Numerical Linear Algebra
Lecture 10

Example

We will construct a lower triangular matrix using Given's rotation from the matrix

$$A = \begin{bmatrix} 5 & 4 & 3 \\ 4 & 6 & 1 \\ 3 & 1 & 7 \end{bmatrix}.$$

Given's matrix for $j < k$

```
function [G] = GivensMatrixLow(A, j,k)

$a = A(k,k)$
$b = A(j,k)$
$r = sqrt(a^2 + b^2);$
$c = a/r;$
$s = -b/r;$
$G = eye(length(A));$
$G(j,j) = c;$
$G(k,k) = c;$
$G(j,k) = s;$
$G(k,j) = -s;$
```

```
>>G1up = GivensMatrixLow(A,2,3)
```

$$G1 = \begin{bmatrix} 1.000000000000000 & 0 & 0 \\ 0 & 0.989949493661166 & -0.141421356237310 \\ 0 & 0.141421356237310 & 0.989949493661166 \end{bmatrix}$$

```
>> A1 =G1*A
```

$$A1 = \begin{bmatrix} 5.000000000000000 & 4.000000000000000 & 3.000000000000000 \\ 3.535533905932737 & 5.798275605729690 & -0.000000000000000 \\ 3.535533905932738 & 1.838477631085023 & 7.071067811865475 \end{bmatrix}$$

>>G2 = GivensMatrixLow(A1,1,3)

$$G2 = \begin{bmatrix} 0.920574617898323 & 0 & -0.390566732942472 \\ 0 & 1.000000000000000 & 0 \\ 0.390566732942472 & 0 & 0.920574617898323 \end{bmatrix}$$

>> A2=G2*A1

$$A2 = \begin{bmatrix} 3.222011162644131 & 2.964250269632601 & -0.000000000000000 \\ 3.535533905932737 & 5.798275605729690 & -0.000000000000000 \\ 5.207556439232954 & 3.254722774520597 & 7.681145747868607 \end{bmatrix}$$

>>G3 = GivensMatrixLow(A2,1,2)

$$G3 = \begin{bmatrix} 0.890391914715406 & -0.455194725594918 & 0 \\ 0.455194725594918 & 0.890391914715406 & 0 \\ 0 & 0 & 1.000000000000000 \end{bmatrix}$$

>> A3=G3*A2

$$A3 = \begin{bmatrix} 1.259496302198541 & 0 & -0.000000000000000 \\ 4.614653291088246 & 6.512048806713364 & -0.000000000000000 \\ 5.207556439232954 & 3.254722774520597 & 7.681145747868607 \end{bmatrix}$$

Example

We will construct an upper triangular matrix using Given's rotation from the matrix

$$A = \begin{bmatrix} 5 & 4 & 3 \\ 4 & 6 & 1 \\ 3 & 1 & 7 \end{bmatrix}.$$

```
» G1=GivensMatrixUpper(A,2,1)
G1 =
0.7809 0.6247 0
-0.6247 0.7809 0
0 0 1.0000
» A1=G1*A
A1 =
6.4031 6.8716 2.9673
0 2.1864 -1.0932
3.0000 1.0000 7.0000

» G2=GivensMatrixUpper(A1,3,1)
G2 =
0.9055 0 0.4243
0 1.0000 0
-0.4243 0 0.9055
```

```
» A2=G2*A1
A2 =
7.0711 6.6468 5.6569
0 2.1864 -1.0932
0.0000 -2.0099 5.0799
» G3=GivensMatrixUpper(A2,3,2)
G3 =
1.0000 0 0
0 0.7362 -0.6768
0 0.6768 0.7362
» A3=G3*A2
A3 =
7.0711 6.6468 5.6569
-0.0000 2.9698 -4.2426
0.0000 0.0000  3.0000
```

Example

Transform the given matrix A to a tridiagonal by a single transformation based on a Given's rotation.

$$\mathbf{A} = \begin{bmatrix} 7 & 4 & 3 \\ 4 & 5 & 2 \\ 3 & 2 & 2 \end{bmatrix}$$

To obtain the tridiagonal matrix using Given's rotation we have to zero out $(1,3),(3,1)$ elements of the matrix A.
To do that we compute elements of Givens matrix G and then compute GAG^T. Values of c, s are computed from the known $a = 4$ and $b = 3$ as

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

Example

We get formulas:

$$
\begin{aligned}
r &= \sqrt{a^2 + b^2} = \sqrt{4^2 + 3^2} = 5,\\
c &= \frac{a}{r} = 4/5 = 0.8,\\
s &= \frac{-b}{r} = -3/5 = -0.6.
\end{aligned}
$$

The Given's matrix will be

$$
\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{bmatrix}
$$

or

$$
\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.8 & 0.6 \\ 0 & -0.6 & 0.8 \end{bmatrix}
$$

Example

Then the tridiagonal matrix will be computed as

$$\mathbf{GAG^T} = \begin{bmatrix} 7 & 5 & 0 \\ 5 & 5.84 & -0.88 \\ 0 & -0.88 & 1.16 \end{bmatrix}$$

