Lecture XII
Numerical Issues in Matrix Theory

- Perturbation theory
- Computational methods for matrices

Sensitivity Analysis

Question:

How will the eigenvalues and eigenvectors of a matrix A change, if A is perturbed into $A+\delta A,$ with δA being small?

Problem Statement

Given $\left(A,\ \delta A,\ u^i,\ \lambda_i\right)$, find the perturbation $\delta\lambda^i$ of the eigenvalue λ^i, and δu^i, such that

$$\left(A+\delta A\right)\left(u^i+\delta u^i\right)=\left(\lambda_i+\delta\lambda^i\right)\left(u_i+\delta u^i\right).$$

Unknown

$Note:\ \ \delta u^i=\varepsilon \bullet u^i$ remains to be a solution of the above eq. if $\delta A=\varepsilon A$, for small ε.

Standing Assumptions

- The eigenvalues λ_i of A are distinct, associated with (linearly independent) eigenvectors u^i.
- The perturbed eigenvectors $u^i + \delta u^i$ of $A + \delta A$ are normalized in the sense that

$$u^i + \delta u^i = \sum_{k=1}^{n} c_{ik} u^k, \quad c_{ii} = 1$$

so, $\delta u^i = \sum_{k=1}^{n} \varepsilon_{ik} u^k, \quad \varepsilon_{ii} = 0, \varepsilon_{ik}, i \neq k,$ unknown.

(This would guarantee $\delta u^i = 0$ if $\delta A = 0$.)

Principle of Biorthogonality

Consider $A \in \mathbb{C}^{n\times n}$ that has distinct eigenvalues $\{\lambda_i\}_{i=1}^{n}$ with associated eigenvectors $\{u^i\}_{i=1}^{n}$. Let $\{\overline{v}^i\}_{i=1}^{n}$ be eigenvectors associated with $\{\overline{\lambda}_i\}_{i=1}^{n}$ of $A^* \doteq \overline{A}^T$.

Then,

$$\left\langle u^i, v^i \right\rangle = (u^i)^T \overline{v}^i \neq 0, \quad \left\langle u^i, v^k \right\rangle = (u^i)^T \overline{v}^k = 0, \ \forall i \neq k.$$

Proof

First, note that

$$\det\left(\bar{\lambda}_i I - A^*\right) = \det\left(\bar{\lambda}_i I - \bar{A}^T\right)$$

$$= \overbrace{\det\left(\lambda_i I - A^T\right)}^{-} = \overbrace{\det\left(\lambda_i I - A\right)^T}^{-} = 0$$

confirming the fact that $\left\{\bar{\lambda}_i\right\}$ are eigenvalues of A^*.

Then, $\forall i \neq k,\ Au^i = \lambda_i u^i$ and $A^* v^i = \bar{\lambda}_i v^i$.

Proof (cont'd)

Clearly, $\langle Au^i, v^k \rangle = \langle \lambda_i u^i, v^k \rangle$, $\langle u^i, A^* v^k \rangle = \langle u^i, \bar{\lambda}_k v^k \rangle$.

Using $\langle Au, v \rangle = \langle u, A^* v \rangle \ \forall u, v$, it follows

$\langle \lambda_i u^i, v^k \rangle = \langle u^i, \bar{\lambda}_k v^k \rangle$ *or* equivalently,

$\lambda_i \langle u^i, v^k \rangle = \lambda_k \langle u^i, v^k \rangle$

$\Rightarrow \langle u^i, v^k \rangle = 0, \ \forall i \neq k$ because $\lambda_i \neq \lambda_k$.

Proof (cont'd)

To prove that $\langle u^i, v^i \rangle \neq 0, \ \forall i,$ it suffices to note that

$$u^i = \sum_{k=1}^{n} \alpha_{ik} v^k, \text{ with } \{v^k\}_{k=1}^{n} \text{ mutually orthogonal}$$

By contradiction, assume that $\langle u^i, v^i \rangle = 0.$

$$Then, \ \langle u^i, u^i \rangle = \sum_{k=1}^{n} \langle u^i, \alpha_{ik} v^k \rangle = \sum_{k=1}^{n} \bar{\alpha}_{ik} \langle u^i, v^k \rangle = 0$$

$\Rightarrow u^i = 0,$ contradiction with u^i being an eigenvector.

Back to our Problem:

Computation of $\delta\lambda_i$, ε_{ik} ?

