Systems of Linear Equations

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

A set of n equations and n unknowns:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n &= b_2 \\
\vdots \qquad\qquad & \vdots \\
a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots + a_{nn}x_n &= b_n
\end{aligned}
$$

Gauss-Seidel /Jacobi Method

- Iterative or approximate methods provide an alternative to the elimination methods. *The Gauss-Seidel* and *Gauss-Jacobi method* are the most commonly used iterative methods.
- The system $[A]\{X\} = \{B\}$ is reshaped by solving the first equation for x_1, the second equation for x_2 and the third for x_3, …and n^{th} equation for x_n. For conciseness, we will limit ourselves to a 3x3 set of equations.

$$x_1 = \frac{b_1 - a_{12}x_2 - a_{13}x_3}{a_{11}}$$

$$x_2 = \frac{b_2 - a_{21}x_1 - a_{23}x_3}{a_{22}}$$

$$x_3 = \frac{b_3 - a_{31}x_1 - a_{32}x_2}{a_{33}}$$

Now we can start the solution process by choosing guesses for the *x*'s. A simple way to obtain initial guesses is to assume that they are zero. These zeros can be substituted into x_1 equation to calculate a new $x_1 = b_1/a_{11}$.

The Gauss-Seidel method **First Iteration** **The Gauss-Jacobi method**

The Gauss-Seidel method

First Iteration

$$x_1 = (b_1 - a_{12}x_2 - a_{13}x_3)/a_{11}$$
$$x_2 = (b_2 - a_{21}x_1 - a_{23}x_3)/a_{22}$$
$$x_3 = (b_3 - a_{31}x_1 - a_{32}x_2)/a_{33}$$

Second Interation

$$x_1 = (b_1 - a_{12}x_2 - a_{13}x_3)/a_{11}$$
$$x_2 = (b_2 - a_{21}x_1 - a_{23}x_3)/a_{22}$$
$$x_3 = (b_3 - a_{31}x_1 - a_{32}x_2)/a_{33}$$

The Gauss-Jacobi method

First Iteration

$$x_1 = (b_1 - a_{12}x_2 - a_{13}x_3)/a_{11}$$
$$x_2 = (b_2 - a_{21}x_1 - a_{23}x_3)/a_{22}$$
$$x_3 = (b_3 - a_{31}x_1 - a_{32}x_2)/a_{33}$$

Second Interation

$$x_1 = (b_1 - a_{12}x_2 - a_{13}x_3)/a_{11}$$
$$x_2 = (b_2 - a_{21}x_1 - a_{23}x_3)/a_{22}$$
$$x_3 = (b_3 - a_{31}x_1 - a_{32}x_2)/a_{33}$$

New x_1 is substituted to calculate x_2 and x_3. The procedure is repeated until the convergence criterion is satisfied:

$$\left|\varepsilon_{a,i}\right| = \left|\frac{x_i^j - x_i^{j-1}}{x_i^j}\right| 100\% < \varepsilon_s$$

For all *i*, where *j* and *j-1* are the present and previous iterations.

Convergence Criterion for Gauss-Seidel Method

- The Gauss-Seidel method has two fundamental problems as any iterative method:
  - It is sometimes non convergent, and
  - If it converges, converges very slowly.
- Recalling that sufficient conditions for convergence of two linear equations, *u*(*x,y*) and *v*(*x,y*) are

$$\left|\frac{\partial u}{\partial x}\right| + \left|\frac{\partial u}{\partial y}\right| < 1 \quad , \quad \left|\frac{\partial v}{\partial x}\right| + \left|\frac{\partial v}{\partial y}\right| < 1$$

- Similarly, in case of two simultaneous equations, the Gauss-Seidel algorithm can be expressed as

$$u(x_1, x_2) = \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}} x_2 \quad , \quad v(x_1, x_2) = \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}} x_1$$

$$\frac{\partial u}{\partial x_1} = 0 \qquad \qquad \frac{\partial u}{\partial x_2} = -\frac{a_{12}}{a_{11}}$$

$$\frac{\partial v}{\partial x_1} = -\frac{a_{21}}{a_{22}} \qquad \qquad \frac{\partial v}{\partial x_2} = 0$$

- Substitution into convergence criterion of two linear equations yield:

$$\left|\frac{a_{12}}{a_{11}}\right| < 1 \quad , \quad \left|\frac{a_{21}}{a_{22}}\right| < 1$$

- In other words, the absolute values of the slopes must be less than unity for convergence: $|a_{11}| > |a_{12}| \quad , \quad |a_{22}| > |a_{21}|$

For *n* equations : $$\boxed{|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{i,j}|}$$

Not all systems of equations will converge. One class of system of equations always converges if it has a *diagonally dominant* coefficient matrix. Diagonally dominant: [*A*] in [*A*][*X*] = [*C*] is diagonally dominant if:

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

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.

