Session 13
Eigenvalue and Eigenvector

Christopher Ting

: cting@hiroshima-u.ac.jp

: +81 082-424-7268

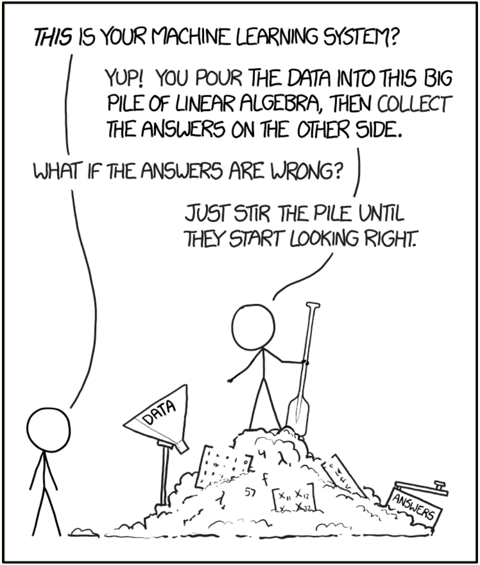
: http://cting.x10host.com/

Christopher Ting

June 3, 2020 Version 0.5

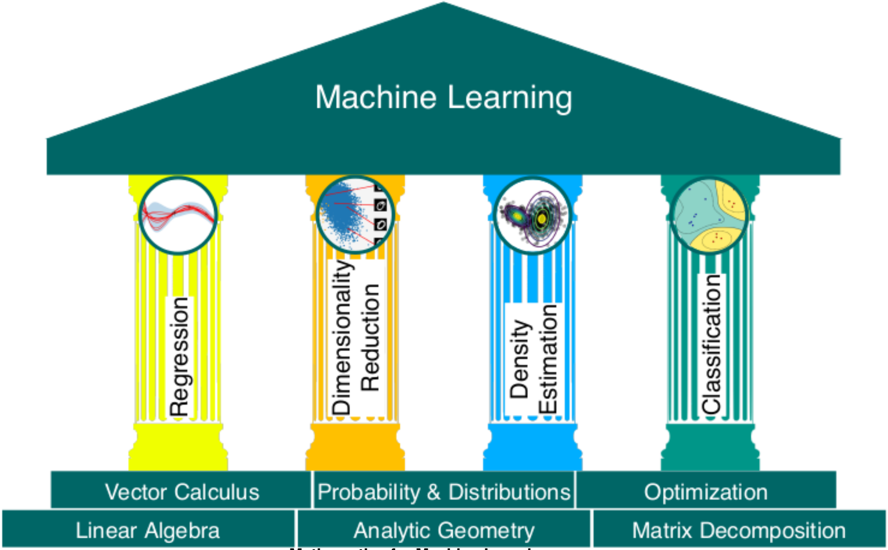
Table of Contents

Funny Inspiration



xkcd.com

Application for Machine Learning



Mathematics for Machine Learning

Learning Outcomes

- Build a case for the role of eigen analysis in machine learning.
- Define and describe eigenvalue and its multiplicity, eigenvector, eigenspace, as well as eigen polynomial.
- Solve the characteristic equations to obtain the eigenvalues.
- Given an eigenvalue, solve the defining equation to obtain a representative eigenvector,
- Explain and illustrate the abstract version of eigen analysis.
- List the properties of eigen polynomial and provide an intuitive meaning for each property.
- Analyze the dimension of a eigenspace in relation to the dimension and rank of the matrix.
- Perform diagonalization and triangularization for a given matrix.

Inspiring Quote

I like–I love calculus. I love linear algebra, probability and statistics, that kind of stuff. I just really like that.

— **Pardis Sabeti**

Preamble

- So far, we have understood that a matrix $\boldsymbol{A}$ is no different from a map (or operator) that transforms a vector $\boldsymbol{v}$ into another vector $\boldsymbol{w}$, i.e., $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{w}$, and $\boldsymbol{v} \neq \boldsymbol{w}$.
- What if $\boldsymbol{A}$ transforms some special vectors $\boldsymbol{v}$ back to its own (eigen) self, up to a proportional constant λ?

