Mathematics Methods for Computer Science

Instructor: Xubo Yang

SJTU-SE DALAB

Lecture

Eigenproblems II: Computation

Setup

$$A \in \mathbb{R}^{n\times n} \text{ symmetric}$$

$$\vec{x}_1, \ldots, \vec{x}_n \in \mathbb{R}^n \text{ eigenvectors}$$

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n| \text{ eigenvalues}$$

Setup

$A \in \mathbb{R}^{n \times n}$ symmetric

$\vec{x}_1, \ldots, \vec{x}_n \in \mathbb{R}^n$ eigenvectors

$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$ eigenvalues

Review (Spectral Theorem):
What do we know about the eigenvectors?

Usual Trick

$$\vec{v} \in \mathbb{R}^n$$

$$\Downarrow$$

$$\vec{v} = c_1\vec{x}_1 + \cdots + c_n\vec{x}_n$$

recall：对于哈密顿矩阵，其特征向量构成一个正交矩阵，回忆线性代数知识，一组正交向量构成的正交基的线性组合可以表示任意向量。

Observation

$$\begin{aligned} A\vec{v} &= c_1 A\vec{x}_1 + \cdots + c_n A\vec{x}_n \\ &= c_1\lambda_1\vec{x}_1 + \cdots + c_n\lambda_n\vec{x}_n \end{aligned}$$

$$A^2\vec{v} = \lambda_1^2\left(c_1\vec{x}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^2 c_2\vec{x}_2 + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^2 c_n\vec{x}_n\right)$$

$$\vdots$$

$$A^k\vec{v} = \lambda_1^k\left(c_1\vec{x}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^k c_2\vec{x}_2 + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^k c_n\vec{x}_n\right)$$

For Large k

在第四页PPT已经进行过人为的排序了

$$A^k\vec{v} \approx \lambda_1^k c_1 \vec{x}_1$$

(assuming $|\lambda_2| < |\lambda_1|$ and $c_1 \neq 0$)

Power Iteration

$$\vec{v}_k = A\vec{v}_{k-1}$$

Power Iteration

$$\vec{v}_k = A\vec{v}_{k-1}$$

Question:

What if $|\lambda_1| > 1$

Normalized Power Iteration

$$\vec{w}_k = A\vec{v}_{k-1}$$
$$\vec{v}_k = \frac{\vec{w}_k}{||\vec{w}_k||}$$

Normalized Power Iteration

$$\vec{w}_k = A\vec{v}_{k-1}$$
$$\vec{v}_k = \frac{\vec{w}_k}{||\vec{w}_k||}$$

Question: Which norm?

一般为2范数

Eigenvalues of Inverse Matrix

$$A\vec{v} = \lambda\vec{v} \Rightarrow A^{-1}\vec{v} = \frac{1}{\lambda}\vec{v}$$

λ _1最大<==>1/ λ _1最小

Eigenvalues of Inverse Matrix

$$A\vec{v} = \lambda\vec{v} \Rightarrow A^{-1}\vec{v} = \frac{1}{\lambda}\vec{v}$$

Question:

What is the largest-magnitude eigenvalue?

Inverse Iteration

$$\vec{w}_k = A^{-1}\vec{v}_{k-1}$$

$$\vec{v}_k = \frac{\vec{w}_k}{||\vec{w}_k||}$$

Inverse Iteration

$$\vec{w}_k = A^{-1}\vec{v}_{k-1}$$
$$\vec{v}_k = \frac{\vec{w}_k}{||\vec{w}_k||}$$

Question: How to make faster?

分解!

