

Eigenvalues and Eigenvectors

Gerardo Toraldo

Università di Napoli Federico II

A.A. 2016-2017

About Matrices: Just to remember...

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a^1, a^2, \dots, a^n)^T \in \mathbb{R}^{m \times n}$$

Any matrix A defines the linear transformation in \mathbb{R}^n

$$f(\mathbf{x}) = Ax = (\mathbf{a}^1)x_1 + (\mathbf{a}^2)x_2, \dots, (\mathbf{a}^n)x_n$$

In particular (if \mathbf{e}_i are the basis vectors (b.v.)) in \mathbb{R}^n

- ❶ $f(\mathbf{e}_i) = \mathbf{a}^i$
- ❷ the linear transformation f is uniquely determined by its values at the (b.v.) \mathbf{e}_i .

Some basic preliminaries

The form of the matrix affects the way to solve linear algebra problems.

- A symmetric if $A = A^T$
- A Hermitian if $A = A^+$
- A Normal if $AA^+ = A^+A$
- U Orthonormal if $U^{-1} = U^T$
- U Unitary if $U^{-1} = U^+$

where

- $A^T \in \mathfrak{R}^{n \times m}$ is the transpose of A
- A^+ is the conjugate transpose or Hermitian transpose of A , obtained from A by taking the transpose and then taking the complex conjugate of each entry.

About Matrices: Just to remember...

Matrix Norm $\|A\|$ is a vector norm on $\mathfrak{R}^{m \times n}$, i.e.

- $\|A\| = 0 \geq 0$
- $\|A\| = 0 \Leftrightarrow A = 0$
- $|\alpha|\|A\| = \|\alpha A\| = |\alpha|\|A\|$, $\alpha \in \mathfrak{R}$, $A \in \mathfrak{R}^{m \times n}$ (absolutely homogeneous)
- $\|A + B\| \leq \|A\| + \|B\|$ (subadditive)

Submultiplicative (Consistent) on $\mathfrak{R}^{n \times n}$ if

- $\|AB\| \leq \|A\|\|B\|$

The set $\mathfrak{R}^{n \times n}$ equipped with a submultiplicative norm is an Banach algebra.

A matrix **norm** $\|\cdot\|$ on $\mathfrak{R}^{n \times n}$ is **compatible** with a vector norm $\|\cdot\|_a$ on \mathfrak{R}^n iff

$$\|A\mathbf{x}\|_a \leq \|A\|\|\mathbf{x}\|_a$$

(**geometrically**: The linear transformation through f can (at most) extend the length of a vector \mathbf{x} of $\|A\|$ times)

Given a vector norm $\|\cdot\|$ in \mathbb{R}^n , the following norm is induced in $\mathbb{R}^{n \times n}$

$$\begin{aligned}\|A\| &= \sup\{\|A\mathbf{x}\| : \mathbf{x} \in K^n \text{ with } \|\mathbf{x}\| = 1\} \\ &= \sup\left\{\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} : \mathbf{x} \in K^n \text{ with } \mathbf{x} \neq 0\right\}.\end{aligned}$$

For any induced norm $\|\cdot\|$ is $\|I\| = 1$

The matrix norm corresponding to the p -norm for vectors, $p \geq 1$, is:

$$\|A\|_p = \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p}.$$

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|, \quad (\text{maximum absolute column sum of the matrix})$$

$$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}| \quad (\text{maximum absolute row sum of the matrix}),$$

$$\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2} = \sqrt{\text{trace}(A^*A)} = \sqrt{\sum_{i=1}^n \sigma_i^2} \quad (\text{Frobenius or l2 norm}).$$

$$\|A\|_{\max} = \max_{ij} |a_{ij}| \quad (\text{max norm}).$$

Norms equivalency

Any two matrix norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ are equivalent. i.e.,

$$r \|A\|_\alpha \leq \|A\|_\beta \leq s \|A\|_\alpha$$

for some positive numbers r and s .

Values of r and s for some norms ($A \in \mathfrak{R}^{n \times n}$ full rank):

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n} \|A\|_2$$

$$\|A\|_F \leq \|A\|_* \leq \sqrt{n} \|A\|_F$$

$$\|A\|_{\max} \leq \|A\|_2 \leq n \|A\|_{\max}$$

$$\frac{1}{\sqrt{n}} \|A\|_\infty \leq \|A\|_2 \leq \sqrt{n} \|A\|_\infty$$

$$\frac{1}{\sqrt{n}} \|A\|_1 \leq \|A\|_2 \leq \sqrt{n} \|A\|_1, \|A\|_2 \leq \sqrt{\|A\|_1 \|A\|_\infty},$$

The eigenproblem

Find scalars λ and vector \mathbf{x} such that

$$A\mathbf{x} = \lambda\mathbf{x} \quad (\text{right}) \quad (1)$$

$$\mathbf{y}^+ A = \lambda \mathbf{y}^+ \quad [Y^T A = \lambda Y^T \text{ if } Y \in \mathbb{R}^n] \quad (\text{left}) \quad (2)$$

or (in block eigenvectors form)

$$AX = XD, \quad Y^+ A = DY^+ \quad [Y^T A = DY^T \text{ if } Y \in \mathbb{R}^{m \times n}], \quad (3)$$

where the columns of X and Y are the right and left eigenvectors. Note:

$$Y^+ A = DY^+ \iff A^+ Y = YD^+ \quad [Y^T A = DY^T \iff A^T Y = YD]$$

(the left eigenvalues/eigenvectors problem is equivalent of computing the right eigenvalues/eigenvectors of A^+ (and taking the conjugate of the eigenvalues))

NOTE: If A is symmetric: $A\mathbf{x} = \lambda\mathbf{x} \iff \mathbf{x}^T A = \lambda\mathbf{x}^T$ (right and left eigenvectors coincide)

In matlab: `A=gallery('grcar',5,1);`

`[X,D]=eig(A)`

`max(A*X-X*D)`

`[Y,D1]=eig(A')`

`max(D-D1)`

`max(Y.'*A-D1*Y.')`

The Eigenvalues are the roots of the **Characteristic Polynomial Equation** (on λ of order n):

$$p_A(\lambda) = \det(A - \lambda I) = 0 \quad (4)$$

If A is a real matrix, all of its complex eigenvalues are in complex conjugate pairs. Therefore, **the eigenvalues of the left and right eigenvectors of a real matrix are the same**. Also, if A is a real matrix, all of its complex eigenvectors are in **complex conjugate pairs**. Therefore, the left eigenvectors simplify to the transpose of the right eigenvectors of A^T if A is real.

