Lecture 11 Outline

Sunday, September 29, 2019 11:02 AM

Numerical Solutions

Matrix + Vector Norms

Condition Number

LU Decomposition

Gaussian Elimination

Numerical Solution of Ax=b

Monday, September 30, 2019 3:00 PM

Examine methods to find solutions $\underline{x}$ to

$$\underline{A}\,\underline{x} = \underline{b}$$

Do not find $\underline{A}^{-1}$ explicitly

Use numerical methods (e.g. LU decomposition) to solve

General solution methods:

(1) Exact (to machine precision)

(2) Approximate / iterative (obtain solution to some user defined tolerance)

Before introducing the methods, we need to define what we mean by a solution "close" to the solution → Need matrix & vector norms

Also need to discuss how $\underline{A}$ itself can amplify the error → Need condition number of $\underline{A}$

Matrix & Vector Norms

Sunday, September 29, 2019 11:07 AM

Recall that the 2-norm of a vector is the "length"

$$\|\underline{x}\|_2 = (x_1^2 + x_2^2 + \dots + x_n^2)^{1/2}$$

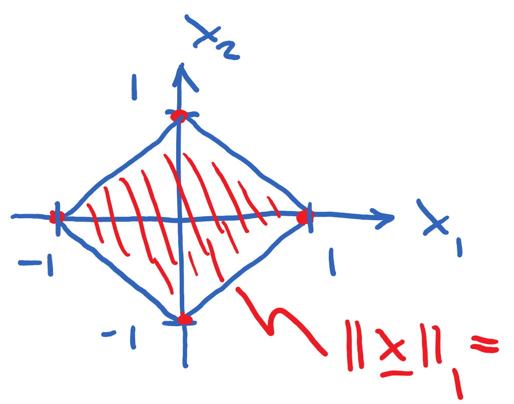
Generalize to the p-norm:

$$\|\underline{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \quad \text{for } 1 \leq p < \infty$$

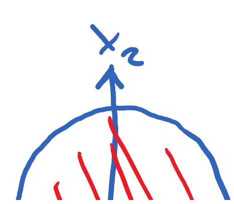
In 2D, all of these norms can be given by areas

Let $\underline{x} = (x_1, x_2)$ with $\cancel{x_1^2 + x_2^2 \leq 1}$ $\|\underline{x}\|_p \leq 1$

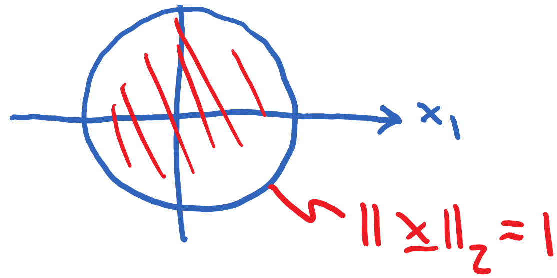
$$\|\underline{x}\|_1 = \sum_{i=1}^{n} |x_i|$$



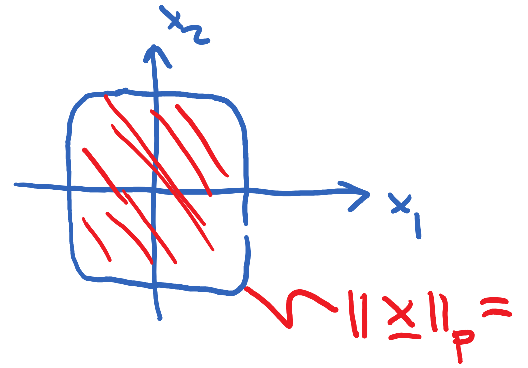
$$\|\underline{x}\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}$$



