Eigenvalues and Eigenvectors

Power Iteration

$x_0 = \alpha_1 u_1 + \alpha_2 u_2 - - - + \alpha_n u_n$

$x_{k+1} = A x_k \qquad A_{n\times n}$

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

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

$\frac{\lambda_2}{\lambda_1} < 1$

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.,

$|\lambda_1| > |\lambda_2| > |\lambda_3| \ldots > |\lambda_n|$

$k \to \infty \implies \boxed{x_k = (\lambda_1)^k \alpha_1 u_1}$

$u_1, \lambda_1 \qquad \|x_k\| \to \text{large}$

How can we now get the eigenvalues?

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

vector

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

vector

then how can we evaluate the corresponding eigenvalue λ?

$$x \cdot Ax = \lambda\, x \cdot x$$

$$\lambda = \frac{x \cdot Ax}{x \cdot x} \quad \text{or} \quad \boxed{\lambda = \frac{x^T A x}{x^T x}}$$

Rayleigh coefficient

Power Iteration

$x_0 = $ ________________

$x = x_0$

for $i = 1, 2, \ldots$

$$x = Ax \longrightarrow \|x\| \to \text{growing}$$

$$\lambda = \frac{x^T A x}{x^T x}$$

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

$$\boldsymbol{x}_0 = \frac{\boldsymbol{x}_0}{\|\boldsymbol{x}_0\|}$$

$for\ k = 1,2, \ldots$

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

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

x_k is normalized

$$\lambda = \frac{x_k^T A x_k}{x_k^T x_k}$$

Normalized Power Iteration

$*\alpha_1 \neq 0$
$*$ finite calculation

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

$$x_0 = \cancel{\alpha_1 u_1} + \cdots + \alpha_n u_n \qquad \boxed{\alpha_1 = 0 \;?}$$

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

$$x_k = \underline{\alpha_2 \lambda_2^k u_2} + \underbrace{\lambda_1^k \left[\alpha_3 \left(\frac{\lambda_3}{\lambda_1}\right)^k u_3 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1}\right)^k u_n\right]}$$

$k \to \infty$

$x_k \to \alpha_2 \boxed{u_2} \lambda_2^k \to \lambda_2, u_2$

<u>In theory</u> (in infinite precision)

→ power iteration will converge u_2

<u>In practice</u>:

→ p. I. will converge to u_1!

Normalized Power Iteration

$|\lambda_1| \geq |\lambda_2|$

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

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

1) λ_1 and λ_2 both positives

$$x_k \rightarrow \lambda^k \alpha_1 u_1 + \lambda^k \alpha_2 u_2 \qquad \lambda_1 = \lambda_2 = \lambda$$

$$\frac{x_k}{\lambda^k} \rightarrow \alpha_1 u_1 + \alpha_2 u_2 \longrightarrow \text{L. combination of } \underset{\sim}{u_1}, \underset{\sim}{u_2}$$

$$\rightarrow \lambda = \lambda_1 = \lambda_2 = \frac{x^T A x}{x^T x} \quad \checkmark$$

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

$$x_k \rightarrow \pm |\lambda|^k (\alpha_1 u_1 + \alpha_2 u_2)$$

$k \rightarrow$ odd

$k \rightarrow$ even

$x_k \rightarrow$ converge L.C. $\underline{u}_1$, $\underline{u}_2$ + flipped signs

$-x_k$

$+x_k$

$$\boxed{\lambda = \frac{x^T A x}{x^T x}}$$

$$|\lambda| = |\lambda_1| = |\lambda_2|$$

Normalized Power Iteration

$$x_k = (\lambda_1)^k \left[\alpha_1 u_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k u_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1}\right)^k u_n\right]$$

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

3) λ_1 and λ_2 opposite signs

$\lambda_1 > 0$, $\lambda_2 < 0$

$$X_k \longrightarrow \lambda_1^k \alpha_1 u_2 + \lambda_2^k \alpha_2 u_2$$

k is even

$$\lambda^k (\alpha_1 u_1 + \alpha_2 u_2)$$

k is odd

$$\pm \lambda^k (\alpha_1 u_1 - \alpha_2 u_2)$$

$\Rightarrow$ no longer have convergence

no longer have λ

Fails

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: $\boxed{|\lambda_1| = |\lambda_2|}$. 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

approx u_1

exact u_1

$$\boxed{\|e\| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)}$$

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

$$\frac{x_k}{\lambda_1^k \alpha_1} = u_1 + \left(\frac{\alpha_2}{\alpha_1}\right)\left(\frac{\lambda_2}{\lambda_1}\right)^k u_2 + \cdots$$

$$\text{error} = \frac{x_k}{\lambda_1^k \alpha_1} - u_1 = \underbrace{\left(\frac{\alpha_2}{\alpha_1}\right)\left(\frac{\lambda_2}{\lambda_1}\right)^k u_2}_{\text{dominant}} + \underbrace{\left(\frac{\alpha_3}{\alpha_1}\right)\left(\frac{\lambda_3}{\lambda_1}\right)^k u_3 + \cdots}_{\text{neglect}}$$

