Computing a Tridiagonal Matrix that is Orthogonally Similar

$A \in \mathbb{R}^{n\times n}$, $A^T = A$

Find $Q_1, Q_2, \dots, Q_{n-2} \in \mathbb{R}^{n\times n}$ all orthogonal such that

$$Q_{n-2}^T \cdots Q_2^T Q_1^T A Q_1 Q_2 \cdots Q_{n-2} = \underset{\text{tridiagonal}}{T}$$

Illustration 4×4 case

$$A = \begin{bmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{bmatrix} \longmapsto \begin{bmatrix} x & x & 0 & 0 \\ x & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{bmatrix} = Q_1^T A Q_1$$

$$\longmapsto \begin{bmatrix} x & x & 0 & 0 \\ x & x & x & 0 \\ 0 & x & x & x \\ 0 & 0 & x & x \end{bmatrix} = Q_2^T Q_1^T A Q_1 Q_2$$

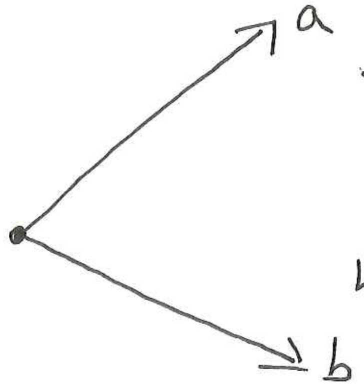
Given $v \in \mathbb{R}^n$, would like to find an orthogonal $Q \in \mathbb{R}^{n\times n}$ such that

$$Qv = \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \exists \alpha \in \mathbb{R}.$$

Employ Householder reflectors for this purpose.

Householder Reflectors

Given $a, b \in \mathbb{R}^n$, find orthogonal Q s.t. $Qa = b$

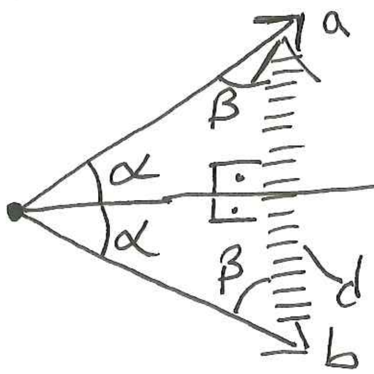


Remark

$$\|Qa\|_2^2 = (Qa)^T(Qa) = a^T Q^T Q a = \|a\|_2^2$$

Hence, $a, b \in \mathbb{R}^n$ must be s.t. $\|a\|_2 = \|b\|_2$.

Let $d := a - b$.



reflect about a hyperplane namely $\text{span}\{d\}^\perp$

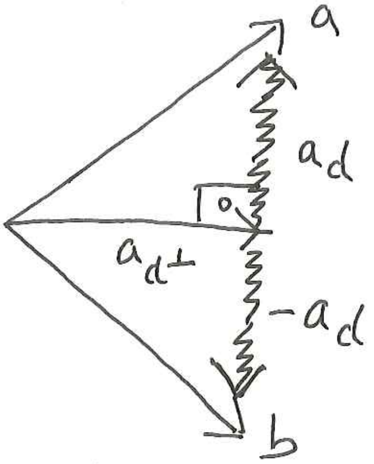
Decompose a

$$a = a_d + a_{d^\perp}$$

where $a_d \in \text{span}\{d\}$,

$a_{d^\perp} \in \text{span}\{d\}^\perp$

(that is a_d is the orthogonal projection of a onto $\text{span}\{d\}$)



$$
\begin{aligned}
b &= -a_d + a_{d^\perp} \\
&= -a_d + (a - a_d) \\
&= a - 2a_d \\
&= a - 2\frac{dd^T}{d^Td}a \\
&= \left(I_n - \frac{2dd^T}{d^Td}\right).a
\end{aligned}
$$

