On the iterative solution of systems of the form

\[ A^T Ax = A^T b + c \]

E. Riccietti (IRIT-INP, Toulouse)

Joint work with: H. Calandra (TOTAL)
S. Gratton (IRIT-INP, Toulouse)
X. Vasseur (ISAE-SUPAERO, Toulouse)

Manchester University
30th July, 2019
Given $A \in \mathbb{R}^{m \times n}$, $m \geq n$ with $\text{rank}(A) = n$, $b \in \mathbb{R}^m$ and $x, c \in \mathbb{R}^n$, solve

$$A^\top Ax = A^\top b + c \quad \text{(SYS)}$$

or

$$\min_x \|Ax - b\|^2 - x^\top c$$

Remarks

- This is a generalization of the normal equations for least-squares problems (case $c = 0$)
Motivating applications (I)

- Multilevel Levenberg-Marquardt method


\[
\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} \| F(x) \|^2.
\]

We have at disposal an approximation to the objective function:

\[
f^H(x^H) = \frac{1}{2} \| F^H(x^H) \|^2, \quad x^H \in \mathbb{R}^{n_H}, \quad n_H < n
\]

Coarse model:

\[
m_k^H(x_k^H, s^H) = \frac{1}{2} \| F^H(x_k^H) + J^H(x_k^H) s^H \|^2 + \frac{\lambda_k}{2} \| s^H \|^2 + (R \nabla f(x_k) - \nabla f^H(x_0^H))^T s^H,
\]

with \( J^H(x_k^H) \) the Jacobian matrix of \( F^H \) at \( x_k^H \), \( R \) a full-rank linear restriction operator and \( x_0^H = Rx_k \).
Motivating applications (II)

- **Penalty function method**

\[
\begin{align*}
\min_{x} & \quad f(x) \\
\text{s.t.} & \quad g(x) = 0,
\end{align*}
\]

Penalty function:
\[
\Phi_\sigma(x) = f(x) - g(x)^T y_\sigma(x),
\]
where \( y_\sigma(x) \in \mathbb{R}^m \) is defined as the solution of the following minimization problem:
\[
\min_{y} \left\| A(x)^T y - \nabla f(x) \right\|^2 + \sigma g(x)^T y,
\]
with \( A(x) \) the Jacobian matrix of \( g(x) \) at \( x \) and \( \sigma > 0 \), a given real-valued penalty parameter.
Introduction

Interesting questions

- What is the **conditioning** of $A^T A x = A^T b + c$?
  - Standard theory for linear systems do not take into account **structured perturbations** and gives underwhelming results.
  - Structured conditioning analysis is necessary. Presence of $c$ results in a different mapping from data to solution.

- What is the **backward error**?
  - Different set of admissible perturbations on the matrix.

- **How to numerically solve** it by an iterative method?
  - Methods for normal equations such as CGLS cannot be used.
THEORETICAL RESULTS
Conditioning, case $c = 0$

Let $\delta x = x - \hat{x}$, $\hat{x}$ a perturbed solution.

Forward error bound

From standard theory on linear systems:

$$\frac{\|\delta x\|}{\|x\|} \leq \kappa(A)^2 u$$

For least squares problems:

$$\frac{\|\delta x\|}{\|x\|} \leq \gamma_m \kappa_{LS} u, \quad \kappa_{LS} = \kappa(A) \left(1 + \frac{\|A^\dagger\| \|r\|}{\|x\|}\right), \quad r = b - Ax$$

Underwhelming result!

