Eigenvalues and Eigenvectors

Linear Algebra, Michaelmas Term 2018
Jonathan Whiteley

Eigenvalues and eigenvectors

Let A be a $N \times N$ matrix

Suppose

$$A\mathbf{u} = \lambda\mathbf{u}$$

for some scalar λ and non-zero vector $\mathbf{u}$

We then say that λ is an eigenvalue of A, with corresponding eigenvector $\mathbf{u}$

We insist that an eigenvector must be a non-zero vector.

This is because if $\mathbf{u} = \mathbf{0}$ then $A\mathbf{u} = \lambda\mathbf{u}$ is trivially true for all λ

The eigenvector corresponding to an eigenvalue λ is not unique

Suppose $A\mathbf{u} = \lambda\mathbf{u}$

If $\mathbf{v} = c\mathbf{u}$ for some scalar $c \neq 0$ then

$$\begin{aligned} A\mathbf{v} &= A(c\mathbf{u}) \\ &= cA\mathbf{u} \\ &= c\lambda\mathbf{u} \\ &= \lambda(c\mathbf{u}) \\ &= \lambda\mathbf{v} \end{aligned}$$

and so $\mathbf{v}$ is also an eigenvector corresponding to the eigenvalue λ

Calculating eigenvalues and eigenvectors

We are seeking λ and (non-zero) $\mathbf{u}$ such that

$$A\mathbf{u} = \lambda\mathbf{u}$$

We may write this as

$$(A - \lambda\mathcal{I})\,\mathbf{u} = \mathbf{0}$$

where $\mathcal{I}$ is the identity matrix

This equation can only have non-zero solutions for $\mathbf{u}$ if

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

This allows us to calculate the eigenvalues (if they exist)

Example of calculating eigenvalues and eigenvectors of a matrix

Calculate the eigenvalues and eigenvectors of the matrix A given by

$$A = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 2 & 3 \\ 1 & 0 & 2 \end{pmatrix}$$

The eigenvalues λ satisfy

$$\det \begin{pmatrix} 2-\lambda & 0 & 1 \\ -1 & 2-\lambda & 3 \\ 1 & 0 & 2-\lambda \end{pmatrix} = 0$$

Expanding the determinant gives

$$(2-\lambda)\left[(2-\lambda)(2-\lambda) - (3)(0)\right] + (1)\left[(-1)(0) - (2-\lambda)(1)\right] = 0$$

This is known as the characteristic equation of the matrix A

A little manipulation gives

$$\begin{aligned} (2-\lambda)\left[(2-\lambda)(2-\lambda) - 1\right] &= 0 \\ (2-\lambda)\left[\lambda^2 - 4\lambda + 3\right] &= 0 \\ (2-\lambda)(\lambda-3)(\lambda-1) &= 0 \end{aligned}$$

and so the eigenvalues of A are $\lambda = 1, 2, 3$

To find the eigenvector corresponding to the eigenvalue $\lambda = 1$ we need to find a non-zero vector $\mathbf{u}$ such that

$$(A - \mathcal{I})\mathbf{u} = \mathbf{0}$$

that is,

$$\begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 3 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Adding the first row to the second row, and then subtracting the first row from the third row gives

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

We then have $\mathbf{u}$ given by

$$\mathbf{u} = r \begin{pmatrix} -1 \\ -4 \\ 1 \end{pmatrix}$$

To find the eigenvector corresponding to the eigenvalue $\lambda = 2$ we need to find a non-zero vector $\mathbf{u}$ such that

$$(A - 2\mathcal{I})\mathbf{u} = \mathbf{0}$$

that is,

$$\begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

We then have $\mathbf{u}$ given by

$$\mathbf{u} = q \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Exercise: show that the eigenvector corresponding to $\lambda = 3$ is given by

$$\mathbf{u} = p \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

We then have the following eigenvalues and eigenvectors of A:

$$\lambda = 1, \qquad \mathbf{u} = \begin{pmatrix} -1 \\ -4 \\ 1 \end{pmatrix}$$

$$\lambda = 2, \qquad \mathbf{u} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\lambda = 3, \qquad \mathbf{u} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

The number of eigenvalues of a matrix

In the previous example, the 3×3 matrix A had 3 eigenvalues

Suppose A is a $N \times N$ matrix

Using the properties of determinants it can be shown that the characteristic equation of A — that is, $\det(A - \lambda\mathcal{I}) = 0$ — is a polynomial in λ of degree N

A can then have at most N eigenvalues

Outline proof:

The characteristic equation, $\det(A - \lambda\mathcal{I}) = 0$ may be written

$$\det\begin{pmatrix} A_{11}-\lambda & A_{12} & \dots & A_{1N} \\ A_{21} & A_{22}-\lambda & \dots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \dots & A_{NN}-\lambda \end{pmatrix}$$

The determinant is equal to a linear sum of all products containing exactly one entry from each row, and one entry from each column