$$\|\underline{x}\|_2 = \left(\sum_{i=1} |x_i| \right)$$

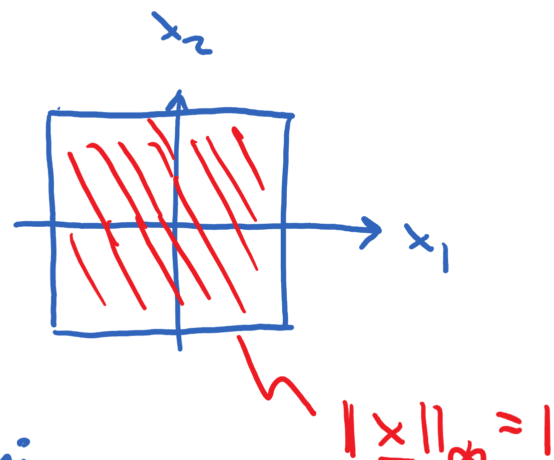


$$\|\underline{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$



Special norm

$$\|\underline{x}\|_\infty = \max |x_i|$$



Each of these norms obey:

1) $\|\underline{x}\| \geq 0$ & $\|\underline{x}\| = 0$ iff $\underline{x} = \underline{0}$

2) $\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\|$

3) $\|\alpha \underline{x}\| = |\alpha| \, \|\underline{x}\|$

Matrix Norms

Vector-induced matrix norms $\rightarrow$ Those that result from the application of a matrix

Let $\underline{A} \in M_{mn}$. The induced matrix norm is the number c, such that

$$\|\underline{A}\,\underline{x}\|_{(m)} \leq c\|\underline{x}\|_{(n)} \text{ for all } \underline{x} \in \mathbb{R}^n$$

Note: $\|\cdot\|_{(m)}$ & $\|\cdot\|_{(n)}$ are not the m-norm or n-norm, but rather the norm in that m or n space.

Example: 1-norm of a matrix

Let $\underline{x} \in \mathbb{R}^n$ such that $\|\underline{x}\|_1 \leq 1$

$$\|\underline{A}\,\underline{x}\|_1 = \left\|\sum_{j=1}^{n} x_j \underline{a}_j\right\|_1 \leq \sum_{j=1}^{n} |x_j| \|\underline{a}_j\|_1$$

$$\leq \max_j \|\underline{a}_j\|_1$$

where $\underline{a}_j$ is the j^{th} column of $\underline{A}$

Let $\underline{x} = \underline{e}_j$, where $\underline{e}_j$ is the vector that maximizes $\|\underline{a}_j\|_1$

$$\Rightarrow \|\underline{A}\|_1 = \max_{1 \leq j \leq n} \|\underline{a}_j\|_1 \leftarrow \text{maximum column sum}$$

Example:

$$\|\underline{A}\|_\infty = \max_{1 \leq j \leq m} \|\underline{a}_j^*\|_1$$

where $\underline{a}_j^*$ is the j^{th} row of $\underline{A}$

$\|\underline{A}\|_\infty$ is the maximum row sum

Other common matrix-norms

Frobenius Norm:

$$\|\underline{A}\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

Matrix Norms also follow:

1) $\|\underline{A}\| \geq 0$, $\|\underline{A}\| = 0$ iff $\underline{A} = \underline{0}$

2) $\|\underline{A} + \underline{B}\| \leq \|\underline{A}\| + \|\underline{B}\|$

3) $\|\alpha \underline{A}\| = |\alpha| \, \|\underline{A}\|$

Example: Matrix Norms

$$\begin{array}{cc} & \Sigma_i \\ \underline{A} = \begin{bmatrix} 5 & 4 & 0 \\ -2 & 1 & 1 \\ 3 & 6 & -1 \\ 0 & -15 & 6 \end{bmatrix} & \begin{array}{c} 9 \\ 4 \\ 10 \\ \boxed{21} \end{array} \\ \Sigma_j \quad \begin{array}{ccc} 10 & \boxed{26} & 8 \end{array} & \end{array}$$

$\therefore \|\underline{A}\|_1 = 26$ Maximum column 1-norm

$\|\underline{A}\|_\infty = 21$ Maximum row 1-norm

Condition Number

Monday, September 30, 2019 3:16 PM

Consider the numerical solution to $\underline{A}\,\underline{x} = \underline{b}$

