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

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

$$A \cdot x = b,$$

⇛ $A = a_{ij} \in \mathbb{R}^{n \times n}$ and $\det(A) \neq 0$
⇛ $b = b_i \in \mathbb{R}^n$.
⇛ 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 radios 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:

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

⇛ 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 $(n+1)(n-1)n! = n^{n+2}$
⇛ Brutal 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

⇛ 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:

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

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

⇛ The first row multiplied by the $d_{i,1}^{[1]}$ we subtract from the i^{th} row. ⇛ 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 $5 \cdot 10^{12}$.

Backup