2/19/2019

Bisection and Gaussian elimination

Bisection:

~~[illegible]~~:

Also used in finding a root of f., $f(x) = 0$.

Idea: Start with some initial interval in which we know exists a root by the ~~[illegible]~~ intermediate value theorem.

i.e. we have (a_0, b_0) so that

$$f(a_0)\, f(b_0) < 0.$$

$\iff$ $f(a_0)$ and $f(b_0)$ have opposite sign.

Define two new intervals: (a_1, c_1), (c_1, b_1)

with $a_1 = a_0$, $c_1 = \frac{1}{2}(a_0 + b_0)$, $b_1 = b_0$.

$$\text{Then } (a_2, b_2) = \begin{cases} (a_1, c_1) & \text{if } \cancel{f(a_1)f(c_1) < 0}\ \ f(c_1)f(b_1) > 0 \\ (c_1, b_1) & \text{if } f(c_1)f(b_1) < 0. \end{cases}$$

Gaussian Elimination

Very simple example:

$$x_1 + x_2 = 1$$
$$x_1 - x_2 = 0.$$

To solve, we would add the equations to get:

$$2x_1 = 1. \iff x_1 = \frac{1}{2}.$$

We can plug $x_1 = \frac{1}{2}$ back into either equation to get $x_2 = \frac{1}{2}$.

This can be formulated in matrix form:

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}}_{\underline{\underline{A}}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{\underline{x}} = \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{=\underline{b}}.$$

Gaussian elimination can be thought of as a process through which we transform $\underline{\underline{A}}$ to <u>upper triangle form</u>.

Definition 2.2 Let n be an integer ≥ 2.

The matrix $\underline{\underline{L}} \in \mathbb{R}^{n \times n}$ is <u>lower triangular</u> if $\underline{\underline{L}} = (l_{ij})$ satisfies:

$$l_{ij} = 0 \quad \forall i,j \text{ with } 1 \leq i < j \leq n.$$

$\underline{\underline{L}}$ is called <u>unit lower triangular</u> if it is lower triangular and all its diagonal elements are equal to 1.

$\underline{\underline{U}}$ is called <u>upper triangular</u> if $\underline{\underline{U}}^T$ is lower triangular.

Ex:

$$\begin{pmatrix} 0 & 5 & 6 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{pmatrix}. \qquad \begin{pmatrix} 1 & 5 & 6 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$$

~~lower~~ upper Δ — unit upper Δ.

Back to the example:

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

If we try to put A is upper Δ form, we can multiply the 1st equation by -1 and add it to the second equation:

$$\begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Now, the second equation reads:

$$-2x_2 = -1$$

implying $x_2 = \frac{1}{2}$. We can then "back-substitute x_2 into the 1st equation:

$$x_1 + x_2 = 1 \Leftrightarrow x_1 + \frac{1}{2} = 1$$

$$\Leftrightarrow x_1 = \frac{1}{2},$$

to get x_1. This is a simple example showing the process of backsubstitution and Gaussian elimination.

Gaussian Elimination : Bigger Example:

History : Introduced by Gauss in 1809 to solve six linear equations with six unknowns. This was a model for the motion of the asteroid Pallas.

Example:

$$x_1 + x_2 + x_3 = 6$$

$$2x_1 + 4x_2 + 2x_3 = 16$$

$$-x_1 + 5x_2 - 4x_3 = -3.$$

We can rewrite this in matrix form

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 16 \\ -3 \end{pmatrix}.$$

~~Multiply 3rd row by 2:~~

$$\cancel{\begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -2 & 10 & -8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ 6 \end{pmatrix}}$$

$x + y = 1$
$x - y = 0.$

$2x = 1.$
$x = \frac{1}{2}.$

Multiply 1st row by -2:

$$\begin{pmatrix} -2 & -2 & -2 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -12 \\ 16 \\ -3 \end{pmatrix}$$

Add 1st and second rows, and put result in the second row

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ -1 & 5 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ -3 \end{pmatrix}$$

Note that above we went ahead and also multiplied the 1st row by $-1/2$.

Now add the 1st row to the 3rd row, and put result in 3rd row:

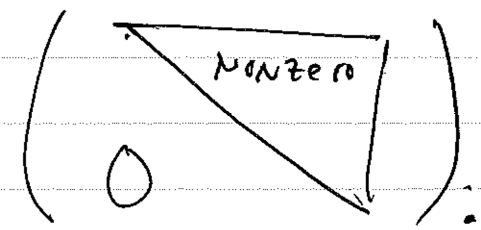
$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 6 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ 3 \end{pmatrix}.$$

Finally, take the second row, multiply by -3, add to the third row, and put the result in the 3rd row:

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ -9 \end{pmatrix}.$$

To solve, we can now use "back substitution."

First, notice the structure of the resulting matrix: it is called "upper triangular"



Let's go back and look at the first operation we did: Multiply first row by -2 and add to second row: This can be expressed in terms of a matrix:

$$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

i.e. $$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & 5 & -4 \end{pmatrix}$$

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

Notice that $\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is "lower triangular."

We can write this sequence of row operations as a product of lower Δ matrices, $L_1, L_2, L_3, \ldots$

$$L_1 L_2 L_3 A x = L_1 L_2 L_3 b$$

where we showed in our example that

$$L_1 L_2 L_3 A$$

is an upper triangular matrix.