The conditioning of the problem depends on $\kappa(A)^2$ only if $\|r\|$ is large!
Theoretical results

Conditioning

Definition

If \( F \) is a continuously differentiable function

\[
F : \mathcal{X} \to \mathcal{Y} \\
x \mapsto F(x),
\]

the absolute condition number of \( F \) at \( x \) is the scalar \( \| F'(x) \|_{op} \). The relative condition number of \( F \) at \( x \) is

\[
\frac{\| F'(x) \|_{op} \| x \|\_{\mathcal{X}}}{\| F(x) \|\_{\mathcal{Y}}}. 
\]

Conditioning, case $c = 0$

Definition of $F$
We consider $F$ as the function that maps $A, b$ to the solution $x$ of a least squares problem:

$$F : \mathbb{R}^{m \times n} \times \mathbb{R}^m \rightarrow \mathbb{R}^n$$

$$(A, b) \mapsto F(A, b) = A^\dagger b.$$

Explicit formula for the conditioning
The absolute condition number of a least-squares problem, with Euclidean norm on the solution and Frobenius norm on the data\(^a\), is given by

$$\kappa_{NE} = \|A^\dagger\| \sqrt{1 + \|x\|^2 + \|A^\dagger\|^2 \|r\|^2}$$

---

\(^a\| [A, b] \|^2_F := \|A\|^2_F + \|b\|^2$$
A formula for the condition number, \( c \neq 0 \)

**Lemma**

The absolute condition number of the problem SYS is given by

\[
\| F'(A, b, c) \|_{\text{op}} = \| (r^T \otimes (A^T A)^{-1}) L_T + x^T \otimes A^\dagger, A^\dagger, (A^T A)^{-1} \|, 
\]

where \( L_T \) is the linear operator such that \( \text{vec}(A^T) = L_T \text{vec}(A) \) and \( r = b - Ax \).

**Case \( c = 0 \)**

\[
\| F'(A, b, c) \|_{\text{op}} = \| (r^T \otimes (A^T A)^{-1}) L_T + x^T \otimes A^\dagger, A^\dagger \|. 
\]
An explicit formula for the condition number, $c \neq 0$

We consider $F$ as the function that maps $A, b, c$ to the solution $x$ of SYS

$$F : \mathbb{R}^{m \times n} \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$$

$$(A, b, c) \mapsto F(A, b, c) = A^\dagger b + A^\dagger (A^\dagger)^\top c.$$

Theorem

The absolute condition number of problem SYS, with Euclidean norm on the solution and Frobenius norm on the data$^a$, is $\sqrt{\|\tilde{M}\|}$, with $\tilde{M} \in \mathbb{R}^{n \times n}$ given by

$$\tilde{M} = (1 + \|r\|^2)(A^\top A)^{-2} + (1 + \|x\|^2)(A^\top A)^{-1} - 2 \text{ sym}(B),$$

with $B = A^\dagger r x^\top (A^\top A)^{-1}$, $\text{sym}(B) = \frac{1}{2}(B + B^\top)$ and $x$ the exact solution of SYS.

$^a \| [A, b, c] \|_F^2 := \|A\|_F^2 + \|b\|^2 + \|c\|^2$

Upper bound for the condition number

$$\sqrt{\|\tilde{M}\|} \leq (1 + \|r\| + 2\sqrt{\|c\| \|x\|})\|A^\dagger\|^2 + (1 + \|x\|)\|A^\dagger\|.$$
Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ and $\tilde{x}$ a perturbed solution to SYS. Find the smallest perturbation $E$ of $A$ such that the vector $\tilde{x}$ exactly solves

$$(A + E)^T(A + E)x = (A + E)^T b + c,$$

i.e. given

$$G := \{E \in \mathbb{R}^{m \times n} : (A + E)^T(A + E) \tilde{x} = (A + E)^T b + c\},$$

we want to compute the quantity:

$$\eta(\tilde{x}) = \min_{E \in G} \|E\|_F.$$
Set of admissible perturbations on the matrix

Theorem

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^{m}$, $c, \tilde{x} \in \mathbb{R}^{n}$ and assume that $\tilde{x} \neq 0$. Let $\tilde{r} = b - A\tilde{x}$ and define two sets $\mathcal{E}, \mathcal{M}$ by

$\mathcal{E} = \{ E \in \mathbb{R}^{m \times n} : (A + E)^\top (b - (A + E)\tilde{x}) = -c \},$

$\mathcal{M} = \{ \nu (\alpha c^\top - \nu^\top A) + (I_m - \nu \nu^\top) (\tilde{r}\tilde{x}^\top + Z(I_n - \tilde{x}\tilde{x}^\top)) : \nu \in \mathbb{R}^m, Z \in \mathbb{R}^{m \times n}, \alpha \in \mathbb{R}, s.t. \alpha \| \nu \|^2 (\nu^\top b - \alpha c^\top \tilde{x}) = -1 \}.$

Then $\mathcal{E} = \mathcal{M}$.