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

- Now, $\boldsymbol{A}$ is a matrix and it is possible, in principle, to have another vector that is invariant under $\boldsymbol{A}$ up to another constant λ'.
- Natural questions:
  - Given a matrix $\boldsymbol{A}$, how many pairs of $(\boldsymbol{v}, \lambda)$ are there?
  - What is the meaning of the proportional constant λ?
  - What is the meaning of $\boldsymbol{v}$?
  - How are λ_i and $\boldsymbol{v}_i$ $(i = 1, 2, \ldots, n)$ linked to the basis of $\boldsymbol{A}$?
  - etc

Eigen Value, Vector, and Space

Definition 2.1 (Eigen Value, Vector, and Space).

With respect to an n-dimensional complex-valued square matrix $\boldsymbol{A}$, suppose there exists λ and vector $\boldsymbol{v}$ such that

$$\boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v} \qquad \left(\boldsymbol{v} \in \mathfrak{C}^n \backslash \{\mathbf{0}\}, \lambda \in \mathfrak{C}\right).$$

Then, λ is said to be the **eigenvalue** of $\boldsymbol{A}$, and $\boldsymbol{v}$ the **eigenvector** of $\boldsymbol{A}$. Moreover,

$$W_{\boldsymbol{A}}(\lambda) = \left\{\boldsymbol{v} \in \mathfrak{C}^n \middle| \boldsymbol{A}\boldsymbol{v} = \lambda \boldsymbol{v}\right\}$$

is said to be the **eigenspace**.

Subspace of $\mathfrak{C}^n$

Theorem 2.2.

The eigenspace $W_{\boldsymbol{A}}(\lambda)$ is a subspace of $\mathfrak{C}^n$.

Proof.

- Note that $\mathbf{0}$, though excluded from the definition of eigenvector, is in $W_{\boldsymbol{A}}(\lambda)$.
- Let T be a linear transformation established by $\boldsymbol{A}$. For any $\boldsymbol{v}_1, \boldsymbol{v}_2 \in W_{\boldsymbol{A}}(\lambda)$, and any scalar $c \in \mathfrak{C}$,

$$T(\boldsymbol{v}_1 + \boldsymbol{v}_2) = T(\boldsymbol{v}_1) + T(\boldsymbol{v}_2) = \lambda \boldsymbol{v}_1 + \lambda \boldsymbol{v}_2 = \lambda(\boldsymbol{v}_1 + \boldsymbol{v}_2),$$
$$T(c\boldsymbol{v}_1) = cT(\boldsymbol{v}_1) = c\lambda \boldsymbol{v}_1 = \lambda(c\boldsymbol{v}_1).$$

- It follows that $\boldsymbol{v}_1 + \boldsymbol{v}_2 \in W_{\boldsymbol{A}}(\lambda)$ and $c\boldsymbol{v}_1 \in W_{\boldsymbol{A}}(\lambda)$, which shows that $W_{\boldsymbol{A}}(\lambda)$ is a subspace of $\mathfrak{C}^n$.

□

Eigen Polynomial of a Square Matrix

Definition 2.3 (Eigen Polynomial).

With respect to an n-dimensional complex-valued square matrix $\boldsymbol{A}$,

$$\Phi_{\boldsymbol{A}}(\lambda) = \det\left(\lambda \boldsymbol{I}_n - \boldsymbol{A}\right) = \det \begin{bmatrix} \lambda - a_{11} & \cdots & -a_{1n} \\ \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & \lambda - a_{nn} \end{bmatrix}$$

$\Phi_{\boldsymbol{A}}$ is called the **eigen polynomial** or **characteristic polynomial**.

Theorem

Theorem 2.4.

For a square matrix $\boldsymbol{A}$, the following statements hold:

1. *λ is the eigenvalue of $\boldsymbol{A} \iff \Phi_{\boldsymbol{A}}(\lambda) = 0$.*
2. *The eigenvector $\boldsymbol{v}$ is a non-trivial solution of*

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

which is a system of first-order equations.

Proof.

