Eigenvalues and Eigenvectors

Gerardo Toraldo

Università di Napoli Federico II

A.A. 2019-2020

About Matrices: Just to remember...

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \left(\mathbf{a}^1, \mathbf{a}^2, \ldots \mathbf{a}^n\right) \in \Re^{m\times n}$$

Any matrix A defines the linear transformation in $\Re^n$

$$f(\mathbf{x}) = A\mathbf{x} = (\mathbf{a}^1)x_1 + (\mathbf{a}^2)x_2, \ldots (\mathbf{a}^n)x_n \text{ where } \mathbf{x} = (x_1, x_2, ..x_n)^T$$

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

1. $f(\mathbf{e}_i) = \mathbf{a}^i$
2. 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 symmetrix if $A = A^T$
- A Hermitian if $A = A^+$
- A Normal if $AA^+ = A^+A$ (A commutes with its adjoint)
- U Orthonormal if $U^{-1} = U^T$
- U Unitary if $U^{-1} = U^+$

where

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

Prove that:

e1 If A is normal, so is $A - \lambda I$ for any scalar λ.

e2 If A is normal, then $AB = BA \Rightarrow A^+B = BA^+$

About Matrices: Just to remember...

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

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

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

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

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

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

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

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

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

$$\|A\| = \sup\{\|A\mathbf{x}\| : \mathbf{x} \in \Re^n \text{ with } \|\mathbf{x}\| = 1\}$$
$$= \sup\left\{\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} : \mathbf{x} \in \Re^n \text{ with } \mathbf{x} \neq 0\right\}.$$

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

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

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

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

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

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

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

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 ($A \in \Re^{n\times n}$ 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

$$A\mathbf{x} = \lambda\mathbf{x} \ \ (\textbf{right}) \tag{1}$$

$$\mathbf{y}^+A = \lambda\mathbf{y}^+ \ [\mathbf{y}^TA = \lambda\mathbf{y}^T \text{ if } \mathbf{y} \in \Re^n] \ \ (\textbf{left}) \tag{2}$$

or (in block eigenvectors form)

$$AX = XD,\ Y^+A = DY^+ \ [Y^TA = DY^T \text{ if } Y \in \Re^{m\times n}], \tag{3}$$

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

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

(the left eigenvelues/eigenvectors problem is equivalent of computing the right eigenvelues/eigenvectors of A^+ (and takig the conjugate of the eigenvalues)
NOTE: If A is symmetric: $A\mathbf{x} = \lambda\mathbf{x} \Leftrightarrow \mathbf{x}^TA = \lambda\mathbf{x}^T$ (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):

$$p_A(\lambda) = \det(A - \lambda I) = 0 \tag{4}$$

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);
- $\lambda(\mathbf{x}) \in \sigma(A) \Rightarrow \lambda(\mathbf{x}) = \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$;
- $\det(A) = \Pi \lambda_i$;
- $\mathrm{tr}(A) = \sum a_{ii} = \sum \lambda_i$;
- $\sigma(A) = \sigma(A^T)$ (and then $\rho(A) = \rho(A^T)$

where the spectral radius of A is defined as

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

and $\sigma(A) = \{\lambda_i,\ i = 1 \ldots n\}$ 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 $\sigma(A^T A) = \{1000001.9999.., 0.000000999998..$, thus the spectral radius is $1000001.9999..$, and thus $\|A\|_2$ is about $1000.000999...$

For a triangular matrix

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

one has

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

For a block triangular matrix

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

one has

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

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 unitay) 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 $\lambda \in \sigma(A) \iff \lambda - \mu \in \sigma(A - \mu I)$

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

Polinomial $\lambda \in \sigma(A) \iff P(\lambda) \in \sigma(P(A))$ where $P(x)$ is a poliynomial.