Case $c = 0$

$\mathcal{E} = \{ E \in \mathbb{R}^{m \times n} : (A + E)^\top (b - (A + E)\tilde{x}) = 0 \},$

$\mathcal{M} = \{ -\nu \nu^\top A + (I_m - \nu \nu^\top) (\tilde{r}\tilde{x}^\top + Z(I_n - \tilde{x}\tilde{x}^\top)) : \nu \in \mathbb{R}^m, Z \in \mathbb{R}^{m \times n} \}.$
Lower bound on the backward error

Lemma

The set of admissible perturbations $\mathcal{E}$ defined in Theorem is such that $\mathcal{E} \subseteq \mathcal{M}_2$, with

$$\mathcal{M}_2 = \left\{ \nu \left( \alpha c^T - \nu^\dagger A \right) + (I_m - \nu \nu^\dagger) (\tilde{r} \tilde{x}^\dagger + Z (I_n - \tilde{x} \tilde{x}^\dagger)) : \nu \in \mathbb{R}^m, Z \in \mathbb{R}^{m \times n}, \alpha \in \mathbb{R} \right\}.$$ 

Then,

$$\min_{\mathcal{E}} \|E\|_F^2 \geq \min_{\mathcal{M}_2} \|E\|_F^2 = \frac{\|\tilde{r}\|^2}{\|\tilde{x}\|^2} + \min\{\lambda_*, 0\},$$

for $\lambda_* = \lambda_{\min} \left( A(I_n - cc^T)A^T - \frac{\tilde{r} \tilde{r}^T}{\|\tilde{x}\|^2} \right)$, with $\lambda_{\min}(M)$ denoting the smallest eigenvalue of the matrix $M$.

Case $c = 0$

$$\min_{\mathcal{E}} \|E\|_F^2 = \frac{\|\tilde{r}\|^2}{\|\tilde{x}\|^2} + \min\{\lambda_*, 0\}, \quad \lambda_* = \lambda_{\min} \left( AA^T - \frac{\tilde{r} \tilde{r}^T}{\|\tilde{x}\|^2} \right).$$
NUMERICAL SOLUTION OF THE SYSTEM
CG vs CGLS for normal equations

Same method in exact arithmetic, different performance in finite precision for some problems:

- in CGLS $d_k = b - Ax_k$ is recurred and $r_k = A^T d_k$.

### Algorithm 1 CG for $A^T Ax = A^T b$

Input: $A, b, x_0$.

Define $r_0 = A^T (b - Ax_0), \ p_1 = r_0$.

for $k = 1, 2, \ldots$ do

\[
\alpha_k = \frac{r_{k-1}^T r_{k-1}}{\|Ap_k\|^2},
\]

\[
x_k = x_{k-1} + \alpha_k p_k,
\]

\[
r_k = r_{k-1} - \alpha_k A^T (Ap_k),
\]

\[
\beta_k = \frac{r_k^T r_k}{r_{k-1}^T r_{k-1}},
\]

\[
p_{k+1} = r_k + \beta_k p_k.
\]

end for

### Algorithm 2 CGLS for $A^T Ax = A^T b$

Input: $A, b, x_0$.

Define $d_0 = b - Ax_0, \ r_0 = A^T d_0, \ p_1 = r_0$.

for $k = 1, 2, \ldots$ do

\[
t_k = Ap_k,
\]

\[
\alpha_k = \frac{r_{k-1}^T r_{k-1}}{\|t_k\|^2},
\]

\[
x_k = x_{k-1} + \alpha_k p_k,
\]

\[
d_k = d_{k-1} - \alpha_k t_k,
\]

\[
r_k = A^T d_k,
\]

\[
\beta_k = \frac{r_k^T r_k}{r_{k-1}^T r_{k-1}},
\]

\[
p_{k+1} = r_k + \beta_k p_k.
\]

end for


Initial rounding error due to the product $r_0 = A^T b + c - A^T A x_0$:

$$\| \delta x \| \leq \kappa(A)^2 u \left( \frac{\| b \|}{\| A \|} + \frac{\| c \|}{\| A \|^2} \right).$$

This initial error cannot be canceled, and the best error bound we can hope for will include the term given above.

