# Iterative methdos

Zecheng Zhang

April 10, 2023

## 1 Eigenvalue problem

We will consider symmetric matrix $A \in \mathbb{R}^{m\times m}$. We define the Rayleigh Quotient,

$$r(x) = \frac{x^t A x}{x^t x}. \tag{1}$$

Note that if $x$ is an eigenvector of $A$, $r(x) = \lambda$ is its eigenvalue.

One way to understand this formula is: given $x$, what is the scale $\alpha$ which acts almost like an eigenvalue of $x$ in the sense that $Ax - \alpha x$ is minimized? This is a least square problem, but $x$ is the matrix $\alpha$ is the unknown vector, and $Ax$ is the right-hand side $b$ vector. We can see that $\alpha = r(x)$ if we consider the normal equation.

Take the derivative of $r(x)$ with respect to all component $x_j$ of $x$, we can easily derive that,

$$\nabla r(x) = \frac{2}{x^t x}(Ax - r(x)x). \tag{2}$$

We can see that when $x$ is the eigenvector, the gradient vanishes. Conversely, if the gradient is trivial with $x \neq 0$, $x$ is an eigenvector with eigenvalue $r(x)$.

**Theorem 1.1.** Let $q_j$ be an eigenvector of $A$, we have

$$r(x) - q_j = \mathcal{O}(\|x - q_j\|^2), \tag{3}$$

as $x \to q_j$.

The Power iteration is expected to return an eigenvector corresponding to the largest eigenvalues.

---

**Algorithm 1:** Power Iteration

---

Set $v_0$ with $\|v_0\| = 1$.
**for** $k = 1$ *to* ... **do**
  $w = Av^k$
  $v^k = w/\|w\|$
  $\lambda^k = (v^k)^T A v^k$

---

**Theorem 1.2.** Suppose $|\lambda_1| > |\lambda_2| \geq ... \geq |\lambda_m| \geq 0$ and $q_1^T v^0 \neq 0$. Then the algorithm satisfies,

$$\|v^{(k)} - q_1\| = \mathcal{O}(|\frac{\lambda_2}{\lambda_1}|^k), \tag{4}$$

$$|\lambda^{(k)} - \lambda_1| = \mathcal{O}(|\frac{\lambda_2}{\lambda_1}|^{2k}), \tag{5}$$

as $k \to \infty$

Rayleigh quotient of $A \in \mathbb{R}^{m\times m}$

$$r(x) = \frac{x^t A x}{x^t x}, \quad x \in \mathbb{R}^m$$

i. Suppose $(x, \lambda)$ is an eigen-pair of $A$, $Ax = \lambda x$,

$$r(x) = \frac{x^t \lambda x}{x^t x} = \lambda$$

ii. $$\nabla r(x) = \begin{pmatrix} \frac{\partial r}{\partial x_1} \\ \cdots \\ \frac{\partial r}{\partial x_m} \end{pmatrix} = \frac{2}{x^t x}\left( Ax - r(x)\, x \right)$$

when $x$ is an eigenvector of $A$,

$\Rightarrow \quad \nabla r(x) = 0$

when $\nabla r(x) = 0$, $\Rightarrow$ $x^{\neq 0}$ is an eigenvector of $A$

Thm 1.1.

Let $(\xi_j, \lambda_j)$ be an eigen-pair of $A$,

$$r(x) - \lambda_j = O\left( \|x - \xi_j\|^2 \right), \quad \text{as } x \to \xi_j$$

$\rightarrow$ Rayleigh quotient is quadratically accurate estimation of the eigenvalue.

$\rightarrow$ Given the approximation to e.vector we have an approximation of the 2nd to the eigenvalue.

Power iteration.

Set $v^{(0)}$, $\|v_0\| = 1$.

for $k = 1, \ldots,$ ①

$w = A v^{(k-1)}$

$v^{(k)} = w / \|w\|$

$\lambda^{(k)} = \underbrace{(v^{(k)})^t A (v^{(k)})}_{\text{Rayleigh quotient of } A.}$

By thm 1.1, $\lambda^{(k)} \to \lambda_1$

① $k \to \infty$, we have the convergence, but to write the code, we need some terminate condition.

---

② $K_n(A)$ Krylov subspace

$= [\,b,\ Ab,\ A^2 b,\ \ldots\ A^{n-1} b\,]$

Here $b = v^{(0)}$

Represent $v^{(k)}$ in $K_n(A)$

---

③ A is symmetric, $\Rightarrow$ all eigenvalues of A are real, e.g., $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_m$

& Power iteration can find $(\lambda_1, v_1)$

---

④ The convergence is linear, Specifically, the algorithm will reduce the error by a factor $\left(\frac{\lambda_2}{\lambda_1}\right)$ in each iteration.