λ being the eigenvalue of $\boldsymbol{A}$ implies that $\lambda \boldsymbol{v} - \lambda \boldsymbol{A} = \boldsymbol{0}$, and vice versa. That is, it is equivalent to $(\lambda \boldsymbol{I}_n - \boldsymbol{A})\boldsymbol{v} = \boldsymbol{0}$. For the system of equations $(\lambda \boldsymbol{I}_n - \boldsymbol{A})\boldsymbol{v} = \boldsymbol{0}$ to have nontrivial solutions, according to the theorems in Lesson 3 about the conditions for trivial and nontrivial solutions, it is equivalent to $\Phi_{\boldsymbol{A}}(\lambda) = \det(\lambda \boldsymbol{I}_n - \boldsymbol{A}) = 0$. □

Eigen Equation and Multiplicity

- The equation $\Phi_{\boldsymbol{A}}(\lambda) = 0$ is called the **eigen equation** of matrix $\boldsymbol{A}$.
- For the n-dimensional matrix $\boldsymbol{A}$, the order of its eigen equation is n.
- The fundamental theorem of algebra indicates that there are n solutions of complex numbers, some of which may be duplicates, for any n-order equation.
- Hence $\Phi_{\boldsymbol{A}}(\lambda) = 0$ has at most n eigenvalues.
- Let $\lambda_1, \ldots, \lambda_r$ $(1 \leqq r \leqq n)$ be the eigenvalues of $\boldsymbol{A}$ that are different from each other, and they can be expressed as

$$\underbrace{\overbrace{\lambda_1, \ldots, \lambda_1}^{n_1}, \ldots\ldots, \overbrace{\lambda_r, \ldots, \lambda_r}^{n_r}}_{n}.$$

- For each $i = 1, 2, \ldots, r$, n_i is called the **multiplicity** of the eigenvalue λ_i.

Numerical Illustration

- Find the eigenvalues of $\begin{bmatrix} 1 & -1 \\ 3 & -2 \end{bmatrix}$

  Answer: The eigen polynomial is

$$\Phi_{\boldsymbol{A}}(\lambda) = \det\left(\lambda \boldsymbol{I}_2 - \boldsymbol{A}\right) = \det \begin{bmatrix} \lambda - 1 & 1 \\ -3 & \lambda + 2 \end{bmatrix} = \lambda^2 + \lambda + 1.$$

  The solutions of the quadratic equation are $\dfrac{-1 \pm \sqrt{3}i}{2}$.

- Find the eigenvalues of $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

  Answer: The eigen polynomial is

$$\Phi_{\boldsymbol{A}}(\lambda) = \det\left(\lambda \boldsymbol{I}_2 - \boldsymbol{A}\right) = \det \begin{bmatrix} \lambda - 1 & -2 \\ -2 & \lambda - 4 \end{bmatrix} = \lambda(\lambda - 5).$$

  The eigenvalues are 0 and 5.

Eigen Polynomial of Triangular Matrix

- The eigen polynomial of an upper triangular matrix
$$\boldsymbol{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ \text{\Large 0} & & a_{nn} \end{bmatrix}$$ is
$$\Phi_{\boldsymbol{A}}(\lambda) = \begin{vmatrix} \lambda - a_{11} & \cdots & -a_{1n} \\ & \ddots & \vdots \\ \text{\Large 0} & & \lambda - a_{nn} \end{vmatrix} = (\lambda - a_{11}) \cdots (\lambda - a_{nn}).$$
- According to Theorem 2.4, all the eigenvalues of $\boldsymbol{A}$ are the diagonal elements $a_{11}, a_{22}, \ldots, a_{nn}$.
- From the property of determinant, the same result is obtained for the lower triangular matrix.

Example 2.5

Example 2.5.

Find the eigenvalue and the eigenspace of the matrix $\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ -2 & 3 \end{bmatrix}$.

- The eigen polynomial of $\boldsymbol{A}$ is

$$\Phi_{\boldsymbol{A}}(\lambda) = \det(\lambda \boldsymbol{I}_2 - \boldsymbol{A}) = \begin{vmatrix} \lambda - 1 & -1 \\ 1 & \lambda - 3 \end{vmatrix} = (\lambda - 2)^2.$$

- Thus, the eigenvalue is 2 with a multiplicity of 2.
- Now, $2\boldsymbol{I}_2 - \boldsymbol{A} = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$, and

$$(2\boldsymbol{I}_2 - \boldsymbol{A}) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Longrightarrow x - y = 0.$$

- Therefore, the eigenspace of $\lambda = 2$ is $W_{\boldsymbol{A}}(2) = \left\{ c \begin{bmatrix} 1 \\ 1 \end{bmatrix} \middle| c \in \Re \right\}$.