Optimal bound:

$$\| \delta x \| \leq \sqrt{\| \tilde{M} \|} \| [A, b, c] \|_F u$$

If

$$\| b \| \| A \| + \| c \| \gg \left[ 1 + \| r \| + 2 \sqrt{\| c \|} \| x \| + \frac{1 + \| x \|}{\| A^\dagger \|} \right] \sqrt{\| A \|_F^2 + \| b \|^2 + \| c \|^2}$$

CG can be expected to produce less than optimal accuracy.
IDEA to design a stable method

- Extend the successful algorithmic procedures to the case \( c \neq 0 \)
- Need to factorize matrix \( A \) in both the left and right hand sides

\[ A^T (A^T x - b) \]

Two solution methods
We propose two iterative methods based on two different reformulations of the problem
Proposed methods (I) CGLS$_\epsilon$

Given $\epsilon > 0$, let us then define

$$A_\epsilon = \begin{bmatrix} A \\ \epsilon c^T \end{bmatrix}, \quad b_\epsilon = \begin{bmatrix} b \\ 1/\epsilon \end{bmatrix}.$$  

We then consider the following linear least squares problem:

$$\min_x \| A_\epsilon x - b_\epsilon \|^2,$$

with normal equations

$$(A^T A + \epsilon^2 cc^T)x = A^T b + c. \quad \text{(SYS}_\epsilon)$$

CGLS$_\epsilon$ solves SYS$_\epsilon$ with CGLS method

**Lemma**

Let $x_\epsilon$ be the solution of SYS$_\epsilon$ and $x$ be the solution of SYS. Then, $\lim_{\epsilon \to 0} x_\epsilon = x$ and the relative norm of the error satisfies

$$\frac{\| x_\epsilon - x \|}{\| x \|} \leq \epsilon^2 \frac{\| c \| \| w \|}{1 + \epsilon^2 c^T w}, \quad w = (A^T A)^{-1} c.$$
Will a really small $\epsilon$ may cause large errors in finite arithmetic?

A perturbed solution $\tilde{x}_\epsilon = x_\epsilon + \delta x_\epsilon$ will be such that:

$$(A_\epsilon^T A_\epsilon)(\delta x_\epsilon) = \delta(A_\epsilon^T b_\epsilon). \quad |\delta(A_\epsilon^T b_\epsilon)| \leq \gamma_{m+1}|A_\epsilon^T||b_\epsilon|$$

This overestimates the error!

$$fl(A_\epsilon^T b_\epsilon) = fl(A^T b) + fl\left(\epsilon c \frac{1}{\epsilon}\right) + \delta_s,$$

with $\delta_s$ error due to the summation.

If $\epsilon = 2^i$ for $i \in \mathbb{Z}$, then $fl\left(\epsilon c \frac{1}{\epsilon}\right) = c$. Then,

$$fl\left(A_\epsilon^T b_\epsilon\right) = A^T b + c + \delta_p + \delta_s, \quad \text{with} \quad |\delta_s| \leq u|fl(A^T b) + c|, \quad |\delta_p| \leq \gamma_m|A||b|,$$

and the bound does not depend on $\epsilon$. 


