# Chapter 9

# Eigenanalysis

**Contents**



This presentation of **matrix eigenanalysis** treats the subject in depth for a $3 \times 3$ matrix $A$. The generalization to an $n \times n$ matrix $A$ is easily supplied by the reader.

## 9.1 Eigenanalysis I

Treated here is eigenanalysis for matrix equations. The topics are *eigenanalysis*, *eigenvalue*, *eigenvector*, *eigenpair* and *diagonalization*.

### What's Eigenanalysis?

Matrix eigenanalysis is a computational theory for the matrix equation $\vec{\mathbf{y}} = A\vec{\mathbf{x}}$. Here, we assume $A$ is a $3 \times 3$ matrix.

The basis of eigenanalysis is **Fourier's Model**:

$$
\begin{aligned}
\vec{\mathbf{x}} &= c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3 \text{ implies} \\
\vec{\mathbf{y}} &= A\vec{\mathbf{x}} \\
&= c_1\lambda_1\vec{\mathbf{v}}_1 + c_2\lambda_2\vec{\mathbf{v}}_2 + c_3\lambda_3\vec{\mathbf{v}}_3.
\end{aligned} \tag{1}
$$

These relations can be written as a single equation:

$$A\left(c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3\right) = c_1\lambda_1\vec{\mathbf{v}}_1 + c_2\lambda_2\vec{\mathbf{v}}_2 + c_3\lambda_3\vec{\mathbf{v}}_3.$$

The scale factors $\lambda_1$, $\lambda_2$, $\lambda_3$ and independent vectors $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ depend only on $A$. Symbols $c_1$, $c_2$, $c_3$ stand for arbitrary numbers. This implies variable $\vec{\mathbf{x}}$ exhausts all possible 3-vectors in $\mathcal{R}^3$ and $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ is a basis for $\mathcal{R}^3$. Fourier's model is a replacement process:

> To compute $A\vec{\mathbf{x}}$ from $\vec{\mathbf{x}} = c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3$, replace each vector $\vec{\mathbf{v}}_i$ by its scaled version $\lambda_i\vec{\mathbf{v}}_i$.

Fourier's model is said to **hold** provided there exist $\lambda_1$, $\lambda_2$, $\lambda_3$ and independent vectors $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ satisfying (1). It is known that Fourier's model fails for certain matrices $A$, for example,

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

**Powers and Fourier's Model**. Equation (1) applies to compute powers $A^n$ of a matrix $A$ using only the basic vector space toolkit. To illustrate, only the vector toolkit for $\mathcal{R}^3$ is used in computing

$$A^5\vec{\mathbf{x}} = x_1\lambda_1^5\vec{\mathbf{v}}_1 + x_2\lambda_2^5\vec{\mathbf{v}}_2 + x_3\lambda_3^5\vec{\mathbf{v}}_3.$$

This calculation does not depend upon finding previous powers $A^2$, $A^3$, $A^4$ as would be the case by using matrix multiply.

Fourier's model can reduce computational effort. Matrix $3 \times 3$ multiplication to produce $\vec{\mathbf{y}}_k = A^k\vec{\mathbf{x}}$ requires $9k$ multiply operations whereas Fourier's $3 \times 3$ model gives the answer with $3k + 9$ multiply operations.

**Fourier's model illustrated**. Let

$$\begin{array}{lll} A = \begin{pmatrix} 1 & 3 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & -5 \end{pmatrix} & & \\ \lambda_1 = 1, & \lambda_2 = 2, & \lambda_3 = -5, \\ \vec{\mathbf{v}}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, & \vec{\mathbf{v}}_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, & \vec{\mathbf{v}}_3 = \begin{pmatrix} 1 \\ -2 \\ -14 \end{pmatrix}. \end{array}$$

Then Fourier's model holds (details later) and

$$\begin{array}{rcl} \vec{\mathbf{x}} & = & c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 1 \\ -2 \\ -14 \end{pmatrix} \quad \text{implies} \\ A\vec{\mathbf{x}} & = & c_1(1) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2(2) \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + c_3(-5) \begin{pmatrix} 1 \\ -2 \\ -14 \end{pmatrix} \end{array}$$

Eigenanalysis might be called *the method of simplifying coordinates*. The nomenclature is justified, because Fourier's model computes $\vec{\mathbf{y}} = A\vec{\mathbf{x}}$ by scaling independent vectors $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$, which is a triad or **coordinate system**.

Success stories for eigenanalysis include geometric problems, systems of differential equations representing mechanical systems, chemical kinetics, electrical networks, and heat and wave partial differential equations.

In summary:

> The subject of **eigenanalysis** discovers a coordinate system and scale factors such that Fourier's model holds. Fourier's model simplifies the matrix equation $\vec{\mathbf{y}} = A\vec{\mathbf{x}}$.

**Differential Equations and Fourier's Model**. Systems of differential equations can be solved using Fourier's model, giving a compact and elegant formula for the general solution. An example:

$$\begin{array}{rcrcrcr} x_1' &=& x_1 &+& 3x_2, && \\ x_2' &=& && 2x_2 &-& x_3, \\ x_3' &=& && &-& 5x_3. \end{array}$$

The matrix form is $\vec{\mathbf{x}}' = A\vec{\mathbf{x}}$, where $A$ is the same matrix used in the Fourier model illustration of the previous paragraph.

Fourier's idea of re-scaling applies as well to differential equations, in the following context. First, expand the initial condition $\vec{\mathbf{x}}(0)$ in terms of basis elements $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$:

$$\vec{\mathbf{x}}(0) = c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3.\vec{\mathbf{v}}_3.$$

Then the general solution of $\vec{\mathbf{x}}' = A\vec{\mathbf{x}}$ is given by replacing each $\vec{\mathbf{v}}_i$ by the re-scaled vector $e^{\lambda_i t}\vec{\mathbf{v}}_i$, giving the formula

$$\vec{\mathbf{x}}(t) = c_1 e^{\lambda_1 t}\vec{\mathbf{v}}_1 + c_2 e^{\lambda_2 t}\vec{\mathbf{v}}_2 + c_3 e^{\lambda_3 t}\vec{\mathbf{v}}_3.$$

For the illustration here, the result is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = c_1 e^{t}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 e^{2t}\begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + c_3 e^{-5t}\begin{pmatrix} 1 \\ -2 \\ -14 \end{pmatrix}.$$

## What's an Eigenvalue?

It is a scale factor. An eigenvalue is also called a *proper value* or a *hidden value*. Symbols $\lambda_1$, $\lambda_2$, $\lambda_3$ used in Fourier's model are eigenvalues.