Another way to construct tridiagonal matrix is to compute the second Givens matrix G_2 and apply it for $A_1 = GA$ where

$$\mathbf{A_1} = \mathbf{GA} = \begin{bmatrix} 7 & 4 & 3 \\ 5 & 5.2 & 2.8 \\ 0 & -1.4 & 0.4 \end{bmatrix}$$

Example

Values of c, s are computed from the known $a = 2.8$ and $b = 3$ as

$$\begin{aligned} r &= \sqrt{a^2 + b^2} = \sqrt{2.8^2 + 3^2} \approx 4.1037, \\ c &= \frac{a}{r} = 4/5 = 0.6823, \\ s &= \frac{-b}{r} = -3/5 = -0.7311. \end{aligned}$$

The second Given's matrix will be

$$\mathbf{G_2} = \begin{bmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The tridiagonal matrix is computed as:

$$\mathbf{A_2} = \mathbf{G_2 A_1} = \mathbf{G_2 G A} = \begin{bmatrix} 1.121 & -1.0722 & 0 \\ 8.529 & 6.4723 & 4.1037 \\ 0 & -1.4 & 0.4 \end{bmatrix}$$

Example

We will construct tridiagonal matrix from the matrix

$$A = \begin{bmatrix} 5 & 4 & 3 \\ 3 & 6 & 1 \\ 4 & 1 & 7 \end{bmatrix}.$$

using Hauseholder transformations.

Example

To perform tridiagonal reduction for the matrix A we use Householder transformation in following steps:

- Choose $x = (3,4)^T$ and compute

$$u = x + \alpha e_1,$$

where $\alpha = -sign(3) \cdot ||x||, \|x\| = \sqrt{25} = 5$, and thus $\alpha = -5$.

- Construct $u = x + \alpha e_1 = (3,4)^T - (5,0)^T = (-2,4)^T$.
- Construct

$$v = \frac{u}{\|u\|}$$

with $\|u\| = \sqrt{20}$.
Therefore $v = (-2/\sqrt(20), 4/\sqrt(20))^T$.

- Compute

$$Q' = I - 2vv^T = \begin{pmatrix} 0.6 & 0.8 \\ 0.8 & -0.6 \end{pmatrix}.$$

- Construct the matrix of the Householder transformation as:

$$Q_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.6 & 0.8 \\ 0 & 0.8 & -0.6 \end{pmatrix}$$

- Then compute

$$A_1 = Q_1 A = \begin{pmatrix} 5 & 4 & 3 \\ 5 & 4.4 & 6.2 \\ 0 & 4.2 & -3.4 \end{pmatrix}.$$

such that Q_1 leaves the first row of $Q_1 A$ unchanged.

- Choose new vector $x = (4,3)^T$ for A_1^T and compute

$$u = x + \alpha e_1,$$

where $\alpha = -sign(4) \cdot ||x||, \|x\| = \sqrt{25} = 5$, and thus $\alpha = -5$.

- Construct $u = x + \alpha e_1 = (4,3)^T - (5,0)^T = (-1,3)^T$.
- Construct
$$v = \frac{u}{\|u\|}$$
with $\|u\| = \sqrt{10}$.
Therefore $v = (-1/\sqrt{10}, 3/\sqrt{10})^T$.
- Compute
$$V' = I - 2vv^T = \begin{pmatrix} 0.8 & 0.6 \\ 0.6 & -0.8 \end{pmatrix}.$$

- Construct the second matrix of the Householder transformation V_1 as:

$$V_1 = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & V' \end{array}\right]$$

to get

$$V_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.8 & 0.6 \\ 0 & 0.6 & -0.8 \end{pmatrix}$$

and then compute

$$Q_1 A V_1 = \begin{pmatrix} 5 & 5 & 0 \\ 5 & 7.24 & -2.32 \\ 0 & 1.32 & 5.24 \end{pmatrix}.$$

such that V_1 leaves the first column of A_1 unchanged.

Canonical Forms

DEFINITION. The polynomial $p(\lambda) = det(A - \lambda I)$ is called the *characteristic polynomial of* A. The roots of $p(\lambda) = 0$ are the eigenvalues of A.
Since the degree of the characteristic polynomial $p(\lambda)$ equals n, the dimension of A, it has n roots, so A has n eigenvalues.

DEFINITION. A nonzero vector x satisfying $Ax = \lambda x$ is a *(right) eigenvector* for the eigenvalue λ. A nonzero vector y such that $y^* A = \lambda y^*$ is a *left eigenvector*. (Recall that $y^* = (\bar{y})^T$ is the *conjugate transpose* of y.)

DEFINITION. Let S be any nonsingular matrix. Then A and $B = S^{-1}AS$ are called *similar* matrices, and S is a similarity transformation. PROPOSITION. Let $B = S^{-1}AS$, so A and B are similar. Then A and B have the same eigenvalues, and x (or y) is a right (or left) eigenvector of A if and only if $S^{-1}x$ (or S^*y) is a right (or left) eigenvector of B.

