and $z'_{nn} = \mu$.

Idea. An unreduced Hessenberg matrix cannot lose more than one dimension of rank: its first $n-1$ columns are independent whatever the matrix does, because the subdiagonal alone makes them so. Shifting by an eigenvalue makes the matrix singular, so the one missing dimension must be the last one, and in a QR factorization “the last dimension is missing” reads as a zero last row of $\mathbf{R}$.

Proof. Write $\mathbf{B} = \mathbf{Z} - \mu\mathbf{I}$, again upper Hessenberg with the same nonzero subdiagonal entries.

First, the columns $\mathbf{b}_1, \dots, \mathbf{b}_{n-1}$ of $\mathbf{B}$ are linearly independent. Delete the first row: what remains is the $(n-1) \times (n-1)$ matrix with entries $b_{i+1,j}$ for $1 \leq i, j \leq n-1$, which is **lower** triangular, since $b_{i+1,j} = 0$ for $i+1 > j+1$, that is for $i > j$; and its diagonal entries $b_{j+1,j}$ are nonzero. A triangular matrix with nonzero diagonal is invertible (Lemma 3.6.2), so those $n-1$ truncated columns are independent, and the full columns are independent a fortiori.

Second, $\text{rank } \mathbf{B} = n-1$: it is at least $n-1$ by the previous paragraph and at most $n-1$

because $\mathbf{B}$ is singular, μ being an eigenvalue of $\mathbf{Z}$.

Now $\mathbf{B} = \mathbf{Q}\mathbf{R}$ with $\mathbf{Q}$ invertible, so $\text{rank } \mathbf{R} = \text{rank } \mathbf{B} = n-1$ and the columns $\mathbf{R}\mathbf{e}_1, \dots, \mathbf{R}\mathbf{e}_{n-1}$ are independent, being $\mathbf{Q}^*\mathbf{b}_1, \dots, \mathbf{Q}^*\mathbf{b}_{n-1}$. Since $\mathbf{R}$ is upper triangular, its last row has the single possibly nonzero entry r_{nn} ; if $r_{nn} \neq 0$, then $\mathbf{R}$ would be invertible (Lemma 3.6.2 applied after noting that independence of the first $n-1$ columns forces $r_{11}, \dots, r_{n-1,n-1} \neq 0$), contradicting $\text{rank } \mathbf{R} = n-1$. Hence $r_{nn} = 0$ and the last row of $\mathbf{R}$ is zero.

A matrix with zero last row keeps it under multiplication on the right: $(\mathbf{R}\mathbf{Q})_{nj} = \sum_k r_{nk}q_{kj} = 0$. So the last row of $\mathbf{R}\mathbf{Q}$ is zero and the last row of $\mathbf{Z}' = \mathbf{R}\mathbf{Q} + \mu\mathbf{I}$ is $\mu\mathbf{e}_n^\top$, as claimed. ■

The proposition is the ideal case, and it explains the practice: the closer μ_k is to an eigenvalue, the smaller $z'_{n,n-1}$ is observed to become, and the algorithm watches that one entry. When it falls below a tolerance it is set to zero, the matrix splits into a smaller Hessenberg block and a known eigenvalue, and the iteration restarts on the block. This is **deflation**, and with it the algorithm is observed to finish after $O(n)$ steps in total rather than converging forever.

Two claims in that paragraph are descriptions of observed behavior, not theorems, and this book proves neither: that a nearer shift always gives a smaller subdiagonal entry, and that the shifted algorithm terminates after $O(n)$ steps. Nothing in this section or later depends on them. What is proved is the exact case, [Proposition 24.5.5](#), and the example below exhibits it.

Example. The shift repairs the rotation, and deflates a symmetric example

- (a) Apply one shifted QR step with $\mu = i$ to $\mathbf{B} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.
- (b) Apply one shifted QR step with $\mu = 2$ to the symmetric tridiagonal matrix $\mathbf{A}$ with diagonal $(2, 2, 2)$ and off-diagonal entries 1.

Solution.