Spectral radius and norms

Theorem

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

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

Moreover

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

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

$$\rho(A) = \lim_{k \to \infty} \left\|A^k\right\|^{\frac{1}{k}} \quad \textit{Gelfand's Formula} \tag{7}$$

Eigendecomposition

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

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

then

$$AQ = \begin{bmatrix} \lambda_1 v_1, & \lambda_2 v_2, & \ldots & \lambda_n v_n \end{bmatrix} = Q\Lambda, \text{ where } \Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots \lambda_n)$$

and then

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

Theorem (diagonalizable matrix)

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

Theorem (Schur decomposition)

If $A \in \Re^{n \times n}$, then there exists 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 (real symmetrix for instance) matrix is diagonalizable.

Examples from Wikipedia

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

$$|A - \lambda I| = \left| \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 4 & 9 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right| = \begin{vmatrix} 2-\lambda & 0 & 0 \\ 0 & 3-\lambda & 4 \\ 0 & 4 & 9-\lambda \end{vmatrix},$$

$$= (2-\lambda)\big[(3-\lambda)(9-\lambda) - 16\big] = -\lambda^3 + 14\lambda^2 - 35\lambda + 22.$$

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

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

.

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

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

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 : $\mathrm{Eig}(\lambda)$ = set of eigenvectors having λ as eigenvalue
($\mathrm{Eig}(\lambda)$ invariant w.r.t. A)
Note: distinct eigenvalues always have linearly independent eigenvectors.

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

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

Theorem

Geometric multiplicity $\leq$ **Algebraic multiplicity**

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

Definition

$A \in \mathcal{C}^{n \times n}$ 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, $e = \begin{bmatrix} 1\, 0 \ldots 0 \end{bmatrix}^T$.

Theorem (Jordan form of a matrix)

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

$$X^{-1} A X = \textit{diag}(J_1, J_2 \ldots J_m) \textit{ with } J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} \in \mathcal{C}^{n_i \times n_i} \tag{8}$$

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

SVD decomposition

Theorem

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

$$U^+ AV = \Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots \sigma_p) \in \Re^{m\times n} \quad (\textit{i.e. } A = U\Sigma V^+) \tag{9}$$

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

NOTE:

$$U^+ AVV^+ A^+ U = U^+ AA^+ U = \Sigma\Sigma^+ \tag{10}$$

$$V^+ A^+ UU^+ AV = V^+ A^+ AV = \Sigma^+\Sigma \tag{11}$$

and then

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

- 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

SVD decomposition: range and inverse of A

If $\sigma_1 \geq \sigma_2 \ldots \geq \sigma_r \geq \sigma_{r+1} = \ldots \sigma_p = 0$ then

$$\textbf{range}(A) = \textbf{span}\{\mathbf{u}^1 \ldots \mathbf{u}^r\}; \;\; \textbf{ker}(A) = \textbf{span}\{\mathbf{v}^{r+1} \ldots \mathbf{v}^n\}$$

Definition

A Moore-Penrose pseudo inverse of $A \in \Re^{m\times n}$ is any matrix A^* such that

$$AA^*A = A; \;\; A^*AA^* = A^*$$
$$(AA^*)^+ = AA^*; \;\; (A^*A)^+ = A^*A$$

Moore-Penrose pseudo inverse based on SVD

If **rank**$(A) = r$ (with SVD $U^+AV = \Sigma$) then a Moore-Penrose pseudo inverse of$A$ is the matrix

$$A^* = V\tilde{\Sigma}U^+$$

where $\tilde{\Sigma} = \textbf{diag}\{1/\sigma_1 \ldots 1/\sigma_r, 0, \ldots 0\}$

If $A \in \Re^{m\times n}$ with **rank**$(A) = n < m$ then $A^* = (A^TA)^{-1}A^T$

SVD decomposition: square matrices

If $A \in C^{n\times n}$ $[A \in \Re^{n\times n}]$, then $U \in C^{n\times m}$ $[U \in \Re^{n\times n}]$ and $V \in C^{n\times n}$ $[V \in \Re^{n\times n}]$ such that

$$U^+ AV = \Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots \sigma_n) \in \Re^{n\times n} \tag{12}$$

with $\sigma_1 \geq \ldots \geq \sigma_n \geq 0$.