Proof. Using the fact that $\det(X \cdot Y) = \det(X) \cdot det(Y)$ for any square matrices X and Y, we can write

$$\det(A - \lambda I) = \det(S^{-1}(A - \lambda I)S) = \det(B - \lambda I).$$

So A and B have the same characteristic polynomials. $Ax = \lambda x$ holds if and only if $\underbrace{S^{-1}AS}_{B}\underbrace{S^{-1}x}_{x^*} = \lambda \underbrace{S^{-1}x}_{x^*}$ or $B(S^{-1}x) = \lambda(S^{-1}x)$. Similarly, $y^*A = \lambda y^*$ if and only if $y^*SS^{-1}AS = \lambda y^*S$ or $(S^*y)^*B = \lambda(S^*y)^*$. □

THEOREM. Jordan canonical form. Given A, there exists a nonsingular S such that $S^{-1}AS = J$, where J is in *Jordan canonical form*. This means that J is block diagonal, with $J = diag(J_{n_1}(\lambda_1), J_{n_2}(\lambda_2), \ldots, J_{n_k}(\lambda_k))$ and

$$J_{n_i}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{bmatrix}^{n_i \times n_i} .$$

J is unique, up to permutations of its diagonal blocks.

For a proof of this theorem, see a book on linear algebra such as [F. Gantmacher. The Theory of Matrices, vol. II (translation). Chelsea, New York, 1959] or [P. Halmos. Finite Dimensional Vector Spaces. Van Nostrand, New York, 1958].

- Each $J_m(\lambda)$ is called a *Jordan block* with eigenvalue λ of *algebraic multiplicity* m.
- If some $n_i = 1$, and λ_i is an eigenvalue of only that one Jordan block, then λ_i is called a *simple eigenvalue*.
- If all $n_i = 1$, so that J is diagonal, A is called *diagonalizable*; otherwise it is called *defective*.
- An n-by-n defective matrix does not have n eigenvectors. Although defective matrices are "rare" in a certain well-defined sense, the fact that some matrices do not have n eigenvectors is a fundamental fact confronting anyone designing algorithms to compute eigenvectors and eigenvalues.
- Symmetric matrices are never defective.

PROPOSITION.

- A Jordan block has one right eigenvector, $e_1 = [1, 0, \dots, 0]^T$, and one left eigenvector, $e_n = [0, \dots, 0, 1]^T$.
- Therefore, a matrix has n eigenvectors matching its n eigenvalues if and only if it is diagonalizable.
- In this case, $S^{-1}AS = diag(\lambda_i)$. This is equivalent to $AS = S\, diag(\lambda_i)$, so the i-th column of S is a right eigenvector for λ_i.
- It is also equivalent to $S^{-1}A = diag(\lambda_i)S^{-1}$, so the conjugate transpose of the ith row of S^{-1} is a left eigenvector for λ_i.
- If all n eigenvalues of a matrix A are distinct, then A is diagonalizable.

Proof. Let $J = J_m(\lambda)$ for ease of notation. It is easy to see $Je_1 = \lambda e_1$ and $e_n^T J = \lambda e_n^T$, so e_1 and e_n are right and left eigenvectors of J, respectively. To see that J has only one right eigenvector (up to scalar multiples), note that any eigenvector x must satisfy $(J - \lambda I)x = 0$, so x is in the null space of

$$J - \lambda I = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

But the null space of $J - \lambda I$ is clearly $\text{span}(e_1)$, so there is just one eigenvector. If all eigenvalues of A are distinct, then all its Jordan blocks must be 1-by-1, so $J = diag(\lambda_1, \ldots, \lambda_n)$ is diagonal. □

Example

We illustrate the concepts of eigenvalue and eigenvector with a problem of *mechanical vibrations*. We will see a defective matrix arise in a natural physical context.
Newton's law $F = ma$ applied to this system yields

$$\begin{aligned} m_i \ddot{x}_i(t) &= k_i(x_{i-1}(t) - x_i(t)) \\ &\quad \text{force on mass from spring } i \\ &\quad + k_{i+1}(x_{i+1}(t) - x_i(t)) \\ &\quad \text{force on mass from spring } i+1 \\ &\quad - b_i \dot{x}_i(t) \\ &\quad \text{force on mass from damper } i \end{aligned}$$

Example

or

$$M\ddot{x}(t) = -B\dot{x}(t) - Kx(t),$$

where $M = diag(m_1, \ldots, m_n)$, $B = diag(b_1, \ldots, b_n)$, and

$$K = \begin{bmatrix} k_1 + k_2 & -k_2 & & & \\ -k_2 & k_2 + k_3 & -k_3 & & \\ & \ddots & \ddots & \ddots & \\ & & -k_{n-1} & k_{n-1} + k_n & -k_n \\ & & & -k_n & k_n \end{bmatrix}.$$

We assume that all the masses m_i are positive. M is called the *mass matrix*, B is the *damping matrix*, and K is the *stiffness matrix*.

Example

Electrical engineers analyzing linear circuits arrive at an analogous equation by applying Kirchoff's and related laws instead of Newton's law. In this case x represents branch currents, M represent inductances, B represents resistances, and K represents admittances (reciprocal capacitances).