Historically, the German term *eigenwert* was used exclusively in literature, because German was the preferred publication language for physics. Due to literature migration into English language journals, a hybrid term *eigenvalue* evolved, the German word *wert* replaced by *value*

**A Key Example**. Let $\vec{\mathbf{x}}$ in $\mathcal{R}^3$ be a data set variable with coordinates $x_1$, $x_2$, $x_3$ recorded respectively in units of meters, millimeters and centimeters. We consider the problem of conversion of the mixed-unit $\vec{\mathbf{x}}$-data into proper MKS units (meters-kilogram-second) $\vec{\mathbf{y}}$-data via the equations

$$\begin{array}{rcl} y_1 &=& x_1, \\ y_2 &=& 0.001x_2, \\ y_3 &=& 0.01x_3. \end{array} \tag{2}$$

Equations (2) are a **model** for changing units. Scaling factors $\lambda_1 = 1$, $\lambda_2 = 0.001$, $\lambda_3 = 0.01$ are the **eigenvalues** of the model. To summarize:

> The **eigenvalues** of a model are **scale factors**. They are normally represented by symbols $\lambda_1, \lambda_2, \lambda_3, \ldots$.

The data conversion problem (2) can be represented as $\vec{\mathbf{y}} = A\vec{\mathbf{x}}$, where the diagonal matrix $A$ is given by

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \quad \lambda_1 = 1,\ \lambda_2 = \frac{1}{1000},\ \lambda_3 = \frac{1}{100}.$$

## What's an Eigenvector?

Symbols $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ in Fourier's model are called eigenvectors, or *proper vectors* or *hidden vectors*. They are assumed independent.

The **eigenvectors** $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ of model (2) are three independent **directions of application** for the respective scale factors $\lambda_1 = 1$, $\lambda_2 = 0.001$, $\lambda_3 = 0.01$. The directions identify the components of the data set, to which the individual scale factors are to be multiplied, to perform the data conversion. Because the scale factors apply individually to the $x_1$, $x_2$ and $x_3$ components of a vector $\vec{\mathbf{x}}$, then

$$\vec{\mathbf{v}}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{\mathbf{v}}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{\mathbf{v}}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3}$$

The data is represented as $\vec{\mathbf{x}} = x_1\vec{\mathbf{v}}_1 + x_2\vec{\mathbf{v}}_2 + x_3\vec{\mathbf{v}}_3$. The answer $\vec{\mathbf{y}} = A\vec{\mathbf{x}}$ is given by the equation

$$\begin{aligned}
\vec{\mathbf{y}} &= \begin{pmatrix} \lambda_1 x_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_2 x_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \lambda_3 x_3 \end{pmatrix} \\
&= \lambda_1 x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\
&= x_1\lambda_1\vec{\mathbf{v}}_1 + x_2\lambda_2\vec{\mathbf{v}}_2 + x_3\lambda_3\vec{\mathbf{v}}_3.
\end{aligned}$$

In summary:

> The **eigenvectors** of a model are independent **directions of application** for the scale factors (eigenvalues).

**History of Fourier's Model.** The subject of **eigenanalysis** was popularized by J. B. Fourier in his 1822 publication on the theory of heat, *Théorie analytique de la chaleur*. His ideas can be summarized as follows for the $n \times n$ matrix equation $\vec{\mathbf{y}} = A\vec{\mathbf{x}}$.

> The vector $\vec{\mathbf{y}} = A\vec{\mathbf{x}}$ is obtained from eigenvalues $\lambda_1$, $\lambda_2$, ..., $\lambda_n$ and eigenvectors $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, ..., $\vec{\mathbf{v}}_n$ by replacing the eigenvectors by their scaled versions $\lambda_1\vec{\mathbf{v}}_1$, $\lambda_2\vec{\mathbf{v}}_2$, ..., $\lambda_n\vec{\mathbf{v}}_n$:
>
> $$\begin{array}{lllllllll} \vec{\mathbf{x}} & = & c_1\vec{\mathbf{v}}_1 & + & c_2\vec{\mathbf{v}}_2 & + & \cdots & + & c_n\vec{\mathbf{v}}_n \quad \text{implies} \\ \vec{\mathbf{y}} & = & x_1\lambda_1\vec{\mathbf{v}}_1 & + & x_2\lambda_2\vec{\mathbf{v}}_2 & + & \cdots & + & c_n\lambda_n\vec{\mathbf{v}}_n. \end{array}$$

**Determining Equations.** The eigenvalues and eigenvectors are determined by homogeneous matrix–vector equations. In Fourier's model

$$A(c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3) = c_1\lambda_1\vec{\mathbf{v}}_1 + c_2\lambda_2\vec{\mathbf{v}}_2 + c_3\lambda_3\vec{\mathbf{v}}_3$$

choose $c_1 = 1$, $c_2 = c_3 = 0$. The equation reduces to $A\vec{\mathbf{v}}_1 = \lambda_1\vec{\mathbf{v}}_1$. Similarly, taking $c_1 = c_2 = 0$, $c_2 = 1$ implies $A\vec{\mathbf{v}}_2 = \lambda_2\vec{\mathbf{v}}_2$. Finally, taking $c_1 = c_2 = 0$, $c_3 = 1$ implies $A\vec{\mathbf{v}}_3 = \lambda_3\vec{\mathbf{v}}_3$. This proves:

**Theorem 1 (Determining Equations in Fourier's Model)**
Assume Fourier's model holds. Then the eigenvalues and eigenvectors are determined by the three equations

$$\begin{aligned} A\vec{\mathbf{v}}_1 &= \lambda_1\vec{\mathbf{v}}_1, \\ A\vec{\mathbf{v}}_2 &= \lambda_2\vec{\mathbf{v}}_2, \\ A\vec{\mathbf{v}}_3 &= \lambda_3\vec{\mathbf{v}}_3. \end{aligned}$$

The three relations of the theorem can be distilled into one homogeneous matrix–vector equation

$$A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}.$$

Write it as $A\vec{\mathbf{x}} - \lambda\vec{\mathbf{x}} = \vec{\mathbf{0}}$, then replace $\lambda\vec{\mathbf{x}}$ by $\lambda I\vec{\mathbf{x}}$ to obtain the standard form[1]

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

Let $B = A - \lambda I$. The equation $B\vec{\mathbf{v}} = \vec{\mathbf{0}}$ has a nonzero solution $\vec{\mathbf{v}}$ if and only if there are infinitely many solutions. Because the matrix is square, infinitely many solutions occurs if and only if $\mathbf{rref}(B)$ has a row of zeros. Determinant theory gives a more concise statement: $\det(B) = 0$ if and only if $B\vec{\mathbf{v}} = \vec{\mathbf{0}}$ has infinitely many solutions. This proves:

**Theorem 2 (Characteristic Equation)**
If Fourier's model holds, then the eigenvalues $\lambda_1$, $\lambda_2$, $\lambda_3$ are roots $\lambda$ of the polynomial equation

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

[1] Identity $I$ is required to factor out the matrix $A - \lambda I$. It is wrong to factor out $A - \lambda$, because $A$ is $3 \times 3$ and $\lambda$ is $1 \times 1$, incompatible sizes for matrix addition.

The equation is called the **characteristic equation**. The **characteristic polynomial** is the polynomial on the left, normally obtained by cofactor expansion or the triangular rule.

