CHAPTER 4

EIGENTHINGS

1. Some matrix algebra

It is easy to compute powers, inverses and other functions of diagonal matrices.

Example.

\[
\begin{pmatrix}
2 & 0 \\
0 & 3
\end{pmatrix}^2 = \begin{pmatrix}
4 & 0 \\
0 & 9
\end{pmatrix} = \begin{pmatrix}
2^2 & 0 \\
0 & 3^2
\end{pmatrix}
\]

\[
\begin{pmatrix}
-1 & 0 \\
0 & \frac{1}{2}
\end{pmatrix}^{-1} = \begin{pmatrix}
-1 & 0 \\
0 & 2
\end{pmatrix} = \begin{pmatrix}
(-1)^{-1} & 0 \\
0 & (\frac{1}{2})^{-1}
\end{pmatrix}
\]

\[
\begin{pmatrix}
1 & 0 \\
0 & 4
\end{pmatrix}^6 = \begin{pmatrix}
1 & 0 \\
0 & 4096
\end{pmatrix} = \begin{pmatrix}
1^6 & 0 \\
0 & 4^6
\end{pmatrix}
\]

Let \( \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \) denote the \( n \times n \) diagonal matrix

\[
\begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \lambda_n
\end{pmatrix}
\]

Exercise 13. Show that

\[
\text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \text{ diag}(\mu_1, \mu_2, \ldots, \mu_n) = \text{diag}(\lambda_1 \mu_1, \lambda_2 \mu_2, \ldots, \lambda_n \mu_n),
\]

and that

\[
\text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)^{-1} = \text{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1}).
\]

Theorem 25. For any integer \( m \), \( \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)^m = \text{diag}(\lambda_1^m, \lambda_2^m, \ldots, \lambda_n^m) \).

Proof. For positive \( m \), use induction and Exercise 13. For negative \( m \), use the fact that \( A^{-m} = (A^{m})^{-1} \) and the exercise.

We can take this further: for a matrix \( A \), we define \( e^A \) or \( \exp(A) \) by

\[
\exp(A) = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots.
\]

Then it follows that

\[
\exp(\text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)) = \text{diag}(e^{\lambda_1}, e^{\lambda_2}, \ldots, e^{\lambda_n}).
\]

In summary, diagonal matrices make for easy computations.
Now take a more general matrix $A$, and suppose that $A = MDM^{-1}$, where $D$ is diagonal and $M$ is invertible. Then

$$A^2 = MDM^{-1}MDM^{-1} = MD^2M^{-1}.$$ 

More generally, if $m \in \mathbb{Z}$,

$$A^m = MD^mM^{-1},$$

and $\exp(A) = M \exp(D)M^{-1}$. Thus, calculations with $A$ are quite easy too—very few matrix multiplications are needed. It is now worth asking whether, given a matrix $A$, we can find a diagonal matrix $D$ and an invertible matrix $M$ so that $A = MDM^{-1}$.

### 2. Geometry of linear transformations

It is easy to understand the effect of multiplying a vector by a diagonal matrix. For instance, suppose that

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

We can understand the effect of multiplying a vector $v$ by $D$ by resolving the vector into components parallel to the axes.

If we consider a rectangle $R$ with sides parallel to the axes, then $DR$, the set of all $Dv$ where $v \in R$, is also a rectangle with sides parallel to the axes.
The lengths of the sides increase by factors of 2 and 3.

Now suppose that $A = MDM^{-1}$, where

$$M = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$ 

Multiplication by $M$ rotates vectors by $\pi/4$ (anticlockwise). If we take a rectangle $R$ with sides parallel to $y = x$ and $y = -x$, then multiplying $R$ by $A = MDM^{-1}$ first rotates $R$ so that the sides are parallel to the axes, then increases the size of the rotated rectangle, and then rotates the rectangle back again to have sides parallel to the lines $y = \pm x$.

Thus, we have some geometric understanding of the effect of multiplying by $A$.

