Eigen Methods
Math 246, Spring 2008, Professor David Levermore

Eigenvalues and Eigenvectors. Let $\mathbf{A}$ be a real $n\times n$ matrix. Recall that a number λ (possibly complex) is an *eigenvalue* of $\mathbf{A}$ if there exists a nonzero vector $\mathbf{v}$ (possibly complex) such that

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}\,. \tag{1}$$

Each such vector is an *eigenvector* associated with λ, and $(\lambda, \mathbf{v})$ is an *eigenpair* of $\mathbf{A}$.

Fact 1: If $(\lambda, \mathbf{v})$ is an eigenpair of $\mathbf{A}$ then so is $(\lambda, \alpha\mathbf{v})$ for every complex $\alpha \neq 0$. In other words, if $\mathbf{v}$ is an eigenvector associated with an eigenvalue λ of $\mathbf{A}$ then so is $\alpha\mathbf{v}$ for every complex $\alpha \neq 0$. In particular, eigenvectors are not unique.

Reason. Because $(\lambda, \mathbf{v})$ is an eigenpair of $\mathbf{A}$ you know that (1) holds. It follows that

$$\mathbf{A}(\alpha\mathbf{v}) = \alpha\mathbf{A}\mathbf{v} = \alpha\lambda\mathbf{v} = \lambda(\alpha\mathbf{v})\,.$$

Because the scalar α and vector $\mathbf{v}$ are nonzero, the vector $\alpha\mathbf{v}$ is also nonzero. Therefore $(\lambda, \alpha\mathbf{v})$ is also an eigenpair of $\mathbf{A}$. □

Recall that the characteristic polynomial of $\mathbf{A}$ is defined by

$$p_{\mathbf{A}}(z) = \det(z\mathbf{I} - \mathbf{A})\,. \tag{2}$$

It has the form

$$p_{\mathbf{A}}(z) = z^n + \pi_1 z^{n-1} + \pi_2 z^{n-2} + \cdots + \pi_{n-1}z + \pi_n\,,$$

where the coefficients π_1, π_2, $\cdots$, π_n are real. In other words, it is a real monic polynomial of degree n. One can show that in general

$$\pi_1 = -\operatorname{tr}(\mathbf{A})\,, \qquad \pi_n = (-1)^n \det(\mathbf{A})\,.$$

In particular, when $n = 2$ one has

$$p_{\mathbf{A}}(z) = z^2 - \operatorname{tr}(\mathbf{A})z + \det(\mathbf{A})\,.$$

Because $\det(z\mathbf{I} - \mathbf{A}) = (-1)^n \det(\mathbf{A} - z\mathbf{I})$, this definition of $p_{\mathbf{A}}(z)$ coincides with the book's definition when n is even, and is its negative when n is odd. Both conventions are common. We have chosen the convention that makes $p_{\mathbf{A}}(z)$ monic. What matters most about $p_{\mathbf{A}}(z)$ is its roots and their multiplicity, which are the same for both conventions.

Fact 2: A number λ is an eigenvalue of $\mathbf{A}$ if and only if $p_{\mathbf{A}}(\lambda) = 0$. In other words, the eigenvalues of $\mathbf{A}$ are the roots of $p_{\mathbf{A}}(z)$.

Reason. If λ is an eigenvalue of $\mathbf{A}$ then by (1) there exists a nonzero vector $\mathbf{v}$ such that

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

It follows that $p_{\mathbf{A}}(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A}) = 0$.

Conversely, if $p_{\mathbf{A}}(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A}) = 0$ then there exists a nonzero vector $\mathbf{v}$ such that $(\lambda\mathbf{I} - \mathbf{A})\mathbf{v} = 0$. It follows that

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

whereby λ and $\mathbf{v}$ satisfy (1), which implies λ is an eigenvalue of $\mathbf{A}$. □

Fact 2 shows that the eigenvalues of a $n\times n$ matrix $\mathbf{A}$ can be found if you can find all the roots of the characteristic polynomial of $\mathbf{A}$. You can then find all the eigenvectors associated with each eigenvalue by finding a general nonzero solution of (1).

