Math 151BH: Honors Applied Numerical Methods

Lecture 11–12: Finding leading eigenvalues

Instructor: Daniel Mckenzie

Dept. of Mathematics, UCLA

April 23, 2021

Unnormalized power method

- Keep assuming $A \in \mathbb{R}^{n \times n}$ is symmetric.
- Suppose eigenvalues satisfy:

$$1 = \lambda_1 > |\lambda_2| \geq \ldots \geq |\lambda_n| \geq 0 \tag{1}$$

- Corresponding eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$.

Consider the following algorithm:

1. Choose $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$ at random.
2. For $k = 1, \ldots, K$ do:
   2.1 $\boldsymbol{x}^{(k)} = A\boldsymbol{x}^{(k-1)}$.

Claim that $\boldsymbol{x}^{(k)} \to \boldsymbol{v}_1$.

Unnormalized power method

- The eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ are an (orthonormal) basis for $\mathbb{R}^n$.
- Write $\boldsymbol{x}^{(0)} = \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \ldots + \alpha_n \boldsymbol{v}_n$.
- Can compute:

$$\begin{aligned}
\boldsymbol{x}^{(1)} &= A\boldsymbol{x}^{(0)} = A\left(\alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \ldots + \alpha_n \boldsymbol{v}_n\right) && (2)\\
&= \alpha_1 A\boldsymbol{v}_1 + \alpha_2 A\boldsymbol{v}_2 + \ldots + \alpha_n A\boldsymbol{v}_n && (3)\\
&= \alpha_1 \lambda_1 \boldsymbol{v}_1 + \alpha_2 \lambda_2 \boldsymbol{v}_2 + \ldots + \alpha_n \lambda_n \boldsymbol{v}_n && (4)\\
\text{Similarly: } \boldsymbol{x}^{(k)} &= \alpha_1 \lambda_1^k \boldsymbol{v}_1 + \alpha_2 \lambda_2^k \boldsymbol{v}_2 + \ldots + \alpha_n \lambda_n^k \boldsymbol{v}_n && (5)\\
& && (6)
\end{aligned}$$

- By assumption $\lambda_1 = 1$ so $\lambda_1^k = 1$ for all k.
- By assumption $|\lambda_i| < 1$ for $i \geq 2$ so $\lambda_i^k \to 0$.
- **Conclusion:**

$$\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \lim_{k\to\infty} \left(\alpha_1 \lambda_1^k \boldsymbol{v}_1 + \alpha_2 \lambda_2^k \boldsymbol{v}_2 + \ldots + \alpha_n \lambda_n^k \boldsymbol{v}_n\right) = \alpha_1 \boldsymbol{v}_1 \qquad (7)$$

Towards the normalized power method

Suppose that eigenvalues do not satisfy:

$$1 = \lambda_1 > |\lambda_2| \geq \ldots \geq |\lambda_n| \geq 0. \tag{8}$$

Several things can go wrong:

1. If $\lambda_1 > 1$ then $\lim_{k \to \infty} \lambda_1^k = \infty$ so algorithm fails to converge.
2. If $\lambda_1 < 0$ then $\lambda_1^k = \begin{cases} \text{positive} & \text{if } k \text{ is even} \\ \text{negative} & \text{if } k \text{ is odd} \end{cases}$ so $\{\boldsymbol{x}^{(k)}\}_k$ will not converge.
3. What happens if A is not diagonalizable?

The Rayleigh Quotient

Assume that $A \in \mathbb{R}^{n \times n}$ is symmetric.

$$\textbf{The Rayleigh Quotient: } \lambda^R(\boldsymbol{x}) := \frac{\boldsymbol{x}^\top A \boldsymbol{x}}{\boldsymbol{x}^\top \boldsymbol{x}} \text{ for } \boldsymbol{x} \neq 0 \tag{9}$$

Some properties:

- If $(\lambda, \boldsymbol{x})$ is an eigenpair then $\lambda^R(\boldsymbol{x}) = \lambda$.
- $\min_{\boldsymbol{x}:\ \|\boldsymbol{x}\|_2=1} \lambda^R(\boldsymbol{x}) = \lambda_{\min}(A)$ and $\max_{\boldsymbol{x}:\ \|\boldsymbol{x}\|_2=1} \lambda^R(\boldsymbol{x}) = \lambda_{\max}(A)$.
- Rayleigh quotient is even: $\lambda^R(-\boldsymbol{x}) = \lambda^R(\boldsymbol{x})$
- $\lambda_R(\cdot)$ is continuous (away from $\boldsymbol{x} = 0$).