In this example, we can check that

$$A = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 5/2 & -1/2 \\ -1/2 & 5/2 \end{pmatrix}.$$ 

If we were just given the matrix $A$, and asked to find a diagonal matrix $D$ and an invertible matrix $M$ so that $A = MDM^{-1}$, it would not be obvious how to proceed. But the geometry gives us a hint. When we multiply vectors by $D$, we change their direction, as well as their magnitude, except when we consider vectors parallel to the axes. Similarly, when we multiply vectors by $A$, we change their directions except when we have vectors parallel to $y = \pm x$. The effect of multiplying by $M^{-1}$ is to move the lines $y = \pm x$ onto the coordinate axes, and multiplying by $M$ takes us back again.

The key idea is that, to find $M$ and $D$, we will look for vectors $v$ so that $Av = \lambda v$ for some $\lambda$. These will give the axes which have to be moved onto the usual coordinate axes. A nonzero vector $v$ such that $Av = \lambda v$ for some $\lambda$ is called an eigenvector for $A$; $\lambda$ is called the associated eigenvalue.

We consider another example. Suppose that

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad \text{and} \quad M = \begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix}.$$ 

and \( A = \text{MDM}^{-1} \). We will describe geometrically the effect of multiplication by \( A \).

In light of the previous examples, we might hope that there are two axes in \( \mathbb{R}^2 \), probably not the coordinate axes, so that multiplication by \( A \) preserves these axes, and multiplication by \( A \) sends geometrical objects with sides parallel to these axes into other geometrical objects with sides parallel to the axes. Let us verify this.

In the previous example, where \( M \) was a rotation, the “natural axes” for \( A \) were obtained by rotating the coordinate axes. In this example too, the “natural axes” are obtained by multiplying the coordinate axes by \( M \). Write \( v_1 = Me_1 \) and \( v_2 = Me_2 \), where \( e_1 \) and \( e_2 \) are the standard basis vectors. Then

\[
Av_1 = (\text{MDM}^{-1}) Me_1 \\
= MDe_1 \\
= M(2e_1) \\
= 2Me_1 \\
= 2v_1.
\]

and similarly \( Av_2 = 3v_2 \).

The vectors of the form

\[ x_1 v_1 + x_2 v_2 \]

where \( c_1 \leq x_1 \leq d_1 \) and \( c_2 \leq x_2 \leq d_2 \) (\( x_1, x_2 \) vary, while \( c_1, c_2, d_1, d_2 \) are fixed) make up a parallelogram, with sides parallel to \( v_1 \) and \( v_2 \). When we multiply the vectors in this parallelogram by \( M^{-1} \), we get the rectangle of vectors of the form \( x_1 e_1 + x_2 e_2 \), where \( c_1 \leq x_1 \leq d_1 \) and \( c_2 \leq x_2 \leq d_2 \). When we multiply by \( D \), we get the rectangle of vectors of the form \( 2x_1 e_1 + 3x_2 e_2 \), where \( x_1 \) and \( x_2 \) vary as before. Finally, when we multiply by \( M \), we obtain the parallelogram of vectors of the form \( 2x_1 v_1 + 3x_2 v_2 \), where \( x_1 \) and \( x_2 \) vary as before. Thus multiplication by \( A \) maps parallelograms with sides parallel to \( v_1 \) and \( v_2 \) into parallelograms with sides parallel to \( v_1 \) and \( v_2 \) as before, but the sides parallel to \( v_1 \) are expanded by a factor of 2 and those parallel to \( v_2 \) are expanded by factor of 3. We could have seen this more directly, since

\[
A(x_1 v_1 + x_2 v_2) = x_1 Av_1 + x_2 Av_2 = 2x_1 v_1 + 3x_2 v_2.
\]

Thus, to understand the geometry, the key fact is that \( Av_1 = 2v_1 \) and \( Av_2 = 3v_2 \).