$k \to \infty$

$$\underset{\sim}{e}_k = \frac{\alpha_2}{\alpha_1}\left(\frac{\lambda_2}{\lambda_1}\right)^k \underset{\sim}{u}_2$$

$$\|\underset{\sim}{e}_k\| = \left|\frac{\alpha_2}{\alpha_1}\right| \left|\frac{\lambda_2}{\lambda_1}\right|^k \|u_2\|$$

Convergence and error

$$\|e_k\| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$$

$$\frac{\|e_{k+1}\|}{\|e_k\|} = \frac{\left|\frac{\lambda_2}{\lambda_1}\right|^{k+1}}{\left|\frac{\lambda_2}{\lambda_1}\right|^{k}} = \left|\frac{\lambda_2}{\lambda_1}\right| = \text{constant}$$

$$\boxed{\begin{aligned}\|e_{k+1}\| &= \left|\frac{\lambda_2}{\lambda_1}\right| \|e_k\| \\ \|e_k\| &= \left|\frac{\lambda_2}{\lambda_1}\right|^k \|e_0\|\end{aligned}} \rightarrow \text{linear convergence}$$

Example

$A = \begin{cases} 3 \longrightarrow \lambda_3 \\ 4 \longrightarrow \lambda_2 \\ 5 \longrightarrow \lambda_1 \Longrightarrow u_1 \end{cases}$

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 $||\mathbf{x} - \mathbf{x}_0|| = 0.3$.

$x_0 = \longrightarrow ||x - x_0|| = 0.3$

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.)

$||e_0|| = ||x - x_0|| = 0.3$

$||e_1|| = ||x - x_1|| = \left|\frac{\lambda_2}{\lambda_1}\right| ||e_0|| = \frac{4}{5}(0.3)$

$||e_2|| = \left|\frac{\lambda_2}{\lambda_1}\right| ||e_1||$

$||e_3|| = \left|\frac{\lambda_2}{\lambda_1}\right| ||e_2||$

$||e_3|| = \left|\frac{\lambda_2}{\lambda_1}\right|^3 ||e_0||$

$= \left(\frac{4}{5}\right)^3 0.3$

$||e_3|| = 0.1536$ //

(Normalized)

Power iteration $\longrightarrow$ $x_{k+1} = \boxed{A} x_k$ $\quad x_0$:

Converges to multiple of eigenvector $\underline{u}_1$

corresponding to $\underline{\underline{\lambda_1}}$ $\rightarrow$ largest eigenvalue in magnitude

Rayleigh coefficient : $\lambda = \dfrac{x^T A x}{x^T x}$

What if I want another eigenvalue ?

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$$

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

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

$\lambda, x \rightarrow A$

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

$$Ax = \lambda x$$

$$\frac{1}{\lambda} A x = x$$

$$\boxed{\frac{1}{\lambda} x = A^{-1} x}$$

$\frac{1}{\lambda}$ is an eigenvalue of A^{-1}

$\lambda \rightarrow$ is the largest

$\frac{1}{\lambda}$? $\longrightarrow$ smallest

Inverse Power Method

smallest eigenvalue

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

Power method

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

Eigenvalues of A^{-1}

$$\left|\frac{1}{\lambda_n}\right| > \left|\frac{1}{\lambda_{n-1}}\right| > \cdots > \left|\frac{1}{\lambda_1}\right|$$