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

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

Example: Given the system of equations

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

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

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

With an initial guess of

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

The coefficient matrix is:

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

Will the solution converge using the Gauss-Seidel method? Checking if the coefficient matrix is diagonally dominant

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

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

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

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

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

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

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

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

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

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

The absolute relative approximate error

$$\left|\in_a\right|_1 = \left|\frac{0.5000 - 1.0000}{0.5000}\right| \times 100 = 100.00\%$$

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

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

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

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

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

Iteration #2 absolute relative approximate error

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

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

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

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. 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 $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$.

Example: 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

$$x_1 = \frac{76 - 7x_2 - 13x_3}{3} \qquad x_2 = \frac{28 - x_1 - 3x_3}{5} \qquad x_3 = \frac{1 - 12x_1 - 3x_2}{-5}$$

Conducting six 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	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? The Gauss-Seidel Method can still be used

The coefficient matrix is not diagonally dominant $[A] = \begin{bmatrix} 3 & 7 & 13 \\ 1 & 5 & 3 \\ 12 & 3 & -5 \end{bmatrix}$

$[A] = \begin{bmatrix} 12 & 3 & -5 \\ 1 & 5 & 3 \\ 3 & 7 & 13 \end{bmatrix}$ But this is the same set of equations used in the previous example, which did converge.

If a system of linear equations is not diagonally dominant, check to see if rearranging the equations can form a diagonally dominant matrix. 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 equations prevent this set of equations from having a diagonally dominant coefficient matrix?

LU-Decompositions

DEFINITION 1 A factorization of a square matrix A as $A = LU$, where L is lower triangular and U is upper triangular is called an *LU-decomposition* (or *LU-factorization*) of A.

The Method of LU-Decomposition

Step 1. Rewrite the system $A\mathbf{x} = \mathbf{b}$ as

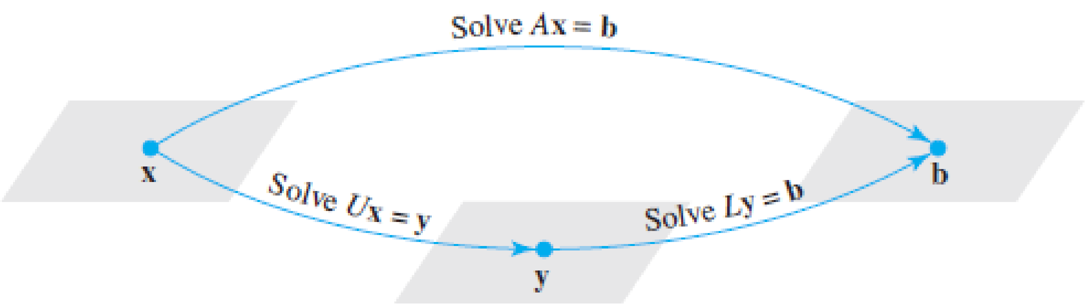
$$LU\mathbf{x} = \mathbf{b} \tag{2}$$

Step 2. Define a new $n \times 1$ matrix $\mathbf{y}$ by

$$U\mathbf{x} = \mathbf{y} \tag{3}$$

Step 3. Use (3) to rewrite (2) as $L\mathbf{y} = \mathbf{b}$ and solve this system for $\mathbf{y}$.

Step 4. Substitute $\mathbf{y}$ in (3) and solve for $\mathbf{x}$.



