Applied Mathematics 205

Unit V: Eigenvalue Problems

Lecturer: Dr. David Knezevic

Unit V: Eigenvalue Problems

Chapter V.2: Fundamentals

Eigenvalues and Eigenvectors

Eigenvalues and eigenvectors of real-valued matrices can be complex

Hence in this Unit we will generally work with complex-valued matrices and vectors, $A \in \mathbb{C}^{n \times n}$, $v \in \mathbb{C}^n$

For $A \in \mathbb{C}^{n \times n}$, we shall consider the eigenvalue problem: find $(\lambda, v) \in \mathbb{C} \times \mathbb{C}^n$ such that

$$\begin{aligned} Av &= \lambda v, \\ \|v\|_2 &= 1 \end{aligned}$$

Note that for $v \in \mathbb{C}^n$, $\|v\|_2 \equiv \left(\sum_{k=1}^n |v_k|^2\right)^{1/2}$, where $|\cdot|$ is the modulus of a complex number

Eigenvalues and Eigenvectors

This problem can be reformulated as:

$$(A - \lambda \mathrm{I})v = 0$$

We know this system has a solution if and only if $(A - \lambda \mathrm{I})$ is singular, hence we must have

$$\det(A - \lambda \mathrm{I}) = 0$$

$p(z) \equiv \det(A - z\mathrm{I})$ is a degree n polynomial, called the characteristic polynomial of A

The eigenvalues of A are exactly the roots of the characteristic polynomial

Characteristic Polynomial

By the fundamental theorem of algebra, we can factorize $p(z)$ as

$$p(z) = c_n(z - \lambda_1)(z - \lambda_2)\cdots(z - \lambda_n),$$

where the roots $\lambda_i \in \mathbb{C}$ need not be distinct

Note also that complex eigenvalues of a matrix $A \in \mathbb{R}^{n\times n}$ must occur as complex conjugate pairs

That is, if $\lambda = \alpha + i\beta$ is an eigenvalue, then so is its complex conjugate $\overline{\lambda} = \alpha - i\beta$

Characteristic Polynomial

This follows from the fact that for a polynomial p with real coefficients, $p(\overline{z}) = \overline{p(z)}$ for any $z \in \mathbb{C}$:

$$p(\overline{z}) = \sum_{k=0}^{n} c_k (\overline{z})^k = \sum_{k=0}^{n} c_k \overline{z^k} = \overline{\sum_{k=0}^{n} c_k z^k} = \overline{p(z)}$$

Hence if $w \in \mathbb{C}$ is a root of p, then so is $\overline{w}$, since

$$0 = p(w) = \overline{p(w)} = p(\overline{w})$$

Companion Matrix

We have seen that every matrix has an associated characteristic polynomial

Similarly, every polynomial has an associated companion matrix

The companion matrix, C_n, of $p \in \mathbb{P}_n$ is a matrix which has eigenvalues that match the roots of p

Divide p by its leading coefficient to get a monic polynomial, i.e. with leading coefficient equal to 1 (this doesn't change the roots)

$$p_{\text{monic}}(z) = c_0 + c_1 z + \cdots + c_{n-1} z^{n-1} + z^n$$

Companion Matrix

Then p_{monic} is the characteristic polynomial of the $n \times n$ companion matrix

$$C_n = \begin{bmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{bmatrix}$$

Companion Matrix

We show this for the $n = 3$ case: Consider

$$p_{3,\mathrm{monic}}(z) \equiv c_0 + c_1 z + c_2 z^2 + z^3,$$

which has companion matrix

$$C_3 \equiv \begin{bmatrix} 0 & 0 & -c_0 \\ 1 & 0 & -c_1 \\ 0 & 1 & -c_2 \end{bmatrix}$$

Recall that for a 3×3 matrix, we have

$$\begin{aligned} \det\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\right) &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \end{aligned}$$

Companion Matrix

Substituting entries of C_3 then gives

$$\det(z\mathrm{I} - C_3) = c_0 + c_1 z + c_2 z^2 + z^3 = p_{3,\mathrm{monic}}(z)$$