- (a) Here $\mu = i$ is an eigenvalue, so [Proposition 24.5.5](#) predicts an exact deflation. Indeed

$$\mathbf{B} - i\mathbf{I} = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix},$$

whose second column is $-i$ times the first, so the matrix has rank 1. Its first column has norm $\sqrt{2}$, so take $\mathbf{q}_1 = \frac{1}{\sqrt{2}}(-i, 1)$ and $\mathbf{q}_2 = \frac{1}{\sqrt{2}}(1, -i)$, which is a unit vector orthogonal to $\mathbf{q}_1$ since $\overline{(-i)} \cdot 1 + \overline{1} \cdot (-i) = i - i = 0$. With $\mathbf{Q} = (\mathbf{q}_1 \mid \mathbf{q}_2)$ one computes $\mathbf{R} = \mathbf{Q}^*(\mathbf{B} - i\mathbf{I})$, which is

$$\mathbf{R} = \begin{pmatrix} \sqrt{2} & -i\sqrt{2} \\ 0 & 0 \end{pmatrix},$$

with zero last row as predicted. Then $\mathbf{R}\mathbf{Q} = \text{diag}(-2i, 0)$, and

$$\mathbf{B}' = \mathbf{R}\mathbf{Q} + i\mathbf{I} = \text{diag}(-i, i).$$

One step, and the eigenvalues are on the diagonal. The unshifted iteration would have run forever.

(b) The matrix is

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad \mathbf{A} - 2\mathbf{I} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

and 2 is an eigenvalue of $\mathbf{A}$, since $(\mathbf{A} - 2\mathbf{I})(1, 0, -1) = \mathbf{0}$. Two Givens rotations triangularize $\mathbf{A} - 2\mathbf{I}$, as in the proof of [Proposition 24.5.4](#). The first uses the entries 0 and 1 in positions (1, 1) and (2, 1), so $r = 1$, $c_1 = 0$, $s_1 = 1$, and $\mathbf{G}_1$ exchanges rows 1 and 2 with a sign, producing rows $(1, 0, 1)$, $(0, -1, 0)$, $(0, 1, 0)$. The second uses the entries -1 and 1 now in positions (2, 2) and (3, 2), so $r = \sqrt{2}$, $c_2 = -1/\sqrt{2}$, $s_2 = 1/\sqrt{2}$, and it turns those two rows into $(0, \sqrt{2}, 0)$ and $\mathbf{0}$. Hence

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{Q} = \mathbf{G}_1^\top \mathbf{G}_2^\top = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix},$$

with the last row of $\mathbf{R}$ zero, as [Proposition 24.5.5](#) predicts. One checks $\mathbf{QR} = \mathbf{A} - 2\mathbf{I}$ column by column. Multiplying back and restoring the shift,

$$\mathbf{A}' = \mathbf{RQ} + 2\mathbf{I} = \begin{pmatrix} 2 & \sqrt{2} & 0 \\ \sqrt{2} & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

The last row and column have split off, carrying the eigenvalue 2, and the remaining 2×2 block has eigenvalues $2 \pm \sqrt{2}$ — the other two eigenvalues of $\mathbf{A}$. One shifted step, and the problem is finished.

24.5.6 Three unshifted steps, watched

Example. The subdiagonal shrinking at the predicted rate

Run the unshifted QR iteration on the symmetric tridiagonal matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix},$$

whose eigenvalues are $2 + \sqrt{2}$, 2, and $2 - \sqrt{2}$. Carry out the first step exactly, and tabulate three steps.

Solution. Step one, by hand. Apply Theorem 11.2.2 to the columns $\mathbf{a}_1 = (2, 1, 0)$, $\mathbf{a}_2 = (1, 2, 1)$, $\mathbf{a}_3 = (0, 1, 2)$.

From $\mathbf{a}_1$: $r_{11} = \|\mathbf{a}_1\| = \sqrt{5}$ and $\mathbf{q}_1 = \frac{1}{\sqrt{5}}(2, 1, 0)$.