Inverse Iteration with LU

Solve $L\vec{y}_k = \vec{v}_{k-1}$

Solve $U\vec{w}_k = \vec{y}_k$

Normalize $\vec{v}_k = \dfrac{\vec{w}_k}{||\vec{w}_k||}$

Eigenvalues of Shifted Matrix

前面的PPT使用迭代的方法求解出了最大特征值以及最小特征值(逆矩阵)，这里开始求解中间的一些特征值。求解的策略是使用一些参数σ去求距离这些σ最近的特征值。

$$A\vec{v} = \lambda\vec{v} \Rightarrow (A - \sigma I)\vec{v} = (\lambda - \sigma)\vec{v}$$

Shifted Inverse Iteration

To find eigenvalue closest to σ:

$$\vec{v}_{k+1} = \frac{(A - \sigma I)^{-1}\vec{v}_k}{||(A - \sigma I)^{-1}\vec{v}_k||}$$

Heuristic: Convergence Rate

Recall power iteration:

$$A^k\vec{v} =$$
$$\lambda_1^k\left(c_1\vec{x}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^k c_2\vec{x}_2 + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^k c_n\vec{x}_n\right)$$

Strategy for Better Convergence

For power iteration, find σ with

$$\left|\frac{\lambda_2-\sigma}{\lambda_1-\sigma}\right| < \left|\frac{\lambda_2}{\lambda_1}\right|$$

Least-Squares Approximation

If $\vec{v}_0$ is approximately an eigenvector:

$$\underset{\lambda}{arg\,min}\, ||A\vec{v}_0 - \lambda\vec{v}_0||_2^2 = \frac{\vec{v}_0^\top A\vec{v}_0}{||\vec{v}_0||_2^2}$$

那么将这个式子带到之前求解中间的特征值的思路中去，实际上就是，σ对应的特征值就相当于估计值，λ就是理论值，所以当偏差最小的对应的σ值就是估计出来的中间的λ的值！

这个式子是在求v_0估计矩阵得出的特征值与理论特征值偏差最小时对应的λ的值

Rayleigh Quotient Iteration

$$\vec{w}_k = (A - \sigma_k I)^{-1} \vec{v}_{k-1}$$

$$\vec{v}_k = \frac{\vec{w}_k}{\|\vec{w}_k\|}$$

$$\sigma_{k+1} = \frac{\vec{v}_k^\top A \vec{v}_k}{\|\vec{v}_k\|_2^2}$$

Rayleigh Quotient Iteration

$$\vec{w}_k = (A - \sigma_k I)^{-1} \vec{v}_{k-1}$$

$$\vec{v}_k = \frac{\vec{w}_k}{\|\vec{w}_k\|}$$

$$\sigma_{k+1} = \frac{\vec{v}_k^\top A \vec{v}_k}{\|\vec{v}_k\|_2^2}$$

Efficiency per iteration vs. number of iterations?

Unlikely Failure Mode for Iteration

What is $\vec{v}_0$?

Unlikely Failure Mode for Iteration

What is $\vec{v}_0$?

What happens when $\vec{v}_0 \cdot \vec{x}_1 = 0$?

Bug or Feature?

1. Compute $\vec{x}_0$ via power iteration.
2. Project $\vec{x}_0$ out of $\vec{v}_0$.
3. Compute $\vec{x}_1$ via power iteration.
4. Project $span\{\vec{x}_0, \vec{x}_1\}$ out of $\vec{v}_0$.
5. $\ldots$

Bug or Feature?

1. Compute $\vec{x}_0$ via power iteration.
2. Project $\vec{x}_0$ out of $\vec{v}_0$.
3. Compute $\vec{x}_1$ via power iteration.
4. Project $span\{\vec{x}_0, \vec{x}_1\}$ out of $\vec{v}_0$.
5. $\ldots$

Assumption: A is symmetric.

Avoiding Numerical Drift

Do power iteration on $P^\top AP$ where P projects out known eigenvectors.

Deflation

Modify A so that power iteration reveals an eigenvector you have not yet computed.

Similarity Transformations

Similar matrices

Two matrices A and B are similar if there exists T with $B = T^{-1}AT$.

Similarity Transformations

Similar matrices

Two matrices A and B are similar if there exists T with $B = T^{-1}AT$.

Proposition

Similar matrices have the same eigenvalues.

