

## Ch 9.3: The power method

Thursday, November 20, 2025

8:09 PM

### The power method

*Use case is fairly specific/limited but it's a part of more sophisticated algorithms, and it's simple, which is why we teach it.*

Simple... but limited. It finds the largest eigenvalue in magnitude (and its associated eigenvector), approximately.  
It's an iterative alg.

Let's take  $A \in \mathbb{C}^{n \times n}$  w/ eigenvalues (sorted)

$$|\lambda_1| > |\lambda_2| > \dots \geq |\lambda_n| \geq 0$$

and assume we have a basis of eigenvectors (ie., diagonalizable)  
 $\{\vec{v}_1, \dots, \vec{v}_n\}$

So any  $\vec{x} \in \mathbb{C}^n$  can be written  $\vec{x} = \sum_{i=1}^n \beta_i \vec{v}_i$

thus

$$A \cdot \vec{x} = A \cdot \left( \sum_i \beta_i \vec{v}_i \right) = \sum_i \beta_i \cdot A \vec{v}_i = \sum_i \beta_i \cdot \lambda_i \vec{v}_i$$

and

$$A^k \vec{x} = \dots = \sum_i \beta_i \cdot \lambda_i^k \vec{v}_i$$

$$\text{i.e. } A^k \vec{x} = \lambda_1^k \cdot \left( \underbrace{\sum_{i=1}^n \beta_i \left( \frac{\lambda_i}{\lambda_1} \right)^k \vec{v}_i}_{\text{For } i=1, \text{ this is } 1} \right)$$

For  $i=1$ , this is 1

For  $i \neq 1$ ,  $\left| \frac{\lambda_i}{\lambda_1} \right|^k < 1$

so

$$A^k \vec{x} \approx \lambda_1^k \cdot \beta_1 \vec{v}_1 \text{ for } k \text{ large}$$

probably goes to 0 or vs or oscillates,

so we normalize to fix that

$$\text{so } \left| \frac{\lambda_i}{\lambda_1} \right|^k \rightarrow 0$$

#### Simple normalization

Initialize at  $\vec{x}_0$  w/ unit norm

for  $k=1, 2, 3, \dots$

$$\vec{y}_k = A \cdot \vec{x}_{k-1}$$

$$\mu_{k-1} = \vec{x}_{k-1}^\top \cdot \vec{y}_k$$

$$\vec{x}_k = \vec{y}_k / \|\vec{y}_k\|_2$$

If  $\vec{x}$  is an eigenvector,

$$A \vec{x} = \lambda \vec{x}$$

$$\text{so } \vec{x}^\top A \vec{x} = \lambda \cdot \vec{x}^\top \vec{x}$$

$$\text{or } \lambda = \frac{\vec{x}^\top A \vec{x}}{\|\vec{x}\|^2}$$

Rayleigh quotient

then  $\mu_k$  converges to  $\lambda_1$  at rate  $O(|\frac{\lambda_2}{\lambda_1}|^k)$

the vector  $\vec{x}_k$  "kind of" converges to the eigenvector  $\vec{v}_i$ :

really,  $\text{dist}(\vec{x}_k, \text{span}(\vec{v}_i)) \rightarrow 0$  (in fact, at rate  $| \frac{\lambda_i}{\lambda_1} |^k$ )

(why this distinction? First,  $\vec{v}_i$  may not be normalized but  $\vec{x}_k$  is. Second,  $\vec{x}_k$  might flip sign,

$$\text{ex: } \lambda_1 = -3, \quad A = \begin{bmatrix} -3 & 0 \\ 0 & 1 \end{bmatrix}, \quad \vec{x}_k = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\vec{y}_k = A\vec{x}_k = \begin{bmatrix} -3 \\ 0 \end{bmatrix}, \quad \vec{x}_{k+1} = \vec{y}_k / \|\vec{y}_k\| = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

sign alternates.

That's OK, both are eigenvalues,

but you can't literally say  $\vec{x}_k$  converges.

To avoid this flip-flopping, the book introduces a ...