The product of entries on the diagonal is a polynomial of degree N in λ, and no other contribution will have a higher degree

The characteristic polynomial is therefore a polynomial in λ of degree N

Distinct eigenvalues

Let A be a $N \times N$ matrix

Suppose A has N distinct eigenvalues, $\lambda_1, \lambda_2, \dots, \lambda_N$, with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$

Under these conditions there are two useful properties:

1. The eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$ are linearly independent
2. Define the matrices S and D by

$$S = \left(\mathbf{v}_1\ \mathbf{v}_2 \dots \mathbf{v}_N\right), \qquad D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_N \end{pmatrix}$$

We may then write $D = S^{-1}AS$, and we say "A can be diagonalised"

A couple of remarks

We know that S^{-1} exists — provided the rows of S are linearly independent it will be non-singular

If A, B, X are $N \times N$ matrices, and

$$B = X^{-1}AX$$

then A and B are known as similar matrices

Proof that the eigenvectors are linearly independent

Suppose, for scalars $\alpha_1, \alpha_2, \dots, \alpha_N$,

$$\sum_{i=1}^{N} \alpha_i \mathbf{v}_i = \mathbf{0} \tag{1}$$

We want to show that we must have $\alpha_1 = \alpha_2 = \dots = \alpha_N = 0$

Multiplying Eq. (1) by A gives

$$\sum_{i=1}^{N} \alpha_i A\mathbf{v}_i = \mathbf{0}$$

By definition of eigenvectors we may write this as

$$\sum_{i=1}^{N} \alpha_i \lambda_i \mathbf{v}_i = \mathbf{0} \tag{2}$$

Multiplying Eq. (1) by λ_N, and then subtracting Eq. (2) gives

$$\sum_{i=1}^{N} \alpha_i(\lambda_N - \lambda_i)\mathbf{v}_i = \mathbf{0}$$

The last term in this sum is zero, and so we can write

$$\sum_{i=1}^{N-1} \alpha_i(\lambda_N - \lambda_i)\mathbf{v}_i = \mathbf{0} \tag{3}$$

Multiplying Eq. (3) by A gives

$$\sum_{i=1}^{N-1} \alpha_i(\lambda_N - \lambda_i)A\mathbf{v}_i = \mathbf{0}$$

$$\text{and so } \sum_{i=1}^{N-1} \alpha_i(\lambda_N - \lambda_i)\lambda_i\mathbf{v}_i = \mathbf{0} \tag{4}$$

Multiplying Eq. (3) by λ_{N-1}, and then subtracting Eq. (4) gives

$$\sum_{i=1}^{N-2} \alpha_i(\lambda_N - \lambda_i)(\lambda_{N-1} - \lambda_i)\mathbf{v}_i = \mathbf{0} \tag{5}$$

Note that the sum is from $i = 1$ to $i = N - 2$

If we repeat this procedure — multiply Eq. (5) by A, use the definition of eigenvectors, etc. — we obtain

$$\sum_{i=1}^{N-3} \alpha_i(\lambda_N - \lambda_i)(\lambda_{N-1} - \lambda_i)(\lambda_{N-2} - \lambda_i)\mathbf{v}_i = \mathbf{0}$$

If we keep going with this procedure we will eventually obtain

$$\sum_{i=1}^{k} \alpha_i(\lambda_N - \lambda_i)(\lambda_{N-1} - \lambda_i)\ldots(\lambda_{k+1} - \lambda_i)\mathbf{v}_i = \mathbf{0} \tag{6}$$

$$\vdots \quad \vdots$$

$$\sum_{i=1}^{2} \alpha_i(\lambda_N - \lambda_i)(\lambda_{N-1} - \lambda_i)\ldots(\lambda_3 - \lambda_i)\mathbf{v}_i = \mathbf{0} \tag{7}$$

$$\alpha_1(\lambda_N - \lambda_i)(\lambda_{N-1} - \lambda_i)\ldots(\lambda_3 - \lambda_i)(\lambda_2 - \lambda_i)\mathbf{v}_1 = \mathbf{0} \tag{8}$$

All the eigenvalues, λ_i are distinct.

None of the terms involving λ_i in Eq. (8) can be zero

$\mathbf{v}_1$ is an eigenvector, and so $\mathbf{v}_1 \neq \mathbf{0}$

We must then have $\alpha_1 = 0$

Eq. (7) then becomes

$$\alpha_2(\lambda_N - \lambda_2)(\lambda_{N-1} - \lambda_2)\ldots(\lambda_3 - \lambda_2)\mathbf{v}_2 = \mathbf{0}$$

Using the same argument as above $\alpha_2 = 0$

If we keep going we eventually find that

$$\alpha_1 = \alpha_2 = \ldots = \alpha_N = 0$$

and so the eigenvectors are linearly independent, as required.