NOTE:

$$U^+ AA^+ U = V^+ A^+ AV = \Sigma^2 \tag{13}$$

and then

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

- If $A \in C^{n\times n}$ with **rank**$(A) = n$ then $A^* = (A)^{-1}$
- since $\sigma_i(A) = \sqrt{\lambda_i(A^+A)} = \sqrt{\lambda_i(AA^+)}$ then, 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

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

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

Condition number and singular values/eigenvalues

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

(where $\alpha/0$ is set to ∞ for any $\alpha \neq 0$) 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.

$$\exists\, \mathbf{v} \in \Re^n : \ (D + X^{-1}EX - \mu I)\mathbf{v} = 0 \Leftrightarrow (D-\mu I)^{-1}[(D-\mu I) + X^{-1}EX]\mathbf{v} = 0$$
$$\Leftrightarrow [I + (D-\mu I)^{-1}X^{-1}EX]\mathbf{v} = 0 \quad \text{and then}$$

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

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

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: Diagonizable matrices

$\mathbf{x}, \mathbf{y}$ right and left eigenvectors corresponding to the eigenvalue λ $(\text{i.e.} A\mathbf{x} = \lambda\mathbf{x},\ y^T A = \lambda \mathbf{y}^T)$. 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 we get

$$(A + \varepsilon\delta A)\mathbf{x}'(\varepsilon) + \delta A\mathbf{x}(\varepsilon) = \lambda'(\varepsilon)\mathbf{x}(\varepsilon) + \lambda(\varepsilon)\mathbf{x}'(\varepsilon)$$

and, if we set $\varepsilon = 0$ and we premultiply by $\mathbf{y}^T$ we get

$$\mathbf{y}^T\delta A\mathbf{x} + \mathbf{y}^T A\mathbf{x}'(0) = \lambda'\mathbf{y}^T\mathbf{x} + \mathbf{y}^T\lambda\mathbf{x}'(0)$$

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

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

or even

$$|\delta\lambda| \simeq \frac{|\mathbf{y}^T\delta A\mathbf{x}|}{|\mathbf{y}^T\mathbf{x}|}|\delta\varepsilon| \tag{14}$$

Individual eigenvalues Stability:

Recalling that $\mathbf{y}^T\mathbf{x} = \|\mathbf{y}\|\|\mathbf{x}\|\cos\theta_\lambda$, assuming $\|\delta A\| = 1$, from (14) it follows

$$|\delta\lambda| \leq \frac{1}{|\cos\theta_\lambda|}|\delta\varepsilon| \tag{15}$$

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]=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

» COND1=1/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]=eig(AE);
» x1E=XE(1:2,1);x2E=XE(1:2,2);
» y1E=YE(1:2,1);y2E=YE(1:2,2);
» 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=diag([1e-1,1e-2,1e-3,1e-4,1e-5])
» Q=gallery('orthog',5);
» B=Q'*B*Q;
» E=diag([1e-3 1e-5 0 0 0])
» [XB,DB,YB]=eig(B);
» B1=B+E;
» B2=B+X'*E*Q;
» [XB1,DB1,YB1]=eig(B1);
» [XB2,DB2,YB2]=eig(B2);
» norm(XB(1:5,1)-XB2(1:5,1))
ans = 0.0058
» 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.

Eigenvecors 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 ($\varepsilon = 0.001$)

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

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

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

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

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

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

$>> [V0, D0] = \mathsf{eig}(A0)$

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

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

$>> [V1, D1] = \mathsf{eig}(A1)$

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

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

The power method

Assumptions: A diagonalizable, $|\lambda_1| > \lambda_2 \geq \lambda_2 \ldots \lambda_n$, with $\mathbf{x}_1 \ldots \mathbf{x}_n$ othonormal 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 $k = 1 \ldots$

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

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

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

Power method

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

for $k = 1 \ldots$

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

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

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

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

- The **normalization** is made to takle the geometric groth 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 poin 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 $x_0 \in \Re^n$.
- Convergence rate depends on $|\lambda_2/\lambda 1|$

Example:

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

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

The inverse power method

Assumptions: A diagonalizable, $|\lambda_1| \geq |\lambda_2| \geq |\lambda_2| \ldots |\lambda_{n-1}| > |\lambda_n|$, with $\mathbf{x}_1 \ldots \mathbf{x}_n$ othonormal 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 $k = 1 \ldots$ $\mathbf{z}_k = A^{-1}\mathbf{q}_{k-1}$
i.e. q_{k-1} is solution of the system $Ax = q_{k-1}$
$\mathbf{q}_k = \mathbf{z}_k / \|\mathbf{z}_k\|$
$\sigma_k = \mathbf{q}_k^T A \mathbf{q}_k$

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 $Hx_1 = \mathbf{e}_1$
- 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 $A_1 \in \Re^{n-1 \times n-1}$, with eigevalues $\lambda_2, ..\lambda_n$
- the power method can be used to compute $\mathbf{x}_2$ 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 $\mathbf{y}_2 \in \Re^{n-1 \times n-1}$ 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

$$A\mathbf{x} = H^{-1} \begin{bmatrix} \lambda_1 & a^T \\ 0 & A_1 \end{bmatrix} H H^{-1} \begin{bmatrix} \alpha \\ \mathbf{y}_2 \end{bmatrix} = H^{-1} \begin{bmatrix} \lambda_1 & a^T \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} \alpha \\ \mathbf{y}_2 \end{bmatrix} = H^{-1} \begin{bmatrix} \lambda_1 \alpha + a^T \mathbf{y}_2 \\ A_1 y_2 \end{bmatrix}$$

$A\mathbf{x} = \lambda_2 \mathbf{x}_2$ 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 $0, \lambda_2, \lambda_3, ..\lambda_n$

The Cholesky factorization

Theorem

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

$$A = R^T R \tag{16}$$

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 $A \in R^{k\times k}$, partitioned as $A = \begin{bmatrix} A_1 & \mathbf{a} \\ \mathbf{a}^T & a_{kk} \end{bmatrix}$ then, there exists $R_1 \in R^{k-1\times k-1}$ upper triangular such that $A_1 = {R_1}^T R_1$. It easy to check that (16) 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 $A \in \Re^{n\times n}$ **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,...:
```

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

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

```
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.

$$A^{(k)} = {Q^{(1)}}^T \dots {Q^{(k-1)}}^T {Q^{(k)}}^T A^{(0)} Q^{(k)} Q^{(k-1)} \dots Q^{(1)}$$

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

The QR Iteration

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

Theorem (Real Schur decomposition)

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

$$Q^T AQ = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1m} \\ 0 & R_{22} & \dots & R_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & R_{mm} \end{bmatrix} = \tilde{R} \tag{17}$$

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 $A \in \Re^{n\times n}$ with eigenvalues λ_i such that $|\lambda_1| \geq |\lambda_2| \ldots \geq |\lambda_n|$.

- *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. $Q^T AQ$ 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 (17))*

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

Theorem (Convergence rate of the QR iteration)

Suppose $A \in \Re^{n\times n}$ with eigenvalues λ_i such that $|\lambda_1| > |\lambda_2| \ldots > |\lambda_n|$. 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).

Householder Transformation (reflection)

Householder matrix

Let $\mathbf{v} \in \Re^n$, $\mathbf{v} \neq 0$, the follwing Householder matrix is associated to $\mathbf{v}$

$$H = I - 2\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$$

H is a rank-1, **orthogonal** and **symmetric** modification of the Identity

The Householder transformation $H(\mathbf{x}) = H\mathbf{x}$ of a vector $\mathbf{x}$ is its reflection (mirror image) with respect to a plane or hyperplane through the origin with unit normal vector $\mathbf{v}/\|\mathbf{v}\|$.

