Eigenanalysis

- What’s Eigenanalysis?
- Fourier’s Eigenanalysis Model is a Replacement Process
- Powers and Fourier’s Model
- Differential Equations and Fourier’s Model
- Fourier’s Model Illustrated
- What is Eigenanalysis? What’s an Eigenvalue? What’s an Eigenvector?
- Data Conversion Example
- History of Fourier’s Model
- Determining Equations. How to Compute Eigenpairs.
- Independence of Eigenvectors.

What's Eigenanalysis?

Matrix eigenanalysis is a computational theory for the matrix equation

$$\mathbf{y} = \boldsymbol{A}\mathbf{x}.$$

Fourier's Eigenanalysis Model

For exposition purposes, we assume $\boldsymbol{A}$ is a $\mathbf{3} \times \mathbf{3}$ matrix.

$$\begin{aligned} \mathbf{x} &= \boldsymbol{c}_1 \mathbf{v}_1 + \boldsymbol{c}_2 \mathbf{v}_2 + \boldsymbol{c}_3 \mathbf{v}_3 \text{ implies} \\ \mathbf{y} &= \boldsymbol{A}\mathbf{x} \\ &= \boldsymbol{c}_1 \boldsymbol{\lambda}_1 \mathbf{v}_1 + \boldsymbol{c}_2 \boldsymbol{\lambda}_2 \mathbf{v}_2 + \boldsymbol{c}_3 \boldsymbol{\lambda}_3 \mathbf{v}_3. \end{aligned} \tag{1}$$

Eigenanalysis Notation

The scale factors $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$, $\boldsymbol{\lambda}_3$ and independent vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ depend only on $\boldsymbol{A}$. Symbols $\boldsymbol{c}_1$, $\boldsymbol{c}_2$, $\boldsymbol{c}_3$ stand for arbitrary numbers. This implies variable $\mathbf{x}$ exhausts all possible fixed vectors in $\boldsymbol{R}^3$.

Fourier's Model is a Replacement Process

$$\boldsymbol{A}\left(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3\right) = c_1\boldsymbol{\lambda}_1\mathbf{v}_1 + c_2\boldsymbol{\lambda}_2\mathbf{v}_2 + c_3\boldsymbol{\lambda}_3\mathbf{v}_3.$$

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

Fourier's model is said to **hold** provided there exist scale factors and independent vectors satisfying (1). Fourier's model is known to fail for certain matrices $\boldsymbol{A}$.

Powers and Fourier's Model

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

$$\boldsymbol{A}^5\mathrm{x} = \boldsymbol{x}_1\boldsymbol{\lambda}_1^5\mathrm{v}_1 + \boldsymbol{x}_2\boldsymbol{\lambda}_2^5\mathrm{v}_2 + \boldsymbol{x}_3\boldsymbol{\lambda}_3^5\mathrm{v}_3.$$

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

Details for $A^3(\mathrm{x})$

Let $\mathrm{x} = \boldsymbol{x}_1\mathrm{v}_1 + \boldsymbol{x}_2\mathrm{v}_2 + \boldsymbol{x}_3\mathrm{v}_3$. Then

$$\begin{aligned}
\boldsymbol{A}^3(\boldsymbol{x}) &= \boldsymbol{A}^2(\boldsymbol{A}(\boldsymbol{x})) \\
&= \boldsymbol{A}^2(\boldsymbol{x}_1\boldsymbol{\lambda}_1\mathrm{v}_1 + \boldsymbol{x}_2\boldsymbol{\lambda}_2\mathrm{v}_2 + \boldsymbol{x}_3\boldsymbol{\lambda}_3\mathrm{v}_3) \\
&= \boldsymbol{A}(\boldsymbol{A}(\boldsymbol{x}_1\boldsymbol{\lambda}_1\mathrm{v}_1 + \boldsymbol{x}_2\boldsymbol{\lambda}_2\mathrm{v}_2 + \boldsymbol{x}_3\boldsymbol{\lambda}_3\mathrm{v}_3)) \\
&= \boldsymbol{A}(\boldsymbol{x}_1\boldsymbol{\lambda}_1^2\mathrm{v}_1 + \boldsymbol{x}_2\boldsymbol{\lambda}_2^2\mathrm{v}_2 + \boldsymbol{x}_3\boldsymbol{\lambda}_3^2\mathrm{v}_3) \\
&= \boldsymbol{x}_1\boldsymbol{\lambda}_1^3\mathrm{v}_1 + \boldsymbol{x}_2\boldsymbol{\lambda}_2^3\mathrm{v}_2 + \boldsymbol{x}_3\boldsymbol{\lambda}_3^3\mathrm{v}_3
\end{aligned}$$