Example 2.6

Example 2.6.

Find all the eigenvalues and the eigenspaces of $\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 2 \\ -1 & 2 & 1 \\ 2 & -1 & 1 \end{bmatrix}$.

- The eigen polynomial of $\boldsymbol{A}$ is

$$\Phi_{\boldsymbol{A}}(\lambda) = \det\left(\lambda \boldsymbol{I}_3 - \boldsymbol{A}\right) = \begin{vmatrix} \lambda - 1 & -1 & -2 \\ 1 & \lambda - 2 & -1 \\ -2 & 1 & \lambda - 1 \end{vmatrix} = \lambda(\lambda - 1)(\lambda - 3).$$

- The eigenvalues are 0, 1, and 3.

Example 2.6 (Cont'd)

- To solve $\Phi_{\boldsymbol{A}}(0) = 0$, we consider

$$0\boldsymbol{I}_3 - \boldsymbol{A} = -\boldsymbol{A} = \begin{bmatrix} -1 & -1 & -2 \\ 1 & -2 & -1 \\ -2 & 1 & -1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

- The eigenspace is

$$W_{\boldsymbol{A}}(0) = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathfrak{R}^3 \,\middle|\, \begin{matrix} x + z = 0 \\ y + z = 0 \end{matrix} \right\} = \left\{ c \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \,\middle|\, c \in \mathfrak{R} \right\}.$$

Example 2.6 (Cont'd)

- To solve $\Phi_{\boldsymbol{A}}(1) = 0$, we consider

$$1\boldsymbol{I}_3 - \boldsymbol{A} = \begin{bmatrix} 0 & -1 & -2 \\ 1 & -1 & -1 \\ -2 & 1 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

- The eigenspace is

$$W_{\boldsymbol{A}}(1) = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathfrak{R}^3 \middle| \begin{matrix} x + z = 0 \\ y + 2z = 0 \end{matrix} \right\} = \left\{ c \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \middle| c \in \mathfrak{R} \right\}.$$

Example 2.6 (Cont'd)

- To solve $\Phi_{\boldsymbol{A}}(3) = 0$, we consider

$$3\boldsymbol{I}_3 - \boldsymbol{A} = \begin{bmatrix} 2 & -1 & -2 \\ 1 & 1 & -1 \\ -2 & 1 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

- The eigenspace is

$$W_{\boldsymbol{A}}(3) = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathfrak{R}^3 \middle| \begin{matrix} x - z = 0 \\ y = 0 \end{matrix} \right\} = \left\{ c \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \middle| c \in \mathfrak{R} \right\}.$$

□

Linear Transformation

Definition 2.7 (Linear Transformation's Eigenvalue & Eigenvector).

Let the linear transformation defined on V be $T : V \longrightarrow V$. With respect to $\boldsymbol{v} \longrightarrow T(\boldsymbol{v})$,

$$T(\boldsymbol{v}) = \lambda \boldsymbol{v},$$

for which the scalar $\lambda \in \mathfrak{C}$ and the vector $\boldsymbol{v} \in V \backslash \{\mathbf{0}\}$ exist. Then, λ is said to be the eigenvalue of the transformation T, and $\boldsymbol{v}$ is the corresponding eigenvector.

Definition 2.8 (Linear Transformation's Eigenspace).

For any linear transformation $T : V \longrightarrow V$ and its any eigenvalue λ, the eigenspace is defined as the subset:

$$W_T(\lambda) = \left\{ \boldsymbol{v} \in V \middle| T(\boldsymbol{v}) = \lambda \boldsymbol{v} \right\}.$$

$W_T(\lambda)$ is a Subspace

Proof.

- Write the identity transformation as $\mathbf{1}$.
- By the definition of eigenspace, for any $\boldsymbol{v} \in W_T(\lambda)$:

$$\lambda \boldsymbol{v} - T(\boldsymbol{v}) = \mathbf{0} \quad \Longleftrightarrow \quad (\lambda \mathbf{1} - T)(\boldsymbol{v}) = \mathbf{0}.$$

- Hence, we find that $W_T(\lambda) = \mathrm{Ker}(\lambda \mathbf{1} - T)$
- It follows that $W_T(\lambda)$ is a subspace of V.