Are there any other nonzero vectors \( v \) so that \( Av = \lambda v \) for some scalar \( \lambda \). No, since any vector \( v \) can be written in the form \( x_1 v_1 + x_2 v_2 \), so that \( Av = 2x_1 v_1 + 3x_2 v_2 \). If \( Av = \lambda v \), then

\[
2x_1 v_1 + 3x_2 v_2 = \lambda(x_1 v_1 + x_2 v_2).
\]

Thus

\[
(2 - \lambda)x_1 v_1 = (\lambda - 3)x_2 v_2.
\]

Now \( v_1 \) cannot be a multiple of \( v_2 \), or vice versa, for if \( v_1 = \mu v_2 \), then

\[
Av_1 = A(\mu v_2) = \mu(Av_2),
\]
and so $2v_1 = 3\mu v_2 = 3v_1$, which is absurd. Thus, equation (7) implies that $(2 - \lambda)x_1 = 0$ and $(\lambda - 3)x_2 = 0$. It follows that

$$\lambda = 2 \quad \text{and} \quad x_2 = 0$$

or

$$\lambda = 3 \quad \text{and} \quad x_1 = 0$$

or

$$\lambda \neq 2, 3 \quad \text{and} \quad x_1 = x_2 = 0.$$ 

So the only eigenvectors (that is, nonzero vectors which satisfy $Av = \lambda v$) are multiples of $v_1$ or $v_2$. And the only eigenvalues, i.e., scalars $\lambda$ such that $Av = \lambda v$ for some nonzero vector $v$, are 2 and 3. Observe that

$$v_1 = M e_1 = \begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and

$$v_2 = M e_2 = \begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$ 

Thus the columns of the matrix $M$ are the eigenvectors $v_1$ and $v_2$, and the diagonal entries of the matrix $D$ are the eigenvalues 2 and 3. But how can we find the vectors $v_1$ and $v_2$ or the numbers 2 and 3 without knowing the factorisation?

Suppose just that

$$A = \begin{pmatrix} -1 & 3 \\ -4 & 6 \end{pmatrix}.$$ 

How do we find vectors $v$ and numbers $\lambda$ so that $Av = \lambda v$? If $v = (x, y)^T$, then

$$-x + 3y = \lambda x$$

$$-4x + 6y = \lambda y.$$ 

It may be possible to solve these equations for the unknowns $x, y$ and $\lambda$, but since there are terms $\lambda x$ and $\lambda y$, these equations are not linear. But if we knew $\lambda$, then the equations would be linear and we could solve them. So the problem is to find the values of $\lambda$, that is, the eigenvalues. Observe that, if $\lambda$ is an eigenvalue of $A$, then there is a nonzero vector $v$ so that

$$Av = \lambda v.$$ 

Thus

$$Av - \lambda Iv = 0,$$

i.e.,

$$(A - \lambda I)v = 0.$$ 

Then $A - \lambda I$ is singular, so $\det(A - \lambda I) = 0$. In our case,

$$0 = \det \left[ \begin{pmatrix} -1 & 3 \\ -4 & 6 \end{pmatrix} - \begin{pmatrix} \lambda \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ \lambda \end{pmatrix} \right]$$

$$= \det \begin{pmatrix} -1 - \lambda & 3 \\ -4 & 6 - \lambda \end{pmatrix}$$

$$= (-1 - \lambda)(6 - \lambda) + 12$$

$$= \lambda^2 - 5\lambda + 6.$$
This equation tells us that $\lambda = 2$ or $3$. Going backwards, we can then find $v_1$ and $v_2$, and hence $M$ (the columns are $v_1$ and $v_2$) and $D$ (the entries are 2 and 3).

3. Some examples

Problem 41. Given the matrix

$$A = \begin{pmatrix} 1 & -4 \\ 0 & -1 \end{pmatrix},$$

find the eigenvalues of $A$ and then find the eigenvectors. Write $A$ in the form $MDM^{-1}$, where $D$ is diagonal.