You can quickly find the eigenvectors for any 2×2 matrix $\mathbf{A}$ with help from the Cayley-Hamilton Theorem, which states that $p_{\mathbf{A}}(\mathbf{A}) = 0$. The eigenvalues λ_1 and λ_2 are the roots of $p_{\mathbf{A}}(z)$, so $p_{\mathbf{A}}(z) = (z-\lambda_1)(z-\lambda_2)$. Hence, by the Cayley-Hamilton Theorem

$$0 = p_{\mathbf{A}}(\mathbf{A}) = (\mathbf{A}-\lambda_1\mathbf{I})(\mathbf{A}-\lambda_2\mathbf{I}) = (\mathbf{A}-\lambda_2\mathbf{I})(\mathbf{A}-\lambda_1\mathbf{I})\,. \tag{3}$$

It follows that every nonzero column of $\mathbf{A}-\lambda_2\mathbf{I}$ is an eigenvector associated with λ_1, and that every nonzero column of $\mathbf{A}-\lambda_1\mathbf{I}$ is an eigenvector associated with λ_2.

Example. Find the eigenpairs of $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$.

Solution. The characteristic polynomial of $\mathbf{A}$ is

$$p_{\mathbf{A}}(z) = z^2 - 6z + 5 = (z-1)(z-5)\,.$$

By Fact 2 the eigenvalues of $\mathbf{A}$ are 1 and 5. Because

$$\mathbf{A}-\mathbf{I} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, \qquad \mathbf{A}-5\mathbf{I} = \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix},$$

Every column of $\mathbf{A}-5\mathbf{I}$ has the form

$$\alpha\begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{for some } \alpha \neq 0\,,$$

while every column of $\mathbf{A}-\mathbf{I}$ has the form

$$\alpha\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{for some } \alpha \neq 0\,.$$

It follows from (3) that the eigenpairs of $\mathbf{A}$ are

$$\left(1\,, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right), \qquad \left(5\,, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right).$$

Example. Find the eigenpairs of $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix}$.

Solution. The characteristic polynomial of $\mathbf{A}$ is

$$p_{\mathbf{A}}(z) = z^2 - 6z + 13 = (z-3)^2 + 4 = (z-3)^2 + 2^2\,.$$

By Fact 2 the eigenvalues of $\mathbf{A}$ are $3+i2$ and $3-i2$. Because

$$\mathbf{A}-(3+i2)\mathbf{I} = \begin{pmatrix} -i2 & 2 \\ -2 & -i2 \end{pmatrix}, \qquad \mathbf{A}-(3-i2)\mathbf{I} = \begin{pmatrix} i2 & 2 \\ -2 & i2 \end{pmatrix}.$$

Every column of $\mathbf{A}-(3-i2)\mathbf{I}$ has the form

$$\alpha\begin{pmatrix} 1 \\ i \end{pmatrix} \quad \text{for some } \alpha \neq 0\,,$$

while every column of $\mathbf{A}-(3+i2)\mathbf{I}$ has the form

$$\alpha\begin{pmatrix} 1 \\ -i \end{pmatrix} \quad \text{for some } \alpha \neq 0\,.$$

It follows from (3) that the eigenpairs of $\mathbf{A}$ are

$$\left(3+i2\,, \begin{pmatrix} 1 \\ i \end{pmatrix}\right), \qquad \left(3-i2\,, \begin{pmatrix} 1 \\ -i \end{pmatrix}\right).$$

Notice that in the above example the eigenvectors associated with $3 - i2$ are complex conjugates to those associated with $3 + i2$. This illustrates is a particular instance of the following general fact.

Fact 3: If $(\lambda, \mathbf{v})$ is an eigenpair of the real matrix $\mathbf{A}$ then so is $(\overline{\lambda}, \overline{\mathbf{v}})$.

Reason. Because $(\lambda, \mathbf{v})$ is an eigenpair of $\mathbf{A}$ you know by (1) that $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Because $\mathbf{A}$ is real, the complex conjugate of this equation is

$$\mathbf{A}\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}\,,$$

where $\overline{\mathbf{v}}$ is nonzero because $\mathbf{v}$ is nonzero. It follows that $(\overline{\lambda}, \overline{\mathbf{v}})$ is an eigenpair of $\mathbf{A}$. □

Both examples given above illustrate particular instances of the following general facts.

Fact 4: Let λ be an eigenvalue of the real matrix $\mathbf{A}$. If λ is real then it has a real eigenvector. If λ is not real then none of its eigenvectors are real.