Proof that A can be diagonalised

Define the matrices S and D by

$$S = \begin{pmatrix}\mathbf{v}_1 \; \mathbf{v}_2 \ldots \mathbf{v}_N\end{pmatrix}, \qquad D = \begin{pmatrix}\lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_N\end{pmatrix}$$

We have shown that the eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$ are linearly independent

The matrix S is therefore non-singular, and the inverse S^{-1} exists

Writing S^{-1} as

$$S^{-1} = \begin{pmatrix}\mathbf{w}_1^\top \\ \mathbf{w}_2^\top \\ \vdots \\ \mathbf{w}_N^\top\end{pmatrix}$$

As $S^{-1}S = \mathcal{I}$, we must have

$$\mathbf{w}_i \cdot \mathbf{v}_j = \delta_{ij} = \begin{cases}1, & i = j \\ 0, & i \neq j\end{cases}$$

Using the definition of eigenvectors we may write

$$\begin{aligned} AS &= A\left(\mathbf{v}_1\ \mathbf{v}_2 \ldots \mathbf{v}_N\right) \\ &= \left(A\mathbf{v}_1\ A\mathbf{v}_2 \ldots A\mathbf{v}_N\right) \\ &= \left(\lambda_1\mathbf{v}_1\ \lambda_2\mathbf{v}_2 \ldots \lambda_N\mathbf{v}_N\right) \end{aligned}$$

We then have

$$S^{-1}AS = \begin{pmatrix} \mathbf{w}_1^\top \\ \mathbf{w}_2^\top \\ \vdots \\ \mathbf{w}_N^\top \end{pmatrix} \left(\lambda_1\mathbf{v}_1\ \lambda_2\mathbf{v}_2 \ldots \lambda_N\mathbf{v}_N\right)$$

$$= \begin{pmatrix} \lambda_1\mathbf{w}_1\cdot\mathbf{v}_1 & \lambda_2\mathbf{w}_1\cdot\mathbf{v}_2 & \ldots & \lambda_N\mathbf{w}_1\cdot\mathbf{v}_N \\ \lambda_1\mathbf{w}_2\cdot\mathbf{v}_1 & \lambda_2\mathbf{w}_2\cdot\mathbf{v}_2 & \ldots & \lambda_N\mathbf{w}_2\cdot\mathbf{v}_N \\ \vdots & \vdots & & \vdots \\ \lambda_1\mathbf{w}_N\cdot\mathbf{v}_1 & \lambda_2\mathbf{w}_N\cdot\mathbf{v}_2 & \ldots & \lambda_N\mathbf{w}_N\cdot\mathbf{v}_N \end{pmatrix}$$

Hence,

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_N \end{pmatrix}$$

We then have $D = S^{-1}AS$ as required

Example: diagonalising a matrix

Suppose the matrix A is given by

$$A = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 2 & 3 \\ 1 & 0 & 2 \end{pmatrix}$$

Find matrices D and S such that $D = S^{-1}AS$, and D is a diagonal matrix

We calculated the eigenvalues and eigenvectors of A earlier.

The eigenvalues are distinct, and so matrices D and S will exist

We have

$$\lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = 3$$

and

$$\mathbf{v}_1 = \begin{pmatrix} -1 \\ -4 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

We may then write

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

$$S = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 \\ -4 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$$

We can then verify that $D = S^{-1}AS$ as required

What happens when the eigenvalues are not distinct?

We have shown that, if A is a $N \times N$ matrix with N distinct eigenvalues, we can find an invertible matrix S and a diagonal matrix D such that $D = S^{-1}AS$

What happens when N distinct eigenvalues don't exist?

Case 1: repeated eigenvalues

Suppose A is the identity matrix, so that

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The characteristic equation for A is

$$(\lambda - 1)^2 = 0$$

and so we have the repeated roots $\lambda = 1, 1$

The eigenvectors then satisfy

$$\begin{pmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{pmatrix} \mathbf{u} = \mathbf{0}, \qquad \text{and so} \qquad \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{u} = \mathbf{0}$$

and so every vector is an eigenvector

Choosing

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

we then set

$$S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and then $D = S^{-1}AS$

Our choice of eigenvectors on the previous slide was arbitrary

Any two linearly independent eigenvectors would have allowed us to diagonalise A

In this case A did not have N distinct eigenvalues, but A could be diagonalised

Suppose B is the matrix given by

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

The characteristic equation for B is

$$(\lambda - 1)^2 = 0$$

and so we again have the repeated roots $\lambda = 1, 1$

The eigenvectors then satisfy

$$\begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \mathbf{0}, \qquad \text{and so} \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \mathbf{0}$$

