Recitation 7

Recitation Instructor: Shivam Verma
Email: shivamverma@nyu.edu
Ph: 718-362-7836
Office hours: WWH 605 (2.50 - 4.50 pm, Tuesdays)

Brief Overview

- Overview of Jacobi's method for finding eigenvalues
- Solving Jacobi's method in MATLAB
- Finding eigenvectors using Jacobi's method (MATLAB)
- Special properties of symmetric matrices
  - Proof(s) of theorem 5.1
  - Orthogonal diagonalization

Jacobi's method: Overview

Idea: Use **orthogonal transformations** (pre- and post- multiply) to convert **symmetric matrix** to diagonal form.

- Use a plane rotation matrix of the form:

$$R(\varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$$

- Can check that this is an orthogonal matrix.
- Choose ϕ to make (p,q) and (q,p) element zero:

$$\varphi = \frac{1}{2}\tan^{-1}\frac{2a_{pq}}{a_{qq} - a_{pp}}$$

- See [2] for example. Also see sec. 5.2 in [1] for good discussion.

Jacobi's method in MATLAB

Take the symmetric matrix:

A = [3 1 2;
1 2 1;
2 1 4]

We will use the *serialized Jacobi* method in MATLAB to obtain the eigenvalues and eigenvectors of A.

```
B = [3,1,2;1,2,1;2,1,4];
A=B;
tol = 1e-6;
count=0;
Q=eye(3);
while true
    count=count+1;
    normerr = norm(A-diag(diag(A)),2);
    if normerr<tol
        break;
    end
    %zero elem. (2,1)
    theta=0.5*atan(2*A(2,1)/(A(1,1)-A(2,2)));
    Q1 = [cos(theta),-sin(theta),0;sin(theta),cos(theta),0;0,0,1];
    A=Q1'*A*Q1;
    %zero elem. (3,1)
    theta=0.5*atan(2*A(3,1)/(A(1,1)-A(3,3)));
    Q2 = [cos(theta),0,-sin(theta);0,1,0;sin(theta),0,cos(theta)];
    A=Q2'*A*Q2;
    A;
    %zero elem. (3,2)
    theta=0.5*atan(2*A(3,2)/(A(2,2)-A(3,3)));
    Q3 = [1,0,0;0,cos(theta),-sin(theta);0,sin(theta),cos(theta)];
    A=Q3'*A*Q3;
    Q=Q*Q1*Q2*Q3;
    A;
end

[V,D]=eig(B);
%compare A with D to check if eigvalues correct
%compare Q with V to check if eigvectors correct (could be diff. generally)
```

Note: The above method is *not* very efficient! Think about how you can make this more efficient. One idea could be:

- Combining the three different serialized iterations for (2,1), (3,1) and (3,3) into a loop. You would not do the same thing for a 10x10 matrix!
- Try implementing the *classical Jacobi* iteration method, where each iteration involves finding the maximum element and then applying the orthogonal transformations. Compare the time it takes for both of these methods with each other. Which one is more computationally intensive? Why?

Eigenvectors in MATLAB

After applying Jacobi's iteration method, look at the matrix $Q = Q_1 Q_2 ... Q_n$, whose columns contain the eigenvalues. Compare them with the eigenvectors obtained from MATLAB's *eig* function. They may or may not be different, depending on the precise method that MATLAB employs.

Applicability of Jacobi's method

- The Jacobi method usually converges in a reasonable number of iterations, and is a satisfactory method for **small or moderate-sized matrices**.

- Many problems, particularly in the numerical solution of PDEs, **give rise to very large matrices that are sparse**, with most of the elements being zero.

- In many practical situations, we **don't need to compute all the eigenvalues**. More common to require a few of the largest eigenvalues and corresponding eigenvectors, or a few of the smallest.

- Jacobi's method **not suitable for such problems**, as it always produces all the eigenvalues, and will not preserve the sparse structure of a matrix during the course of the iteration. For example, symmetric tridiagonal matrix doesn't remain tridiagonal after one Jacobi iteration.

Properties of symmetric matrices

Theorem 5.1 *Suppose that $A \in \mathbb{R}^{n\times n}_{\text{sym}}$; then, the following statements are valid.*

(i) *There exist n linearly independent eigenvectors $\boldsymbol{x}^{(i)} \in \mathbb{R}^n$ and corresponding eigenvalues $\lambda_i \in \mathbb{R}$ such that $A\boldsymbol{x}^{(i)} = \lambda_i \boldsymbol{x}^{(i)}$ for all $i = 1, 2, \ldots, n$.*

(ii) *The function*

$$\lambda \mapsto \det(A - \lambda I) \tag{5.2}$$

is a polynomial of degree n with leading term $(-1)^n\lambda^n$, called the **characteristic polynomial of** *A. The eigenvalues of A are the zeros of the characteristic polynomial.*

(iii) *If the eigenvalues λ_i and λ_j of A are distinct, then the corresponding eigenvectors $\boldsymbol{x}^{(i)}$ and $\boldsymbol{x}^{(j)}$ are orthogonal in $\mathbb{R}^n$, i.e.,*

$$\boldsymbol{x}^{(i)\mathrm{T}}\boldsymbol{x}^{(j)} = 0 \qquad \text{if } \lambda_i \neq \lambda_j\,, \qquad i, j \in \{1, 2, \ldots, n\}\,.$$

(iv) *If λ_i is a root of multiplicity m of (5.2), then there is a linear subspace in $\mathbb{R}^n$ of dimension m, spanned by m mutually orthogonal eigenvectors associated with the eigenvalue λ_i.*