Reason. Let $\mathbf{v}$ be any eigenvector associated with λ, so that $(\lambda, \mathbf{v})$ is an eigenpair of $\mathbf{A}$. Let $\lambda = \mu + i\nu$ and $\mathbf{v} = \mathbf{u} + i\mathbf{w}$ where μ and ν are real numbers and $\mathbf{u}$ and $\mathbf{w}$ are real vectors. One then has

$$\mathbf{A}\mathbf{u} + i\mathbf{A}\mathbf{w} = \mathbf{A}\mathbf{v} = \lambda\mathbf{v} = (\mu + i\nu)(\mathbf{u} + i\mathbf{w}) = (\mu\mathbf{u} - \nu\mathbf{w}) + i(\mu\mathbf{w} + \nu\mathbf{u})\,,$$

which is equivalent to

$$\mathbf{A}\mathbf{u} - \mu\mathbf{u} = -\nu\mathbf{w}\,, \qquad \text{and} \qquad \mathbf{A}\mathbf{w} - \mu\mathbf{w} = \nu\mathbf{u}\,.$$

If $\nu = 0$ then $\mathbf{u}$ and $\mathbf{w}$ will be real eigenvectors associated with λ whenever they are nonzero. But at least one of $\mathbf{u}$ and $\mathbf{w}$ must be nonzero because $\mathbf{v} = \mathbf{u} + i\mathbf{w}$ is nonzero. Conversely, if $\nu \neq 0$ and $\mathbf{w} = 0$ then the second equation above implies $\mathbf{u} = 0$ too, which contradicts the fact that at least one of $\mathbf{u}$ and $\mathbf{w}$ must be nonzero. Hence, if $\nu \neq 0$ then $\mathbf{w} \neq 0$. □

Solutions of First-Order Systems. We are now ready to use eigenvalues and eigenvectors to construct solutions of first-order differential systems with a constant coefficient matrix. The system we study is

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{A}\mathbf{x}\,, \tag{4}$$

where $\mathbf{x}(t)$ is a vector and $\mathbf{A}$ is a real $n{\times}n$ matrix. We begin with the following basic fact.

Fact 5: If $(\lambda, \mathbf{v})$ is an eigenpair of $\mathbf{A}$ then a solution of (4) is

$$\mathbf{x}(t) = e^{\lambda t}\mathbf{v}\,. \tag{5}$$

Reason. By direct calculation we see that

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\big(e^{\lambda t}\mathbf{v}\big) = e^{\lambda t}\lambda\mathbf{v} = e^{\lambda t}\mathbf{A}\mathbf{v} = \mathbf{A}\big(e^{\lambda t}\mathbf{v}\big) = \mathbf{A}\mathbf{x}\,,$$

whereby $\mathbf{x}(t)$ given by (5) solves (4). □

If $(\lambda, \mathbf{v})$ is a real eigenpair of $\mathbf{A}$ then recipe (5) will yield a real solution of (4). But if λ is an eigenvalue of $\mathbf{A}$ that is not real then recipe (5) will not yield a real solution. However, if we also use the solution associated with the conjugate eigenpair $(\overline{\lambda}, \overline{\mathbf{v}})$ then we can construct two real solutions.

Fact 6: Let $(\lambda, \mathbf{v})$ be an eigenpair of $\mathbf{A}$ with $\lambda = \mu + i\nu$ and $\mathbf{v} = \mathbf{u} + i\mathbf{w}$ where μ and ν are real numbers while $\mathbf{u}$ and $\mathbf{w}$ are real vectors. Then two real solutions of (4) are

$$
\begin{aligned}
\mathbf{x}_1(t) &= \operatorname{Re}\big(e^{\lambda t}\mathbf{v}\big) = e^{\mu t}\big(\mathbf{u}\cos(\nu t) - \mathbf{w}\sin(\nu t)\big)\,, \\
\mathbf{x}_2(t) &= \operatorname{Im}\big(e^{\lambda t}\mathbf{v}\big) = e^{\mu t}\big(\mathbf{w}\cos(\nu t) + \mathbf{u}\sin(\nu t)\big)\,.
\end{aligned}
\tag{6}
$$

