TANA15/Lecture 4 - Contents

The Eigenvalue Decomposition

- Definitions and Basics.
- Localization of Eigenvalues. Sensitivity.
- Applications: Roots of Polynomials. Functions of Matrices.

Computing Eigenvalues

- Rayleigh Quotient.
- The Power iteration. Inverse Iteration.

Definition and Basics

Definition Let $A \in \mathbb{C}^{n \times n}$. If there is a scalar $\lambda \in \mathbb{C}$ and a vector $x \neq 0$ in $\mathbb{C}^n$ such that

$$Ax = \lambda x$$

then λ is an *eigenvalue* and x is an *eigenvector*.

Remark Real matrices can have complex eigenvalues. It is easier to treat the complex case.

Lemma If λ is an eigenvalue then $A - \lambda I$ is *singular*. This means that $\mathrm{Null}(A - \lambda I) \neq \{0\}$.

Remark Eigenvectors are never unique. If x is an eigenvector so is αx. Only the subspaces $\mathrm{Null}(A - \lambda I)$ are unique.

Definition Let $X = (x_1, x_2, \ldots, x_k)$ be all the linearly independent eigenvectors associated with A and $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k)$. Then $AX = XD$.

Definition If $A \in \mathbb{C}^{n \times n}$ has a full set of linearly independent eigenvectors so $X = (x_1, \ldots, x_n)$. Then

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

is called the *eigenvalue decomposition*.

Remark In the above case A is called *non-defective*. In this case the matrix X provides a *basis* for $\mathbb{R}^n$.

Complex Matrices

Defitinion Let $x, y \in \mathbb{C}^n$. The scalar product is

$$(x, y) = x^H y = \sum_{i=1}^{n} \bar{x}_i y_i,$$

and the *Euclidean norm* is

$$\|x\|_2^2 = x^H x = \sum_{i=1}^{n} |x_i|^2,$$

where x^H is the Hermitean transpose of x.

Remark The real case is a special case of the complex one.

Definition If a matrix $X \in \mathbb{C}^{n\times n}$ satisfies $X^H X = I$ then $X = (x_1, \ldots, x_n)$ is *unitary* and its column vectors form an orthonormal basis for $\mathbb{C}^n$.

Lemma If Q is *unitary* then $\|Qx\|_2 = \|x\|_2$.

Remark In the real case the matrix Q is *orthogonal*.

Lemma If $A \in \mathbb{C}^{n\times n}$ is *Hermitean*, i.e. $A = A^H$, then its eigenvectors are *unitary* so $X^{-1} = X^H$ and $A = XDX^H$.

Corollary If A is real and symmetric, i.e. $A = A^T$, then its eigenvectors are *orthogonal* so $X^{-1} = X^T$ and $A = XDX^T$.

Remark Both Hermitean and Symmetric matrices have *real* eigenvalues. Anti-Hermitean, i.e. $A^H = -A$, have *pure imaginary* eigenvalues.

Lemma If (λ, x) is an eigenpair then $(A - \lambda I)x = 0$, $x \neq 0$, and therefore the eigenvalues are the roots of

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

where $p_A(\lambda)$ is called the *Characteristic polynomial* of A.

Remark If A^{-1} exists so that $Ax = b$ has a unique solution for every b. Then A is *non-singular*. This is equivalent to zero not beeing an eigenvalue of A.

Suppose $p_A(\lambda)$ is the characteristic polynomial of A. Since $p_A(\lambda)$ has n roots we may write

$$p_A(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\dots(\lambda_n - \lambda).$$

Definition The *Algebraic multiplicity* $\gamma_1(\lambda)$ of λ is its multiplicity as a root of $p_A(\lambda)$.

Definition The *Geometric multiplicity* $\gamma_2(\lambda)$ is given by $\gamma_2(\lambda) = \dim(\text{null}(A - \lambda I))$

Remark This is the number of linearly independent eigenvectors associated with λ.

Example

Consider the *Jordan block*

$$J = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

What are the eigenvalues? Algebraic and Geometric multiplicity?

