chapter 4 : computing eigenvalues

e-value : frequency , growth rate , energy level

e-vector : normal mode , principal component , bound state

note : We assume that $A$ is symmetric. This implies that the e-values $\lambda_i$ are real and the e-vectors $q_i$ form an orthonormal basis, i.e. $q_i^T q_j = 0, ||q_i||_2 = 1$, and any $x$ can be written as a linear combination of the $q_i$. (pf : omit) We further assume that $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$.

ex : $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$

$$f_A(\lambda) = \det \begin{pmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{pmatrix} = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = (\lambda-3)(\lambda-1)$$

$$\lambda_1 = 3 \ : \ Ax = 3x \ \Rightarrow \ \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

choose $x_1 = 1$, then $2 - x_2 = 3, -1 + 2x_2 = 3x_2 \ \Rightarrow \ x_2 = -1$

normalize : $q_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \ \Rightarrow \ ||q_1||_2 = 1$

$$\lambda_2 = 1 \ : \ Ax = x \ \Rightarrow \ \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

choose $x_1 = 1$, then $2 - x_2 = 1, -1 + 2x_2 = x_2 \ \Rightarrow \ x_2 = 1$

normalize : $q_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \ \Rightarrow \ ||q_2||_2 = 1$ , check : $q_1^T q_2 = 0$ ok

obvious approach for computing e-values

form $f_A(\lambda) = \det(A - \lambda I)$ , solve $f_A(\lambda) = 0$ by the methods of chapter 2

ex

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ , \ \tilde{A} = \begin{pmatrix} 1+\epsilon & 0 \\ 0 & 1-\epsilon \end{pmatrix} \ : \ \text{perturbed matrix}$$

$$f_A(\lambda) = (1-\lambda)^2 = \lambda^2 - 2\lambda + 1 \ \Rightarrow \ \lambda = 1$$

$$f_{\tilde{A}}(\lambda) = (1+\epsilon-\lambda)(1-\epsilon-\lambda) = \lambda^2 - 2\lambda + 1 - \epsilon^2 \ \Rightarrow \ \lambda = 1 \pm \epsilon$$

This shows that the e-values of the perturbed matrix are close to the e-values of the original matrix (this is true for symmetric matrices in general). However, note that a small change of size $\epsilon^2$ in the coefficients of the characteristic polynomial leads to a relatively large change of size $\epsilon$ in the roots. Hence the roots of $f_A(\lambda)$ depend very sensitively on the coefficients.

ex (Wilkinson)

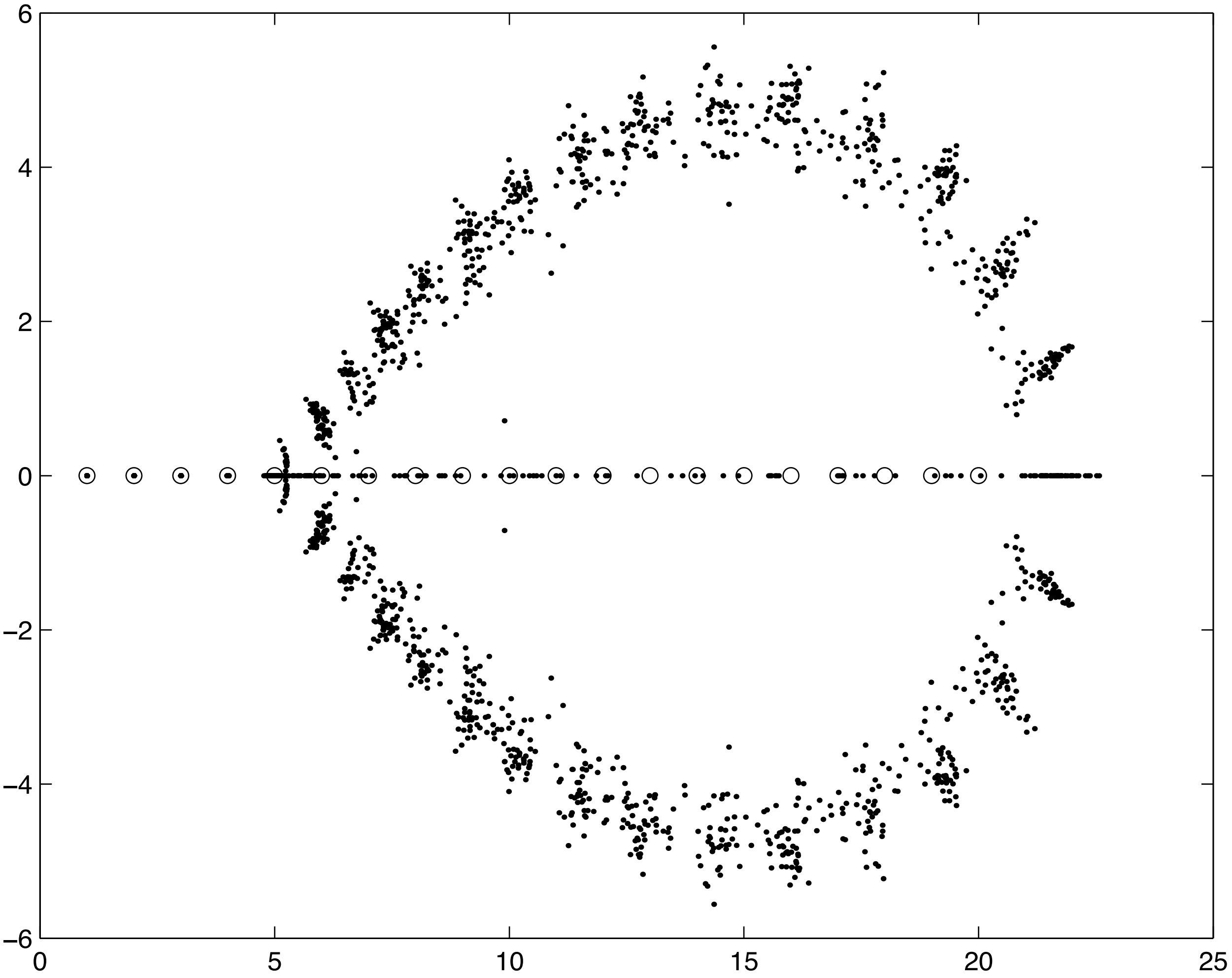
$A = \operatorname{diag}(1, 2, \ldots, 20)$