We must have $q = 0$, and so only one linearly independent eigenvector exists,

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

In this case B does not have N distinct eigenvalues, and B can't be diagonalised

Case 2: fewer than N eigenvalues exist

Suppose C is the matrix given by

$$C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

The characteristic equation for C is

$$\lambda^2 + 1 = 0$$

and so no real eigenvalues exist

We then can't diagonalise C

Symmetric matrices

Let A be a real, symmetric matrix of size $N \times N$, that is $A = A^\top$

The matrix A will then have the following properties

- A will have N real eigenvalues (possibly including repeated eigenvalues)
- Eigenvectors corresponding to different eigenvalues are orthogonal
- A diagonal matrix D, and a matrix P exist such that

$$P^\top P = \mathcal{I}, \quad D = P^\top AP$$

where the columns of P are normalised eigenvectors of A

A matrix P that satisfies $P^\top P = \mathcal{I}$ is known as an orthogonal matrix

Suppose λ and μ are (possibly complex) eigenvalues of A

We then have

$$A\mathbf{v} = \lambda\mathbf{v}, \quad A\mathbf{w} = \mu\mathbf{w}$$

for eigenvectors $\mathbf{v}$ and $\mathbf{w}$

We may then write

$$\begin{aligned} \mu\mathbf{v}^\top\mathbf{w} &= \mathbf{v}^\top(\mu\mathbf{w}) \\ &= \mathbf{v}^\top A\mathbf{w} \\ &= \mathbf{v}^\top A^\top\mathbf{w} \quad \text{as } A \text{ is symmetric} \\ &= (A\mathbf{v})^\top\mathbf{w} \\ &= \lambda\mathbf{v}^\top\mathbf{w} \end{aligned}$$

We then have

$$(\mu - \lambda)\mathbf{v}^\top\mathbf{w} = 0$$

which may be written

$$(\mu - \lambda)\mathbf{v}\cdot\mathbf{w} = 0$$

If $\mu \neq \lambda$ we then have

$$\mathbf{v}\cdot\mathbf{w} = 0$$

and so the eigenvectors corresponding to λ and μ are orthogonal, as required

Now, suppose λ is complex. We then have

$$A\mathbf{v} = \lambda\mathbf{v}, \quad A\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$$

where $\bar{\lambda}$ denotes the complex conjugate of λ

We can then set $\mu = \bar{\lambda}$ and $\mathbf{w} = \bar{\mathbf{v}}$ in the analysis on the previous two slides to give

$$(\bar{\lambda} - \lambda)\mathbf{v} \cdot \bar{\mathbf{v}} = 0$$

Noting that $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \bar{\mathbf{v}}$, and that this cannot be zero as $\mathbf{v}$ is an eigenvector, we deduce that $\bar{\lambda} = \lambda$

In other words, λ must be real

We will now prove that a diagonal matrix D, and a matrix P exist such that

$$P^\top P = \mathcal{I}, \quad D = P^\top AP$$

We will prove this by induction on N, the size of the matrix

When $N = 1$ we have $A = \begin{pmatrix} A_{11} \end{pmatrix}$, and then we may write

$$D = \begin{pmatrix} A_{11} \end{pmatrix}, \qquad P = \begin{pmatrix} 1 \end{pmatrix}$$

In this case we then have $P^\top P = \mathcal{I}$ and $D = P^\top AP$

We will now assume that the claim is true for matrices of size $(N-1) \times (N-1)$, and will show that this implies it is true for matrices of size $N \times N$

Let A be a symmetric matrix of size $N \times N$

Let λ_1 be an eigenvalue of A, with eigenvector $\mathbf{v}_1$, where $\|\mathbf{v}_1\| = 1$.

Extend to an orthonormal basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$ for $\Re^n$, so that

$$\mathbf{v}_i \cdot \mathbf{v}_j = \mathbf{v}_i^\top \mathbf{v}_j = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & \text{otherwise} \end{cases}$$

Define the matrix Q by

$$Q = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_N \end{pmatrix}$$

As $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$ is an orthonormal basis,

$$Q^\top Q = \mathcal{I}$$

We also have

$$Q^\top A Q = \begin{pmatrix} \mathbf{v}_1^\top \\ \mathbf{v}_2^\top \\ \vdots \\ \mathbf{v}_N^\top \end{pmatrix} \begin{pmatrix} A\mathbf{v}_1 & A\mathbf{v}_2 & \dots & A\mathbf{v}_N \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{v}_1^\top \\ \mathbf{v}_2^\top \\ \vdots \\ \mathbf{v}_N^\top \end{pmatrix} \begin{pmatrix} \lambda_1\mathbf{v}_1 & A\mathbf{v}_2 & \dots & A\mathbf{v}_N \end{pmatrix}$$

This becomes

$$Q^\top A Q = \begin{pmatrix} \lambda_1 & \mathbf{w}^\top \\ \mathbf{0} & B \end{pmatrix}$$

where B is a $(N-1) \times (N-1)$ matrix

Note that

$$\left(Q^\top A Q\right)^\top = Q^\top A^\top \left(Q^\top\right)^\top$$
$$= Q^\top A Q$$

We then deduce that $Q^\top A Q$ is symmetric, and so $\mathbf{w} = \mathbf{0}$, and $B = B^\top$