Note

- There are n eigenvalues, not necessarily distinct (Fundamental Theorem of Algebra);
- $\lambda(\mathbf{x}) \in \sigma(A) \Rightarrow \lambda(\mathbf{x}) = \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$;
- $\det(A) = \prod \lambda_i$;
- $\text{tr}(A) = \sum a_{ii} = \sum \lambda_i$;
- $\sigma(A) = \sigma(A^T)$ (and then $\rho(A) = \rho(A^T)$)

where the spectral radius of A is defined as

$$\rho(A) = \max \{ |\lambda_1|, \dots, |\lambda_n| \}.$$

and $\sigma(A) = \{ \lambda_i, i = 1 \dots n \}$ is the spectrum of A .

Note: for any compatible norm $\|\cdot\|$

$$\rho(A) \leq \|A\|$$

About the 2-norm

$$\|A\|_2 = \sup \frac{\|Ax\|_2}{\|x\|_2} = \sup \sqrt{\frac{x^T A^T A x}{x^T x}} = \sqrt{\rho(A^T A)}$$

If A is symmetric, then

$$\|A\|_2 = \rho(A)$$

This is not true if there is not symmetry.

Example (from wikipedia)

$$A = \begin{pmatrix} 1 & 1000 \\ 0 & 1 \end{pmatrix}$$

$\sigma(A) = \{1\}$ and then $\rho(A) = 1$.

$$A^T A = \begin{pmatrix} 1 & 1000 \\ 1000 & 1000001 \end{pmatrix}$$

and $\sigma(A^T A) = \{1000001.9999\dots, 0.000000999998\dots\}$, thus the spectral radius is $1000001.9999\dots$, and thus $\|A\|_2$ is about $1000.000999\dots$

For a triangular matrix

$$A = \begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{12} & \cdots & \mathbf{a}_{1q} \\ 0 & \mathbf{a}_{22} & \cdots & \mathbf{a}_{2q} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \mathbf{a}_{qq} \end{bmatrix}$$

one has

$$\det(A) = \prod_{i=1 \dots q} a_{ii}; \quad \sigma(A) = \{a_{ii}\}$$

For a block triangular matrix

$$A = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1q} \\ 0 & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2s} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \mathbf{A}_{qq} \end{bmatrix}$$

one has

$$\det(A) = \prod_{i=1 \dots q} \det(A_{ii}); \quad \sigma(A) = \bigcup \sigma(A_{ii})$$

Similarity (and other) transformation

We recall that B is **similar** to A if there is a nonsingular matrix T such that $B = T^{-1}AT$ (**unitarily similar** if T is unitary) Then

$$B\mathbf{y} = \lambda\mathbf{y} \Leftrightarrow T^{-1}AT\mathbf{y} = \lambda\mathbf{y} \Leftrightarrow A(T\mathbf{y}) = \lambda(T\mathbf{y})$$

so A and B have same eigenvalues, and if \mathbf{y} is eigenvector of B , then $\mathbf{x} = T\mathbf{y}$ is eigenvector of A

Similarity transformations preserve eigenvalues and eigenvectors are easily recovered

Moreover

Shift If $B = A - \mu I$, then $\lambda \in \sigma(A) \Leftrightarrow \lambda - \mu \in \sigma(A - \mu I)$

Inverse If A is nonsingular then $\lambda \in \sigma(A) \Leftrightarrow 1/\lambda \in \sigma(A^{-1})$

Polynomial $\lambda \in \sigma(A) \Leftrightarrow P(\lambda) \in \sigma(P(A))$ where $P(x)$ is a polynomial.

Theorem (Schur decomposition)

If $A \in \mathbb{R}^{n \times n}$, then there a unitary matrix Q such that A can be expressed as $A = QUQ^{-1} = QUQ^T$ where

$$U = \begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} & \dots & u_{1,n} \\ & u_{2,2} & u_{2,3} & \dots & u_{2,n} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & u_{n-1,n} \\ 0 & & & & u_{n,n} \end{bmatrix}$$

and therefore A is similar to a triangular matrix, since

$$U = Q^{-1}AQ$$

As a consequence every normal (symmetric for instance) matrix is diagonalizable.

Spectral radius and norms

Theorem

Let $A \in \mathbb{R}^{n \times n}$ with spectral radius $\rho(A)$ and a consistent matrix norm $\|\cdot\|$. Then for each $k \in \mathbb{N}$

$$\rho(A) \leq \|A^k\|^{\frac{1}{k}}$$

Moreover

$$\rho(A) < 1 \Leftrightarrow \lim_{k \rightarrow \infty} A^k = 0 \quad (5)$$

$$\rho(A) > 1 \Leftrightarrow \lim_{k \rightarrow \infty} \|A^k\| = \infty \quad (6)$$

$$\rho(A) = \lim_{k \rightarrow \infty} \|A^k\|^{\frac{1}{k}} \quad \text{Gelfand's Formula} \quad (7)$$

Eigendecomposition

Let assume A has n **linearly independent eigenvectors** v_1, v_2, \dots, v_n with (not necessarily distinct) eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. Let consider the matrix

$$Q = [v_1, v_2, \dots, v_n]$$

then

$$AQ = [\lambda_1 v_1, \quad \lambda_2 v_2, \quad \dots \quad \lambda_n v_n] = Q\Lambda, \text{ where } \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

and then

$$A = Q\Lambda Q^{-1}, \quad Q^{-1}AQ = \Lambda$$

Theorem (diagonalizable matrix)

A matrix A is diagonalizable (i.e. similar to a diagonal matrix) iff the eigenvectors form a basis.

