Simultaneous Linear Equations

GongYi龚怡

2167570874@qq.com

Link

Slides：

https://pan.baidu.com/s/1uqWdtLXJhJ9TvKN1CsK_9w

Access code ：math

Software:

https://pan.baidu.com/s/1iXhXryPJG-YNYF-RedTZ1Q

Access code : 57fs

LU Decomposition

LU Decomposition

LU Decomposition is another method to solve a set of simultaneous linear equations

Which is better, Gauss Elimination or LU Decomposition?

To answer this, a closer look at LU decomposition is needed.

LU Decomposition

Method

For most non-singular matrix $[A]$ that one could conduct Naïve Gauss Elimination forward elimination steps, one can always write it as

$$[A] = [L][U]$$

where

$[L]$ = lower triangular matrix

$[U]$ = upper triangular matrix

How does LU Decomposition work?

If solving a set of linear equations	$[A][X] = [C]$
If $[A] = [L][U]$ then	$[L][U][X] = [C]$
Multiply by	$[L]^{-1}$
Which gives	$[L]^{-1}[L][U][X] = [L]^{-1}[C]$
Remember $[L]^{-1}[L] = [I]$ which	$[I][U][X] = [L]^{-1}[C]$
leads to	$[U][X] = [L]^{-1}[C]$
Now, if $[I][U] = [U]$ then	$[L]^{-1}[C]=[Z]$
Now, let	$[L][Z] = [C] \quad (1)$
Which ends with	$[U][X] = [Z] \quad (2)$
and	

LU Decomposition

How can this be used?

Given $[A][X] = [C]$

1. Decompose $[A]$ into $[L]$ and $[U]$
2. Solve $[L][Z] = [C]$ for $[Z]$
3. Solve $[U][X] = [Z]$ for $[X]$

Is LU Decomposition better than Gaussian Elimination?

$$\text{Solve } [A][X] = [B]$$

T = clock cycle time and n x n = size of the matrix

Forward Elimination

$$CT|_{FE} = T\left(\frac{8n^3}{3} + 8n^2 - \frac{32n}{3}\right)$$

Back Substitution

$$CT|_{BS} = T\left(4n^2 + 12n\right)$$

Decomposition to LU

$$CT|_{DE} = T\left(\frac{8n^3}{3} + 4n^2 - \frac{20n}{3}\right)$$

Forward Substitution

$$CT|_{FS} = T\left(4n^2 - 4n\right)$$

Back Substitution

$$CT|_{BS} = T\left(4n^2 + 12n\right)$$

Is LU Decomposition better than Gaussian Elimination?

To solve $[A][X] = [B]$

Time taken by methods

Gaussian Elimination	LU Decomposition
$T\left(\frac{8n^3}{3}+12n^2+\frac{4n}{3}\right)$	$T\left(\frac{8n^3}{3}+12n^2+\frac{4n}{3}\right)$

T = clock cycle time and nxn = size of the matrix

So both methods are equally efficient.

To find inverse of [A]

Time taken by Gaussian Elimination

$$= n\left(CT|_{FE} + CT|_{BS}\right)$$

$$= T\left(\frac{8n^4}{3} + 12n^3 + \frac{4n^2}{3}\right)$$

Time taken by LU Decomposition

$$= CT|_{DE} + n \times CT|_{FS} + n \times CT|_{BS}$$

$$= T\left(\frac{32n^3}{3} + 12n^2 - \frac{20n}{3}\right)$$

To find inverse of [A]

<u>Time taken by Gaussian Elimination</u>

$$T\left(\frac{8n^4}{3}+12n^3+\frac{4n^2}{3}\right)$$

<u>Time taken by LU Decomposition</u>

$$T\left(\frac{32n^3}{3}+12n^2-\frac{20n}{3}\right)$$

Table 1 Comparing computational times of finding inverse of a matrix using LU decomposition and Gaussian elimination.

n	10	100	1000	10000
$CT\|_{\text{inverse GE}} / CT\|_{\text{inverse LU}}$	3.288	25.84	250.8	2501

For large n, $CT|_{\text{inverse GE}} / CT|_{\text{inverse LU}} \approx n/4$

Method: [A] Decomposes to [L] and [U]

$$[A]=[L][U]=\begin{bmatrix}1 & 0 & 0\\ \ell_{21} & 1 & 0\\ \ell_{31} & \ell_{32} & 1\end{bmatrix}\begin{bmatrix}u_{11} & u_{12} & u_{13}\\ 0 & u_{22} & u_{23}\\ 0 & 0 & u_{33}\end{bmatrix}$$

[*U*] is the same as the coefficient matrix at the end of the forward elimination step.

[*L*] is obtained using the *multipliers* that were used in the forward elimination process

Finding the [U] matrix

Using the Forward Elimination Procedure of Gauss Elimination

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

Step 1: $$\frac{64}{25} = 2.56; \quad Row2 - Row1(2.56) = \begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 144 & 12 & 1 \end{bmatrix}$$

$$\frac{144}{25} = 5.76; \quad Row3 - Row1(5.76) = \begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & -16.8 & -4.76 \end{bmatrix}$$

Finding the [U] Matrix

Matrix after Step 1: $\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & -16.8 & -4.76 \end{bmatrix}$

Step 2: $\dfrac{-16.8}{-4.8} = 3.5; \quad Row3 - Row2(3.5) = \begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix}$

$$[U] = \begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix}$$

Finding the [L] matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix}$$

Using the multipliers used during the Forward Elimination Procedure

From the first step of forward elimination

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

$$\ell_{21} = \frac{a_{21}}{a_{11}} = \frac{64}{25} = 2.56$$

$$\ell_{31} = \frac{a_{31}}{a_{11}} = \frac{144}{25} = 5.76$$

Finding the [L] Matrix

From the second step of forward elimination

$$\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & -16.8 & -4.76 \end{bmatrix} \qquad \ell_{32} = \frac{a_{32}}{a_{22}} = \frac{-16.8}{-4.8} = 3.5$$

$$[L] = \begin{bmatrix} 1 & 0 & 0 \\ 2.56 & 1 & 0 \\ 5.76 & 3.5 & 1 \end{bmatrix}$$

Does [L][U] = [A]?

$$[L][U] = \begin{bmatrix} 1 & 0 & 0 \\ 2.56 & 1 & 0 \\ 5.76 & 3.5 & 1 \end{bmatrix} \begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix} = \ ?$$

Using LU Decomposition to solve SLEs

Solve the following set of linear equations using LU Decomposition

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

Using the procedure for finding the [L] and [U] matrices

$$[A] = [L][U] = \begin{bmatrix} 1 & 0 & 0 \\ 2.56 & 1 & 0 \\ 5.76 & 3.5 & 1 \end{bmatrix} \begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix}$$

Example

Set $[L][Z] = [C]$

$$\begin{bmatrix} 1 & 0 & 0 \\ 2.56 & 1 & 0 \\ 5.76 & 3.5 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

Solve for $[Z]$

$$z_1 = 10$$

$$2.56 z_1 + z_2 = 177.2$$

$$5.76 z_1 + 3.5 z_2 + z_3 = 279.2$$

Example

Complete the forward substitution to solve for [Z]

$$
\begin{aligned}
z_1 &= 106.8 \\
z_2 &= 177.2 - 2.56 z_1 \\
&= 177.2 - 2.56(106.8) \\
&= -96.2 \\
z_3 &= 279.2 - 5.76 z_1 - 3.5 z_2 \\
&= 279.2 - 5.76(106.8) - 3.5(-96.21) \\
&= 0.735
\end{aligned}
$$

$$
[Z] = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ -96.21 \\ 0.735 \end{bmatrix}
$$

Example

Set $[U][X] = [Z]$

$$\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ -96.21 \\ 0.735 \end{bmatrix}$$

Solve for $[X]$

The 3 equations become

$$25a_1 + 5a_2 + a_3 = 106.8$$

$$-4.8a_2 - 1.56a_3 = -96.21$$

$$0.7a_3 = 0.735$$

Example

From the 3rd equation

$$0.7a_3 = 0.735$$

$$a_3 = \frac{0.735}{0.7}$$

$$a_3 = 1.050$$

Substituting in a_3 and using the second equation

$$-4.8a_2 - 1.56a_3 = -96.21$$

$$a_2 = \frac{-96.21 + 1.56a_3}{-4.8}$$

$$a_2 = \frac{-96.21 + 1.56(1.050)}{-4.8}$$

$$a_2 = 19.70$$

Example

Substituting in a_3 and a_2 using the first equation

$$25a_1 + 5a_2 + a_3 = 106.8$$

$$\begin{aligned} a_1 &= \frac{106.8 - 5a_2 - a_3}{25} \\ &= \frac{106.8 - 5(19.70) - 1.050}{25} \\ &= 0.2900 \end{aligned}$$

Hence the Solution Vector is:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0.2900 \\ 19.70 \\ 1.050 \end{bmatrix}$$

Finding the inverse of a square matrix

The inverse [B] of a square matrix [A] is defined as

$$[A][B] = [I] = [B][A]$$

Finding the inverse of a square matrix

How can LU Decomposition be used to find the inverse?