Answer. First we find the eigenvalues:

$$\det(A - \lambda I) = \det \begin{pmatrix} 1 - \lambda & -4 \\ 0 & -1 - \lambda \end{pmatrix}$$

$$= (1 - \lambda)(-1 - \lambda)$$

$$= (\lambda + 1)(\lambda - 1),$$

and $\det(A - \lambda I) = 0$ if and only if $\lambda = \pm 1$. These are the eigenvalues.

Suppose that $\lambda = 1$. We want a nonzero vector $v$ such that

$$(A - I)v = 0$$

i.e.,

$$\begin{pmatrix} 0 & -4 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0.$$

It suffices to take $v = (1, 0)^T$.

Suppose that $\lambda = -1$. We want a nonzero vector $v$ such that

$$(A + I)v = 0,$$

i.e.,

$$\begin{pmatrix} 2 & -4 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

It suffices to take $v = (2, 1)^T$.

Thus the eigenvalues are $+1$ and $-1$ and the eigenvectors are $(1, 0)^T$ and $(2, 1)^T$ respectively. Let

$$M = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Then $A = MDM^{-1}$.

Problem 42. Find the eigenvalues and eigenvectors of $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ and $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. △
Answer. First, 
\[ \det \begin{pmatrix} 2 - \lambda & 0 \\ 0 & 2 - \lambda \end{pmatrix} = (2 - \lambda)^2 = (\lambda - 2)^2, \]
so the eigenvalues are 2, 2.

The equation
\[ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \]
has all \((x, y)^T\) as solutions. There are no particularly good choices of \((x, y)^T\) for the eigenvectors.

Next,
\[ \det \begin{pmatrix} 2 - \lambda & 1 \\ 0 & 2 - \lambda \end{pmatrix} = (2 - \lambda)^2 = (\lambda - 2)^2. \]
Again, the eigenvalues are 2, 2.

The equation
\[ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \]
has solutions \((x, y)^T = (a, 0)^T\), for any \(a\). Thus there is a 1-dimensional space of eigenvectors. △

The moral of the story of the last problem is that when eigenvalues are repeated, the possibilities for the eigenvectors may be complicated.

Problem 43. Find the eigenvalues of \( \begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix} \).

Answer. We see that
\[ \det \begin{pmatrix} \sqrt{3}/2 - \lambda & -1/2 \\ 1/2 & \sqrt{3}/2 - \lambda \end{pmatrix} = (\sqrt{3}/2 - \lambda)^2 + 1/4. \]
This cannot be zero for any real \(\lambda\), but is zero when \(\lambda = \sqrt{3}/2 \pm i/2\). △ When there are complex eigenvalues, there are also complex eigenvectors.

4. Some theory

The characteristic polynomial of a square matrix \(A\) is the polynomial \(p\) given by 
\[ p(\lambda) = \det(A - \lambda I). \]
The trace of \(A\) is defined to be the sum of the diagonal entries.

Theorem 26. The characteristic polynomial of an \(n \times n\) matrix \(A\) is of degree \(n\), and of the form 
\[ \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0. \]
The terms \(a_{n-1}\) and \(a_0\) are equal to the negative trace and ± the determinant of \(A\).

Proof. This works by expansion of the determinant and careful bookkeeping. □
Theorem 27. Suppose that $A$ is an $n \times n$ matrix, with entries in $\mathbb{R}$.

a) If the characteristic polynomial of $A$ has $n$ distinct roots, then $A$ has $n$ corresponding eigenvectors, which are real if the eigenvalue is real and complex otherwise, and linearly independent.

b) If the characteristic polynomial of $A$ has repeated roots, then $A$ may or may not have $n$ corresponding eigenvectors.

c) If $A$ has (possibly repeated) eigenvalues $\lambda_1, \ldots, \lambda_n$, with corresponding eigenvectors $v_1, v_2, \ldots, v_n$ which are linearly independent, then

$$A = MDM^{-1},$$

where $M$ is the matrix whose columns are the vectors $v_1, \ldots, v_n$ (in order), and $D$ is the diagonal matrix $\text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.