We will use a standard trick to change this second-order differential equation to a first-order differential equation, changing variables to

$$y(t) = \begin{bmatrix} \dot{x}(t) \\ x(t) \end{bmatrix}.$$

Example

This yields

$$\begin{aligned}\dot{y}(t) &= \begin{bmatrix} \ddot{x}(t) \\ \dot{x}(t) \end{bmatrix} = \begin{bmatrix} -M^{-1}B\dot{x}(t) - M^{-1}Kx(t) \\ \dot{x}(t) \end{bmatrix} \\ &= \begin{bmatrix} -M^{-1}B & -M^{-1}K \\ I & 0 \end{bmatrix} \cdot \begin{bmatrix} \dot{x}(t) \\ x(t) \end{bmatrix} \\ &= \begin{bmatrix} -M^{-1}B & -M^{-1}K \\ I & 0 \end{bmatrix} \cdot y(t) \equiv Ay(t).\end{aligned}$$

Example

- To solve $\dot{y}(t) = Ay(t)$, we assume that $y(0)$ is given (i.e., the initial positions $x(0)$ and velocities $\dot{x}(0)$ are given).
- One way to write down the solution of this differential equation is $y(t) = e^{At}y(0)$, where e^{At} is the matrix exponential. We will give another more elementary solution in the special case where A is diagonalizable; this will be true for almost all choices of m_i, k_i, and b_i.
- When A is diagonalizable, we can write $A = S\Lambda S^{-1}$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.
- Then $\dot{y}(t) = Ay(t)$ is equivalent to $\dot{y}(t) = S\Lambda S^{-1}y(t)$ or $S^{-1}\dot{y}(t) = \Lambda S^{-1}y(t)$ or $\dot{z}(t) = \Lambda z(t)$, where $z(t) \equiv S^{-1}y(t)$.
- This diagonal system of differential equations $\dot{z}_i(t) = \lambda_i z_i(t)$ has solutions $z_i(t) = e^{\lambda_i t}z_i(0)$. Since $z(t) \equiv S^{-1}y(t)$ and $z(0) \equiv S^{-1}y(0)$ then $Sz(t) \equiv y(t)$ and so $y(t) = S\mathrm{diag}(e^{\lambda_1 t}, \ldots, e^{\lambda_n t})S^{-1}y(0) = Se^{\Lambda t}S^{-1}y(0)$.

Disadvantages of Jordan form

First reason: It is a discontinuous function of A, so any rounding error can change it completely.

Example

Let

$$J_n(0) = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix},$$

which is in Jordan form. For arbitrarily small ϵ, adding $i \cdot \epsilon$ to the (i,i) entry changes the eigenvalues to the n distinct values $i \cdot \epsilon$, and so the Jordan form changes from $J_n(0)$ to $\mathrm{diag}(\epsilon, 2\epsilon, \ldots, n\epsilon)$. $\diamond$

Disadvantages of Jordan form

Second reason: It cannot be computed stably in general. In other words, when we have finished computing S and J, we cannot guarantee that $S^{-1}(A+\delta A)S = J$ for some small δA.

Example

Suppose $S^{-1}AS = J$ exactly, where S is very ill-conditioned. ($\kappa(S) = ||S|| \cdot ||S^{-1}||$ is very large.) Suppose that we are extremely lucky and manage to compute S *exactly* and J with just a tiny error δJ with $||\delta J|| = O(\varepsilon)||A||$. How big is the backward error? In other words, how big must δA be so that $S^{-1}(A+\delta A)S = J + \delta J$? We get

$$\underbrace{S^{-1}AS}_{J} + \underbrace{S^{-1}\delta AS}_{\delta J} = J + \delta J \rightarrow \delta A = S\delta JS^{-1},$$

and all that we can conclude is that $||\delta A|| \leq ||S|| \cdot ||\delta J|| \cdot ||S^{-1}|| = \kappa(S) \cdot ||\delta J|| = O(\varepsilon)\kappa(S)||A||$. Thus $||\delta A||$ may be much larger than $\varepsilon||A||$ because of $\kappa(S)$, which prevents backward stability. $\diamond$

From Jordan to Schur canonical form

- Instead of computing $S^{-1}AS = J$, where S can be an arbitrarily ill-conditioned matrix, we will restrict S to be orthogonal (so $\kappa_2(S) = 1$) to guarantee stability.
- We cannot get a canonical form as simple as the Jordan form any more.

THEOREM. *Schur canonical form.* Given A, there exists a unitary matrix Q ($Q^*Q = QQ^* = I$) and an upper triangular matrix T such that $Q^*AQ = T$. The eigenvalues of A are the diagonal entries of T.

Proof. We use induction on n. It is obviously true if A is 1 by 1. Now let λ be any eigenvalue and u a corresponding eigenvector normalized so $||u||_2 = 1$. Choose $\tilde{U}$ so $U = [u, \tilde{U}]$ is a square unitary matrix. (Note that λ and u may be complex even if A is real.) Then

$$U^* \cdot A \cdot U = \left[\begin{array}{c} u^* \\ \tilde{U}^* \end{array} \right] \cdot A \cdot [u, \tilde{U}] = \left[\begin{array}{cc} u^*Au & u^*A\tilde{U} \\ \tilde{U}^*Au & \tilde{U}^*A\tilde{U} \end{array} \right].$$