Method

LU Decomposition is another method to solve a set of simultaneous linear equations. For most non-singular matrix [*A*], one can always write it as

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

where [*L*] = lower triangular matrix, [*U*] = upper triangular matrix

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*]

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

Finding the [*L*][*U*] matrix

- Using multiplication property

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

$$=\begin{bmatrix}u_{11} & u_{12} & u_{13}\\ \ell_{21}u_{11} & \ell_{21}u_{12}+u_{22} & \ell_{21}u_{13}+u_{23}\\ \ell_{31}u_{11} & \ell_{31}u_{12}+\ell_{32}u_{22} & \ell_{31}u_{13}+\ell_{32}u_{23}+u_{33}\end{bmatrix}$$

Example:

$$\begin{bmatrix}1 & -3 & 7\\ 2 & 4 & -3\\ -3 & 7 & 2\end{bmatrix}=\begin{bmatrix}u_{11} & u_{12} & u_{13}\\ \ell_{21}u_{11} & \ell_{21}u_{12}+u_{22} & \ell_{21}u_{13}+u_{23}\\ \ell_{31}u_{11} & \ell_{31}u_{12}+\ell_{32}u_{22} & \ell_{31}u_{13}+\ell_{32}u_{23}+u_{33}\end{bmatrix}$$

$$=\begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ -3 & -1/5 & 1\end{bmatrix}\begin{bmatrix}1 & -3 & 7\\ 0 & 10 & -17\\ 0 & 0 & 98/5\end{bmatrix}$$

Finding the [*U*] matrix

Example:

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

Step 1: Using the Forward Gauss Elimination

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

$$\frac{144}{25} = 5.76; \quad Row3 - Row1(5.76) = \begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & -16.8 & -4.76 \end{bmatrix} \qquad (2)$$

$$\frac{-16.8}{-4.8} = 3.5; \quad Row3 - Row2(3.5) = \begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & 0 & 0.7 \end{bmatrix} = [U]$$

Finding the [*L*] matrix

Step 2: Using the multipliers during the Forward Elimination Procedure.

From (1) in the forward elimination

$$\begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix}$$

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

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

From (2) in the forward elimination

$$\begin{bmatrix} 25 & 5 & 1 \\ 0 & -4.8 & -1.56 \\ 0 & -16.8 & -4.76 \end{bmatrix} \Rightarrow [L] = \begin{bmatrix} 1 & 0 & 0 \\ 2.56 & 1 & 0 \\ 5.76 & 3.5 & 1 \end{bmatrix}$$

$$\ell_{32} = \frac{a_{32}}{a_{22}} = \frac{-16.8}{-4.8} = 3.5$$

[L][U] = [A]

Example: 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

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

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*]

$$z_1 = 106.8$$

$$2.56 z_1 + z_2 = 177.2$$

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

Complete the forward substitution to solve for $[Z]$

$$z_1 = 106.8$$

$$\begin{aligned} z_2 &= 177.2 - 2.56 z_1 \\ &= 177.2 - 2.56(106.8) \\ &= -96.2 \end{aligned}$$

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

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

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 $[X]$ becomes

$$\begin{aligned} 25a_1 + 5a_2 + a_3 &= 106.8 \\ -4.8a_2 - 1.56a_3 &= -96.21 \\ 0.7a_3 &= 0.735 \end{aligned}$$

From the 3rd equation

$$0.7a_3 = 0.735$$

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

$$a_3 = 1.050$$

Substituting a_3 in the second equation

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

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

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

$$a_2 = 19.70$$

Substituting a_3 and a_2 in the first equation

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

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

$$= \frac{106.8 - 5(19.70) - 1.050}{25}$$

$$= 0.2900$$

Hence the Solution Vector is:

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

The Power Method Approximating Eigen values

DEFINITION

Dominant Eigen value

Let $\lambda_1, \lambda_2, \cdots, \lambda_k, \cdots, \lambda_n$ denote the eigen values of an $n \times n$ matrix **A**. The eigen values λ_k is said to be the **dominant eigen values** of **A** if

$$|\lambda_k| > |\lambda_i|, \quad i = 1, 2, 3, \cdots, n, \text{ but } i \neq k$$

An eigen vector corresponding to λ_k is called the **dominant eigen vector** of **A**.

The Power Method

The Power Method with Maximum Entry Scaling

Step 1. Choose an arbitrary nonzero vector $\mathbf{x}_0$.