**An Illustration.**

$$\begin{aligned}\det\left(\begin{pmatrix}1 & 3\\ 1 & 2\end{pmatrix} - \lambda\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right) &= \begin{vmatrix}1-\lambda & 3\\ 1 & 2-\lambda\end{vmatrix}\\ &= (1-\lambda)(2-\lambda)-6\\ &= \lambda^2-3\lambda-4\\ &= (\lambda+1)(\lambda-4).\end{aligned}$$

The characteristic equation $\lambda^2 - 3\lambda - 4 = 0$ has roots $\lambda_1 = -1$, $\lambda_2 = 4$. The characteristic polynomial is $\lambda^2 - 3\lambda - 4$.

**Theorem 3 (Finding Eigenvectors of $A$)**
For each root $\lambda$ of the characteristic equation, write the frame sequence for $B = A - \lambda I$ with last frame $\mathbf{rref}(B)$, followed by solving for the general solution $\vec{\mathbf{v}}$ of the homogeneous equation $B\vec{\mathbf{v}} = \vec{\mathbf{0}}$. Solution $\vec{\mathbf{v}}$ uses invented parameter names $t_1$, $t_2$, .... The vector basis answers $\partial_{t_1}\vec{\mathbf{v}}$, $\partial_{t_2}\vec{\mathbf{v}}$, ...are independent **eigenvectors** of $A$ paired to eigenvalue $\lambda$.

**Proof**: The equation $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$ is equivalent to $B\vec{\mathbf{v}} = \vec{\mathbf{0}}$. Because $\det(B) = 0$, then this system has infinitely many solutions, which implies the frame sequence starting at $B$ ends with $\mathbf{rref}(B)$ having at least one row of zeros. The general solution then has one or more free variables which are assigned invented symbols $t_1$, $t_2$, ..., and then the vector basis is obtained by from the corresponding list of partial derivatives. Each basis element is a nonzero solution of $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$. By construction, the basis elements (eigenvectors for $\lambda$) are collectively independent. The proof is complete.

The theorem implies that a $3 \times 3$ matrix $A$ with eigenvalues 1, 2, 3 causes three frame sequences to be computed, each sequence producing one eigenvector. In contrast, if $A$ has eigenvalues 1, 1, 1, then only one frame sequence is computed.

**Definition 1 (Eigenpair)**
An **eigenpair** is an eigenvalue $\lambda$ together with a matching eigenvector $\vec{\mathbf{v}} \neq \vec{\mathbf{0}}$ satisfying the equation $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$. The pairing implies that scale factor $\lambda$ is applied to direction $\vec{\mathbf{v}}$.

A $3\times 3$ matrix $A$ for which Fourier's model holds has eigenvalues $\lambda_1$, $\lambda_2$, $\lambda_3$ and corresponding eigenvectors $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$. The **eigenpairs** of $A$ are

$$(\lambda_1, \vec{\mathbf{v}}_1), (\lambda_2, \vec{\mathbf{v}}_2), (\lambda_3, \vec{\mathbf{v}}_3).$$

**Theorem 4 (Independence of Eigenvectors)**
If $(\lambda_1, \vec{\mathbf{v}}_1)$ and $(\lambda_2, \vec{\mathbf{v}}_2)$ are two eigenpairs of $A$ and $\lambda_1 \neq \lambda_2$, then $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$ are independent.

More generally, if $(\lambda_1, \vec{\mathbf{v}}_1)$, ..., $(\lambda_k, \vec{\mathbf{v}}_k)$ are eigenpairs of $A$ corresponding to distinct eigenvalues $\lambda_!$, ..., $\lambda_k$, then $\vec{\mathbf{v}}_1$, ..., $\vec{\mathbf{v}}_k$ are independent.

**Proof**: Let's solve $c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 = \vec{\mathbf{0}}$ for $c_1$, $c_2$. Apply $A$ to this equation, then $c_1\lambda_1\vec{\mathbf{v}}_1 + c_2\lambda_2\vec{\mathbf{v}}_2 = \vec{\mathbf{0}}$. Multiply the first equation by $\lambda_2$ and subtract from the second equation to get $c_1(\lambda_1 - \lambda_2)\vec{\mathbf{v}}_1 = \vec{\mathbf{0}}$. Because $\lambda_1 \neq \lambda_2$, cancellation gives $c_1\vec{\mathbf{v}}_1 = \vec{\mathbf{0}}$. The assumption $\vec{\mathbf{v}}_1 \neq \vec{\mathbf{0}}$ implies $c_1 = 0$. Similarly, $c_2 = 0$. This proves $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$ are independent.

The general case is proved by induction on $k$. The case $k = 1$ follows because a nonzero vector is an independent set. Assume it holds for $k - 1$ and let's prove it for $k$, when $k > 1$. We solve

$$c_1\vec{\mathbf{v}}_1 + \cdots + c_k\vec{\mathbf{v}}_k = \vec{\mathbf{0}}$$

for $c_1$, ..., $c_k$. Apply $A$ to this equation, which effectively replaces each $c_i$ by $\lambda_i c_i$. Then multiply the first equation by $\lambda_1$ and subtract the two equations to get

$$c_2(\lambda_1 - \lambda_2)\vec{\mathbf{v}}_1 + \cdots + c_k(\lambda_1 - \lambda_k)\vec{\mathbf{v}}_k = \vec{\mathbf{0}}.$$

By the induction hypothesis, all coefficients are zero. Because $\lambda_1 - \lambda_i \neq 0$ for $i > 1$, then $c_2$ through $c_k$ are zero. Return to the first equation to obtain $c_1\vec{\mathbf{v}}_1 = \vec{\mathbf{0}}$. Because $\vec{\mathbf{v}}_1 \neq \vec{\mathbf{0}}$, then $c_1 = 0$. This finishes the induction.

**Definition 2 (Diagonalizable Matrix)**
A square matrix $A$ for which Fourier's model holds is called **diagonalizable**. The $n \times n$ matrix $A$ has $n$ eigenpairs with independent eigenvectors.

## Eigenanalysis Facts.

1. An **eigenvalue** $\lambda$ of a triangular matrix $A$ is one of the diagonal entries. If $A$ is non-triangular, then an eigenvalue is found as a root $\lambda$ of $\det(A - \lambda I) = 0$.

2. An eigenvalue of $A$ can be zero, positive, negative or even complex. It is a pure number, with a physical meaning inherited from the model, e.g., a scale factor.

3. An **eigenvector** for eigenvalue $\lambda$ (a scale factor) is a nonzero direction $\vec{\mathbf{v}}$ of application satisfying $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$. It is found from a frame sequence starting at $B = A - \lambda I$ and ending at $\mathbf{rref}(B)$. Independent eigenvectors are computed from the general solution as partial derivatives $\partial/\partial t_1$, $\partial/\partial t_2$, ....

4. If a $3 \times 3$ matrix has three independent eigenvectors, then they collectively form in $\mathcal{R}^3$ a **basis** or **coordinate system**.