From $\mathbf{a}_2$: $r_{12} = \langle \mathbf{a}_2, \mathbf{q}_1 \rangle = 4/\sqrt{5}$, and

$$\mathbf{w}_2 = \mathbf{a}_2 - \frac{4}{5}(2, 1, 0) = \frac{1}{5}(-3, 6, 5), \quad r_{22} = \|\mathbf{w}_2\| = \sqrt{14/5},$$

since $9 + 36 + 25 = 70$ and $\sqrt{70}/5 = \sqrt{14/5}$; so $\mathbf{q}_2 = \frac{1}{\sqrt{70}}(-3, 6, 5)$.

From $\mathbf{a}_3$: $r_{13} = 1/\sqrt{5}$ and $r_{23} = 16/\sqrt{70}$, and

$$\mathbf{w}_3 = \mathbf{a}_3 - \frac{1}{5}(2, 1, 0) - \frac{16}{70}(-3, 6, 5) = \frac{2}{7}(1, -2, 3), \quad r_{33} = \frac{2\sqrt{14}}{7},$$

so $\mathbf{q}_3 = \frac{1}{\sqrt{14}}(1, -2, 3)$. Now $\mathbf{A}^{(2)} = \mathbf{RQ}$. Its $(1, 1)$ entry is the first row of $\mathbf{R}$ against the first column of $\mathbf{Q}$, namely

$$\sqrt{5} \cdot \frac{2}{\sqrt{5}} + \frac{4}{\sqrt{5}} \cdot \frac{1}{\sqrt{5}} + \frac{1}{\sqrt{5}} \cdot 0 = 2 + \frac{4}{5} = \frac{14}{5},$$

and its $(2, 1)$ entry is $\sqrt{14/5} \cdot \frac{1}{\sqrt{5}} + \frac{16}{\sqrt{70}} \cdot 0 = \sqrt{14}/5$. Continuing,

$$\mathbf{A}^{(2)} = \begin{pmatrix} \frac{14}{5} & \frac{\sqrt{14}}{5} & 0 \\ \frac{\sqrt{14}}{5} & \frac{82}{35} & \frac{2\sqrt{5}}{7} \\ 0 & \frac{2\sqrt{5}}{7} & \frac{6}{7} \end{pmatrix}.$$

Three checks: the matrix is symmetric and tridiagonal, as the Quick check after [Proposition 24.5.4](#) predicted; its trace is $\frac{14}{5} + \frac{82}{35} + \frac{6}{7} = \frac{98+82+30}{35} = 6 = \text{tr } \mathbf{A}$; and the entry $(3, 1)$ is zero, not merely small.

Three steps, to eight places. The diagonal entries and the two subdiagonal entries run as follows.

k	$a_{11}^{(k)}$	$a_{22}^{(k)}$	$a_{33}^{(k)}$	$ a_{21}^{(k)} $	$ a_{32}^{(k)} $
1	2.00000000	2.00000000	2.00000000	1.000000	1.000000
2	2.80000000	2.34285714	0.85714286	0.748331	0.638877
3	3.14285714	2.24844720	0.60869565	0.559397	0.187848
4	3.30841121	2.10394692	0.58764187	0.372193	0.052177

The three diagonal entries are heading for $2 + \sqrt{2} = 3.41421356$, 2, and $2 - \sqrt{2} = 0.58578644$, largest modulus first — the order the convergence theorem quoted above predicts, exhibited here rather than proved. The two subdiagonal entries shrink by factors

$$0.748, 0.748, 0.665, \dots \quad \text{and} \quad 0.639, 0.294, 0.278, \dots$$

The predicted asymptotic factors are $|\lambda_2/\lambda_1| = 2/(2 + \sqrt{2}) = 0.585786$ and $|\lambda_3/\lambda_2| = (2 - \sqrt{2})/2 = 0.292893$. The second entry has already settled onto its rate; the first is still approaching it, and by step 10 its factor is 0.5859. Convergence is linear, and slow: each step buys about a quarter of a decimal digit in the $(2, 1)$ entry. That is the cost of not shifting.