The normalized power method

1. Choose $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$ at random with $\|\boldsymbol{x}^{(0)}\|_2 = 1$.
2. For $k = 1, \ldots, K$ do:
   2.1 $\boldsymbol{y}^{(k)} = A\boldsymbol{x}^{(k-1)}$.
   2.2 $\boldsymbol{x}^{(k)} = \dfrac{1}{\|\boldsymbol{y}^{(k)}\|_2}\boldsymbol{y}^{(k)}$.
   2.3 $\lambda^{(k)} = \lambda^R(\boldsymbol{x}^{(k)})$.

Some remarks:

- Can use either $\boldsymbol{x}^{(k)}$ or $\boldsymbol{y}^{(k)}$ in the Rayleigh quotient.
- if $\boldsymbol{y}^{(k)} = 0$ terminate. Have found an eigenpair.

To show convergence:

1. First show **if** power method converges **then** limit point is an eigenpair.
2. Then show that[1] power method does converge and prove rate.

[1] under some assumptions on the eigenvalues

Limit points are eigenpairs

For any $\alpha \in \mathbb{R}$ define:

$$\text{sign}(\alpha) = \begin{cases} +1 & \text{if } \alpha > 0 \\ -1 & \text{if } \alpha < 0 \\ 0 & \text{if } \alpha = 0 \end{cases} \tag{10}$$

Lemma

Suppose power method converges:

$$\lim_{k\to\infty} \lambda^{(k)} = \lambda \neq 0 \textit{ and } \lim_{k\to\infty} (\textit{sign}(\lambda))^k \boldsymbol{x}^{(k)} \to \boldsymbol{x} \tag{11}$$

Then $A\boldsymbol{x} = \lambda \boldsymbol{x}$.

Proof.

$$\begin{aligned} \lambda = \lim_{k\to\infty} \lambda^{(k)} &= \lim_{k\to\infty} \lambda^R(\boldsymbol{x}^{(k)}) && \textit{by assumption} \\ &= \lim_{k\to\infty} \lambda^R \left(\text{sign}(\lambda)^k \boldsymbol{x}^{(k)}\right) && \textit{Rayleigh quotient is even} \\ &= \lambda^R \left(\lim_{k\to\infty} \text{sign}(\lambda)^k \boldsymbol{x}^{(k)}\right) && \textit{Rayleigh quotient is continuous} \\ &= \lambda^R(\boldsymbol{x}) && \textit{by assumption.} \end{aligned}$$

□

Limit points are eigenpairs

Proof continued.

- Have established:

$$\lambda = \lim_{k\to\infty} \lambda^{(k)} = \lambda^R \left(\lim_{k\to\infty} \text{sign}(\lambda)^k \boldsymbol{x}^{(k)} \right) = \lambda^R(\boldsymbol{x}) \tag{12}$$

- Will now show: $A\boldsymbol{x} = \mu\boldsymbol{x}$ for some μ.
- Conclude that $\mu = \lambda^R(\boldsymbol{x}) = \lambda$ and so $(\boldsymbol{x}, \lambda)$ is an eigenpair.
- *(HW)* Show that $\lim_{k\to\infty} \text{sign}(\lambda)^k \boldsymbol{y}^{(k+1)} = A\boldsymbol{x}$.
- *(HW)* Show that $\lim_{k\to\infty} \|\boldsymbol{y}^{(k)}\|_2 = \|A\boldsymbol{x}\|_2$.
- Conclude that:

$$A\boldsymbol{x} = \lim_{k\to\infty} \text{sign}(\lambda)^k \boldsymbol{y}^{(k+1)} \tag{13}$$

$$= \lim_{k\to\infty} \text{sign}(\lambda)^k \left[\|\boldsymbol{y}^{(k+1)}\|_2 \boldsymbol{x}^{(k+1)} \right] \quad \textit{def. of } \boldsymbol{y}^{(k+1)} \tag{14}$$

$$= \text{sign}(\lambda) \lim_{k\to\infty} \left[\|\boldsymbol{y}^{k+1}\|_2 \right] \lim_{k\to\infty} \left[\text{sign}(\lambda)^{k+1} \boldsymbol{x}^{(k+1)} \right] \quad \textit{limit of products} \tag{15}$$

$$= \left[\text{sign}(\lambda) \|A\boldsymbol{x}\|_2 \right] \boldsymbol{x} := \alpha\boldsymbol{x} \tag{16}$$

