1 The Power Method (5.3)

Let $A \in \mathbb{C}^{n\times n}$, and assume that A is semisimple. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote the eigenvalues associated with the linearly independent eigenvectors, $v_1, \ldots, v_n$, respectively. Assume that the vectors are ordered so that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$. If $|\lambda_1| > |\lambda_2|$, then λ_1 is called the **dominant eigenvalue**[1] and v_1 is called the **dominant eigenvector** of A.

1.1 The Iterative Power Method

Assuming we have $|\lambda_1| > |\lambda_2|$ as described above (otherwise, this method may not work), the general idea behind the iterative power method is that we can pick $q \in \mathbb{R}^n$ randomly. Then, we can form the sequence of vectors

$$q, Aq, A^2q, A^3q, \ldots.$$

To calculate this sequence, we don't necessarily need to form the powers of A explicitly. Each vector in the sequence can be obtained by multiplying the previous vector by A, e.g., $A^{j+1}q = A(A^jq)$. It's easy to show that the sequence converge, in a sense, to a dominant eigenvector, for almost all choices of q. Since $v_1, \ldots, v_n$ form a basis for $\mathbb{C}^n$, there exists constants $c_1, \ldots, c_n$ such that

$$q = c_1v_1 + c_2v_2 + \ldots + c_nv_n.$$

We don't know what $v_1, \ldots, v_n$ are, so we don't know what $c_1, \ldots, c_n$ are, either. However, it's clear that, for any choice of q, c_1 will be nonzero. The argument that follows is valid for every q for which $c_1 \neq 0$; multiplying by A, we have

$$\begin{aligned} Aq &= c_1Av_1 + c_2Ac_2 + \ldots + c_nAv_n \\ &= c_1\lambda_1v_1 + c_2\lambda_2v_2 + \ldots + c_n\lambda_nv_n. \end{aligned}$$

Similarly,

$$\begin{aligned} A^2q &= A(c_1\lambda_1v_1 + c_2\lambda_2v_2 + \ldots + c_n\lambda_nv_n) \\ &= c_1\lambda_1(Av_1) + c_2\lambda_2(Av_2) + \ldots + c_n\lambda_n(Av_n) \\ &= c_1\lambda_1(\lambda_1v_1) + c_2\lambda_2(\lambda_2v_2) + \ldots + c_n\lambda_n(\lambda_nv_n) \\ &= c_1\lambda_1^2v_1 + c_2\lambda_2^2v_2 + \ldots + c_n\lambda_n^2v_n. \end{aligned}$$

In general, we have

$$\begin{aligned} A^jq &= c_1\lambda_1^jv_1 + c_2\lambda_2^jv_2 + \ldots + c_n\lambda_n^jv_n \\ &= \lambda_1^j\left(c_1v_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^jv_2 + \ldots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^jv_n\right). \end{aligned}$$

So,

$$\frac{1}{\lambda_1^j}A^iq = c_1v_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^jv_2 + \ldots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^jv_n.$$

Notice that $\lim_{j\mapsto\infty}\left(\frac{\lambda_i}{\lambda_1}\right)^j = 0$ (because $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|$), so

$$\lim_{j\mapsto\infty}\frac{1}{\lambda_1^j}A^jq = c_1v_1,$$

the dominant eigenvector.

Remark: In fact, when we use the power method to converge to a dominant eigenvector, we need to know all the eigenvalues and then check whether they're strictly greater or not. So, this only works if λ_1 is known and it is strictly greater than λ_2 and so on.

[1] Basically, the largest absolute eigenvalue.

1.2 Scaling

Notice how we went from $A_j q$ to $\frac{1}{\lambda_1^j} A^j q$? In some sense, we can say we're scaling $A_j q$. Now, what if we don't know what the value of λ_1 is, but we need to do *some* scaling to get a reasonable convergence? In this case, we can still do some scaling, but not necessarily with λ_1.

Let's start with a random $q \in \mathbb{R}^n$; let $q_0 = q$. We want to use the iterative power formula method,

$$\boxed{q_{j+1} = \frac{1}{s_{j+1}} A q_j \qquad j = 0, 1, 2, \dots},$$

where $\frac{1}{s_{j+1}}$ is a scalar. Here,

$$s_{j+1} = ||Aq_j||_\infty.$$

In particular, all entries of q_{j+1} have absolute value ≤ 1. With this scaling, as $j \mapsto \infty$,

$$q_j \mapsto \text{Dominant eigenvector.}$$

$$s_j \mapsto \text{Dominant eigenvalue (in absolute value sense).}$$

Notice how s_j will eventually converge to the absolute value of the dominant eigenvalue. What if we want the actual value of λ_1? There is another version of the scalar. which is just

$$s_{j+1} = \text{sgn}((Aq_j)_i) \cdot ||Aq_j||_\infty,$$

where i is the index of the first entry of the vector Aq_j such that $|(Aq_j)_i| = ||Aq_j||_\infty$, i.e., the absolute value of the entry at index i is equal to the infinity norm of Aq_j. Then, $\text{sgn}((Aq_j)_i)$ is the sign function, which returns either -1 or 1 based on the sign of $(Aq_j)_i$.