24.5.7 Exercises

24.5.7.1 A. Check your understanding

Exercise. A1

- (a) Write down one step of the unshifted QR iteration, and one step of the shifted one, and prove in one line each that both are similarities.

- (b) State [Theorem 24.5.2](#), and say in one sentence what it has to do with the power method.
- (c) True or false: the QR iterates of a matrix always converge to an upper triangular matrix. Justify your answer.
- (d) Why is the iteration run on an upper Hessenberg matrix rather than on a full one? Give the two costs.
- (e) In [Proposition 24.5.5](#), which hypothesis on $\mathbf{Z}$ is doing the work, and where exactly is it used?

Solution.

- (a) Unshifted: $\mathbf{A}^{(k)} = \mathbf{QR}$, $\mathbf{A}^{(k+1)} = \mathbf{RQ}$; then $\mathbf{RQ} = \mathbf{Q}^* \mathbf{QRQ} = \mathbf{Q}^* \mathbf{A}^{(k)} \mathbf{Q}$. Shifted: $\mathbf{A}^{(k)} - \mu \mathbf{I} = \mathbf{QR}$, $\mathbf{A}^{(k+1)} = \mathbf{RQ} + \mu \mathbf{I}$; then $\mathbf{RQ} + \mu \mathbf{I} = \mathbf{Q}^* (\mathbf{A}^{(k)} - \mu \mathbf{I}) \mathbf{Q} + \mu \mathbf{Q}^* \mathbf{Q} = \mathbf{Q}^* \mathbf{A}^{(k)} \mathbf{Q}$.
- (b) It says $\mathbf{A}^k = \mathbf{U}_k \mathbf{T}_k$ with $\mathbf{U}_k = \mathbf{Q}^{(1)} \cdots \mathbf{Q}^{(k)}$ unitary and $\mathbf{T}_k = \mathbf{R}^{(k)} \cdots \mathbf{R}^{(1)}$ upper triangular, and that $\mathbf{A}^{(k+1)} = \mathbf{U}_k^* \mathbf{A} \mathbf{U}_k$. Comparing first columns, the first column of $\mathbf{U}_k$ is $\mathbf{A}^k \mathbf{e}_1$ normalized, which is the k -th power-method iterate started at $\mathbf{e}_1$.
- (c) False. The rotation $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is a fixed point of the unshifted iteration ([Example \(Two matrices the iteration cannot move\)](#)) and is not triangular; its two eigenvalues have the same modulus, which is exactly the hypothesis [Theorem 24.5.3](#) needs.
- (d) Because a QR step on a full matrix costs $O(n^3)$ flops, with the count genuinely growing like n^3 , while a step on a Hessenberg matrix costs about $6n^2$ ([Proposition 24.5.4](#)), and the pattern is preserved, so the saving is repeated at every step. The reduction itself is $O(n^3)$, paid once ([Theorem 11.8.5](#)).
- (e) That every subdiagonal entry of $\mathbf{Z}$ is nonzero. It is used exactly once, to show that the first $n - 1$ columns of $\mathbf{Z} - \mu \mathbf{I}$ are independent: deleting the first row leaves a lower triangular matrix whose diagonal is that subdiagonal. Without it $\mathbf{Z} - \mu \mathbf{I}$ could drop rank in an earlier column, and the zero row of $\mathbf{R}$ need not be the last.

24.5.7.2 B. Practice

Exercise. B1: One step by hand

Let $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Compute the QR factorization of $\mathbf{A}$ in the sense of [Theorem 11.8.1](#), then $\mathbf{A}^{(2)} = \mathbf{RQ}$. Verify that $\mathbf{A}^{(2)}$ is symmetric, has trace 4 and determinant 3, and that its off-diagonal entry is smaller than that of $\mathbf{A}$. Hence estimate how many steps are needed to make the off-diagonal entry smaller than 10^{-3} .