Find the perturbation $\delta\lambda_i$ of the eigenvalue λ_i and the unknown ε_{ik}, $i \neq k$, such that

$$(A+\delta A)\left(u^i+\delta u^i\right)=\left(\lambda_i+\delta\lambda_i\right)\left(u^i+\delta u^i\right)$$

with

$$\delta u^i=\sum_k \varepsilon_{ik} u^k .$$

Detailed Solution

Ignoring the (smaller) second-order terms $\delta A \delta u^i$ and $\delta\lambda^i \delta u^i$ in

$$\left(A+\delta A\right)\left(u^i+\delta u^i\right)=\left(\lambda_i+\delta\lambda_i\right)\left(u^i+\delta u^i\right),$$

we have

$$A\delta u^i+(\delta A)u^i=\lambda_i\delta u^i+\left(\delta\lambda_i\right)u^i$$

Our second Assumption, i.e. $\delta u^i=\sum_{k=1}^{n}\varepsilon_{ik}u^k$, $(\varepsilon_{ii}=0)$

$\Rightarrow$ $\left\langle \delta u^i, v^i \right\rangle = 0$ using Principle of Biorthogonality

Detailed Solution (cont'd)

$$\left\langle A\delta u^i, v^i \right\rangle + \left\langle (\delta A)u^i, v^i \right\rangle = \left\langle \lambda_i \delta u^i, v^i \right\rangle + \left\langle \left(\delta\lambda_i\right) u^i, v^i \right\rangle$$

$$or,\ 0 + \left\langle (\delta A)u^i, v^i \right\rangle = 0 + \left(\delta\lambda_i\right)\left\langle u^i, v^i \right\rangle$$

$$\text{noting } \left\langle A\delta u^i, v^i \right\rangle = \left\langle \delta u^i, A^* v^i \right\rangle = \left\langle \delta u^i, \overline{\lambda}_i v^i \right\rangle = 0.$$

Therefore,

$$\delta\lambda_i = \frac{\left\langle (\delta A)u^i, v^i \right\rangle}{\left\langle u^i, v^i \right\rangle}$$

Detailed Solution (cont'd)

$$A\delta u^i + (\delta A)u^i = \lambda_i \delta u^i + \left(\delta\lambda_i\right)u^i$$

$$\Rightarrow \quad \forall k \neq i,$$

$$\left\langle A\delta u^i, v^k \right\rangle + \left\langle (\delta A)u^i, v^k \right\rangle = \left\langle \lambda_i \delta u^i, v^k \right\rangle + 0.$$

For the same reason,

$$\left\langle A\delta u^i, v^k \right\rangle = \left\langle \delta u^i, A^* v^k \right\rangle = \left\langle \delta u^i, \bar{\lambda}_k v^k \right\rangle = \lambda_k \left\langle \delta u^i, v^k \right\rangle$$

$$= \lambda_k \left\langle \sum \varepsilon_{ij} u^j, v^k \right\rangle = \lambda_k \varepsilon_{ik} \left\langle u^k, v^k \right\rangle.$$

So, $\lambda_k \varepsilon_{ik} \left\langle u^k, v^k \right\rangle + \left\langle (\delta A)u^i, v^k \right\rangle = \left\langle \lambda_i \delta u^i, v^k \right\rangle.$

Detailed Solution (cont'd)

$$\lambda_k \varepsilon_{ik} \left\langle u^k, v^k \right\rangle + \left\langle (\delta A) u^i, v^k \right\rangle = \left\langle \lambda_i \delta u^i, v^k \right\rangle$$

together with $\delta u^i = \sum_k \varepsilon_{ik} u^k$,

implies

$$\varepsilon_{ik} = \frac{\left\langle (\delta A) u^i, v^k \right\rangle}{(\lambda_i - \lambda_k) \left\langle u^k, v^k \right\rangle}, \quad \forall i \neq k.$$

An Example

Consider $A = diag(\lambda_i)$, with $\lambda_i \neq \lambda_j$, $\forall i \neq j$.

For the pertuabtion $\delta A = \varepsilon B$, with $B = (b_{ij})$

and ε a small scalar.

In this case, the chosen eigenvectors of A, A^*

are

$$u^i = v^i = col(0, \cdots, 1, \cdots, 0) \doteq e^i.$$

By our formulae, $\delta\lambda_i = \left\langle \varepsilon B e^i, e^i \right\rangle = \varepsilon b_{ii}$,

$$\varepsilon_{ik} = \frac{\left\langle \varepsilon B e^i, e^k \right\rangle}{\lambda_i - \lambda_k} = \frac{\varepsilon b_{ki}}{\lambda_i - \lambda_k}, \quad \forall i \neq k.$$

In other words,

$$\delta u^i = \varepsilon \sum_{\substack{k \neq i \\ k=1}}^{n} b_{ki} \left(\lambda_i - \lambda_k \right)^{-1} e^k$$

where $e^k = col\left(0, \cdots, 1, \cdots 0\right)$ with "1" as the k-th element.

