Iterative methdos

Zecheng Zhang

April 11, 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 α 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 α 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 $k \to \infty$

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 μ 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 λ_j and the corresponding eigenvectors of A. Suppose we know any estimate of λ_j and denote it as μ. $(\mu - \lambda_j)^{-1}$ 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 the 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 Reduction to Hessenberg form

Schur factorization returns $A = QTQ^*$, where T is a triangular matrix, i.e., we would like to apply unitary similarity transformation to introduce zeros below the diagonal. The natural first idea is to use the Householder.

The first Householder reflector Q_1^* multiplied on the left of A would introduce zeros below the diagonal in the first column, and the Householder reflector will change all rows of A. This is

Inverse iteration is also linear.

Power iteration: approximation of the eigenvector $\xrightarrow[\text{Thm 1.1}]{\text{Rayleigh quotient}}$ approximation of the eigenvalue

Inverse iteration: approximation of the eigenvalue (μ) $\longrightarrow$ approximation of the eigenvector.

Algorithm Rayleigh quotient iteration.

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

$\lambda^{(0)} = (v^{(0)})^{+} A v^{(0)}$ (Rayleigh quotient)

for $k = 1, 2, \ldots$

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

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

$\lambda^{(k)} = (v^{(k)})^{+} A v^{(k)}$ Rayleigh quotient

Rayleigh quotient iteration has a 3rd order convergence.

$$\| v^{(k+1)} - q_j \| = O\left(\| v^{(k)} - q_j \|^{3} \right)$$

good up to now; however, if we complete the process of multiplying Q_1 on the right, all zeros previously introduced are destroyed. We will verify this in class.

The good idea in step 1 is to choose a unitary matrix Q_1^* that will leave the first row unchanged. It will change the second row to the last row and introduce zeros below the second entry in the first column. It can be verified that the right multiplication by Q_1 will not change the zeros introduced by Q_1^*. After repeating this process for $m-2$ times, the resulting matrix is in the Hessenberg form, denoted as H.

Algorithm 4: Reduction to Hessenberg

```
for k = 1 to m − 2 do
    x = A_{k+1:m,k}
    v_k = (sign(x_1))‖x‖_2 e_1 + x
    v_k = v_k/‖v_k‖
    A_{k+1:m,k:m} = A_{k+1:m,k:m} − 2 v_k v_k^* A_{k+1:m,k:m}
    A_{1:m,k+1:m} = A_{1:m,k+1:m} − 2 A_{1:m,k+1:m} v_k v_k^*
```

When A is Hermitian, H is symmetric, then H is a tridiagonal matrix.

Schur factorization.

$A \in \mathbb{R}^{m \times m}$, $A = QTQ^*$,

Q is unitary, T is upper triangular

diagonal entries of T are

eivals of A.

$$Q^* A Q = T$$

Everytime, we can introduce some zeros below the diagonal entries of A.

At step 1,

$$\underset{A}{\begin{bmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{bmatrix}} \xrightarrow[\text{Householder reflector}]{Q_1^*} \underset{Q_1^* A}{\begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{bmatrix}}$$

Q_1^* can be the Householder reflector

however all 4 (all) rows will be modified.

Now $Q_1^* A Q_1$, what is the structure of $Q_1^* A Q_1$?

To study this, let us consider [consider the transpose of $Q_1^* A Q$ target]

$$Q_1^* (A^* Q)$$

$$\begin{bmatrix} x & 0 & 0 & 0 \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{bmatrix} \xrightarrow{Q_1^*} \begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{bmatrix}$$

$$(Q_1^* A)^* \qquad\qquad Q_1^* (Q_1^* A)^*$$

Finally, if transpose,

$$Q_1^* A Q_1 = \begin{pmatrix} x & 0 & 0 & 0 \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} \quad X$$

Hessenberg form

At step 1

$$\underset{A}{\begin{pmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{pmatrix}} \xrightarrow[\text{Householder reflector}]{Q_1^*} \underset{Q_1^* A.}{\begin{pmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{pmatrix}}$$

3 rows will be modified but the 1st row is unchanged.

study $(Q_1^* A Q_1)^*$

$$\begin{array}{c} \ 12 \\ \underset{(Q_1^* A)^*}{\begin{pmatrix} \times & \times & 0 & 0 \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{pmatrix}} \end{array} \xrightarrow{Q_1^*} \underset{(Q_1^* A Q_1)^*}{\begin{pmatrix} \times & \times & 0 & 0 \\ \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{pmatrix}}$$

mm

only row 2 3 4 will be modified.

