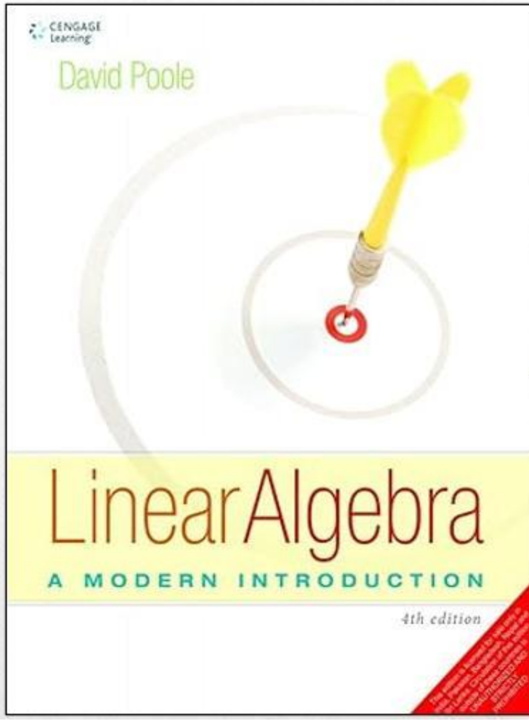

114-2

線性代數 (二)

國立政治大學
統計學系
吳漢銘

4.5 Iterative Methods for Computing Eigenvalues

https://hmwu.idv.tw

computing the eigenvalues of a matrix is to solve the characteristic equation.

However, there are several problems

The first problem is that
it depends on the computation of a determinant,
which is a very time-consuming process for large matrices.

The second problem is that the characteristic equation is a polynomial equation, and there are no formulas for solving polynomial equations of degree higher than 4 (polynomials of degrees 2, 3, and 4 can be solved using the quadratic formula and its analogues).

The Power Method

The power method applies to an $n \times n$ matrix that has a ***dominant eigenvalue*** λ_1—that is, an eigenvalue that is larger in absolute value than all of the other eigenvalues.

For example, if a matrix has eigenvalues -4, -3, 1, and 3, then -4 is the dominant eigenvalue, since $4 = |-4| > |-3| \geq |3| \geq |1|$. On the other hand, a matrix with eigenvalues -4, -3, 3, and 4 has no dominant eigenvalue.

Theorem 4.28

Let A be an $n \times n$ diagonalizable matrix with dominant eigenvalue λ_1. Then there exists a nonzero vector $\mathbf{x}_0$ such that the sequence of vectors $\mathbf{x}_k$ defined by

$$\mathbf{x}_1 = A\mathbf{x}_0, \mathbf{x}_2 = A\mathbf{x}_1, \mathbf{x}_3 = A\mathbf{x}_2, \ldots, \mathbf{x}_k = A\mathbf{x}_{k-1}, \ldots$$

approaches a dominant eigenvector of A.

Example 4.30

Approximate the dominant eigenvector of $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$ using the method of Theorem 4.28.

Solution We will take $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ as the initial vector.

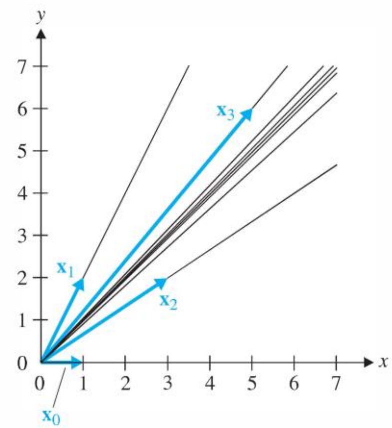

Figure 4.14

k	0	1	2	3	4	5	6	7	8
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 3 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 5 \\ 6 \end{bmatrix}$	$\begin{bmatrix} 11 \\ 10 \end{bmatrix}$	$\begin{bmatrix} 21 \\ 22 \end{bmatrix}$	$\begin{bmatrix} 43 \\ 42 \end{bmatrix}$	$\begin{bmatrix} 85 \\ 86 \end{bmatrix}$	$\begin{bmatrix} 171 \\ 170 \end{bmatrix}$
r_k	—	0.50	1.50	0.83	1.10	0.95	1.02	0.99	1.01
l_k	—	1.00	3.00	1.67	2.20	1.91	2.05	1.98	2.01

the ratio r_k of the first to the second component of $\mathbf{x}_k$ gets very close to 1 as k increases.

the ratio l_k of the first component of $\mathbf{x}_{k+1}$ to that of $\mathbf{x}_k$ will approach λ_1 as k increases.

