SI231b: Matrix Computations

Lecture 15: Eigenvalue Computations

Yue Qiu

qiuyue@shanghaitech.edu.cn

School of Information Science and Technology

ShanghaiTech University

Nov. 04, 2021

Recap: Eigenvalue Revealing Decomposition

Factorize a matrix to a form in which eigenvalues are explicitly displayed

- Diagonalization, $A = V\Lambda V^{-1}$, exists if and only if A is nondefective.
- Schur decompositon, $A = QTQ^H$ always exists.
- Jordan canonical form, $A = SJS^{-1}$ always exists (will not be introduced in our lecture), where

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix}$$

with

$$J_i = \begin{bmatrix} \lambda_i & & & \\ & \lambda_i & & \\ & & \ddots & \\ & & & \lambda_i \end{bmatrix}, \quad \text{or} \quad J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

Outline

- Facts About Eigenvalues
- Power Iteration
- Inverse Iteration
- Subspace Iteration

Some Facts About Eigenvalues

- Eigenvalues of Hermitian matrices are real
- Eigenvalues of real symmetric matrices are real
- Eigenvectors of real symmetric matrices are also real
- Complex eigenvalues of real matrices appear in conjugate pair.
  - For $\mathsf{A} \in \mathbb{R}^{n\times n}$, if (λ, v) is an eigenpair, then also $(\lambda^*, \mathsf{v}^*)$
- Skew-Hermitian matrices ($\mathsf{A} = -\mathsf{A}^H$) have only pure imaginary eigenvalues
- Hermitian/real symmetric matricres are diagonalizable.

The Largest Eigenvalue and Associated Eigenvector

Let $A \in \mathbb{C}^{n\times n}$ be diagonalizable, i.e., $A = V\Lambda V^{-1}$ with $V = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$, and $\Lambda = \mathrm{diag}(\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n)$. Assume that

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|.$$

The following iteration generates a sequence of $(\lambda^{(k)},\ q^{(k)})$ that converges to $(\lambda_1,\ v_1)$.

Power Iteration:

random selection $q^{(0)} \in \mathbb{C}^n$

for $k = 1,\ 2,\ \cdots$

$$z^{(k)} = Aq^{(k-1)}$$

$$q^{(k)} = \frac{z^{(k)}}{\|z^{(k)}\|_2}$$

$$\lambda^{(k)} = (q^{(k)})^H A q^{(k)}$$

end

Convergence of Power Iteration

The Power Iteration can only compute the largest eigenvalue and associated eigenvector with convergence rate

- $|\lambda^{(k)} - \lambda_1| = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$
- $\|\mathrm{q}^{(k)} - \mathrm{v}_1\| = \mathcal{O}\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$
- can have slow convergence when λ_2 is close to λ_1 in magnititude, i.e., $\left|\frac{\lambda_2}{\lambda_1}\right|$ is close to 1.
- convergence can be made faster by using a shift μ with

$$\left|\frac{\lambda_1 - \mu}{\mu - \lambda_j}\right| < \left|\frac{\lambda_2}{\lambda_1}\right|,$$

together with *Inverse Iteration*. Here λ_1 and λ_j are the closest and second closest eigenvalues to μ.

Suppose μ is not an eigenvaue of A, the inverse iteration is given by

Inverse Iteration:

```
random selection q^(0) ∈ ℂ^n
for k = 1, 2, ···
        z = (A − μI)^{-1} q^(k−1)      solve (A − μI)z = q^(k−1)
        q^(k) = z / ||z||_2
        λ^(k) = (q^(k))^H A q^(k)
end
```

random selection $q^{(0)} \in \mathbb{C}^n$

for $k = 1, 2, \cdots$

$$z = (A - \mu I)^{-1} q^{(k-1)} \qquad \texttt{solve } (A - \mu I) z = q^{(k-1)}$$

$$q^{(k)} = \frac{z}{\|z\|_2}$$

$$\lambda^{(k)} = (q^{(k)})^H A q^{(k)}$$

end

- compute the eigenvalue closest to μ
- convergence rate

$$\left| \frac{\mu - \lambda_j}{\mu - \lambda_k} \right|$$

where λ_j and λ_k are the closest and second closest eigenvalues to μ.

Efficiency per iteration vs Number of iterations?

Power Iterations for a Set of Vectors

From the Power Iteration, we know that

- $A^k q_0$ converges to the eigenvector associated with the largest eigenvalue in magnititude.
- if we start with a set of linearly independent vectors $\{q_1, \ q_2, \ \cdots, q_r\}$, then $A^k\{q_1, \ q_2, \ \cdots, q_r\}$ should converge (under suitable assumptions) to a subspace spanned by eigenvectors of A associated with r largest eigenvalues in magnititude.

Suppose there is a gap between the r largest eigenvalues in magnititude and λ_{r+1}, i.e, $|\lambda_1| \geq |\lambda_2| \geq \cdots |\lambda_r| > |\lambda_{r+1}|$

Subspace Iteration:

```
random selection Q^(0) with orthonormal columns
for k = 1, 2, ...
    Z_k = A Q^(k-1)
    Z_k = Q^(k) R^(k)   reduced QR factorization
end
```

- Z_k and $Q^{(k)}$ has the same column space
- equal to the column space of $A^k Q^{(0)}$

Subspace Iteration

- $Q^{(k)}$ converge to subspace associated with r largest eigenvalues in magnititude (dominant invariant subspace).
- $\operatorname{diag}\left(\left(Q^{(k)}\right)^H AQ^{(k)}\right) \to \{\lambda_1,\ \lambda_2,\ \cdots,\lambda_r\}$
- $\|q_i^{(k)} - v_i\| = \mathcal{O}\left(\left|\frac{\lambda_{r+1}}{\lambda_i}\right|^k\right),\ i = 1,\ 2,\ \cdots,\ r$
- $\left|\lambda_i^{(k)} - \lambda_i\right| = \mathcal{O}\left(\left|\frac{\lambda_{r+1}}{\lambda_i}\right|^k\right),\ i = 1,\ 2,\ \cdots,\ r$
- also called simultaneously iteration or orthogonal iteration
- when $r = n$, it coincides with QR iteration

Readings

You are supposed to read

- Lloyd N. Trefethen and David Bau III. *Numerical Linear Algebra*, SIAM, 1997.

  Lecture 25, 26