As B is symmetric and of size $(N-1) \times (N-1)$, by the induction hypothesis there exists an orthogonal matrix $\hat{Q}$, of size $(N-1) \times (N-1)$, and such that

$$\hat{Q}^\top B \hat{Q} = \begin{pmatrix} \lambda_2 & 0 & \dots & 0 \\ 0 & \lambda_3 & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & \lambda_N \end{pmatrix}$$

Define R by

$$R = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{Q} \end{pmatrix}$$

Then R is an orthogonal matrix

Set $P = QR$

We then have

$$P^\top P = R^\top Q^\top Q R$$
$$= R^\top R$$
$$= \mathcal{I}$$

Hence, P is an orthogonal matrix

We then have

$$
\begin{aligned}
P^\top AP &= (QR)^\top AQR = R^\top Q^\top AQR \\
&= R^\top \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & B \end{pmatrix} R \\
&= \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{Q}^\top \end{pmatrix} \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & B \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{Q} \end{pmatrix} = \begin{pmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & \hat{Q}^\top B \hat{Q} \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & \lambda_N \end{pmatrix}
\end{aligned}
$$

as required

Example

The matrix A is given by

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

The eigenvalues satisfy

$$
\det \begin{pmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{pmatrix} = 0
$$

Some manipulation gives

$$
\lambda^3 - 6\lambda^2 + 9\lambda - 4 = 0
$$

By inspection $\lambda = 1$ is an eigenvalue and so

$$
\begin{aligned}
(\lambda - 1)(\lambda^2 - 5\lambda + 4) &= 0 \\
(\lambda - 1)(\lambda - 1)(\lambda - 4) &= 0
\end{aligned}
$$

Hence the eigenvalues are $\lambda = 1, 1, 4$

When $\lambda = 1$ the eigenvectors satisfy

$$
\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
$$

and so $p + q + r = 0$.

Two orthonormal eigenvectors are

$$
\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}
$$

When $\lambda = 4$ the eigenvectors satisfy

$$\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

A normalised eigenvector is

$$\mathbf{v}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

We then write

$$P = \begin{pmatrix} \mathbf{v}_1 & v_2 & \mathbf{v}_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

Then $D = P^\top AP$, and $P^\top P = \mathcal{I}$ as required

Note that, as $P^{-1} = P^\top$, we could write this as $D = P^{-1}AP$

Powers of matrices

Let A be a matrix

We will see later in the course that we often want to calculate powers of a matrix, for example A^9, or the limit as n tends to infinity of A^n

If A can be diagonalised — that is, we can write $D = S^{-1}AS$ for a diagonal matrix D and known invertible matrix S — we can easily calculate any power of a matrix

Writing $A = SDS^{-1}$ we see that

$$\begin{aligned} A^2 &= SDS^{-1}SDS^{-1} \\ &= SD^2S^{-1} \end{aligned}$$

and

$$\begin{aligned} A^3 &= A^2A \\ &= SD^2S^{-1}SDS^{-1} \\ &= SD^3S^{-1} \end{aligned}$$

Clearly

$$A^n = SD^nS^{-1}$$

This can be proved by induction

Example: If A is given by

$$A = \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix}$$

calculate A^n for $n = 2, 3, 4, \ldots$

The eigenvalues of A are given by

$$\begin{aligned} \det(A - \lambda\mathcal{I}) &= 0 \\ (0.8 - \lambda)(0.9 - \lambda) - (0.1)(0.2) &= 0 \\ \lambda^2 - 1.7\lambda + 0.7 &= 0 \end{aligned}$$

and so $\lambda = 1, 0.7$

A has two distinct eigenvalues, and so it can be diagonalised

The eigenvector corresponding to $\lambda = 1$ satisfies

$$\begin{pmatrix} -0.2 & 0.1 \\ 0.2 & -0.1 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which is satisfied for $p = 1, q = 2$

The eigenvector corresponding to $\lambda = 0.7$ satisfies

$$\begin{pmatrix} 0.1 & 0.1 \\ 0.2 & 0.2 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which is satisfied for $s = 1, t = -1$

We may therefore write $D = S^{-1}AS$ where

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 0.7 \end{pmatrix}, \qquad S = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$$

Note that

$$D^n = \begin{pmatrix} 1 & 0 \\ 0 & 0.7^n \end{pmatrix}$$

We then have

$$\begin{aligned} A^n &= SD^nS^{-1} \\ &= S \begin{pmatrix} 1 & 0 \\ 0 & 0.7^n \end{pmatrix} S^{-1} \end{aligned}$$

In the limit that $n \to \infty$, the quantity $0.7^n \to 0$.