## Eigenpair Packages

The eigenpairs of a $3 \times 3$ matrix for which Fourier's model holds are labeled

$$(\lambda_1, \vec{\mathbf{v}}_1), \quad (\lambda_2, \vec{\mathbf{v}}_2), \quad (\lambda_3, \vec{\mathbf{v}}_3).$$

An **eigenvector package** is a matrix package $\mathcal{P}$ of eigenvectors $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ given by

$$\mathcal{P} = \mathbf{aug}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3).$$

An **eigenvalue package** is a matrix package $\mathcal{D}$ of eigenvalues given by

$$\mathcal{D} = \mathbf{diag}(\lambda_1, \lambda_2, \lambda_3).$$

Important is the *pairing* that is inherited from the eigenpairs, which dictates the packaging order of the eigenvectors and eigenvalues. Matrices $\mathcal{P}$, $\mathcal{D}$ are **not unique**: possible are $3!$ $(= 6)$ column permutations.

**An Example**. The eigenvalues for the data conversion problem (2) are $\lambda_1 = 1$, $\lambda_2 = 0.001$, $\lambda_3 = 0.01$ and the eigenvectors $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ are the columns of the identity matrix $I$, given by (3). Then the eigenpair packages are

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.001 & 0 \\ 0 & 0 & 0.01 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

**Theorem 5 (Eigenpair Packages)**
Let $\mathcal{P}$ be a matrix package of eigenvectors and $\mathcal{D}$ the corresponding matrix package of eigenvalues. Then for all vectors $\vec{\mathbf{c}}$,

$$A\mathcal{P}\vec{\mathbf{c}} = \mathcal{P}\mathcal{D}\vec{\mathbf{c}}.$$

**Proof**: The result is valid for $n \times n$ matrices. We prove it for $3 \times 3$ matrices. The two sides of the equation are expanded as follows.

$$\begin{aligned}
\mathcal{P}\mathcal{D}\vec{\mathbf{c}} &= \mathcal{P}\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} && \text{Expand RHS.} \\
&= \mathcal{P}\begin{pmatrix} \lambda_1 c_1 \\ \lambda_2 c_2 \\ \lambda_3 c_3 \end{pmatrix} \\
&= \lambda_1 c_1 \vec{\mathbf{v}}_1 + \lambda_2 c_2 \vec{\mathbf{v}}_2 + \lambda_3 c_3 \vec{\mathbf{v}}_3 && \text{Because } \mathcal{P} \text{ has columns } \vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3. \\
A\mathcal{P}\vec{\mathbf{c}} &= A(c_1 \vec{\mathbf{v}}_2 + c_2 \vec{\mathbf{v}}_2 + c_3 \vec{\mathbf{v}}_3) && \text{Expand LHS.} \\
&= c_1 \lambda_1 \vec{\mathbf{v}}_1 + c_2 \lambda_2 \vec{\mathbf{v}}_2 + c_3 \lambda_3 \vec{\mathbf{v}}_3 && \text{Fourier's model.}
\end{aligned}$$

## The Equation $AP = PD$

The question of Fourier's model holding for a given $3 \times 3$ matrix $A$ is settled here. According to the result, a matrix $A$ for which Fourier's model holds can be constructed by the formula $A = PDP^{-1}$ where $D$ is any diagonal matrix and $P$ is an invertible matrix.

**Theorem 6 ($AP = PD$)**
Fourier's model $A(c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3) = c_1\lambda_1\vec{\mathbf{v}}_1 + c_2\lambda_2\vec{\mathbf{v}}_2 + c_3\lambda_3\vec{\mathbf{v}}_3$ holds for eigenpairs $(\lambda_1, \vec{\mathbf{v}}_1)$, $(\lambda_2, \vec{\mathbf{v}}_2)$, $(\lambda_3, \vec{\mathbf{v}}_3)$ if and only if the packages

$$P = \mathbf{aug}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3), \quad D = \mathbf{diag}(\lambda_1, \lambda_2, \lambda_3)$$

satisfy the two requirements

**1**. Matrix $P$ is invertible, e.g., $\det(\mathcal{P}) \neq 0$.

**2**. Matrix $A = PDP^{-1}$, or equivalently, $AP = PD$.

**Proof**: Assume Fourier's model holds. Define $P = \mathcal{P}$ and $D = \mathcal{D}$, the eigenpair packages. Then **1** holds, because the columns of $P$ are independent. By Theorem 5, $AP\vec{\mathbf{c}} = PD\vec{\mathbf{c}}$ for all vectors $\vec{\mathbf{c}}$. Taking $\vec{\mathbf{c}}$ equal to a column of the identity matrix $I$ implies the columns of $AP$ and $PD$ are identical, that is, $AP = PD$. A multiplication of $AP = PD$ by $P^{-1}$ gives **2**.

Conversely, let $P$ and $D$ be given packages satisfying **1**, **2**. Define $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ to be the columns of $P$. Then the columns pass the rank test, because $P$ is invertible, proving independence of the columns. Define $\lambda_1$, $\lambda_2$, $\lambda_3$ to be the diagonal elements of $D$. Using $AP = PD$, we calculate the two sides of $AP\vec{\mathbf{c}} = PD\vec{\mathbf{c}}$ as in the proof of Theorem 5, which shows that $\vec{\mathbf{x}} = c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_2\vec{\mathbf{v}}_3$ implies $A\vec{\mathbf{x}} = c_1\lambda_1\vec{\mathbf{v}}_1 + c_2\lambda_2\vec{\mathbf{v}}_2 + c_3\lambda_3\vec{\mathbf{v}}_3$. Hence Fourier's model holds.

## The Matrix Eigenanalysis Method

The preceding discussion of data conversion now gives way to abstract definitions which distill the essential theory of eigenanalysis. All of this is algebra, devoid of motivation or application.

**Definition 3 (Eigenpair)**
A pair $(\lambda, \vec{\mathbf{v}})$, where $\vec{\mathbf{v}} \neq \vec{\mathbf{0}}$ is a vector and $\lambda$ is a complex number, is called an **eigenpair** of the $n \times n$ matrix $A$ provided

$$A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}} \quad (\vec{\mathbf{v}} \neq \vec{\mathbf{0}} \text{ required}). \tag{4}$$

The **nonzero** requirement in (4) results from seeking directions for a coordinate system: the zero vector is not a direction. Any vector $\vec{\mathbf{v}} \neq \vec{\mathbf{0}}$ that satisfies (4) is called an **eigenvector** for $\lambda$ and the value $\lambda$ is called an **eigenvalue** of the square matrix $A$. The algorithm:

**Theorem 7 (Algebraic Eigenanalysis)**
Eigenpairs $(\lambda, \vec{\mathbf{v}})$ of an $n \times n$ matrix $A$ are found by this two-step algorithm:

> **Step 1 (College Algebra)**. Solve for eigenvalues $\lambda$ in the $n$th order polynomial equation $\det(A - \lambda I) = 0$.
>
> **Step 2 (Linear Algebra)**. For a given root $\lambda$ from **Step 1**, a corresponding eigenvector $\vec{\mathbf{v}} \neq \vec{\mathbf{0}}$ is found by applying the **rref** method[2] to the homogeneous linear equation
>
> $$(A - \lambda I)\vec{\mathbf{v}} = \vec{\mathbf{0}}.$$
>
> The reported answer for $\vec{\mathbf{v}}$ is routinely the list of partial derivatives $\partial_{t_1}\vec{\mathbf{v}}$, $\partial_{t_2}\vec{\mathbf{v}}$, $\ldots$, where $t_1$, $t_2$, $\ldots$ are invented symbols assigned to the free variables.