At a minimum, $\underline{b}$ has some error in it. Thus, we actually are solving

$$\underline{A}\left(\underline{x} + \underline{\Delta x}\right) = \underline{b} + \underline{\Delta b}$$

where $\underline{\Delta b}$ is error in $\underline{b}$

$\underline{\Delta x}$ is error in $\underline{x}$

$$\underline{A}\left(\underline{x} + \underline{\Delta x}\right) = \underline{b} + \underline{\Delta b}$$

$$\cancel{\underline{A}\,\underline{x}} + \underline{A}\,\underline{\Delta x} = \cancel{\underline{b}} + \underline{\Delta b} \quad , \text{ but } \underline{A}\,\underline{x} = \underline{b}$$

$$\Rightarrow \quad \underline{A}\,\underline{\Delta x} = \underline{\Delta b}$$

Assume $\underline{A}^{-1}$ exists $\Rightarrow$ $\underline{\Delta x} = \underline{A}^{-1}\underline{\Delta b}$

$$\|\underline{\Delta x}\| = \|\underline{A}^{-1}\underline{\Delta b}\| \leq \|\underline{A}^{-1}\|\,\|\underline{\Delta b}\| \qquad (a)$$

Now look at $\underline{A}\,\underline{x} = \underline{b}$ in exact math

$$\|\underline{b}\| = \|\underline{A}\,\underline{x}\| \leq \|\underline{A}\|\,\|x\|$$

$$\Rightarrow \frac{1}{\|x\|} \leq \|\underline{A}\| \frac{1}{\|\underline{b}\|} \qquad (b)$$

Normalized error (from (a) + (b))

$$\frac{\|\underline{\Delta x}\|}{\|\underline{x}\|} \leq \|\underline{A}\|\,\|\underline{A}^{-1}\| \frac{\|\underline{\Delta b}\|}{\|\underline{b}\|}$$

If the normalized error of $\underline{b}$ is

$\frac{\|\underline{\Delta b}\|}{\|\underline{b}\|}$, then the normalized error

of the solution will scale by $\|\underline{A}\|\,\|\underline{A}^{-1}\|$

This is the condition number:

$$\kappa(\underline{A}) = \|\underline{A}\|\,\|\underline{A}^{-1}\|$$

Example: Let $\|\underline{\Delta b}\| / \|\underline{b}\| \sim 10^{-16}$, but $\kappa(\underline{A}) \sim 10^{6}$

Then relative error in the solution $\|\underline{\Delta x}\| / \|\underline{x}\| \sim 10^{-10}$, which is six orders higher than the relative error in $\underline{b}$

In general, $\underline{A}$ also will have errors. Then

$$(\underline{A} + \underline{\Delta A})(\underline{x} + \underline{\Delta x}) = \underline{b} + \underline{\Delta b}$$

which leads to relative error in the solution

$$\frac{\|\underline{\Delta x}\|}{\|\underline{x}\|} \lesssim \kappa(\underline{A}) \left(\frac{\|\underline{\Delta b}\|}{\|\underline{b}\|} + \frac{\|\underline{\Delta A}\|}{\|\underline{A}\|} \right)$$

LU Decomposition

Tuesday, October 1, 2019 12:44 AM

Motivate by examining row echelon form

Take a generic square matrix $\underline{A}$ and convert to an upper triangular matrix $\underline{U}$

Example:

$$\underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 6 & 10 \\ 3 & 14 & 28 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 3 & 14 & 28 \end{bmatrix}$$

Each step of this process, called Gaussian elimination, can be written as a matrix-matrix product. Here,

$$\underline{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\underline{E}_1 \underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 3 & 14 & 28 \end{bmatrix}$$

All of the elimination matrices are lower triangular

Characteristics of lower triangular matrices:

