Numerical Determination of Eigenvalues and Eigenvectors

22.4

Introduction

In Section 22.1 it was shown how to obtain eigenvalues and eigenvectors for low order matrices, 2×2 and 3×3. This involved firstly solving the characteristic equation $\det(A - \lambda I) = 0$ for a given $n \times n$ matrix A. This is an n^{th} order polynomial equation and, even for n as low as 3, solving it is not always straightforward. For large n even obtaining the characteristic equation may be difficult, let alone solving it.

Consequently in this Section we give a brief introduction to alternative methods, essentially **numerical** in nature, of obtaining eigenvalues and perhaps eigenvectors.

We would emphasize that in some applications such as Control Theory we might only require one eigenvalue of a matrix A, usually the one largest in magnitude which is called the **dominant** eigenvalue. It is this eigenvalue which sometimes tells us how a control system will behave.

Prerequisites

Before starting this Section you should ...

- have a knowledge of determinants and matrices
- have a knowledge of linear first order differential equations

Learning Outcomes

On completion you should be able to ...

- use the power method to obtain the dominant eigenvalue (and associated eigenvector) of a matrix
- state the main advantages and disadvantages of the power method

1. Numerical determination of eigenvalues and eigenvectors

Preliminaries

Before discussing **numerical** methods of calculating eigenvalues and eigenvectors we remind you of three results for a matrix A with an eigenvalue λ and associated eigenvector X.

- A^{-1} (if it exists) has an eigenvalue $\dfrac{1}{\lambda}$ with associated eigenvector X.
- The matrix $(A - kI)$ has an eigenvalue $(\lambda - k)$ and associated eigenvector X.
- The matrix $(A - kI)^{-1}$, i.e. the inverse (if it exists) of the matrix $(A - kI)$, has eigenvalue $\dfrac{1}{\lambda - k}$ and corresponding eigenvector X.

  Here k is any real number.

Task

The matrix $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 5 \end{bmatrix}$ has eigenvalues $\lambda = 5,\ 3,\ 1$ with associated eigenvectors $\begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ respectively.

The inverse A^{-1} exists and is $\qquad A^{-1} = \dfrac{1}{3}\begin{bmatrix} 2 & -1 & -5 \\ -1 & 2 & -5 \\ 0 & 0 & \dfrac{3}{5} \end{bmatrix}$

Without further calculation write down the eigenvalues and eigenvectors of the following matrices:

(a) A^{-1} (b) $\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 0 & 0 & 6 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 3 \end{bmatrix}^{-1}$

Your solution

Answer

(a) The eigenvalues of A^{-1} are $\frac{1}{5},\ \frac{1}{3},\ 1$. (Notice that the dominant eigenvalue of A yields the smallest magnitude eigenvalue of A^{-1}.)

(b) The matrix here is $A + I$. Thus its eigenvalues are the same as those of A increased by 1 i.e. $6,\ 4,\ 2$.

(c) The matrix here is $(A - 2I)^{-1}$. Thus its eigenvalues are the reciprocals of the eigenvalues of $(A - 2I)$. The latter has eigenvalues $3,\ 1,\ -1$ so $(A - 2I)^{-1}$ has eigenvalues $\frac{1}{3},\ 1,\ -1$.

In each of the above cases the eigenvectors are the same as those of the original matrix A.

The power method

This is a **direct iteration** method for obtaining the **dominant** eigenvalue (i.e. the largest in magnitude), say λ_1, for a given matrix A and also the corresponding eigenvector.

We will not discuss the theory behind the method but will demonstrate it in action and, equally importantly, point out circumstances when it **fails**.

Task

Let $A = \begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix}$. By solving $\det(A - \lambda I) = 0$ obtain the eigenvalues of A and also obtain the eigenvector associated with the dominant eigenvalue.

Your solution

Answer

$$\det(A-\lambda I) = \begin{vmatrix} 4-\lambda & 2 \\ 5 & 7-\lambda \end{vmatrix} = 0$$

which gives

$$\lambda^2 - 11\lambda + 18 = 0 \qquad \Rightarrow \qquad (\lambda-9)(\lambda-2) = 0$$

so

$$\lambda_1 = 9 \qquad \text{(the dominant eigenvalue)} \quad \text{and} \quad \lambda_2 = 2.$$

The eigenvector $X = \begin{bmatrix} x \\ y \end{bmatrix}$ for $\lambda_1 = 9$ is obtained as usual by solving $AX = 9X$, so

$$\begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 9x \\ 9y \end{bmatrix} \quad \text{from which } 5x = 2y \quad \text{so} \quad X = \begin{bmatrix} 2 \\ 5 \end{bmatrix} \text{ or any multiple.}$$

If we normalize here such that the largest component of X is 1

$$X = \begin{bmatrix} 0.4 \\ 1 \end{bmatrix}$$

We shall now demonstrate how the power method can be used to obtain $\lambda_1 = 9$ and $X = \begin{bmatrix} 0.4 \\ 1 \end{bmatrix}$ where $A = \begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix}$.