The reader is asked to apply the algorithm to the identity matrix $I$; then **Step 1** gives $n$ duplicate answers $\lambda = 1$ and **Step 2** gives $n$ answers, the columns of the identity matrix $I$.

**Proof**: The equation $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$ is equivalent to $(A - \lambda I)\vec{\mathbf{v}} = \vec{\mathbf{0}}$, which is a set of homogeneous equations, consistent always because of the solution $\vec{\mathbf{v}} = \vec{\mathbf{0}}$.

Fix $\lambda$ and define $B = A - \lambda I$. We show that an eigenpair $(\lambda, \vec{\mathbf{v}})$ exists with $\vec{\mathbf{v}} \neq \vec{\mathbf{0}}$ if and only if $\det(B) = 0$, i.e., $\det(A - \lambda I) = 0$. There is a unique solution $\vec{\mathbf{v}}$ to the homogeneous equation $B\vec{\mathbf{v}} = \vec{\mathbf{0}}$ exactly when Cramer's rule applies, in which case $\vec{\mathbf{v}} = \vec{\mathbf{0}}$ is the unique solution. All that Cramer's rule requires is $\det(B) \neq 0$. Therefore, an eigenpair exists exactly when Cramer's rule fails to apply, which is when the determinant of coefficients is zero: $\det(B) = 0$.

Eigenvectors for $\lambda$ are found from the general solution to the system of equations $B\vec{\mathbf{v}} = 0$ where $B = A - \lambda I$. The **rref** method produces systematically a reduced echelon system from which the general solution $\vec{\mathbf{v}}$ is written, depending on invented symbols $t_1, \ldots, t_k$. Since there is never a unique solution, at least one free variable exists. In particular, the last frame **rref**$(B)$ of the sequence has a row of zeros, which is a useful sanity test.

The **basis of eigenvectors** for $\lambda$ is obtained from the general solution $\vec{\mathbf{v}}$, which is a linear combination involving the parameters $t_1, \ldots, t_k$. The **basis elements** are reported as the list of partial derivatives $\partial_{t_1}\vec{\mathbf{v}}, \ldots, \partial_{t_k}\vec{\mathbf{v}}$.

## Diagonalization

A square matrix $A$ is called **diagonalizable** provided $AP = PD$ for some diagonal matrix $D$ and invertible matrix $P$. The preceding discussions imply that $D$ must be a package of eigenvalues of $A$ and $P$ must be the corresponding package of eigenvectors of $A$. The requirement on $P$

[2] For $B\vec{\mathbf{v}} = \vec{\mathbf{0}}$, the frame sequence begins with $B$, instead of **aug**$(B, \vec{\mathbf{0}})$. The sequence ends with **rref**$(B)$. Then the reduced echelon system is written, followed by assignment of free variables and display of the general solution $\vec{\mathbf{v}}$.

to be invertible is equivalent to asking that the eigenvectors of $A$ be independent and equal in number to the column dimension of $A$.

The matrices $A$ for which Fourier's model is valid is precisely the class of diagonalizable matrices. This class is not the set of all square matrices: there are matrices $A$ for which Fourier's model is invalid. They are called **non-diagonalizable matrices**.

Theorem 4 implies that the construction for eigenvector package $P$ always produces independent columns. Even if $A$ has fewer than $n$ eigenpairs, the construction still produces independent eigenvectors. In such **non-diagonalizable** cases, caused by insufficient columns to form $P$, matrix $A$ must have an eigenvalue of multiplicity greater than one.

If all eigenvalues are distinct, then the correct number of independent eigenvectors were found and $A$ is then **diagonalizable** with packages $D$, $P$ satisfying $AP = PD$. This proves the following result.

**Theorem 8 (Distinct Eigenvalues)**
If an $n \times n$ matrix $A$ has $n$ distinct eigenvalues, then it has $n$ eigenpairs and $A$ is diagonalizable with eigenpair packages $D$, $P$ satisfying $AP = PD$.

## Examples

1 **Example (Computing $2 \times 2$ Eigenpairs)**

Find all eigenpairs of the $2 \times 2$ matrix $A = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}$.

**Solution**:
**College Algebra**. The eigenvalues are $\lambda_1 = 1$, $\lambda_2 = -1$. Details:

$$0 = \det(A - \lambda I) \quad \text{Characteristic equation.}$$
$$= \begin{vmatrix} 1-\lambda & 0 \\ 2 & -1-\lambda \end{vmatrix} \quad \text{Subtract } \lambda \text{ from the diagonal.}$$
$$= (1-\lambda)(-1-\lambda) \quad \text{Sarrus' rule.}$$

**Linear Algebra**. The eigenpairs are $\left(1, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$, $\left(-1, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$. Details:

**Eigenvector for** $\lambda_1 = 1$.

$$A - \lambda_1 I = \begin{pmatrix} 1-\lambda_1 & 0 \\ 2 & -1-\lambda_1 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 \\ 2 & -2 \end{pmatrix}$$
$$\approx \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \quad \text{Swap and multiply rules.}$$
$$= \mathbf{rref}(A - \lambda_1 I) \quad \text{Reduced echelon form.}$$

The partial derivative $\partial_{t_1}\vec{\mathbf{v}}$ of the general solution $x = t_1$, $y = t_1$ is eigenvector $\vec{\mathbf{v}}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

**Eigenvector for** $\lambda_2 = -1$.

$$\begin{aligned} A - \lambda_2 I &= \begin{pmatrix} 1-\lambda_2 & 0 \\ 2 & -1-\lambda_2 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 \\ 2 & 0 \end{pmatrix} \\ &\approx \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} && \text{Combination and multiply.} \\ &= \mathbf{rref}(A - \lambda_2 I) && \text{Reduced echelon form.} \end{aligned}$$

The partial derivative $\partial_{t_1}\vec{\mathbf{v}}$ of the general solution $x = 0$, $y = t_1$ is eigenvector $\vec{\mathbf{v}}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

**2 Example (Computing $2 \times 2$ Complex Eigenpairs)**

Find all eigenpairs of the $2 \times 2$ matrix $A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$.

**Solution**:

**College Algebra**. The eigenvalues are $\lambda_1 = 1 + 2i$, $\lambda_2 = 1 - 2i$. Details:

$$\begin{aligned} 0 &= \det(A - \lambda I) && \text{Characteristic equation.} \\ &= \begin{vmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{vmatrix} && \text{Subtract } \lambda \text{ from the diagonal.} \\ &= (1-\lambda)^2 + 4 && \text{Sarrus' rule.} \end{aligned}$$

The roots $\lambda = 1 \pm 2i$ are found from the quadratic formula after expanding $(1-\lambda)^2 + 4 = 0$. Alternatively, use $(1-\lambda)^2 = -4$ and take square roots.