Now we can write $u^* A u = \lambda u^* u = \lambda$, and $\tilde{U}^* A u = \lambda \tilde{U}^* u = 0$ so $U^* A U \equiv \left[\begin{array}{cc} \lambda & \tilde{a}_{12} \\ 0 & \tilde{A}_{22} \end{array} \right]$. By induction, there is a unitary P, so $P^* \tilde{A}_{22} P = \tilde{T}$ is upper triangular. Then

$$U^* A U = \left[\begin{array}{cc} \lambda & \tilde{a}_{12} \\ 0 & P \tilde{T} P^* \end{array} \right] = \left[\begin{array}{cc} 1 & 0 \\ 0 & P \end{array} \right] \left[\begin{array}{cc} \lambda & \tilde{a}_{12} P \\ 0 & \tilde{T} \end{array} \right] \left[\begin{array}{cc} 1 & 0 \\ 0 & P^* \end{array} \right],$$

so

$$\underbrace{\left[\begin{array}{cc} 1 & 0 \\ 0 & P^* \end{array} \right] U^*}_{Q^*} A \underbrace{U \left[\begin{array}{cc} 1 & 0 \\ 0 & P \end{array} \right]}_{Q} = \left[\begin{array}{cc} \lambda & \tilde{a}_{12} P \\ 0 & \tilde{T} \end{array} \right] = T$$

is upper triangular and $Q = U[\begin{array}{cc} 1 & 0 \\ 0 & P \end{array}]$ is unitary as desired. $\diamond$

Remarks on Schur canonical form

- Notice that the Schur form is not unique, because the eigenvalues may appear on the diagonal of T in any order.
- This introduces complex numbers even when A is real. When A is real, we prefer a canonical form that uses only real numbers, because it will be cheaper to compute.
- This means that we will have to sacrifice a triangular canonical form and settle for a block-triangular canonical form.

Real Schur canonical form

THEOREM. *Real Schur canonical form.* If A is real, there exists a real orthogonal matrix V such that $V^TAV = T$ is quasi-upper triangular. This means that T is block upper triangular with 1-by-1 and 2-by-2 blocks on the diagonal. Its eigenvalues are the eigenvalues of its diagonal blocks. The 1-by-1 blocks correspond to real eigenvalues, and the 2-by-2 blocks to complex conjugate pairs of eigenvalues.

Computing Eigenvectors from the Schur Form

Let $Q^*AQ = T$ be the Schur form. Then if $Tx = \lambda x$, we have

$$Tx = \lambda x \rightarrow (Q^*AQ)x = Tx = \lambda x \rightarrow A{\color{red}Qx} = QTx = \lambda {\color{red}Qx}$$

, so Qx is an eigenvector of A. So to find eigenvectors Qx of A, it suffices to find eigenvectors x of T.
Suppose that $\lambda = t_{ii}$ has multiplicity 1 (i.e., it is simple). Write $(T - \lambda I)x = 0$ as

$$\begin{aligned} 0 &= \begin{bmatrix} T_{11} - \lambda I & T_{12} & T_{13} \\ 0 & 0 & T_{23} \\ 0 & 0 & T_{33} - \lambda I \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\ &= \begin{bmatrix} (T_{11} - \lambda I)x_1 + T_{12}x_2 + T_{13}x_3 \\ T_{23}x_3 \\ (T_{33} - \lambda I)x_3 \end{bmatrix}, \end{aligned}$$

where T_{11} is $(i-1)$-by-$(i-1)$, $T_{22} = \lambda$ is 1-by-1, T_{33} is $(n-i)$-by-$(n-i)$, and x is partitioned correspondingly.

Since λ is simple, both $T_{11} - \lambda I$ and $T_{33} - \lambda I$ are nonsingular, so $(T_{33} - \lambda I)x_3 = 0$ implies $x_3 = 0$. Therefore $(T_{11} - \lambda I)x_1 = -T_{12}x_2$. Choosing (arbitrary) $x_2 = 1$ means $x_1 = -(T_{11} - \lambda I)^{-1} T_{12}$, so

$$x = \begin{bmatrix} (\lambda I - T_{11})^{-1} T_{12} \\ 1 \\ 0 \end{bmatrix},$$

Condition number of the eigenvalue λ

THEOREM. Let λ be a simple eigenvalue of A with right eigenvector x and left eigenvector y, normalized so that $||x||_2 = ||y||_2 = 1$. Let $\lambda + \delta\lambda$ be the corresponding eigenvalue of $A + \delta A$. Then

$$\begin{aligned} \delta\lambda &= \frac{y^* \delta A x}{y^* x} + O(||\delta A||^2) \text{ or} \\ |\delta\lambda| &\leq \frac{||\delta A||}{|y^* x|} + O(||\delta A||^2) = \sec\Theta(y,x)||\delta A|| + O(||\delta A||^2), \end{aligned}$$

where $\Theta(y,x)$ is the acute angle between y and x. In other words,$\sec\Theta(y,x) = 1/|y^* x|$ is the condition number of the eigenvalue λ.