Assume the first column of $[B]$ to be $[b_{11}\ \ b_{12}\ \ \ldots\ b_{n1}]^T$

Using this and the definition of matrix multiplication

First column of $[B]$

$$[A]\begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Second column of $[B]$

$$[A]\begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

The remaining columns in $[B]$ can be found in the same manner

Example: Inverse of a Matrix

Find the inverse of a square matrix $[A]$

$$[A]=\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

Using the decomposition procedure, the $[L]$ and $[U]$ matrices are found to be

$$[A]=[L][U]=\begin{bmatrix} 1 & 0 & 0 \\ 2.56 & 1 & 0 \\ 5.76 & 3.5 & 1 \end{bmatrix}\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix}$$

Example: Inverse of a Matrix

Solving for the each column of [*B*] requires two steps

1) Solve [*L*] [*Z*] = [*C*] for [*Z*]

2) Solve [*U*] [*X*] = [*Z*] for [*X*]

Step 1: $$[L][Z]=[C]\rightarrow\begin{bmatrix}1 & 0 & 0\\2.56 & 1 & 0\\5.76 & 3.5 & 1\end{bmatrix}\begin{bmatrix}z_1\\z_2\\z_3\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}$$

This generates the equations:

$$z_1 = 1$$

$$2.56z_1 + z_2 = 0$$

$$5.76z_1 + 3.5z_2 + z_3 = 0$$

Example: Inverse of a Matrix

Solving for $[Z]$

$$
\begin{aligned}
z_1 &= 1 \\
z_2 &= 0 - 2.56 z_1 \\
&= 0 - 2.56(1) \\
&= -2.56 \\
z_3 &= 0 - 5.76 z_1 - 3.5 z_2 \\
&= 0 - 5.76(1) - 3.5(-2.56) \\
&= 3.2
\end{aligned}
$$

$$
[Z] = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -2.56 \\ 3.2 \end{bmatrix}
$$

Example: Inverse of a Matrix

Solving $[U][X] = [Z]$ for $[X]$

$$\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ -2.56 \\ 3.2 \end{bmatrix}$$

$$\begin{aligned} 25b_{11} + 5b_{21} + b_{31} &= 1 \\ -4.8b_{21} - 1.56b_{31} &= -2.56 \\ 0.7b_{31} &= 3.2 \end{aligned}$$

Example: Inverse of a Matrix

Using Backward Substitution

$$b_{31} = \frac{3.2}{0.7} = 4.571$$

$$b_{21} = \frac{-2.56 + 1.560 b_{31}}{-4.8}$$

$$= \frac{-2.56 + 1.560(4.571)}{-4.8} = -0.9524$$

$$b_{11} = \frac{1 - 5b_{21} - b_{31}}{25}$$

$$= \frac{1 - 5(-0.9524) - 4.571}{25} = 0.04762$$

So the first column of the inverse of [A] is:

$$\begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} = \begin{bmatrix} 0.04762 \\ -0.9524 \\ 4.571 \end{bmatrix}$$

Example: Inverse of a Matrix

Repeating for the second and third columns of the inverse

Second Column

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} b_{12} \\ b_{22} \\ b_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} b_{12} \\ b_{22} \\ b_{32} \end{bmatrix} = \begin{bmatrix} -0.08333 \\ 1.417 \\ -5.000 \end{bmatrix}$$

Third Column

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} b_{13} \\ b_{23} \\ b_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} b_{13} \\ b_{23} \\ b_{33} \end{bmatrix} = \begin{bmatrix} 0.03571 \\ -0.4643 \\ 1.429 \end{bmatrix}$$

Example: Inverse of a Matrix

The inverse of $[A]$ is

$$[A]^{-1} = \begin{bmatrix} 0.04762 & -0.08333 & 0.03571 \\ -0.9524 & 1.417 & -0.4643 \\ 4.571 & -5.000 & 1.429 \end{bmatrix}$$

To check your work do the following operation

$$[A][A]^{-1} = [I] = [A]^{-1}[A]$$

Gauss-Seidel Method

Gauss-Seidel Method

An *iterative* method.

Basic Procedure:

-Algebraically solve each linear equation for x_i

-Assume an initial guess solution array

-Solve for each x_i and repeat

-Use absolute relative approximate error after each iteration to check if error is within a pre-specified tolerance.

Gauss-Seidel Method

Why?

The Gauss-Seidel Method allows the user to control round-off error.

Elimination methods such as Gaussian Elimination and LU Decomposition are prone to prone to round-off error.

Also: If the physics of the problem are understood, a close initial guess can be made, decreasing the number of iterations needed.

Gauss-Seidel Method

Algorithm

A set of *n* equations and *n* unknowns:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n = b_2$$

$$\vdots \qquad \vdots$$

$$a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots + a_{nn}x_n = b_n$$

If: the diagonal elements are non-zero

Rewrite each equation solving for the corresponding unknown

ex:

First equation, solve for x_1

Second equation, solve for x_2

Gauss-Seidel Method

Algorithm

Rewriting each equation

$$x_1 = \frac{c_1 - a_{12}x_2 - a_{13}x_3 \ldots\ldots - a_{1n}x_n}{a_{11}} \longleftarrow \text{From Equation 1}$$

$$x_2 = \frac{c_2 - a_{21}x_1 - a_{23}x_3 \ldots\ldots - a_{2n}x_n}{a_{22}} \longleftarrow \text{From equation 2}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$x_{n-1} = \frac{c_{n-1} - a_{n-1,1}x_1 - a_{n-1,2}x_2 \ldots\ldots - a_{n-1,n-2}x_{n-2} - a_{n-1,n}x_n}{a_{n-1,n-1}} \longleftarrow \text{From equation n-1}$$

$$x_n = \frac{c_n - a_{n1}x_1 - a_{n2}x_2 - \ldots\ldots - a_{n,n-1}x_{n-1}}{a_{nn}} \longleftarrow \text{From equation n}$$

Gauss-Seidel Method

Algorithm

General Form of each equation

$$x_1 = \frac{c_1 - \sum\limits_{\substack{j=1 \\ j \neq 1}}^{n} a_{1j} x_j}{a_{11}}$$

$$x_2 = \frac{c_2 - \sum\limits_{\substack{j=1 \\ j \neq 2}}^{n} a_{2j} x_j}{a_{22}}$$

$$x_{n-1} = \frac{c_{n-1} - \sum\limits_{\substack{j=1 \\ j \neq n-1}}^{n} a_{n-1,j} x_j}{a_{n-1,n-1}}$$

$$x_n = \frac{c_n - \sum\limits_{\substack{j=1 \\ j \neq n}}^{n} a_{nj} x_j}{a_{nn}}$$

Gauss-Seidel Method

Algorithm

General Form for any row 'i'

$$x_i = \frac{c_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j}{a_{ii}}, i = 1,2,\ldots,n.$$

How or where can this equation be used?

Gauss-Seidel Method

Solve for the unknowns

Assume an initial guess for [X]

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}$$

Use rewritten equations to solve for each value of x_i.

Important: Remember to use the most recent value of x_i. Which means to apply values calculated to the calculations remaining in the **current** iteration.

$$x_i^{(k+1)} = \frac{c_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)}}{a_{ii}}, i = 1,2,\ldots,n.$$

Gauss-Seidel Method

If the most recent value of x_i is not used just like Gauss-Seidel method, the algorithm is called *Jacobi Method*

$$x_i^{(k+1)} = \frac{c_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j^{(k)}}{a_{ii}}, i = 1, 2, \ldots, n.$$

Gauss-Seidel Method

Calculate the Absolute Relative Approximate Error

$$\left|\varepsilon_a\right|_i = \left|\frac{x_i^{new} - x_i^{old}}{x_i^{new}}\right| \times 100$$

So when has the answer been found?