Reason. Because $(\lambda, \mathbf{v})$ is an eigenpair of $\mathbf{A}$, by Fact 3 so is $(\overline{\lambda}, \overline{\mathbf{v}})$. By recipe (5) two solutions of (4) are $e^{\lambda t}\mathbf{v}$ and $e^{\overline{\lambda} t}\overline{\mathbf{v}}$, which are complex conjugates of each other. Because equation (4) is linear, it follows that two real solutions of (4) are given by

$$
\mathbf{x}_1(t) = \operatorname{Re}\big(e^{\lambda t}\mathbf{v}\big) = \frac{e^{\lambda t}\mathbf{v} + e^{\overline{\lambda} t}\overline{\mathbf{v}}}{2}\,, \qquad \mathbf{x}_2(t) = \operatorname{Im}\big(e^{\lambda t}\mathbf{v}\big) = \frac{e^{\lambda t}\mathbf{v} - e^{\overline{\lambda} t}\overline{\mathbf{v}}}{i2}\,.
$$

Because $\lambda = \mu + i\nu$ and $\mathbf{v} = \mathbf{u} + i\mathbf{w}$ we see that

$$
\begin{aligned}
e^{\lambda t}\mathbf{v} &= e^{\mu t}\big(\cos(\nu t) + i\sin(\nu t)\big)(\mathbf{u} + i\mathbf{v}) \\
&= e^{\mu t}\Big[\big(\mathbf{u}\cos(\nu t) - \mathbf{w}\sin(\nu t)\big) + i\big(\mathbf{w}\cos(\nu t) + \mathbf{u}\sin(\nu t)\big)\Big]\,,
\end{aligned}
$$

whereby $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$ are read off from the real and imaginary parts, yielding (6). □

Example. Find two linearly independent real solutions of

$$
\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{A}\mathbf{x}\,, \qquad \text{where} \quad \mathbf{A} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}.
$$

Solution. By a previous example we know that $\mathbf{A}$ has the real eigenpairs

$$
\left(1\,, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right), \qquad \left(5\,, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right).
$$

By recipe (5) the equation has the real solutions

$$
\mathbf{x}_1(t) = e^{t}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad \mathbf{x}_2(t) = e^{5t}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.
$$

These solutions are linearly independent because

$$
W[\mathbf{x}_1, \mathbf{x}_2](0) = \det\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = 2 \neq 0\,.
$$

Example. Find two linearly independent real solutions of

$$
\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{A}\mathbf{x}\,, \qquad \text{where} \quad \mathbf{A} = \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix}.
$$

Solution. By a previous example we know that $\mathbf{A}$ has the conjugate eigenpairs

$$
\left(3 + i2\,, \begin{pmatrix} 1 \\ i \end{pmatrix}\right), \qquad \left(3 - i2\,, \begin{pmatrix} 1 \\ -i \end{pmatrix}\right).
$$

Because

$$
e^{(3+i2)t}\begin{pmatrix} 1 \\ i \end{pmatrix} = e^{3t}\big(\cos(2t) + i\sin(2t)\big)\begin{pmatrix} 1 \\ i \end{pmatrix} = e^{3t}\begin{pmatrix} \cos(2t) + i\sin(2t) \\ -\sin(2t) + i\cos(2t) \end{pmatrix},
$$

by recipe (6) the equation has the real solutions

$$
\mathbf{x}_1(t) = e^{3t}\begin{pmatrix} \cos(2t) \\ -\sin(2t) \end{pmatrix}, \qquad \mathbf{x}_2(t) = e^{3t}\begin{pmatrix} \sin(2t) \\ \cos(2t) \end{pmatrix}.
$$

These solutions are linearly independent because

$$W[\mathbf{x}_1, \mathbf{x}_2](0) = \det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \neq 0\,.$$

Matrix Exponentials. If recipe (5) yields n linearly independent solutions of the first-order system (4) then they can be used to construct the matrix exponential $e^{t\mathbf{A}}$. The key to this construction is the following fact from linear algebra.

Fact 7: If a real $n{\times}n$ matrix $\mathbf{A}$ has n eigenpairs, $(\lambda_1, \mathbf{v}_1)$, $(\lambda_2, \mathbf{v}_2)$, $\cdots$, $(\lambda_n, \mathbf{v}_n)$, such that the eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\cdots$, $\mathbf{v}_n$ are linearly independent then

$$\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1} \quad , \tag{7}$$

where $\mathbf{V}$ is the $n{\times}n$ matrix whose columns are the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\cdots$, $\mathbf{v}_n$ — i.e.