Power method with A^{-1} will converge $\left|\frac{1}{\lambda_n}\right|$

$$\boxed{x_{k+1} = A^{-1} x_k}$$

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

$$x_{k+1} = A^{-1} x_k \longrightarrow \boxed{A x_{k+1} = x_k}$$

solve! $\longrightarrow x_{k+1}$

① Factorize $A = PLU$ la.lu(A)

$$PLU \underbrace{x_{k+1}}_{y} = x_k \longrightarrow O(n^3)$$

② $Ly = P^T x_k$ $\longrightarrow$ solve for y $\longrightarrow O(n^2)$

③ $U x_{k+1} = y$ $\longrightarrow$ solve for x_{k+1} $\longrightarrow O(n^2)$

Cost of computing eigenvalues using inverse power iteration

$A x_{k+1} = x_k$

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

$O(n^3)$

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

$O(n^3)$

```
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)
```

$O(n^3)$

$O(n^2)$

```
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)
```

$O(n^3)$

$O(n^2)$

λ_n

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

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

$A \longrightarrow \lambda_1, x$

$B \longrightarrow \lambda_2, x$

$(A + \frac{1}{2}B)^{-1} \longrightarrow \lambda, x$

$$(A + \tfrac{1}{2}B)^{-1} x = \lambda x$$

$$\frac{1}{\lambda} x = (A + \tfrac{1}{2}B) x$$

$$\frac{1}{\lambda} x = \underbrace{Ax} + \frac{1}{2} Bx$$

$$\frac{1}{\lambda} x = \lambda_1 x + \frac{1}{2}\lambda_2 x$$

$$= (\lambda_1 + \tfrac{1}{2}\lambda_2) x$$

$$\frac{1}{\lambda} = \lambda_1 + \frac{1}{2}\lambda_2$$

$$\boxed{\lambda = \frac{1}{\lambda_1 + \frac{1}{2}\lambda_2}}$$

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 $\boldsymbol{A}^2 + \sigma\boldsymbol{B}$?

$$\lambda_1, x \longrightarrow A$$

$$\lambda_2, x \longrightarrow B$$

$$\lambda, x \longrightarrow A^2 + \sigma B$$

$$(A^2 + \sigma B)x = \lambda x$$

$$A^2 x + \sigma B x = \lambda x$$

$$\lambda_1^2 x + \sigma \lambda_2 x = \lambda x$$

$$\boxed{\lambda = \lambda_1^2 + \sigma \lambda_2}$$

Eigenvalues of a Shifted Inverse Matrix

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

$(x, \lambda) \longrightarrow A$

$(x, \bar{\lambda}) \longrightarrow (A - \sigma I)^{-1}$? What is $\bar{\lambda}$?

$$(A - \sigma I)^{-1} x = \bar{\lambda}\, x$$

$$\frac{1}{\bar{\lambda}} x = (A - \sigma I) x$$

$$= \underbrace{Ax} - \sigma I x$$

$$\frac{1}{\bar{\lambda}} x = \lambda x - \sigma x = (\lambda - \sigma) x$$

$$\boxed{\frac{1}{\bar{\lambda}} = \lambda - \sigma \qquad \bar{\lambda} = \frac{1}{\boxed{\lambda} - \sigma}}$$

eigenvalue of $(A - \sigma I)$

eigenvalue of A

Eigenvalues of a Shifted Inverse Matrix

$$x_{k+1} = \underbrace{(A - \sigma I)^{-1}} x_k$$

$$\boxed{\bar{\lambda} = \frac{1}{\lambda - \sigma}}$$

$$(A - \sigma I) x_{k+1} = x_k \longrightarrow \text{Inverse Power iteration } (A - \sigma I)$$

converge to eigenvector

to largest $\bar{\lambda}$ $\rightarrow$ $\frac{1}{\lambda - \sigma}$ or $\boxed{\text{to smallest } (\lambda - \sigma)}$

$\lambda \to \sigma$

$\rightarrow$ converging to eigenvalue of A which is closer to σ

$$\underline{\underline{(A-\sigma I)}}\, x_{k+1} = x_k$$

- $\underline{\text{define } \sigma}$

- random $\underline{x}_0$ / • normalize $x = x_0 / \|x_0\|$
- $B = (A - \sigma I)$
- $P, L, U = \text{la.lu}(A - \sigma I)$ $\rightarrow O(n^3)$

$$\boxed{\lambda = \frac{x^T A x}{x^T x}}$$

for $i = 1, 2, \ldots$

$$L y = P^T x \longrightarrow \text{solve for } y$$

$$U x_{new} = y \longrightarrow \text{solve for } x_{new}$$

$\Big\}\ O(n^2)$

$$x = x_{new} / \| x_{new} \|$$

x: eigenvector corresponding λ which is the eig. A closer to σ

Convergence summary

get λ_1 $O(n^2)$ linear

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

$\lambda_c \rightarrow$ eigenvalue closer to σ

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