The iterations are stopped when the absolute relative approximate error is less than a pre-specified tolerance for all unknowns.

Gauss-Seidel Method: Example 1

The upward velocity of a rocket is given at three different times

Table 1 Velocity vs. Time data.

Time, t (s)	**Velocity** v (m/s)
5	106.8
8	177.2
12	279.2

The velocity data is approximated by a polynomial as:

$$v(t) = a_1 t^2 + a_2 t + a_3 \, , 5 \le \mathrm{t} \le 12.$$

Gauss-Seidel Method: Example 1

Using a Matrix template of the form
$$\begin{bmatrix} t_1^2 & t_1 & 1 \\ t_2^2 & t_2 & 1 \\ t_3^2 & t_3 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

The system of equations becomes
$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

Initial Guess: Assume an initial guess of
$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$

Gauss-Seidel Method: Example 1

Rewriting each equation

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

$$a_1 = \frac{106.8 - 5a_2 - a_3}{25}$$

$$a_2 = \frac{177.2 - 64a_1 - a_3}{8}$$

$$a_3 = \frac{279.2 - 144a_1 - 12a_2}{1}$$

Gauss-Seidel Method: Example 1

Applying the initial guess and solving for a_i

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}$$

Initial Guess

$$a_1 = \frac{106.8 - 5(2) - (5)}{25} = 3.6720$$

$$a_2 = \frac{177.2 - 64(3.6720) - (5)}{8} = -7.8510$$

$$a_3 = \frac{279.2 - 144(3.6720) - 12(-7.8510)}{1} = -155.36$$

When solving for a_2, how many of the initial guess values were used?

Gauss-Seidel Method: Example 1

Finding the absolute relative approximate error

$$\left|\varepsilon_a\right|_i = \left|\frac{x_i^{new} - x_i^{old}}{x_i^{new}}\right| \times 100$$

$$\left|\varepsilon_a\right|_1 = \left|\frac{3.6720 - 1.0000}{3.6720}\right| x100 = 72.76\%$$

$$\left|\varepsilon_a\right|_2 = \left|\frac{-7.8510 - 2.0000}{-7.8510}\right| x100 = 125.47\%$$

$$\left|\varepsilon_a\right|_3 = \left|\frac{-155.36 - 5.0000}{-155.36}\right| x100 = 103.22\%$$

At the end of the first iteration

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 3.6720 \\ -7.8510 \\ -155.36 \end{bmatrix}$$

The maximum absolute relative approximate error is 125.47%

Gauss-Seidel Method: Example 1

Iteration #2

Using

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 3.6720 \\ -7.8510 \\ -155.36 \end{bmatrix}$$

from iteration #1

the values of a_i are found:

$$a_1 = \frac{106.8 - 5(-7.8510) - 155.36}{25} = 12.056$$

$$a_2 = \frac{177.2 - 64(12.056) - 155.36}{8} = -54.882$$

$$a_3 = \frac{279.2 - 144(12.056) - 12(-54.882)}{1} = -798.34$$

Gauss-Seidel Method: Example 1

Finding the absolute relative approximate error

$$\left|\epsilon_{a}\right|_1 = \left|\frac{12.056-3.6720}{12.056}\right| x100 = 69.543\%$$

$$\left|\epsilon_{a}\right|_2 = \left|\frac{-54.882-(-7.8510)}{-54.882}\right| x100 = 85.695\%$$

$$\left|\epsilon_{a}\right|_3 = \left|\frac{-798.34-(-155.36)}{-798.34}\right| x100 = 80.540\%$$

At the end of the second iteration

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 12.056 \\ -54.882 \\ -798.54 \end{bmatrix}$$

The maximum absolute relative approximate error is 85.695%

Gauss-Seidel Method: Example 1

Repeating more iterations, the following values are obtained

Iteration	a_1	$\left\|\in_a\right\|_1 \%$	a_2	$\left\|\in_a\right\|_2 \%$	a_3	$\left\|\in_a\right\|_3 \%$
1	3.6720	72.767	−7.8510	125.47	−155.36	103.22
2	12.056	69.543	−54.882	85.695	−798.34	80.540
3	47.182	74.447	−255.51	78.521	−3448.9	76.852
4	193.33	75.595	−1093.4	76.632	−14440	76.116
5	800.53	75.850	−4577.2	76.112	−60072	75.963
6	3322.6	75.906	−19049	75.972	−249580	75.931

Notice – The relative errors are not decreasing at any significant rate

Also, the solution is not converging to the true solution of

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0.29048 \\ 19.690 \\ 1.0857 \end{bmatrix}$$

Gauss-Seidel Method: Pitfall

What went wrong?

Even though done correctly, the answer is not converging to the correct answer

This example illustrates a pitfall of the Gauss-Siedel method: not all systems of equations will converge.

Is there a fix?

One class of system of equations always converges: One with a *diagonally dominant* coefficient matrix.

Diagonally dominant: [A] in [A] [X] = [C] is diagonally dominant if:

$$\left|a_{ii}\right| \geq \sum_{\substack{j=1 \\ j\neq i}}^{n} \left|a_{ij}\right| \quad \text{for all } `i' \qquad \text{and} \qquad \left|a_{ii}\right| > \sum_{\substack{j=1 \\ j\neq i}}^{n} \left|a_{ij}\right| \quad \text{for at least one } `i'$$

Gauss-Seidel Method: Pitfall

Diagonally dominant: The coefficient on the diagonal must be at least equal to the sum of the other coefficients in that row and at least one row with a diagonal coefficient greater than the sum of the other coefficients in that row.

Which coefficient matrix is diagonally dominant?

$$[A] = \begin{bmatrix} 2 & 5.81 & 34 \\ 45 & 43 & 1 \\ 123 & 16 & 1 \end{bmatrix} \qquad [B] = \begin{bmatrix} 124 & 34 & 56 \\ 23 & 53 & 5 \\ 96 & 34 & 129 \end{bmatrix}$$

Most physical systems do result in simultaneous linear equations that have diagonally dominant coefficient matrices.

Gauss-Seidel Method: Example 2

Given the system of equations

$$12x_1 + 3x_2 - 5x_3 = 1$$

$$x_1 + 5x_2 + 3x_3 = 28$$

$$3x_1 + 7x_2 + 13x_3 = 76$$

The coefficient matrix is:

$$[A] = \begin{bmatrix} 12 & 3 & -5 \\ 1 & 5 & 3 \\ 3 & 7 & 13 \end{bmatrix}$$

With an initial guess of

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Will the solution converge using the Gauss-Siedel method?

Gauss-Seidel Method: Example 2

Checking if the coefficient matrix is diagonally dominant

$$[A] = \begin{bmatrix} 12 & 3 & -5 \\ 1 & 5 & 3 \\ 3 & 7 & 13 \end{bmatrix}$$

$$|a_{11}| = |12| = 12 \geq |a_{12}| + |a_{13}| = |3| + |-5| = 8$$

$$|a_{22}| = |5| = 5 \geq |a_{21}| + |a_{23}| = |1| + |3| = 4$$

$$|a_{33}| = |13| = 13 \geq |a_{31}| + |a_{32}| = |3| + |7| = 10$$

The inequalities are all true and at least one row is *strictly* greater than:

Therefore: The solution should converge using the Gauss-Siedel Method

Gauss-Seidel Method: Example 2

Rewriting each equation

$$\begin{bmatrix} 12 & 3 & -5 \\ 1 & 5 & 3 \\ 3 & 7 & 13 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 28 \\ 76 \end{bmatrix}$$

$$x_1 = \frac{1 - 3x_2 + 5x_3}{12}$$

$$x_2 = \frac{28 - x_1 - 3x_3}{5}$$

$$x_3 = \frac{76 - 3x_1 - 7x_2}{13}$$

With an initial guess of

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$x_1 = \frac{1 - 3(0) + 5(1)}{12} = 0.50000$$

$$x_2 = \frac{28 - (0.5) - 3(1)}{5} = 4.9000$$

$$x_3 = \frac{76 - 3(0.50000) - 7(4.9000)}{13} = 3.0923$$

Gauss-Seidel Method: Example 2

The absolute relative approximate error

$$\left|\varepsilon_a\right|_1 = \left|\frac{0.50000 - 1.0000}{0.50000}\right| \times 100 = 100.00\%$$

$$\left|\varepsilon_a\right|_2 = \left|\frac{4.9000 - 0}{4.9000}\right| \times 100 = 100.00\%$$

$$\left|\varepsilon_a\right|_3 = \left|\frac{3.0923 - 1.0000}{3.0923}\right| \times 100 = 67.662\%$$