$$\mathbf{V} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{pmatrix}\,, \tag{8}$$

while $\mathbf{D}$ is the $n{\times}n$ diagonal matrix

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}\,. \tag{9}$$

Reason. Underlying this result is the fact that

$$\begin{aligned} \mathbf{A}\mathbf{V} &= \mathbf{A}\begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{pmatrix} = \begin{pmatrix} \mathbf{A}\mathbf{v}_1 & \mathbf{A}\mathbf{v}_2 & \cdots & \mathbf{A}\mathbf{v}_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} = \mathbf{V}\mathbf{D}\,. \end{aligned} \tag{10}$$

Once we show that $\mathbf{V}$ is inverible then (7) follows upon multiplying the above relation on the left by $\mathbf{V}^{-1}$.

We claim that $\det(\mathbf{V}) \neq 0$ because the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\cdots$, $\mathbf{v}_n$ are linearly independent. Suppose otherwise. Because $\det(\mathbf{V}) = 0$ there exists a nonzero vector $\mathbf{c}$ such that $\mathbf{V}\mathbf{c} = 0$. This means that

$$0 = \mathbf{V}\mathbf{c} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n\,.$$

Because vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\cdots$, $\mathbf{v}_n$ are linearly independent, this implies $c_1 = c_2 = \cdots = c_n = 0$, which contradicts the fact $\mathbf{c}$ is nonzero. Therefore $\det(\mathbf{V}) \neq 0$. Hence, the matrix $\mathbf{V}$ is invertible and (7) follows upon multiplying relation (10) on the left by $\mathbf{V}^{-1}$. □

We call a real $n{\times}n$ matrix $\mathbf{A}$ *diagonalizable* when there exists an invertible matrix $\mathbf{V}$ and a diagonal matrix $\mathbf{D}$ such that $\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$. To *diagonalize* $\mathbf{A}$ means to find such a $\mathbf{V}$ and $\mathbf{D}$. Fact 7 states that $\mathbf{A}$ is diagonalizable when it has n linearly independent eigenvectors. The converse of this statement is also true.

Fact 8: If a real $n\times n$ matrix $\mathbf{A}$ is diagonalizable then it has n linearly independent eigenvectors.

Reason. Because $\mathbf{A}$ is diagonalizable it has the form $\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$ where the matrix $\mathbf{V}$ is invertible and the matrix $\mathbf{D}$ is diagonal.

Let the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\cdots$, $\mathbf{v}_n$ be the columns of $\mathbf{V}$. We claim these vectors are linearly independent. Indeed, if $0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ then because $\mathbf{V} = \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\end{pmatrix}$ we see that

$$0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\end{pmatrix}\begin{pmatrix}c_1\\ c_2\\ \vdots\\ c_n\end{pmatrix} = \mathbf{V}\mathbf{c}\,.$$

Because $\mathbf{V}$ is invertible, this implies that $\mathbf{c} = 0$. The vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\cdots$, $\mathbf{v}_n$ are therefore linearly independent.

Because $\mathbf{V} = \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\end{pmatrix}$ and because $\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$ where $\mathbf{D}$ has the form (9), we see that

$$\begin{aligned}
\begin{pmatrix}\mathbf{A}\mathbf{v}_1 & \mathbf{A}\mathbf{v}_2 & \cdots & \mathbf{A}\mathbf{v}_n\end{pmatrix} &= \mathbf{A}\begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\end{pmatrix}\\
&= \mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}\mathbf{V} = \mathbf{V}\mathbf{D}\\
&= \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\end{pmatrix}\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & \lambda_n\end{pmatrix}\\
&= \begin{pmatrix}\lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n\end{pmatrix}\,.
\end{aligned}$$

Because the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\cdots$, $\mathbf{v}_n$ are linearly independent, they are all nonzero. It then follows from the above relation that $(\lambda_1, \mathbf{v}_1)$, $(\lambda_2, \mathbf{v}_2)$, $\cdots$, $(\lambda_n, \mathbf{v}_n)$ are eigenpairs of $\mathbf{A}$, such that the eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\cdots$, $\mathbf{v}_n$ are linearly independent. □

Example. Show that $\mathbf{A} = \begin{pmatrix}3 & 2\\ 2 & 3\end{pmatrix}$ is diagonalizable, and diagonalize it.