□

Limit points are eigenpairs

- Thus power method is reasonable to apply.
- In practice, don't have $\text{sign}(\lambda)$. If $\lambda < 0$ observe limit cycle:

$$\lim_{k \to \infty} \boldsymbol{x}^{(2k)} = +\boldsymbol{x} \text{ and } \lim_{k \to \infty} \boldsymbol{x}^{(2k+1)} = -\boldsymbol{x} \tag{17}$$

Power method converges

- Assume that A is symmetric and that:

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n| \geq 0 \tag{18}$$

- Let $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ be corresponding normalized eigenvectors and $Q = \begin{bmatrix} \boldsymbol{v}_1 & \ldots & \boldsymbol{v}_n \end{bmatrix}$.
- Write $A = Q\Lambda Q^\top$ with $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.
- Observe that:

$$A^k = \left(Q\Lambda Q^\top\right)^k = \underbrace{\left(Q\Lambda Q^\top\right)\left(Q\Lambda Q^\top\right)\ldots\left(Q\Lambda Q^\top\right)}_{k \text{ times}} \tag{19}$$

$$= Q\Lambda^k Q^\top \quad \text{where } \Lambda^k = \operatorname{diag}\left(\lambda_1^k, \ldots, \lambda_n^k\right) \tag{20}$$

Power method converges

- Recall that $\|\boldsymbol{x}^{(0)}\|_2 = 1$.
- Define auxiliary sequence as:
  1. $\boldsymbol{X}^{(0)} = \boldsymbol{x}^{(0)}$.
  2. $\boldsymbol{X}^{(k)} = A\boldsymbol{X}^{(k-1)}$.

  Note that:
  1. $\boldsymbol{X}^{(k)}$ will diverge if $|\lambda_1| > 1$!
  2. $\boldsymbol{X}^{(k)} = A^k \boldsymbol{x}^{(0)}$.
- *(HW)* Show that $\boldsymbol{x}^{(k)} = \|\boldsymbol{X}^{(k)}\|_2^{-1} \boldsymbol{X}^{(k)}$.

Power method converges

- Although $\boldsymbol{X}^{(k)}$ diverges $\lambda_1^{-k}\boldsymbol{X}^{(k)}$ will converge:

$$\lim_{k\to\infty} \lambda_1^{-k}\boldsymbol{X}^{(k)} = \lim_{k\to\infty} \lambda_1^{-k} A^k \boldsymbol{x}^{(0)} \tag{21}$$

$$= \lim_{k\to\infty} \lambda_1^{-k} \left(Q\Lambda Q^\top\right)^k \boldsymbol{x}^{(0)} = \lim_{k\to\infty} Q\lambda_1^{-k}\Lambda^k Q^\top \boldsymbol{x}^{(0)} \tag{22}$$

$$= \lim_{k\to\infty} Q\tilde{\Lambda}^k Q^\top \boldsymbol{x}^{(0)} \ \textit{where}\ \tilde{\Lambda} = \operatorname{diag}\left(1, \lambda_2/\lambda_1, \ldots, \lambda_n/\lambda_1\right) \tag{23}$$

$$= Q\left[\lim_{k\to\infty} \tilde{\Lambda}^k\right] Q^\top \boldsymbol{x}^{(0)} \tag{24}$$

- Note that:

$$\lim_{k\to\infty} \tilde{\Lambda}^k = \lim_{k\to\infty} \operatorname{diag}\left(1, (\lambda_2/\lambda_1)^k, \ldots, (\lambda_n/\lambda_1)^k\right) \tag{25}$$

$$= \operatorname{diag}\left(1, \lim_{k\to\infty} (\lambda_2/\lambda_1)^k, \ldots, \lim_{k\to\infty} (\lambda_n/\lambda_1)^k\right) \tag{26}$$

$$= \operatorname{diag}\left(1, 0, \ldots, 0\right) \quad \textit{as}\ |\lambda_1| > |\lambda_i|\ \textit{for}\ i \neq 1 \tag{27}$$

$$=: E_{11} \tag{28}$$

- Check that: $QE_{11}Q^\top = \boldsymbol{v}_1\boldsymbol{v}_1^\top$.