**Linear Algebra**. The eigenpairs are $\left(1+2i, \begin{pmatrix} -i \\ 1 \end{pmatrix}\right)$, $\left(1-2i, \begin{pmatrix} i \\ 1 \end{pmatrix}\right)$.

**Eigenvector for** $\lambda_1 = 1 + 2i$.

$$\begin{aligned} A - \lambda_1 I &= \begin{pmatrix} 1-\lambda_1 & 2 \\ -2 & 1-\lambda_1 \end{pmatrix} \\ &= \begin{pmatrix} -2i & 2 \\ -2 & -2i \end{pmatrix} \\ &\approx \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} && \text{Multiply rule.} \\ &\approx \begin{pmatrix} 0 & 0 \\ 1 & i \end{pmatrix} && \text{Combination rule, multiplier}{=}{-i}. \\ &\approx \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} && \text{Swap rule.} \\ &= \mathbf{rref}(A - \lambda_1 I) && \text{Reduced echelon form.} \end{aligned}$$

The partial derivative $\partial_{t_1}\vec{\mathbf{v}}$ of the general solution $x = -it_1$, $y = t_1$ is eigenvector $\vec{\mathbf{v}}_1 = \begin{pmatrix} -i \\ 1 \end{pmatrix}$.

**Eigenvector for** $\lambda_2 = 1 - 2i$.
The problem $(A - \lambda_2 I)\vec{\mathbf{v}} = \vec{\mathbf{0}}$ has solution $\vec{\mathbf{v}} = \overline{\vec{\mathbf{v}}_1}$, because taking conjugates across the equation gives $(A - \overline{\lambda_2} I)\overline{\vec{\mathbf{v}}} = \vec{\mathbf{0}}$; then $\lambda_1 = \overline{\lambda_2}$ gives $\vec{\mathbf{v}} = \overline{\vec{\mathbf{v}}_1} = \begin{pmatrix} i \\ 1 \end{pmatrix}$.

## 3 Example (Computing $3 \times 3$ Eigenpairs)

Find all eigenpairs of the $3 \times 3$ matrix $A = \begin{pmatrix} 1 & 2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

**Solution**:
**College Algebra**. The eigenvalues are $\lambda_1 = 1 + 2i$, $\lambda_2 = 1 - 2i$, $\lambda_3 = 3$. Details:

$$\begin{aligned}
0 &= \det(A - \lambda I) && \text{Characteristic equation.} \\
&= \begin{vmatrix} 1-\lambda & 2 & 0 \\ -2 & 1-\lambda & 0 \\ 0 & 0 & 3-\lambda \end{vmatrix} && \text{Subtract } \lambda \text{ from the diagonal.} \\
&= ((1-\lambda)^2 + 4)(3 - \lambda) && \text{Cofactor rule and Sarrus' rule.}
\end{aligned}$$

Root $\lambda = 3$ is found from the factored form above. The roots $\lambda = 1 \pm 2i$ are found from the quadratic formula after expanding $(1-\lambda)^2 + 4 = 0$. Alternatively, take roots across $(\lambda - 1)^2 = -4$.

**Linear Algebra**.

The eigenpairs are $\left(1 + 2i, \begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix}\right)$, $\left(1 - 2i, \begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix}\right)$, $\left(3, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)$.

**Eigenvector for** $\lambda_1 = 1 + 2i$.

$$\begin{aligned}
A - \lambda_1 I &= \begin{pmatrix} 1-\lambda_1 & 2 & 0 \\ -2 & 1-\lambda_1 & 0 \\ 0 & 0 & 3-\lambda_1 \end{pmatrix} \\
&= \begin{pmatrix} -2i & 2 & 0 \\ -2 & -2i & 0 \\ 0 & 0 & 2-2i \end{pmatrix} \\
&\approx \begin{pmatrix} i & -1 & 0 \\ 1 & i & 0 \\ 0 & 0 & 1 \end{pmatrix} && \text{Multiply rule.} \\
&\approx \begin{pmatrix} 0 & 0 & 0 \\ 1 & i & 0 \\ 0 & 0 & 1 \end{pmatrix} && \text{Combination rule, factor}=-i. \\
&\approx \begin{pmatrix} 1 & i & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} && \text{Swap rule.} \\
&= \mathbf{rref}(A - \lambda_1 I) && \text{Reduced echelon form.}
\end{aligned}$$

The partial derivative $\partial_{t_1}\vec{\mathbf{v}}$ of the general solution $x = -it_1$, $y = t_1$, $z = 0$ is eigenvector $\vec{\mathbf{v}}_1 = \begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix}$.

**Eigenvector for** $\lambda_2 = 1 - 2i$.

The problem $(A - \lambda_2 I)\vec{\mathbf{v}}_2 = \vec{\mathbf{0}}$ has solution $\vec{\mathbf{v}}_2 = \overline{\vec{\mathbf{v}}_1}$. To see why, take conjugates across the equation to give $(\overline{A} - \overline{\lambda_2} I)\overline{\vec{\mathbf{v}}_2} = \overline{\vec{\mathbf{0}}}$. Then $\overline{A} = A$ ($A$ is real) and $\lambda_1 = \overline{\lambda_2}$ gives $(A - \lambda_1 I)\overline{\vec{\mathbf{v}}_2} = 0$. Then $\overline{\vec{\mathbf{v}}_2} = \vec{\mathbf{v}}_1$. Finally, $\vec{\mathbf{v}}_2 = \overline{\overline{\vec{\mathbf{v}}_2}} = \overline{\vec{\mathbf{v}}_1} = \begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix}$.

**Eigenvector for** $\lambda_3 = 3$.

$$\begin{aligned} A - \lambda_3 I &= \begin{pmatrix} 1-\lambda_3 & 2 & 0 \\ -2 & 1-\lambda_3 & 0 \\ 0 & 0 & 3-\lambda_3 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 2 & 0 \\ -2 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &\approx \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} && \text{Multiply rule.} \\ &\approx \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} && \text{Combination and multiply.} \\ &= \mathbf{rref}(A - \lambda_3 I) && \text{Reduced echelon form.} \end{aligned}$$

The partial derivative $\partial_{t_1}\vec{\mathbf{v}}$ of the general solution $x = 0$, $y = 0$, $z = t_1$ is eigenvector $\vec{\mathbf{v}}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

**4 Example (Decomposition $A = PDP^{-1}$)**

Decompose $A = PDP^{-1}$ where $P$, $D$ are eigenvector and eigenvalue packages, respectively, for the $3 \times 3$ matrix

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

Write explicitly Fourier's model in vector-matrix notation.