Solution. The columns are $\mathbf{a}_1 = (2, 1)$ and $\mathbf{a}_2 = (1, 2)$. Then $r_{11} = \sqrt{5}$, $\mathbf{q}_1 = \frac{1}{\sqrt{5}}(2, 1)$, $r_{12} = \langle \mathbf{a}_2, \mathbf{q}_1 \rangle = 4/\sqrt{5}$, and

$$\mathbf{w}_2 = (1, 2) - \frac{4}{5}(2, 1) = \frac{3}{5}(-1, 2), \quad r_{22} = \frac{3}{\sqrt{5}}, \quad \mathbf{q}_2 = \frac{1}{\sqrt{5}}(-1, 2).$$

So

$$\mathbf{Q} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{R} = \frac{1}{\sqrt{5}} \begin{pmatrix} 5 & 4 \\ 0 & 3 \end{pmatrix},$$

and

$$\mathbf{A}^{(2)} = \mathbf{RQ} = \frac{1}{5} \begin{pmatrix} 5 & 4 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 14 & 3 \\ 3 & 6 \end{pmatrix}.$$

It is symmetric; its trace is $20/5 = 4 = \text{tr } \mathbf{A}$; its determinant is $(84 - 9)/25 = 3 = \det \mathbf{A}$. The off-diagonal entry fell from 1 to $3/5$.

The eigenvalues are 3 and 1, so [Theorem 24.5.3](#) gives the asymptotic factor $t = 1/3$ per step. To take $3/5$ below 10^{-3} needs $(1/3)^m \leq 10^{-3}/(3/5)$, that is $m \geq \log(600)/\log 3 \approx 5.8$: about six further steps, so seven in all. (The first steps are a little slower than the asymptotic rate, so eight is a safer count.)

Exercise. B2: Fixed points

Determine which of the following matrices are fixed points of the unshifted QR iteration, that is, satisfy $\mathbf{A}^{(2)} = \mathbf{A}^{(1)}$. Justify your answer; for those that are not, name what changes.

- (a) any upper triangular $\mathbf{A}$ with positive real diagonal entries;
- (b) any unitary $\mathbf{A}$;
- (c) $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$;
- (d) any normal $\mathbf{A}$.

Solution.

- (a) Yes. Then $\mathbf{A} = \mathbf{IA}$ is a factorization with $\mathbf{I}$ unitary and $\mathbf{A}$ upper triangular with positive diagonal, hence the unique one (Theorem 11.8.1), and $\mathbf{A}^{(2)} = \mathbf{AI} = \mathbf{A}$.
- (b) Yes. Then $\mathbf{A} = \mathbf{AI}$ is a factorization with $\mathbf{A}$ unitary and $\mathbf{I}$ upper triangular with positive diagonal, hence the unique one, and $\mathbf{A}^{(2)} = \mathbf{IA} = \mathbf{A}$.
- (c) Yes, by (b): the matrix is a permutation matrix, so its columns are orthonormal. This is worth noticing because its eigenvalues are 1 and -1 , of equal modulus — another instance of the obstruction of [Example \(The unshifted algorithm on a rotation\)](#), and this one is symmetric, so unlike the rotation it *is* orthogonally similar to a diagonal matrix. The unshifted iteration simply will not find it.
- (d) No. A normal matrix need not be fixed: $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ is symmetric, hence normal, and

[Exercise \(B1: One step by hand\)](#) computed $\mathbf{A}^{(2)} = \frac{1}{5} \begin{pmatrix} 14 & 3 \\ 3 & 6 \end{pmatrix} \neq \mathbf{A}$. What is preserved is normality, not the matrix: $\mathbf{A}^{(2)} = \mathbf{Q}^* \mathbf{AQ}$ is unitarily similar to $\mathbf{A}$, and normality is invariant under unitary similarity.