Solution. By a previous example we know that $\mathbf{A}$ has the real eigenpairs

$$\left(1\,,\begin{pmatrix}1\\ -1\end{pmatrix}\right)\,,\qquad \left(5\,,\begin{pmatrix}1\\ 1\end{pmatrix}\right)\,.$$

Because we also know the eigenvectors are linearly independent, $\mathbf{A}$ is diagonalizable. Then (8) and (9) yield

$$\mathbf{V} = \begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}\,,\qquad \mathbf{D} = \begin{pmatrix}1 & 0\\ 0 & 5\end{pmatrix}\,.$$

Because $\det(\mathbf{V}) = 2$, one has

$$\mathbf{V}^{-1} = \frac{1}{2}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\,.$$

It follows from (7) that $\mathbf{A}$ is diagonalized as

$$\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1} = \frac{1}{2}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 5\end{pmatrix}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\,.$$

We are now ready to give a construction of the matrix exponential $e^{t\mathbf{A}}$.

Fact 9: If the real $n\times n$ matrix $\mathbf{A}$ has n eigenpairs, $(\lambda_1, \mathbf{v}_1)$, $(\lambda_2, \mathbf{v}_2)$, $\cdots$, $(\lambda_n, \mathbf{v}_n)$, such that the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\cdots$, $\mathbf{v}_n$ are linearly independent then

$$e^{t\mathbf{A}} = \mathbf{V}e^{t\mathbf{D}}\mathbf{V}^{-1}\,, \tag{11}$$

where $\mathbf{V}$ and $\mathbf{D}$ are the $n\times n$ matrices given by (8) and (9).

Reason. Set $\mathbf{\Phi}(t) = \mathbf{V}e^{t\mathbf{D}}\mathbf{V}^{-1}$. It then follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{\Phi}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\big(\mathbf{V}e^{t\mathbf{D}}\mathbf{V}^{-1}\big) = \mathbf{V}\frac{\mathrm{d}}{\mathrm{d}t}e^{t\mathbf{D}}\mathbf{V}^{-1} = \mathbf{V}\mathbf{D}e^{t\mathbf{D}}\mathbf{V}^{-1} = \mathbf{A}\mathbf{V}e^{t\mathbf{D}}\mathbf{V}^{-1} = \mathbf{A}\mathbf{\Phi}(t)\,,$$

whereby the matrix-valued function $\mathbf{\Phi}(t)$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{\Phi}(t) = \mathbf{A}\mathbf{\Phi}(t)\,.$$

Moreover, because $e^{0\mathbf{D}} = \mathbf{I}$ we see that $\mathbf{\Phi}(t)$ also satisfies the initial condition

$$\mathbf{\Phi}(0) = \mathbf{V}e^{0\mathbf{D}}\mathbf{V}^{-1} = \mathbf{V}\mathbf{I}\mathbf{V}^{-1} = \mathbf{V}\mathbf{V}^{-1} = \mathbf{I}\,.$$

It follows that $\mathbf{\Phi}(t) = e^{t\mathbf{A}}$, whereby (11) follows. $\square$

Formula (11) is the book's method for computing $e^{t\mathbf{A}}$ when $\mathbf{A}$ is diagonalizable. Because not every matrix is diagonalizable, it cannot always be applied. When it can be applied, most of the work needed to apply it goes into computing $\mathbf{V}$ and $\mathbf{V}^{-1}$. The matrix $e^{t\mathbf{D}}$ is simply given by

$$e^{t\mathbf{D}} = \begin{pmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{\lambda_n t} \end{pmatrix}. \tag{12}$$

Once you have $\mathbf{V}$, $\mathbf{V}^{-1}$, and $e^{t\mathbf{D}}$, formula (11) requires two matrix multiplications.

Example. Compute $e^{t\mathbf{A}}$ for $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}$.

Solution. By a previous example we know that $\mathbf{A}$ has the real eigenpairs

$$\left(1\,, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)\,, \qquad \left(5\,, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)\,,$$

and that $\mathbf{A}$ is diagonalizable. By (8) and (9) we also know that

$$\mathbf{V} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\,, \qquad \mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}\,, \qquad \mathbf{V}^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\,.$$

By formulas (11) and (12) we therefore have

$$\begin{aligned} e^{t\mathbf{A}} = \mathbf{V}e^{t\mathbf{D}}\mathbf{V}^{-1} &= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} e^{t} & 0 \\ 0 & e^{5t} \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} e^{t} & -e^{t} \\ e^{5t} & e^{5t} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} e^{5t}+e^{t} & e^{5t}-e^{t} \\ e^{5t}-e^{t} & e^{5t}+e^{t} \end{pmatrix}\,. \end{aligned}$$

Example. Compute $e^{t\mathbf{A}}$ for $\mathbf{A} = \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix}$.