Examples from Wikipedia

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 4 & 9 \end{bmatrix}.$$

$$\begin{aligned} |A - \lambda I| &= \left| \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 4 & 9 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right| = \begin{vmatrix} 2 - \lambda & 0 & 0 \\ 0 & 3 - \lambda & 4 \\ 0 & 4 & 9 - \lambda \end{vmatrix}, \\ &= (2 - \lambda)[(3 - \lambda)(9 - \lambda) - 16] = -\lambda^3 + 14\lambda^2 - 35\lambda + 22. \end{aligned}$$

$\sigma(A) = \{1, 2, 11\}$, with eigenvectors $v_1 = [1 \ 0 \ 0]^T$, $v_2 = [0 \ 2 \ -1]^T$, $v_3 = [0 \ 1 \ 2]^T$,

$$B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\lambda_1 = 1, \lambda_2 = -1/2 + \mathbf{i}\sqrt{3}/2, \lambda_3 = -1/2 - \mathbf{i}\sqrt{3}/2$$

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ -1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

0.28868+9.7145e-17i	0.28868-9.7145e-17i	0.57735	0.5	0.5
-1.1102e-16+0.5i	-1.1102e-16-0.5i	1.3752e-16	2.3592e-16+0.5i	2.3592e-16-0.5i
-0.57735	-0.57735	0.57735	1.1093e-31+1.1102e-16i	1.1093e-31-1.1102e-16i
1.6653e-16-0.5i	1.6653e-16+0.5i	-9.3225e-18	1.249e-16+0.5i	1.249e-16-0.5i
0.28868+9.7145e-17i	0.28868-9.7145e-17i	0.57735	-0.5+4.1633e-17i	-0.5-4.1633e-17i

0.29+0.00i	0.29-0.00i	0.58	0.50	0.50
-0.00+0.50i	-0.00-0.50i	0.00	0.00+0.50i	0.00-0.50i
-0.58	-0.58	0.58	0.00+0.00i	0.00-0.00i
0.00-0.50i	0.00+0.50i	-0.00	0.00+0.50i	0.00-0.50i
0.29+0.00i	0.29-0.00i	0.58	-0.50+0.00i	-0.50-0.00i

Eigenspace of λ : $\text{Eig}(\lambda) =$ set of eigenvectors having λ as eigenvalue
($\text{Eig}(\lambda)$ invariant w.r.t. A)

Note: distinct eigenvalues always have linearly independent eigenvectors.

Geometric multiplicity (Gm) of λ = $\dim \text{Eig}(\lambda) = \dim[\ker(A - \lambda I)] = n - \text{rank}(A - \lambda I)$

Algebraic multiplicity (Am) of λ = multiplicity of λ as root of $P_A(\lambda)$

Theorem

Geometric multiplicity \leq Algebraic multiplicity

(λ defective iff $\text{Gm}(\lambda) < \text{Am}(\lambda)$)

Definition

$A \in \mathbb{C}^{n \times n}$ is **non-defective** if it has a set of n independent eigenvectors

Theorem

*A non defective iff A is **diagonalizable** ($D = X^{-1}AX$ for some X , with D diagonal).*

Example: Any Jordan matrix of size 2×2 or larger is defective.

$$J_n = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

J_n has only one eigenvalue, λ , with algebraic multiplicity n , but only one distinct eigenvector, $e = [1 \ 0 \ \dots \ 0]^T$.

Theorem (Jordan form of a matrix)

$A \in \mathbb{C}^{n \times n} \Rightarrow \exists X$ nonsingular such that

$$X^{-1}AX = \text{diag}(J_1, J_2 \dots J_m) \text{ with } J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} \in \mathbb{C}^{n_i \times n_i} \quad (8)$$

with $\sum n_i = n$; note that if $n_i > 1$ then λ_i defective.

SVD decomposition

Theorem

Let $A \in C^{m \times n}$. There exist two unitary matrices $U \in C^{m \times m}$ and $V \in C^{n \times n}$ such that

$$U^+AV = \Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_p) \in \mathbb{R}^{m \times n} \quad (9)$$

with $p = \min(m, n)$ and $\sigma_1 \geq \dots \geq \sigma_p \geq 0$. Formula (9) is called **Singular Value Decomposition or (SVD) of A** and the numbers σ_i (or $\sigma_i(A)$) are called singular values of A .

NOTE: $U^+AVV^+A^+U = U^+AA^+U = \Sigma^2 = V^+A^+AV$ and then

$$AA^+U = U\Sigma^2, \quad A^+AV = V\Sigma^2$$

- If A is a real-valued matrix, U and V will also be real-valued
- The squared singular values are eigenvalues of the normal matrix, i.e. $\sigma_i(A) = \sqrt{\lambda_i(A^+A)} = \sqrt{\lambda_i(AA^+)}$
- the columns of U , (left singular vectors) of A are the eigenvectors of AA^+
- the columns of V , (right singular vectors) of A are the eigenvectors of A^+A
- if A is symmetric the singular values are the eigenvalues of the matrix (in absolute value).

The condition number of a matrix $A \in C^{n \times n}$ is defined as

$$\kappa(A) = \|A\| \|A^{-1}\|$$

where $\|\cdot\|$ is an induced matrix norm. obviously $\kappa(A) \geq \|AA^{-1}\| = 1$, $\kappa(A^{-1}) = \kappa(A)$, $\kappa(A) = 1$ if A is orthogonal, $K_2(A) = 1$. The condition number of a singular matrix is set equal to infinity.