Exercise. B3: A shift that deflates

Let $\mathbf{Z} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Take μ to be either eigenvalue of $\mathbf{Z}$ and carry out one shifted QR step. Verify the conclusion of [Proposition 24.5.5](#). *Hint: you need only the last row of $\mathbf{R}$.*

Solution. The characteristic polynomial is $x^2 - 5x - 2$, with roots $\mu_{\pm} = (5 \pm \sqrt{33})/2$. Every 2×2 matrix is upper Hessenberg, and the single subdiagonal entry 3 is nonzero, so [Proposition 24.5.5](#) applies with $n = 2$. Take $\mu = \mu_+$. Then

$$\mathbf{Z} - \mu\mathbf{I} = \begin{pmatrix} 1 - \mu & 2 \\ 3 & 4 - \mu \end{pmatrix}$$

is singular, of rank 1 since it is nonzero. Its first column has norm $r = \sqrt{(1 - \mu)^2 + 9} \neq 0$, so a QR factorization has $\mathbf{q}_1 = \frac{1}{r}(1 - \mu, 3)$ and $r_{11} = r \neq 0$. Because the rank is 1, the second column of $\mathbf{R}$ is a multiple of $\mathbf{e}_1$, so $r_{22} = 0$ and the last row of $\mathbf{R}$ vanishes. Then the last row of $\mathbf{RQ}$ vanishes too, and $\mathbf{Z}' = \mathbf{RQ} + \mu\mathbf{I}$ has last row $(0, \mu)$. Its $(1, 1)$ entry is then $\text{tr } \mathbf{Z} - \mu = 5 - \mu = \mu_-$, since the trace is preserved. One step produced both eigenvalues exactly.

24.5.7.3 C. Going deeper

Exercise. C1: Hessenberg form is not enough by itself

- Prove that if $\mathbf{Z}$ is upper Hessenberg with $z_{m+1,m} = 0$ for some $1 \leq m \leq n - 1$, then $\text{spec}(\mathbf{Z}) = \text{spec}(\mathbf{Z}_{11}) \cup \text{spec}(\mathbf{Z}_{22})$, where $\mathbf{Z}_{11}$ is the leading $m \times m$ block and $\mathbf{Z}_{22}$ the trailing $(n - m) \times (n - m)$ block.
- Deduce that a QR implementation may, at every step, replace the problem by two smaller ones as soon as one subdiagonal entry vanishes.
- Give a 3×3 upper Hessenberg matrix with no zero subdiagonal entry for which the unshifted iteration never reduces any subdiagonal entry at all.

Solution.

- With $z_{m+1,m} = 0$ and $\mathbf{Z}$ upper Hessenberg, every entry z_{ij} with $i > m \geq j$ vanishes: for $i > m + 1$ this is $i > j + 1$, and for $i = m + 1$ it is either $j < m$, again $i > j + 1$, or $j = m$, the assumed entry. So $\mathbf{Z}$ is block upper triangular with diagonal blocks $\mathbf{Z}_{11}$ and $\mathbf{Z}_{22}$. By Theorem 8.1.1 and the determinant of a block triangular matrix (Theorem 7.8.4), $p_{\mathbf{Z}}(x) = p_{\mathbf{Z}_{11}}(x)p_{\mathbf{Z}_{22}}(x)$, so the eigenvalues of $\mathbf{Z}$ are those of the two blocks together.
- The two blocks are again upper Hessenberg, and by (a) computing their spectra computes that of $\mathbf{Z}$. Each step on a block of size m costs about $6m^2$ rather than $6n^2$ ([Proposition 24.5.4](#)), so splitting is a strict gain; and part (a) guarantees no eigenvalue is lost. This is deflation.
- Take the cyclic permutation matrix

$$\mathbf{Z} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

which is upper Hessenberg with subdiagonal entries 1, 1, and is orthogonal. By [Exercise \(B2: Fixed points\)](#) (b) it is a fixed point of the unshifted iteration, so no subdiagonal entry ever moves. Its eigenvalues are the three cube roots of 1, all of modulus 1 — the obstruction of [Theorem 24.5.3](#) in its purest form.

Exercise. C2: Hermitian iterates and the Rayleigh quotient

Let $\mathbf{A} \in M_n(\mathbb{C})$ be Hermitian and let $\mathbf{A}^{(k)}$ be its unshifted QR iterates.