□

- The subspace $W_T(\lambda)$ is analogous to the subspace defined by $(\lambda \boldsymbol{I}_n - \boldsymbol{A}) = \mathbf{0}$.

Eigen Polynomial of a Transformation

Definition 2.9.

- With respect to the linear transformation $T : V \longrightarrow V$, the corresponding matrix of T in relation to the basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is denoted by $\boldsymbol{A}$.
- In other words, $\boldsymbol{A}$ is a complex-valued square matrix such that

$$\big(T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_n)\big) = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)\boldsymbol{A}.$$

- Then, the **eigen polynomial** or **characteristic polynomial** $\Phi_T(\lambda)$ is defined as

$$\Phi_T(\lambda) = \Phi_{\boldsymbol{A}}(\lambda) = \det\big(\lambda \boldsymbol{I}_n - \boldsymbol{A}\big),$$

with $\lambda \in \mathfrak{C}$.

Polynomial Equation

Theorem 2.10.

For a linear transformation T and a scalar λ,

$$\lambda \text{ is an eigenvalue of } T \iff \Phi_T(\lambda) = 0.$$

Proof ($\Longrightarrow$)

- If λ is an eigenvalue of T, there exists $\boldsymbol{v} \in V$ such that $T(\boldsymbol{v}) = \lambda \boldsymbol{v}$, which is equivalent to $\lambda \boldsymbol{v} - T(\boldsymbol{v}) = \mathbf{0}$, which implies that $\Phi_T(\lambda) = \mathbf{0}$ by definition.

Proof ($\Longleftarrow$)

- Let $\boldsymbol{A}$ be the representation matrix of T in relation to the basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ of V.

Polynomial Equation (Cont'd)

- Then, $\boldsymbol{A}$ can be expressed as $\boldsymbol{v} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)\boldsymbol{x}$, where $\boldsymbol{x} \in \mathfrak{C}^n$. With $\lambda \in \mathfrak{C}$ denoting a scalar, we have

$$\begin{aligned} \lambda\boldsymbol{v} - T(\boldsymbol{v}) &= \lambda\left(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\right)\boldsymbol{x} - \left(T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_n)\right)\boldsymbol{x} \\ &= \left(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\right)(\lambda\boldsymbol{x} - \boldsymbol{A}\boldsymbol{x}) \\ &= \left(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\right)(\lambda\boldsymbol{I}_n - \boldsymbol{A})\boldsymbol{x} \end{aligned}$$

- By assumption, $\left(\lambda\boldsymbol{I}_n - \boldsymbol{A}\right)\boldsymbol{x} = \boldsymbol{0}$
- Therefore, we obtain a system of linear equations.
- Since $\boldsymbol{x}$ is arbitrary, for λ to be the eigenvalue, the system of linear equations must have non-trivial solution.
- The necessary and sufficient condition is that $\det\left(\lambda\boldsymbol{I}_n - \boldsymbol{A}\right) = 0$. □

Eigenspace of a Linear Transformation

Theorem 2.11.

For the linear transformation $T : V \longrightarrow V$, let $\boldsymbol{A}$ be the representation matrix of T in relation to a basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ of V. The eigenspace of T is given by

$$W_T(\lambda) = \left\{ \boldsymbol{v} = x_1\boldsymbol{v}_1 + \cdots + x_n\boldsymbol{v}_n \middle| \boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathfrak{C}^n, (\lambda \boldsymbol{I}_n - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}. \right\}$$

Proof.

By construction, we have $T(\boldsymbol{v}) = \lambda \boldsymbol{v}$. It is equivalent to $\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$. In turn, it is equivalent to $(\lambda \boldsymbol{I}_n - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$. We just need to solve for $\boldsymbol{x}$ and use its elements as the coefficients in the linear combination involving the basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ of V. □

Example

Example 2.12.