Householder reflector : $H = I_n - 2\frac{dd^T}{d^Td}$

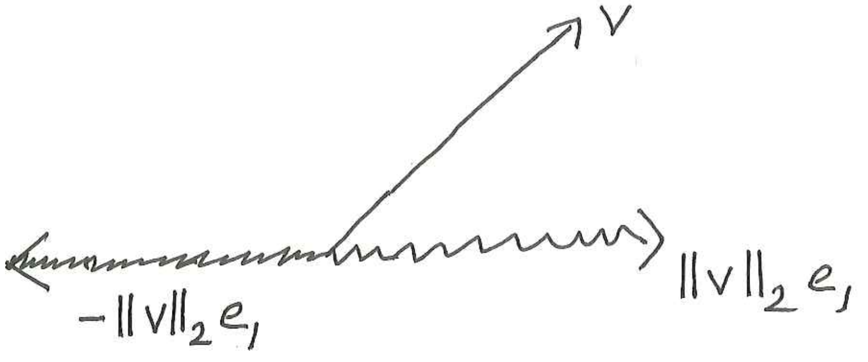
Remark

H is symmetric and orthogonal

$$
\begin{aligned}
H^T &= I_n^T - 2\frac{(dd^T)^T}{d^Td} \\
&= I_n - 2\frac{dd^T}{d^Td} = H_n
\end{aligned}
$$

$$
\begin{aligned}
H_n^T H_n &= \left(I_n - 2\frac{dd^T}{d^Td}\right)\left(I_n - 2\frac{dd^T}{d^Td}\right) \\
&= I_n - 4\frac{dd^T}{d^Td} + \underbrace{\left(2\frac{dd^T}{d^Td}\right)\left(2\frac{dd^T}{d^Td}\right)}_{4\frac{dd^T}{d^Td}} \\
&= 0
\end{aligned}
$$

Would like to make $v \in \mathbb{R}^n$, a multiple of e_1, indeed one of $\pm \|v\|_2 e_1$



Choose the one so that $\|d\|_2$ is larger

$$d = \begin{cases} v + \|v\|_2 e_1 & \text{if } v_1 \geqslant 0 \\ v - \|v\|_2 e_1 & \text{if } v_1 < 0 \end{cases}$$

$$= v + \text{sign}(v_1) \|v\|_2 e_1$$

where

$$\text{sign}(v_1) = \begin{cases} 1 & \text{if } v_1 \geqslant 0 \\ -1 & \text{if } v_1 < 0 \end{cases}$$

This is to reduce the effect of rounding errors.

Ex

Calculate a QR factorization for

$$A = \begin{bmatrix} 4 & 1 \\ 3 & 5 \end{bmatrix}$$

using Householder reflectors.

$$\begin{bmatrix} 4 \\ 3 \end{bmatrix} \xmapsto[H]{} \begin{bmatrix} -5 \\ 0 \end{bmatrix} \qquad d = \begin{bmatrix} 9 \\ 3 \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{2}{\underbrace{d^T d}_{90}} \begin{bmatrix} 9 \\ 3 \end{bmatrix} \begin{bmatrix} 9 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 9/5 & 3/5 \\ 3/5 & 1/5 \end{bmatrix} = \begin{bmatrix} -4/5 & -3/5 \\ -3/5 & 4/5 \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} -4/5 & -3/5 \\ -3/5 & 4/5 \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} 4 & 1 \\ 3 & 5 \end{bmatrix}}_{A} = \begin{bmatrix} -5 & -19/5 \\ 0 & 17/5 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 4 & 1 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} -4/5 & -3/5 \\ -3/5 & 4/5 \end{bmatrix} \begin{bmatrix} -5 & -19/5 \\ 0 & 17/5 \end{bmatrix} //$$

The algorithm

| L – left multiply by Q_1
| R – right multiply by Q_1

1st step

$$A = \begin{bmatrix} x & x & x & \cdots & x \\ \boxed{x} & x & & & \\ \boxed{x} & x & & & \\ \vdots & & & & \\ \boxed{x} & x & & & x \end{bmatrix} \xmapsto{L} \underbrace{\begin{bmatrix} x & x & x & \cdots & x \\ x & x & & & \\ 0 & x & & & \\ \vdots & & & & \\ 0 & x & & & x \end{bmatrix}}_{Q_1^T A = Q_1 A}$$

