Eigenvalues and eigenvectors

Let $A \in \mathbb{R}^{n \times n}$. If $0 \neq v \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$ satisfy

$$Av = \lambda v$$

then λ is called **eigenvalue**, and v is called **eigenvector**.

Eigenvalues and eigenvectors

Let $A \in \mathbb{R}^{n \times n}$. If $0 \neq v \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$ satisfy

$$Av = \lambda v$$

then λ is called **eigenvalue**, and v is called **eigenvector**.

Given a matrix, we want to approximate its eigenvalues and eigenvectors. Some applications:

- Structural engineering (natural frequency, heartquakes)
- Electromagnetics (resonance cavity)
- Google's Pagerank algorithm
- ...

The characteristic polynomial

The eigenvalues of a matrix are the roots of **the characteristic polynomial**

$$p(\lambda) := \det(\lambda I - A) = 0$$

However, computing the roots of a polynomial is a very ill-conditioned problem! We cannot use this approach to compute the eigenvalues.

Eigenvalues and eigenvectors

Algorithms that compute the eigenvalues/eigenvectors of a matrix are divided into two categories:

1. Methods that compute all the eigenvalues/eigenvectors at once.
2. Methods that compute only a few (possibly one) eigenvalues/eigenvectors.

The methods are also different whether the matrix is symmetric or not. In this lesson we will discuss methods of type 2.

Diagonalizable matrices

Definition

We say that a matrix $A \in \mathbb{C}^{n \times n}$ is diagonalizable if there exists a non singular matrix U and a diagonal matrix D such that $U^{-1}AU = D$.

The diagonal element of D are the eignevalue of A and the column u_i of U is an eigenvector of A relative to the eigenvalue $D_{i,i}$.

Since a scalar multiple of an eigenvector is still an eigenvector, we can choose U such that $\|u_i\|_2 = 1$ for $i = 1, \dots, n$.

Finally, we observe that if A is diagonalizable, since U is non singular, then the vectors $\{u_1, \dots, u_n\}$ form a basis of $\mathbb{C}^n$.

From now on, we assume that the eigenvalues are numbered in decreasing order (in module), i.e.

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

Eigenvalues/eigenvectors of a symmetric matrix

Theorem

All the eigenvalues of a real symmetric matrix are **real**. Moreover, there exists a basis of eigenvectors $u_1, \ldots, u_n$, i.e.

$$Au_i = \lambda_i u_i$$

that have real entries and are orthonormal, i.e.

$$(u_i, u_j) = \delta_{ij}$$

The power method

We want to approximate the eigenvalue of A that is largest in module.

$v_0 =$ some vector with $\|v_0\| = 1$.
for k $= 1, 2, \ldots$

$w = Av_{k-1}$ — apply A

$v_k = w / \|w\|$ — normalize

$\mu_k = (v_k)^H Av_k$ — Reyleigh quotient

end

- $(v_k)^H$ denotes the transpose conjugate of the vector v_k
- if A is real and symmetric, since eigenvalues and eigenvectors are real, we can just use real numbers in the algorithm above and $(v_k)^H = (v_k)^T$ is the transpose of the vector v_k. This is the case we will consider in all examples.

The power method

We want to approximate the eigenvalue of A that is largest in module.

$v_0 =$ some vector with $\|v_0\| = 1$.
for k $= 1, 2, \ldots$

$w = Av_{k-1}$ — apply A

$v_k = w / \|w\|$ — normalize

$\mu_k = (v_k)^H A v_k$ — Reyleigh quotient

end

Theorem

Let $A \in \mathbb{C}^{n\times n}$ be a diagonalizable matrix. Assume $|\lambda_1| > |\lambda_2|$ and $v_0 = \sum_{i=1}^n \alpha_i u_i$, with $\alpha_1 \neq 0$. Then there exists $C > 0$, independent of k, such that

$$\|\widetilde{v}_k - u_1\|_2 \leq C \left|\frac{\lambda_2}{\lambda_1}\right|^k, \qquad \text{where } \widetilde{v}_k = \frac{\|A^k v_0\|}{\alpha_1 \lambda_1^k} v_k.$$

Proof

We expand v_0 on the eigenvector basis $\{u_1, \ldots, u_n\}$ choosen s.t. $\|u_i\| = 1$ for $i = 1, \ldots, n$:

$$v_0 = \sum_{i=1}^{n} \alpha_i u_i, \qquad \text{with } \alpha_1 \neq 0$$

It holds

$$A^k v_0 = \sum_{i=1}^{n} \alpha_i \lambda_i^k u_i \qquad \text{and} \qquad v_k = \frac{A^k v_0}{\|A^k v_0\|}$$

Hence, we can write

$$\widetilde{v}_k = \frac{A^k v_0}{\alpha_1 \lambda_1^k} = u_1 + \sum_{i=2}^{n} \frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1}\right)^k u_i$$

At this point, it holds

$$\|\widetilde{v}_k - u_1\|_2 = \left\| \sum_{i=2}^{n} \frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1}\right)^k u_i \right\|_2 \leq \sum_{i=2}^{n} \left\| \frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1}\right)^k u_i \right\|_2 = \sum_{i=2}^{n} \left|\frac{\alpha_i}{\alpha_1}\right| \left|\frac{\lambda_i}{\lambda_1}\right|^k$$

So, we obtain

$$\|\widetilde{v}_k - u_1\|_2 \leq \sum_{i=2}^{n} \left|\frac{\alpha_i}{\alpha_1}\right| \left|\frac{\lambda_i}{\lambda_1}\right|^k \leq (n-1) \cdot \max_{i=2,\dots,n} \left(\left|\frac{\alpha_i}{\alpha_1}\right|\right) \left|\frac{\lambda_2}{\lambda_1}\right|^k = C \left|\frac{\lambda_2}{\lambda_1}\right|^k,$$

where we have defined $C = (n-1) \cdot \max_{i=2,\dots,n} \left(\left|\frac{\alpha_i}{\alpha_1}\right|\right)$. Since C does not depend on k, this concludes the proof.

