MATH 125 Lecture 22: Part B

Eigenvalue Problems
Iterative Methods

Power iteration

Computing an eigenvalue given an eigenvector

Given an eigenvector, resulting from power iteration, how do we find the eigenvalue?

Consider the following minimization program

$$\min_{\lambda} \|Ax - \lambda x\|_2^2$$

$x \equiv$ approximate eigenvalue

$$\|Ax - \lambda x\|_2^2 = x^T A^T A x - \lambda x^T (A + A^T) x + \lambda^2 x^T x$$
$$= x^T A^T A x - \lambda x^T A x - \lambda x^T A^T x + \lambda^2 x^T x$$

use the following identity:

$$\underbrace{(x^T A x)}_{\text{It is just scalar}} = (x^T A x)^T = x^T A^T x$$

$$\|Ax - \lambda x\|_2^2 = x^T A^T A x - 2\lambda x^T A^T x + \lambda^2 x^T x$$

$$\Downarrow$$

Quadratic in λ, differentiate!

$$2\lambda x^T x - 2x^T A^T x = 0$$

$$\lambda = \frac{x^T A x}{x^T x} \equiv \text{Rayleigh quotient associated with } x$$

Note $Ax \approx \lambda x$ (Not exactly λx)

Theorem Suppose $|\lambda_1| > |\lambda_2| > \dots |\lambda_m| \geq 0$ and $q_1^T v^{(0)} \neq 0$
($q_1 \equiv$ dominant eigenvector)
Then $\|v^k - (\pm q_1)\| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$ $\quad |\lambda^{(k)} - \lambda_1| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right)$

as $k \to \infty$

Remarks

- Linear convergence with factor $\approx \left|\frac{\lambda_2}{\lambda_1}\right|$
- If largest two eigenvalues are close, slow convergence
- only limited to finding dominating eigenvalue
- Eg Fails if λ and $-\lambda$ are largest eigenvalues (magnitude)

Inverse iteration

$A \in \mathbb{R}^{m\times m}$ has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$

$\Rightarrow$ Let $\tilde{\lambda}_j$ be an approximation for λ_j

Now we consider the following matrix:

$$(A - \tilde{\lambda}_j I)^{-1}$$

<u>Question</u> what are eigenvalues of $(A - \tilde{\lambda}_j I)^{-1}$?

<u>Answer</u> Eigenvalues of $A - \tilde{\lambda}_j I$ are

$$\lambda_1 - \tilde{\lambda}_j, \lambda_2 - \tilde{\lambda}_j, \ldots, \lambda_m - \tilde{\lambda}_j$$

Eigenvalues of $(A - \tilde{\lambda}_j I)^{-1}$ are

$$\frac{1}{\lambda_1 - \tilde{\lambda}_j}, \frac{1}{\lambda_2 - \tilde{\lambda}_j}, \ldots, \frac{1}{\lambda_m - \tilde{\lambda}_j}$$

If $\tilde{\lambda}_j$ is a good estimate, power iteration converges fast and yields the dominating eigenvalue:

$$\frac{1}{\lambda_j - \tilde{\lambda}_j}$$

<u>Question</u> what are eigenvectors of $(A - \tilde{\lambda}_j I)^{-1}$?

<u>Answer</u> i) Eigenvectors of $(A - \tilde{\lambda}_j I)^{-1}$ are the same as eigenvectors of $(A - \tilde{\lambda}_j I)$

ii) Eigenvectors of $(A - \tilde{\lambda}_j I)$ are the same as eigenvectors of A

Therefore, eigenvectors of $(A - \tilde{\lambda}_j I)^{-1} \equiv$ eigenvectors of A

. So we find the dominating eigenvalue $\frac{1}{\lambda_j - \tilde{\lambda}_j}$, λ_j and the associated eigenvector (j^{th} eigenvector)