Table 4.2

k	0	1	2	3	4	5	6	7	8
$\mathbf{x}_k$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\2\end{bmatrix}$	$\begin{bmatrix}1.5\\1\end{bmatrix}$	$\begin{bmatrix}1.67\\2\end{bmatrix}$	$\begin{bmatrix}1.83\\1.67\end{bmatrix}$	$\begin{bmatrix}1.91\\2\end{bmatrix}$	$\begin{bmatrix}1.95\\1.91\end{bmatrix}$	$\begin{bmatrix}1.98\\2\end{bmatrix}$	$\begin{bmatrix}1.99\\1.98\end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0.5\\1\end{bmatrix}$	$\begin{bmatrix}1\\0.67\end{bmatrix}$	$\begin{bmatrix}0.83\\1\end{bmatrix}$	$\begin{bmatrix}1\\0.91\end{bmatrix}$	$\begin{bmatrix}0.95\\1\end{bmatrix}$	$\begin{bmatrix}1\\0.98\end{bmatrix}$	$\begin{bmatrix}0.99\\1\end{bmatrix}$	$\begin{bmatrix}1\\0.99\end{bmatrix}$
m_k	1	2	1.5	2	1.83	2	1.95	2	1.99

divide each $\mathbf{x}_k$ by the component with the maximum absolute value, so that the largest component is now 1. This method is called ***scaling.***

if m_k denotes the component of $\mathbf{x}_k$ with the maximum absolute value,

we will replace $\mathbf{x}_k$ by $\mathbf{y}_k = (1/m_k)\mathbf{x}_k$.

The Power Method

Let A be a diagonalizable $n \times n$ matrix with a corresponding dominant eigenvalue λ_1.

1. Let $\mathbf{x}_0 = \mathbf{y}_0$ be any initial vector in $\mathbb{R}^n$ whose largest component is 1.
2. Repeat the following steps for $k = 1, 2, \ldots$:
   (a) Compute $\mathbf{x}_k = A\mathbf{y}_{k-1}$.
   (b) Let m_k be the component of $\mathbf{x}_k$ with the largest absolute value.
   (c) Set $\mathbf{y}_k = (1/m_k)\mathbf{x}_k$.

For most choices of $\mathbf{x}_0$, m_k converges to the dominant eigenvalue λ_1 and $\mathbf{y}_k$ converges to a dominant eigenvector.

Example 4.31 Use the power method to approximate the dominant eigenvalue and a dominant eigenvector of

$$A = \begin{bmatrix} 0 & 5 & -6 \\ -4 & 12 & -12 \\ -2 & -2 & 10 \end{bmatrix}$$

Solution Taking as our initial vector $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

Table 4.3

k	0	1	2	3	4	5	6	7
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -1 \\ -4 \\ 6 \end{bmatrix}$	$\begin{bmatrix} -9.33 \\ -19.33 \\ 11.67 \end{bmatrix}$	$\begin{bmatrix} 8.62 \\ 17.31 \\ -9.00 \end{bmatrix}$	$\begin{bmatrix} 8.12 \\ 16.25 \\ -8.20 \end{bmatrix}$	$\begin{bmatrix} 8.03 \\ 16.05 \\ -8.04 \end{bmatrix}$	$\begin{bmatrix} 8.01 \\ 16.01 \\ -8.01 \end{bmatrix}$	$\begin{bmatrix} 8.00 \\ 16.00 \\ -8.00 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.17 \\ -0.67 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.48 \\ 1 \\ -0.60 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1 \\ -0.52 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1 \\ -0.50 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1 \\ -0.50 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1 \\ -0.50 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1 \\ -0.50 \end{bmatrix}$
m_k	1	6	−19.33	17.31	16.25	16.05	16.01	16.00

The Shifted Power Method and the Inverse Power Method

The ***shifted power method*** uses the observation that, if λ is an eigenvalue of A, then $\lambda - \alpha$ is an eigenvalue of $A - \alpha I$ for any scalar α (Exercise 22 in Section 4.3).