This link between matrices and polynomials is used by Matlab's `roots` function

`roots` computes all roots of a polynomial by using algorithms considered in this Unit to find eigenvalues of the companion matrix

Eigenvalue Decomposition

Let λ be an eigenvalue of $A \in \mathbb{C}^{n \times n}$; the set of all eigenvalues is called the spectrum of A

The algebraic multiplicity of λ is the multiplicity of the corresponding root of the characteristic polynomial

The geometric multiplicity of λ is the number of linearly independent eigenvectors corresponding to λ

For example, for $A = \mathrm{I}$, $\lambda = 1$ is an eigenvalue with algebraic and geometric multiplicity of n

(Char. poly. for $A = \mathrm{I}$ is $p(z) = (z - 1)^n$, and $e_i \in \mathbb{C}^n$, $i = 1, 2, \ldots, n$ are eigenvectors)

Eigenvalue Decomposition

Theorem: The geometric multiplicity of an eigenvalue is less than or equal to its algebraic multiplicity

If λ has geometric multiplicity $<$ algebraic multiplicity, then λ is said to be defective

We say a matrix is defective if it has at least one defective eigenvalue

Eigenvalue Decomposition

For example,

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

has one eigenvalue with algebraic multiplicity of 3, geometric multiplicity of 1

```
>> A = [2 1 0; 0 2 1; 0 0 2];
>> [V,D] = eig(A)

V =
   1.000000000000000  -1.000000000000000   1.000000000000000
                   0   0.000000000000000  -0.000000000000000
                   0                   0   0.000000000000000
D =
     2     0     0
     0     2     0
     0     0     2
```

Eigenvalue Decomposition

Let $A \in \mathbb{C}^{n \times n}$ be a nondefective matrix, then it has a full set of n linearly independent eigenvectors $v_1, v_2, \ldots, v_n \in \mathbb{C}^n$

Suppose $V \in \mathbb{C}^{n \times n}$ contains the eigenvectors of A as columns, and let $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$

Then $Av_i = \lambda_i v_i$, $i = 1, 2, \ldots, n$ is equivalent to $AV = VD$

Since we assumed A is nondefective, we can invert V to obtain

$$A = VDV^{-1}$$

This is the eigendecomposition of A

This shows that for a non-defective matrix, A is diagonalized by V

Eigenvalue Decomposition

We introduce the conjugate transpose, $A^* \in \mathbb{C}^{n\times m}$, of a matrix $A \in \mathbb{C}^{m\times n}$

$$(A^*)_{ij} = \overline{A_{ji}}, \quad i = 1, 2, \ldots, m, \ j = 1, 2, \ldots, n$$

A matrix is said to be hermitian if $A = A^*$ (this generalizes matrix symmetry)

A matrix is said to be unitary if $AA^* = \mathrm{I}$ (this generalizes the concept of an orthogonal matrix)

Also, for $v \in \mathbb{C}^n$, $\|v\|_2 = \sqrt{v^* v}$

Eigenvalue Decomposition

The ' operator in Matlab actually performs conjugate transpose:

```
>> [1+i i; -2*i 3]'

ans =

   1.0000 - 1.0000i        0 + 2.0000i
        0 - 1.0000i   3.0000
```

To get the raw transpose, do .':

```
>> [1 i; -2*i 3].'

   1.0000 + 1.0000i        0 - 2.0000i
        0 + 1.0000i   3.0000
```

Eigenvalue Decomposition

In some cases, the eigenvectors of A can be chosen such that they are orthonormal

$$v_i^* v_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

In such a case, the matrix of eigenvectors, Q, is unitary, and hence A can be unitarily diagonalized

$$A = QDQ^*$$

Eigenvalue Decomposition

Theorem: A hermitian matrix is unitarily diagonalizable, and its eigenvalues are real