- (a) Prove that every $\mathbf{A}^{(k)}$ is Hermitian with the same eigenvalues as $\mathbf{A}$.
- (b) Prove that $a_{nn}^{(k)} = R_{\mathbf{A}}(\mathbf{u}_n^{(k)})$, where $\mathbf{u}_n^{(k)}$ is the last column of $\mathbf{U}_{k-1}$ and $R_{\mathbf{A}}$ is the Rayleigh quotient of Definition 17.1.1.
- (c) Hence explain, citing Theorem 20.6.1, why the Rayleigh shift $\mu_k = a_{nn}^{(k)}$ is a certified eigenvalue estimate, and say what the certificate is.

Solution.

- (a) By Theorem 24.5.2, $\mathbf{A}^{(k)} = \mathbf{U}_{k-1}^* \mathbf{A} \mathbf{U}_{k-1}$, with $\mathbf{U}_0 = \mathbf{I}$ at $k = 1$, so $(\mathbf{A}^{(k)})^* = \mathbf{U}_{k-1}^* \mathbf{A}^* \mathbf{U}_{k-1} = \mathbf{A}^{(k)}$, and unitary similarity preserves eigenvalues.
- (b) Writing $\mathbf{u}_n = \mathbf{U}_{k-1} \mathbf{e}_n$, which is a unit vector since $\mathbf{U}_{k-1}$ is unitary,

$$a_{nn}^{(k)} = \mathbf{e}_n^* \mathbf{A}^{(k)} \mathbf{e}_n = \mathbf{e}_n^* \mathbf{U}_{k-1}^* \mathbf{A} \mathbf{U}_{k-1} \mathbf{e}_n = \mathbf{u}_n^* \mathbf{A} \mathbf{u}_n = R_{\mathbf{A}}(\mathbf{u}_n),$$

the last equality because $R_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^* \mathbf{A} \mathbf{x} / \mathbf{x}^* \mathbf{x}$ and $\mathbf{u}_n^* \mathbf{u}_n = 1$.

- (c) Put $\mu = a_{nn}^{(k)}$ and $\mathbf{r} = \mathbf{A} \mathbf{u}_n - \mu \mathbf{u}_n$. Then μ is real, $\mathbf{A}$ being Hermitian, and Theorem 20.6.1 says that $\mathbf{A}$ has an eigenvalue in $[\mu - \|\mathbf{r}\|_2, \mu + \|\mathbf{r}\|_2]$. The certificate is $\|\mathbf{r}\|_2$, and it costs nothing extra: by the computation in (b), $\mathbf{U}_{k-1}^* \mathbf{r} = \mathbf{A}^{(k)} \mathbf{e}_n - \mu \mathbf{e}_n$ is the last column of $\mathbf{A}^{(k)}$ with its diagonal entry removed, so $\|\mathbf{r}\|_2$ is the Euclidean norm of the entries of that column other than the diagonal one. If $\mathbf{A}$ was first reduced to Hessenberg form, so that every $\mathbf{A}^{(k)}$ is Hermitian and upper Hessenberg, hence tridiagonal, that norm is the single entry $|a_{n-1,n}^{(k)}|$. The algorithm is already watching that number.

Exercise. C3: Why the powers are the right object

Let $\mathbf{A} \in M_n(\mathbb{C})$ be invertible, with unshifted QR iterates as in Definition 24.5.1.

- (a) Prove that for $1 \leq j \leq n$, the subspace spanned by the first j columns of $\mathbf{U}_k$ equals $\mathbf{A}^k(\text{span}(\mathbf{e}_1, \dots, \mathbf{e}_j))$.
- (b) Deduce that if the first j columns of $\mathbf{U}_k$ converge to a basis of an $\mathbf{A}$ -invariant subspace W , then the entries of $\mathbf{A}^{(k+1)}$ in rows $j+1, \dots, n$ of columns $1, \dots, j$ tend to 0.

Hint for (b): write the entries as inner products of $\mathbf{A}$ applied to one column against another.

Solution.