---

⑤ When eigen-gap is small the algo converges slowly.

---

Inverse iteration.

Motivation, we want to design an algorithm which can identify any e.val & e.vectors of A.

Suppose $\mu$ is not an eigenvalue of $A$.

suppose $(v, \alpha)$ is an eigen-pair of $(A-\mu I)^{-1}$

$$(A-\mu I)^{-1} v = \alpha v$$

$$v = \alpha (A-\mu I) v$$

$$v = \alpha A v - \alpha \mu v$$

$$\alpha A v = (1+\alpha\mu) v \qquad (\text{assume } \alpha \neq 0)$$

$$A v = (1+\alpha\mu)/\alpha \; v$$

$\Rightarrow$ $v$ is also an eigenvector of $A$.

Now let $A v_i = \lambda_i v_i$

$$\Rightarrow \quad \alpha = \frac{1}{\lambda_i - \mu}$$

(eigenvalue of $(A-\mu I)^{-1}$)

Here, if $\mu \to \lambda_i$, $\Rightarrow$ $\alpha$ is huge, $\alpha$ should be larger than the eigenvalues of $(A-\mu I)^{-1}$

$\Rightarrow$ we can use the power iteration to identify eigen-pair $(\alpha, v)$ of $(A - \mu I)^{-1}$.

$\Rightarrow$ we can get the eigen-vector of $A$.

Rmk, why is this called the inverse iteration ?

approximation $\mu$ to $\lambda_i$ $\longrightarrow$ eigenvectors of $A$.

Inverse iteration

$v^{(0)}$, $\|v^{(0)}\| = 1$

For $k = 1$, to $\infty$

solve $(A - \mu I) w = v^{(k-1)}$ for $w$.

$v^{(k)} = w / \|w\|$

$\lambda^{(k)} = (v^{(k)})^t A v^{(k)}$

eigenval of $A$.

$A$ & $(A - \mu I)^{-1}$ share the eigenvectors.

**Remark 1.** Power iteration has some limitations.

1. It can only find the largest eigenvectors corresponding to the largest eigenvalues.
2. The convergence is linear, i.e., the algorithm reduces the error by a factor $|\frac{\lambda_2}{\lambda_1})$ in every iteration.
3. The quality of the convergence depends on the quotient. If there is no huge eigen-gap, the convergence is slow.

### 1.1 Inverse Iteration

Let $\mu$ be a number which is not an eigenvalue of $A$, the eigenvectors of $(A - \mu I)^{-1}$ are the same as the eigenvectors of $A$, and the coresponding eigenvalues are $(\lambda_j - \mu)^{-1}$, where $\{\lambda_j\}$ are the eigenvalues of $A$.

This motivates us to design an algorithm to identify $\lambda_j$ and the corresponding eigenvectors of $A$. Suppose we know any estimate of $\lambda_j$ and denote it as $\mu$. $(\mu - \lambda_j)^{-1}$ will be very large. According to the Remark, the power iteration can identify $q_j$, which are the eigenvectors of $(A - \mu I)^{-1}$ (also the eigenvectors of $A$). This idea is called the inverse iteration.

**Algorithm 2:** Inverse iteration

$v^0$ = some vectors with norm 1
**for** $k = 1$ *to* ... **do**
  Solve $(A - \mu I)w = v^{k-1}$ for $w$
  $v^k = w/\|w\|$
  $\lambda^k = (v^k)^T A v^k$.

Rayleigh quotient is one method to estimate eigenvalues from an eigenvector estimation. Inverse iteration is an estimate of eigenvector from the eigenvalues.

**Algorithm 3:** RQ iteration

$v^0$ = some vectors with norm 1
$\lambda^0 = v^0 A v^0$ = coresponding Rayleigh quotient. **for** $k = 1$ *to* ... **do**
  Solve $(A - \lambda^{k-1} I)w = v^{k-1}$ for $w$
  $v^k = w/\|w\|$
  $\lambda^k = (v^k)^T A v^k$.

Without proof, the Rayleigh Quotient iteration has cubic convergence.

## 2 Iterative methods

In this section, let us consider matrix $A \in \mathbb{R}^{m \times m}$. Broadly speaking, the idea of iterative methods is to:

1. Gradually refine the solution iteratively.
2. Each iteration should be (a lot) cheaper than direct methods.