$f_A(\lambda) = (1-\lambda)(2-\lambda)\cdots(20-\lambda) = \sum_{k=0}^{20} a_k \lambda^k$

set $\tilde{a}_k = a_k(1 + 10^{-10}\epsilon_k)$ , $\epsilon_k \in (0,1)$ : random , $p(\lambda) = \sum_{k=0}^{20} \tilde{a}_k \lambda^k$ , roots = ?

Matlab

```
plot( zeros(1,20) , 'o' ); hold on;
for i=1:100
  r = roots( poly(1:20) .* (ones(1,21)+1e-10*randn(1,21)) );
  plot( r , '.'  ); axis([0 , 25 , -6 , 6 ]);
end
```



This example emphasizes that the roots of the characteristic polynomial can depend very sensitively on the coefficients, even if the e-value problem for the matrix is well-conditioned. Hence, solving $f_A(\lambda) = 0$ is not a suitable approach for computing e-values (in general).

def : Given any $x \neq 0$, define $R_A(x) = \dfrac{x^TAx}{x^Tx}$ : Rayleigh quotient.

note

1. For $x = q_j$, $R_A(q_j) = \dfrac{q_j^TAq_j}{q_j^Tq_j} = \dfrac{q_j^T\lambda_jq_j}{q_j^Tq_j} = \lambda_j$.

2. For $x \approx q_j$, $R_A(x)$ is an approximation to $\lambda_j$ and we can derive an error estimate by Taylor expansion. First recall some notation.

$$f(x_1,x_2) = f(a_1,a_2) + \frac{\partial f}{\partial x_1}(a_1,a_2)(x_1-a_1) + \frac{\partial f}{\partial x_2}(a_1,a_2)(x_2-a_2) + \cdots$$

$$f(x) = f(a) + \nabla f(a)\cdot(x-a) + O(||x-a||^2) \quad , \quad \nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)$$

Now apply this to the Rayleigh quotient.

$$R_A(x) \,=\, R_A(q_j) \,+\, \nabla R_A(q_j)\cdot(x-q_j) \,+\, O(||x-q_j||^2)$$

$$\nabla R_A(x) \,=\, \nabla\left(\frac{x^TAx}{x^Tx}\right) = \frac{x^Tx\cdot\nabla(x^TAx) \,-\, x^TAx\cdot\nabla(x^Tx)}{(x^Tx)^2}$$

$$\nabla(x^Tx) \,=\, \nabla(x_1^2+x_2^2) \,=\, (2x_1,2x_2) \,=\, 2x^T$$

$$x^TAx \,=\, (x_1,x_2)\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22}\end{pmatrix}\begin{pmatrix} x_1 \\ x_2\end{pmatrix} \,=\, a_{11}x_1^2 \,+\, 2a_{12}x_1x_2 \,+\, a_{22}x_2^2$$

$$\nabla(x^TAx) \,=(2a_{11}x_1 \,+\, 2a_{12}x_2, 2a_{12}x_1 \,+\, 2a_{22}x_2) \,=\, 2(Ax)^T$$

$$\nabla R_A(x) \,=\, \frac{x^Tx\cdot 2(Ax)^T -\, x^TAx\cdot 2x^T}{(x^Tx)^2} \,=\, \frac{2}{x^Tx}((Ax)^T - R_A(x)x^T)$$

$$\nabla R_A(q_j) \,=\, \frac{2}{q_j^Tq_j}((Aq_j)^T - R_A(q_j)q_j^T) \,=\, 2((\lambda_jq_j)^T - \lambda_jq_j^T) \,=\, 0$$

$R_A(x) \,=\, \lambda_j \,+\, O(||x-q_j||^2)$ : quadratic appoximation

section 4.1 : power method

idea : $v$ , $Av$ , $A^2v$ , ...

algorithm

```
1. $v^{(0)}$ : given , $||v^{(0)}||_2 = 1$
2. for $k = 1,2,\ldots$
3.     $w = Av^{(k-1)}$            % this can be done efficiently if $A$ is sparse
4.     $v^{(k)} = w/||w||_2$       % this is done to avoid overflow/underflow
5.     $\lambda^{(k)} = (v^{(k)})^TAv^{(k)}$
```

thm

Assume that $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$ and $q_1^T v^{(0)} \neq 0$.

