Linear equation systems: exact methods

Marcin Chrząszcz, Danny van Dyk
mchrzasz@cern.ch,
danny.van.dyk@gmail.com

University of Zurich[UZH]

Numerical Methods,
10 October, 2016

Linear eq. system

$\Rightarrow$ This and the next lecture will focus on a well known problem. Solve the following equation system:

$$A \cdot x = b,$$

$\Rightarrow$ $A = a_{ij} \in \mathbb{R}^{n \times n}$ and $\det(A) \neq 0$
$\Rightarrow$ $b = b_i \in \mathbb{R}^n$.
$\Rightarrow$ The problem: Find the x vector.

Error digression

$\Rrightarrow$ There is enormous amount of ways to solve the linear equation system.
$\Rrightarrow$ The choice of one over the other of them should be gathered by the *condition* of the matrix A denoted at $cond(A)$. $\Rrightarrow$ If the $cond(A)$ is small we say that the problem is well conditioned, otherwise we say it's ill conditioned.
$\Rrightarrow$ The *condition* relation is defined as:

$$cond(A) = \|A\| \cdot \|A^{-1}\|$$

$\Rrightarrow$ Now there are many definitions of different norms... The most popular one (so-called "column norm"):

$$\|A\|_1 = \max_{1 \leqslant j \leqslant n} \sum_{i=1}^{n} |a_{i,j}|,$$

where n -is the dimension of A, i, j are columns and rows numbers.

More norms

⇛ A different norm is a spectral norm:

$$\|A\|_2 = \sqrt{\rho(A^T A)}$$
$$\rho(M) = \max\{|\lambda_i| : \det M - \lambda I = 0,\ i = 1, ...n\}$$

where $\rho(M)$ - spectral radius of M matrix, I unit matrix, λ_i eigenvalues of M.

⇛ Row norm:

$$\|A\|_\infty = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{i,j}|,$$

Digression:

⇛ Calculation of the matrix norms are not a simple process at all. There are certain class of matrices that make the calculations easier.

⇛ The spectral norm can be also defined:

$$cond_2(A) = \frac{\max_{1 \leqslant i \leqslant n} |\lambda_i|}{\min_{1 \leqslant i \leqslant n} |\lambda_i|},$$

Example, ill-conditioned matrix

⇛ The text-book example of wrongly conditioned matrix is the Hilbert matrix:

$$h_{i,j} = \frac{1}{i+j-1}$$

⇛ Example:

$$h_{i,j}^{4\times 4} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}$$

⇛ The condition of this matrix:

$$cond(A) = \mathcal{O}\left(\frac{e^{3.52N}}{\sqrt{N}}\right)$$

⇛ For 8×8 matrix we get:

$$cond_1(A) = 3.387 \cdot 10^{10}, \quad cond_2(A) = 1.526 \cdot 10^{10}, \quad cond_\infty(A) = 3.387 \cdot 10^{10}$$

⇛ Clearly large numbers ;)

Exact methods: Cramer method

⇛ If $\det A \neq 0$ then the solutions are given by:

$$x_i = \frac{\det A_i}{\det A}$$

⇛ So calculate the solutions one needs to calculate $n+1$ determinants. To calculate each determinate one needs $(n-1)n!$ multiplications.
⇛ Putting it all together one needs $(n+1)(n-1)n! = n^{n+2}$
⇛ Brute force but works ;)

Example, ill-conditioned matrix

⇛ The text-book example of wrongly conditioned matrix is the Hilbert matrix:

$$h_{i,j} = \frac{1}{i+j-1}$$

⇛ Example:

$$h_{i,j}^{4\times 4} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}$$

⇛ The condition of this matrix:

$$cond(A) = \mathcal{O}\left(\frac{e^{3.52N}}{\sqrt{N}}\right)$$

⇛ For 8×8 matrix we get:

$$cond_1(A) = 3.387 \cdot 10^{10}, \quad cond_2(A) = 1.526 \cdot 10^{10}, \quad cond_\infty(A) = 3.387 \cdot 10^{10}$$

⇛ Clearly large numbers ;)

Exact methods: Cramer method

$\Rrightarrow$ If $\det A \neq 0$ then the solutions are given by:

$$x_i = \frac{\det A_i}{\det A}$$

$\Rrightarrow$ So calculate the solutions one needs to calculate $n+1$ determinants. To calculate each determinate one needs $(n-1)n!$ multiplications.
$\Rrightarrow$ Putting it all together one needs $(n+1)(n-1)n! = n^{n+2}$
$\Rrightarrow$ Brutal force but works ;)