The maximum absolute relative error after the first iteration is 100%

Gauss-Seidel Method: Example 2

After Iteration #1

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.5000 \\ 4.9000 \\ 3.0923 \end{bmatrix}$$

Substituting the x values into the equations

$$x_1 = \frac{1 - 3(4.9000) + 5(3.0923)}{12} = 0.14679$$

$$x_2 = \frac{28 - (0.14679) - 3(3.0923)}{5} = 3.7153$$

$$x_3 = \frac{76 - 3(0.14679) - 7(4.900)}{13} = 3.8118$$

After Iteration #2

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.14679 \\ 3.7153 \\ 3.8118 \end{bmatrix}$$

Gauss-Seidel Method: Example 2

Iteration #2 absolute relative approximate error

$$\left|\varepsilon_a\right|_1 = \left|\frac{0.14679 - 0.50000}{0.14679}\right| \times 100 = 240.61\%$$

$$\left|\varepsilon_a\right|_2 = \left|\frac{3.7153 - 4.9000}{3.7153}\right| \times 100 = 31.889\%$$

$$\left|\varepsilon_a\right|_3 = \left|\frac{3.8118 - 3.0923}{3.8118}\right| \times 100 = 18.874\%$$

The maximum absolute relative error after the first iteration is 240.61%

This is much larger than the maximum absolute relative error obtained in iteration #1. Is this a problem?

Gauss-Seidel Method: Example 2

Repeating more iterations, the following values are obtained

Iteration	a_1	$\left\|\in_a\right\|_1 \%$	a_2	$\left\|\in_a\right\|_2 \%$	a_3	$\left\|\in_a\right\|_3 \%$
1	0.50000	100.00	4.9000	100.00	3.0923	67.662
2	0.14679	240.61	3.7153	31.889	3.8118	18.876
3	0.74275	80.236	3.1644	17.408	3.9708	4.0042
4	0.94675	21.546	3.0281	4.4996	3.9971	0.65772
5	0.99177	4.5391	3.0034	0.82499	4.0001	0.074383
6	0.99919	0.74307	3.0001	0.10856	4.0001	0.00101

The solution obtained $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.99919 \\ 3.0001 \\ 4.0001 \end{bmatrix}$ is close to the exact solution of $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$

Gauss-Seidel Method: Example 3

Given the system of equations

$$3x_1 + 7x_2 + 13x_3 = 76$$

$$x_1 + 5x_2 + 3x_3 = 28$$

$$12x_1 + 3x_2 - 5x_3 = 1$$

With an initial guess of

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Rewriting the equations

$$x_1 = \frac{76 - 7x_2 - 13x_3}{3}$$

$$x_2 = \frac{28 - x_1 - 3x_3}{5}$$

$$x_3 = \frac{1 - 12x_1 - 3x_2}{-5}$$

Gauss-Seidel Method: Example 3

Conducting six iterations, the following values are obtained

Iteration	a_1	$\left\|\varepsilon_a\right\|_1 \%$	A_2	$\left\|\varepsilon_a\right\|_2 \%$	a_3	$\left\|\varepsilon_a\right\|_3 \%$
1	21.000	95.238	0.80000	100.00	50.680	98.027
2	−196.15	110.71	14.421	94.453	−462.30	110.96
3	−1995.0	109.83	−116.02	112.43	4718.1	109.80
4	−20149	109.90	1204.6	109.63	−47636	109.90
5	2.0364×10^5	109.89	−12140	109.92	4.8144×10^5	109.89
6	-2.0579×10^5	109.89	1.2272×10^5	109.89	-4.8653×10^6	109.89

The values are not converging.

Does this mean that the Gauss-Seidel method cannot be used?

Gauss-Seidel Method

The Gauss-Seidel Method can still be used

The coefficient matrix is not diagonally dominant

$$[A]=\begin{bmatrix} 3 & 7 & 13 \\ 1 & 5 & 3 \\ 12 & 3 & -5 \end{bmatrix}$$

But this is the same set of equations used in example #2, which did converge.

$$[A]=\begin{bmatrix} 12 & 3 & -5 \\ 1 & 5 & 3 \\ 3 & 7 & 13 \end{bmatrix}$$

If a system of linear equations is not diagonally dominant, check to see if rearranging the equations can form a diagonally dominant matrix.

Gauss-Seidel Method

Not every system of equations can be rearranged to have a diagonally dominant coefficient matrix.

Observe the set of equations

$$x_1 + x_2 + x_3 = 3$$

$$2x_1 + 3x_2 + 4x_3 = 9$$

$$x_1 + 7x_2 + x_3 = 9$$

Which equation(s) prevents this set of equation from having a diagonally dominant coefficient matrix?

Gauss-Seidel Method

Summary

-Advantages of the Gauss-Seidel Method

-Algorithm for the Gauss-Seidel Method

-Pitfalls of the Gauss-Seidel Method

Adequacy of Solutions

Objectives

1. *differentiate between ill-conditioned and well-conditioned systems of equations,*
2. *define the norm of a matrix,*
3. *define the condition number of a square matrix,*
4. *relate the condition number to the ill or well conditioning of a system of equations, that is, determine how much trust you can trust the solution of a set of equations.*

Well-conditioned and ill-conditioned

What do you mean by ill-conditioned and well-conditioned system of equations?

A system of equations is considered to be **well-conditioned** if a small change in the coefficient matrix or a small change in the right hand side results in a small change in the solution vector.

A system of equations is considered to be **ill-conditioned** if a small change in the coefficient matrix or a small change in the right hand side results in a large change in the solution vector.

Example 1

Is this system of equations well-conditioned?

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

Example 1 (cont.)

Solution

The solution to the set of equations is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Make a small change in the right hand side vector of the equations

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7.998 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -3.999 \\ 4.000 \end{bmatrix}$$

Example 1 (cont.)

Make a small change in the coefficient matrix of the equations

$$\begin{bmatrix} 1.001 & 2.001 \\ 2.001 & 3.998 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3.994 \\ 0.001388 \end{bmatrix}$$

This last systems of equation “looks” ill-conditioned because a small change in the coefficient matrix or the right hand side resulted in a large change in the solution vector.

Example 2

Is this system of equations well-conditioned?

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

Example 2 (cont.)

Solution

The solution to the previous equations is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Make a small change in the right hand side vector of the equations.

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7.001 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1.999 \\ 1.001 \end{bmatrix}$$

Example 2 (cont.)

Make a small change in the coefficient matrix of the equations.

$$\begin{bmatrix} 1.001 & 2.001 \\ 2.001 & 3.001 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2.003 \\ 0.997 \end{bmatrix}$$

This system of equation “looks” well conditioned because small changes in the coefficient matrix or the right hand side resulted in small changes in the solution vector.

Well-conditioned and ill-conditioned

So what if the system of equations is ill conditioned or well conditioned?

Well, if a system of equations is ill-conditioned, we cannot trust the solution as much. Revisit the velocity problem:

$$\begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

The values in the coefficient matrix $[A]$ are squares of time, etc. For example, if instead of $a_{11} = 25,$ you used $a_{11} = 24.99,$ would you want this small change to make a huge difference in the solution vector. If it did, would you trust the solution?

Later we will see how much (quantifiable terms) we can trust the solution in a system of equations. Every invertible square matrix has a **condition number** and coupled with the **machine epsilon**, we can quantify how many significant digits one can trust in the solution.

Condition number

To calculate the condition number of an invertible square matrix, I need to know what the norm of a matrix means. How is the norm of a matrix defined?

Just like the determinant, the norm of a matrix is a simple unique scalar number. However, the norm is always positive and is defined for all matrices – square or rectangular, and invertible or noninvertible square matrices.