Householder Asymmetric Deflation

这里的矩阵H为正交矩阵

$$
\begin{aligned}
& H\vec{x}_1 = \vec{e}_1 \\
\Longrightarrow\ & HAH^\top \vec{e}_1 = HAH\vec{e}_1 \text{ by symmetry} \\
& = HA\vec{x}_1 \text{ since } H^2 = I \\
& = \lambda_1 H\vec{x}_1 \\
& = \lambda_1 \vec{e}_1
\end{aligned}
$$

通过这种方式得到了e_1的一个特征值一定是λ_1

Householder Asymmetric Deflation

$$HAH^\top = \begin{pmatrix} \lambda_1 & \vec{b}^\top \\ \vec{0} & B \end{pmatrix}$$

Similarity transform of $A \Rightarrow$ same eigenvalues.

Householder Asymmetric Deflation

$$HAH^\top = \begin{pmatrix} \lambda_1 & \vec{b}^\top \\ \vec{0} & B \end{pmatrix}$$

Similarity transform of $A \Rightarrow$ same eigenvalues.

Do power iteration on B.

方式与上述类似

Householder Asymmetric Deflation

$$HAH^\top = \begin{pmatrix} \lambda_1 & \vec{b}^\top \\ \vec{0} & B \end{pmatrix}$$

Similarity transform of $A \Rightarrow$ same eigenvalues.

Do power iteration on B.

Reveals eigenvalues + vectors one at a time.

Conjugation without Inversion

$$
\begin{aligned}
& Q^{-1} = Q^\top \\
\Rightarrow\ & Q^{-1}AQ = Q^\top AQ
\end{aligned}
$$

Conjugation without Inversion

$$Q^{-1} = Q^{\top}$$
$$\Rightarrow Q^{-1}AQ = Q^{\top}AQ$$

But which Q ?

Should involve matrix structure but be easy to compute.

要选择与原矩阵有一定的结构联系并且为正交矩阵的矩阵应该想到QR分解。

Experiment

$$A = QR$$

$$Q^{-1}AQ = \ ?$$ RQ

QR Iteration

$$\begin{aligned} A_1 &= A \\ \text{Factor } A_k &= Q_k R_k \\ \text{Multiply } A_{k+1} &= R_k Q_k \end{aligned}$$

Commutativity

Lemma

Take $A, B \in \mathbb{R}^{n \times n}$. Suppose that the eigenvectors of A span $\mathbb{R}^n$ and have distinct eigenvalues. Then, $AB = BA$ if and only if A and B have the same set of eigenvectors (with possibly different eigenvalues).

If QR Iteration Converges

$$A_\infty = Q_\infty R_\infty = R_\infty Q_\infty$$

(Convergence proof in book.)

Starting Point

$$(A + \delta A)(\vec{x} + \delta\vec{x}) = (\lambda + \delta\lambda)(\vec{x} + \delta\vec{x})$$

Starting Point

$$(A + \delta A)(\vec{x} + \delta\vec{x}) = (\lambda + \delta\lambda)(\vec{x} + \delta\vec{x})$$

What are the independent and dependent variables?

Starting Point

$$(A + \delta A)(\vec{x} + \delta\vec{x}) = (\lambda + \delta\lambda)(\vec{x} + \delta\vec{x})$$

What are the independent and dependent variables?

Approximation:

$$A\delta\vec{x} + \delta A \cdot \vec{x} \approx \lambda\delta\vec{x} + \delta\lambda \cdot \vec{x}$$

Trick: Left Eigenvector

$$A\vec{x} = \lambda\vec{x}, \vec{x} \neq \vec{0} \Rightarrow$$
$$\exists \vec{y} \neq \vec{0} \text{ such that } A^\top \vec{y} = \lambda \vec{y}$$

Change in Eigenvalue

$$|\delta\lambda| \lesssim \frac{\|\delta A\|_2}{|\vec{y}\cdot\vec{x}|}$$

Change in Eigenvalue

$$|\delta\lambda| \lesssim \frac{\|\delta A\|_2}{|\vec{y}\cdot\vec{x}|}$$

What about symmetric A ?