Exact methods: Gauss method

$\Rrightarrow$ The idea besides the Gauss method is simple: transform the $Ax = b$ to get the equvalent matrix $A^{[n]}x = b^{[n]}$ where $A^{[n]}$ is triangular matrix:

$$A^{[n]} = \begin{pmatrix} a_{11}^{[n]} & a^{[n]}{}_{12} & ... & a_{1n}^{[n]} \\ 0 & a_{22}^{[n]} & ... & a_{2n}^{[n]} \\ ... & & & \\ 0 & 0 & ... & a_{nn}^{[n]} \end{pmatrix}$$

$\Rrightarrow$ The algorithm: $\Rrightarrow$ To do so we calculate the: $d_{i,1}^{[1]} = \frac{a_{i1}^{[1]}}{a_{11}^{[1]}}$

$\Rrightarrow$ The first row multiplied by the $d_{i,1}^{[1]}$ we subtract from the i^{th} row. $\Rrightarrow$ After this we get:

$$\begin{pmatrix} a_{11}^{[1]} & a^{[1]}{}_{12} & ... & a_{1n}^{[1]} \\ 0 & a_{22}^{[1]} & ... & a_{2n}^{[1]} \\ ... & & & \\ 0 & a_{n2}^{[1]} & ... & a_{nn}^{[1]} \end{pmatrix} \overrightarrow{x} = \begin{pmatrix} b_1^{[1]} \\ b_1^{[1]} \\ ... \\ b_1^{[1]} \end{pmatrix}$$

Exact methods: Gauss method 2

⇛ Now one needs to repeat the above n times moving each time row down.

Cons:

⇛ The algoright can be stooped if you divide by zero.
⇛ The method is very efficient to accumulate numerical errors.

Pros:

⇛ The number of needed floating point operations is less then Cramer.
⇛ Example for 15 equations: 1345 vs $5 \cdot 10^{12}$.

Exact methods: modified Gauss method

⇛ The biggest disadvantage of Gauss method is the fact that we can have zero elements :(
⇛ The modified method fixes this problem :)
⇛ The modification is as follows:

- In each step before we do the elimination we look for main element:

$$|a_{mk}^{|k|} = \max\{|a_{jk}^{|k|} : j = k, k+1, ..., n\}$$

- We exchange the rows m and k.
- We do the standard elimination.

Exact methods: Jordan elimination method

$\Rrightarrow$ The Jordan elimination method is similar to Guass method but the idea is to transform the matrix from $Ax = b$ to $Ix = b^{\lfloor n+1 \rfloor}$.

1. We start from eliminating the as in Gauss method in the first row.
2. When we move to the second row we eliminate the x_2 element also from the first raw.
3. The third raw we eliminate the first and second raw.
4. We repeat this n times.

$\Rrightarrow$ After this we will get new system $Ix = b^{\lfloor n+1 \rfloor}$.
$\Rrightarrow$ The $b^{\lfloor n+1 \rfloor}$ is already the solution! No need to do more.

Exact methods: LU method

⇛ The most popular of the exact methods is so-called LU method.

⇛ The idea is very simple; we represent the matrix in a form:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{2n} \\ & & \dots & & \\ {}_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ l_{21} & 1 & 0 & \dots & 0 \\ l_{31} & l_{32} & 1 & \dots & 0 \\ & & \dots & & \\ l_{n1} & l_{n2} & l_{n3} & \dots & l_{nn} \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u12 & u_{13} & \dots & u_{1n} \\ 0 & u22 & u_{23} & \dots & u_{2n} \\ 0 & 0 & u_{33} & \dots & u_{3n} \\ & & \dots & & \\ 0 & 0 & 0 & \dots & u_{nn} \end{pmatrix} = L \cdot U$$

⇛ After this decomposition we need to solve:

$$\begin{cases} Ly & = b \\ Ux & = y \end{cases}$$

Exact methods: LU method, algorithms

⇛ We solve the following matrix: $A^{\lfloor 1 \rfloor} x = b^{\lfloor 1 \rfloor}$

⇛ We start by preparing the matrix $L^{\lfloor 1 \rfloor}$:

$$L^{\lfloor 1 \rfloor} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -l_{21} & 1 & 0 & \dots & 0 \\ -l_{31} & 0 & 1 & \dots & 0 \\ & & \dots & & \\ -l_{n1} & 0 & 0 & \dots & 1 \end{pmatrix}$$

where the l_{i1} are defined:

$$l_{i1} = \frac{a_{i1}^{\lfloor 1 \rfloor}}{a_{11}^{\lfloor 1 \rfloor}}$$

⇛ Now we take our base equation by $L^{\lfloor 1 \rfloor}$:

$$L^{\lfloor 1 \rfloor} A^{\lfloor 1 \rfloor} x = L^{\lfloor 1 \rfloor} b^{\lfloor 1 \rfloor}$$