Proof. Subtract $Ax = \lambda x$ from $(A + \delta A)(x + \delta x) = (\lambda + \delta\lambda)(x + \delta x)$ to get

$$A\delta x + \delta A x + \underbrace{\delta A \delta x}_{0} = \lambda\delta x + \delta\lambda x + \underbrace{\delta\lambda\delta x}_{0}.$$

Ignore the second-order terms (those with two "δ terms" as factors: $\delta A \delta x$ and $\delta\lambda\delta x$) and multiply by y^* to get

$$\underbrace{y^* A\delta x}_{cancels} + y^* \delta A x = \underbrace{y^* \lambda\delta x}_{cancels} + y^* \delta\lambda x.$$

Now $y^* A\delta x$ cancels $y^* \lambda\delta x$, so we can solve for $\delta\lambda = (y^* \delta A x)/(y^* x)$ as desired. □

COROLLARY. Let A be symmetric (or normal: $AA^* = A^*A$). Then $|\delta\lambda| \leq ||\delta A|| + O(||\delta A||^2)$.

Proof. If A is symmetric or normal, then its eigenvectors are all orthogonal, i.e., $Q^*AQ = \Lambda$ with $QQ^* = I$. So the right eigenvectors x (columns of Q) and left eigenvectors y (conjugate transposes of the rows of Q^*) are identical, and $1/|y^*x| = 1$. □

Bauer-Fike Theorem.

THEOREM. Bauer-Fike. Let A have all simple eigenvalues (i.e., be diagonalizable). Call them λ_i, with right and left eigenvectors x_i and y_i, normalized so $||x_i||_2 = ||y_i||_2 = 1$. Then the eigenvalues of $A + \delta A$ lie in disks B_i, where B_i has center λ_i and radius $n\frac{||\delta A||_2}{|y_i^* x_i|}$.
Our proof will use Gershgorin's theorem.

GERSHGORIN'S THEOREM

GERSHGORIN'S THEOREM. Let B be an arbitrary matrix. Then the eigenvalues λ of B are located in the union of the n disks defined by $|\lambda - b_{ii}| \leq \sum_{j \neq i} |b_{ij}|$ for $i = 1$ to n.
We will also need two lemmas which we present on the next slides.

LEMMA. Let $S = [x_1, \ldots, x_n]$ the nonsingular matrix of right eigenvectors. Then

$$S^{-1} = \begin{bmatrix} y_1^* / y_1^* x_1 \\ y_2^* / y_2^* x_2 \\ \vdots \\ y_n^* / y_n^* x_n \end{bmatrix}.$$

Proof of Lemma. We know that A is diagonalizible and thus $S^{-1}AS = \Lambda$, or $AS = S\Lambda$, where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, since the columns x_i of S are eigenvectors. This is equivalent to $S^{-1}A\underbrace{SS^{-1}}_{I} = \Lambda S^{-1}$ or

$S^{-1}A = \Lambda S^{-1}$, so the rows of S^{-1} are conjugate transposes of the left eigenvectors y_i. So

$$S^{-1} = \begin{bmatrix} y_1^* \cdot c_1 \\ \vdots \\ y_n^* \cdot c_n \end{bmatrix}$$

for some constants c_i. But $I = S^{-1}S$, so $1 = (S^{-1}S)_{ii} = y_i^* x_i \cdot c_i$, and $c_i = \frac{1}{y_i^* x_i}$ as desired. □

LEMMA. If each column of (any matrix) S has two-norm equal to 1, $||S||_2 \leq \sqrt{n}$. Similarly, if each row of a matrix has two-norm equal to 1, its two-norm is at most $\sqrt{n}$.

Proof of Lemma. $||S||_2 = ||S^T||_2 = \max_{||x||_2=1} ||S^T x||_2$. Each component of $S^T x$ is bounded by 1 by the Cauchy-Schwartz inequality, so $||S^T x||_2 \leq ||[\underbrace{1, \ldots, 1}_{n}]^T||_2 = \sqrt{n}$. $\square$

Proof of the Bauer-Fike theorem.
We will apply Gershgorin's theorem to $S^{-1}(A+\delta A)S = \Lambda + F$, where $\Lambda = S^{-1}AS = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ and $F = S^{-1}\delta AS$. The idea is to show that the eigenvalues of $A + \delta A$ lie in balls centered at the λ_i with the given radii. To do this, we take the disks containing the eigenvalues of $\Lambda + F$ that are defined by Gershgorin's theorem,

$$|\lambda - (\lambda_i + f_{ii})| \leq \sum_{j \neq i}^{n} |f_{ij}|,$$

and enlarge them slightly to get the disks

$$\begin{aligned}
|\lambda - \lambda_i| &\leq \sum_j^n 1 \cdot |f_{ij}| \leq (\sum_j^n 1)^{1/2} \left(\sum_j^n |f_{ij}|^2 \right)^{1/2} \quad \text{by Cauchy} - \text{Schwarz} \\
&\leq n^{1/2} \left(\sum_j^n |f_{ij}|^2 \right)^{1/2} \quad \text{by Cauchy} - \text{Schwarz} \\
&= n^{1/2} \cdot ||F(i,:)||_2.
\end{aligned}$$

Now we need to bound the two-norm of the i-th row $F(i,:)$ of $F = S^{-1}\delta A S$:

$$\begin{aligned}
||F(i,:)||_2 &= ||(S^{-1}\delta A S)(i,:)||_2 \\
&\leq ||(S^{-1})(i,:)||_2 \cdot ||\delta A||_2 \cdot \underbrace{||S||_2}_{\leq\sqrt{n}} \\
&\leq \frac{n^{1/2}}{|y_i^* x_i|} \cdot ||\delta A||_2 \quad \text{by Lemmas.}
\end{aligned}$$