5. Another example

Problem 44. Suppose that

$$A = \begin{pmatrix} -4 & 2 & 7 \\ -2 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}.$$  

Find the eigenvalues and eigenvectors of $A$, and find $A^{101}$.

Answer. Consider

$$p(\lambda) = \det(A - \lambda I)$$

$$= \det \begin{pmatrix} -4 - \lambda & 2 & 7 \\ -2 & 1 - \lambda & 2 \\ 0 & 0 & 3 - \lambda \end{pmatrix}$$

$$= (3 - \lambda) \det \begin{pmatrix} -4 - \lambda & 2 \\ -2 & 1 - \lambda \end{pmatrix}$$

$$= (3 - \lambda) [(\lambda + 4)(\lambda - 1) + 4]$$

$$= (3 - \lambda) [\lambda^2 + 3\lambda]$$

$$= -\lambda(\lambda - 3)(\lambda + 3).$$

Therefore the eigenvalues are $\pm 3, 0$.

Suppose that $\lambda = 3$. The corresponding eigenvector $v_1$ satisfies

$$(A - 3I)v_1 = 0,$$

i.e.,

$$\begin{pmatrix} -7 & 2 & 7 \\ -2 & -2 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$  

We may take $v_1 = (0, 0, 1)^T$, or more generally, $(0, 0, a)^T$ where $a \neq 0$.

Suppose that $\lambda = -3$. The corresponding eigenvector $v_2$ satisfies

$$(A + 3I)v_2 = 0,$$
i.e.,
\[
\begin{pmatrix}
-1 & 2 & 7 \\
-2 & 4 & 2 \\
0 & 0 & 6
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
= \begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix}
.
\]

Putting
\[
\begin{pmatrix}
-1 & 2 & 7 \\
-2 & 4 & 2 \\
0 & 0 & 6
\end{pmatrix}
\]
into row-echelon form, we get
\[
\begin{pmatrix}
-1 & 2 & 7 \\
0 & 0 & 7 \\
0 & 0 & 6
\end{pmatrix},
\]
and hence we may take \(v_2 = (2, 1, 0)^T\).

Finally, consider the eigenvalue 0. The corresponding eigenvector \(v_3\) satisfies
\[
A v_3 = 0,
\]
i.e.,
\[
\begin{pmatrix}
-4 & 2 & 7 \\
-2 & 1 & 2 \\
0 & 0 & 3
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
= \begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix}
.
\]

Again by going to row-echelon form, we see that we may take \(v_3 = (1, 2, 0)^T\). Then \(A = MDM^{-1}\), where
\[
M = \begin{pmatrix}
0 & 2 & 1 \\
0 & 1 & 2 \\
1 & 0 & 0
\end{pmatrix}
\quad \text{and} \quad
D = \begin{pmatrix}
3 & 0 & 0 \\
0 & -3 & 0 \\
0 & 0 & 0
\end{pmatrix}.
\]

To find \(A^{101}\), we observe that
\[
A^{101} = (MDM^{-1})^{101}
= MD^{101}M^{-1}
= M \begin{pmatrix}
3^{101} & 0 & 0 \\
0 & -3^{101} & 0 \\
0 & 0 & 0
\end{pmatrix} M^{-1}
= 3^{100}M \begin{pmatrix}
3 & 0 & 0 \\
0 & -3 & 0 \\
0 & 0 & 0
\end{pmatrix} M^{-1}
= 3^{100}MDM^{-1}
= 3^{100}A,
\]
which is surprisingly simple!

From this example, you can see that being able to find the eigenvalues and eigenvectors depends on being able to solve the characteristic equation. When \(n\) is big, it can
be impossible to solve this exactly. In some numerical packages, things are turned on their head, and polynomial equations are solved (approximately) by finding eigenvalues (approximately).

