Eigenvalues and Eigenvectors

Eigenvalue problem

Let $\boldsymbol{A}$ be an $n \times n$ matrix:

$\boldsymbol{x} \neq \boldsymbol{0}$ is an **eigenvector** of $\boldsymbol{A}$ if there exists a scalar λ such that

$$\boldsymbol{A}\,\boldsymbol{x} = \lambda\,\boldsymbol{x}$$

where λ is called an **eigenvalue**.

If $\boldsymbol{x}$ is an eigenvector, then $\alpha\boldsymbol{x}$ is also an eigenvector. Therefore, we will usually seek for **normalized eigenvectors**, so that

$$\|\boldsymbol{x}\| = 1$$

Note: When using Python, numpy.linalg.eig will normalize using $p=2$ norm.

How do we find eigenvalues?

Linear algebra approach:

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

Therefore the matrix $(\boldsymbol{A} - \lambda\,\boldsymbol{I})$ is singular $\Longrightarrow det(\boldsymbol{A} - \lambda\,\boldsymbol{I}) = 0$

$p(\lambda) = det(\boldsymbol{A} - \lambda\,\boldsymbol{I})$ is the characteristic polynomial of degree n.

In most cases, there is no analytical formula for the eigenvalues of a matrix (Abel proved in 1824 that there can be no formula for the roots of a polynomial of degree 5 or higher) $\Longrightarrow$ **Approximate the eigenvalues numerically**!

Example

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}$$

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

Diagonalizable Matrices

A $n \times n$ matrix $\boldsymbol{A}$ with n linearly independent eigenvectors $\boldsymbol{u}$ is said to be **diagonalizable**.

$$\boldsymbol{A}\,\boldsymbol{u_1} = \lambda_1\,\boldsymbol{u_1},$$
$$\boldsymbol{A}\,\boldsymbol{u_2} = \lambda_2\,\boldsymbol{u_2},$$
$$\dots$$
$$\boldsymbol{A}\,\boldsymbol{u_n} = \lambda_n\,\boldsymbol{u_n},$$

Example $\boldsymbol{A} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}$ $det\begin{pmatrix} 2-\lambda & 1 \\ 4 & 2-\lambda \end{pmatrix} = 0$

Solution of characteristic polynomial gives: $\lambda_1 = 4, \lambda_2 = 0$

To get the eigenvectors, we solve: $\boldsymbol{A}\,\boldsymbol{x} = \lambda\,\boldsymbol{x}$

$$\begin{pmatrix} 2-(4) & 1 \\ 4 & 2-(4) \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \boldsymbol{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 2-(0) & 1 \\ 4 & 2-(0) \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \boldsymbol{x} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Example

The eigenvalues of the matrix:

$$\boldsymbol{A} = \begin{pmatrix} 3 & -18 \\ 2 & -9 \end{pmatrix}$$

are $\lambda_1 = \lambda_2 = -3$.

Select the **incorrect** statement:

A) Matrix $\boldsymbol{A}$ is diagonalizable
B) The matrix $\boldsymbol{A}$ has only one eigenvalue with multiplicity 2
C) Matrix $\boldsymbol{A}$ has only one linearly independent eigenvector
D) Matrix $\boldsymbol{A}$ is not singular

Let's look back at diagonalization…

1) If a $n \times n$ matrix $\boldsymbol{A}$ has n linearly independent eigenvectors $\boldsymbol{x}$ then $\boldsymbol{A}$ is diagonalizable, i.e.,

$$\boldsymbol{A = UDU^{-1}}$$

where the columns of $\boldsymbol{U}$ are the linearly independent normalized eigenvectors $\boldsymbol{x}$ of $\boldsymbol{A}$ (which guarantees that $\boldsymbol{U^{-1}}$ exists) and $\boldsymbol{D}$ is a diagonal matrix with the eigenvalues of $\boldsymbol{A}$.

2) If a $n \times n$ matrix $\boldsymbol{A}$ has less then n linearly independent eigenvectors, the matrix is called defective (and therefore not diagonalizable).

3) If a $n \times n$ **symmetric** matrix $\boldsymbol{A}$ has n distinct eigenvalues then $\boldsymbol{A}$ is diagonalizable.

A $n \times n$ symmetric matrix A with n distinct eigenvalues is diagonalizable.