If $\mathbf{u} = \mathbf{v}/\|\mathbf{v}\|$,
(*) $H^T H = (I - 2\mathbf{u}\mathbf{u}^T)(I - 2\mathbf{u}\mathbf{u}^T) = I - 2\mathbf{u}\mathbf{u}^T - 2\mathbf{u}\mathbf{u}^T + 4\mathbf{u}\mathbf{u}^T\mathbf{u}\mathbf{u}^T = I$
(**) $\mathbf{x} \in \Re^n$ can be decomposed as $\mathbf{x} = \tilde{\mathbf{x}} + \mathbf{x}^{\mathbf{u}}$, where $\mathbf{x}^{\mathbf{u}} = (\mathbf{u}^T\mathbf{x})\mathbf{u}$ is the orthogonal projection of $\mathbf{x}$ into Span$\{\mathbf{v}\}$, and $\tilde{\mathbf{x}}$ belongs to the hyperplane through the origin with unit normal vector $\mathbf{u}$, then

$$H\mathbf{x} = \mathbf{x} - 2\mathbf{u}\mathbf{u}^T\mathbf{x} = \tilde{\mathbf{x}} - \mathbf{x}^{\mathbf{u}}$$

Householder Transformation

Householder matrices never formed explicitly:

$$H\mathbf{x} = \left(I - 2\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} \right) \mathbf{x} = \mathbf{x} - 2\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}\mathbf{x} = \mathbf{x} - 2\frac{\mathbf{x}^T\mathbf{v}}{\mathbf{v}^T\mathbf{v}}\mathbf{v}$$

They can be used to zero selected components of a vector. If $\mathbf{v} = \mathbf{x} + \alpha\mathbf{e}_i$ with $\alpha = \pm\|\mathbf{x}\|$ then

$$H\mathbf{x} = -\alpha\mathbf{e}_i$$

Note that $v_i = x_i + \alpha$, choice of the sign to avoid cancellation.

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

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

$$A = A_1 = \begin{array}{c} \begin{matrix} \mathtt{1} & \mathtt{n-1} \end{matrix} \\ \begin{bmatrix} a_{11} & \mathbf{a}_1^T \\ \mathbf{a}_1 & B_1 \end{bmatrix} \end{array} ; Q_1 = \begin{array}{c} \begin{matrix} \mathtt{1} & \mathtt{n-1} \end{matrix} \\ \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & H_1 \end{bmatrix} \end{array} ; \; Q_1 \cdot A_1 = \begin{array}{c} \begin{matrix} \mathtt{1} & \mathtt{n-1} \end{matrix} \\ \begin{bmatrix} a_{11} & \mathbf{a}_1^T \\ H_1\mathbf{a}_1 & H_1B_1 \end{bmatrix} \end{array} ;$$

$$Q_1 \cdot A_1 \cdot Q_1 = \begin{array}{c} \begin{matrix} \mathtt{1} & \mathtt{n-1} \end{matrix} \\ \begin{bmatrix} a_{11} & \mathbf{a}_1^T H_1 \\ H_1\mathbf{a}_1 & H_1B_1H_1 \end{bmatrix} \end{array} = \begin{array}{c} \begin{matrix} \mathtt{1} & \mathtt{n-1} \end{matrix} \\ \left[\begin{array}{c|c} a_{11} & \alpha\mathbf{e}_1^T \\ \hline \alpha\mathbf{e}_1 & H_1B_1H_1 \end{array}\right] \end{array} = A_2$$

At the $k-th$ step $A_k = \begin{matrix} \\ \mathtt{k-1} \\ 1 \\ \mathtt{n-k} \end{matrix} \begin{matrix} \begin{matrix} \mathtt{k-1} & 1 & \mathtt{n-k} \end{matrix} \\ \begin{bmatrix} T_k & b_k & \mathbf{0}^T \\ b_k^T & a_{kk}^{(.)} & \mathbf{a}_k^T \\ \mathbf{0} & \mathbf{a}_k & B_k \end{bmatrix} \end{matrix} \begin{matrix} size \\ \mathtt{k-1} \\ 1 \\ \mathtt{n-k} \end{matrix}$ where T_k is tridiagonal,

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

$Q_k = \begin{matrix} \\ \mathtt{k} \\ \mathtt{n-k} \end{matrix} \begin{matrix} \begin{matrix} \mathtt{k} & \mathtt{n-k} \end{matrix} \\ \begin{bmatrix} I_k & \mathbf{0}^T \\ \mathbf{0} & H_K \end{bmatrix} \end{matrix}$ is such that

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

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

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

$>> [H, Q] = houshess(A)$

$H =$

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.