Eigenvalues and Eigenvectors

CS/SE 4X03

Ned Nedialkov

McMaster University

November 26, 2021

Outline

Introduction

- Given an $n \times n$ matrix A, a nonzero vector v is an eigenvector of A if

$$Av = \lambda v, \qquad \lambda \text{ is scalar} \tag{1}$$

  That is, v does not change direction under the transformation Av

- We can write (1) as

$$Av = \lambda v \Leftrightarrow Av = \lambda I v \Leftrightarrow (A - \lambda I)v = 0$$

  I is the $n \times n$ identity matrix

- $(A - \lambda I)v = 0$ has a nonzero solution v when $\det(A - \lambda I) = 0$
- $\det(A - \lambda I)$ is *characteristic polynomial*, $\det(A - \lambda I) = 0$ is *characteristic equation*

Introduction cont.

Example 1. Consider

$$A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$$

Then

$$A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 3 & -4-\lambda \end{bmatrix}$$

$$\begin{aligned} \det(A - \lambda I) &= (-4-\lambda)(1-\lambda) + 6 = -4 + 4\lambda - \lambda + \lambda^2 - 6 \\ &= \lambda^2 + 3\lambda - 10 = 0 \end{aligned}$$

has roots $\lambda_1 = 2$ and $\lambda_2 = -5$.

Introduction cont.

Example 1. cont.

For $\lambda_1 = 2$,

$$(A - 2I)v = \begin{bmatrix} 1-2 & 2 \\ 3 & -4-2 \end{bmatrix} v = \begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix} v = 0$$

has solution $v_1 = [2, 1]^T$, and

$$Av_1 = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

For $\lambda_1 = -5$,

$$(A + 5I)v = \begin{bmatrix} 1+5 & 2 \\ 3 & -4+5 \end{bmatrix} v = \begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix} v = 0$$

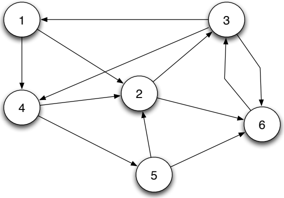
has solution $v_2 = [1, -3]^T$, and

$$Av_1 = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \end{bmatrix} = \begin{bmatrix} -5 \\ 15 \end{bmatrix} = -5 \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

Introduction cont.

Example 2. Example form p. 223–225 of U. Ascher and C. Greig, A First Course in Numerical Methods

- Graph with n nodes, $1, 2, \ldots, n$, each node corresponds to a webpage
- If node i links to node j, directed edge from i to j



- Assume j points to N_j pages

  E.g. page $j = 3$ points to $N_3 = 3$ pages $1, 4, 6$

Introduction cont.

Example 2. cont.

- Denote the importance of node (page) i by x_i
- j contributes $\frac{1}{N_j}x_j$ to the importance of each page it points to

  E.g. page $j = 3$ contributes $\frac{1}{3}x_3$ to each of $1, 4$, and 6

  page $j = 1$ contributes $\frac{1}{2}x_1$ to each of 2 and 4
- Can be represented as a matrix

	1	2	3	4	5	6
1			$\frac{1}{3}$			
2	$\frac{1}{2}$			$\frac{1}{2}$	$\frac{1}{2}$	
3		$\frac{1}{2}$				1
4	$\frac{1}{2}$		$\frac{1}{3}$			
5				$\frac{1}{2}$		
6		$\frac{1}{2}$	$\frac{1}{3}$		$\frac{1}{2}$	

a blank denotes 0

- Column: out links, row: in links

Introduction cont.

Example 2. cont.

- Then

$$x_i = \sum_{j:j \to i} \frac{1}{N_j} x_j, \quad i = 1, \ldots, n$$

  E.g.

$$x_4 = \frac{1}{N_1} x_1 + \frac{1}{N_3} x_3 = \frac{1}{2} x_1 + \frac{1}{3} x_3$$

- We have n equations. As a system:

$$Ax = x$$

- Eigenvalue problem!

Introduction cont.

Example 2. cont.

- Given page i, the number of links to it is $<< n$, total number of pages. n can be in the billions
- A is very large and sparse
- The sum in each column is 1
- This matrix is *column stochastic*
- There is a unique largest eigenvalue 1
- Entries in the corresponding eigenvector are positive
- How to find this eigenvector?