(1) Multiplication of lower triangular matrices results in a lower triangular matrix, i.e.,

$$\underline{L}_1 \underline{L}_2 = \underline{L}_3$$

(2) Inverse of a lower triangular matrix is lower triangular

$$\underline{L}_1^{-1} = \underline{L}_4$$

Gaussian elimination is nothing but repeated multiplication by elimination matrices, each of which is lower triangular

$$\underline{E}_n \underline{E}_{n-1} \cdots \underline{E}_3 \underline{E}_2 \underline{E}_1 \underline{A} = \underline{U}$$

row echelon form (upper triangular)

$$\underbrace{\underline{E}_n \underline{E}_{n-1} \cdots \underline{E}_3 \underline{E}_2 \underline{E}_1}_{\underline{L}^{-1} \sim \text{a lower triangular matrix}} \underline{A} = \underline{U} \sim \text{(upper triangular)}$$

$$\underline{L}^{-1} \underline{A} = \underline{U}$$

$$\underline{L}\,\underline{L}^{-1} \underline{A} = \underline{L}\,\underline{U}$$

$$\therefore \underline{A} = \underline{L}\,\underline{U} \Leftarrow \text{LU decomposition!}$$

(1) The diagonal of $\underline{L}$ must have all 1's

Consider

$$\underline{A} = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{bmatrix}$$

$$\left.\begin{aligned} l_{11} u_{11} &= 1 \\ l_{11} u_{12} &= 2 \\ l_{21} u_{11} &= 3 \\ l \; u + l \; u &= 5 \end{aligned}\right\} \text{6 unknowns, only 4 eqns}$$

$$l_{21} u_{12} + l_{22} u_{22} = 5 \quad \Big)$$

Set $l_{11} = 1$, $l_{22} = 1 \Rightarrow$ 6 eqns total

12) Why do we care?

Solve $\underline{A}\,\underline{x} = \underline{b}$ by finding $\underline{A}^{-1}$

Then $\underline{x} = \underline{A}^{-1}\,\underline{b} \leftarrow$ Very expensive!

Look at $\underline{A}\,\underline{x} = \underline{b}$

$$\underline{A}\,\underline{x} = \underline{L}\,\underline{U}\,\underline{x} = \underline{b}$$

$$\underline{x} = (\underline{L}\,\underline{U})^{-1}\,\underline{b}$$

$$\underline{x} = \underbrace{\underline{U}^{-1}\,\underline{L}^{-1}}\,\underline{b}$$

$\underline{U}^{-1}$ + $\underline{L}^{-1}$ are cheap!

because both are triangular

$$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \end{bmatrix} \quad \begin{bmatrix} b_1 \end{bmatrix}$$

$$\underline{L}\,\underline{y} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

(1) Compute $\underline{A} = \underline{L}\,\underline{U}$

(2) Solve $\underline{L}\,\underline{y} = \underline{b} \Rightarrow \underline{y} = \underline{L}^{-1}\underline{b}$

(3) Solve $\underline{U}\,\underline{x} = \underline{y} \Rightarrow \underline{x} = \underline{U}^{-1}\underline{L}^{-1}\underline{b}$

Example: Compute LU of

$$\underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 6 & 10 \\ 3 & 14 & 28 \end{bmatrix}$$

1) Eliminate a_{21}

Multiply $\underline{A}$ by $\underline{E}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$\underline{E}_1\underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 3 & 14 & 28 \end{bmatrix}$$

2) Eliminate location (3,1)

$$\underline{E}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

$$\Rightarrow \underline{E}_2 \underline{E}_1 \underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 8 & 19 \end{bmatrix}$$

3) Eliminate location (3,2)

$$\underline{E}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix}$$

$$\Rightarrow \underline{E}_3 \underline{E}_2 \underline{E}_1 \underline{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix} = \underline{U}$$

$$\underline{A} = (\underline{E}_3 \underline{E}_2 \underline{E}_1)^{-1} \underline{U}$$

$$\underline{A} = \underline{E}_1^{-1} \underline{E}_2^{-1} \underline{E}_3^{-1} U$$

where

$$\underline{E}_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ +2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\underline{E}_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +3 & 0 & 1 \end{bmatrix}$$

$$\underline{E}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & +4 & 1 \end{bmatrix}$$

$$\therefore \underline{E}_1^{-1} \underline{E}_2^{-1} \underline{E}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix}$$

The negative of the operation,
<u>or</u> factor of the value divided
by the pivot

$$\underline{A} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix}}_{\underline{L}} \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix}}_{\underline{U}}$$