Suppose $\lambda, \boldsymbol{u}$ and $\mu, \boldsymbol{v}$ are eigenpairs of $\boldsymbol{A}$

$\lambda \boldsymbol{u} = \boldsymbol{A}\boldsymbol{u}$
$\mu \boldsymbol{v} = \boldsymbol{A}\boldsymbol{v}$

Some things to remember about eigenvalues:

- Eigenvalues can have zero value
- Eigenvalues can be negative
- Eigenvalues can be real or complex numbers
- A $n \times n$ real matrix can have complex eigenvalues
- The eigenvalues of a $n \times n$ matrix are not necessarily unique. In fact, we can define the multiplicity of an eigenvalue.
- If a $n \times n$ matrix has n linearly independent eigenvectors, then the matrix is diagonalizable

How can we get eigenvalues numerically?

Assume that $\boldsymbol{A}$ is diagonalizable (i.e., it has n linearly independent eigenvectors $\boldsymbol{u}$). We can propose a vector $\boldsymbol{x}$ which is a linear combination of these eigenvectors:

$$\boldsymbol{x} = \alpha_1 \boldsymbol{u}_1 + \alpha_2 \boldsymbol{u}_2 + \cdots + \alpha_n \boldsymbol{u}_n$$

Power Iteration

Our goal is to find an eigenvector $\boldsymbol{u}_i$ of $\boldsymbol{A}$. We will use an iterative process, where we start with an initial vector, where here we assume that it can be written as a linear combination of the eigenvectors of $\boldsymbol{A}$.

$$\boldsymbol{x}_0 = \alpha_1 \boldsymbol{u}_1 + \alpha_2 \boldsymbol{u}_2 + \cdots + \alpha_n \boldsymbol{u}_n$$

Power Iteration

$$\boldsymbol{x}_k = (\lambda_1)^k \left[\alpha_1 \boldsymbol{u}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k \boldsymbol{u}_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1}\right)^k \boldsymbol{u}_n \right]$$

Assume that $\alpha_1 \neq 0$, the term $\alpha_1 \boldsymbol{u}_1$ dominates the others when k is very large.

Since $|\lambda_1| > |\lambda_2|$, we have $\left(\frac{\lambda_2}{\lambda_1}\right)^k \ll 1$ when k is large

Hence, as k increases, $\boldsymbol{x}_k$ converges to a multiple of the first eigenvector $\boldsymbol{u}_1$, i.e.,

How can we now get the eigenvalues?

If $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$ such that

$$\boldsymbol{A}\,\boldsymbol{x} = \lambda\,\boldsymbol{x}$$

then how can we evaluate the corresponding eigenvalue λ?

Power Iteration

Normalized Power Iteration

$$\boldsymbol{x}_k = (\lambda_1)^k \left[\alpha_1 \boldsymbol{u}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \boldsymbol{u}_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \boldsymbol{u}_n \right]$$

$\boldsymbol{x_0}$ = arbitrary nonzero vector

$$\boldsymbol{x_0} = \frac{\boldsymbol{x_0}}{\|\boldsymbol{x_0}\|}$$

for $k = 1,2, \ldots$

$$\boldsymbol{y}_k = \boldsymbol{A}\, \boldsymbol{x}_{k-1}$$

$$\boldsymbol{x}_k = \frac{\boldsymbol{y}_k}{\|\boldsymbol{y}_k\|}$$

Normalized Power Iteration

$$\boldsymbol{x}_k = (\lambda_1)^k \left[\alpha_1 \boldsymbol{u}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \boldsymbol{u}_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \boldsymbol{u}_n \right]$$

What if the starting vector $\boldsymbol{x_0}$ have no component in the dominant eigenvector $\boldsymbol{u}_1$ ($\alpha_1 = 0$)?

Normalized Power Iteration

$$\boldsymbol{x}_k = (\lambda_1)^k \left[\alpha_1 \boldsymbol{u}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \boldsymbol{u}_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \boldsymbol{u}_n \right]$$

What if the first two largest eigenvalues (in magnitude) are the same, $|\lambda_1| = |\lambda_2|$?

1) λ_1 and λ_2 both positives

Normalized Power Iteration

$$\boldsymbol{x}_k = (\lambda_1)^k \left[\alpha_1 \boldsymbol{u}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \boldsymbol{u}_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \boldsymbol{u}_n \right]$$

What if the first two largest eigenvalues (in magnitude) are the same, $|\lambda_1| = |\lambda_2|$?