With respect to the 3-dimensional space,

$$\mathbb{R}[x]_2 = \{a_0 + a_1 x + a_2 x^2 \big| a_0, a_1, a_2 \in \Re\},$$

find the eigenvalues and eigenspaces for the following linear transformation $T : \mathbb{R}[x]_2 \longrightarrow \mathbb{R}[x]_2$:

$$T\big(f(x)\big) = f''(x) - 2xf'(x) - f(x)$$

- The basis of $\mathbb{R}[x]_2$ is $\{1, x, x^2\}$. By definition, $T(1) = -1$, $T(x) = -3x$, and $T(x^2) = 2 - 5x^2$.
- Hence, in terms of columns, the representation matrix is given by

$$\big(T(1), T(x), T(x^2)\big) = (1, x, x^2) \begin{bmatrix} -1 & 0 & 2 \\ 0 & -3 & 0 \\ 0 & 0 & -5 \end{bmatrix} =: (1, x, x^2)\boldsymbol{A}.$$

Example (Cont'd)

- It is easy to see that the eigenvalues are $-1, -3$, and -5.
- To find the eigenvector, we need to solve $(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{v} = \boldsymbol{0}$.
- For $\lambda = -1$, the eigenvector is found to be $a_0 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, where $a_0 \in \Re$.
- For $\lambda = -3$, the eigenvector is found to be $a_1 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, where $a_1 \in \Re$.
- For $\lambda = -5$, the eigenvector is found to be $\widetilde{a}_2 \begin{bmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{bmatrix} = a_2 \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$,

  where $a_2 = \dfrac{1}{2}\widetilde{a} \in \Re$.

Example (Cont'd)

- Consequently, the eigenspaces are

$$W_T(-1) = \{f(x) = a_0 \big| a_0 \in \Re\},$$

$$W_T(-3) = \{f(x) = a_1 x \big| a_1 \in \Re\},$$

$$\begin{aligned} W_T(-5) &= \{f(x) = -a_2 + 2a_2 x^2 \big| a_2 \in \Re\} \\ &= \{f(x) = a_2(2x^2 - 1) \big| a_2 \in \Re\}. \end{aligned}$$

Properties of Eigen Polynomial

- For any n-dimensional square matrix $\boldsymbol{A}$ and an n-dimensional regular matrix $\boldsymbol{P}$,

$$\Phi_{\boldsymbol{A}}(\lambda) = \Phi_{\boldsymbol{P}^{-1}\boldsymbol{AP}}(\lambda).$$

  That is, $\boldsymbol{A}$ and $\boldsymbol{P}^{-1}\boldsymbol{AP}$ have the same eigenvalues.

- $\Phi_{\boldsymbol{A}}(\lambda) = \Phi_{\boldsymbol{A}'}(\lambda)$.

- The eigen polynomial $\Phi_T(\lambda)$ does not depend on the basis of V and it is uniquely determined.

Dimension of Eigenspace

Theorem 3.1.

For a linear transformation $T : V \longrightarrow V$ with eigenvalue λ,

$$\dim W_T(\lambda) = \dim W_{\boldsymbol{A}}(\lambda) = n - \operatorname{rank}(\lambda \boldsymbol{I}_n - \boldsymbol{A}), \tag{1}$$

where $\boldsymbol{A}$ is the matrix in correspondence with the transformation T for any arbitrary basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ of V. Moreover,

$$\dim W_T(\lambda) = \dim W_{\boldsymbol{A}}(\lambda) > 0 \quad \Longleftrightarrow \quad 0 \leqq \operatorname{rank}(\lambda \boldsymbol{I}_n - \boldsymbol{A}) < n.$$

Theorem

Theorem 3.2.

With respect to a transformation T on an n-dimensional linear space V, i.e., $T : V \longrightarrow V$, let $\lambda_1, \ldots \lambda_r$ $(1 \leqq r \leqq n)$ be the eigenvalues that are different from each other. Then

(1) *For each $i = 1, \ldots, r$, let $\boldsymbol{v}_i \in W_T(\lambda_i) \backslash \{\mathbf{0}\}$. Then $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ are linearly independent.*

(2) $\displaystyle\sum_{i=1}^{r} \dim W_T(\lambda_i) \leqq n$.

