Eigenvalues and eigenvectors

Peter Rowlett

1 Eigenvalues and eigenvectors

1.1 Background

Consider a system of equations of the form

$$\begin{aligned} ax + by &= 0 \\ cx + dy &= 0. \end{aligned}$$

Here, $(x, y) = (0, 0)$ is a *trivial* solution of any such system. We might prefer to investigate situations where there are *non-trivial* solutions.

1.1.1 Example 1

$$\begin{aligned} 5x - 3y &= 0 \\ 10x - 2y &= 0. \end{aligned}$$

Representing this as an augmented matrix

$$\left[\begin{array}{cc|c} 5 & -3 & 0 \\ 10 & -2 & 0 \end{array}\right]$$

we can replace row 2 with 2×row 1 minus row 2 to obtain

$$\left[\begin{array}{cc|c} 5 & -3 & 0 \\ 0 & -8 & 0 \end{array}\right]$$

Row 2 gives $-8y = 0 \implies y = 0$, so the only solution possible is the trivial one, $(x, y) = (0, 0)$.

1.1.2 Example 2

$$\begin{aligned} 5x - 3y &= 0 \\ 10x - 6y &= 0. \end{aligned}$$

Again, as an augmented matrix:

$$\left[\begin{array}{cc|c} 5 & -3 & 0 \\ 10 & -6 & 0 \end{array}\right]$$

Replacing row 2 with 2×row 1 minus row 2 gives

$$\left[\begin{array}{cc|c} 5 & -3 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

This is consistent and tells us that the second equation is a linear combination of the first (in fact, it is double the first). If we let $y = t$ then by row 1 we have $x = \frac{3}{5}t$. There is an infinite family of solutions of the form $(x, y) = (\frac{3}{5}t, t)$ for $t \in \mathbb{R}$.

1.2 Eigenvalues

Consider the system

$$\begin{aligned} 2x + 4y &= \lambda x \\ x + 5y &= \lambda y \end{aligned}$$

where λ is some unknown constant.

This system has the trivial solution $(x, y) = (0, 0)$. We are interested in cases where there is not a trivial solution.

We can write the system in matrix form using

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 1 & 5 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \text{and} \quad \mathbf{Av} = \lambda\mathbf{v}.$$

1.3 Definition

Let $\mathbf{A}$ be an $n \times n$ matrix. Then the number λ is an *eigenvalue* of $\mathbf{A}$ if there exists a non-zero vector $\mathbf{v}$ such that

$$\mathbf{Av} = \lambda\mathbf{v}.$$

Then $\mathbf{v}$ is an *eigenvector* of $\mathbf{A}$ corresponding to λ.

Rewrite $\mathbf{Av} = \lambda\mathbf{v}$ as

$$\begin{aligned} \mathbf{Av} &= \lambda\mathbf{v} \\ \mathbf{Av} - \lambda\mathbf{v} &= \mathbf{0} \\ (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{v} &= \mathbf{0}. \end{aligned}$$

Now, we require $\mathbf{A} - \lambda\mathbf{I}$ must not be invertible, otherwise

$$
\begin{aligned}
(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{v} &= \mathbf{0} \\
(\mathbf{A}-\lambda\mathbf{I})^{-1}(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{v} &= (\mathbf{A}-\lambda\mathbf{I})^{-1}\,\mathbf{0} \\
\mathbf{v} &= \mathbf{0},
\end{aligned}
$$

and we wanted non-trivial solutions.

Recall that a square matrix $\mathbf{A}$ is invertible if and only if $\det(\mathbf{A}) \neq 0$.

So we require

$$\det(\mathbf{A}-\lambda\mathbf{I}) = 0.$$

We call $p(\lambda) = \det(\mathbf{A}-\lambda\mathbf{I})$ the *characteristic polynomial* of $\mathbf{A}$.

Example

Find the eigenvalues corresponding to

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 1 & 5 \end{bmatrix}.$$

$$
\begin{aligned}
\mathbf{A}-\lambda\mathbf{I} &= \begin{bmatrix} 2 & 4 \\ 1 & 5 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 2 & 4 \\ 1 & 5 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \\
&= \begin{bmatrix} 2-\lambda & 4 \\ 1 & 5-\lambda \end{bmatrix}. \\
\text{Then } p(\lambda) = |\mathbf{A}-\lambda\mathbf{I}| &= \begin{vmatrix} 2-\lambda & 4 \\ 1 & 5-\lambda \end{vmatrix} \\
&= (2-\lambda)(5-\lambda) - 4 \times 1 \\
&= \lambda^2 - 7\lambda + 6 = (\lambda-6)(\lambda-1) = 0
\end{aligned}
$$

when $\lambda = 6$ or 1. Say $\lambda_1 = 6$ & $\lambda_2 = 1$.

To find the eigenvector corresponding to each eigenvalue, simply solve the system of linear equations given by $(\mathbf{A}-\lambda\mathbf{I})\mathbf{v} = \mathbf{0}$.