### Complicated normalization scheme

Subroutine  $\text{signed-l}^\infty\text{-normalization}(\vec{x})$

Compute  $\|\vec{x}\|_\infty = \max_{i \leq n} |x_i|$

Let  $p$  be the index,  $1 \leq p \leq n$ , such that

$|x_p| = \|\vec{x}\|_\infty$ . If there's more than 1 such index, choose the first index

Return  $\vec{x}/x_p, x_p$  so this has  $\|\cdot\|_\infty = 1$   
but the sign is consistent

### power-iter( $\vec{x}_0$ )

$\vec{x}_0, \mu_0 \leftarrow \text{signed-l}^\infty\text{-norm}(\vec{x}_0)$

for  $k = 1, 2, 3, \dots$

$$\vec{y}_k = A \cdot \vec{x}_k$$

$$\vec{y}_k, \mu_k \leftarrow \text{signed-l}^\infty\text{-norm}(\vec{y}_k)$$

if  $\|\vec{y}_k - \vec{x}_k\| < \text{tolerance}$

break

else

$$\vec{x}_{k+1} \leftarrow \vec{y}_k$$

return  $\vec{y}_k, \mu_k$  // estimates of eigenvector, eigenvalue

The power method is guaranteed to work if

- 1)  $A$  is diagonalizable
  - 2)  $|\lambda_1| > |\lambda_2|$  strict separation
  - 3)  $\beta_i \neq 0$ , i.e.  $\vec{x}_0$  not chosen unlucky
- } can be slightly relaxed  
} due to floating pt. numbers,  
this isn't a big worry

 The book claims the  $\|\cdot\|_2$  normalized version, using  $\mu_k = \frac{\vec{x}_k^T A \vec{x}_k}{\|\vec{x}_k\|^2}$ , is only applicable for symmetric matrices (or Hermitian, if complex). That's not true.

But it is true that:

$A = A^T$ (or $A = A^*$ if complex) $A \neq A^T$ (or $A \neq A^*$ if complex)	$O(\frac{1}{\lambda_1} (\frac{1}{\lambda_2})^{2k})$ $O(\frac{1}{\lambda_1} (\frac{1}{\lambda_2})^k)$ convergence
--	---

So it works better for symmetric/Hermitian matrices

proof sketches of convergence rates:

(1) if  $A = A^T$  then  $\mu_k \rightarrow \lambda_1$  at rate  $O(\frac{1}{\lambda_1} (\frac{1}{\lambda_2})^{2k})$

let's simplify and not give a general proof, but still do enough to convey the idea, so assume  $n=3$  and  $\vec{x} = \vec{v}_1 + \vec{v}_2 + \vec{v}_3$  (i.e.  $\beta_1 = \beta_2 = \beta_3$ , for simplicity)

$$\text{So } A \cdot \vec{x} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \lambda_3 \vec{v}_3$$

Since  $A = A^T$  we can choose  $\{\vec{v}_i\}$  to be orthonormal.

$$A \vec{x}_k = \lambda_1^k \vec{v}_1 + \lambda_2^k \vec{v}_2 + \lambda_3^k \vec{v}_3$$

then

$$\mu_k = \frac{\vec{x}_k^T A \vec{x}_k}{\vec{x}_k^T \vec{x}_k} = \frac{\lambda_1^{2k+1} + \lambda_2^{2k+1} + \lambda_3^{2k+1}}{\lambda_1^{2k} + \lambda_2^{2k} + \lambda_3^{2k}}$$

using orthogonality

To show  $|\mu_k - \lambda_1| = O(\frac{1}{\lambda_1} (\frac{1}{\lambda_2})^{2k})$ , we'll show

$$\frac{|\mu_k - \lambda_1|}{|\frac{1}{\lambda_2} (\frac{1}{\lambda_1})^{2k}|} \leq \underbrace{\text{bound}}_{\text{independent of } k}$$

$$\left| \frac{\lambda_1^{2k+1} + \lambda_2^{2k+1} + \lambda_3^{2k+1}}{\lambda_1^{2k} + \lambda_2^{2k} + \lambda_3^{2k}} - \lambda_1 \right| \cdot \left| \frac{\lambda_1^{2k}}{\lambda_2^{2k}} \right|$$

$$= \left| \frac{\lambda_1^{2k+1} + \lambda_2^{2k+1} + \lambda_3^{2k+1} - \lambda_1 (\lambda_1^{2k} + \lambda_2^{2k} + \lambda_3^{2k})}{\lambda_2^{2k}} \right| \cdot \frac{|\lambda_1^{2k}|}{|\lambda_1^{2k} + \lambda_2^{2k} + \lambda_3^{2k}|}$$

made a common denominator,  
rearranged products

(SUBTLE)