- We choose an *arbitrary* 2×1 column vector

$$X^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

- We premultiply this by A to give a new column vector $X^{(1)}$:

$$X^{(1)} = \begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

- We 'normalize' $X^{(1)}$ to obtain a column vector $Y^{(1)}$ with largest component 1: thus

$$Y^{(1)} = \frac{1}{12} \begin{bmatrix} 6 \\ 12 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$$

- We continue the process

$$X^{(2)} = AY^{(1)} = \begin{bmatrix} 4 & 2 \\ 6 & 7 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 9.5 \end{bmatrix}$$

$$Y^{(2)} = \frac{1}{9.5} \begin{bmatrix} 4 \\ 9.5 \end{bmatrix} = \begin{bmatrix} 0.421053 \\ 1 \end{bmatrix}$$

Task

Continue this process for a further step and obtain $X^{(3)}$ and $Y^{(3)}$, quoting values to 6 d.p.

Your solution

Answer

$$X^{(3)} = AY^{(2)} = \begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix} \begin{bmatrix} 0.421053 \\ 1 \end{bmatrix} = \begin{bmatrix} 3.684210 \\ 9.105265 \end{bmatrix}$$

$$Y^{(3)} = \frac{1}{9.105265} \begin{bmatrix} 0.404624 \\ 1 \end{bmatrix}$$

The first 8 steps of the above **iterative process** are summarized in the following table (the first three rows of which have been obtained above):

Table 1

Step r	$X_1^{(r)}$	$X_2^{(r)}$	α_r	$Y_1^{(r)}$	$Y_2^{(r)}$
1	6	12	12	0.5	1
2	4	9.5	9.5	0.421053	1
3	3.684211	9.105265	9.105265	0.404624	1
4	3.618497	9.023121	9.023121	0.401025	1
5	3.604100	9.005125	9.005125	0.400228	1
6	3.600911	9.001138	9.001138	0.400051	1
7	3.600202	9.000253	9.000253	0.400011	1
8	3.600045	9.000056	9.000056	0.400002	1

In Table 1, α_r refers to the largest magnitude component of $X^{(r)}$ which is used to normalize $X^{(r)}$ to give $Y^{(r)}$. We can see that α_r is converging to 9 which we know is the dominant eigenvalue λ_1 of A. Also $Y^{(r)}$ is converging towards the associated eigenvector $[0.4,\ 1]^T$.

Depending on the accuracy required, we could decide when to stop the iterative process by looking at the difference $|\alpha_r - \alpha_{r-1}|$ at each step.

Task

Using the power method obtain the dominant eigenvalue and associated eigenvector of

$$A = \begin{bmatrix} 3 & -1 & 0 \\ -2 & 4 & -3 \\ 0 & -1 & 1 \end{bmatrix} \quad \text{using a starting column vector} \quad X^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Calculate $X^{(1)}$, $Y^{(1)}$ and α_1:

Your solution

Answer

$$X^{(1)} = AX^{(0)} = \begin{bmatrix} 3 & -1 & 0 \\ -2 & 4 & -3 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$$

so $Y^{(1)} = \frac{1}{2}\begin{bmatrix} 1 \\ -0.5 \\ 0 \end{bmatrix}$ using $\alpha_1 = 2$, the largest magnitude component of $X^{(1)}$.

Carry out the next two steps of this iteration to obtains $X^{(2)},\ Y^{(2)},\ \alpha_2$ and $X^{(3)},\ Y^{(3)},\ \alpha_3$:

Your solution

Answer

$$X^{(2)} = \begin{bmatrix} 3 & -1 & 0 \\ -2 & 4 & -3 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 3.5 \\ -4 \\ 0.5 \end{bmatrix} \qquad Y^{(2)} = -\frac{1}{4}\begin{bmatrix} -0.875 \\ 1 \\ -0.125 \end{bmatrix} \qquad \alpha_2 = -4$$

$$X^{(3)} = \begin{bmatrix} 3 & -1 & 0 \\ -2 & 4 & -3 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} -0.875 \\ 1 \\ -0.125 \end{bmatrix} = \begin{bmatrix} -3.625 \\ 6.125 \\ -1.125 \end{bmatrix} \qquad Y^{(3)} = \frac{1}{6.125}\begin{bmatrix} -0.5918 \\ 1 \\ -0.1837 \end{bmatrix} \qquad \alpha_3 = 6.125$$

After just 3 iterations there is little sign of convergence of the normalizing factor α_r. However the next two values obtained are

$$\alpha_4 = 5.7347 \qquad \alpha_5 = 5.4774$$

and, after 14 iterations, $|\alpha_{14} - \alpha_{13}| < 0.0001$ and the power method converges, albeit slowly, to

$$\alpha_{14} = 5.4773$$

which (correct to 4 d.p.) is the dominant eigenvalue of A. The corresponding eigenvector is

$$\begin{bmatrix} -0.4037 \\ 1 \\ -0.2233 \end{bmatrix}$$

It is clear that the power method requires, for its practical execution, a computer.

