Numerical methods

There are two classes of methods for solving the linear system

$$A\underline{x} = \underline{b} \tag{1}$$

Direct methods: they give the exact solution (up to computer precision) in

a finite number of operations.

Iterative methods: starting from an initial guess $\underline{x}^{(0)}$ they construct a sequence $\{\underline{x}^{(k)}\}$ such that

$$\underline{x} = \lim_{k \to \infty} \underline{x}^{(k)}.$$

Direct methods

The most important direct method is GEM (Gaussian elimination method).

With GEM, the original system (1) is transformed into an equivalent system (i.e., having the same solution)

$$U\underline{x} = \widetilde{\underline{b}} \qquad (*)$$

with U upper triangular matrix. This is done in $n-1$ steps where, at each step, one column is eliminated, that is, all the coefficients of that column below the diagonal are transformed into zeros. The cost for computing U is of the order of $\frac{2}{3}n^3$, see later... the cost for computing $\widetilde{b}$ is of the order of n^2; the cost for solving the upper triangular system is n^2, so that the total cost is $\sim \frac{2}{3}n^3 + 2n^2 \sim \frac{2}{3}n^3$.

GEM algorithm

$A^{(1)} := A,\ b^{(1)} := b$

$$A^{(1)}x = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1^{(1)} \\ b_2^{(1)} \\ \vdots \\ b_n^{(1)} \end{bmatrix} = b^{(1)}$$

- We eliminate the first column:

$$A^{(2)}x = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(2)} & \dots & a_{nn}^{(2)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(2)} \end{bmatrix} = b^{(2)}$$

$$l_{i1} = a_{i1}^{(1)} / a_{11}^{(1)}, \quad a_{ij}^{(2)} = a_{ij}^{(1)} - l_{i1} a_{1j}^{(1)}, \quad b_i^{(2)} = b_i^{(1)} - l_{i1} b_1^{(1)}$$

GEM algorithm

- We go on eliminating columns in the same way. At the $k-$th step:

$$A^{(k)}x = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & \dots & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & & & & a_{2n}^{(2)} \\ \vdots & & \ddots & & & \vdots \\ 0 & \dots & 0 & a_{kk}^{(k)} & \dots & a_{kn}^{(k)} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{nk}^{(k)} & \dots & a_{nn}^{(k)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_k^{(k)} \\ \vdots \\ b_n^{(k)} \end{bmatrix} = b^{(k)}$$

- At the last step $A^{(n)}x = b^{(n)}$ with U upper triangular,

$$\Longrightarrow U := A^{(n)},\ \widetilde{b} := b^{(n)}$$

GEM pseudocode

Input: $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$

for $k = 1, \ldots, n-1$

  for $i = k+1, \ldots, n$

    $l_{i,k} = a_{i,k}/a_{k,k}$

    $a_{i,k:n} = a_{i,k:n} - l_{i,k}a_{k,k:n}$

    $b_i = b_i - l_{i,k}b_k$

  end

end

define $U = A$, $\widetilde{b} = b$, then solve $Ux = \widetilde{b}$ with back substitution

remark: we do not need to define U and $\widetilde{b}$, it is just to be consistent with the notation of the previous slides

Homework

Write the above algorithm in MATLAB

Computational cost of GEM

Let us start by analysing the inner loop:

for $i = k+1, \dots, n$

$l_{i,k} = a_{i,k}/a_{k,k}$ — 1 FLOP

$a_{i,k:n} = a_{i,k:n} - l_{i,k} a_{k,k:n}$ — $2(n-k+1) \sim 2(n-k)$ FLOPs

$b_i = b_i - l_{i,k} b_k$ — 2 FLOPs

end

Inner loop cost $\sim 2(n-k)^2$ FLOPs

Let us focus on the outer loop:

- For $k = 1$, the inner loop costs $\sim 2(n-1)(n-1) \sim 2(n-1)^2$ FLOPs
- For $k = 2$, the inner loop costs $\sim 2(n-2)(n-2) \sim 2(n-2)^2$ FLOPs
- $\dots$
- For $k = n-1$, the inner loop costs $\sim 2 \sim 2 \cdot (1)^2$ FLOPs

Finally, recalling that:

$$\sum_{t=1}^{T} t^2 = \frac{T \cdot (T+1) \cdot (2T+1)}{6},$$

the cost of GEM is:

$$\sim 2 \cdot \sum_{s=1}^{n-1} s^2 = 2\frac{(n-1) \cdot n \cdot (2(n-1)+1)}{6} \sim \frac{2}{3} n^3 \text{ FLOPs}$$

If we also include the cost for solving the resulting upper triangular system $U\underline{x} = \underline{\widetilde{b}}$, the total cost is $\sim \frac{2}{3}n^3 + n^2 \sim \frac{2}{3}n^3$ FLOPs.