We then have

$$A^n \to S \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} S^{-1}$$

where

$$S^{-1} = \begin{pmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{pmatrix}$$

Markov Chains

In a given year, a person may have a normal cholesterol level, or a high cholesterol level

Someone with a high cholesterol level in a given year has a 20% chance of losing high cholesterol (and, therefore, is 80% likely to remain with high cholesterol)

Someone with a normal cholesterol level in a given year has a 10% chance of gaining high cholesterol (and, therefore, is 90% likely to remain with normal cholesterol)

Suppose we know the probability that a given person has high cholesterol in a given year. We may then use linear algebra to calculate the probability that this person has high cholesterol in future years.

Define:

- X_k to be the event that the person has a high cholesterol level in year k
- Y_k to be the event that the person has a normal cholesterol level in year k

The law of total probability then tells us that

$$P(X_k) = P(X_k|A)P(A) + P(X_k|A^c)P(A^c)$$

If we let A be the event that the person has a high cholesterol level in year $k-1$

A^c is then the event that the person has a normal cholesterol level in year $k-1$

We then have

$$\begin{aligned} P(X_k|A) &= 0.8 \\ P(A) &= P(X_{k-1}) \\ P(X_k|A^c) &= 0.1 \\ P(A^c) &= P(Y_{k-1}) \end{aligned}$$

and so

$$P(X_k) = 0.8P(X_{k-1}) + 0.1P(Y_{k-1})$$

Similarly,

$$P(Y_k) = P(Y_k|A)P(A) + P(Y_k|A^c)P(A^c)$$

with

$$\begin{aligned} P(Y_k|A) &= 0.2 \\ P(A) &= P(X_{k-1}) \\ P(Y_k|A^c) &= 0.9 \\ P(A^c) &= P(Y_{k-1}) \end{aligned}$$

and so

$$P(Y_k) = 0.2P(X_{k-1}) + 0.9P(Y_{k-1})$$

Writing

$$x_k = P(X_k), \qquad y_k = P(Y_k)$$

and

$$\begin{aligned}\mathbf{x}_k &= \begin{pmatrix} x_k \\ y_k \end{pmatrix} \\ &= \begin{pmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{pmatrix} \begin{pmatrix} x_{k-1} \\ y_{k-1} \end{pmatrix} \\ &= A\mathbf{x}_{k-1} \\ &= A^k \mathbf{x}_0\end{aligned}$$

The matrix A is the same matrix we used earlier when looking at powers of a matrix

Using the powers of A derived earlier, we may write

$$\mathbf{x}_k = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0.7^k \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{pmatrix} \mathbf{x}_0$$

As $k \to \infty$ we obtain

$$\begin{aligned}\mathbf{x}_\infty &= \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{pmatrix} \mathbf{x}_0 \\ &= \begin{pmatrix} 1/3 & 1/3 \\ 2/3 & 2/3 \end{pmatrix} \mathbf{x}_0\end{aligned}$$

Another way to view the Markov chain on previous slides is to write the initial vector $\mathbf{x}_0$ in terms of the basis of eigenvectors:

$$\mathbf{x}_0 = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2$$

We then have

$$\begin{aligned}A\mathbf{x}_0 &= A(\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2) \\ &= \alpha_1 A\mathbf{v}_1 + \alpha_2 A\mathbf{v}_2 \\ &= \alpha_1 \lambda_1 \mathbf{v}_1 + \alpha_2 \lambda_2 \mathbf{v}_2\end{aligned}$$

Similarly,

$$\begin{aligned}A^2\mathbf{x}_0 &= \alpha_1 \lambda_1^2 \mathbf{v}_1 + \alpha_2 \lambda_2^2 \mathbf{v}_2 \\ A^3\mathbf{x}_0 &= \alpha_1 \lambda_1^3 \mathbf{v}_1 + \alpha_2 \lambda_2^3 \mathbf{v}_2 \\ \vdots \quad & \vdots \\ A^k\mathbf{x}_0 &= \alpha_1 \lambda_1^k \mathbf{v}_1 + \alpha_2 \lambda_2^k \mathbf{v}_2\end{aligned}$$

For our example $\lambda_1 = 1$ and $\lambda_2 = 0.7$ and so, as $k \to \infty$, we have

$$A^k \mathbf{x}_0 \to \alpha_1 \mathbf{v}_1$$

We need to calculate α_1 to use the result on the previous slide

Recall that

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

If

$$\mathbf{x}_0 = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2$$

then

$$\mathbf{x}_0 \cdot \mathbf{v}_1 = \alpha_1 \mathbf{v}_1 \cdot \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 \cdot \mathbf{v}_1$$
$$\mathbf{x}_0 \cdot \mathbf{v}_2 = \alpha_1 \mathbf{v}_1 \cdot \mathbf{v}_2 + \alpha_2 \mathbf{v}_2 \cdot \mathbf{v}_2$$

We can write this as the linear system

$$\begin{pmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \mathbf{v}_2 \cdot \mathbf{v}_1 \\ \mathbf{v}_1 \cdot \mathbf{v}_2 & \mathbf{v}_2 \cdot \mathbf{v}_2 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \mathbf{x}_0 \cdot \mathbf{v}_1 \\ \mathbf{x}_0 \cdot \mathbf{v}_2 \end{pmatrix}$$

allowing us to evaluate α_1 (and α_2)