6. Applications to differential equations

Solving certain differential equations involves matrix methods. For instance, suppose that \( u \) and \( v \) are unknown functions and

\[
\frac{du(x)}{dx} = 4u(x) + 3v(x) \\
\frac{dv(x)}{dx} = -3u(x) + 2v(x).
\]

We can write this as

\[
\frac{dw(x)}{dx} = \begin{pmatrix} 4 & 3 \\ -3 & 2 \end{pmatrix} w(x) = Aw(x),
\]

say, where \( w(x) = (u(x), v(x))^T \). The solution of the usual differential equation

\[
\frac{dw(x)}{dx} = aw(x)
\]

is \( w(x) = e^{ax}c \), where \( c \) is a constant. Similarly, the solution of

\[
\frac{dw(x)}{dx} = Aw(x)
\]

is \( w(x) = e^{x\mathbf{A}}c \), where \( c \) is a constant vector. We prove this by a series argument. As argued before, if

\[
\mathbf{A} = \mathbf{MDM}^{-1},
\]

where \( \mathbf{D} = \text{diag}(\lambda_1, \lambda_2) \), then

\[
x\mathbf{A} = \mathbf{MDM}^{-1}
\]

so that

\[
(x\mathbf{A})^n = \mathbf{MD}^n\mathbf{M}^{-1}
\]

and

\[
e^{x\mathbf{A}} = \sum_{n=0}^{\infty} \frac{(x\mathbf{A})^n}{n!} \\
= \mathbf{M} \left[ \sum_{n=0}^{\infty} \frac{(x\mathbf{D})^n}{n!} \right] \mathbf{M}^{-1} \\
= \mathbf{M}e^{x\mathbf{D}}\mathbf{M}^{-1}.
\]

So we can compute \( e^{x\mathbf{A}} \) effectively if we can compute \( \mathbf{M} \) and \( \mathbf{D} \).
We see that
\[
\frac{d}{dx} e^{xA} = \frac{d}{dx} M \begin{pmatrix} e^{x\lambda_1} & 0 \\ 0 & e^{x\lambda_2} \end{pmatrix} M^{-1} \\
= M \left[ \frac{d}{dx} \begin{pmatrix} e^{x\lambda_1} & 0 \\ 0 & e^{x\lambda_2} \end{pmatrix} \right] M^{-1} \\
= M \begin{pmatrix} \lambda_1 e^{x\lambda_1} & 0 \\ 0 & \lambda_2 e^{x\lambda_2} \end{pmatrix} M^{-1} \\
= M \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} e^{x\lambda_1} & 0 \\ 0 & e^{x\lambda_2} \end{pmatrix} M^{-1} \\
= M \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} M^{-1} M \begin{pmatrix} e^{x\lambda_1} & 0 \\ 0 & e^{x\lambda_2} \end{pmatrix} M^{-1} \\
= A e^{xA}.
\]

Hence
\[
\frac{d}{dx} (e^{xA} c) = A e^{xA} c = A (e^{xA} c).
\]

**Challenge Problem.** Compute \( \exp \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \).

**Answer.** First we compute \( \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n \). When \( n \) is small, we see that
\[
\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix},
\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix},
\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^4 = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}.
\]

It seems reasonable to guess that
\[
\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix},
\]
and this may be proved by induction. Now
\[
\exp \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \frac{1}{0!} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^0 + \frac{1}{1!} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^1 + \cdots + \frac{1}{n!} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n + \cdots \\
= \frac{1}{0!} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{1!} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} + \cdots + \frac{1}{n!} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} + \cdots \\
= \begin{pmatrix} e^1 \sum_{n=1}^{\infty} \frac{n}{n!} \\ 0 \end{pmatrix} e^1.
\]
To compute \( \sum_{n=1}^{\infty} \frac{n}{n!} \), we change variable \((m = n - 1)\)

\[
\sum_{n=1}^{\infty} \frac{n}{n!} = \sum_{n=1}^{\infty} \frac{1}{(n-1)!}
\]

\[
= \sum_{m=0}^{\infty} \frac{1}{m!}
\]

\[
= e.
\]