Condition number and eigenvalues

$$\kappa_2(A) = \sigma_1(A)/\sigma_n(A)$$

and for symmetric definite matrices

$$\kappa_2(A) = \frac{\lambda_{max}}{\lambda_{min}} = \rho(A)\rho(A^{-1})$$

for that reason, $\kappa_2(A)$ is called **spectral condition number**.

Eigenvalues sensitivity (a priori and a posteriori estimates)

Assumption: A is diagonalizable with $D = \text{diag}(\lambda_i) = X^{-1}AX$

Theorem (Bauer Fike 1)

If μ is an eigenvalue of $A + E$ (E is a perturbation matrix), then

$$\min_{\lambda \in \sigma(A)} |\lambda - \mu| \leq \kappa_p(X) \|E\|$$

Theorem (Bauer Fike 2)

Let $(\tilde{\lambda}, \tilde{v})$ be an approximate eigenvalue-eigenvector couple, and let $\mathbf{r} = A\tilde{v} - \tilde{\lambda}\tilde{v}$ be the residual (with $\|\tilde{v}\| = 1$). Then, there exists $\lambda \in \sigma(A)$ such that

$$|\lambda - \tilde{\lambda}| \leq \kappa_p(X) \|\mathbf{r}\|$$

Theorem (Bauer Fike 3)

Let assume A be invertible. If μ is an eigenvalue of $A + E$ (E is a perturbation matrix), then

$$\min_{\lambda \in \sigma(A)} \frac{|\lambda - \mu|}{|\lambda|} \leq \kappa_p(X) \frac{\|E\|}{\|A\|}$$

Eigenvalues sensitivity (a priori and a posteriori estimates)

Theorem (Bauer Fike 1)

If μ is an eigenvalue of $A + E$ (E is a perturbation matrix), then

$$\min_{\lambda \in \sigma(A)} |\lambda - \mu| \leq \kappa_p(X) \|E\|$$

proof: let assume $\mu \notin \sigma(A)$;

$\mu \in \sigma(A + E) \Rightarrow X^{-1}(A + E - \mu I)X = D + X^{-1}EX - \mu I$ is a singular matrix, i.e.

$$\begin{aligned} \exists \mathbf{v} \in \mathfrak{R}^n : (D + X^{-1}EX - \mu I)\mathbf{v} = 0 &\Leftrightarrow (D - \mu I)^{-1}[(D - \mu I) + X^{-1}EX]\mathbf{v} = 0 \\ &\Leftrightarrow [I + (D - \mu I)^{-1}X^{-1}EX]\mathbf{v} = 0 \quad \text{and then} \end{aligned}$$

$$\|\mathbf{v}\| = \|(D - \mu I)^{-1}X^{-1}EX\mathbf{v}\| \leq \|(D - \mu I)^{-1}\| \kappa(X) \|\mathbf{v}\| \|E\|$$

$$\Leftrightarrow 1 \leq \left\| \begin{bmatrix} (\lambda_1 - \mu)^{-1} & & \\ & \ddots & \\ & & (\lambda_n - \mu)^{-1} \end{bmatrix} \right\| \kappa(X) \|E\|$$

Eigenvalues sensitivity (a priori and a posteriori estimates)

Numerical stability of the problem:

- The condition number of matrix of eigenvectors is the condition number of eigenvalues;
- Eigenvalues may be sensitive if eigenvectors are nearly linearly dependent (i.e., matrix is nearly defective)
- **For normal matrix** ($A^*A = AA^*$): $\min_{\lambda \in \sigma(A)} |\lambda - \mu| \leq \|E\|$. Eigenvectors are orthogonal, so **eigenvalues are well-conditioned** (small perturbation in A leads to small perturbation on the spectrum): .

Individual eigenvalues Stability: Diagonalizable matrices

\mathbf{x}, \mathbf{y} right and left eigenvectors corresponding to the eigenvalue λ (i.e. $A\mathbf{x} = \lambda\mathbf{x}$, $\mathbf{y}^T A = \lambda\mathbf{y}^T$). Let consider the perturbed problem

$$A(\varepsilon)\mathbf{x}(\varepsilon) = \lambda(\varepsilon)\mathbf{x}(\varepsilon), \text{ where } A(\varepsilon) = A + \varepsilon\delta A$$

If we differentiate w.r.t. ε and premultiply by \mathbf{y}^T we get

$$\mathbf{y}^T \delta A \mathbf{x} + \mathbf{y}^T A \mathbf{x}' = \lambda' \mathbf{y}^T \mathbf{x} + \mathbf{y}^T \lambda \mathbf{x}'$$

and since $\mathbf{y}^T A = \lambda \mathbf{y}^T$ we get

$$\frac{d\lambda}{d\varepsilon} = \frac{\mathbf{y}^T \delta A \mathbf{x}}{\mathbf{y}^T \mathbf{x}}$$

or even

$$|\delta\lambda| \simeq \frac{|\mathbf{y}^T \delta A \mathbf{x}|}{|\mathbf{y}^T \mathbf{x}|} |\delta\varepsilon| \quad (10)$$

Individual eigenvalues Stability:

Recalling that $\mathbf{y}^T \mathbf{x} = \|\mathbf{y}\| \|\mathbf{x}\| \cos \theta_\lambda$, assuming $\|\delta A\| = 1$, from (10) it follows

$$|\delta \lambda| \leq \frac{1}{|\cos \theta_\lambda|} |\delta \varepsilon| \quad (11)$$

For small perturbations E on A (say of size ε), $\lambda \in \sigma(A)$, then its perturbation $\tilde{\lambda} \in \sigma(A + E)$, is such that

$$|\lambda - \tilde{\lambda}| \leq \frac{\varepsilon}{|\mathbf{y}^T \mathbf{x}|},$$

where \mathbf{x} and \mathbf{y} are the (unitary) left and right eigenvectors associated to λ .

