Lecture 38

$(A-\lambda I)x=0$

$\Updownarrow$

Ch 4: Computing eigenvalues

Problem Given A, find λ and $x \neq 0$ such that $Ax = \lambda x$

λ: e'value, frequency, growth rate, energy level

x: e'vector, normal mode, principal component, bound state

Note We assume that A is real and symmetric. This implies that the e'values λ_i are real and the e'vectors q_i form an orthonormal basis, i.e. $q_i^T q_j = 0$ for $i \neq j$, $\|q_i\|_2 = 1$, and any x can be written as a linear combination of the q_i (Pf: omit).

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

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

$\lambda_1 = 1 \Rightarrow (A - \lambda_1 I)x = 0 \Rightarrow \begin{pmatrix} 2-1 & -1 \\ -1 & 2-1 \end{pmatrix} x = 0 \Rightarrow \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$x_1 - x_2 = 0 \Rightarrow$ let $x_2 = s$: free parameter $\Rightarrow x_1 = x_2 = s$

$\Rightarrow x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = s\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ take $s = 1 \Rightarrow x = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Rightarrow \boxed{q_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}}$

$\|x\| = \sqrt{2}$ $\quad \|q_1\| = 1$

$\lambda_2 = 3 \Rightarrow (A - \lambda_2 I)x = 0 \quad \begin{pmatrix} 2-3 & -1 \\ -1 & 2-3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \sim \begin{pmatrix} -1 & -1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$-x_1 - x_2 = 0$ Take $x_1 = 1 \Rightarrow x_2 = -1 \Rightarrow x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ $\quad \|x\|_2 = \sqrt{2}$

$\Rightarrow \boxed{q_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}}$ $\quad \|q_2\| = 1$

<u>Check</u>: $q_1^T q_2 = \frac{1}{\sqrt{2}}(1 \;\; 1) \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0 \Rightarrow q_1 \perp q_2$ ✓

<u>obvious method for computing e'values</u>

<u>Step 1</u> form $f_A(\lambda) = \det(A - \lambda I)$

<u>Step 2</u> solve $f_A(\lambda) = 0$ by the methods of Chapter 2 (rootfinding methods)

<u>Ex</u>

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

$$f_A(\lambda) = (1-\lambda)^2 = \lambda^2 - 2\lambda + 1 = 0 \Rightarrow \lambda = 1$$

$$f_{\tilde{A}}(\lambda) = (1+\varepsilon-\lambda)(1-\varepsilon-\lambda) = \lambda^2 - 2\lambda + 1 - \varepsilon^2 = 0 \Rightarrow \lambda = 1 \pm \varepsilon$$

Hence a change of size ε in the matrix elements leads to a change of size ε in the e'values, but a change of size ε^2 in the coefficients of the characteristic polynomial leads to a change of size ε in the roots. Hence the roots of $f_A(\lambda)$ depend sensitively on the coefficients and this implies that the obvious method for computing e'values is unstable.

Ex (Wilkinson)

$$A = \mathrm{diag}(1, 2, \dots, 20) = \begin{pmatrix} 1 & & & 0 \\ & 2 & & \\ & & \ddots & \\ 0 & & & 20 \end{pmatrix}$$

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

Set $\tilde{a}_k = a_k(1 + 10^{-10}\varepsilon_k)$, $\quad \varepsilon_k \in (0,1)$: random

$$p(\lambda) = \sum_{k=0}^{20} \tilde{a}_k \lambda^k, \quad \text{roots} = ?$$

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
```

See printout of the program.

This example shows that roots of the characteristic polynomial can be very sensitive to perturbations in the coefficients. Hence, solving $f_A(\lambda) = 0$ is not a practical method for computing e'values (in general).

Def Given any $x \neq 0$, define $R_A(x) = \dfrac{x^T A x}{x^T x} = \dfrac{\langle x, Ax \rangle}{\langle x, x \rangle}$:

<u>Rayleigh Quotient</u>

Note

1. For $x = q_i$, $R_A(q_i) = \dfrac{q_i^T A q_i}{\underbrace{q_i^T q_i}_{=1}} = \dfrac{q_i^T \lambda_i q_i}{1} = \lambda_i \underbrace{q_i^T q_i}_{=1} = \lambda_i$

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

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

$\nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$: gradient

$$f(x) = f(a) + \nabla f(a) \cdot (x - a) + O(\|x - a\|^2)$$

$$R_A(x) = \underbrace{R_A(q_j)}_{=\lambda_j} + \nabla R_A(q_j)(x - q_j) + O(\|x - q_j\|^2)$$

$$\nabla R_A(x) = \nabla \left(\frac{x^T A x}{x^T x} \right) = \frac{x^T x \cdot \nabla (x^T A x) - x^T A x \cdot \nabla (x^T x)}{(x^T x)^2}$$

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

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

$$x^T A x = (x_1 \ \ x_2) \begin{pmatrix} a_{11} & a_{12} \\ \underset{a_{12}}{\overset{\shortparallel}{a_{21}}} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a_{11} x_1^2 + 2 a_{12} x_1 x_2 + a_{22} x_2^2$$

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

$$\nabla R_A(x) = \frac{x^T x \cdot 2 (Ax)^T - x^T A x \cdot 2 x^T}{(x^T x)^2} = \frac{2}{x^T x} \left((Ax)^T - R_A(x)\, x^T \right)$$