Computational Methods

for solving inhomogeneous equations

- Gaussian elimination method
- LR factorization method
- Gauss-Seidel iterative method

Linear Inhomogeneous Equations

As seen previously, many problems (such as computing eigenvalues and eigenvectors) reduce down to solving for unknown x :

$$Ax = b, \quad A = \left(a_{ij}\right)_{n\times n}, \quad b = \left(b_i\right)_{n\times 1}.$$

Gaussian Elimination Method

Consider the following example

$$
\begin{cases}
4x_2 - x_3 = 5, \\
x_1 + x_2 + x_3 = 6, \\
2x_1 - 2x_2 + x_3 = 1.
\end{cases}
$$

Problem:

Solve for unknown $x = col\left(x_1, x_2, x_3\right)$.

Systematic Procedure

Step 1: Set up the augmented matrix

$$[A,\ b] = \begin{pmatrix} 0 & 4 & -1 & 5 \\ 1 & 1 & 1 & 6 \\ 2 & -2 & 1 & 1 \end{pmatrix}$$

Step 2: Interchange the first two rows:

$$\begin{pmatrix} 1 & 1 & 1 & 6 \\ 0 & 4 & -1 & 5 \\ 2 & -2 & 1 & 1 \end{pmatrix}$$

Systematic Procedure

Step 3 : Subtract twice the first row from the last:

$$\begin{pmatrix} 1 & 1 & 1 & 6 \\ 0 & 4 & -1 & 5 \\ 0 & -4 & -1 & -11 \end{pmatrix}$$

and then, add the second row to the last row:

$$\begin{pmatrix} 1 & 1 & 1 & 6 \\ 0 & 4 & -1 & 5 \\ 0 & 0 & -2 & -6 \end{pmatrix},$$

where A has become a upper-triangular matrix.

Systematic Procedure

Step 4: From the special form of

$$\begin{pmatrix} 1 & 1 & 1 & 6 \\ 0 & 4 & -1 & 5 \\ 0 & 0 & -2 & -6 \end{pmatrix},$$

$(bottom-up)$ we can read out the solutions:

$$x_3 = 3,\ x_2 = \frac{5+x_3}{4} = 2,$$

$$x_1 = 6 - x_2 - x_3 = 1.$$

Comment 1

The elimination method only involves algebraic operations to the rows (!).

It is also useful for solving the standard linear programming (LP) problem:

$$\min_{x} P = \sum_{i=1}^{n} c_i x_i$$

subject to: $Ax = b, \quad x \geq 0.$

Comment 2

Consider a general equation of the form $AX = B$,

where $\det A \neq 0, \;\; A: n\times n, \;\; B: n\times k.$

The elimination method solves the equation after the following nos. of algebraic operations:

$$\mu_n = n^2k + \frac{1}{3}(n-1)n(n+1), \text{ multiplications/divisions}$$

$$\alpha_n = n(n-1)k + \frac{1}{6}(n-1)n(2n-1),$$

additions/subtractions.

Comment 2 (cont'd)

If we use Cramer's Rule to solve $AX = B$,
where $\det A \neq 0$, $A : n \times n$, $B : n \times k$,
the computational complexity is of the order
$k \bullet (n!)$.

LR factorization method

It consists of decomposing the matrix A into a left-triangular matrix L with 1's on the diagonal, and a right-triangular matrix R with nonzero diagonal elements $\{r_{ii}\}$:

$$L = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ * & \cdots & 1 \end{pmatrix}, \qquad R = \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & r_{nn} \end{pmatrix}$$

LR factorization method

That is,

$$A = LR.$$

Then, the equation $Ax = b$ becomes two more easily solvable linear equations:

$$Lc = b$$

and

$$Rx = c.$$

An illustrative example

Consider the linear equation $Ax = col\left(1,3\right)$

$$with \quad A = \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}.$$

First, decompose A into the form LR, i.e.,

$$\begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ l_{21} & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}$$

An illustrative example

Simple computation leads to

$$l_{21} = -\frac{1}{3},\ r_{11} = 3,\ r_{22} = \frac{5}{3},\ r_{12} = -1,\ i.e.,$$

$$\begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{3} & 1 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ 0 & \frac{5}{3} \end{pmatrix}$$

An illustrative example

Now, solve for c:

$$\begin{pmatrix} 1 & 0 \\ -\frac{1}{3} & 1 \end{pmatrix} c = b = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \Rightarrow c = \begin{bmatrix} 1 \\ \frac{10}{3} \end{bmatrix}$$

and then solve for x:

$$\begin{pmatrix} 3 & -1 \\ 0 & \frac{5}{3} \end{pmatrix} x = c = \begin{bmatrix} 1 \\ \frac{10}{3} \end{bmatrix} \Rightarrow x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Comment

The necessary and sufficient condition for LR decomposition is that all leading principal minors of A are nonzero.