(Example.) Suppose

$$Aq_j = \begin{bmatrix} -1 \\ 0 \\ 1/2 \\ 1 \\ 0 \end{bmatrix}.$$

We know that

$$||Aq_j||_\infty = 1.$$

So,

$$s_{j+1} = ||Aq_j||_\infty = 1.$$

To find the sign, we note that there are two values in Aq_j such that its absolute value equals $||Aq_j||_\infty = 1$;

- Value -1 at index $i = 1$ (top value),
- Value 1 at index $i = 4$ (second-to-bottom value).

We want the *first* entry, so $i = 1$. Therefore, $\text{sgn}((Aq_j)_1) = \text{sgn}(-1) = -1$ and so

$$s_{j+1} = \text{sgn}((Aq_j)_i) \cdot ||Aq_j||_\infty = -1 \cdot 1 = -1.$$

1.2.1 Stopping Criterion

Because this method is an iterative method, we need to stop at some point. We can set a threshold at $\epsilon > 0$. Stop the iteration when $||q_{j+1} - q_j||_\infty < \epsilon$, basically $q_{j+1} \approx q_j$. So,

$$\frac{1}{s_{j+1}} A q_j \approx q_j \implies Aq_j \approx s_{j+1} q_j.$$

Notice how this formula looks very similar to $Av = \lambda v$; in that sense, we can say that q_j is the approximated eigenvector and s_{j+1} is the approximated eigenvalue.

1.2.2 Flop Count and Rate of Convergence

Recall again $|\lambda_1| > |\lambda_2| \geq \ldots \geq |\lambda_n|$. Let's look at the flop count; notice that we have $\mathcal{O}(n^2)$ in every step of the iteration (matrix-vector multiplication). If we do N iterations, then the overall flop count is $\mathcal{O}(Nn^2)$.

Additionally, the convergence of the power method can be *slow*. In particular,

- If $|\lambda_2/\lambda_1|$ is small (e.g., $|1/1000|$), this means that $|\lambda_1| \gg |\lambda_2$ and convergence is fast.
- If $|\lambda_2/\lambda_1| \approx 1$ (e.g., $|0.99/1|$), then $|\lambda_1| \approx |\lambda_2|$ and convergence is slow.

(Example.) Suppose

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}.$$

Suppose you need to apply 3 steps of the power method to approximate λ_1 and v_1.

- $j = 0$: Let's start with[a] $q_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then,

$$Aq_0 = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}.$$

  We're now interested in the sign of s_1. To find the sign, we need to find the index i of the first entry of the vector Aq_0 such that $|(Aq_0)_i| = ||Aq_0||_\infty$. We know that $||Aq_0||_\infty = \max\{4, 2\} = 4$, so we find that $i = 1$ and so $\text{sgn}((Aq_0)_1) = \text{sgn}(4) = 1$.

  Thus, $s_1 = 1 \cdot 4 = 4$ and so

$$q_1 = \frac{1}{s_1} Aq_0 = \frac{1}{4} Aq_0 = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix}.$$

- $j = 1$: With $q_1 = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix}$, we have

$$Aq_1 = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 7/2 \\ 3/2 \end{bmatrix}.$$

  We know that $||Aq_1||_\infty = \frac{7}{2}$ and so we find the index at $i = 1$ and $\text{sgn}((Aq_1)_1) = \text{sgn}(7/2) = 1$. Thus, $s_2 = 1 \cdot \frac{7}{2} = \frac{7}{2}$ and so

$$q_2 = \frac{1}{s_2} Aq_1 = \frac{2}{7} \begin{bmatrix} 7/2 \\ 3/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3/7 \end{bmatrix}.$$

- $j = 3$: With $q_2 = \begin{bmatrix} 1 \\ 3/7 \end{bmatrix}$, we have

$$Aq_2 = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3/7 \end{bmatrix} = \begin{bmatrix} 24/7 \\ 10/7 \end{bmatrix}.$$

  We find that $||Aq_2||_\infty = 24/7$ and so, again, $i = 1$ and $\text{sgn}((Aq_2)_1) = \text{sgn}(24/7) = 1$. Thus, $s_3 = 1 \cdot \frac{24}{7} = \frac{24}{7}$ and

$$q_3 = \frac{1}{s_3} Aq_2 = \frac{7}{24} \begin{bmatrix} 24/7 \\ 10/7 \end{bmatrix} = \begin{bmatrix} 1 \\ 5/12 \end{bmatrix}.$$

With this, we find that $s_3 = 24/7 \approx 3.429$ and $q_3 = \begin{bmatrix} 1 \\ 5/12 \end{bmatrix} \approx \begin{bmatrix} 1 \\ 0.4167 \end{bmatrix}$. In actuality, the eigenvalue is $\lambda_1 = 3.4142$ and the eigenvector is $v_1 = \begin{bmatrix} 1 \\ 0.4142 \end{bmatrix}$, so after *three* steps, the approximation is very close.

[a]Remember that the initial vector is randomly chosen.

Remarks:

- The power method is easy to implement.
- The power method does *not* converge if $|\lambda_1| = |\lambda_2|$.