Power method converges

- So we observe that:

$$\lim_{k\to\infty} \lambda_1^{-k} \boldsymbol{X}^{(k)} = \left[\boldsymbol{v}_1 \boldsymbol{v}_1^\top\right] \boldsymbol{x}^{(0)} = \left[\boldsymbol{v}_1^\top \boldsymbol{x}^{(0)}\right] \boldsymbol{v}_1 \tag{29}$$

- We also have a convergence of norms:

$$\begin{aligned}
\lim_{k\to\infty} |\lambda_1|^{-k} \|\boldsymbol{X}^{(k)}\|_2 &= \lim_{k\to\infty} |\lambda_1^{-k}| \|\boldsymbol{X}^{(k)}\|_2 && (30)\\
&= \lim_{k\to\infty} \|\lambda_1^{-k} \boldsymbol{X}^{(k)}\|_2 && (31)\\
&= \|\lim_{k\to\infty} \lambda_1^{-k} \boldsymbol{X}^{(k)}\|_2 = \left\| \left[\boldsymbol{v}_1^\top \boldsymbol{x}^{(0)}\right] \boldsymbol{v}_1 \right\|_2 && (32)\\
&= \left|\boldsymbol{v}_1^\top \boldsymbol{x}^{(0)}\right| \quad \textit{as } \|\boldsymbol{v}_1\|_2 = 1 && (33)
\end{aligned}$$

Power method converges

- The final trick:

$$
\begin{aligned}
\lim_{k\to\infty} \operatorname{sign}(\lambda_1)^k \boldsymbol{x}^{(k)} &= \lim_{k\to\infty} \frac{\operatorname{sign}(\lambda_1)^k \boldsymbol{X}^{(k)}}{\|\boldsymbol{X}^{(k)}\|_2} \\
&= \lim_{k\to\infty} \frac{\operatorname{sign}(\lambda_1)^k |\lambda_1|^{-k} \boldsymbol{X}^{(k)}}{|\lambda_1|^{-k}\|\boldsymbol{X}^{(k)}\|_2} \quad \textit{multiplying top and bottom} \\
&= \lim_{k\to\infty} \frac{\lambda_1^{-k} \boldsymbol{X}^{(k)}}{|\lambda_1|^{-k}\|\boldsymbol{X}^{(k)}\|_2} \quad \textit{as } \operatorname{sign}(\alpha)|\alpha| = \alpha \\
&= \frac{\lim_{k\to\infty} \lambda_1^{-k} \boldsymbol{X}^{(k)}}{\lim_{k\to\infty} |\lambda_1|^{-k}\|\boldsymbol{X}^{(k)}\|_2} \quad \textit{limit of quotient} \\
&= \frac{\left[\boldsymbol{v}_1^\top \boldsymbol{x}^{(0)}\right] \boldsymbol{v}_1}{|\boldsymbol{v}_1^\top \boldsymbol{x}^{(0)}|} \quad \textit{By earlier work} \\
&= \operatorname{sign}\left(\boldsymbol{v}_1^\top \boldsymbol{x}^{(0)}\right) \boldsymbol{v}_1
\end{aligned}
$$

Power method converges: conclusion

Theorem

Suppose that:

1. *A is symmetric.*
2. $|\lambda_1| > |\lambda_2|$.
3. $\boldsymbol{v}_1^\top \boldsymbol{x}^{(0)} \neq 0$.

Let $\boldsymbol{x}^{(k)}, \lambda^{(k)}$ be the sequence generated by the power method. Then:

1. $\lim_{k\to\infty} \lambda^{(k)} = \lambda_1$.
2. $\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \operatorname{sign}\left(\boldsymbol{v}_1^\top \boldsymbol{x}^{(0)}\right) \boldsymbol{v}_1$.
3. $\left\|\operatorname{sign}\left(\boldsymbol{v}_1^\top \boldsymbol{x}^{(0)}\right) \boldsymbol{v}_1 - \boldsymbol{x}^{(k)}\right\|_2 \leq C\left(|\lambda_2|/|\lambda_1|\right)^k$

- *HW:* Show the convergence to λ_1.
- For rate of convergence, see textbook.
- Can remove assumption that A is symmetric; see texbook.