Iterative methods for linear algebra

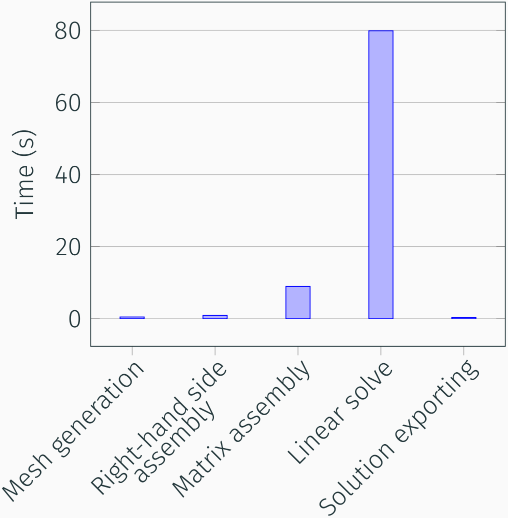
Pierre Jolivet `pierre@joliv.et`
Ronan Guivarch `ronan.guivarch@toulouse-inp.fr`

1/2 *Basic iterative methods*

Table of contents

Introduction

The big picture of implicit methods



Mathematically speaking

Projection methods Relaxation-based methods		Multifrontal factorizations Supernodal factorizations
Iterative methods		**Direct methods**
Sparse MV & dot products	Level of parallelism	BLAS 3
Low	Memory consumption	High
✗	Robustness w.r.t. $\kappa(A)$	✓
✓ (if need be)	Intervention of the user	✗

Mathematically speaking

Domain decomposition methods

Multigrid methods

“Hybrid” methods

Projection methods

Relaxation-based methods

Iterative methods

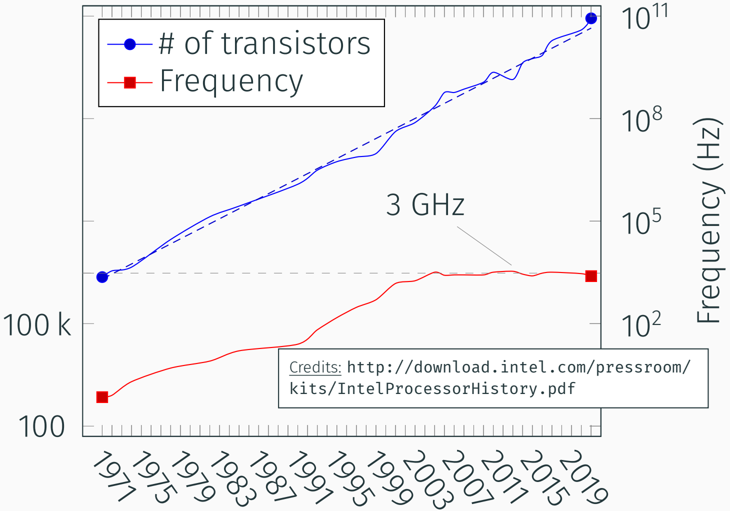
Multifrontal factorizations

Supernodal factorizations

Direct methods

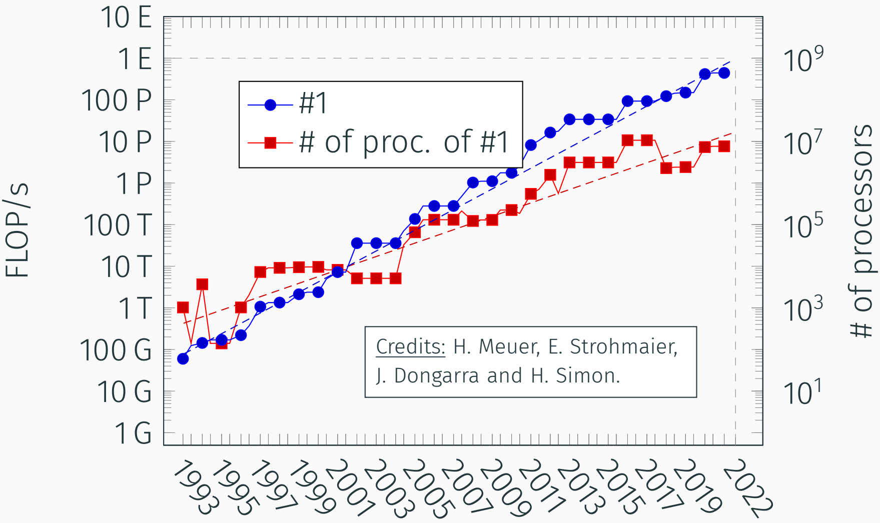
Sparse MV & dot products	Level of parallelism	BLAS 3
Low	Memory consumption	High
✗	Robustness w.r.t. $\kappa(A)$	✔
✔ (if need be)	Intervention of the user	✗