\* it's not automatically true:  
 1) normalization isn't automatic  
 2) orthogonality is possible  
 but only automatic if  
 all eigenvalues are distinct

$$\left| \frac{\lambda_1^{2k+1} + \lambda_2^{2k+1} + \lambda_3^{2k+1} - \lambda_1(\lambda_1^{2k} + \lambda_2^{2k} + \lambda_3^{2k})}{\lambda_2^{2k}} \right| \cdot \frac{|\lambda_1^{2k}|}{|\lambda_1^{2k} + \lambda_2^{2k} + \lambda_3^{2k}|} =$$

$$\frac{|(\lambda_2 - \lambda_1)\lambda_2^{2k} + (\lambda_3 - \lambda_1)\lambda_3^{2k}|}{\lambda_2^{2k}} \cdot \frac{\lambda_1^{2k}}{\lambda_1^{2k} + \lambda_2^{2k} + \lambda_3^{2k}} \quad \text{drop } 1 \cdot 1 \text{ since } \lambda \in \mathbb{R} \text{ so } \lambda^2 > 0$$

$$\leq \left[ |\lambda_2 - \lambda_1| \cdot \underbrace{\frac{\lambda_2^{2k}}{\lambda_2^{2k}}}_{=1} + |\lambda_3 - \lambda_1| \underbrace{\frac{\lambda_3^{2k}}{\lambda_2^{2k}}}_{\leq 1} \right] \cdot \left[ \frac{\lambda_1^{2k}}{\lambda_1^{2k}} \right] \leq |\lambda_2 - \lambda_1| + |\lambda_3 - \lambda_1|$$

which is finite and doesn't grow with  $k$ .  $\square$

(2) If  $A \neq A^T$  ( $\text{or } A \neq A^*$  if complex),

$$|\mu_k - \lambda_1| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right) \text{ but not } O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right)$$

let's show this (via a counterexample)

The difference is that eigenvectors need not be orthogonal:

take  $\vec{x} = \vec{v}_1 + \vec{v}_2$ , say  $\|\vec{v}_1\|_2 = \|\vec{v}_2\|_2$  but  $\vec{v}_1^\top \vec{v}_2 = c \neq 0$

so

$$\vec{x}_k = \lambda_1^k \vec{v}_1 + \lambda_2^k \vec{v}_2, \text{ as before,}$$

$$\begin{aligned} \text{but } \mu_k &= \frac{\vec{x}_k^\top A \vec{x}_k}{\|\vec{x}_k\|_2^2} = \frac{(\lambda_1^k \vec{v}_1 + \lambda_2^k \vec{v}_2)^\top (\lambda_1^{k+1} \vec{v}_1 + \lambda_2^{k+1} \vec{v}_2)}{(\lambda_1^k \vec{v}_1 + \lambda_2^k \vec{v}_2)^\top (\lambda_1^k \vec{v}_1 + \lambda_2^k \vec{v}_2)} \\ &= \frac{\lambda_1^{2k+1} + \lambda_2^{2k+1} + c \lambda_1^k \lambda_2^{k+1} + c \lambda_1^{k+1} \lambda_2^k}{\lambda_1^{2k} + \lambda_2^{2k} + 2c \lambda_1^k \lambda_2^k} \end{aligned}$$

$$\begin{aligned} \frac{|\mu_k - \lambda_1|}{|\lambda_2/\lambda_1|^{2k}} &= \frac{\cancel{\lambda_1^{2k+1}} + \cancel{\lambda_2^{2k+1}} + c \lambda_1^k \lambda_2^{k+1} + c \lambda_1^{k+1} \lambda_2^k - \lambda_1(\lambda_1^{2k} + \lambda_2^{2k} + 2c \lambda_1^k \lambda_2^k)}{\lambda_2^{2k}} \cdot \frac{1}{\lambda_1^{2k}} \\ &= \frac{\lambda_2^{2k+1} + c[\lambda_1^k \lambda_2^{k+1} + \lambda_1^{k+1} \lambda_2^k]}{\lambda_2^{2k}} \cdot \frac{\lambda_1^{2k}}{\lambda_1^{2k} + \lambda_2^{2k} + 2c \lambda_1^k \lambda_2^k} \end{aligned}$$

Problematic

$$\lim_{k \rightarrow \infty} \frac{\lambda_1^{k+1} \lambda_2^k}{\lambda_2^{2k}} = \lim_{k \rightarrow \infty} \lambda_1 \cdot \left(\frac{\lambda_1}{\lambda_2}\right)^k$$

$\Rightarrow$  So raising  $^k$  means it's not bounded.