2) λ_1 and λ_2 both negative

Normalized Power Iteration

$$\boldsymbol{x}_k = (\lambda_1)^k \left[\alpha_1 \boldsymbol{u}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \boldsymbol{u}_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \boldsymbol{u}_n \right]$$

What if the first two largest eigenvalues (in magnitude) are the same, $|\lambda_1| = |\lambda_2|$?

3) λ_1 and λ_2 opposite signs

Potential pitfalls

1. Starting vector $\boldsymbol{x_0}$ may have no component in the dominant eigenvector $\boldsymbol{u}_1$ ($\alpha_1 = 0$). This is usually unlikely to happen if $\boldsymbol{x_0}$ is chosen randomly, and in practice not a problem because rounding will usually introduce such component.

2. Risk of eventual overflow (or underflow): in practice the approximated eigenvector is normalized at each iteration (Normalized Power Iteration)

3. First two largest eigenvalues (in magnitude) may be the same: $|\lambda_1| = |\lambda_2|$. In this case, power iteration will give a vector that is a linear combination of the corresponding eigenvectors:
   - If signs are the same, the method will converge to correct magnitude of the eigenvalue. If the signs are different, the method will not converge.
   - This is a "real" problem that cannot be discounted in practice.

Error

$$\boldsymbol{x}_k = (\lambda_1)^k \left[\alpha_1 \boldsymbol{u}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \boldsymbol{u}_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \boldsymbol{u}_n \right]$$

Convergence and error

Example

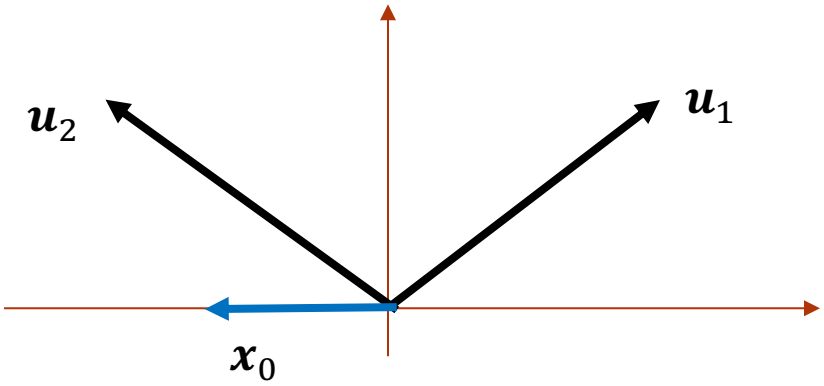
Suppose you are given a matrix with eigenvalues 3, 4, and 5.
You use (normalized) power iteration to approximate one of the eigenvectors $\|\mathbf{x}\|$. For simplicity, assume $\|\mathbf{x}\| = 1$. Your initial guess $\mathbf{x}_0$ has a norm of the error $\|\mathbf{x} - \mathbf{x}_0\| = 0.3$.

How big will the error be after three rounds of normalized power iteration?

(Note that for normalized power iteration, all vectors under consideration have norm 1, so the absolute and the relative error are the same.)

Iclicker question

The matrix $\boldsymbol{A} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ has eigenvalues $(4,2)$ and corresponding eigenvectors $\boldsymbol{u_1} = (1,1)$ and $\boldsymbol{u_2} = (-1,1)$.



Suppose we want to use the normalized power iteration, starting from $x_0 = (-0.5,0)$. Select the correct statement

A) Normalized power iteration will not converge

B) Normalized power iteration will converge to the eigenvector corresponding to the eigenvalue 2.

C) Normalized power iteration will converge to the eigenvector corresponding to the eigenvalue 4.

Suppose $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$ such that

$$\boldsymbol{A}\,\boldsymbol{x} = \lambda\,\boldsymbol{x}$$

What is an eigenvalue of $\boldsymbol{A}^{-1}$?

Inverse Power Method

Previously we learned that we can use the Power Method to obtain the largest eigenvalue and corresponding eigenvector, by using the update

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\,\boldsymbol{x}_k$$

Suppose there is a single smallest eigenvalue of $\boldsymbol{A}$. With the previous ordering

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots > |\lambda_n|$$

Think about this question...