Solution. By a previous example we know that $\mathbf{A}$ has the conjugate eigenpairs

$$\left(3 + i2\,, \begin{pmatrix} 1 \\ i \end{pmatrix}\right), \qquad \left(3 - i2\,, \begin{pmatrix} 1 \\ -i \end{pmatrix}\right).$$

By (8) and (9) we know that

$$\mathbf{V} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}, \qquad \mathbf{D} = \begin{pmatrix} 3+i2 & 0 \\ 0 & 3-i2 \end{pmatrix}.$$

Because $\det(\mathbf{V}) = -i2$, we have

$$\mathbf{V}^{-1} = \frac{1}{-i2} \begin{pmatrix} -i & -1 \\ -i & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

By formula (12) we have

$$e^{t\mathbf{D}} = \begin{pmatrix} e^{(3+i2)t} & 0 \\ 0 & e^{(3-i2)t} \end{pmatrix} = e^{3t} \begin{pmatrix} e^{i2t} & 0 \\ 0 & e^{-i2t} \end{pmatrix}.$$

By formula (11) we therefore have

$$\begin{aligned} e^{t\mathbf{A}} &= \mathbf{V} e^{t\mathbf{D}} \mathbf{V}^{-1} = \frac{e^{3t}}{2} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} e^{i2t} & 0 \\ 0 & e^{-i2t} \end{pmatrix} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \\ &= \frac{e^{3t}}{2} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} e^{i2t} & -ie^{i2t} \\ e^{-i2t} & ie^{-i2t} \end{pmatrix} = \frac{e^{3t}}{2} \begin{pmatrix} e^{i2t} + e^{-i2t} & -ie^{i2t} + ie^{-i2t} \\ ie^{i2t} - ie^{-i2t} & e^{i2t} + e^{-i2t} \end{pmatrix} \\ &= \frac{e^{3t}}{2} \begin{pmatrix} 2\cos(2t) & 2\sin(2t) \\ -2\sin(2t) & 2\cos(2t) \end{pmatrix} = e^{3t} \begin{pmatrix} \cos(2t) & \sin(2t) \\ -\sin(2t) & \cos(2t) \end{pmatrix}. \end{aligned}$$

Remark. Because $\mathbf{A}$ is real, $e^{t\mathbf{A}}$ must be real. As the above example illustrates, the matrices $\mathbf{V}$ and $\mathbf{D}$ may not be real, but will always combine in formula (11) to yield the real result.

Remark. While not every matrix is diagonalizable, most matrices are. Here we give four criteria that insure a real $n\times n$ matrix $\mathbf{A}$ is diagonalizable.

- If $\mathbf{A}$ has n distinct eigenvalues then it is diagonalizable.
- If $\mathbf{A}$ is symmetric ($\mathbf{A}^T = \mathbf{A}$) then its eigenvalues are real ($\overline{\lambda} = \lambda$), and it will have n real eigenvectors $\mathbf{v}_j$ that can be normalized so that $\mathbf{v}_j^T \mathbf{v}_k = \delta_{jk}$. With this normalization $\mathbf{V}^{-1} = \mathbf{V}^T$.
- If $\mathbf{A}$ is skew-symmetric ($\mathbf{A}^T = -\mathbf{A}$) then its eigenvalues are imaginary ($\overline{\lambda_j} = -\lambda_j$), and it will have n eigenvectors $\mathbf{v}_j$ that can be normalized so that $\mathbf{v}_j^* \mathbf{v}_k = \delta_{jk}$. With this normalization $\mathbf{V}^{-1} = \mathbf{V}^*$.
- If $\mathbf{A}$ is normal ($\mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A}^T$) then it will have n eigenvectors $\mathbf{v}_j$ that can be normalized so that $\mathbf{v}_j^* \mathbf{v}_k = \delta_{jk}$. With this normalization $\mathbf{V}^{-1} = \mathbf{V}^*$.

If $\mathbf{A}$ is either symmetric or skew-symmetric then it is normal. Both of the examples we have worked have distinct eigenvalues. The first example is symmetric. The second is normal.