One of the popular definitions of a norm is the row sum norm (also called the uniform-matrix norm). For a $m \times n$ matrix $[A]$, the row sum norm of $[A]$ is defined as

$$\|A\|_{\infty} = \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|$$

that is, find the sum of the absolute value of the elements of each row of the matrix . The maximum out of the such values is the row sum norm of the matrix .

Example 3

Find the row sum norm of the following matrix [A].

$$A = \begin{bmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix}$$

Solution

$$\|A\|_\infty = \max_{1 \le i \le 3} \sum_{j=1}^{3} |a_{ij}|$$

$$= \max[(|10| + |-7| + |0|), (|-3| + |2.099| + |6|), (|5| + |-1| + |5|)]$$

$$= \max[(10 + 7 + 0), (3 + 2.099 + 6), (5 + 1 + 5)]$$

$$= \max[17, 11.099, 11]$$

$$= 17.$$

How is the norm related to the conditioning of the matrix?

Let us start answering this question using an example. Go back to the *ill-conditioned* system of equations,

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

that gives the solution as

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Denoting the above set of equations as

$$[A][X] = [C]$$

$$\|X\|_\infty = 2$$

$$\|C\|_\infty = 7.999$$

How is the norm related to the conditioning of the matrix? (cont.)

Making a small change in the right hand side,

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7.998 \end{bmatrix}$$

gives,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -3.999 \\ 4.000 \end{bmatrix}$$

Denoting the above set of equations by

$$[A][X'] = [C']$$

right hand side vector is found by

$$[\Delta C] = [C'] - [C]$$

and the change in the solution vector is found by

$$[\Delta X] = [X'] - [X]$$

How is the norm related to the conditioning of the matrix? (cont.)

then

$$[\Delta C] = \begin{bmatrix} 4.001 \\ 7.998 \end{bmatrix} - \begin{bmatrix} 4 \\ 7.999 \end{bmatrix}$$

$$= \begin{bmatrix} 0.001 \\ -0.001 \end{bmatrix}$$

and

$$[\Delta X] = \begin{bmatrix} -3.999 \\ 4.000 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -5.999 \\ 3.000 \end{bmatrix}$$

then

$$\|\Delta C\|_\infty = 0.001$$

$$\|\Delta X\|_\infty = 5.999$$

How is the norm related to the conditioning of the matrix? (cont.)

The relative change in the norm of the solution vector is

$$\frac{\|\Delta X\|_\infty}{\|X\|_\infty} = \frac{5.999}{2}$$
$$= 2.9995$$

The relative change in the norm of the right hand side vector is

$$\frac{\|\Delta C\|_\infty}{\|C\|_\infty} = \frac{0.001}{7.999}$$
$$= 1.250 \times 10^{-4}$$

How is the norm related to the conditioning of the matrix? (cont.)

See the small relative change of 1.250×10^{-4} in the right hand side vector results in a large relative change in the solution vector as 2.9995.

In fact, the ratio between the relative change in the norm of the solution vector and the relative change in the norm of the right hand side vector is

$$\frac{\|\Delta X\|_\infty / \|X\|_\infty}{\|\Delta C\|_\infty / \|C\|_\infty} = \frac{2.9995}{1.250 \times 10^{-4}}$$

$$= 23993$$

How is the norm related to the conditioning of the matrix? (cont.)

Let us now go back to the *well-conditioned* system of equations.

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

Gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Denoting the system of equations by

$$[A][X] = [C]$$

$$\|X\|_\infty = 2$$

$$\|C\|_\infty = 7$$

How is the norm related to the conditioning of the matrix? (cont.)

Making a small change in the right hand side vector

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.001 \\ 7.001 \end{bmatrix}$$

Gives

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1.999 \\ 1.001 \end{bmatrix}$$

Denoting the above set of equations by

$$[A][X'] = [C']$$

the change in the right hand side vector is then found by

$$[\Delta C] = [C'] - [C]$$

How is the norm related to the conditioning of the matrix? (cont.)

and the change in the solution vector is

$$[\Delta X] = [X'] - [X]$$

then

$$[\Delta C] = \begin{bmatrix} 4.001 \\ 7.001 \end{bmatrix} - \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

$$= \begin{bmatrix} 0.001 \\ 0.001 \end{bmatrix}$$

and

$$[\Delta X] = \begin{bmatrix} 1.999 \\ 1.001 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -0.001 \\ 0.001 \end{bmatrix}$$

How is the norm related to the conditioning of the matrix? (cont.)

then

$$\|\Delta C\|_\infty = 0.001$$

$$\|\Delta X\|_\infty = 0.001$$

The relative change in the norm of solution vector is

$$\frac{\|\Delta X\|_\infty}{\|X\|_\infty} = \frac{0.001}{2}$$

$$= 5 \times 10^{-4}$$

The relative change in the norm of the right hand side vector is

$$\frac{\|\Delta C\|_\infty}{\|C\|_\infty} = \frac{0.001}{7}$$

$$= 1.429 \times 10^{-4}$$

How is the norm related to the conditioning of the matrix? (cont.)

See the small relative change the right hand side vector of 1.429×10^{-4} results in the small relative change in the solution vector of 5×10^{-4}

In fact, the ratio between the relative change in the norm of the solution vector and the relative change in the norm of the right hand side vector is

$$\frac{\|\Delta X\|_\infty / \|X\|_\infty}{\|\Delta C\|_\infty / \|C\|_\infty} = \frac{5\times10^{-4}}{1.429\times10^{-4}}$$

$$= 3.5$$

Properties of norms

What are some properties of norms?

1. For a matrix $[A]$, $\|A\| \geq 0$
2. For a matrix $[A]$ and a scalar k, $\|kA\| = |k|\|A\|$
3. For two matrices $[A]$ and $[B]$ of same order, $\|A + B\| \leq \|A\| + \|B\|$
4. For two matrices $[A]$ and $[B]$ that can be multiplied as $[A][B]$, $\|AB\| \leq \|A\|\|B\|$

Identifying well-conditioned and ill conditioned system of equations

Is there a general relationship that exists between $\|\Delta X\|/\|X\|$ and $\|\Delta C\|/\|C\|$ or between $\|\Delta X\|/\|X\|$ and $\|\Delta A\|/\|A\|$? If so, it could help us identify well-conditioned and ill conditioned system of equations.

If there is such a relationship, will it help us quantify the conditioning of the matrix? That is, will it tell us how many significant digits we could trust in the solution of a system of simultaneous linear equations?

Identifying well-conditioned and ill conditioned system of equations (cont.)

there is a relationship that exists between

$$\frac{\|\Delta X\|}{\|X\|} \text{ and } \frac{\|\Delta C\|}{\|C\|}$$

and between

$$\frac{\|\Delta X\|}{\|X\|} \text{ and } \frac{\|\Delta A\|}{\|A\|}$$

these relationships are

$$\frac{\|\Delta X\|}{\|X + \Delta X\|} \leq \|A\| \, \|A^{-1}\| \frac{\|\Delta C\|}{\|C\|}$$

Identifying well-conditioned and ill conditioned system of equations (cont.)

and

$$\frac{\|\Delta X\|}{\|X\|} \le \|A\| \left\|A^{-1}\right\| \frac{\|\Delta A\|}{\|A\|}$$

the above two inequalities show that the relative change in the norm of the right hand side vector or the coefficient matrix can be amplified by as much as $\|A\| \left\|A^{-1}\right\|$.

This number $\|A\| \left\|A^{-1}\right\|$ is called the **condition number** of the matrix and coupled with the machine epsilon, we can quantify the accuracy of the solution of $[A][X] = [C]$

Proof

Proof for $[A][X]=[C]$

that
$$\frac{\|\Delta X\|}{\|X+\Delta X\|} \le \|A\|\,\|A^{-1}\|\frac{\|\Delta A\|}{\|A\|}$$

Proof

let

$$[A][X]=[C] \qquad (1)$$

then $[A]$ is changed to $[A']$ the $[X]$ will change to $[X']$ such that

$$[A'][X']=[C] \qquad (2)$$

Proof (cont.)

From Equations (1) and (2),

$$[A][X]=[A'][X']$$

Denoting change in $[A]$ and $[X]$ matrices as $[\Delta A]$ and $[\Delta X]$, respectively

$$[\Delta A]=[A']-[A]$$

$$[\Delta X]=[X']-[X]$$

then

$$[A][X]=([A]+[\Delta A])([X]+[\Delta X])$$

Proof (cont.)