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 general solution is given by the formula [Fourier's theorem, proved later]

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

which is related to Fourier's model by the symbolic formulas

$$\begin{array}{rcl} \mathbf{x}(0) &=& c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 \\ && \text{undergoes replacements } \mathbf{v}_i \longrightarrow e^{\lambda_i t}\mathbf{v}_i \text{ to obtain} \\ \mathbf{x}(t) &=& c_1 e^{\lambda_1 t}\mathbf{v}_1 + c_2 e^{\lambda_2 t}\mathbf{v}_2 + c_3 e^{\lambda_3 t}\mathbf{v}_3. \end{array}$$

Fourier's model illustrated

Let

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

$$\lambda_1 = 1, \qquad \lambda_2 = 2, \qquad \lambda_3 = -5,$$

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ -2 \\ -14 \end{pmatrix}.$$

Then Fourier's model holds (details later) and

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

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

What is Eigenanalysis?

The subject of **eigenanalysis** discovers a coordinate system $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and scale factors $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$, $\boldsymbol{\lambda}_3$ such that Fourier's model holds. Fourier's model simplifies the matrix equation $\mathbf{y} = \boldsymbol{A}\mathbf{x}$, through the formula

$$\boldsymbol{A}(\boldsymbol{c}_1\mathbf{v}_1 + \boldsymbol{c}_2\mathbf{v}_2 + \boldsymbol{c}_3\mathbf{v}_3) = \boldsymbol{c}_1\boldsymbol{\lambda}_1\mathbf{v}_1 + \boldsymbol{c}_2\boldsymbol{\lambda}_2\mathbf{v}_2 + \boldsymbol{c}_3\boldsymbol{\lambda}_3\mathbf{v}_3.$$

What's an Eigenvalue?

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

> The **eigenvalues** of a model are scale factors. Think of them as a system of units *hidden* in the matrix $\boldsymbol{A}$.

What's an Eigenvector?

Symbols $\mathbf{v_1}$, $\mathbf{v_2}$, $\mathbf{v_3}$ in Fourier's model are called eigenvectors, or *proper vectors* or *hidden vectors* or *characteristic vectors*. They are assumed independent.

> The **eigenvectors** of a model are independent **directions of application** for the scale factors (eigenvalues). Think of each eigenpair $(\boldsymbol{\lambda}, \mathbf{v})$ as a coordinate axis $\mathbf{v}$ where the action of matrix $\boldsymbol{A}$ is to move $\boldsymbol{\lambda}$ units along $\mathbf{v}$.

Data Conversion Example

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

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

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

Data Conversion Example – Continued

Problem (2) can be represented as $\mathbf{y} = \boldsymbol{A}\mathbf{x}$, where the diagonal matrix $\boldsymbol{A}$ is given by

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{\lambda}_1 & 0 & 0 \\ 0 & \boldsymbol{\lambda}_2 & 0 \\ 0 & 0 & \boldsymbol{\lambda}_3 \end{pmatrix}, \quad \boldsymbol{\lambda}_1 = 1,\ \boldsymbol{\lambda}_2 = \frac{1}{1000},\ \boldsymbol{\lambda}_3 = \frac{1}{100}.$$

Fourier's model for this matrix $\boldsymbol{A}$ is

$$\boldsymbol{A}\left(\boldsymbol{c}_1 \begin{pmatrix}1\\0\\0\end{pmatrix} + \boldsymbol{c}_2 \begin{pmatrix}0\\1\\0\end{pmatrix} + \boldsymbol{c}_3 \begin{pmatrix}0\\0\\1\end{pmatrix}\right) = \boldsymbol{c}_1\boldsymbol{\lambda}_1 \begin{pmatrix}1\\0\\0\end{pmatrix} + \boldsymbol{c}_2\boldsymbol{\lambda}_2 \begin{pmatrix}0\\1\\0\end{pmatrix} + \boldsymbol{c}_3\boldsymbol{\lambda}_3 \begin{pmatrix}0\\0\\1\end{pmatrix}$$

The **eigenvectors** $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ of the model are the columns of the identity matrix.

