35 Power method

By the end of this section, you should be able to answer the following questions:

- What is the power method and what does it do?
- Under what conditions can it fail?
- What is deflation, and how does it work in conjunction with the power method?

In applications we sometimes need to find eigenvalues and eigenvectors of a large square matrix. In these cases it is usually impractical, or more to the point not computationally feasible, to find the roots of the characteristic polynomial. Instead, we are forced to rely on computational techniques which estimate eigenvalues and eigenvectors. The power method is one such technique which estimates the largest eigenvalue (provided it is unique) and its corresponding eigenvector.

35.1 Dominant eigenvalue

Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ such that

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

The eigenvalue λ_1 of the matrix A is called the *dominant eigenvalue* of A. The eigenvector $\boldsymbol{v}_1$ corresponding to λ_1 is called the *dominant eigenvector.*

35.1.1 Example

Identify the dominant eigenvalue and eigenvector of the matrix $\begin{pmatrix} -3 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$.

eigenvalues $-1, -3, -4$.

$\Rightarrow -4$ is dominant

Since $|-4| > |-3|, |-1|$

& $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ is a dominant eigenvector.

35.2 The algorithm (power method)

Form a sequence of vectors $\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \ldots$ where $\boldsymbol{u}_0$ is an (almost!) arbitrarily chosen vector, $\boldsymbol{u}_{k+1} = A\boldsymbol{u}_k$ (for $k \geq 0$). Then (usually) for k large, → Conditions on $\underline{u}_0$

(i) The dominant eigenvalue is $\lambda_1 \approx \dfrac{(\boldsymbol{u}_{k+1})_j}{(\boldsymbol{u}_k)_j}$, any $j \leq n$ with $(\boldsymbol{u}_k)_j \neq 0$ (usually we choose j so that $|(\boldsymbol{u}_{k+1})_j|$ is the largest possible),

(ii) $\boldsymbol{u}_k \approx$ dominant eigenvector. (i.e. last vector in sequence)

35.2.1 Example

For $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$, find the exact value of the dominant eigenvalue and eigenvector, then apply the power method approximation.

char. poly. $\det(A - \lambda I) = (3-\lambda)(3-\lambda) - 1$

$= \lambda^2 - 6\lambda + 8$

$= (\lambda - 4)(\lambda - 2)$

$\lambda = 4$ is dominant, corr. eigenvector $= \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$\underline{u}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$\underline{u}_1 = A\underline{u}_0 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$

$\underline{u}_2 = A\underline{u}_1 = \begin{pmatrix} 10 \\ 6 \end{pmatrix} \approx 6\begin{pmatrix} 1.667 \\ 1 \end{pmatrix}$

$\underline{u}_3 = A\underline{u}_2 = \begin{pmatrix} 36 \\ 28 \end{pmatrix} \simeq 28\begin{pmatrix} 1.286 \\ 1 \end{pmatrix}$

$\underline{u}_4 = \begin{pmatrix} 136 \\ 120 \end{pmatrix} \simeq 120\begin{pmatrix} 1.133 \\ 1 \end{pmatrix}$

$\underline{u}_5 = \begin{pmatrix} 528 \\ 496 \end{pmatrix} \simeq 496\begin{pmatrix} 1.065 \\ 1 \end{pmatrix}$

$\underline{u}_6 = \begin{pmatrix} 2080 \\ 2016 \end{pmatrix} \simeq 2016\begin{pmatrix} 1.033 \\ 1 \end{pmatrix}$

$\underline{u}_7 = \begin{pmatrix} 8256 \\ 8128 \end{pmatrix} \simeq 8128\begin{pmatrix} 1.016 \\ 1 \end{pmatrix}$ ...

$\Rightarrow$ dominant eigenvector $\approx \underset{\sim}{u}_7 = \begin{pmatrix} 8256 \\ 8128 \end{pmatrix}$

or just $\begin{pmatrix} 1.016 \\ 1 \end{pmatrix}$

$\Rightarrow A\underline{u}_7 \simeq \lambda \underset{\sim}{u}_7$

$\Rightarrow \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 8256 \\ 8128 \end{pmatrix} = \begin{pmatrix} 32896 \\ 32640 \end{pmatrix} = \underline{u}_8$

Take largest component of $\underline{u}_8$ & Corresponding entry in $\underline{u}_7$

$\Rightarrow$ dominant eigenvalue $\simeq \dfrac{32896}{8256} \simeq 3.984$.