But hermitian matrices are not the only matrices that can be unitarily diagonalized... $A \in \mathbb{C}^{n \times n}$ is normal if

$$A^*A = AA^*$$

Theorem: A matrix is unitarily diagonalizable if and only if it is normal

Gershgorin's Theorem

Due to the link between eigenvalues and polynomial roots, in general one has to use iterative methods to compute eigenvalues

However, it is possible to gain some information about eigenvalue locations more easily from Gershgorin's Theorem

Let $D(c, r) \equiv \{x \in \mathbb{C} : |x - c| \leq r\}$ denote a disk in the complex plane centered at c with radius r

For a matrix $A \in \mathbb{C}^{n \times n}$, $D(a_{ii}, R_i)$ is called a Gershgorin disk, where

$$R_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|,$$

Gershgorin's Theorem

Theorem: All eigenvalues of $A \in \mathbb{C}^{n \times n}$ are contained within the union of the n Gershgorin disks of A

Proof: See lecture

Gershgorin's Theorem

Note that a matrix is diagonally dominant if

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \quad \text{for } i = 1, 2, \ldots, n$$

It follows from Gershgorin's Theorem that a diagonally dominant matrix cannot have a zero eigenvalue, hence must be invertible

For example, the finite difference discretization matrix of the differential operator $-\Delta + \mathrm{I}$ is diagonally dominant

In 2-dimensions, $(-\Delta + \mathrm{I})u = -u_{xx} - u_{yy} + u$

Each row of the corresponding discretization matrix contains diagonal entry $4/h + 1$, and four off-diagonal entries of $-1/h$

Sensitivity of Eigenvalue Problems

We shall now consider the sensitivity of the eigenvalues to perturbations in the matrix A

Suppose A is nondefective, and hence $A = VDV^{-1}$

Let δA denote a perturbation of A, and let $E \equiv V^{-1}\delta AV$, then

$$V^{-1}(A + \delta A)V = V^{-1}AV + V^{-1}\delta AV = D + E$$

Sensitivity of Eigenvalue Problems

For a nonsingular matrix X, the map $A \to X^{-1}AX$ is called a similarity transformation of A

Theorem: A similarity transformation preserves eigenvalues

Proof: We can equate the characteristic polynomials of A and $X^{-1}AX$ (denoted $p_A(z)$ and $p_{X^{-1}AX}(z)$, respectively) as follows:

$$\begin{aligned} p_{X^{-1}AX}(z) &= \det(z\mathrm{I} - X^{-1}AX) \\ &= \det(X^{-1}(z\mathrm{I} - A)X) \\ &= \det(X^{-1})\det(z\mathrm{I} - A)\det(X) \\ &= \det(z\mathrm{I} - A) \\ &= p_A(z), \end{aligned}$$

where we have used the identities $\det(AB) = \det(A)\det(B)$, and $\det(X^{-1}) = 1/\det(X)$ $\square$

Sensitivity of Eigenvalue Problems

The identity $V^{-1}(A + \delta A)V = D + E$ is a similarity transformation

Therefore $A + \delta A$ and $D + E$ have the same eigenvalues

Let λ_k, $k = 1, 2, \ldots, n$ denote the eigenvalues of A, and $\tilde{\lambda}$ denote an eigenvalue of $A + \delta A$

Then for some $w \in \mathbb{C}^n$, $(\tilde{\lambda}, w)$ is an eigenpair of $(D + E)$, i.e.

$$(D + E)w = \tilde{\lambda}w$$

Sensitivity of Eigenvalue Problems

This can be rewritten as

$$w = (\tilde{\lambda}\mathrm{I} - D)^{-1} E w$$

This is a promising start because:

- we want to bound $|\tilde{\lambda} - \lambda_k|$ for some k
- $(\tilde{\lambda}\mathrm{I} - D)^{-1}$ is a diagonal matrix with entries $1/(\tilde{\lambda} - \lambda_k)$ on the diagonal