**Solution**: By the preceding example, the eigenpairs are

$$\left(1+2i, \begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix}\right), \quad \left(1-2i, \begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix}\right), \quad \left(3, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right).$$

The packages are therefore

$$D = \mathbf{diag}(1+2i, 1-2i, 3), \quad P = \begin{pmatrix} -i & i & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

**Fourier's model**. The action of $A$ in the model

$$A\left(c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3\right) = c_1\lambda_1\vec{\mathbf{v}}_1 + c_2\lambda_2\vec{\mathbf{v}}_2 + c_3\lambda_3\vec{\mathbf{v}}_3$$

is to replace the basis $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ by scaled vectors $\lambda_1\vec{\mathbf{v}}_1$, $\lambda_2\vec{\mathbf{v}}_2$, $\lambda_3\vec{\mathbf{v}}_3$. In vector form, the model is

$$AP\vec{\mathbf{c}} = PD\vec{\mathbf{c}}, \quad \vec{\mathbf{c}} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

Then the action of $A$ is to replace eigenvector package $P$ by the re-scaled package $PD$. Explicitly,

$$\begin{aligned} \vec{\mathbf{x}} &= c_1\begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix} + c_2\begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix} + c_3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ implies} \\ A\vec{\mathbf{x}} &= c_1(1+2i)\begin{pmatrix} -i \\ 1 \\ 0 \end{pmatrix} + c_2(1-2i)\begin{pmatrix} i \\ 1 \\ 0 \end{pmatrix} + c_3(3)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \end{aligned}$$

## 5 Example (Diagonalization I)

Report **diagonalizable** or **non-diagonalizable** for the $4 \times 4$ matrix

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

If $A$ is diagonalizable, then assemble and report eigenvalue and eigenvector packages $D$, $P$.

**Solution**: The matrix $A$ is **non-diagonalizable**, because it fails to have 4 eigenpairs. The details:

**Eigenvalues**.

$$\begin{aligned} 0 &= \det(A - \lambda I) && \text{Characteristic equation.} \\ &= \begin{vmatrix} 1-\lambda & 2 & 0 & 0 \\ -2 & 1-\lambda & 0 & 0 \\ 0 & 0 & 3-\lambda & 1 \\ 0 & 0 & 0 & 3-\lambda \end{vmatrix} \\ &= \begin{vmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{vmatrix}(3-\lambda)^2 && \text{Cofactor expansion applied twice.} \\ &= \left((1-\lambda)^2 + 4\right)(3-\lambda)^2 && \text{Sarrus' rule.} \end{aligned}$$

The roots are $1 \pm 2i$, 3, 3, listed according to multiplicity.

**Eigenpairs**. They are

$$\left(1+2i, \begin{pmatrix} -i \\ 1 \\ 0 \\ 0 \end{pmatrix}\right), \quad \left(1-2i, \begin{pmatrix} i \\ 1 \\ 0 \\ 0 \end{pmatrix}\right), \quad \left(3, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}\right).$$

Because only three eigenpairs exist, instead of four, then the matrix $A$ is **non-diagonalizable**. Details:

**Eigenvector for** $\lambda_1 = 1 + 2i$.

$$\begin{aligned}
A - \lambda_1 I &= \begin{pmatrix} 1-\lambda_1 & 2 & 0 & 0 \\ -2 & 1-\lambda_1 & 0 & 0 \\ 0 & 0 & 3-\lambda_1 & 1 \\ 0 & 0 & 0 & 3-\lambda_1 \end{pmatrix} \\
&= \begin{pmatrix} -2i & 2 & 0 & 0 \\ -2 & -2i & 0 & 0 \\ 0 & 0 & 2-2i & 1 \\ 0 & 0 & 0 & 2-2i \end{pmatrix} \\
&\approx \begin{pmatrix} -i & 1 & 0 & 0 \\ -1 & -i & 0 & 0 \\ 0 & 0 & 2-2i & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} && \text{Multiply rule, three times.} \\
&\approx \begin{pmatrix} -i & 1 & 0 & 0 \\ -1 & -i & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} && \text{Combination and multiply rule.} \\
&\approx \begin{pmatrix} 1 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} && \text{Combination and multiply rule.} \\
&= \mathbf{rref}(A - \lambda_1 I) && \text{Reduced echelon form.}
\end{aligned}$$

The general solution is $x_1 = -it_1$, $x_2 = t_1$, $x_3 = 0$, $x_4 = 0$. Then $\partial_{t_1}$ applied to this solution gives the reported eigenpair.

**Eigenvector for** $\lambda_2 = 1 - 2i$.
Because $\lambda_2$ is the conjugate of $\lambda_1$ and $A$ is real, then an eigenpair for $\lambda_2$ is found by taking the complex conjugate of the eigenpair reported for $\lambda_1$.

**Eigenvector for** $\lambda_3 = 3$. In theory, there can be one or two eigenpairs to report. It turns out there is only one, because of the following details.

$$\begin{aligned}
A - \lambda_3 I &= \begin{pmatrix} 1-\lambda_3 & 2 & 0 & 0 \\ -2 & 1-\lambda_3 & 0 & 0 \\ 0 & 0 & 3-\lambda_3 & 1 \\ 0 & 0 & 0 & 3-\lambda_3 \end{pmatrix} \\
&= \begin{pmatrix} -2 & 2 & 0 & 0 \\ -2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}
\end{aligned}$$

$$\approx \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{Multiply rule, two times.}$$

$$\approx \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{Combination and multiply rule.}$$

$$= \mathbf{rref}(A - \lambda_3 I) \qquad \text{Reduced echelon form.}$$

Apply $\partial_{t_1}$ to the general solution $x_1 = 0$, $x_2 = 0$, $x_3 = t_1$, $x_4 = 0$ to give the eigenvector matching the eigenpair reported above.

**6 Example (Diagonalization II)**
Report **diagonalizable** or **non-diagonalizable** for the $4 \times 4$ matrix

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

If $A$ is diagonalizable, then assemble and report eigenvalue and eigenvector packages $D$, $P$.

**Solution**: The matrix $A$ is **diagonalizable**, because it has 4 eigenpairs

$$\left(1+2i, \begin{pmatrix} -i \\ 1 \\ 0 \\ 0 \end{pmatrix}\right), \quad \left(1-2i, \begin{pmatrix} i \\ 1 \\ 0 \\ 0 \end{pmatrix}\right), \quad \left(3, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}\right), \quad \left(3, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\right).$$

Then the eigenpair packages are given by

$$D = \begin{pmatrix} -1+2i & 0 & 0 & 0 \\ 0 & 1-2i & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \quad P = \begin{pmatrix} -i & i & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The details parallel the previous example, except for the calculation of eigenvectors for $\lambda_3 = 3$. In this case, the reduced echelon form has two rows of zeros and parameters $t_1$, $t_2$ appear in the general solution. The answers given above for eigenvectors correspond to the partial derivatives $\partial_{t_1}$, $\partial_{t_2}$.

**7 Example (Non-diagonalizable Matrices)**
Verify that the matrices

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

are all non-diagonalizable.

**Solution**: Let $A$ denote any one of these matrices and let $n$ be its dimension.

First, we will decide on diagonalization, without computing eigenpairs. Assume, in order to reach a contradiction, that eigenpair packages $D$, $P$ exist with $D$ diagonal and $P$ invertible such that $AP = PD$. Because $A$ is triangular, its eigenvalues appear already on the diagonal of $A$. Only 0 is an eigenvalue and its multiplicity is $n$. Then the package $D$ of eigenvalues is the zero matrix and an equation $AP = PD$ reduces to $AP = 0$. Multiply $AP = 0$ by $P^{-1}$ to obtain $A = 0$. But $A$ is not the zero matrix, a contradiction. We conclude that $A$ is not diagonalizable.