- (a) By Theorem 24.5.2, $\mathbf{A}^k = \mathbf{U}_k \mathbf{T}_k$ with $\mathbf{T}_k$ upper triangular and invertible (Quick check above), so $\mathbf{A}^k \mathbf{e}_i = \mathbf{U}_k \mathbf{T}_k \mathbf{e}_i$ lies in $\text{span}(\mathbf{U}_k \mathbf{e}_1, \dots, \mathbf{U}_k \mathbf{e}_i)$ for each i , since $\mathbf{T}_k \mathbf{e}_i$ has zeros below row i . Hence $\mathbf{A}^k(\text{span}(\mathbf{e}_1, \dots, \mathbf{e}_j)) \subseteq \text{span}(\mathbf{U}_k \mathbf{e}_1, \dots, \mathbf{U}_k \mathbf{e}_j)$. Both sides have dimension j : the left because $\mathbf{A}^k$ is invertible, the right because the columns of a unitary matrix are independent. One inclusion plus equal dimension gives equality.

(b) Write $\mathbf{u}_i = \mathbf{U}_k \mathbf{e}_i$. For $p > j \geq q$,

$$a_{pq}^{(k+1)} = \mathbf{e}_p^* \mathbf{U}_k^* \mathbf{A} \mathbf{U}_k \mathbf{e}_q = \mathbf{u}_p^* \mathbf{A} \mathbf{u}_q = \langle \mathbf{A} \mathbf{u}_q, \mathbf{u}_p \rangle.$$

Suppose $\mathbf{u}_1, \dots, \mathbf{u}_j$ converge to an orthonormal basis $\mathbf{w}_1, \dots, \mathbf{w}_j$ of an $\mathbf{A}$ -invariant subspace W , let P be the orthogonal projection onto W (Definition 11.3.2) and put $\delta_k = \max_{q \leq j} \|\mathbf{u}_q - \mathbf{w}_q\|_2$, so $\delta_k \rightarrow 0$. Since $\mathbf{A} \mathbf{w}_q \in W$, we have $(\mathbf{I} - P) \mathbf{A} \mathbf{w}_q = \mathbf{0}$ and hence $\|(\mathbf{I} - P) \mathbf{A} \mathbf{u}_q\|_2 = \|(\mathbf{I} - P) \mathbf{A} (\mathbf{u}_q - \mathbf{w}_q)\|_2 \leq \|\mathbf{A}\|_2 \delta_k$. Since $\langle \mathbf{u}_p, \mathbf{u}_q \rangle = 0$ for $q \leq j < p$, we have $|\langle \mathbf{u}_p, \mathbf{w}_q \rangle| = |\langle \mathbf{u}_p, \mathbf{w}_q - \mathbf{u}_q \rangle| \leq \delta_k$ by Cauchy–Schwarz (Theorem 11.1.5), so $\|P \mathbf{u}_p\|_2 \leq j \delta_k$. Splitting $\mathbf{u}_p$ and using that P is self-adjoint,

$$\begin{aligned} |\langle \mathbf{A} \mathbf{u}_q, \mathbf{u}_p \rangle| &= |\langle (\mathbf{I} - P) \mathbf{A} \mathbf{u}_q, \mathbf{u}_p \rangle + \langle \mathbf{A} \mathbf{u}_q, P \mathbf{u}_p \rangle| \\ &\leq \|(\mathbf{I} - P) \mathbf{A} \mathbf{u}_q\|_2 + \|\mathbf{A}\|_2 \|P \mathbf{u}_p\|_2 \leq (1 + j) \|\mathbf{A}\|_2 \delta_k, \end{aligned}$$

by Cauchy–Schwarz again and $\|\mathbf{u}_p\|_2 = \|\mathbf{u}_q\|_2 = 1$. The right-hand side tends to 0, so $a_{pq}^{(k+1)} \rightarrow 0$. This is [Theorem 24.5.3](#) for $j = 1$, read as a statement about invariant subspaces rather than about eigenvectors, and it is the form in which the general convergence theorem quoted above is proved.