$\underset{\sim}{u}_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$?

$\underset{\sim}{u}_1 = A\underline{u}_0 = \begin{pmatrix} 2 \\ -2 \end{pmatrix} = 2\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

35.2.2 Assumptions

The power method depends on several assumptions: (e.g. $\lambda = 1, -1$ & all other $|\lambda| < 1$)

1. There is a dominant eigenvalue.
2. The eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ are linearly independent and hence form a basis for $\mathbb{R}^n$.
3. The chosen vector $\boldsymbol{u}_0$ that starts the iteration is non-zero and when written as a linear combination of the basis of eigenvectors, has a non-zero component of the dominant eigenvector. → difficult to enforce. Failure is rare in applications. More common in artificial classroom examples.

35.2.3 Understanding the power method

Suppose λ_1 is the dominant eigenvalue of an $n \times n$ matrix A, so that

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

and hence $\lambda_1 \neq 0$. For simplicity, suppose that A has n linearly independent eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathbb{R}^n$. With n linearly independent vectors in $\mathbb{R}^n$, we have a basis, so any vector $\boldsymbol{u} \in \mathbb{R}^n$ can be written as a linear combination of the vectors in the basis. In particular, set (assumption 2.)

v_1 is dominant.

$$\boldsymbol{u}_0 = t_1\boldsymbol{v}_1 + t_2\boldsymbol{v}_2 + \cdots + t_n\boldsymbol{v}_n$$

for some scalars $t_1, \ldots, t_n$. $\boxed{\text{Suppose } t_1 \neq 0}$ (this turns out to be crucial). Then

$$\begin{aligned}\boldsymbol{u}_1 &= A\boldsymbol{u}_0 = t_1A\boldsymbol{v}_1 + t_2A\boldsymbol{v}_2 + \cdots + t_nA\boldsymbol{v}_n = t_1\lambda_1\boldsymbol{v}_1 + t_2\lambda_2\boldsymbol{v}_2 + \cdots + t_n\lambda_n\boldsymbol{v}_n,\\ \boldsymbol{u}_2 &= A\boldsymbol{u}_1 = t_1\lambda_1A\boldsymbol{v}_1 + \cdots + t_n\lambda_nA\boldsymbol{v}_n = t_1\lambda_1^2\boldsymbol{v}_1 + \cdots + t_n\lambda_n^2\boldsymbol{v}_n\end{aligned}$$

and in general

$$\begin{aligned}A^k u_0 = A u_{k-1} = \boldsymbol{u}_k &= t_1\lambda_1^k\boldsymbol{v}_1 + t_2\lambda_2^k\boldsymbol{v}_2 + \cdots + t_n\lambda_n^k\boldsymbol{v}_n\\ &= \lambda_1^k\left[t_1\boldsymbol{v}_1 + t_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{v}_2 + \cdots + t_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{v}_n\right].\end{aligned}$$

Since $|\lambda_1| > |\lambda_2|, \ldots, |\lambda_n|$, $\left|\frac{\lambda_2}{\lambda_1}\right| < 1, \ldots, \left|\frac{\lambda_n}{\lambda_1}\right| < 1$ so that $\left(\frac{\lambda_i}{\lambda_1}\right)^k \to 0$ as $k \to \infty$. So for large k,

$$\boxed{\boldsymbol{u}_k \approx \lambda_1^k t_1\boldsymbol{v}_1}$$

i.e. $\boldsymbol{u}_k \approx$ eigenvector corresponding to λ_1.

if $|\lambda_1| > |\lambda_2| > \cdots$

→ Convergence depends on $|\lambda_2|$

Also, $u_{k+1} \approx \lambda_1^{k+1} t_1 v_1$, so

$$\frac{(u_{k+1})_j}{(u_k)_j} \approx \frac{(\lambda_1^{k+1} t_1 v_1)_j}{(\lambda_1^k t_1 v_1)_j} = \frac{\lambda_1^{k+1} t_1 (v_1)_j}{\lambda_1^k t_1 (v_1)_j} = \lambda_1.$$

Note this does not work if $\boxed{t_1 = 0,}$ i.e. if u_0 is a linear combination of only non-dominant eigenvectors. $\hookrightarrow$ might converge to v_2.