Thus \( \exp \left( \begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array} \right) = \left( \begin{array}{cc} e & e \\ 0 & e \end{array} \right) \).

\[\triangle\]

More generally,

\[
\exp \left( t \begin{array}{cc} a & b \\ 0 & a \end{array} \right) = \exp \left( at \begin{array}{cc} 1 & b/a \\ 0 & 1 \end{array} \right)
\]

\[
= \left( \begin{array}{cccc}
\sum_{n=0}^{\infty} \frac{(at)^n}{n!} & \sum_{n=1}^{\infty} \frac{(at)^n nb}{n! a} \\
0 & \sum_{n=0}^{\infty} \frac{(at)^n}{n!}
\end{array} \right)
\]

\[
= \begin{pmatrix} e^{at} & e^{at}b/a \\ 0 & e^{at} \end{pmatrix}.
\]

When we exponentiate upper triangular \( n \times n \) matrices, we can get polynomials of degree up to \( n - 1 \) appearing.

**EXAMPLE.** Consider the differential equation

\[
y'' + 4y' + 4y = 0.
\]

We can solve this by matrix methods. First, let \( u(x) = y(x) \) and \( v(x) = y'(x) \). Then

\[
u'(x) = v(x)
\]

\[
v'(x) = y''(x)
\]

\[
= -4y'(x) - 4y(x)
\]

\[
= -4v(x) - 4u(x).
\]

We conclude that, if \( w(x) = (u(x), v(x))^T \), then

\[
w'(x) = \begin{pmatrix} v(x) \\ -4v(x) - 4u(x) \end{pmatrix}
\]

\[
= \begin{pmatrix} 0 & 1 \\ -4 & -4 \end{pmatrix} \begin{pmatrix} u(x) \\ v(x) \end{pmatrix}
\]

\[
= \begin{pmatrix} 0 & 1 \\ -4 & -4 \end{pmatrix} w(x).
\]
The solution is
\[ w(x) = \exp \left( x \begin{pmatrix} 0 & 1 \\ -4 & -4 \end{pmatrix} \right) c, \]
where \( c \) is a constant vector. It can be shown that
\[ \begin{pmatrix} 0 & 1 \\ -4 & -4 \end{pmatrix} = M \begin{pmatrix} -2 & 1 \\ 0 & -2 \end{pmatrix} M^{-1}. \]
By arguments like those above, we see that
\[ \exp \left( x \begin{pmatrix} 0 & 1 \\ -4 & -4 \end{pmatrix} \right) = M \begin{pmatrix} e^{-2x} & xe^{-2x} \\ 0 & e^{-2x} \end{pmatrix} M^{-1}. \]
Then
\[ w(x) = Me^{-2x} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} M^{-1} c \]
\[ = Me^{-2x} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \]
\[ = Me^{-2x} \begin{pmatrix} d_1 + d_2x \\ d_2 \end{pmatrix} \]
\[ = \begin{pmatrix} k_1e^{-2x} + k_2e^{-2x}x \\ k_3e^{-2x} + k_4e^{-2x}x \end{pmatrix}, \]
where \( d_1, d_2, k_1, k_2, k_3 \) and \( k_4 \) are constants. In particular, this implies that \( u(x) = k_1e^{-2x} + k_2e^{-2x}x \). Thus the polynomial terms which arise in differential equations are “related” to those which arise in exponentiating matrices. Note also that the characteristic polynomial of
\[ y'' + 4y' + 4y = 0 \]
is \( \lambda^2 + 4\lambda + 4 = 0 \), while the characteristic polynomial of the matrix \( \begin{pmatrix} 0 & 1 \\ -4 & -4 \end{pmatrix} \) is
\[ \det \begin{pmatrix} -\lambda & 1 \\ -4 & -\lambda \end{pmatrix} = \lambda^2 + 4\lambda + 4. \]
It is not a coincidence that these coincide!