The quantity $\kappa(\lambda) = 1/|\mathbf{y}^T \mathbf{x}| = 1/\cos \theta_\lambda$ is the **condition number of the eigenvalue** λ (θ_λ is the angle between \mathbf{x} and \mathbf{y}).

- For symmetric or Hermitian matrix, right and left eigenvectors are the same, $\kappa(\lambda) = 1$ and eigenvalues are inherently well-conditioned
- Eigenvalues of nonnormal matrices may be sensitive,
- For multiple/closely clustered eigenvalues, eigenvectors may be sensitive

In matlab try this:

```

> A=gallery('grcar',5,1);
> cond(A)
ans = 2
> x=L(1:5,1); > y=R(1:5,1);
> [L,DL]=eig(A);
> [R,DR]=eig(A');
> klambda = 1/(x'*y)
ans = -9.0072e+15

```

Eigenvalues and eigenvectors stability: examples

$$E = \begin{bmatrix} 101 & -90 \\ 110 & -980 \end{bmatrix}$$

- » $[X,D,Y]=\text{eig}(A)$;
 - » $x1=X(1:2,1);x2=X(1:2,2)$;
 - » $y1=Y(1:2,1);y2=Y(1:2,2)$;
- Conditioning of the first eigenvalue**

- » $\text{COND1}=1/\text{norm}(x1'*y1)$;
- ans:COND1=2.00e+02

Build the Perturbation matrix

$$E = \begin{bmatrix} 1.0e-03 & 1.0e-03 \\ 0 & 0 \end{bmatrix}$$

- » $E=[1e-3 \ 0; \ 0 \ 1e-3]$; » $AE=A+E$;
 - » $[XE,DE,YE]=\text{eig}(AE)$;
 - » $x1E=XE(1:2,1);x2E=XE(1:2,2)$;
 - » $y1E=YE(1:2,1);y2E=YE(1:2,2)$;
 - » $\text{norm}(D(1,1)-DE(1,1))$ ans = 0.3
- The perturbation in the eigenvalue is 2 order of magnitude bigger that the perturbation on the matrix**

- » $B=\text{diag}([1e-1,1e-2,1e-3,1e-4,1e-5])$
 - » $Q=\text{gallery}('orthog',5)$;
 - » $B=Q'*B*Q$;
 - » $E=\text{diag}([1e-3 \ 1e-5 \ 0 \ 0 \ 0])$
 - » $[XB,DB,YB]=\text{eig}(B)$;
 - » $B1=B+E$;
 - » $B2=B+X'*E*Q$;
 - » $[XB1,DB1,YB1]=\text{eig}(B1)$;
 - » $[XB2,DB2,YB2]=\text{eig}(B2)$;
 - » $\text{norm}(XB(1:5,1)-XB2(1:5,1))$
- ans = 0.0058
- » $\text{norm}(XB(1:5,4)-XB2(1:5,4))$
- ans = 1.6905e-13

Why In the second perturbation matrix the 5th eigenvector does not change?

Eigenvalues sensitivity (a posteriori estimates)

Theorem

Let assume A hermitian, and \tilde{x} , $\tilde{\lambda}$ be **couple of eigenvector-eigenvalue computed**. Then

$$\min_{\lambda_i \in \sigma(A)} |\lambda_i - \tilde{\lambda}| \leq \frac{\|\tilde{r}\|}{\|\tilde{x}\|}$$

where $\tilde{r} = A\tilde{x} - \tilde{\lambda}\tilde{x}$ (residual)

Theorem

Let assume A diagonalizable, with X matrix of eigenvectors, and \tilde{x} , $\tilde{\lambda}$ be **couple of eigenvector-eigenvalue computed**. Then

$$\frac{\|\tilde{r}\|}{\|\tilde{x}\|} \leq \varepsilon \Rightarrow \min_{\lambda_i \in \sigma(A)} |\lambda_i - \tilde{\lambda}| \leq \varepsilon \|X^{-1}\| \cdot \|X\|.$$

where $\tilde{r} = A\tilde{x} - \tilde{\lambda}\tilde{x}$

Eigenvectors sensitivity

Theorem

Eigenvector conditioning If \mathbf{x}_k is an eigenvector of A , there exists an eigenvector $\mathbf{x}_k(\varepsilon)$ of $A(\varepsilon) = A + \varepsilon E$ such that

$$\|x_k - \mathbf{x}_k(\varepsilon)\| \leq \frac{\varepsilon}{\min_{j \neq k} \|\lambda_k - \lambda_j\|}$$

The condition number of an eigenvector depends on the **gap between its eigenvalue and the closest other eigenvalue**: the smaller the gap the more sensitive the eigenvectors.

Eigenvectors sensitivity (from netlib)

Eigenvectors and eigenspaces can be ill-conditioned. For example, changing

$$A_0 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & \epsilon \\ 0 & \epsilon & 1 \end{bmatrix} \text{ to } A_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 + \epsilon \end{bmatrix}$$

rotates the two eigenvectors corresponding to the two eigenvalues near 1 by $\pi/4$, no matter how small ϵ is. Thus they are very sensitive to small changes. In this example the two eigenvalues near 1 are very close, so their gaps are small and their eigenvectors are sensitive. But the two-dimensional invariant subspace they span is very insensitive to changes in A (because their eigenvalues, both near 1, are very far from the next closest eigenvalue, at 2).