Expanding the previous expression

$$[A][X]=[A][X]+[A][\Delta X]+[\Delta A][X]+[\Delta A][\Delta X]$$

$$[0]=[A][\Delta X]+[\Delta A]([X]+[\Delta X])$$

$$-[A][\Delta X]=[\Delta A]([X]+[\Delta X])$$

$$[\Delta X]=-[A]^{-1}[\Delta A]([X]+[\Delta X])$$

Applying the theorem of norms, that the norm of multiplied matrices is less than the multiplication of the individual norms of the matrices,

$$\|\Delta X\|\leq\left\|A^{-1}\right\|\|\Delta A\|\|X+\Delta X\|$$

Proof (cont.)

Multiplying both sides by $\|A\|$

$$\|A\|\|\Delta X\| \le \|A\|\|A^{-1}\|\|\Delta A\|\|X+\Delta X\|$$

$$\frac{\|\Delta X\|}{\|X+\Delta X\|} \le \|A\|\|A^{-1}\|\frac{\|\Delta A\|}{\|A\|}$$

How do I use the above theorems to find how many significant digits are correct in my solution vector?

the relative error in a solution vector is Cond (A) relative error in right hand side.
the possible relative error in the solution vector is $\le$ Cond (A) $\times\ \varepsilon_{mach}$

Hence Cond (A) $\times\ \varepsilon_{mach}$ should give us the number of significant digits, *m* at least correct in our solution by comparing it with 0.5×10^{-m}

Example 4

How many significant digits can I trust in the solution of the following system of equations?

$$\begin{bmatrix} 1 & 2 \\ 2 & 3.999 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

Example 4 (cont.)

Solution

For

$$[A]=\begin{bmatrix}1 & 2\\ 2 & 3.999\end{bmatrix}$$

it can be show

$$[A]^{-1}=\begin{bmatrix}-3999 & 2000\\ 2000 & -1000\end{bmatrix}$$

$$\|A\|_{\infty}=5.999$$

$$\|A^{-1}\|_{\infty}=5999$$

$$Cond(A)=\|A\|_{\infty}\|A^{-1}\|_{\infty}$$

$$=5.999\times 5999.4$$

$$=35990$$

Example 4 (cont.)

Assuming single precision with 24 bits used in the mantissa for real numbers, the machine epsilon is

$$\varepsilon_{mach} = 2^{1-24}$$
$$= 0.119209 \times 10^{-6}$$

$$Cond(A) \times \varepsilon_{mach} = 35990 \times 0.119209 \times 10^{-6}$$

$$= 0.4290 \times 10^{-2}$$

comparing it with 0.5×10^{-m}

$$0.5 \times 10^{-m} < 0.4290 \times 10^{-2}$$

$$m \leq 2$$

So two significant digits are at least correct in the solution vector.

Example 5

How many significant digits can I trust in the solution of the following system of equations?

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$$

Example 5 (cont.)

Solution

For

$$[A]=\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix}$$

It can be shown

$$[A]^{-1}=\begin{bmatrix}-3 & 2\\ 2 & -1\end{bmatrix}$$

Then

$$\|A\|_\infty = 5$$

$$\|A^{-1}\|_\infty = 5$$

$$\begin{aligned}\text{Cond}\,(\text{A}) &= \|A\|_\infty \|A^{-1}\|_\infty \\ &= 5\times 5 \\ &= 25\end{aligned}$$

Example 5 (cont.)

Assuming single precision with 24 bits used in the mantissa for real numbers, the machine epsilon

$$\begin{aligned} \varepsilon_{mach} &= 2^{1-24} \\ &= 0.119209 \times 10^{-6} \end{aligned}$$

$$\begin{aligned} Cond(A) \times \varepsilon_{mach} &= 25 \times 0.119209 \times 10^{-6} \\ &= 0.2980 \times 10^{-5} \end{aligned}$$

Comparing it with 0.5×10^{-m}

$$0.5 \times 10^{-m} \leq 0.2980 \times 10^{-5}$$

$$m \leq 5$$

So five significant digits are at least correct in the solution vector.

Key terms

- *Ill-Conditioned matrix*
- *Well-Conditioned matrix*
- *Norm*
- *Condition Number*
- *Machine Epsilon*
- *Significant Digits*

Eigenvalues and Eigenvectors

Objectives

1. *Define eigenvalues and eigenvectors of a square matrix*
2. *Find eigenvalues and eigenvectors of a square matrix*
3. *Relate eigenvalues to the singularity of a square matrix, and*
4. *Use the power method to numerically find the largest eigenvalue in magnitude of a square matrix and the corresponding eigenvector.*

Eigenvalue

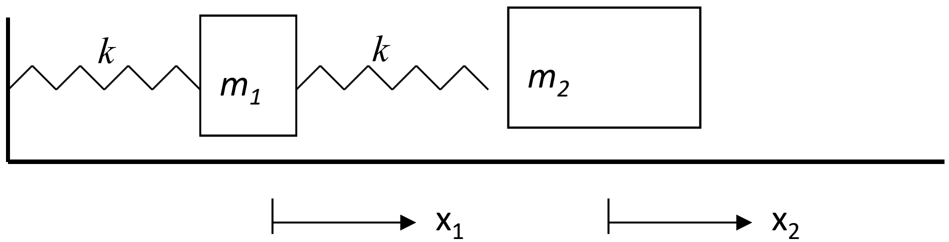
What does eigenvalue mean?

The word eigenvalue comes from the German word *Eigenwert* where Eigen means *characteristic* and Wert means *value*. However, what the word means is not on your mind! You want to know why I need to learn about eigenvalues and eigenvectors. Once I give you an example of an application of eigenvalues and eigenvectors, you will want to know how to find these eigenvalues and eigenvectors.

Physical example

Can you give me a physical example application of eigenvalues and eigenvectors?

Look at the spring-mass system as shown in the picture below.



Assume each of the two mass-displacements to be denoted by x_1 and x_2, and let us assume each spring has the same spring constant k.

Physical example (cont.)

Then by applying Newton's 2nd and 3rd law of motion to develop a force-balance for each mass we have

$$m_1 \frac{d^2 x_1}{dt^2} = -kx_1 + k(x_2 - x_1) \quad \text{and} \quad m_2 \frac{d^2 x_2}{dt^2} = -k(x_2 - x_1)$$

Rewriting the equations, we have

$$m_1 \frac{d^2 x_1}{dt^2} - k(-2x_1 + x_2) = 0 \quad \text{and} \quad m_2 \frac{d^2 x_2}{dt^2} - k(x_1 - x_2) = 0$$

Let $m_1 = 10, m_2 = 20, k = 15$ which gives,

$$10 \frac{d^2 x_1}{dt^2} - 15(-2x_1 + x_2) = 0 \quad \text{and} \quad 20 \frac{d^2 x_2}{dt^2} - 15(x_1 - x_2) = 0$$

Physical example (cont.)

From vibration theory, the solutions can be of the form

$$x_i = A_i \sin\left(\omega t - \emptyset\right)$$

Where

A_i = amplitude of the vibration of mass ,
ω = frequency of vibration,
$\emptyset$ = phase shift.

Then

$$\frac{d^2 x_i}{dt^2} = -A_i w^2 Sin(\omega t - \emptyset)$$

Physical example (cont.)

Substituting x_i and $\frac{d^2x_i}{dt^2}$ in equations,

$$-10A_1\omega^2 - 15(-2A_1 + A_2) = 0 \quad \text{and} \quad -20A_2\omega^2 - 15(A_1 - A_2) = 0$$

gives

$$(-10\omega^2 + 30)A_1 - 15A_2 = 0 \quad \text{and} \quad -15A_1 + (-20\omega^2 + 15)A_2 = 0$$

or

$$(-\omega^2 + 3)A_1 - 1.5A_2 = 0 \quad \text{and} \quad -0.75A_1 + (-\omega^2 + 0.75)A_2 = 0$$

Physical example (cont.)