The previous theorem implies that the sequence $\{\widetilde{v}_k\}$ converges to the eigenvector u_1. Since $\widetilde{v}_k$ is a scalar multiple of v_k, they have the same direction and this direction converges to the direction of u_1. As a result, for k that goes to $+\infty$ the vector v_k tends to have the same direction of u_1. Thus v_k tends to be an eigenvector relaltive to λ_1.

Remark

if $|\lambda_2| \ll |\lambda_1|$ the convergence will be fast. On the other hand, if $\lambda_2 \approx \lambda_1$ the convergence will be slow.

We also have a convergence results for the approximation of the eigenvalue λ_1.

Corollary

Let $A \in \mathbb{C}^{n\times n}$ be a diagonalizable matrix. Assume $|\lambda_1| > |\lambda_2|$ and $v_0 = \sum_{i=1}^{n} \alpha_i u_i$, with $\alpha_1 \neq 0$. Then it holds

$$|\mu_k - \lambda_1| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right), \qquad \text{for } k \to +\infty.$$

For symmetric real matrices, we have a better convergence results:

Corollary

Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Assume $|\lambda_1| > |\lambda_2|$ and $v_0 = \sum_{i=1}^{n} \alpha_i u_i$, with $\alpha_1 \neq 0$. Then it holds

$$|\mu_k - \lambda_1| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right), \qquad \text{for } k \to +\infty.$$

Some observations

One of the hypothesis of the previous results is $\alpha_1 \neq 0$, where α_i are defined such that $v_0 = \sum_{i=1}^{n} \alpha_i u_i$. Clearly, $u_1, \dots, u_n$ are unknown and we cannot check if v_0 satisfies this hypothesis.

Practically this is not a real obstacle. Consider for simplicity the case of $A \in \mathbb{R}^{n \times n}$ symmetric. If we choose v_0 s.t $\alpha_1 = 0$ then:

- in *exact arithmetic*, we get $\lim_{k \to +\infty} \widetilde{v}_k = u_2$ and $\lim_{k \to +\infty} \mu_k = \lambda_2$, as long as $|\lambda_2| > |\lambda_3|$ and $\alpha_2 \neq 0$.
- in *finite arithmetic*, during the iterations of the Power Method, round-off errors cause the appearance of a non-zero component in the direction of u_1, in a certain v_k. When this happens, the method starts to converge towards the dominant eigenvalue λ_1 and its corresponding eigenvector u_1.

For more general $A \in \mathbb{C}^{n \times n}$ (possibly, real and non symmetric) the same happens but one has to use complex finite arithmetic and initialize v_0 as a vector with nonzero real and imaginary entries.

Stopping criterion

A simple stopping criterion for the power method is based on the residual:

$$\text{Stop when} \qquad \|Av_k - \mu_k v_k\| \leq \texttt{tol}$$

How can we compute other eigenvalues and eigenvectors?

Let $\mu \in \mathbb{C}$ a user-specified parameter that is not an eigenvalue of A, we want to approximate the closest eigenvalue of A to μ, i.e.

$$\lambda_J = \operatorname*{argmin}_i |\mu - \lambda_i|$$

Inverse Power method

Input: $A \in \mathbb{C}^{n\times n}$, $v_0 \in \mathbb{C}^n$ with $\|v_0\| = 1$, MAXITER $\in \mathbb{N}$, tol $\in \mathbb{R}^+$.

for $\mathrm{k} = 1, 2, \ldots, \mathrm{MAXITER}$

$$w = (A - \mu I)^{-1} v_{k-1} \qquad \text{(equivalently, solve } (A - \mu I)\, w = v_{k-1}\text{)}$$

$$v_k = w / \|w\|$$

$$\mu_k = (v_k)^H A v_k \qquad \text{(Rayleigh quotient with } A\text{)}$$

Check the Stopping criterion

end

Output: μ_k and v_k.

Since μ is not an eigenvalue of A, the matrix $A - \mu I$ is non singular.

Since $Au_i = \lambda_i u_i$, then $(A - \mu I)u_i = (\lambda_i - \mu)u_i$, and then $\frac{1}{\lambda_i - \mu} u_i = (A - \mu I)^{-1} u_i$. Let λ_J be the eigenvalue of A closest to μ, the largest (in module) eigenvalue of $(A - \mu I)^{-1}$ is then $\frac{1}{\lambda_J - \mu}$, and the relative eigenvector is u_J. The inverse power method is just a power method applied to $(A - \mu I)^{-1}$, and the previous results apply: $\widetilde{v}_k$ converges to u_J. Since the Rayleigh quotient μ_k is computed with A instead of $(A - \mu I)^{-1}$, it converges to λ_J.

Theorem

Assume $|\mu - \lambda_J| < |\mu - \lambda_i| \ \forall \ i = 1, \ldots, n, \ i \neq J$ and $v_0 = \sum_{i=1}^{n} \alpha_i u_i$, with $\alpha_J \neq 0$. Then

$$\lim_{k \to +\infty} \mu_k = \lambda_J$$

and

$$\lim_{k \to +\infty} \|\widetilde{v}_k - u_J\|_2 = 0, \qquad \text{where } \widetilde{v}_k = \frac{\|A^k v_0\|}{\alpha_1 \lambda_1^k} v_k.$$

Note that if $\mu = 0$, the method approximates the eigenvalue of A that is smallest in module.