With matlab ($\epsilon = 0.001$)

```
>> [V0, D0] = eig(A0)
```

```
>> [V1, D1] = eig(A1)
```

$$V0 = \begin{bmatrix} 0 & 0 & 1.0000 \\ -0.7071 & 0.7071 & 0 \\ 0.7071 & 0.7071 & 0 \end{bmatrix}$$

$$V1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$D0 = \begin{bmatrix} 0.9990 & 0 & 0 \\ 0 & 1.0010 & 0 \\ 0 & 0 & 2.0000 \end{bmatrix}$$

$$D1 = \begin{bmatrix} 1.0000 & 0 & 0 \\ 0 & 1.0010 & 0 \\ 0 & 0 & 2.0000 \end{bmatrix}$$

$\gg [V0, D0] = \text{eig}(A0)$

$$V0 = \begin{bmatrix} 0 & 0 & 1.0000 \\ -0.7071 & 0.7071 & 0 \\ 0.7071 & 0.7071 & 0 \end{bmatrix}$$

$$D0 = \begin{bmatrix} 0.9990 & 0 & 0 \\ 0 & 1.0010 & 0 \\ 0 & 0 & 2.0000 \end{bmatrix}$$

 $\gg [V1, D1] = \text{eig}(A1)$

$$V1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$D1 = \begin{bmatrix} 1.0000 & 0 & 0 \\ 0 & 1.0010 & 0 \\ 0 & 0 & 2.0000 \end{bmatrix}$$

The power method

Assumptions: A diagonalizable, $|\lambda_1| > \lambda_2 \geq \lambda_3 \dots \lambda_n$, with $\mathbf{x}_1 \dots \mathbf{x}_n$ orthonormal basis of eigenvectors

Power method

pick \mathbf{q}_0 : $\mathbf{q}_0 = \sum \alpha_i \mathbf{x}_i$, $\alpha_1 \neq 0$,
 $\|\mathbf{q}_0\| = 1$

for $k = 1 \dots$

$$\mathbf{z}_k = A\mathbf{q}_{k-1}$$

$$\mathbf{q}_k = \mathbf{z}_k / \|\mathbf{z}_k\|$$

$$\sigma_k = \mathbf{q}_k^T A \mathbf{q}_k$$

Power method

pick \mathbf{q}_0 : $\mathbf{q}_0 = \sum \alpha_i \mathbf{x}_i$, $\alpha_1 \neq 0$, $\|\mathbf{q}_0\| = 1$
 for $k = 1 \dots$

$$\mathbf{q}_k = \frac{A^k \mathbf{q}_0}{\prod_{i=1..k-1} \|A^i \mathbf{q}_i\|}$$

$$\sigma_k = \mathbf{q}_k^T A \mathbf{q}_k$$

since $A^k \mathbf{q}_0 = \alpha_1 \lambda_1^k \left(\mathbf{x}_1 + \sum_{i=2}^n \frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1} \right)^k \mathbf{x}_i \right)$, $\left| \frac{\lambda_i}{\lambda_1} \right|^k \rightarrow 0$ then

$$\|\mathbf{q}_k - \mathbf{x}_1\| \rightarrow 0, \quad |\sigma_k - \lambda_1| \rightarrow 0$$

- The **normalization** is made to tackle the geometric growth of components at each iteration (risk of possible overflow/underflow)
- The **normalization** has just a **numerical effect** (in infinite precision it has no effect)

The power method: limitations

- The power method fails if the starting point is wrong ($\alpha_1 = 0$)
- The power method fails if there are several eigenvalues of maximum modulus (the method may converge to a linear combination of eigenvectors)
- For real matrices, complex eigenvectors will never be reached if $x_0 \in \mathbb{R}^n$.
- Convergence rate depends on $|\lambda_2/\lambda_1|$

Example:

$$A = \begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.3 & 0 \\ 0 & 0 & 0 & 1.0 \end{bmatrix} x_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0.5 \end{bmatrix} \dots x_5 = \begin{bmatrix} 1.0000 \\ 0.0312 \\ 0.0024 \\ 0.5000 \end{bmatrix} \dots x_{10} = \begin{bmatrix} 1.0000 \\ 0.0010 \\ 0.0000 \\ 0.5000 \end{bmatrix} \quad \sigma = 1.118$$

$$A = \begin{bmatrix} 2.0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.3 & 0 \\ 0 & 0 & 0 & 1.0 \end{bmatrix} x_0 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0.5 \end{bmatrix} \dots x_{100} = \begin{bmatrix} 0. \\ 0. \\ 0. \\ 0.5000 \end{bmatrix}$$

The inverse power method

Assumptions: A diagonalizable, $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \dots |\lambda_{n-1}| > |\lambda_n|$, with $\mathbf{x}_1 \dots \mathbf{x}_n$ orthonormal basis of eigenvectors

Inverse Power method

pick \mathbf{q}_0 : $\mathbf{q}_0 = \sum \alpha_i \mathbf{x}_i$, $\alpha_n \neq 0$, $\|\mathbf{q}_0\| = 1$

for $k = 1 \dots \mathbf{z}_k = A^{-1} \mathbf{q}_{k-1}$

i.e. \mathbf{z}_k is solution of the system $A\mathbf{x} = \mathbf{q}_{k-1}$

$$\mathbf{q}_k = \mathbf{z}_k / \|\mathbf{z}_k\|$$

$$\sigma_k = \mathbf{q}_k^T A \mathbf{q}_k$$

The inverse iteration:

- computes the smallest eigenvalue.
- at each iteration it requires the solution of a system (but the matrix factorization is made just once)

Computing more than one eigenvalue: Deflation

After λ_1 and x_1 have been computed, then x_2 and λ_2 , can be computed by deflation.