The power method

The ideas on the previous slides may be extended to estimate the largest eigenvalue (assumed unique) and corresponding eigenvector

Let $\mathbf{x}_0$ be an arbitrary vector. For $k = 1, 2, 3, \ldots$ define

$$\mathbf{y} = A\mathbf{x}_{k-1}, \qquad M = \max_i |y_i|$$
$$\mathbf{x}_k = \frac{1}{M}\mathbf{y}$$

As $k \to \infty$, $\mathbf{x}_k$ approaches the eigenvector corresponding to the largest eigenvalue. This eigenvalue may be estimated by the Raleigh quotient

$$\lambda_{\max} = \frac{\mathbf{x}_k^\top A \mathbf{x}_k}{\mathbf{x}_k^\top \mathbf{x}_k}$$

Proof

Assume that the eigenvectors of A form a basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$, where λ_1 is the eigenvalue with the largest modulus

We may write $\mathbf{x}_0$ in terms of the basis of eigenvectors

Write

$$\mathbf{x}_0 = \sum_{i=1}^{N} \alpha_i \mathbf{v}_i$$

We then have

$$
\begin{aligned}
\mathbf{x}_1 &= L_1 \lambda_1 \sum_{i=1}^{N} \alpha_i \frac{\lambda_i}{\lambda_1} \mathbf{v}_i \\
\mathbf{x}_2 &= L_2 \lambda_1 \sum_{i=1}^{N} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^2 \mathbf{v}_i \\
\vdots \quad & \vdots \\
\mathbf{x}_k &= L_k \lambda_1 \sum_{i=1}^{N} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \mathbf{v}_i
\end{aligned}
$$

Recall that, for $i = 2, 3, \ldots, N$, we have assumed that $\frac{\lambda_i}{\lambda_1} < 1$.

As $k \to \infty$ the sum on the previous slide then becomes

$$\mathbf{x}_k \to L_k \lambda_1 \alpha_1 \mathbf{v}_1$$

which is a multiple of the eigenvector corresponding to λ_1

We then have, as $k \to \infty$,

$$
\begin{aligned}
\frac{\mathbf{x}_k^\top A \mathbf{x}_k}{\mathbf{x}_k^\top \mathbf{x}_k} \to \frac{\mathbf{v}_1^\top A \mathbf{v}_1}{\mathbf{v}_1^\top \mathbf{v}_1}
&= \frac{\mathbf{v}_1^\top \lambda_1 \mathbf{v}_1}{\mathbf{v}_1^\top \mathbf{v}_1} \\
&= \lambda_1 \frac{\mathbf{v}_1^\top \mathbf{v}_1}{\mathbf{v}_1^\top \mathbf{v}_1} \\
&= \lambda_1
\end{aligned}
$$

A potential flaw with this method is that the coefficient of $\mathbf{v}_1$ in our initial vector $\mathbf{x}_0$, that is α_1 may take the value zero

Rounding errors, normally a potential source of problems, come to our aid on this occasion

The rounding errors will introduce some component of $\mathbf{v}_1$ into $\mathbf{x}_k$, and this will ultimately dominate the computation

Positive and non-negative matrices

The previous examples show that the eigenvalue with largest magnitude can play an important role

The following two theorems extend these ideas

Perron's theorem

Let A be a $N \times N$ matrix with all entries positive. We will refer to this as a positive matrix

The matrix A has a real eigenvalue λ_1 satisfying

1. $\lambda_1 > 0$
2. λ_1 has a corresponding positive eigenvector
3. For any other eigenvalue λ of A we have $|\lambda| \leq \lambda_1$

Recall that the Markov matrix from the last section is a positive matrix

Reducible matrices

A matrix A is called reducible if, subject to some permutation of the rows, P, and the same permutation of the columns, $P^\top$, A can be written in the block form

$$PAP^T = \begin{bmatrix} B & C \\ O & D \end{bmatrix},$$

where B and D are square matrices. Otherwise the matrix is said to be irreducible.

Perron-Frobenius theorem

Let A be a $N \times N$ irreducible matrix with all entries non-negative. Then A has a unique real eigenvalue λ_1 satisfying

1. $\lambda_1 > 0$
2. λ_1 has a corresponding positive eigenvector.
3. For any other eigenvalue λ of A we have $|\lambda| \leq \lambda_1$.
4. If $|\lambda| = \lambda_1$ then λ is a complex root of the equation $\lambda^n - \lambda_1^n = 0$.