f-SCALING



Credits: `http://download.intel.com/pressroom/kits/IntelProcessorHistory.pdf`

TOP500



10 E
1 E
100 P
10 P
1 P
100 T
10 T
1 T
100 G
10 G
1 G
FLOP/s
10^9
10^7
10^5
10^3
10^1
# of processors
#1
# of proc. of #1
Credits: H. Meuer, E. Strohmaier,
J. Dongarra and H. Simon.
1993
1995
1997
1999
2001
2003
2005
2007
2009
2011
2013
2015
2017
2019
2022

Stationary iterative methods

Motivation

Let us consider the system $Ax = f$

It is assumed that it is too expensive to compute $x = A^{-1}f$

Motivation

Let us consider the system $Ax = f$

It is assumed that it is too expensive to compute $x = A^{-1}f$

Let x_k be an approximate solution, then: $x = x_k + e_k$

What equation does the error verify?

Motivation

Let us consider the system $Ax = f$

It is assumed that it is too expensive to compute $x = A^{-1}f$

Let x_k be an approximate solution, then: $x = x_k + e_k$

What equation does the error verify?

$$\begin{aligned} Ae_k &= f - Ax_k \\ &= r_k \end{aligned}$$

Let us consider the system $Ax = f$

It is assumed that it is too expensive to compute $x = A^{-1}f$

Let x_k be an approximate solution, then: $x = x_k + e_k$

What equation does the error verify?

$$\begin{aligned} Ae_k &= f - Ax_k \\ &= r_k \end{aligned}$$

Thus, $x = x_k + A^{-1}r_k$

Basic preconditioning

Basic principles

The equation $x = x_k + A^{-1} r_k$ cannot be computed

Basic principles

The equation $x = x_k + A^{-1} r_k$ cannot be computed

Replace A^{-1} by M^{-1}, where M is a suitable preconditioner:

$$\text{given } x_0 \qquad r_k = f - A x_k$$

$$x_{k+1} = x_k + M^{-1} r_k$$

Basic principles

The equation $x = x_k + A^{-1} r_k$ cannot be computed

Replace A^{-1} by M^{-1}, where M is a suitable preconditioner:

$$\text{given } x_0 \qquad r_k = f - Ax_k$$
$$x_{k+1} = x_k + M^{-1} r_k$$

The preconditioner M may be:

- explicit (M^{-1} is known, e.g., SPAI)
- implicit (action of M^{-1} on a vector is known, e.g., MG)

Basic principles

The equation $x = x_k + A^{-1} r_k$ cannot be computed

Replace A^{-1} by M^{-1}, where M is a suitable preconditioner:

$$\text{given } x_0 \qquad \begin{aligned} r_k &= f - Ax_k \\ x_{k+1} &= x_k + M^{-1} r_k \end{aligned}$$

The preconditioner M may be:

- explicit (M^{-1} is known, e.g., SPAI)
- implicit (action of M^{-1} on a vector is known, e.g., MG)

Matrix splitting

$$\text{given } x_0 \qquad x_{k+1} = (I - M^{-1}A)x_k + M^{-1}f$$

Jacobi method

Approximate $A = L + D + U$ by its diagonal, $M^{-1} = D^{-1}$

$$\text{given } x_0 \qquad x_{k+1} = -D^{-1}(L+U)x_k + D^{-1}f$$

The ith coefficient of x_{k+1} is given by:

$$x_{k+1_i} = \frac{1}{A_{ii}}\left(f_i - \sum_{\substack{j=1\\ j\neq i}}^{n} A_{ij}x_{k_j}\right)$$