We can move the $L_1 L_2 L_3$ multiplying b to the left hand side by ~~[illegible]~~ multiplying by their inverses:

$$L_3^{-1} L_2^{-1} L_1^{-1} L_1 L_2 L_3 A x = b$$

$$\Leftrightarrow (L_1 L_2 L_3)^{-1} (L_1 L_2 L_3 A) x = b.$$

What is the structure of $(L_1 L_2 L_3)^{-1}$?

Exercise: Show $(L_1 L_2 L_3)^{-1}$ is lower triangular.

We've Factored $A = (L_1 L_2 L_3)^{-1} (L_1 L_2 L_3 A)$

Into the product of a lower Δ and upper Δ Matrix in the process of solving the linear system. This is called the LU Factorization of A.

Thm: 2.1 These are true for any integer $n \geq 2$.

(i) The product of two lower Δ matrices of order n is lower Δ of order n.

(ii) The product of two unit lower Δ matrices of order n is unit lower Δ of order n.

(iii) a lower Δ matrix is nonsingular $\Leftrightarrow$ all the diagonal elements are nonzero.

(iv) The inverse of a nonsingular lower Δ matrix of order n is lower Δ of order n.

(v) The inverse of a unit lower Δ matrix of order n is unit lower Δ of order n.

Pf of iv (By induction).

Base case: $n=2$.

let $L = \begin{pmatrix} a & 0 \\ c & b \end{pmatrix}$. Note that we can write down L^{-1} explicitly:

$$L^{-1} = \frac{1}{\det L} \begin{pmatrix} b & 0 \\ -c & a \end{pmatrix}.$$

Since $\det L = ab$, we are implicitly assuming $a \neq 0$ and $b \neq 0$. Thus, L^{-1} is lower Δ.

Inductive hypothesis: Assume true for $2 \leq n < k$.

i.e. the inverse of a nonsingular lower Δ matrix of order k is lower Δ of order k.

Prove true for $k+1$. Let L be a lower Δ triangular matrix of order $k+1$.

We can partition L and its inverse L^{-1} as:

$$L = \begin{pmatrix} L_1 & \underline{0} \\ \underline{r}^T & \alpha \end{pmatrix}, \quad L^{-1} = \begin{pmatrix} X & \underline{y} \\ \underline{z}^T & \beta \end{pmatrix}.$$

By definition, $L L^{-1} = I_{k+1}$, so

$$L_1 X = I_k, \quad L_1 \underline{y} = \underline{0}, \quad r^T X + \alpha z^T = \underline{0}^T$$

$$\text{and} \quad \underline{r}^T \underline{y} + \alpha \beta = 1.$$

By induction, since $X = L_1^{-1}$, X is lower Δ. Also, since L_1 is invertible, $\underline{y} = \underline{0}$. This shows L^{-1} is lower Δ.

Some comments:

Rmk: In our first example, we reduced the system of equations as:

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}}_{A}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\downarrow$$

$$\begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Equivalently, we multiplied A by $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} A = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}.$$

So we have $L_1 A = U_1$, where

$$L_1 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad U_1 = \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix}.$$

Now, we can write $A = L_1^{-1} U_1$,
with $L_1^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

$$A = L_1^{-1} U_1' = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix},$$

This is $\underline{\underline{\text{an}}}$ LU factorization ~~for~~ A,

but it is not unique! In particular,
we can squeeze in $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{a} \end{pmatrix}$
For $a \neq 0$ to get

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

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

$$= \begin{pmatrix} a & 0 \\ a & a \end{pmatrix} \begin{pmatrix} \frac{1}{a} & \frac{1}{a} \\ 0 & -\frac{2}{a} \end{pmatrix}.$$

<u>Fact</u>: (I think, please check ü). An LU factorization is unique if L is required to be $\underline{\underline{\text{unit}}}$ lower Δ.

Base on the nonuniquess of a factorization of A into ~~uppe~~ a product of upper Δ and lower Δ matrices, we explictly define the (unique) LU Factorization of A as follows:

Definition The **LU Factorization** of A, if it exists, is the product $A = LU$, where L is unit lower Δ and U is upper Δ.

Rmk: let's say we are given an arbitrary matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and we want to transform this to upper Δ, using the techniques just described.
We can multiply the 1^{st} row by $-\frac{c}{a}$ and add to the second row:

$$\begin{pmatrix} 1 & 0 \\ -\frac{c}{a} & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & -\frac{c}{a}(b+d) \end{pmatrix}.$$

But, what if $a \approx 0$. This is not stable. (we will define "not stable" later in the class).

It suffices to say that we need to be careful about dividing by numbers close to zero when we do Gaussian elimination.

Example: $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.

If we try to use (blindly) our process to convert A to an upper Δ matrix, we are in trouble since $a_{11} = 0$.

But, we can multiply A by a <u>permutation matrix</u> P which rearranges the rows.

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

$$PA = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Trivially, we have $PA = LU$,

where $L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

<u>Thm</u>: let $n \geq 2$ and $A \in \mathbb{R}^{n \times n}$. There exists a permutation matrix P, a unit lower Δ matrix L, and an upper Δ matrix U, so that $P, L, U \in \mathbb{R}^{n \times n}$ and

$$PA = LU.$$