Power method

- Method for finding the largest eigenvalue and corresponding eigenvector
- Denote an eigen pair by (λ_i, x_i)
- Assume λ_1 real and $|\lambda_1| > |\lambda_i|$ for all $i = 2, \ldots, n$
- Assume A has n linearly independent eigenvectors
- Any $v \in \mathbb{R}^n$ can be written as

$$v = \sum_{j=1}^{n} \alpha_j x_j, \quad \alpha_j \text{ scalar}$$

- Compute Av, A^2v, …, A^kv

Power method cont.

We have

$$
\begin{aligned}
Av &= A\sum_{j=1}^{n}\alpha_j x_j = \sum_{j=1}^{n}\alpha_j (Ax_j) = \sum_{j=1}^{n}\alpha_j\lambda_j x_j \\
A^2v &= A(Av) = A\sum_{j=1}^{n}\alpha_j\lambda_j x_j = \sum_{j=1}^{n}\alpha_j\lambda_j (Ax_j) = \sum_{j=1}^{n}\alpha_j\lambda_j^2 x_j \\
&\vdots \\
A^kv &= A(A^{k-1}v) = A\sum_{j=1}^{n}\alpha_j\lambda_j^{k-1} x_j = \sum_{j=1}^{n}\alpha_j\lambda_j^{k-1} (Ax_j) \\
&= \sum_{j=1}^{n}\alpha_j\lambda_j^{k} x_j
\end{aligned}
$$

Power method cont.

$$
\begin{aligned}
A^k v &= \lambda_1^k \alpha_1 x_1 + \lambda_2^k \alpha_2 x_2 + \cdots + \lambda_n^k \alpha_n x_n \\
&= \lambda_1^k \left(\alpha_1 x_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^k \alpha_2 x_2 + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^k \alpha_n x_n \right)
\end{aligned}
$$

- Since $|\lambda_1| > |\lambda_j|$, $\left(\frac{\lambda_j}{\lambda_1}\right)^k \to 0$ as $k \to \infty$ for all $j \geq 2$
- Then
$$A^k v \to (\lambda_1^k \alpha_1) x_1$$
- $A^k v$ converges to a multiple of x_1, the eigenvector corresponding to λ_1

Power method cont.

- Rate of convergence depends on $|\lambda_2|/|\lambda_1|$
- If $|\lambda_1| > 1$, $A^k v \approx (\lambda_1^k \alpha_1) x_1$ can overflow
- If $|\lambda_1| < 1$, $A^k v \approx (\lambda_1^k \alpha_1) x_1$ can underflow
- How to avoid over/underflow? Normalize
- Compute
  - start with any v
  - for $k = 1, 2, \ldots,$ until convergence
    - $\widetilde{v} = Av$
    - $v = \widetilde{v}/\|\widetilde{v}\|$

Rayleigh quotient

- Rayleigh quotient

$$\mu(v) = \frac{v^T A v}{v^T v}$$

- If v is an eigenvector

$$\mu(v) = \frac{v^T A v}{v^T v} = \frac{v^T \lambda v}{v^T v} = \lambda$$

- If v is an approximation to an eigenvector $\mu(v) \approx \lambda$

Power method
v_0 initial guess
for $k = 1, 2, \ldots$ until termination
  $\widetilde{v} = A v_{k-1}$
  $v_k = \widetilde{v} / \|\widetilde{v}\|$
  $\lambda_1^{(k)} = v_k^T A v_k$

Example 2. cont.

- Start with $v = (1/6, 1/6, 1/6, 1/6, 1/6, 1/6)^T$, $\|v\|_1 = 1$
- $\widetilde{v} = Av$, $\|\widetilde{v}\|_1 = 1$, need to normalize
- The x for the PageRank example is

$$x \approx \begin{bmatrix} 0.0994 \\ 0.1615 \\ 0.2981 \\ 0.1491 \\ 0.0745 \\ 0.2174 \end{bmatrix}$$

- Ranking is

page	1	2	3	4	5	6
rank	5	3	1	4	6	2