Diagonalization

- For a square matrix $\boldsymbol{A}$, there exists regular matrix $\boldsymbol{P}$, if $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is a triangular matrix, $\boldsymbol{A}$ is said to be triangularizable. In particular, if $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is a diagonal matrix, then $\boldsymbol{A}$ is said to be **diagonalizable**.

Theorem 4.1.

For any n-dimensional square matrix $\boldsymbol{A}$, let the solutions of $\Phi_{\boldsymbol{A}}(\lambda) = 0$ be $\lambda_1, \ldots, \lambda_n \in \mathfrak{C}$, and let $\boldsymbol{P} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_n \end{bmatrix}$ where $\boldsymbol{x}_i$ is the eignvector corresponding to the eigenvalue λ_i. Then,

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\begin{bmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{bmatrix}$$

Diagonalizability and Eigenvalues

Theorem 4.2.

For any n-dimensional square matrix $\boldsymbol{A}$, denote the mutually different eigenvalues of $\boldsymbol{A}$ by $\lambda_1, \ldots, \lambda_r$ $(1 \leqq r \leqq n)$, and the multiplicity of λ_i by n_i. Then, the following 3 statements are equivalent.

1. $\dim W_{\boldsymbol{A}}(\lambda_i) = n_i \quad (i = 1, 2, \ldots, r)$
2. $\displaystyle\sum_{i=1}^{r} \dim W_{\boldsymbol{A}}(\lambda_i) = n$
3. $\boldsymbol{A}$ *is diagonalizable.*

Theorem 4.3.

If the n eigenvalues of an n-dimensional square matrix $\boldsymbol{A}$ are all different, then $\boldsymbol{A}$ is diagonalizable.

Example

- Let $\boldsymbol{A} = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$.
- $\det\left(\lambda \boldsymbol{I}_2 - \boldsymbol{A}\right) = \begin{vmatrix} \lambda - 4 & 3 \\ -2 & \lambda + 1 \end{vmatrix} = (\lambda - 4)(\lambda + 1) + 6 = (\lambda - 1)(\lambda - 2) = 0$.
- Having found the eigenvalues 1 and 2, we proceed to find the eigenvectors:

$$\begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 1 \begin{bmatrix} x \\ y \end{bmatrix} \quad \Longrightarrow \quad W_A(1) = \left\{ a \begin{bmatrix} 1 \\ 1 \end{bmatrix} \middle| a \in \Re \right\}.$$

$$\begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2 \begin{bmatrix} x \\ y \end{bmatrix} \quad \Longrightarrow \quad W_A(2) = \left\{ b \begin{bmatrix} 3 \\ 2 \end{bmatrix} \middle| b \in \Re \right\}.$$

Example (Cont'd)

- Let $\boldsymbol{P} = \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix}$.
- We compute $\boldsymbol{P}^{-1} = \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}$.
- It follows that the diagonal matrix is obtained as follows:

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Example

- Consider the matrix $\boldsymbol{A} = \begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix}$
- The eigen polynomial is $\Phi_{\boldsymbol{A}}(\lambda) = (\lambda - 3)^2$.
- The eigenvalue is 3. Hence,

$$3\boldsymbol{I}_2 - \boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

- In view of (1) in Theorem 3.1

$$\dim W_{\boldsymbol{A}}(3) = 2 - \operatorname{rank}\left(3\boldsymbol{I}_2 - \boldsymbol{A}\right) = 2 - 1 = 1.$$

- The multiplicity of $\lambda = 3$ is 2. Since $\dim W_{\boldsymbol{A}}(3) \neq 2$, and in view of Theorem 4.2, we conclude that $\boldsymbol{A}$ is not diagonalizable.

Triangularizable

Theorem 4.4.

For any n-dimensional square matrix $\boldsymbol{A}$, denote the eigenvalues of $\boldsymbol{A}$ by $\lambda_1, \ldots, \lambda_n$ $(\in \mathfrak{C})$ Then, there exists a regular matrix $\boldsymbol{P}$ such that

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{bmatrix} \lambda_1 & & \bigstar \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{bmatrix}.$$

That is, any arbitrary complex-valued square matrix $\boldsymbol{A}$ is triangularizable.