⇛ We get the a new system:

$$A^{\lfloor 2 \rfloor} x = b^{\lfloor 2 \rfloor} \qquad \Leftrightarrow \qquad L^{\lfloor 1 \rfloor} A^{\lfloor 1 \rfloor} x = L^{\lfloor 1 \rfloor} b^{\lfloor 1 \rfloor}$$

Exact methods: LU method, algorithms

$\Rrightarrow$ In the second step we construct the $L^{\lfloor 2 \rfloor}$ in the form:

$$L^{\lfloor 2 \rfloor} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & -l_{32} & 1 & \dots & 0 \\ & & \dots & & \\ 0 & -l_{n2} & 0 & \dots & 1 \end{pmatrix}$$

where :

$$l_{i2} = \frac{a_{i2}^{\lfloor 2 \rfloor}}{a_{22}^{\lfloor 2 \rfloor}}$$

$\Rrightarrow$ Now we take the $A^{\lfloor 2 \rfloor} x = b^{\lfloor 2 \rfloor}$ and multiply it by $L^{\lfloor 2 \rfloor}$:

$$L^{\lfloor 2 \rfloor} A^{\lfloor 2 \rfloor} x = L^{\lfloor 2 \rfloor} b^{\lfloor 2 \rfloor}$$

$\Rrightarrow$ We get the a new system:

$$A^{\lfloor 3 \rfloor} x = b^{\lfloor 3 \rfloor} \qquad \Leftrightarrow \qquad L^{\lfloor 2 \rfloor} L^{\lfloor 1 \rfloor} A^{\lfloor 1 \rfloor} x = L^{\lfloor 2 \rfloor} L^{\lfloor 1 \rfloor} b^{\lfloor 1 \rfloor}$$

Exact methods: LU method, algorithms

⇛ Now the we repeat the above steps $n-1$ times after which we get:

$$L^{\lfloor n-1 \rfloor} L^{\lfloor n-2 \rfloor} ... L^{\lfloor 2 \rfloor} L^{\lfloor 1 \rfloor} A^{\lfloor 1 \rfloor} = A^{\lfloor n \rfloor} = U$$

$$L^{\lfloor n-1 \rfloor} L^{\lfloor n-2 \rfloor} ... L^{\lfloor 2 \rfloor} L^{\lfloor 1 \rfloor} b^{\lfloor 1 \rfloor} = b^{\lfloor n \rfloor}$$

⇛ From the above we can calculate:

$$A^{\lfloor 1 \rfloor} = \left(L^{\lfloor 1 \rfloor}\right)^{-1} \left(L^{\lfloor 2 \rfloor}\right)^{-1} ... \left(L^{\lfloor n-1 \rfloor}\right)^{-1} A^{\lfloor n \rfloor}$$

⇛ So the matrix L we search for is:

$$L = \left(L^{\lfloor 1 \rfloor}\right)^{-1} \left(L^{\lfloor 2 \rfloor}\right)^{-1} ... \left(L^{\lfloor n-1 \rfloor}\right)^{-1}$$

⇛ The $\left(L^{\lfloor k \rfloor}\right)^{-1}$ can be easy calculated:

$$\left(L^{\lfloor k \rfloor}\right)^{-1} = \begin{pmatrix} 1 & 0 & ... & 0 & ... & 0 \\ 0 & 1 & ... & 0 & ... & 0 \\ & & & ... & & \\ 0 & 0 & ... & 1 & ... & 0 \\ 0 & 0 & ... & l_{k+1k} & ... & 0 \\ 0 & 0 & ... & l_{k+2k} & ... & 0 \\ & & & ... & & \\ 0 & 0 & ... & l_{nk} & ... & 1 \end{pmatrix}$$

Exact methods: LU method, algorithms

⇛ Now the only thing left is to solve the simple linear system:

$$\begin{cases} Ly &= b \\ Ux &= y \end{cases}$$

⇛ Because of the triangular matrix the solution is straightforward:

$$y_1 = \frac{b_1}{L_1 1}, \qquad y_i = \frac{b_1 - \sum_{j=1^{i-1}} L_{ij} y_j}{L_{ii}} \quad i \geqslant 2$$

Summary

- $\Rrightarrow$ This lecture we learn the exact methods of solving linear equation system.
- $\Rrightarrow$ The three most popular one are: Gauss, Jordan, LU.
- $\Rrightarrow$ By default the LU method should be used.
- $\Rrightarrow$ And remember: be sure the system is well conditioned!

Backup