Lemma It holds that $1 \leq \gamma_2(\lambda) \leq \gamma_1(\lambda)$.

Definition If $\gamma_1(\lambda) = \gamma_2(\lambda)$ for all eigenvalues λ then A is *non–defective*.

Remark In this case we can diagonalize $A = XDX^{-1}$! For a *defective eigenvalue* we have $\gamma_2(\lambda) < \gamma_1(\lambda)$.

Applications

Example Suppose A is non–defective and consider the first order system of ODEs,

$$y'(t) = Ay(t), \quad t > 0, y(0) = b.$$

The solution can be written

$$y(t) = c_1 x_1 e^{\lambda_1 t} + c_2 x_2 e^{\lambda_2 t} + \dots + c_n x_n e^{\lambda_n t}, \quad \text{with}, \quad c = X^{-1} b.$$

Example Roots of the polynomial $p(x) = x^3 + c_2x^2 + c_1x + c_0$ are the eigenvalues of the *companion matrix*

$$\begin{pmatrix} -c_2 & -c_1 & -c_0 \\ 1 & & \\ & 1 & \end{pmatrix}.$$

Remark The matlab code `roots` is based on the companion matrix.

Functions of Matrices

The Taylor series representation of the *scalar* function $f(t)$ is

$$f(t) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} t^k.$$

For any matrix $A \in \mathbb{R}^{n\times n}$ we define

$$f(A) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} A^k.$$

Remark If the power series for $f(t)$ is absolute convergent for $|t| < L$ then the series $f(A)$ is absolute convergent for $\|A\| < L$.

Proposition If A is non–defective then

$$f(A) = Xf(D)X^{-1},$$

where X is the eigenvector matrix, $D = \mathrm{diag}(\lambda_i)$ are the eigenvalues, and $f(D) = \mathrm{diag}(f(\lambda_i))$.

Remarks The eigenvalue decomposition offers a cheap and stable way to compute $f(A)$.

Matlab has functions `expm`, `cosm`, etc. There is also a function `funm`.

Localization of Eigenvalues

Theorem (Gershgorin I) The eigenvalues of A are located in the union of the n discs.

$$|\lambda - a_{ii}| \le r_i = \sum_{j\neq i} |a_{ij}|, \qquad i = 1, 2, \ldots, n.$$

Remark Since $\lambda(A)$ and $\lambda(A^T)$ we can replace r_i by $c_i = \sum_{j\neq i} |a_{ji}|$.

Theorem (Gershgorin II) Every isolated subset of discs contains exactly as many eigenvalues as the number of discs.

Example Locate the eigenvalues of the matrix

$$A = \begin{pmatrix} 3.3 & 0.4 & -0.7 \\ 0.3 & 2.8 & 0.2 \\ 0.2 & -0.3 & -4 \end{pmatrix}$$

as accurately as possible. Can you conclude that the eigenvalues are real?

The Rayleigh Quotient

Definition Let $A \in \mathbb{R}^{n\times n}$ and $u \in \mathbb{R}^n$ be a non–zero vector. The function

$$\rho(u)=\frac{u^T Au}{u^T u},$$

is called the *Rayleigh quotient*.

Remark The Rayleight qoutient is obtained by treating $Au = \rho u$ as a least squares problem. Thus if (x, λ) is an eigenpair of A then $\rho(x) = \lambda$.

The power method

Suppose A is real, non–defective, $|\lambda_1|>|\lambda_2|\geq\ldots\geq|\lambda_n|$, and $\{x_i\}$ are the eigenvectors.

Algorithm Take $q^{(0)}$ such that $\|q^{(0)}\|_2 = 1$ and form for $k = 1, 2, \ldots,$

$$\begin{aligned} w^{(k)} &= Aq^{(k-1)}, \\ \rho_{k-1} &= (q^{(k-1)})^T w^{(k)}, \\ q^{(k)} &= w^{(k)}/\|w^{(k)}\|_2. \end{aligned}$$

Then $(\rho_k, q^{(k)})$ converge to the eigenpair (λ_1, x_1).