<u>Question</u> How expensive is this? (Rough estimate)

$$z_{j+1} = (A - \tilde{\lambda}_j I)^{-1} x_j \qquad x_{j+1} = \frac{x_j}{\|x_j\|_2}$$

$$\downarrow$$

$$(A - \tilde{\lambda}_j I)\, z_{j+1} = x_j$$

(Solve a linear system, expensive)

<u>Remarks</u> What if μ is an eigenvalue of A ($\mu = \tilde{\lambda}_j$)? $(A - \mu I)$ is singular. what if μ is nearly eigenvalue? (leading to ill conditioning)

$\Rightarrow$ Not trivial but turns out this is ok for inverse iteration

- The convergence depends on the estimate μ
- Standard methods for calculating multiple eigenvectors given estimate of eigenvalues

* Extension to Rayleigh quotient iteration

QR iteration

A and B are similar if

$$B = T^{-1} A T$$

Similar matrices have the same eigenvalues.

$\Rightarrow$ To find eigenvalues of A, use similarity transformation T, and find eigenvalues of B

Cons we have to invert T. Expensive!

Goal Find a good T where the inverse could be readily computed.

$\Rightarrow$ orthogonal matrix : $Q^{-1} = Q^T$

$\Rightarrow$ Q should not be agnostic of A

Two observations

i) $A = QR$

ii) $Q^T A Q$ has the same eigenvalues as A

$$Q^T A Q = Q^T Q R Q = R Q$$

$\Rightarrow$ $A_2 = RQ$ is not equal to A but has same eigenvalues

Factor A_2 as $A_2 = Q_2 R_2$

$$Q_2^T A_2 Q_2 = R_2 Q_2$$

$\Rightarrow$ $A_3 = R_2 Q_2$ is not equal to A_2 but has same eigenvalues

Factor A_3 as $A_3 = Q_3 R_3$

$$Q_3^T A_3 Q_3 = R_3 Q_3$$

$\Rightarrow$ For many matrices A, as $k \to \infty$, A_k converges to a triangular matrix

Basic QR Algorithm

$A_0 = A$

for $k = 1, 2, \ldots$

$$Q^{(k)} R^{(k)} = A^{(k-1)}$$

$$A^{(k)} = R^{(k)} Q^{(k)}$$

$$A^{(k)} = R^{(k)} Q^{(k)}$$

$$Q^{(k)} A^{(k)} = \underline{Q^{(k)} R^{(k)}} Q^{(k)}$$

$$Q^{(k)} A^{(k)} = A^{(k-1)} Q^{(k)}$$

$$A^{(k)} = [Q^{(k)}]^T A^{(k-1)} [Q^{(k)}]$$

Wilkinson Convergence Theorem

Let $A \in \mathbb{R}^{m \times m}$. Assume that different eigenvalues of A have different absolute values. Let V be the matrix whose columns are linearly independent eigenvectors of A. $V = LU$ (has an LU decomposition) [in exact arithmetic]. Then $A^{(j)}$ converges to an upper triangular matrix.

Question When does this fail?

Question Does this always work for symmetric matrices?

Deflation

Hotelling's deflation

consider $(A - \lambda_1 v_1 v_1^T) v_j$ [A symmetric]

$j = 1$ $\quad (A - \lambda_1 v_1 v_1^T) v_1 = \lambda_1 v_1 - \lambda_1 v_1 (v_1^T v_1) = 0 v_1$

$j \neq 1$ $\quad (A - \lambda_1 v_1 v_1^T) v_j = \lambda_j v_j - \lambda_1 v_1 (v_1^T v_j) = \lambda_j v_j$

Householder deflation

$Hx_1 = e_1$ where $Ax_1 = \lambda_1 x_1$; $\|x_1\|_2 = 1$

$$\begin{aligned} HAH^T e_1 &= HAHe_1 \\ &= HAx_1 \\ &= \lambda_1 Hx_1 = \lambda_1 e_1 \end{aligned}$$

First column of HAH^T is $\lambda_1 e_1$

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

Apply