Gauss–Seidel method

Approximate A by its diagonal + lower triang. $M^{-1} = (L + D)^{-1}$

$$\text{given } x_0 \qquad x_{k+1} = (L + D)^{-1}(f - Ux_k)$$

The ith coefficient of x_{k+1} is given by:

$$x_{k+1_i} = \frac{1}{A_{ii}} \left(f_i - \sum_{1 \leqslant j < i}^{n} A_{ij} x_{k+1_j} - \sum_{i < j \leqslant n}^{n} A_{ij} x_{k_j} \right)$$

Krylov methods

Cayley–Hamilton theorem

The characteristic polynomial χ of A defined as:

$$\chi(\lambda) = \det(\lambda I_n - A)$$

verifies $\chi(A) = 0$

Cayley–Hamilton theorem

The characteristic polynomial χ of A defined as:

$$\chi(\lambda) = \det(\lambda I_n - A)$$

verifies $\chi(A) = 0$

This may be used to show that $\exists!\ p \in \mathbb{K}_{n-1}[X] : p(A) = A^{-1}$

Indeed, $p(X) = \dfrac{(-1)^{n+1}}{\det(A)} q(X) : \chi(X) = (-1)^n \det(A) I_n + Xq(X)$

Krylov subspaces

Given x_0, the solution verifies:

$$\begin{aligned} x &= x_0 + e_0 \\ &= x_0 + A^{-1} r_0 \\ &= x_0 + p(A) r_0 \end{aligned}$$

Krylov subspaces

Given x_0, the solution verifies:

$$\begin{aligned} x &= x_0 + e_0 \\ &= x_0 + A^{-1} r_0 \\ &= x_0 + p(A) r_0 \end{aligned}$$

Definition

$$\mathcal{K}_m(A, r_0) = \operatorname{span}(r_0, Ar_0, \ldots, A^{m-1} r_0)$$

Let the grade ν of r_0 w.r.t. A be defined as:

$$\dim \mathcal{K}_m(A, r_0) = \begin{cases} m & \text{if } m < \nu \\ \nu & \text{if } \nu \leqslant m \end{cases}$$

It verifies $\nu = \min \left\{ m \in [\![1; n]\!] : A^{-1} r_0 \in \mathcal{K}_m(A, r_0) \right\}$

Polynomial projection methods

Subspace + Petrov–Galerkin conditions:

- $x_m \in x_0 + \mathcal{K}_m(A, r_0)$
- $\forall u \in \mathcal{K}_m(A, r_0) \in (Be_m)^T u = 0$

If B is SPD,

$$\forall u \in \mathcal{K}_m(A, r_0), (Be_m)^T u = 0 \iff ||e_m||_B = \min_{y \in \mathcal{K}_m(A, r_0)} ||e_0 - y||_B$$

GMRES

Let $B = A^T A$, then:

$$||r_m||_2 = \min_{y \in x_0 + \mathcal{K}_m(A, r_0)} ||f - Ay||_2$$

GMRES

Let $B = A^T A$, then:

$$||r_m||_2 = \min_{y \in x_0 + \mathcal{K}_m(A, r_0)} ||f - Ay||_2$$

Arnoldi process $\implies$ basis V_m and Hessenberg matrix H_m:

$$AV_m = V_{m+1} H_m$$

GMRES

Let $B = A^T A$, then:

$$||r_m||_2 = \min_{y \in x_0 + \mathcal{K}_m(A, r_0)} ||f - Ay||_2$$

Arnoldi process $\implies$ basis V_m and Hessenberg matrix H_m:

$$AV_m = V_{m+1} H_m$$

For all $y \in x_0 + \mathcal{K}_m(A, r_0)$, $\exists !\, z \in \mathbb{K}^m$:

$$\begin{aligned} f - Ay &= f - A(x_0 + V_m z) \\ &= r_0 - AV_m z \\ &= \beta v_1 - V_{m+1} H_m z \\ &= V_{m+1}(\beta e_1 - H_m z) \end{aligned}$$

Left- and right-preconditioning

- $M^{-1}Ax = M^{-1}f$
- $AM^{-1}y = f$ and $M^{-1}y = x$

Different subspace conditions:

- $x_m \in x_0 + \mathcal{K}_m(M^{-1}A, M^{-1}r_0)$
- $x_m \in x_0 + M^{-1}\mathcal{K}_m(AM^{-1}, r_0)$

Conclusion

General framework

Basic preconditioners often used as building blocks

Krylov methods are widely used with preconditioners:

- domain decomposition
- multigrid

Theoretical analysis (most often in the symmetric case)