Which code snippet is the best option to compute the smallest eigenvalue of the matrix $\boldsymbol{A}$?

A)
```
x = x0/la.norm(x0)
for k in range(30):
    x = la.solve(A, x)
    x = x/la.norm(x)
```

B)
```
x = x0/la.norm(x0)
for k in range(30):
    x = la.inv(A)@x
    x = x/la.norm(x)
```

C)
```
x = x0/la.norm(x0)
for k in range(30):
    P, L, U = sla.lu(A)
    y = sla.solve_triangular(L, np.dot(P.T, x), lower=True)
    x = sla.solve_triangular(U, y)
    x = x/la.norm(x)
```

D)
```
x = x0/la.norm(x0)
P, L, U = sla.lu(A)
for k in range(30):
    y = sla.solve_triangular(L, np.dot(P.T, x), lower=True)
    x = sla.solve_triangular(U, y)
    x = x/la.norm(x)
```

E) I have no idea!

Inverse Power Method

Cost of computing eigenvalues using inverse power iteration

```
x = x0/la.norm(x0)
for k in range(30):
    x = la.solve(A, x)
    x = x/la.norm(x)
```

```
x = x0/la.norm(x0)
for k in range(30):
    x = la.inv(A)@x
    x = x/la.norm(x)
```

```
x = x0/la.norm(x0)
for k in range(30):
    P, L, U = sla.lu(A)
    y = sla.solve_triangular(L, np.dot(P.T, x), lower=True)
    x = sla.solve_triangular(U, y)
    x = x/la.norm(x)
```

```
x = x0/la.norm(x0)
P, L, U = sla.lu(A)
for k in range(30):
    y = sla.solve_triangular(L, np.dot(P.T, x), lower=True)
    x = sla.solve_triangular(U, y)
    x = x/la.norm(x)
```

Suppose $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$ such that $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{\lambda_1}\,\boldsymbol{x}$ and also $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{B}$ such that $\boldsymbol{B}\,\boldsymbol{x} = \boldsymbol{\lambda_2}\,\boldsymbol{x}$. What is an eigenvalue of

What is an eigenvalue of $\left(\boldsymbol{A} + \boldsymbol{\frac{1}{2}}\boldsymbol{B}\right)^{-1}$?

Suppose $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{A}$ such that $\boldsymbol{A}\,\boldsymbol{x} = \lambda_1\,\boldsymbol{x}$ and also $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{B}$ such that $\boldsymbol{B}\,\boldsymbol{x} = \lambda_2\,\boldsymbol{x}$. What is an eigenvalue of

What is an eigenvalue of $\boldsymbol{A}^2 + \sigma\boldsymbol{B}$?

Eigenvalues of a Shifted Inverse Matrix

Suppose the eigenpairs $(\boldsymbol{x}, \lambda)$ satisfy $\boldsymbol{A}\boldsymbol{x} = \lambda\, \boldsymbol{x}$.

Eigenvalues of a Shifted Inverse Matrix

Convergence summary

	Method	Cost	Convergence $\|\boldsymbol{e_{k+1}}\|/\|\boldsymbol{e_k}\|$
Power Method	$\boldsymbol{x}_{k+1} = \boldsymbol{A}\,\boldsymbol{x}_k$	$k\,n^2$	$\left\|\frac{\lambda_2}{\lambda_1}\right\|$
Inverse Power Method	$\boldsymbol{A}\,\boldsymbol{x}_{k+1} = \boldsymbol{x}_k$	$n^3 + k\,n^2$	$\left\|\frac{\lambda_n}{\lambda_{n-1}}\right\|$
Shifted Inverse Power Method	$(\boldsymbol{A} - \sigma\boldsymbol{I})\boldsymbol{x}_{k+1} = \boldsymbol{x}_k$	$n^3 + k\,n^2$	$\left\|\frac{\lambda_c - \sigma}{\lambda_{c2} - \sigma}\right\|$

λ_1: largest eigenvalue (in magnitude)
λ_2: second largest eigenvalue (in magnitude)
λ_n: smallest eigenvalue (in magnitude)
λ_{n-1}: second smallest eigenvalue (in magnitude)
λ_c: closest eigenvalue to σ
λ_{c2}: second closest eigenvalue to σ