Ranking Sports Teams

Five squash players play each other in a round robin tournament. The following table of results was recorded

$$\begin{array}{c} \\ P1 \\ P2 \\ P3 \\ P4 \\ P5 \end{array} \begin{bmatrix} P1 & P2 & P3 & P4 & P5 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

You may assume that the matrix above is irreducible

We would now like to rank the players, r_i, in such a way that $r_i > r_j$ indicates that player i is ranked higher than player j

We will let the r_i's be "probabilities"

That is $0 \leq r_i \leq 1$ and $\sum_{i=1}^{5} r_i = 1$ and have the corresponding ranking vector, $\mathbf{r} = [r_1, r_2, r_3, r_4, r_5]^T$.

We will assume that player i's ranking be proportional, with constant of proportionality α, to the sum of the rankings of the players defeated by i

This allows us to write

$$r_1 = \alpha(r_2 + r_4 + r_5), \quad \text{etc}$$

We may write this in the form

$$\begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{bmatrix} = \alpha \begin{bmatrix} 0&1&0&1&1 \\ 0&0&1&1&1 \\ 1&0&0&1&0 \\ 0&0&0&0&1 \\ 0&0&1&0&0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{bmatrix}, \quad A\mathbf{r} = \frac{1}{\alpha}\mathbf{r}$$

That is, $\mathbf{r}$ is an eigenvector of the matrix A, with eigenvalue $\lambda = 1/\alpha$

Using the Perron-Frobenius Theorem we have a guaranteed unique ranking vector associated with the largest eigenvalue $\frac{1}{\alpha}$ of A

$$\begin{bmatrix} 0.29 \\ 0.27 \\ 0.22 \\ 0.08 \\ 0.14 \end{bmatrix}$$

We see from the results that both players 1 and 2 beat three players, but player 1 beat player 2 giving them a slightly better ranking

Internet Search Engines

When using an internet search engine we want the search results returned in a sensible order

Suppose we have five results returned from a search and if the i-th results references the $j-$th then we place a 1 in the (i, j) entry of a matrix A, for example

$$\begin{array}{c} \\ S1 \\ S2 \\ S3 \\ S4 \\ S5 \end{array} \left[\begin{array}{ccccc} S1 & S2 & S3 & S4 & S5 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right] := A$$

You may assume that the matrix above is irreducible

In this $S3$ refers to $S1$, $S2$ refers to $S3$ and so on

We now want the ranking of i to be proportional to the sum of the rankings that refer to i e.g.

$$r_4 = \alpha(r_1 + r_2 + r_3).$$

This leads to the eigenvalue problem find $\mathbf{r}$ such that

$$A^T\mathbf{r} = \frac{1}{\alpha}\mathbf{r}$$

In this case we have

$$\begin{bmatrix} 0.14 \\ 0.08 \\ 0.22 \\ 0.27 \\ 0.29 \end{bmatrix}.$$

The fifth result is ranked top. Like the fourth result it is referenced three times in total, but the fourth references include the fifth giving it the edge

Differential equations

Suppose we want to find the general solution of the system of differential equations given by

$$\frac{\mathrm{d}y_1}{\mathrm{d}t} = 2y_1 + y_3$$
$$\frac{\mathrm{d}y_2}{\mathrm{d}t} = -y_1 + 2y_2 + 3y_3$$
$$\frac{\mathrm{d}y_3}{\mathrm{d}t} = y_1 + 2y_3$$

We may write this in matrix form as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 2 & 3 \\ 1 & 0 & 2 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

In an earlier lecture we saw that we could write the matrix on the previous slide as SDS^{-1} where

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

$$S = \begin{pmatrix} -1 & 0 & 1 \\ -4 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$$

We may then write the system of differential equations as

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}t} = SDS^{-1}\mathbf{y}$$

As the entries of S are constant this may be written

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(S^{-1}\mathbf{y}\right) = DS^{-1}\mathbf{y}$$

Setting $\mathbf{z} = S^{-1}\mathbf{y}$, we obtain

$$\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}t} = D\mathbf{z}$$

which may be written

$$\frac{\mathrm{d}z_1}{\mathrm{d}t} = z_1, \qquad \frac{\mathrm{d}z_2}{\mathrm{d}t} = 2z_2, \qquad \frac{\mathrm{d}z_3}{\mathrm{d}t} = 3z_3$$

We then have, for arbitrary constants A, B, C:

$$z_1 = A\mathrm{e}^{t}, \qquad z_2 = B\mathrm{e}^{2t}, \qquad z_3 = C\mathrm{e}^{3t}$$

Finally,

$$\begin{aligned} \mathbf{y} &= S\mathbf{z} \\ &= \begin{pmatrix} -1 & 0 & 1 \\ -4 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} A\mathrm{e}^{t} \\ B\mathrm{e}^{2t} \\ C\mathrm{e}^{3t} \end{pmatrix} \\ &= \begin{pmatrix} -A\mathrm{e}^{t} + C\mathrm{e}^{3t} \\ -4A\mathrm{e}^{t} + B\mathrm{e}^{2t} + 2C\mathrm{e}^{3t} \\ A\mathrm{e}^{t} + C\mathrm{e}^{3t} \end{pmatrix} \end{aligned}$$