Substituting x_i and $\dfrac{d^2 x_i}{dt^2}$ in equations,

$$-10A_1\omega^2 - 15(-2A_1 + A_2) = 0 \quad \text{and} \quad -20A_2\omega^2 - 15(A_1 - A_2) = 0$$

gives

$$(-10\omega^2 + 30)A_1 - 15A_2 = 0 \quad \text{and} \quad -15A_1 + (-20\omega^2 + 15)A_2 = 0$$

or

$$(-\omega^2 + 3)A_1 - 1.5A_2 = 0 \quad \text{and} \quad -0.75A_1 + (-\omega^2 + 0.75)A_2 = 0$$

In matrix form, these equations can be rewritten as

$$\begin{bmatrix} -\omega^2 + 3 & -1.5 \\ -0.75 & -\omega^2 + 0.75 \end{bmatrix}\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 3 & -1.5 \\ -0.75 & 0.75 \end{bmatrix}\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} - \omega^2\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Physical example (cont.)

Let $\omega^2 = \lambda$

$$[A] = \begin{bmatrix} 3 & -1.5 \\ -0.75 & 0.75 \end{bmatrix}$$

$$[X] = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

$$[A][X] - \lambda[X] = 0$$

$$[A][X] = \lambda[X]$$

In the above equation, λ is the eigenvalue and $[X]$ is the eigenvector corresponding to λ . As you can see, if we know λ for the above example we can calculate the natural frequency of the vibration $\omega = \sqrt{\lambda}$

Physical example (cont.)

Why are the natural frequencies of vibration important? Because you do not want to have a forcing force on the spring-mass system close to this frequency as it would make the amplitude A_i very large and make the system unstable.

General definition of eigenvalues and eigenvectors of a square matrix

What is the general definition of eigenvalues and eigenvectors of a square matrix?

If $[A]$ is a $n \times n$ matrix, then $[X] \neq \vec{0}$ is an eigenvector of $[A]$ if

$$[A][X] = \lambda[X]$$

where λ is a scalar and $[X] \neq 0$. The scalar λ is called the eigenvalue of $[A]$ and $[A]$ is called the eigenvector corresponding to the eigenvalue λ .

How do I find eigenvalues of a square matrix?

To find the eigenvalues of a $n \times n$ matrix $[A]$, we have

$$[A][X] = \lambda[X]$$

$$[A][X] - \lambda[X] = 0$$

$$[A][X] - \lambda[I][X] = 0$$

$$([A] - [\lambda][I])[X] = 0$$

Now for the above set of equations to have a nonzero solution,

$$\det([A] - \lambda[I]) = 0$$

How do I find eigenvalues of a square matrix? (cont.)

This left hand side can be expanded to give a polynomial in λ solving the above equation would give us values of the eigenvalues. The above equation is called the characteristic equation of $[A]$.

For a $[A]$ $n \times n$ matrix, the characteristic polynomial of A is of degree n as follows

$$\det([A] - \lambda[I]) = 0$$

giving

$$\lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + - - + c_n = 0$$

Hence. this polynomial has n roots

Example 1

Find the eigenvalues of the physical problem discussed in the beginning of this chapter, that is, find the eigenvalues of the matrix

$$[A] = \begin{bmatrix} 3 & -1.5 \\ -0.75 & 0.75 \end{bmatrix}$$

Solution

$$[A] - \lambda[I] = \begin{bmatrix} 3-\lambda & -1.5 \\ -0.75 & 0.75-\lambda \end{bmatrix}$$

$$\det([A] - \lambda[I]) = (3-\lambda)(0.75-\lambda) - (-0.75)(-1.5) = 0$$

$$2.25 - 0.75\lambda - 3\lambda + \lambda^2 - 1.125 = 0$$

Example 1 (cont.)

$$\lambda^2 - 3.75\lambda + 1.125 = 0$$

$$\lambda = \frac{-(-3.75) \pm \sqrt{(-3.75)^2 - 4(1)(1.125)}}{2(1)}$$

$$= \frac{3.75 \pm 3.092}{2}$$

$$= 3.421, 0.3288$$

So the eigenvalues are 3.421 and 0.3288.

Example 2

Find the eigenvectors of

$$A = \begin{bmatrix} 3 & -1.5 \\ -0.75 & 0.75 \end{bmatrix}$$

Solution

The eigenvalues have already been found in Example 1 as

$$\lambda_1 = 3.421, \lambda_2 = 0.3288$$

Let

$$[X] = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

be the eigenvector corresponding to

$$\lambda_1 = 3.421$$

Example 2 (cont.)

Hence

$$([A]-\lambda_1[I])[X]=0$$

$$\left\{\begin{bmatrix} 3 & -1.5 \\ -0.75 & 0.75 \end{bmatrix} - 3.421\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} -0.421 & -1.5 \\ -0.75 & -2.671 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

If

$$x_1 = s$$

then

$$-0.421s - 1.5x_2 = 0$$

$$x_2 = -0.2808s$$

Example 2 (cont.)

The eigenvector corresponding to $\lambda_1 = 3.421$ then is,

$$[X] = \begin{bmatrix} s \\ -0.2808s \end{bmatrix}$$

$$= s\begin{bmatrix} 1 \\ -0.2808 \end{bmatrix}$$

The eigenvector corresponding to $\lambda_1 = 3.421$ is

$$\begin{bmatrix} 1 \\ -0.2808 \end{bmatrix}$$

Similarly, the eigenvector corresponding to $\lambda_2 = 0.3288$ is

$$\begin{bmatrix} 1 \\ 1.781 \end{bmatrix}$$

Example 3

Find the eigenvalues and eigenvectors of

$$[A] = \begin{bmatrix} 1.5 & 0 & 1 \\ -0.5 & 0.5 & -0.5 \\ -0.5 & 0 & 0 \end{bmatrix}$$

Solution

The characteristic equation is given by

$$\det([A] - \lambda[I]) = 0$$

$$\det\begin{bmatrix} 1.5-\lambda & 0 & 1 \\ -0.5 & 0.5-\lambda & -0.5 \\ -0.5 & 0 & -\lambda \end{bmatrix} = 0$$

$$(1.5-\lambda)[(0.5-\lambda)(-\lambda)-(-0.5)(0)]+(1)[(-0.5)(0)-(-0.5)(0.5-\lambda)]=0$$

Example 3 (cont.)

$$-\lambda^3 + 2\lambda^2 - 1.25\lambda = 0$$

The roots of the above equation are

$$\lambda = 0.5, 0.5, 1.0$$

Note that there are eigenvalues that are repeated. Since there are only two distinct eigenvalues, there are only two eigenspaces. But, corresponding to $\lambda = 0.5$ there should be two eigenvectors that form a basis for the eigenspace.

To find the eigenspaces, let

$$[X] = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Example 3 (cont.)

Given

$$[(A-\lambda I)][X]=0$$

then

$$\begin{bmatrix} 1.5-\lambda & 0 & 1 \\ -0.5 & 0.5-\lambda & -0.5 \\ -0.5 & 0 & -\lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

For $\lambda = 0.5$,

$$\begin{bmatrix} 1 & 0 & 1 \\ -0.5 & 0 & -0.5 \\ -0.5 & 0 & -0.5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving this system gives

$$x_1 = -a, x_2 = b\, x_3 = a$$

Example 3 (cont.)

So

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -a \\ b \\ a \end{bmatrix}$$

$$= \begin{bmatrix} a \\ 0 \\ -a \end{bmatrix} + \begin{bmatrix} 0 \\ b \\ 0 \end{bmatrix}$$

$$= a\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

So the vectors $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ form a basis for the eigenspace for the eigenvalue $\lambda = 0.5$

Example 3 (cont.)

For $\lambda = 1$

$$\begin{bmatrix} 0.5 & 0 & 1 \\ -0.5 & -0.5 & -0.5 \\ -0.5 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving this system gives

$$x_1 = a, x_2 = -0.5a, x_3 = -0.5a$$

The eigenvector corresponding to $\lambda = 1$ is

$$\begin{bmatrix} a \\ -0.5a \\ -0.5a \end{bmatrix} = a \begin{bmatrix} 1 \\ -0.5 \\ -0.5 \end{bmatrix}$$

Hence the vector $\begin{bmatrix} 1 \\ -0.5 \\ -0.5 \end{bmatrix}$ is a basis for the eigenspace for the eigenvalue of $\lambda = 1$

Theorems of eigenvalues and eigenvectors

Theorem 1: If $[A]$ is a $n \times n$ triangular matrix – upper triangular, lower triangular or diagonal, the eigenvalues of $[A]$ are the diagonal entries of $[A]$.

Theorem 2: $\lambda = 0$ is an eigenvalue of $[A]$ if $[A]$ is a singular (noninvertible) matrix.