In this case, say $(\mathbf{A}-6\mathbf{I})\,\mathbf{v}_1 = \mathbf{0}$ & $(\mathbf{A}-\mathbf{I})\,\mathbf{v}_2 = \mathbf{0}$.

1. $(\mathbf{A}-6\mathbf{I})\mathbf{v}_1 = \mathbf{0}$.

$$
\begin{aligned}
\mathbf{A}-6\mathbf{I} &= \begin{bmatrix} 2-6 & 4 \\ 1 & 5-6 \end{bmatrix} \\
&= \begin{bmatrix} -4 & 4 \\ 1 & -1 \end{bmatrix}.
\end{aligned}
$$

Let $\mathbf{v}_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$.
Then

$$(\mathbf{A} - 6\mathbf{I})\,\mathbf{v}_1 = \begin{bmatrix} -4 & 4 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Looking at row 2, we have

$$\begin{aligned} x_1 - y_1 &= 0 \\ x_1 &= y_1, \end{aligned}$$

which is satisfied by, e.g., $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
So say $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
(You should check $\mathbf{A}\mathbf{v_1} = 6\mathbf{v_1}$.)

2. $(\mathbf{A} - \mathbf{I})\mathbf{v}_2 = \mathbf{0}$.

$$\begin{aligned} \mathbf{A} - \mathbf{I} &= \begin{bmatrix} 2-1 & 4 \\ 1 & 5-1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 4 \\ 1 & 4 \end{bmatrix}. \end{aligned}$$

Let $\mathbf{v}_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$.
Then

$$(\mathbf{A} - \mathbf{I})\,\mathbf{v}_2 = \begin{bmatrix} 1 & 4 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Looking at row 1, we have

$$\begin{aligned} x_2 + 4y_2 &= 0 \\ x_2 &= -4y_2, \end{aligned}$$

which is satisfied by, e.g., $\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$.
So say $\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$.
(You should check $\mathbf{A}\mathbf{v_2} = \mathbf{v_2}$.)

2 Diagonalisation

2.1 Definition

Sometimes it can be easier to handle problems involving square matrices by transforming them to easier problems about diagonal matrices.

A *diagonal matrix* is one where the only non-zero entries are on the main diagonal (top left to bottom right).

To *diagonalise* a matrix $\mathbf{A}$ means to find an invertible matrix $\mathbf{U}$ and a diagonal matrix $\mathbf{D}$ such that

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^{-1}.$$

If we have a set of eigenvalues and eigenvectors for $\mathbf{A}$, we can achieve this by forming $\mathbf{U}$ using the eigenvectors as columns and forming $\mathbf{D}$ by including the corresponding eigenvalues as the entries along the main diagonal.

2.2 Example

Earlier we saw that

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 1 & 5 \end{bmatrix}$$

had eigenvalues $\lambda_1 = 6$ and $\lambda_2 = 1$ with corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$.

So we can diagonalise $\mathbf{A}$ using

$$\mathbf{U} = \begin{bmatrix} 1 & -4 \\ 1 & 1 \end{bmatrix}$$
$$\mathbf{D} = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$$

We must compute $\mathbf{U}^{-1}$. This is

$$\mathbf{U}^{-1} = \frac{1}{1 \times 1 - -4 \times 1} \begin{bmatrix} 1 & 4 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{5} & \frac{4}{5} \\ -\frac{1}{5} & \frac{1}{5} \end{bmatrix}.$$

We can check

$$\begin{bmatrix} 1 & -4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{5} & \frac{4}{5} \\ -\frac{1}{5} & \frac{1}{5} \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 1 & 5 \end{bmatrix}.$$

2.3 Why is this useful?

One example is if you want to compute $\mathbf{A}^n$ for arbitrary n.

Given $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^{-1}$, consider

$$\begin{aligned} \mathbf{A^2} &= \left(\mathbf{U}\mathbf{D}\mathbf{U}^{-1}\right)\left(\mathbf{U}\mathbf{D}\mathbf{U}^{-1}\right) \\ &= \mathbf{U}\mathbf{D}\left(\mathbf{U}^{-1}\mathbf{U}\right)\mathbf{D}\mathbf{U}^{-1} \\ &= \mathbf{U}\mathbf{D}\mathbf{I}\mathbf{D}\mathbf{U}^{-1} \\ &= \mathbf{U}\mathbf{D}\mathbf{D}\mathbf{U}^{-1} \\ &= \mathbf{U}\mathbf{D}^2\mathbf{U}^{-1}. \end{aligned}$$

This generalises, so we can say that $\mathbf{A}^n = \mathbf{U}\mathbf{D}^n\mathbf{U}^{-1}$.
Now consider $\mathbf{D}^n$.
If $\mathbf{D} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ then

$$\mathbf{D}^2 = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} = \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix}.$$

This process generalises, so if

$$\mathbf{D} = \begin{bmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_k \end{bmatrix},$$

then

$$\mathbf{D}^n = \begin{bmatrix} a_1^n & 0 & \dots & 0 \\ 0 & a_2^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_k^n \end{bmatrix}.$$