Combined this equation with equation above, this proves the theorem.

$$\begin{aligned}
|\lambda - \underbrace{\lambda_i}_{centers}| &\leq n^{1/2} \cdot ||F(i,:)||_2 \\
&\leq n^{1/2} \cdot \frac{n^{1/2}}{|y_i^* x_i|} \cdot ||\delta A||_2 = \underbrace{\frac{n}{|y_i^* x_i|} \cdot ||\delta A||_2}_{radius}
\end{aligned}$$

□

THEOREM. Let λ be a simple eigenvalue of A, with unit right and left eigenvectors x and y and condition number $c = 1/|y^* x|$. Then there is a δA such that $A + \delta A$ has a multiple eigenvalue at λ, and

$$\frac{||\delta A||_2}{||A||_2} \leq \frac{1}{\sqrt{c^2 - 1}}.$$

When $c \gg 1$, i.e., the eigenvalue is ill-conditioned, then the upper bound on the distance is $1/\sqrt{c^2 - 1} \approx \frac{1}{c}$, the reciprocal of the condition number.

We assume that A is real.

- Power method

  This method can find only the largest eigenvalue for A and the corresponding eigenvector.

- Inverse iteration

  We find all other eigenvalues and eigenvectors applying method for $(A - \sigma I)^{-1}$ for some shift σ.

- Orthogonal iteration

  Lets compute entire invariant subspace.

- QR iteration

  reorganized orthogonal iteration, ultimate algorithm.

- Hessenberg reduction

- Tridiagonal and bidiagonal reduction

Power Method

ALGORITHM. Power method: Given x_0, we iterate

```
$i = 0$
repeat
    $y_{i+1} = Ax_i$
    $x_{i+1} = y_{i+1}/||y_{i+1}||_2$    (approximate eigenvector)
    $\tilde{\lambda}_{i+1} = x_{i+1}^T A x_{i+1}$    (approximate eigenvalue)
    $i = i + 1$
until convergence
```

Let us first apply this algorithm in the very simple case when $A = \mathrm{diag}(\lambda_1, ..., \lambda_n)$, with $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|)$. In this case the eigenvectors are just the columns e_i of the identity matrix. Note that x_i can also be written $x_i = A^i x_0 / ||A^i x_0||_2$, since the factors $1/||y_{i+1}||_2$ only scale x_{i+1} to be a unit vector and do not change its direction. Taking $S = I$ lets us write $x_0 = S(S^{-1}x_0) = [\xi_1, \ldots, \xi_n]^T$, or

$$A^i x_0 \equiv A^i \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \begin{bmatrix} \xi_1 \lambda_1^i \\ \xi_2 \lambda_2^i \\ \vdots \\ \xi_n \lambda_n^i \end{bmatrix} = \xi_1 \lambda_1^i \begin{bmatrix} 1 \\ \frac{\xi_2}{\xi_1}(\frac{\lambda_2}{\lambda_1})^i \\ \vdots \\ \frac{\xi_n}{\xi_1}(\frac{\lambda_n}{\lambda_1})^i \end{bmatrix},$$

where we have assumed $\xi_1 \neq 0$. Since all the fractions λ_j/λ_1 are less than 1 in absolute value, $A^i x_0$ becomes more and more nearly parallel to e_1, so $x_i = A^i x_0 / ||A^i x_0||_2$ becomes closer and closer to $\pm e_i$, the eigenvector corresponding to the largest eigenvalue λ_1. The rate of convergence depends on how much smaller than 1 the ratios $|\lambda_2/\lambda_1| \geq \cdots \geq |\lambda_n/\lambda_1|)$ are, the smaller the faster. Since x_i converges to $\pm e_1$, $\tilde{\lambda}_i = x_i^T A x_i$ converges to λ_1, the largest eigenvalue.

Assumptions on power method

- In the simplest case we applied algorithm when A is diagonal.
- Consider now general case: assume that $A = S\Lambda S^{-1}$ is diagonalizable, with $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and the eigenvalues sorted so that $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$. Write $S = [s_1, \ldots, s_n]$, where the columns s_i are the corresponding eigenvectors and also satisfy $||s_i||_2 = 1$; in the previous slide we had $S = I$. This lets us write $\mathbf{x_0} = \mathbf{S}(\mathbf{S}^{-1}\mathbf{x_0}) \equiv \mathbf{S}([\xi_1, \ldots, \xi_\mathbf{n}]^\mathbf{T})$. Also, since $A = S\Lambda S^{-1}$, we can write

$$A^i = \underbrace{(S\Lambda S^{-1}) \cdots (S\Lambda S^{-1})}_{i \text{ times}} = S\Lambda^i S^{-1}$$

since all the $S^{-1} \cdot S$ pairs cancel.