Summary

The **eigenvalues** of a model are **scale factors**, normally represented by symbols

$$\boldsymbol{\lambda}_1, \quad \boldsymbol{\lambda}_2, \quad \boldsymbol{\lambda}_3, \quad \ldots .$$

The **eigenvectors** of a model are independent **directions of application** for the scale factors (eigenvalues). They are normally represented by symbols

$$\mathbf{v}_1, \quad \mathbf{v}_2, \quad \mathbf{v}_3, \quad \ldots .$$

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 $\mathbf{y} = A\mathbf{x}$.

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

Determining Equations

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

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

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

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

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

Determining Equations and Linear Algebra

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

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

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

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

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

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

Theorem 2 (Characteristic Equation)
If Fourier's model holds, then the eigenvalues λ_1, λ_2, λ_3 are roots λ of the polynomial equation

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

The equation $\det(A - \lambda I) = 0$ is called the **characteristic equation**. The **characteristic polynomial** is the polynomial on the left, $\det(A - \lambda I)$, normally obtained by cofactor expansion or the triangular rule.

Theorem 3 (Finding Eigenvectors of A)
For each root λ 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 $\mathbf{v}$ of the homogeneous equation $B\mathbf{v} = \mathbf{0}$. Solution $\mathbf{v}$ uses invented symbols t_1, t_2, ... The vector basis answers $\partial_{t_1}\mathbf{v}$, $\partial_{t_2}\mathbf{v}$, ... are independent **eigenvectors** of A paired to eigenvalue λ.

Proof: The equation $A\mathbf{v} = \lambda\mathbf{v}$ is equivalent to $B\mathbf{v} = \mathbf{0}$. 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 $A\mathbf{v} = \lambda\mathbf{v}$. By construction, the basis elements (eigenvectors for λ) are collectively independent. The proof is complete.

Eigenpairs of a Matrix

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

An applied view of an eigenpair is a coordinate axis $\mathbf{v}$ and a unit system along this axis. The action of the matrix $\boldsymbol{A}$ is to move $\boldsymbol{\lambda}$ units along this axis.

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

$$(\boldsymbol{\lambda}_1, \mathbf{v}_1), (\boldsymbol{\lambda}_2, \mathbf{v}_2), (\boldsymbol{\lambda}_3, \mathbf{v}_3).$$

Eigenvectors are Independent

Theorem 4 (Independence of Eigenvectors)
If $(\boldsymbol{\lambda}_1, \mathbf{v}_1)$ and $(\boldsymbol{\lambda}_2, \mathbf{v}_2)$ are two eigenpairs of $\boldsymbol{A}$ and $\boldsymbol{\lambda}_1 \neq \boldsymbol{\lambda}_2$, then $\mathbf{v}_1$, $\mathbf{v}_2$ are independent.
More generally, if $(\boldsymbol{\lambda}_1, \mathbf{v}_1)$, ..., $(\boldsymbol{\lambda}_k, \mathbf{v}_k)$ are eigenpairs of $\boldsymbol{A}$ corresponding to distinct eigenvalues $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_k$, then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are independent.

Theorem 5 (Distinct Eigenvalues)
If an $\boldsymbol{n} \times \boldsymbol{n}$ matrix $\boldsymbol{A}$ has $\boldsymbol{n}$ distinct eigenvalues, then its eigenpairs $(\boldsymbol{\lambda}_1, \mathbf{v}_1)$, ..., $(\boldsymbol{\lambda}_n, \mathbf{v}_n)$ produce independent eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Therefore, Fourier's model holds:

$$\boldsymbol{A}\left(\sum_{i=1}^{n} \boldsymbol{c}_i \mathbf{v}_i\right) = \sum_{i=1}^{n} \boldsymbol{c}_i (\boldsymbol{\lambda}_i \mathbf{v}_i).$$