35.3 Deflation

The power method gives only the dominant eigenvalue. For symmetric matrices, we can find the next most dominant one by *deflation*, based on the following.

If A is $n \times n$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and v_1 is an eigenvector corresponding to λ_1, then set

$$B = A - \left(\frac{\lambda_1}{v_1^T v_1}\right) v_1 v_1^T.$$

For any matrix X, $(XX^T)^T = (X^T)^T X^T = XX^T$ i.e. XX^T is symmetric

Note that $v_1 v_1^T$ is a symmetric $n \times n$ matrix, and hence B is symmetric.

If A is symmetric and v_i is an eigenvector of A corresponding to $\lambda_i \neq 0$, then v_i is also an eigenvector of B corresponding to λ_i. $|\lambda_1| > |\lambda_2| > |\lambda_3| > \cdots \geq |\lambda_n|$ 0

The eigenvalues of B are $0, \lambda_2, \ldots, \lambda_n$.

For symmetric A with eigenvalues $\lambda_1, \ldots, \lambda_n$ where $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$, having used the power method to find an approx'n to λ_1 and v_1, form

$$B = A - \left(\frac{\lambda_1}{v_1^T v_1}\right) v_1 v_1^T$$

and repeat the power method on B to find an approximation for λ_2 and v_2.

In theory, you could repeat the power-deflation combination for other eigenvalues, but because the power method only approximates the dominant eigenvalue, we will be introducing some error into the method of deflation. Each time we repeat the process, the error not only propagates, but grows substantially.

35.3.1 Example

Apply deflation to the previous example of $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$, then use the power method on the new matrix to approximate the next most dominant eigenvalue and corresponding eigenvector.

DEFLATION $\quad \underline{v}_1 = \begin{pmatrix} 1.016 \\ 1 \end{pmatrix}, \lambda_1 = 3.969$

$$\underline{v}_1^T \underline{v}_1 = 2.032$$

$$\underline{v}_1 \underline{v}_1^T = \begin{pmatrix} 1.016 \\ 1 \end{pmatrix} \begin{pmatrix} 1.016 & 1 \end{pmatrix} = \begin{pmatrix} 1.032 & 1.016 \\ 1.016 & 1 \end{pmatrix}$$

$$B = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} - 1.953 \begin{pmatrix} 1.032 & 1.016 \\ 1.016 & 1 \end{pmatrix}$$

$$\simeq \begin{pmatrix} 0.985 & -0.984 \\ -0.984 & 1.047 \end{pmatrix}$$

POWER METHOD $\quad$ try $\underline{u}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

$$\underline{u}_1 = \begin{pmatrix} 0.985 \\ -0.984 \end{pmatrix}, \quad \underline{u}_2 = \begin{pmatrix} 1.938 \\ -1.999 \end{pmatrix}$$

$$\underline{u}_3 = \begin{pmatrix} 3.876 \\ -4.000 \end{pmatrix}, \quad \underline{u}_4 = \begin{pmatrix} 7.754 \\ -8.002 \end{pmatrix} = 7.754 \begin{pmatrix} 1 \\ -1.032 \end{pmatrix}$$

$$\Rightarrow \text{eigenvector} \approx \begin{pmatrix} 1 \\ -1.032 \end{pmatrix}$$

$$\& \begin{pmatrix} 0.985 & -0.984 \\ -0.984 & 1.047 \end{pmatrix} \begin{pmatrix} 1 \\ -1.032 \end{pmatrix} \approx \begin{pmatrix} 2.000 \\ -2.065 \end{pmatrix}$$

$$\Rightarrow \text{eigenvalue} \approx \frac{-2.065}{-1.032} \approx 2.001$$