This finally lets us write

$$A^i x_0 = \underbrace{S \Lambda^i S^{-1}}_{A^i} \cdot \underbrace{S \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}}_{x_0} = S \begin{bmatrix} \xi_1 \lambda_1^i \\ \xi_2 \lambda_2^i \\ \vdots \\ \xi_n \lambda_n^i \end{bmatrix} = \xi_1 \lambda_1^i S \begin{bmatrix} 1 \\ \frac{\xi_2}{\xi_1} (\frac{\lambda_2}{\lambda_1})^i \\ \vdots \\ \frac{\xi_n}{\xi_1} (\frac{\lambda_n}{\lambda_1})^i \end{bmatrix} .$$

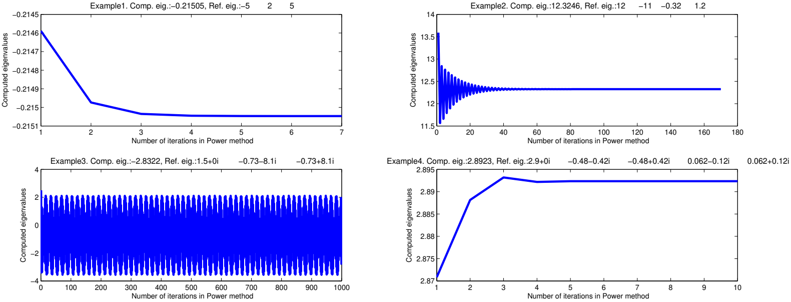
As before, the vector j in brackets converges to e_1, so $A^i x_0$ gets closer and closer to a multiple of $Se_1 = s_1$, the eigenvector corresponding to λ_1. Therefore, $\tilde{\lambda}_i = x_i^T A x_i$ converges to $s_1^T A s_1 = s_1^T \lambda_1 s_1 = \lambda_1$.

- A minor drawback of this method is the assumption that $\xi_1 \neq 0$, i.e., that x_0 is not the invariant subspace $\mathrm{span}\{s_2, \dots, s_n\}$ this is true with very high probability if x_0 is chosen at random.

- A major drawback is that it **converges to the eigenvalue/eigenvector pair only for the eigenvalue of largest absolute magnitude**, and its convergence rate depends on $|\lambda_2/\lambda_1|$, a quantity which may be close to 1 and thus cause very slow convergence. Indeed, if A is real and the largest eigenvalue is complex, there are two complex conjugate eigenvalues of largest absolute value $|\lambda_1| = |\lambda_2|$, and so the above analysis does not work at all. In the extreme case of an orthogonal matrix, all the eigenvalues have the same absolute value, namely, 1.

- To plot the convergence of the power method, see HOMEPAGE/Matlab/powerplot.m.

Examples of running of Power method in Matlab



Results are described in the next two slides.

- Example 1. In this example we test the matrix

$$A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -5 \end{bmatrix}$$

with exact eigenvalues $(5, 2, -5)$. The Power method can converge as to the exact first eigenvalue 5 as well as to the completely erroneous eigenvalue, see Figure. This is because two eigenvalues of this matrix, 5 and -5, have the same absolute values: $|5| = |-5|$. Thus, assumption 2 about convergence of the Power method is not fulfilled.

- Example 2. In this example the matrix A is given by

$$A = \begin{bmatrix} 3 & 7 & 8 & 9 \\ 5 & -7 & 4 & -7 \\ 1 & -1 & 1 & -1 \\ 9 & 3 & 2 & 5 \end{bmatrix}$$

This matrix has four different reference real eigenvalues $\lambda = (\lambda_1, ..., \lambda_4)$ given by

$\lambda = (12.3246, -11.1644, -0.3246, 1.1644)$.

Now all assumptions about matrix A are fulfilled and the Power method has converged to the reference eigenvalue 12.3246, see Figure. We note that reference eigenvalues on this figure are rounded.

- Example 3. Now we take the matrix

$$A = \begin{bmatrix} 0 & 5 & -2 \\ 6 & 0 & -12 \\ 1 & 3 & 0 \end{bmatrix}$$

  with one real and two complex eigenvalues:

  $\lambda = (1.4522, -0.7261 + 8.0982i, -0.7261 - 8.0982i)$.

  We observe at Figure that Power method does not converge in this case again. This is because assumption 4 on the convergence of the power method is not fulfilled.

- Example 4. In this example the matrix A has size 5×5. Elements of this matrix are uniformly distributed pseudorandom numbers on the open interval $(0, 1)$. Using the Figure we observe that in the first round of our computations we have obtained good approximation to the eigenvalue 2.9 thought not all assumptions about Power method are fulfilled. This means that on the second round of computations we can get completely different erroneous eigenvalue. This example is similar to the example 1 where convergence was not achieved.

Inverse Iteration

We will overcome the drawbacks of the power method just described by applying the power method to $(A - \sigma I)^{-1}$ instead of A, where σ is called a *shift*. This will let us converge to the eigenvalue closest to σ, rather than just λ_1. This method is called *inverse iteration* or the *inverse power method*.

ALGORITHM. Inverse iteration: Given x_0, we iterate

```
$i = 0$
repeat
    $y_{i+1} = (A - \sigma I)^{-1} x_i$
    $x_{i+1} = y_{i+1}/||y_{i+1}||_2$      (approximate eigenvector)
    $\tilde{\lambda}_{i+1} = x_{i+1}^T A x_{i+1}$      (approximate eigenvalue)
    $i = i + 1$
until convergence
```