Sensitivity of Eigenvalue Problems

Taking norms yields

$$\|w\|_2 \leq \|(\tilde{\lambda}\mathrm{I} - D)^{-1}\|_2 \|E\|_2 \|w\|_2,$$

or

$$\|(\tilde{\lambda}\mathrm{I} - D)^{-1}\|_2^{-1} \leq \|E\|_2$$

Note that the norm of a diagonal matrix is given by its largest entry (in abs. val.)[1]

$$\begin{aligned} \max_{v \neq 0} \frac{\|Dv\|}{\|v\|} &= \max_{v \neq 0} \frac{\|(D_{11}v_1, D_{22}v_2, \ldots, D_{nn}v_n)\|}{\|v\|} \\ &\leq \left\{ \max_{i=1,2,\ldots,n} |D_{ii}| \right\} \max_{v \neq 0} \frac{\|v\|}{\|v\|} \\ &= \max_{i=1,2,\ldots,n} |D_{ii}| \end{aligned}$$

[1] This holds for any induced matrix norm, not just the 2-norm

Sensitivity of Eigenvalue Problems

Hence $\|(\tilde{\lambda}\mathrm{I} - D)^{-1}\|_2 = 1/|\tilde{\lambda} - \lambda_{k^*}|$, where λ_{k^*} is the eigenvalue of A closest to $\tilde{\lambda}$

Therefore it follows from $\|(\tilde{\lambda}\mathrm{I} - D)^{-1}\|_2^{-1} \leq \|E\|_2$ that

$$
\begin{aligned}
|\tilde{\lambda} - \lambda_{k^*}| &= \|(\tilde{\lambda}\mathrm{I} - D)^{-1}\|_2^{-1} \\
&\leq \|E\|_2 \\
&= \|V^{-1}\delta A V\|_2 \\
&\leq \|V^{-1}\|_2 \|\delta A\|_2 \|V\|_2 \\
&= \mathrm{cond}(V)\|\delta A\|_2
\end{aligned}
$$

This result is known as the Bauer-Fike Theorem

Sensitivity of Eigenvalue Problems

Hence suppose we compute the eigenvalues, $\tilde{\lambda}_i$, of the perturbed matrix $A + \delta A$

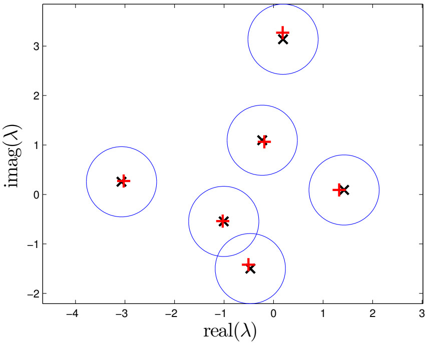
Then Bauer-Fike tells us that each $\tilde{\lambda}_i$ must reside in a disk of radius $\mathrm{cond}(V)\|\delta A\|_2$ centered on some eigenvalue of A

If V is poorly conditioned, then even for small perturbations δA, the disks can be large: sensitivity to perturbations

If A is normal then $\mathrm{cond}(V) = 1$, in which case the Bauer-Fike disk radius is just $\|\delta A\|_2$

Sensitivity of Eigenvalue Problems

Matlab example: Eigenvalue perturbation and Bauer-Fike Theorem



Sensitivity of Eigenvalue Problems

Note that a limitation of Bauer-Fike is that it does not tell us which disk $\tilde{\lambda}_i$ will reside in

Therefore, this doesn't rule out the possibility of, say, all $\tilde{\lambda}_i$ clustering in just one Bauer-Fike disk

In the case that A and $A + \delta A$ are hermitian, we have a stronger result

Sensitivity of Eigenvalue Problems

Weyl's Theorem: Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \cdots \leq \tilde{\lambda}_n$ be the eigenvalues of hermitian matrices A and $A + \delta A$, respectively. Then $\max_{i=1,\ldots,n} |\lambda_i - \tilde{\lambda}_i| \leq \|\delta A\|_2$.

Hence in the hermitian case, each perturbed eigenvalue must be in the disk[2] of its corresponding unperturbed eigenvalue!

[2] In fact, eigenvalues of a hermitian matrix are real, so disk here is actually an interval in $\mathbb{R}$