$v^{(1)}$

$$Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & H_1 \end{bmatrix} \left(\Leftarrow H_1 \text{ is the reflector turning } v^{(1)} \text{ into } -\mathrm{sign}(v_1^{(1)}) \|v^{(1)}\|_2 e_1 \right.$$

$$H_1 = I_{n-1} - 2 \frac{d^{(1)} d^{(1)T}}{d^{(1)T} d^{(1)}}, \quad d^{(1)} = v^{(1)} + \mathrm{sign}(v_1^{(1)}) \|v^{(1)}\|_2 e_1$$

$$Q_1^T A = \begin{array}{c} \quad\; v^{(1)T} \\ \begin{bmatrix} x & \boxed{x \;\; x \;\; \cdots \;\; x} \\ x & x \;\; x \;\; \cdots \;\; x \\ 0 & x \qquad\qquad \\ \vdots & \\ 0 & x \qquad\quad\; x \end{bmatrix} \end{array} \xmapsto{R} \begin{bmatrix} x & x & 0 & \cdots & 0 \\ x & x & x & \cdots & x \\ 0 & x & & & \\ \vdots & & & & \\ 0 & x & & & x \end{bmatrix} = Q_1^T A Q_1 \quad (=: A^{(2)})$$

$$\left(H_1 v^{(1)} = -\mathrm{sign}(v_1^{(1)}) \|v^{(1)}\|_2 e_1 \Rightarrow v^{(1)T} H_1 = -\mathrm{sign}(v_1^{(1)}) \|v^{(1)}\|_2 e_1^T \right)$$

kth step $(k = 1, \dots, n-2)$

$$A^{(k)} = Q_{k-1}^T \cdots Q_1^T A Q_1 \cdots Q_{k-1}$$

$$= \begin{bmatrix} \times & \times & \cdots & & & & 0 \\ \times & \ddots & \times & \times & 0 & \cdots & 0 \\ & \ddots & \times & \times & \times & \cdots & \times \\ 0 & 0 & 0 & \times & \times & \cdots & \times \\ & \vdots & \vdots & \vdots & \vdots & & \vdots \\ & 0 & 0 & \times & \times & \cdots & \times \end{bmatrix}$$

$v^{(k)T}$

(k,k) entry

$v^{(k)}$

$$Q_k = \begin{bmatrix} I_k & 0 \\ 0 & H_k \end{bmatrix}$$

$$H_k = I_{n-k} - \frac{2 d^{(k)} d^{(k)T}}{d^{(k)T} d^{(k)}}, \quad d^{(k)} = v^{(k)} + \mathrm{sign}(v_1^{(k)}) \|v^{(k)}\|_2 e_1$$

$$A^{(k)} \xmapsto{L} Q_k^T A^{(k)} = Q_k A^{(k)} \quad \left(\begin{array}{l} A^{(k)}(k+2:n, k) \\ \text{becomes } 0. \end{array} \right)$$

$$\xmapsto{R} Q_k^T A^{(k)} Q_k =: A^{(k+1)} \quad \left(\begin{array}{l} A^{(k)}(k, k+2:n) \\ \text{becomes } 0. \end{array} \right)$$

Termination

$Q_{n-2}^T \cdots Q_1^T A Q_1 \cdots Q_{n-2}$ is tridiagonal.

The QR Algorithm

$A \in \mathbb{R}^{n\times n}, \quad A^T = A$

Stage ①

Reduction into tridiagonal form

$$Q_{n-2}^T \cdots Q_1^T A Q_1 \cdots Q_{n-2} = T$$

tridiagonal

$Q_1, \ldots, Q_{n-2} \in \mathbb{R}^{n\times n}$ - orthogonal

(just discussed in Ch 5 - P3)

Stage ②

$$\hat{Q}_k^T \cdots \hat{Q}_1^T T \hat{Q}_1 \cdots \hat{Q}_k = T^{(k)}$$

Generate a sequence $\{T^{(k)}\}$ s.t.
$T^{(k)}$ becomes diagonal as $k \to \infty$. (under mild assumptions)

Applicable to nonsymmetric A as well with following changes.

Stage ① - T is Hessenberg (rather than tridiagonal)

e.g. $\begin{bmatrix} x & x & x & x \\ x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \end{bmatrix}$

i.e. $t_{ij} = 0$ whenever $i > j+1$.