Gauss-Seidel iterative method

Carl F. Gauss, 1777-1855

The main idea is to construct a sequence $\left\{x^i\right\}_{i=0}^{\infty}$,
defined by a recursive relation, with initial guess x^0,
such that $x^i \to x$, solution of $Ax = b$.

Illustration of the Original Idea

For the purpose of illustration, consider $Ax = b$

$with\ A = \left(a_{ij}\right)_{3\times 3}$, all diagonal $a_{ii} \neq 0.$

Then, $Ax = b$ implies

$$\begin{cases} x_1 = a_{11}^{-1}\left(b_1 - a_{12}x_2 - a_{13}x_3\right) \\ x_2 = a_{22}^{-1}\left(b_2 - a_{21}x_1 - a_{23}x_3\right) \\ x_3 = a_{33}^{-1}\left(b_3 - a_{31}x_1 - a_{32}x_2\right) \end{cases}$$

Illustration of the Original Idea

If $x^k = col\left(x_j^k\right)$ is an estimate at Step k, then a good guess $x^{k+1} = col\left(x_j^{k+1}\right)$ at Step $k+1$ should be:

$$\begin{cases} x_1^{k+1} = a_{11}^{-1}\left(b_1 - a_{12}x_2^k - a_{12}x_3^k\right) \\ x_2^{k+1} = a_{22}^{-1}\left(b_2 - a_{21}x_1^k - a_{23}x_3^k\right) \\ x_3^{k+1} = a_{33}^{-1}\left(b_3 - a_{31}x_1^k - a_{32}x_2^k\right) \end{cases}$$

Illustration of the Original Idea

In fact, Gauss-Seidel proved that such a sequence

$\left\{x^k\right\}_{k=0}^{\infty}$ converges to $x = A^{-1}b \;\forall x^0,$ if (and only if)

all roots λ of the equation

$$\det\begin{pmatrix} \lambda a_{11} & a_{12} & a_{13} \\ \lambda a_{21} & \lambda a_{22} & a_{23} \\ \lambda a_{31} & \lambda a_{32} & \lambda a_{33} \end{pmatrix} = 0$$

are inside the unit disk, i.e., $|\lambda| < 1.$

General Iterative Algorithm

$$x_i^{k+1} = a_{ii}^{-1}\left(b_i - \sum_{j<i} a_{ij} x_j^k - \sum_{j>i} a_{ij} x_j^k\right)$$

where

$$x^k = col\left(x_j^k\right),\ x^{k+1} = col\left(x_j^{k+1}\right).$$

Exercises

1. Solve $AX = B$, if

$$\begin{bmatrix} A & B \end{bmatrix} = \begin{pmatrix} 0 & 2 & -1 & \vdots & 1 & -3 \\ -3 & 1 & 4 & \vdots & 2 & 27 \\ 1 & 6 & -5 & \vdots & 2 & -22 \end{pmatrix}.$$

2. Apply the LR factorization method to solve

$$\begin{pmatrix} 1 & 0 & 7 \\ 3 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix} x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Answers

$$1) \qquad X = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 7 \end{pmatrix}$$

$$2) \qquad x = \begin{pmatrix} -26/9 \\ 16/3 \\ 5/9 \end{pmatrix}$$

More Exercises
for Previous Lectures

1. Let A be a lower-triangular matrix with nonzero diagonal elements. Is A^{-1} a triangular matrix?

2. Give a simple close-form expression for

$$\det \begin{pmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{pmatrix}$$

More Exercises

3. Let A be an $m \times n$ matrix. Show that $Ax = 0$ has a nontrivial solution $x \neq 0$ if and only if the columns are linearly dependent.

4. Does the following equation have a solution:

$$\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} x = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

How about $\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} x = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$? If yes, general solutions?

More Exercises

5. Solve the coupled difference equations:

$$\begin{cases} u_{k+1} = -7u_k + 4v_k \\ v_{k+1} = -8u_k + v_k \end{cases}$$

with initial values $u_0 = 1,\ v_0 = 2.$

(Hint: use the theory of canonical forms.)

6. If A is similar to B (i.e., $B = P^{-1}AP$ for some nonsingular P), and if B is similar to C, then A is similar to C.

More Exercises

7. Can you bring the following matrix into a Jordan form

$$A = \begin{pmatrix} 17 & 0 & -25 \\ 0 & 3 & 0 \\ 9 & 0 & -13 \end{pmatrix}$$

8. For the above matrix A, solve the differential equation $\dot{x} = Ax$, with initial value $x^0 \in \mathbb{R}^3$.