- let compute H , nonsingular matrix such that $Hx_1 = e_1$
- H is such that

$$HAH^{-1} = \begin{bmatrix} \lambda_1 & a^T \\ 0 & A_1 \end{bmatrix} \Leftrightarrow A = H^{-1} \begin{bmatrix} \lambda_1 & a^T \\ 0 & A_1 \end{bmatrix} H$$

where $A_1 \in \mathfrak{R}^{(n-1) \times (n-1)}$, with eigenvalues $\lambda_2, \dots, \lambda_n$

- the power method can be used to compute x_2 and λ_2 (provided is $\lambda_2 > \lambda_3$)
- by iterating this process, the first h eigenvalues can be computed

How to compute H ? An orthogonal matrix is the most convenient choice: **Householder**.

Deflation

How to compute the eigenvector \mathbf{x}_2 ?

If $\mathbf{y}_2 \in \mathbb{R}^{n-1 \times n-1}$ such that $A_1 \mathbf{y}_2 = \lambda_2 \mathbf{y}_2$, then if $\mathbf{x} = H^{-1} \begin{bmatrix} \alpha \\ \mathbf{y}_2 \end{bmatrix}$ then

$$A\mathbf{x} = H^{-1} \begin{bmatrix} \lambda_1 & a^T \\ 0 & A_1 \end{bmatrix} H H^{-1} \begin{bmatrix} \alpha \\ \mathbf{y}_2 \end{bmatrix} = H^{-1} \begin{bmatrix} \lambda_1 & a^T \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} \alpha \\ \mathbf{y}_2 \end{bmatrix} = H^{-1} \begin{bmatrix} \lambda_1 \alpha + a^T \mathbf{y}_2 \\ A_1 \mathbf{y}_2 \end{bmatrix}$$

$A\mathbf{x} = \lambda_2 \mathbf{x}_2$ if and only if $\alpha = \frac{a^T \mathbf{y}_2}{\lambda_2 - \lambda_1}$

An alternative deflation approach.

- pick \mathbf{u} such that $\mathbf{u}^T \mathbf{x}_1 = \lambda_1$;
- consider the matrix $A - \mathbf{x}_1 \mathbf{u}^T$ whose eigenvalues are $0, \lambda_2, \lambda_3, \dots, \lambda_n$

The Cholesky factorization

Theorem

Every (strictly) positive definite matrix $A \in \mathbb{R}^{n \times n}$ can be factored as

$$A = R^T R \quad (12)$$

with R (Cholesky factor) upper triangular with positive diagonal elements.

proof: by induction on n . Trivially the theorem holds for $n = 1$. Let assume it holds for

$n = k - 1$. If $A \in \mathbb{R}^{k \times k}$, partitioned as $A = \begin{bmatrix} A_1 & \mathbf{a}^T \\ \mathbf{a} & a_{kk} \end{bmatrix}$ then, there exists

$R_1 \in \mathbb{R}^{(k-1) \times (k-1)}$ upper triangular such that $A_1 = R_1^T R_1$. It easy to check that (12) holds for

$$R = \begin{bmatrix} R_1 & \mathbf{r} \\ 0^T & r_{kk} \end{bmatrix}$$

where $r_{kk} = \sqrt{a_{kk} - \mathbf{r}^T \mathbf{r}}$ and $R_1^T \mathbf{r} = \mathbf{a}$

- (time) complexity order for computing R is $(1/6)n^3$ flops
- The Cholesky factorization gives a practical method for testing positive definiteness

The QR Iteration

Iterative techniques for approximating **all the eigenvalues** of a matrix $A \in \mathbb{R}^{n \times n}$ **simultaneously**.

The basic idea: reducing A , by means of suitable similarity transformations, into a form for which the calculation of the eigenvalues is easier than on the starting matrix

QR iteration

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

for $k = 1, 2, 3, \dots$:

$$Q^{(k)} R^{(k)} = A^{(k-1)};$$

$$A^{(k)} = R^{(k)} Q^{(k)}$$

endfor

At each step

- one QR factorization;
- one matrix matrix product

Notice that $A^{(k)} = R^{(k)} Q^{(k)}$; $R^{(k)} = Q^{(k)T} A^{(k-1)}$
and then $A^{(k)} = Q^{(k)T} A^{(k-1)} Q^{(k)}$, i.e.

$$A^{(k)} = Q^{(1)T} \dots Q^{(k-1)T} Q^{(k)T} A^{(k-1)} Q^{(k)} Q^{(k-1)} \dots Q^{(1)}$$

and every matrix $A^{(k)}$ is orthogonally similar to A and belongs to $\mathbb{R}^{n \times n}$

The QR Iteration

Note: in general, a real matrix can have complex eigenvalues. However

Theorem (Real Schur decomposition)

Given a real matrix $A \in \mathfrak{R}^{n \times n}$, there exists an orthogonal matrix $Q \in \mathfrak{R}^{n \times n}$ such that

$$Q^T A Q = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1m} \\ 0 & R_{22} & \dots & R_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & R_{mm} \end{bmatrix} = \tilde{R} \quad (13)$$

where each block R_{ii} is either a real number or a matrix of order 2 having complex conjugate eigenvalues.

Theorem (Convergence of the QR iteration)

Suppose $A \in \mathbb{R}^{n \times n}$ with eigenvalues λ_i such that $|\lambda_1| \geq |\lambda_2| \dots \geq |\lambda_n|$.

- If the eigenvalues are real, with $|\lambda_i| > |\lambda_{i+1}|$ for $i \leq n - 1$, then the sequence $\{A^{(k)}\}$ generated by the algorithm converges to an upper triangular matrix. i.e. $Q^T A Q$ is a Schur decomposition of A , where $\lim_k Q^{(k)} = Q$
- If the eigenvalues are not real, but simple, with $|\lambda_i| > |\lambda_{i+1}|$ for $i \leq n - 1$ unless λ_i and λ_{i+1} are a complex conjugate pair, then the sequence $\{A^{(k)}\}$ has a pseudo-convergence, i.e. it converges to a upper block triangular matrix of size 1 or 2 (real Schur decomposition of the form (13))

The QR algorithm is a numerical way to realize the Schur decomposition.