Stage ② - $T^{(k)}$ becomes upper triangular as $k \to \infty$

Stage 2

$T^{(k)}$ is such that

* $T^{(0)} = T$

* (++) $T^{(k+1)} = \hat{R}^{(k+1)} \hat{Q}^{(k+1)} + \mu_{k+1} I_n$

where $\hat{R}^{(k+1)}, \hat{Q}^{(k+1)}$ are s.t.

$$(+) \left(T^{(k)} - \mu_{k+1} I_n\right) = \hat{Q}^{(k+1)} \hat{R}^{(k+1)}$$

is a QR factorization.

($\mu_{k+1} \in \mathbb{R}$ above is a given shift)

Claim:

$T^{(k)}$ and $T^{(k+1)}$ are orthogonally similar.

Proof

From (+)

$$\hat{R}^{(k+1)} = \left[\hat{Q}^{(k+1)}\right]^T \left(T^{(k)} - \mu_{k+1} I_n\right)$$

Substitute this in (++)

$$T^{(k+1)} = \left[\hat{Q}^{(k+1)}\right]^T \left(T^{(k)} - \mu_{k+1} I_n\right) \hat{Q}^{(k+1)} + \mu_{k+1} I_n$$

$$= \left[\hat{Q}^{(k+1)}\right]^T T^{(k)} \hat{Q}^{(k+1)}$$

QR Factorization for a Tridiagonal Matrix by Rotators

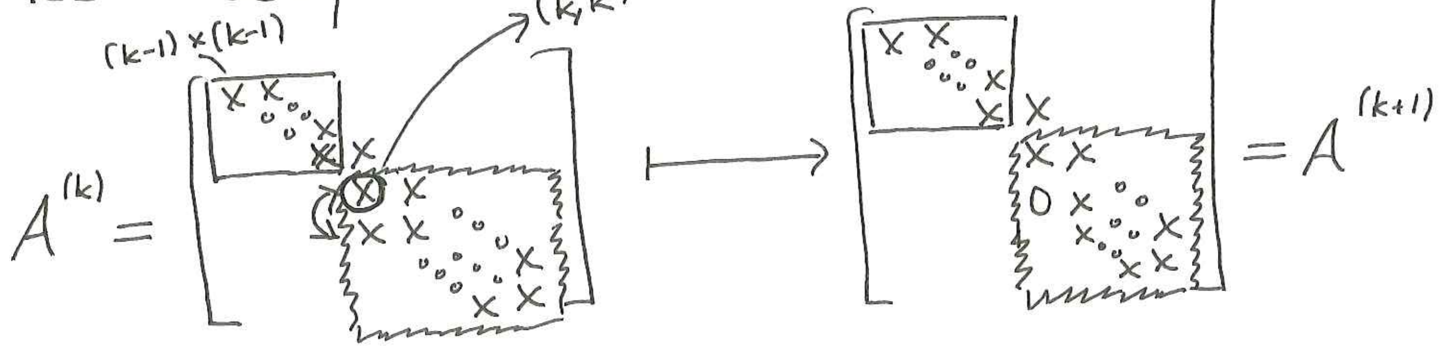
$$\underbrace{\begin{bmatrix} x & x & & \\ x & x & x & \\ & x & x & x \\ & & x & x \end{bmatrix}}_{T} \xmapsto{R^{(12)}(\theta_1)} \begin{bmatrix} x & x & & \\ 0 & x & x & \\ & x & x & x \\ & & x & x \end{bmatrix} \xmapsto{R^{(23)}(\theta_2)} \begin{bmatrix} x & x & & \\ 0 & x & x & \\ & 0 & x & x \\ & & x & x \end{bmatrix}$$

$$\xmapsto{R^{(34)}(\theta_3)} \underbrace{\begin{bmatrix} x & x & & \\ & x & x & \\ & & x & x \\ & & & x \end{bmatrix}}_{R}$$

$$R^{(34)}(\theta_3)R^{(23)}(\theta_2)R^{(12)}(\theta_1)T = R$$

$$\text{(QR factorization)}\quad T = \underbrace{\left\{R^{(12)}(\theta_1)^T R^{(23)}(\theta_2)^T R^{(34)}(\theta_3)^T\right\}}_{Q} R$$

kth step



$$R^{(k,k+1)}(\theta_k)\, A^{(k)} = A^{(k+1)}$$

$$\begin{array}{r} \\ \text{kth row} \leftarrow \\ \text{(k+1)st row} \leftarrow \\ \\ \end{array} \begin{bmatrix} I & & & \\ & c & s & \\ & -s & c & \\ & & & I \end{bmatrix}$$