Step 2. Compute $A\mathbf{x}_0$ and multiply it by the factor $1/\max(A\mathbf{x}_0)$ to obtain the first approximation $\mathbf{x}_1$ to a dominant eigenvector. Compute the Rayleigh quotient of $\mathbf{x}_1$ to obtain the first approximation to the dominant eigenvalue.

Step 3. Compute $A\mathbf{x}_1$ and scale it by the factor $1/\max(A\mathbf{x}_1)$ to obtain the second approximation $\mathbf{x}_2$ to a dominant eigenvector. Compute the Rayleigh quotient of $\mathbf{x}_2$ to obtain the second approximation to the dominant eigenvalue.

Step 4. Compute $A\mathbf{x}_2$ and scale it by the factor $1/\max(A\mathbf{x}_2)$ to obtain the third approximation $\mathbf{x}_3$ to a dominant eigenvector. Compute the Rayleigh quotient of $\mathbf{x}_3$ to obtain the third approximation to the dominant eigenvalue.

Continuing in this way will generate a sequence of better and better approximations to the dominant eigenvalue and a corresponding eigenvector.

Example: Consider the follow matrix A

$$A = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

Assume an arbitrary vector $x_0 = \{1\ 1\ 1\}^T$

Multiply the matrix by the matrix [A] by {x}

$$\begin{bmatrix} 4 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} 5 \\ 3 \\ -1 \end{Bmatrix} \quad \Rightarrow \begin{Bmatrix} 5 \\ 3 \\ -1 \end{Bmatrix} = 5 \begin{Bmatrix} 1 \\ 0.6 \\ -0.2 \end{Bmatrix}$$

Normalize the result of the product

$$\begin{bmatrix} 4 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{Bmatrix} 1 \\ 0.6 \\ -0.2 \end{Bmatrix} = \begin{Bmatrix} 4.6 \\ 1 \\ 0.2 \end{Bmatrix} \quad \Rightarrow \begin{Bmatrix} 4.6 \\ 1 \\ 0.2 \end{Bmatrix} = 4.6 \begin{Bmatrix} 1 \\ 0.217 \\ 0.0435 \end{Bmatrix}$$

$$\begin{bmatrix} 4 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{Bmatrix} 1 \\ 0.217 \\ 0.0435 \end{Bmatrix} = \begin{Bmatrix} 4.2174 \\ 0.4783 \\ -0.0435 \end{Bmatrix} \Rightarrow \begin{Bmatrix} 4.2174 \\ 0.4783 \\ -0.0435 \end{Bmatrix} = 4.2174 \begin{Bmatrix} 1 \\ 0.1134 \\ -0.0183 \end{Bmatrix}$$

$$\begin{bmatrix} 4 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{Bmatrix} 1 \\ 0.1134 \\ -0.0183 \end{Bmatrix} = \begin{Bmatrix} 4.1134 \\ 0.2165 \\ 0.0103 \end{Bmatrix} \Rightarrow \begin{Bmatrix} 4.1134 \\ 0.2165 \\ 0.0103 \end{Bmatrix} = 4.1134 \begin{Bmatrix} 1 \\ 0.0526 \\ 0.0025 \end{Bmatrix}$$

As you continue to multiple each successive vector = 4 and the vector $u_k=\{1\ 0\ 0\}^T$

The special advantage is that the eigen vector corresponds to the dominant eigen value and is generated at the same time.

The disadvantage is that the method only supplies obtains only one eigen value

Power Method

There are 2 ways of finding the other eigen values of **A**.

1. Using the Power method to find the eigen value of "**A^{-1}**" which becomes the smallest eigen value of **A**.

$$[A]\{x\} = \lambda\{x\} \Rightarrow [A]^{-1}[A]\{x\} = \lambda[A]^{-1}\{x\} \Longrightarrow \frac{1}{\lambda}\{x\} = [A]^{-1}\{x\} \Rightarrow \mu\{x\} = [B]\{x\}$$

2. Using the Power method to find the largest eigen value of "**A**". Then using the Power method to find the eigen value of "**B**" from

$$\frac{1}{\lambda_B}\{x\} = [[A\} - \lambda_A[I]]^{-1}\ \{x\} \Rightarrow \mu\{x\} = [B]\{x\}$$