Take the transpose,

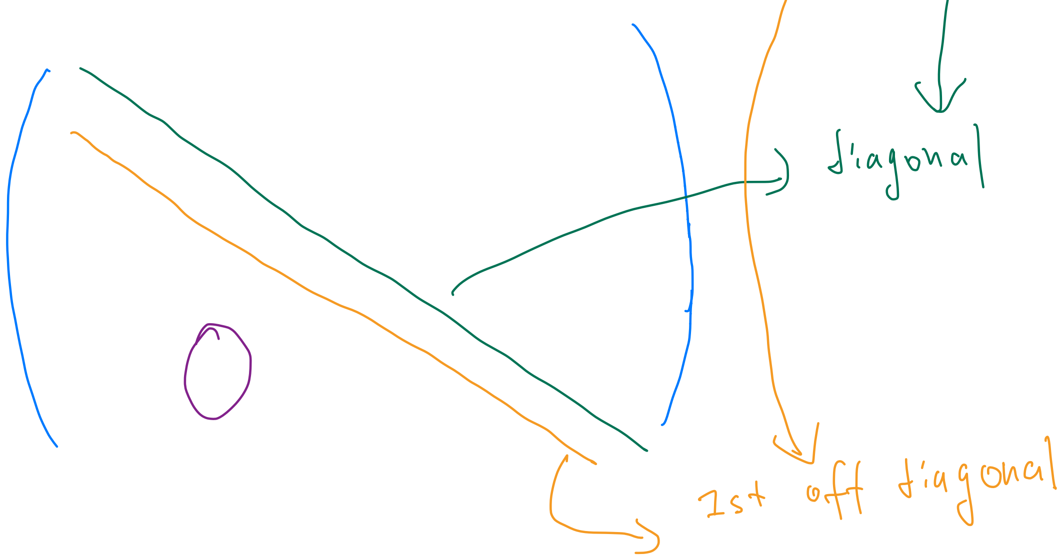
$$Q_1^* A Q_1 = \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ 0 & x & x & x \\ 0 & x & x & x \end{pmatrix}$$

At step 2.

$$Q_2^* \left(Q_1^* A Q_1 \right) Q_2 \rightarrow \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \end{pmatrix}$$

$= 0$

Hessenberg form



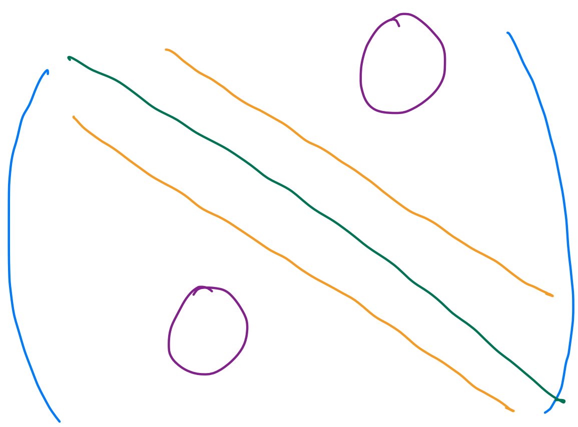
A = is symmetric. $A \in \mathbb{R}^{m \cdot m}$

total iteration = $m-2$.

$$\underbrace{Q_{m-2}^* \cdots Q_2^* Q_1^*}_{Q^*} A \underbrace{Q_1 Q_2 \cdots Q_{m-2}}_{Q} = H$$

$$H^* = (Q^* A Q)^* = Q^* A Q = H$$

$\Rightarrow$ H is symmetric (Hermitian if $\mathbb{C}^{n \cdot n}$)

$H =$  $\rightarrow$ tridiagonal matrix.