$\downarrow$ kth col, $\downarrow$ (k+1)st col

Would like

$$0 = a^{(k+1)}_{k+1,k} = -s\, a^{(k)}_{kk} + c\, a^{(k)}_{k+1,k}$$

Choose

$$s = \frac{a^{(k)}_{k+1,k}}{\rho}, \quad c = \frac{a^{(k)}_{kk}}{\rho} \quad \text{where} \quad \rho := \sqrt{\left[a^{(k)}_{kk}\right]^2 + \left[a^{(k)}_{k+1,k}\right]^2}$$

Observe also

$$a^{(k+1)}_{kk} = c\, a^{(k)}_{kk} + s\, a^{(k)}_{(k+1),k}$$

$$a^{(k+1)}_{k,k+1} = c\, a^{(k)}_{k,k+1} + s\, a^{(k)}_{(k+1),k+1}$$

$$a^{(k+1)}_{k+1,k+1} = -s\, a^{(k)}_{k,k+1} + c\, a^{(k)}_{k+1,k+1}$$

Pseudocode

for $k = 1, \ldots, n-1$

  $\rho \leftarrow \sqrt{a_{kk}^2 + a_{k+1,k}^2}$

  $s \leftarrow a_{k+1,k} / \rho \qquad c \leftarrow a_{kk} / \rho$

  $a_{kk} \leftarrow c\, a_{kk} + s\, a_{k+1,k}$

  $\tilde{a}_{k,k+1} \leftarrow c\, a_{k,k+1} + s\, a_{k+1,k+1}$

  $a_{k+1,k+1} \leftarrow -s\, a_{k,k+1} + c\, a_{k+1,k+1}$

  $a_{k+1,k} \leftarrow 0 \qquad a_{k,k+1} \leftarrow \tilde{a}_{k,k+1}$

end

arithmetic operations $= 14(n-1) \sim 14n$

Note : R factor in the end is bidiagonal, i.e., only r_{ii} and $r_{i,i+1}$ are nonzero (possibly)

Claim

Given $T^{(k)}$ in Stage 2 of QR algorithm is tridiagonal, $T^{(k+1)}$ is also tridiagonal.

Proof outline

In the QR factorization

$$T^{(k)} - \mu_{k+1} I_n = \hat{Q}^{(k+1)} \hat{R}^{(k+1)},$$

$\hat{R}^{(k+1)}$ is bidiagonal.

It turns out $[\hat{R}^{(k+1)}]^{-1}$ is also bidiagonal, turns out $\hat{Q}^{(k+1)} = (T^{(k)} - \mu_{k+1} I_n) [\hat{R}^{(k+1)}]^{-1}$ is tridiagonal. Finally, $\hat{R}^{(k+1)} \hat{Q}^{(k+1)} + \mu_{k+1} I_n$ is tridiagonal. $\blacksquare$

Choice of shifts (widely used ones)

$\mu_{k+1} = t^{(k)}_{n,n-1}$ Rayleigh shift

$\mu_{k+1} =$ eigenvalue of $T^{(k)}(n-1:n, n-1:n)$ closest to $t^{(k)}_{nn}$ Wilkinson shift

Convergence (when $\mu_k = 0 \ \forall k$)

POWER ITERATION

$q^{(0)}$ – unit vector

$$q^{(k+1)} = \frac{T q^{(k)}}{\|T q^{(k)}\|}$$

$$\mu_k = [q^{(k)}]^T T q^{(k)}$$

SIMULTANEOUS POWER ITERATION

$\tilde{Q}^{(0)}$ – orthogonal matrix

$$\tilde{Q}^{(k+1)} \tilde{R}^{(k+1)} = T \tilde{Q}^{(k)}$$

$$\tilde{T}^{(k)} = [\tilde{Q}^{(k)}]^T T \tilde{Q}^{(k)}$$

Power Iteration

under mild assumptions

* $q^{(k)}$ becomes aligned with $\text{span}\{v_n\}$. as $k \to \infty$
* $\mu_k \to \lambda_n$ as $k \to \infty$

(Recall $\lambda_n, \dots, \lambda_1$ are eigenvalues of A s.t. $|\lambda_n| \geq \lambda_{n-1} \geq \dots \geq |\lambda_1|$ and v_k is eigenvector corr. to λ_k)

Simultaneous Power Iteration

under mild assumptions

* $\tilde{Q}^{(k)}(:,j)$ becomes aligned with $\text{span}\{v_{n-j+1}\}$ as $k \to \infty$ for $j = 1, 2, \dots, n$
* $\tilde{T}^{(k)}_{jj} \to \lambda_{n-j+1}$ as $k \to \infty$ for $j = 1, 2, \dots, n$.