Stopping rule If $A = A^T$ and $r = Aq^{(k)} - \rho_k q^{(k)}$. Then $\|r\|_2 < \varepsilon$ ensures that $|\lambda_1 - \rho_k| < \varepsilon$.

Proposition The power method computes estimates $(\rho_k, q^{(k)})$ of (λ_1, x_1) that satisfy

$$q^{(k)} = \pm x_1 + O(\gamma^k), \qquad \text{and,} \qquad \rho_k = \lambda_1 + O(\gamma^{k_1}),$$

where $\gamma = |\lambda_2|/|\lambda_1|$ and $k_1 = 2k$ if A is symmetric and $k_1 = k$ otherwise.

Remark The speed of convergence depend on the *quotient* $\gamma = |\lambda_2/\lambda_1|$. The factor γ can be improved by linear transformations.

Lemma Suppose $B = A - sI$. Then $\lambda(B) = \lambda(A) - s$.

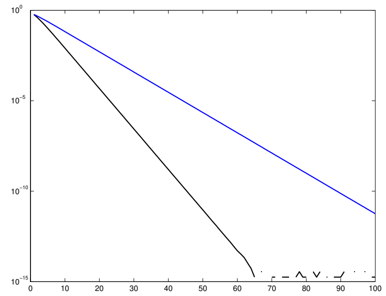
Example Suppose

```
>> A=[ 3 4 1 ; 4 5 -1 ; 1 -1 6]
```

The eigenvalues and eigenvectors are

```
>> [X,D]=eig(A)
X =                                D =
-0.7674    0.2242   -0.6007   -0.4297         0         0
 0.6046   -0.0587   -0.7943         0    6.2909         0
 0.2134    0.9728    0.0905         0         0    8.1388
```


Example Power iteration results.



The errors $\|\bar{x}^{(k)} - x_3\|_2$ (blue) and $|\rho_k - \lambda_3|$ (black). A is symmetric and $\gamma = |\lambda_2/\lambda_3| = 0.7729$. This is *slow* converence.

Inverse iteration

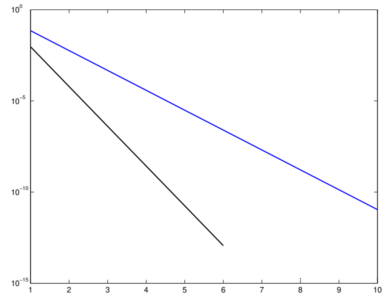
Proposition Let (λ, x) be an eigenpair of A. Put $B = (A - sI)^{-1}$. Then (μ, x), with $\mu = 1/(\lambda - s)$, is an eigenpair of B.

Example If A has eigenvalues $\lambda_1 = -0.4297$, $\lambda_2 = 6.2909$, and $\lambda_3 = 8.1388$. Then with $s = 8$ we get,

$$\gamma = |\frac{\lambda_3 - s}{\lambda_2 - s}| = |\frac{8.1388 - 8}{6.2909 - 8}| \approx 0.0812$$

Much faster convergence than the power method ($\gamma = |\lambda_2/\lambda_3| = 0.7729$).

Example Inverse iteration results with $s = 8$.



The errors $\|\bar{x}^{(k)} - x_3\|_2$ (blue) and $|\rho_k - \lambda_3|$ (black). This is *fast* converence.

Lemma Let (x, λ) be an eigenpair of a symmetric matrix A, and $\bar{x}$ an approximation of x. If $\|x - \bar{x}\|_2 = O(\varepsilon)$ then $|\lambda - \rho(\bar{x})| = O(\varepsilon^2)$.

Lemma If A is symmetric then $B = (A - sI)^{-1}$ is also symmetric.

Remark This means the faster convergence.

Lemma If λ is an eigenvalue then $A - \lambda I$ is *singular*. We can compute the eigenvector by, e.g., setting $x_1 = 1$, and solving $Ax = \lambda x$.

Remark This means that we only need to compute eigenvalues. Eigenvectors can easily be obtained.

Question How to find more efficient methods for computing eigenvalues.