Problems with the power method

1. If the initial column vector $X^{(0)}$ is an eigenvector of A other than that corresponding to the dominant eigenvalue, say λ_1, then the method will fail since the iteration will converge to the wrong eigenvalue, say λ_2, after only one iteration (because $AX^{(0)} = \lambda_2 X^{(0)}$ in this case).

2. It is possible to show that the speed of convergence of the power method depends on the ratio

$$\frac{\text{magnitude of dominant eigenvalue } \lambda_1}{\text{magnitude of next largest eigenvalue}}$$

   If this ratio is small the method is slow to converge.

   In particular, if the dominant eigenvalue λ_1 is **complex** the method will fail completely to converge because the complex conjugate $\overline{\lambda}_1$ will also be an eigenvalue and $|\lambda_1| = |\overline{\lambda}_1|$

3. The power method only gives one eigenvalue, the dominant one λ_1 (although this is often the most important in applications).

Advantages of the power method

1. It is simple and easy to implement.

2. It gives the eigenvector corresponding to λ_1 as well as λ_1 itself. (Other numerical methods require separate calculation to obtain the eigenvector.)

Finding eigenvalues other than the dominant

We discuss this topic only briefly.

1. Obtaining the smallest magnitude eigenvalue

If A has dominant eigenvalue λ_1 then its inverse A^{-1} has an eigenvalue $\dfrac{1}{\lambda_1}$ (as we discussed at the beginning of this Section.) Clearly $\dfrac{1}{\lambda_1}$ will be the smallest magnitude eigenvalue of A^{-1}. Conversely if we obtain the **largest** magnitude eigenvalue, say λ_1', of A^{-1} by the power method then the **smallest** eigenvalue of A is the reciprocal, $\dfrac{1}{\lambda_1'}$.

This technique is called the **inverse power method**.

Example

If $A = \begin{bmatrix} 3 & -1 & 0 \\ -2 & 4 & -3 \\ 0 & -1 & 1 \end{bmatrix}$ then the inverse is $A^{-1} = \begin{bmatrix} 1 & 1 & 3 \\ 2 & 3 & 9 \\ 2 & 3 & 10 \end{bmatrix}$.

Using $X^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ in the power method applied to A^{-1} gives $\lambda_1' = 13.4090$. Hence the smallest magnitude eigenvalue of A is $\dfrac{1}{13.4090} = 0.0746$. The corresponding eigenvector is $\begin{bmatrix} 0.3163 \\ 0.9254 \\ 1 \end{bmatrix}$.

In practice, finding the inverse of a large order matrix A can be expensive in computational effort. Hence the inverse power method is implemented without actually obtaining A^{-1} as follows.

As we have seen, the power method applied to A utilizes the scheme:

$$X^{(r)} = AY^{(r-1)} \qquad r = 1, 2, \ldots$$

where $Y^{(r-1)} = \dfrac{1}{\alpha_{r-1}} X^{(r-1)}$, α_{r-1} being the largest magnitude component of $X^{(r-1)}$.
For the inverse power method we have

$$X^{(r)} = A^{-1}Y^{(r-1)}$$

which can be re-written as

$$AX^{(r)} = Y^{(r-1)}$$

Thus $X^{(r)}$ can actually be obtained by solving this system of linear equations without needing to obtain A^{-1}. This is usually done by a technique called LU decomposition i.e. writing A (once and for all) in the form

$A = LU \qquad L$ being a lower triangular matrix and U upper triangular.

2. Obtaining the eigenvalue closest to a given number p

Suppose λ_k is the (unknown) eigenvalue of A closest to p . We know that if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A then $\lambda_1 - p,\ \lambda_2 - p, \ldots, \lambda_n - p$ are the eigenvalues of the matrix $A - pI$. Then $\lambda_k - p$ will be the **smallest** magnitude eigenvalue of $A - pI$ but $\dfrac{1}{\lambda_k - p}$ will be the **largest** magnitude eigenvalue of $(A - pI)^{-1}$. Hence if we apply the power method to $(A - pI)^{-1}$ we can obtain λ_k. The method is called the **shifted inverse power** method.

3. Obtaining all the eigenvalues of a large order matrix

In this case neither solving the characteristic equation $\det(A - \lambda I) = 0$ nor the power method (and its variants) is efficient.

The commonest method utilized is called the **QR technique**. This technique is based on **similarity transformations** i.e. transformations of the form

$$B = M^{-1}AM$$

where B has the same eigenvalues as A. (We have seen earlier in this Workbook that one type of similarity transformation is $D = P^{-1}AP$ where P is formed from the eigenvectors of A. However, we are now, of course, dealing with the situation where we are trying to find the eigenvalues and eigenvectors of A.)

In the QR method A is reduced to upper (or lower) triangular form. We have already seen that a triangular matrix has its eigenvalues on the diagonal.

For details of the QR method, or more efficient techniques, one of which is based on what is called a **Householder** transformation, the reader should consult a text on numerical methods.