Theorem (Convergence rate of the QR iteration)

Suppose $A \in \mathbb{R}^{n \times n}$ with eigenvalues λ_i such that $|\lambda_1| > |\lambda_2| \dots > |\lambda_n|$. Then, for the sequence $\{A^{(k)}\}$ generated by the QR algorithm one has:

$$a_{i,i-1}^{(k)} = O \left| \frac{\lambda_{i-1}}{\lambda_i} \right|^k$$

Numerical issue: If the eigenvalues of A are not well-separated, the convergence of $A^{(k)}$ towards a triangular matrix can be quite slow (use of shift techniques).

The Symmetric QR Algorithm: Householder(**H**) reduction to tridiagonal form

The basic idea: Given $A \in \mathbb{R}^{n \times n}$, symmetric, compute a tridiagonal matrix $T = Q^T A Q$ where Q is product of Householder matrices. The first step: compute the **H** matrix $H_1 \in \mathbb{R}^{(n-1) \times (n-1)}$ such that $H_1 \mathbf{a}_1 = \alpha \mathbf{e}_1$

$$A = A_1 = \begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{a}_1 & B_1 \end{array}; Q_1 = \begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & H_1 \end{array}; Q_1 \cdot A_1 = \begin{array}{c|c} 1 & \mathbf{0} \\ \hline H_1 \mathbf{a}_1 & H_1 B_1 \end{array};$$

$$Q_1 \cdot A_1 \cdot Q_1 = \begin{array}{c|c} 1 & \mathbf{0} \\ \hline H_1 \mathbf{a}_1 & H_1 B_1 H_1 \end{array} = \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \alpha \mathbf{e}_1 & H_1 B_1 H_1 \end{array} \right] = A_2$$

At the k -th step $A_k = \begin{array}{ccc|c} & k-1 & 1 & n-k \\ \begin{array}{c} T_k \\ b_k^T \\ \mathbf{0} \end{array} & \begin{array}{c} b_k \\ a_{kk}^{(\cdot)} \\ \mathbf{a}_k \end{array} & \begin{array}{c} \mathbf{0}^T \\ \mathbf{a}_k^T \\ B_k \end{array} & \begin{array}{c} \text{size} \\ k-1 \\ 1 \\ n-k \end{array} \end{array}$ where T_k is tridiagonal,

$b_k = (0, 0, \dots, *)$. If H_k is the Householder matrix such that $H_k \mathbf{a}_k = \alpha_k \mathbf{e}_1 \in \mathbb{R}^{n-k}$, then

$Q_k = \begin{bmatrix} I_k & \mathbf{0}^T \\ \mathbf{0} & H_k \end{bmatrix}$ is such that

$$A_{k+1} = Q_k \cdot A_k \cdot Q_k = \begin{bmatrix} T_k & b_k & \mathbf{0}^T \\ b_k^T & a_{kk}^{(\cdot)} & \mathbf{a}_k^T H_k \\ \mathbf{0} & H_k \mathbf{a}_k & H_k B_k H_k \end{bmatrix} = \begin{bmatrix} T_k & b_k & \mathbf{0}^T \\ b_k^T & a_{kk}^{(\cdot)} & \mathbf{a}_k^T H_k \\ \mathbf{0} & H_k \mathbf{a}_k & H_k B_k H_k \end{bmatrix}$$

Reduction to tridiagonal form

```

function [H,Q] = houshess(A)
% REDUCTION OF A MATRIX
% TO A SIMILAR HESSENBERG ONE.
% SEE QUARTERONI, SACCO, SALERI
n = max(size(A));
Q = eye(n);
H = A;

for k = 1 : (n - 2)

    [v,beta] = vhouse(H(k + 1 : n, k));

    I = eye(k);
    N = zeros(k, n - k);
    m = length(v);
    R = eye(m) - beta * v * v';
    H(k + 1 : n, k : n) = R * H(k + 1 : n, k : n);
    H(1 : n, k + 1 : n) = H(1 : n, k + 1 : n) * R;
    P = [I, N; N', R];
    Q = Q * P;
end

```

Householder vector

```

function [v,beta] = vhouse(x)
% BUILDING HOUSEHOLDER
% VECTOR.
% SEE QUARTERONI, SACCO,
% SALERI
n = length(x); x = x/norm(x);
s = x(2 : n)' * x(2 : n);
v = [1; x(2 : n)];
if (s == 0)
    beta = 0;
else
    mu = sqrt(x(1)^2 + s);
if (x(1) <= 0)
    v(1) = x(1) - mu;
else
    v(1) = -s/(x(1) + mu);
end
beta = 2 * v(1)^2 / (s + v(1)^2);
v = v/v(1);
end

```

Example

This example reduces the real symmetric matrix A to tridiagonal form, where

$$A = \begin{bmatrix} 2.07 & 3.87 & 4.20 & -1.15 \\ 3.87 & -0.21 & 1.87 & 0.63 \\ 4.20 & 1.87 & 1.15 & 2.06 \\ -1.15 & 0.63 & 2.06 & -1.81 \end{bmatrix}$$

>> $[H, Q] = \text{houshess}(A)$

$H =$

```
2.0700e+000  5.8258e+000  2.1288e-016  1.8350e-015
5.8258e+000  1.4741e+000  2.6240e+000  4.1155e-016
1.8085e-015  2.6240e+000 -6.4916e-001 -9.1627e-001
-3.7668e-016  6.3359e-016 -9.1627e-001 -1.6949e+000
```

$Q =$

```
1.00000  0.00000  0.00000  0.00000
0.00000  0.66429  0.04004  0.74640
0.00000  0.72094  0.22939  -0.65393
0.00000  -0.19740  0.97251  0.12352
```

Pay attention to the numerical values of the *diagonal and off-diagonal elements* of the tridiagonal matrix.