Theorem 3: $[A]$ and $[A]^{\mathrm{T}}$ have the same eigenvalues.

Theorem 4: Eigenvalues of a symmetric matrix are real.

Theorem 5: Eigenvectors of a symmetric matrix are orthogonal, but only for distinct eigenvalues.

Theorem 6: $|\det(A)|$ is the product of the absolute values of the eigenvalues of $[A]$

Example 4

What are the eigenvalues of

$$[A] = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 7 & 3 & 0 & 0 \\ 9 & 5 & 7.5 & 0 \\ 2 & 6 & 0 & -7.2 \end{bmatrix}$$

Solution

Since the matrix $[A]$ is a lower triangular matrix, the eigenvalues of $[A]$ are the diagonal elements of $[A]$. The eigenvalues are

$$\lambda_1 = 6, \lambda_2 = 3, \lambda_3 = 7.5, \lambda_4 = -7.2$$

Example 5

One of the eigenvalues of

$$[A]=\begin{bmatrix}5 & 6 & 2\\ 3 & 5 & 9\\ 2 & 1 & -7\end{bmatrix}$$

is zero. Is $[A]$ invertible?

Solution

$\lambda = 0$ is an eigenvalue of $[A]$, that implies $[A]$ is singular and is not invertible.

Example 6

Given the eigenvalues of

$$[A]=\begin{bmatrix} 2 & -3.5 & 6 \\ 3.5 & 5 & 2 \\ 8 & 1 & 8.5 \end{bmatrix}$$

are

$$\lambda_1 = -1.547, \lambda_2 = 12.33, \lambda_3 = 4.711$$

What are the eigenvalues of $[B]$ if

$$[B]=\begin{bmatrix} 2 & 3.5 & 8 \\ -3.5 & 5 & 1 \\ 6 & 2 & 8.5 \end{bmatrix}$$

Example 6 (cont.)

Solution

Since $[B]=[A]^T$, the eigenvalues of $[A]$ and $[B]$ are the same. Hence eigenvalues of $[B]$ also are

$$\lambda_1 = -1.547, \lambda_2 = 12.33, \lambda_3 = 4.711$$

Example 7

Given the eigenvalues of

$$[A]=\begin{bmatrix} 2 & -3.5 & 6 \\ 3.5 & 5 & 2 \\ 8 & 1 & 8.5 \end{bmatrix}$$

are

$$\lambda_1 = -1.547, \lambda_2 = 12.33, \lambda_3 = 4.711$$

Calculate the magnitude of the determinant of the matrix.

Example 7 (cont.)

Solution

Since

$$|\det[A]| = |\lambda_1||\lambda_2||\lambda_3|$$

$$= |-1.547||12.33||4.711|$$

$$= 89.88$$

Finding eigenvalues and eigenvectors numerically

How does one find eigenvalues and eigenvectors numerically?

One of the most common methods used for finding eigenvalues and eigenvectors is the power method. It is used to find the largest eigenvalue in an absolute sense. Note that if this largest eigenvalues is repeated, this method will not work. Also this eigenvalue needs to be distinct. The method is as follows:

Finding eigenvalues and eigenvectors numerically (cont.)

1. Assume a guess $[X^{(0)}]$ for the eigenvector in $[A][X]=\lambda[X]$ equation. one of the entries of $[X^{(0)}]$ needs to be unity.

2. Find

$$[Y^{(1)}]=[A][X^{(0)}]$$

3. Scale $[Y^{(1)}]$ so that the chosen unity component remains unity

$$[Y^{(1)}]=\lambda^{(1)}[X^{(1)}]$$

4. Repeat steps (2) and (3) with

$[X]=[X^{(1)}]$ to get $[X^{(2)}]$

5. Repeat the steps 2 and 3 until the value of the eigenvalue converges.

Finding eigenvalues and eigenvectors numerically (cont.)

If E_s is the pre-specified percentage relative error tolerance to which you would like the answer to converge to, keep iterating until

$$\left|\frac{\lambda^{(i+1)} - \lambda^{(i)}}{\lambda^{(i+1)}}\right| \times 100 \le E_s$$

where the left hand side of the above inequality is the definition of absolute percentage relative approximate error, denoted generally by E_s A pre-specified percentage relative tolerance of $0.5 \times 10^{2-m}$ Implies at least m significant digits are current in your answer. When the system converges, the value of λ is the largest (in absolute value) eigenvalue of $[A]$.

Example 8

Using the power method, find the largest eigenvalue and the corresponding eigenvector of

$$[A] = \begin{bmatrix} 1.5 & 0 & 1 \\ -0.5 & 0.5 & -0.5 \\ -0.5 & 0 & 0 \end{bmatrix}$$

Example 8 (cont.)

Solution

Assume

$$[X^{(0)}] = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$[A][X^{(0)}] = \begin{bmatrix} 1.5 & 0 & 1 \\ -0.5 & 0.5 & -0.5 \\ -0.5 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2.5 \\ -0.5 \\ -0.5 \end{bmatrix}$$

$$Y^{(1)} = 2.5\begin{bmatrix} 1 \\ -0.2 \\ -0.2 \end{bmatrix}$$

$$\lambda^{(1)} = 2.5$$

Example 8 (cont.)

We will choose the first element of $[X^{(0)}]$ to be unity.

$$[X^{(1)}] = \begin{bmatrix} 1 \\ -0.2 \\ -0.2 \end{bmatrix}$$

$$[A][X^{(1)}] = \begin{bmatrix} 1.5 & 0 & 1 \\ -0.5 & 0.5 & -0.5 \\ -0.5 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -0.2 \\ -0.2 \end{bmatrix}$$

$$= \begin{bmatrix} 1.3 \\ -0.5 \\ -0.5 \end{bmatrix}$$

$$[X^{(2)}] = 1.3 \begin{bmatrix} 1 \\ -0.3846 \\ -0.3846 \end{bmatrix}$$

$$\lambda^{(2)} = 1.3$$

Example 8 (cont.)

$$[X^{(2)}] = \begin{bmatrix} 1 \\ -0.3846 \\ -0.3846 \end{bmatrix}$$

The absolute relative approximate error in the eigenvalues is

$$\left|\varepsilon_a\right| = \left|\frac{\lambda^{(2)} - \lambda^{(1)}}{\lambda^{(2)}}\right| \times 100$$

$$= \left|\frac{1.3 - 1.5}{1.5}\right| \times 100$$

$$= 92.307\%$$

Example 8 (cont.)

Conducting further iterations, the values of $\lambda^{(i)}$ and the corresponding eigenvectors is given in the table below

i	$\lambda^{(i)}$	$[X^{(i)}]$	$\lvert\varepsilon_a\rvert$ (%)
1	2.5	$\begin{bmatrix} 1 \\ -0.2 \\ -0.2 \end{bmatrix}$	———
2	1.3	$\begin{bmatrix} 1 \\ -0.38462 \\ -0.38462 \end{bmatrix}$	92.307
3	1.1154	$\begin{bmatrix} 1 \\ -0.44827 \\ -0.44827 \end{bmatrix}$	16.552
4	1.0517	$\begin{bmatrix} 1 \\ -0.47541 \\ -0.47541 \end{bmatrix}$	6.0529
5	1.02459	$\begin{bmatrix} 1 \\ -0.48800 \\ -0.48800 \end{bmatrix}$	1.2441

Example 8 (cont.)

The exact value of the eigenvalue is $\lambda = 1$

and the corresponding eigenvector is

$$[X] = \begin{bmatrix} 1 \\ -0.5 \\ -0.5 \end{bmatrix}$$

Keyterms

- *Eigenvalue*
- *Eigenvectors*
- *Power method*

Homework

1. Write a LU decomposition code to solve equations $Ax = b$:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 1 & -2 \\ 2 & 5 & 3 & -2 \\ -2 & -2 & 3 & 5 \\ 1 & 3 & 2 & 3 \end{bmatrix} \qquad \boldsymbol{b} = (4, 7, -1, 0)^{\mathrm{T}}$$

2. Write a Gauss-Seidel code to solve equations:

$$\begin{cases} 11x_1 - 3x_2 - 2x_3 = 3 \\ -x_1 + 5x_2 - 3x_3 = 6 \\ -2x_1 - 12x_2 + 19x_3 = -7 \end{cases}$$

The initial guess is $(0, 0, 0)^{\mathrm{T}}$, write the results of the first 3 iterations