(v) *Suppose that each of the eigenvectors $\boldsymbol{x}^{(i)}$ of A is* **normalised**, *in other words, $\boldsymbol{x}^{(i)\mathrm{T}}\boldsymbol{x}^{(i)} = 1$ for $i = 1, 2, \ldots, n$, and let X denote the square matrix whose columns are the normalised (orthogonal) eigenvectors; then, the matrix $\Lambda = X^{\mathrm{T}}AX$ is diagonal, and the diagonal elements of Λ are the eigenvalues of A.*

(vi) *Let $Q \in \mathbb{R}^{n\times n}$ be an orthogonal matrix and define $B \in \mathbb{R}^{n\times n}_{\text{sym}}$ by $B = Q^{\mathrm{T}}AQ$; then, $\det(B - \lambda I) = \det(A - \lambda I)$ for each $\lambda \in \mathbb{R}$. The eigenvalues of B are the same as the eigenvalues of A, and the eigenvectors of B are the vectors $Q^{\mathrm{T}}\boldsymbol{x}^{(i)}$, $i = 1, 2, \ldots, n$.*

(vii) *Any vector $\boldsymbol{v} \in \mathbb{R}^n$ can be expressed as a linear combination of the (ortho)normalised eigenvectors $\boldsymbol{x}^{(i)}$, $i = 1, 2, \ldots, n$, of A, i.e.,*

$$\boldsymbol{v} = \sum_{i=1}^{n} \alpha_i \boldsymbol{x}^{(i)}, \qquad \alpha_i = \boldsymbol{x}^{(i)\mathrm{T}}\boldsymbol{v}\,.$$

(viii) *The trace of A, $\mathrm{Trace}(A) = \sum_{i=1}^{n} a_{ii}$, is equal to the sum of the eigenvalues of A.*

[Image source: Chap. 5, Introduction to Numerical Analysis, E. Suli & D. Mayers]

Summary of important properties of symmetric matrices (from Th. 5.1):

- **Eigenvalues are real**. How?
  - Think about the characteristic polynomial of a 2x2 matrix. Can you prove that its roots are real?
  - General proof:
    - $\lambda_1 < q_1, q_1 > \; = < \lambda_1 q_1, q_1 > \; = \; < Aq_1, q_1 > \; = \; < q_1, Aq_1 > \; = \lambda_1^* < q_1, q_1 >$
    - $\Rightarrow \lambda_1 = \lambda_1^*$

- **Eigenvectors are linearly dependent, and orthogonal for distinct eigenvalues.**
  - Orthogonal:
    - Let $\lambda_1\ (q_1)$ and $\lambda_2\ (q_2)$ be two *distinct* eigenvalues (and respective eigenvectors) of A.
    - $\lambda_1 q_2{}^T q_1 = q_2{}^T \lambda_1 q_1 = q_2{}^T A q_1 = (Aq_2)^T q_1 = \lambda_2 q_2{}^T q_1 \Rightarrow \lambda_1 = \lambda_2$
    - Thus, it must be $q_2{}^T q_1 = 0 \Rightarrow q_1 \textit{ and } q_2 \textit{ are orthogonal}$
    - If λ_1 is an eigenvalue with *multiplicity > 1*, can construct orthonormal eigenvectors using Gram-Schmidt process.
    - Linearly independent: Easily proved since has *n* orthogonal eigenvectors, which form a basis.

- They can always **be diagonalized by orthogonal matrices**.
  - Let A have all distinct eigenvalues.
  - $D = Q^T A Q = [q_1{}^T A q_1 \; q_1{}^T A q_2 \ldots q_1{}^T A q_n; \; \ldots \ldots \ldots; \; q_n{}^T A q_1 \; q_n{}^T A q_2 \ldots q_n{}^T A q_n]$
  - Using previous property, $q_1{}^T A q_2 = (Aq_1)^T q_2 = \lambda_1 q_1{}^T q_2 = 0$
  - Thus, A is diagonalized by Q.
  - If an eigenvalue λ has multiplicity *m*, then we can always find a set of *m* orthonormal eigenvectors for λ using Gram-schmidt process.
    - $Aq_1 = \lambda_1 q_1, \; Aq_2 = \lambda_1 q_2 \Rightarrow A(q_1 + q_2) = \lambda_1(q_1 + q_2)$
    - $q_1{}' = q_1, \; q_2{}' = q_2 - (q_1{}^T q_2 / q_1{}^T q_1) \cdot q_1 \Rightarrow q_1{}'^T q_2{}' = 0$
    -

- **Orthogonal matrix** which diagonalizes a given symmetric matrix has that matrix's **eigenvectors as its columns**.
  - Sanity check as follows:
  - Open MATLAB and define a random symmetric matrix
    - B = rand(3)
    - B = B + B'

- Find out eigenvectors for B
  - [V,D] = eig(B)
- Try: V'*B*V
- You should get a matrix similar to D, i.e. diagonal matrix with eigenvalues.

- Eigenvalues of A are same as those of D = Q'AQ, where Q is an orthogonal matrix.
  - $Q^T Q = QQ^T = I,\ \lambda_i = q_i^T A q_i = q_i^T QQ^T AQQ^T q_i = (Qq_i)^T Q^T AQ(Qq_i)$
  - $Qq_i = \{q_i^T q_j\} = \{\delta_{ij}\} = e_i$
  - $\lambda_i = e_i^T diag(\mu_1, ..., \mu_n) e_i = \mu_i$

Helpful links

1. See p. 1 & 2 on bad eigenvalue problems
2. See sec. 2 on Jacobi method
3. More about Jacobi's method
4. Quick overview of linear algebra and relevant numerical algorithms
5. Jacobi convergence and eigenvalue problem examples
6. Eigenvalues of symmetric matrices