Where λ_A is the dominant eigen value of **A**. Then the nearest eigen value of **A** is found from

$$\lambda_{new} = \frac{1}{\mu} + \lambda_A$$

Power Method for Lowest Eigenvalue

Problem Statement. Employ the power method to determine the lowest eigenvalue

Solution.
$$\begin{bmatrix} 3.556 & -1.778 & 0 \\ -1.778 & 3.556 & -1.778 \\ 0 & -1.778 & 3.556 \end{bmatrix}$$
its matrix inverse can be evaluated as

$$[A]^{-1} = \begin{bmatrix} 0.422 & 0.281 & 0.141 \\ 0.281 & 0.562 & 0.281 \\ 0.141 & 0.281 & 0.422 \end{bmatrix}$$

Using the same format the power method can be applied to this matrix.

First iteration:

$$\begin{bmatrix} 0.422 & 0.281 & 0.141 \\ 0.281 & 0.562 & 0.281 \\ 0.141 & 0.281 & 0.422 \end{bmatrix} \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} 0.884 \\ 1.124 \\ 0.884 \end{Bmatrix} = 1.124 \begin{Bmatrix} 0.751 \\ 1 \\ 0.751 \end{Bmatrix}$$

Second iteration:

$$\begin{bmatrix} 0.422 & 0.281 & 0.141 \\ 0.281 & 0.562 & 0.281 \\ 0.141 & 0.281 & 0.422 \end{bmatrix} \begin{Bmatrix} 0.751 \\ 1 \\ 0.751 \end{Bmatrix} = \begin{Bmatrix} 0.704 \\ 0.984 \\ 0.704 \end{Bmatrix} = 0.984 \begin{Bmatrix} 0.715 \\ 1 \\ 0.715 \end{Bmatrix}$$

where $|\varepsilon_a| = 14.6\%$.

Third iteration:

$$\begin{bmatrix} 0.422 & 0.281 & 0.141 \\ 0.281 & 0.562 & 0.281 \\ 0.141 & 0.281 & 0.422 \end{bmatrix} \begin{Bmatrix} 0.715 \\ 1 \\ 0.715 \end{Bmatrix} = \begin{Bmatrix} 0.684 \\ 0.964 \\ 0.684 \end{Bmatrix} = 0.964 \begin{Bmatrix} 0.709 \\ 1 \\ 0.709 \end{Bmatrix}$$

where $|\varepsilon_a| = 4\%$.

Thus, after only three iterations, the result is converging on the value of 0.9602, which is the reciprocal of the smallest eigenvalue, 1.04145

EXAMPLE

To illustrate the utility of shifts, we again look at

$$A = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix}.$$

We now apply the shifted inverse power method, using the same initial guess as before, this time with a shift of $\lambda_* = 2.5$, chosen because we know that the smallest eigenvalue is at least as large as 2. From power method we have that the approximate eigenvalue, after two iterations, was $\frac{1}{\mu_2} = 4.61604920078069$, which is not a very good approximation to the correct value of 2.58578643762691. This time the first two iterations (using the same initial vector as for the unshifted iteration) are

$$y^{(1)} = (A - 2.5I)^{-1} z^{(0)} = \begin{bmatrix} 1.45298541901634 \\ -1.64486177807920 \\ 1.54400744143763 \end{bmatrix}, \quad \mu_1 = -1.64486177807920,$$

$$z^{(1)} = y^{(1)}/\mu_1 = \begin{bmatrix} -0.88334803469813 \\ 1.00000000000000 \\ -0.93868522085829 \end{bmatrix}$$

$$y^{(2)} = (A - 2.5I)^{-1} z^{(1)} = \begin{bmatrix} -9.44765403794921 \\ 13.28813302222568 \\ -9.48454549538931 \end{bmatrix}, \quad \mu_2 = 13.28813302222568,$$

$$z^{(2)} = y^{(2)}/\mu_2 = \begin{bmatrix} -0.71098430623377 \\ 1.00000000000000 \\ -0.71376057716502 \end{bmatrix}.$$

The approximate eigenvalue is now

$$\lambda \approx \lambda_* + \frac{1}{\mu_2} = 2.57525511660121,$$

which is substantially better than we achieved in two iterations with the unshifted method.