Then $||v^{(k)} - (\pm q_1)|| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$ , $|\lambda^{(k)} - \lambda_1| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right)$.

The ± depends on sign of $\lambda_1$.

pf

$$v^{(0)} = \alpha_1 q_1 + \alpha_2 q_2 + \cdots + \alpha_n q_n \quad , \quad \text{where } \alpha_i = q_i^T v^{(0)}$$

$$\begin{aligned} v^{(k)} = \beta_k A^k v^{(0)} &= \beta_k \left(\alpha_1 \lambda_1^k q_1 + \alpha_2 \lambda_2^k q_2 + \cdots + \alpha_n \lambda_n^k q_n\right) \\ &= \beta_k \lambda_1^k \left(\alpha_1 q_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k q_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1}\right)^k q_n\right) \end{aligned}$$

If $q_1^T v^{(0)} = 0$, then the scheme converges to $q_2, \lambda_2$. ok

note : The power method has some limitations.

1. it only gives the largest e-value $\lambda_1$
2. $v^{(k)}$ , $\lambda^{(k)}$ converge linearly and the convergence factor $\left|\frac{\lambda_2}{\lambda_1}\right|$ may not be small

section 4.2 : inverse iteration

idea : apply power method to $A^{-1}$ , $(A - \mu I)^{-1}$ , $\mu$ : shift

1. $w = A^{-1} v \Rightarrow Aw = v$ , $w = (A - \mu I)^{-1} v \Rightarrow (A - \mu I) w = v$

2. $A q_i = \lambda_i q_i \Rightarrow A^{-1} q_i = \lambda_i^{-1} q_i$

The largest e-value of $A^{-1}$ is $\lambda_n^{-1}$, so the vectors $v^{(k)}$ converge to $q_n$.

3. $(A - \mu I) q_i = (\lambda_i - \mu) q_i \Rightarrow (A - \mu I)^{-1} q_i = (\lambda_i - \mu)^{-1} q_i$

The largest e-value of $(A - \mu I)^{-1}$ is $|\lambda_J - \mu|^{-1}$, where $\lambda_J$ is the e-value of $A$ closest to $\mu$, so the vectors $v^{(k)}$ converge to $q_J$.

algorithm

```
1. $v^{(0)}$ : given , $||v^{(0)}||_2 = 1$
2. for $k = 1, 2, \ldots$
3.     solve $(A - \mu I) w = v^{(k-1)}$        % e.g. $LU$ factorization
4.     $v^{(k)} = w / ||w||_2$
5.     $\lambda^{(k)} = (v^{(k)})^T A v^{(k)}$              % converges to $\lambda_J$
```

note : Using a suitable shift $\mu$, any e-value of $A$ can be obtained and the rate of convergence can be accelerated.

## Rayleigh quotient iteration

idea : update $\mu$

algorithm

```
1. $v^{(0)}$ : given , $||v^{(0)}||_2 = 1$
2. $\lambda^{(0)} = (v^{(0)})^T A v^{(0)}$
3. for $k = 1, 2, \ldots$
4.     solve $(A - \lambda^{(k-1)} I) w = v^{(k-1)}$
5.     $v^{(k)} = w / ||w||_2$
6.     $\lambda^{(k)} = (v^{(k)})^T A v^{(k)}$
```

thm

If $v^{(0)}$ is sufficiently close to an e-vector $q_J$, then

$$\left.\begin{aligned} ||v^{(k+1)} - (\pm q_J)|| &= O(||v^{(k)} - (\pm q_J)||^3) \\ |\lambda^{(k+1)} - \lambda_J| &= O(|\lambda^{(k)} - \lambda_J|^3) \end{aligned}\right\} \; : \; \underline{\text{cubic convergence}}$$

pf : omit

ex

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix} \; , \quad \lambda_1 = 5.214319743377 \quad , \quad v^{(0)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \frac{1}{\sqrt{3}}$$

| $k$ | power method | shifted inverse iteration , $\mu = 5$ | Rayleigh quotient iteration |
|---|---|---|---|
| 0 | 5.0 | 5.0 | 5.0 |
| 1 | 5.181818 | 5.213114 | 5.213114 |
| 2 | $\underline{5.2}08192$ | $\underline{5.214312}617$ | $\underline{5.214319743}184$ |
|  | 2 | 6 | 10 |