Second, we attack the diagonalization question directly, by solving for the eigenvectors corresponding to $\lambda = 0$. The frame sequence has first frame $B = A - \lambda I$, but $B$ equals $\mathbf{rref}(B)$ and no computations are required. The resulting reduced echelon system is just $x_1 = 0$, giving $n-1$ free variables. Therefore, the eigenvectors of $A$ corresponding to $\lambda = 0$ are the last $n-1$ columns of the identity matrix $I$. Because $A$ does not have $n$ independent eigenvectors, then $A$ is not diagonalizable.

Similar examples of non-diagonalizable matrices $A$ can be constructed with $A$ having from 1 up to $n-1$ independent eigenvectors. The examples with ones on the super-diagonal and zeros elsewhere have exactly one eigenvector.

**8 Example (Fourier's 1822 Heat Model)**

Fourier's 1822 treatise *Théorie analytique de la chaleur* studied dissipation of heat from a laterally insulated welding rod with ends held at 0°C. Assume the initial heat distribution along the rod at time $t = 0$ is given as a linear combination

$$f = c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + c_3\vec{\mathbf{v}}_3.$$

Symbols $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ are in the vector space $V$ of all twice continuously differentiable functions on $0 \le x \le 1$, given explicitly as

$$\vec{\mathbf{v}}_1 = \sin \pi x, \quad \vec{\mathbf{v}}_2 = \sin 2\pi x, \quad \vec{\mathbf{v}}_3 = \sin 3\pi x.$$

**Fourier's heat model** re-scales[3] each of these vectors to obtain the temperature $u(t, x)$ at position $x$ along the rod and time $t > 0$ as the model equation

$$u(t, x) = c_1 e^{-\pi^2 t}\vec{\mathbf{v}}_1 + c_2 e^{-4\pi^2 t}\vec{\mathbf{v}}_2 + c_3 e^{-9\pi^2 t}\vec{\mathbf{v}}_3.$$

Verify that $u(t, x)$ solves Fourier's partial differential equation heat model

$$\begin{array}{rcl} \dfrac{\partial u}{\partial t} & = & \dfrac{\partial^2 u}{\partial x^2}, \\ u(0, x) & = & f(x), \quad 0 \le x \le 1, \\ u(t, 0) & = & 0, \quad \text{zero Celsius at rod's left end}, \\ u(t, 1) & = & 0, \quad \text{zero Celsius at rod's right end}. \end{array}$$

[3]The scale factors are not constants nor are they eigenvalues, but rather, they are exponential functions of $t$, as was the case for matrix differential equations $\vec{\mathbf{x}}' = A\vec{\mathbf{x}}$

**Solution**: First, we prove that the partial differential equation is satisfied by Fourier's solution $u(t,x)$. This is done by expanding the left side (LHS) and right side (RHS) of the differential equation, separately, then comparing the answers for equality.

Trigonometric functions $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ are solutions of three different linear ordinary differential equations: $u'' + \pi^2 u = 0$, $u'' + 4\pi^2 u = 0$, $u'' + 9\pi^2 u = 0$. Because of these differential equations, we can compute directly

$$\frac{\partial^2 u}{\partial x^2} = -\pi^2 c_1 e^{-\pi^2 t}\vec{\mathbf{v}}_1 - 4\pi^2 c_2 e^{-4\pi^2 t}\vec{\mathbf{v}}_2 - 9\pi^2 c_3 e^{-9\pi^2 t}\vec{\mathbf{v}}_3.$$

Similarly, computing $\partial_t u(t,x)$ involves just the differentiation of exponential functions, giving

$$\frac{\partial u}{\partial t} = -\pi^2 c_1 e^{-\pi^2 t}\vec{\mathbf{v}}_1 - 4\pi^2 c_2 e^{-4\pi^2 t}\vec{\mathbf{v}}_2 - 9\pi^2 c_3 e^{-9\pi^2 t}\vec{\mathbf{v}}_3.$$

Because the second display is exactly the first, then LHS = RHS, proving that the partial differential equation is satisfied.

The relation $u(0,x) = f(x)$ is proved by observing that each exponential factor becomes $e^0 = 1$ when $t = 0$.

The two relations $u(t,0) = u(t,1) = 0$ hold because each of $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$, $\vec{\mathbf{v}}_3$ vanish at $x = 0$ and $x = 1$. The verification is complete.

## Exercises 9.1

**Eigenanalysis.** Classify as true or false. If false, then correct the text to make it true.

**1.** The purpose of eigenanalysis is to find a coordinate system.

**2.** Diagonal matrices have all their eigenvalues on the last row.

**3.** Eigenvalues are scale factors.

**4.** Eigenvalues of a diagonal matrix cannot be zero.

**5.** Eigenvectors $\vec{\mathbf{v}}$ of a diagonal matrix can be zero.

**6.** Eigenpairs $(\lambda, \vec{\mathbf{v}})$ of a diagonal matrix $A$ satisfy the equation $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$.

**Eigenpairs of a Diagonal Matrix.** Find the eigenpairs of $A$.

**7.** $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$

**8.** $\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$

**9.** $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

**10.** $\begin{pmatrix} 0 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

**11.** $\begin{pmatrix} 7 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -6 \end{pmatrix}$

**12.** $\begin{pmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -1 \end{pmatrix}$

**Fourier's Model.**

**13.**

**Eigenanalysis Facts.**

**14.**

**Eigenpair Packages.**

**15.**

The Equation $AP = PD$.

**16.**

Matrix Eigenanalysis Method.

**17.**

Basis of Eigenvectors.

**18.**

Independence of Eigenvectors.

**19.**

Diagonalization Theory.

**20.**

Non-diagonalizable Matrices.

**21.**

Distinct Eigenvalues.

**22.** $\begin{pmatrix} 2 & 6 \\ 5 & 3 \end{pmatrix}$

**23.** $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$

**24.** $\begin{pmatrix} 2 & 6 & 2 \\ 9 & 3 & 9 \\ 1 & 3 & 1 \end{pmatrix}$

**25.** $\begin{pmatrix} 0 & 2 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 3 \end{pmatrix}$

**26.** $\begin{pmatrix} 7 & 12 & 6 \\ 2 & 2 & 2 \\ -7 & -12 & -6 \end{pmatrix}$

**27.** $\begin{pmatrix} 2 & 2 & -6 \\ -3 & -4 & 3 \\ -3 & -4 & -1 \end{pmatrix}$

Computing $2 \times 2$ Eigenpairs.

**28.**

Computing $2 \times 2$ Complex Eigenpairs.

**29.**

Computing $3 \times 3$ Eigenpairs.

**30.**

Decomposition $A = PDP^{-1}$.

**31.**

Diagonalization I

**32.**

Diagonalization II

**33.**

Non-diagonalizable Matrices

**34.**

Fourier's Heat Model

**35.**