Followup of Example in Slide 36

- Earlier, we have shown that $\boldsymbol{A} = \begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix}$ is not diagonalizable.
- It is triangularizable? According to Theorem 4.4, it should be.
- An eigenvector of eigenvalue of 3 is $\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.
- Next, find a vector $\boldsymbol{x}_2 = \begin{bmatrix} p \\ q \end{bmatrix}$ that is linearly independent to $\boldsymbol{x}_1$.

$$\boldsymbol{x}_1' \begin{bmatrix} p \\ q \end{bmatrix} = 0 \qquad \Longrightarrow \qquad p = q.$$

- Hence, we set $p = q = 1$, and the basis of $\Re^2$ is obtained as $\{\boldsymbol{x}_1, \boldsymbol{x}_2\}$.

Followup of Example in Slide 36 (Con'td)

- We calculate

$$\boldsymbol{A}\boldsymbol{x}_2 - 3\boldsymbol{x}_2 = \left(\boldsymbol{A} - 3\boldsymbol{I}_2\right)\boldsymbol{x}_2 = \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} = -2\boldsymbol{x}_1.$$

  Consequently, $\boldsymbol{A}\boldsymbol{x}_2 = -2\boldsymbol{x}_1 + 3\boldsymbol{x}_2$.

- To consolidate, $\boldsymbol{A}\begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} 3\boldsymbol{x}_1 & -2\boldsymbol{x}_1 + 3\boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ 0 & 3 \end{bmatrix}$.

- Accordingly, we let $\boldsymbol{P} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, and

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\begin{bmatrix} 3 & -2 \\ 0 & 3 \end{bmatrix} \quad \Longleftrightarrow \quad \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{bmatrix} 3 & -2 \\ 0 & 3 \end{bmatrix}.$$

Theorems

Theorem 4.5.

For any n-dimensional square matrix $\boldsymbol{A}$, denote the eigenvalues of $\boldsymbol{A}$ by $\lambda_1, \ldots, \lambda_n$ $(\in \mathfrak{C})$ Then, the following two equations are true.

1. $\mathrm{Tr}\boldsymbol{A} = \lambda_1 + \cdots + \lambda_n$.
2. $\det\boldsymbol{A} = \lambda_1 \cdots \lambda_n$.

Theorem 4.6 (The Caley-Hamilton Theorem).

For any arbitrary n-dimensional square matrix $\boldsymbol{A}$, when its eigen polynomial is $\Phi_{\boldsymbol{A}}(\lambda)$, then $\Phi_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{0}_n$.

Theorem 4.7.

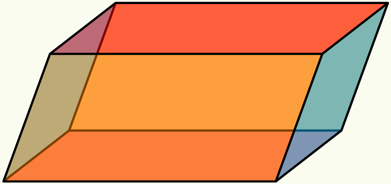
Given a real-valued n-dimensional square matrix $\boldsymbol{A}$, the following two statements are equivalent:

1. *$\boldsymbol{A}$ is symmetric.*
2. *$\boldsymbol{A}$ is diagonalizable.*

Illustration of $\det A_{3\times 3} = \lambda_1 \lambda_2 \lambda_3$

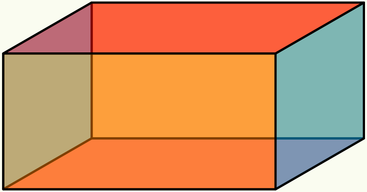
A full rank square matrix:

$$\boldsymbol{A}_{3\times 3} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{bmatrix}$$



Result of diagonalization:

$$\boldsymbol{D} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$



Takeaways

- Eigen analysis is essentially a change of basis of a matrix (or transformation) to one that is orthogonal in the same space.
- An application of eigen analysis is dimensionality reduction in machine learning.
- Eigenvalue may be interpreted as the “length” of a particular dimension in the multi-dimensional space.
- Eigenvalues are found by solving the eigen equations.
- The eigenvector for an eigenvalue is not uniquely determined; a scalar multiple of the eigenvector is also an eigenvector.
- When the n-dimensional matrix is full rank, the multiplicity of every eigenvalue is 1.
- The eigen polynomial is invariant under transpose, matrix transformation, diagonalization and triangularization in particular, by a regular matrix.
- The eigenvectors are linearly independent.

Keywords