QR Algorithm (with $\mu_k = 0$)

$$T^{(0)} = T$$

$$T^{(k)} = \hat{Q}^{(k+1)} \hat{R}^{(k+1)}$$

$$T^{(k+1)} = \hat{R}^{(k+1)} \hat{Q}^{(k+1)}$$

<u>THM</u>

The sequences $\{T^{(k)}\}$, $\{\hat{Q}^{(k)}\}$ by the QR algorithm, and $\{\tilde{T}^{(k)}\}$, $\{\tilde{Q}^{(k)}\}$ by the simultaneous power iteration with $\tilde{Q}^{(0)} = I_n$ satisfy the following:

(1) $\tilde{Q}^{(k)} = \hat{Q}^{(1)} \hat{Q}^{(2)} \dots \hat{Q}^{(k)} \quad \forall k,\ k \geq 1.$

(2) $\tilde{T}^{(k)} = T^{(k)} \quad \forall k.$

<u>Proof</u>

By induction on k.

Base case $k=1$

First observe $T^{(0)} = T = [\tilde{Q}^{(0)}]^T T \tilde{Q}^{(0)} = \tilde{T}^{(0)}$,

so

$$\hat{Q}^{(1)} \hat{R}^{(1)} = T^{(0)} = T = \tilde{Q}^{(1)} \tilde{R}^{(1)}$$

and we have $\underline{\hat{Q}^{(1)} = \tilde{Q}^{(1)}}$ by the uniqueness of the QR factorization (assuming QR factorizations are computed so that the diagonal entries of the R factors are positive).

Furthermore,

$$T^{(1)} = \hat{R}^{(1)}\hat{Q}^{(1)} = [\hat{Q}^{(1)}]^T T^{(0)} \hat{Q}^{(1)}$$
$$= [\tilde{Q}^{(1)}] T \tilde{Q}^{(1)} = \tilde{T}^{(1)}.$$

Inductive case

Suppose (1) and (2) hold for a given k, would like to establish the validity of (1) and (2) for k+1.

By the inductive hypothesis

$$T^{(k)} = \tilde{T}^{(k)} = [\tilde{Q}^{(k)}]^T T \tilde{Q}^{(k)}$$
$$= \{[\tilde{Q}^{(k)}]^T \tilde{Q}^{(k+1)}\} \tilde{R}^{(k+1)},$$

where the last expression is a QR factorization. As we also have $T^{(k)} = \hat{Q}^{(k+1)}\hat{R}^{(k+1)}$ and by the uniqueness of the QR factorization

$$(*) \quad \hat{Q}^{(k+1)} = [\tilde{Q}^{(k)}]^T \tilde{Q}^{(k+1)} \quad \left(\begin{array}{l} \text{as well as} \\ \hat{R}^{(k+1)} = \tilde{R}^{(k+1)} \end{array} \right)$$

$$\Longrightarrow \tilde{Q}^{(k+1)} = \tilde{Q}^{(k)} \hat{Q}^{(k+1)}$$

$$\xRightarrow[\text{hypothesis}]{\text{by the inductive}} \tilde{Q}^{(k+1)} = \hat{Q}^{(1)} \cdots \hat{Q}^{(k)} \hat{Q}^{(k+1)}.$$

Additionally,

$$\tilde{T}^{(k+1)} = [\tilde{Q}^{(k+1)}]^T T \tilde{Q}^{(k+1)} = \tilde{R}^{(k+1)} [\tilde{Q}^{(k)}]^T \tilde{Q}^{(k+1)}$$
$$\overset{\text{(see (*))}}{=} \hat{R}^{(k+1)} \hat{Q}^{(k+1)} = T^{(k+1)}.$$

Hence, the result follows from induction. $\square$