Thus, if λ_1 is the dominant eigenvalue of A, the eigenvalues of $A - \lambda_1 I$ will be 0, $\lambda_2 - \lambda_1, \lambda_3 - \lambda_1, \ldots, \lambda_n - \lambda_1$. We can then apply the power method to compute $\lambda_2 - \lambda_1$, and from this value we can find λ_2. Repeating this process will allow us to compute all of the eigenvalues.

Example 4.32 Use the shifted power method to compute the second eigenvalue of the matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$ from Example 4.30.

Solution In Example 4.30, we found that $\lambda_1 = 2$. To find λ_2, we apply the power method to

$A - 2I = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}$ We take $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, but other choices will also work.

Our choice of $\mathbf{x}_0$ has produced the eigenvalue -3 after only two iterations.

Therefore, $\lambda_2 - \lambda_1 = -3$,

so $\lambda_2 = \lambda_1 - 3 = 2 - 3 = -1$ is the second eigenvalue of A.

inverse power method

Theorem 4.18 (b) if A is invertible with eigenvalue λ, then A^{-1} has eigenvalue $1/\lambda$.

dominant eigenvalue will be the *reciprocal of the smallest* (in magnitude) eigenvalue λ, then A^{-1} has eigenvalue $1/\lambda$.

apply the power method to A^{-1}, dominant eigenvalue will be

the *reciprocal of the smallest* (in magnitude) eigenvalue of A.

we compute $\mathbf{x}_k = A^{-1}\mathbf{y}_{k-1}$.

Example 4.33 Use the inverse power method to compute the second eigenvalue of the matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$ from Example 4.30.

Solution We start, as in Example 4.30, with $\mathbf{x}_0 = \mathbf{y}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. To solve $A\mathbf{x}_1 = \mathbf{y}_0$,

Table 4.5

k	0	1	2	3	4	5	6	7	8	9
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix}$	$\begin{bmatrix} -0.5 \\ 1.5 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -0.83 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -1.1 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -0.95 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -1.02 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -0.99 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -1.01 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} -0.33 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.6 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.45 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.52 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.49 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.51 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.50 \\ 1 \end{bmatrix}$
m_k	1	1	0.5	1.5	−0.83	−1.1	−0.95	−1.02	−0.99	−1.01

the values m_k are converging to -1.

smallest eigenvalue of A is the reciprocal of -1 (which is also -1).

The Shifted Inverse Power Method

find an approximation for *any* eigenvalue, provided

we have a close approximation to that eigenvalue.

if a scalar α is given,
the ***shifted inverse power method*** will find the eigenvalue λ of A that is closest to α.

If λ is an eigenvalue of A and $\alpha \neq \lambda$, then $A - \alpha I$ is invertible if α is not an eigenvalue of A and $1/(\lambda - \alpha)$ is an eigenvalue of $(A - \alpha I)^{-1}$.

If α is close to λ, then $1/(\lambda - \alpha)$ will be a dominant eigenvalue of $(A - \alpha I)^{-1}$.

if α is *very* close to λ, then $1/(\lambda - \alpha)$ will be *much* bigger in magnitude than the next eigenvalue, so the convergence will be very rapid.

CAS

Exercises 4.5

2, 3, 6, 9, 12, 21

Example 4.34 Use the shifted inverse power method to approximate the eigenvalue of

$$A = \begin{bmatrix} 0 & 5 & -6 \\ -4 & 12 & -12 \\ -2 & -2 & 10 \end{bmatrix}$$

that is closest to 5.