What about the conditioning of the problem?
Due to the presence of small $\epsilon$ in the right-hand side the residual will generally be really large.
Standard conditioning analysis of least squares problems is not well-suited in this case.
We can show that the conditioning does not depend on $\|b_\epsilon - A_\epsilon x_\epsilon\|$, that will be really large, but rather on $\|r_\epsilon\| = \|b - Ax_\epsilon\|$, that will be indeed much smaller.
Let $F_\varepsilon$ be the function that maps $A, b, c$ to the solution $x_\varepsilon$ of SYS$\varepsilon$

$$
F_\varepsilon : \mathbb{R}^{m \times n} \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n
$$

$$(A, b, c) \mapsto F_\varepsilon(A, b, c) = (A_\varepsilon^\top A_\varepsilon)^{-1}(A^\top b + c),$$

and let $r_\varepsilon = b - Ax_\varepsilon$.

The absolute condition number of problem SYS$\varepsilon$, with Euclidean norm on the solution and Frobenius norm on the data, is then given by:

$$
\|F'_\varepsilon(A, b, c)\|_{\text{op}} = \|(r_\varepsilon^\top \otimes (A_\varepsilon^\top A_\varepsilon)^{-1}) L_T + x_\varepsilon^\top \otimes (A_\varepsilon^\top A_\varepsilon)^{-1} A^\top, (A_\varepsilon^\top A_\varepsilon)^{-1} A^\top, (1 - 2\varepsilon c^\top x_\varepsilon)(A_\varepsilon^\top A_\varepsilon)^{-1}\|.
$$

Computable formula: $\sqrt{\|\tilde{M}_\varepsilon\|}$, with

$$
\tilde{M}_\varepsilon = ((1 - 2\varepsilon c^\top x_\varepsilon)^2 + \|r_\varepsilon\|^2)(A_\varepsilon^\top A_\varepsilon)^{-2}
+ (1 + \|x_\varepsilon\|^2)(A_\varepsilon^\top A_\varepsilon)^{-1} A^\top A(A_\varepsilon^\top A_\varepsilon)^{-1} - 2 \text{ sym}(B_\varepsilon)
$$

with $B_\varepsilon = (A_\varepsilon^\top A_\varepsilon)^{-1} A^\top r_\varepsilon x_\varepsilon^\top (A_\varepsilon^\top A_\varepsilon)^{-1}$ and $\text{sym}(B_\varepsilon) = \frac{1}{2}(B_\varepsilon + B_\varepsilon^\top)$.
Proposed method (II) CGLS/

Given $\hat{I} \in \mathbb{R}^{(m+1) \times (m+1)}$, we define $\hat{A} \in \mathbb{R}^{(m+1) \times n}$ and $\hat{b} \in \mathbb{R}^{m+1}$ as:

$$
\hat{A} = \begin{bmatrix} A \\ c^T \end{bmatrix}, \quad \hat{I} = \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}, \quad \hat{b} = \begin{bmatrix} b \end{bmatrix}.
$$

We then reformulate SYS as:

$$
\hat{A}^T \hat{I} \hat{A} x = \hat{A}^T \hat{b}
$$

Possible to factorize $\hat{A}^T$ in both the right and the left-hand sides:

- no need of recurring the residual $r = \hat{A}^T (\hat{I} \hat{A} x - \hat{b})$ (simply update $\hat{d} = \hat{I} \hat{A} x - \hat{b}$ along the iterations and form $r$ by multiplication with $\hat{A}^T$)
- computation of $p_k^T A^T A p_k$ as $\| \hat{I} \hat{A} p_k \|^2$

We can therefore expect the same benefits of CGLS as compared to CG.
Algorithm 3 CGLS for $A^T A x = A^T b + c$

Input: $\hat{A}, \hat{b}, x_0$
Define $\hat{d}_0 = \hat{b} - \hat{A} x_0$, $r_0 = \hat{A}^T (\hat{b} - \hat{A} x_0)$, $p_1 = r_0$.

for $k = 1, 2, \ldots$ do

$\hat{t}_k = \hat{A} p_k,$

$\alpha_k = \frac{r_{k-1}^T r_{k-1}}{\hat{t}_k^T \hat{t}_k},$

$x_k = x_{k-1} + \alpha_k p_k,$

$\hat{d}_k = \hat{d}_{k-1} - \alpha_k \hat{t}_k,$

$r_k = \hat{A}^T \hat{d}_k,$

$\beta_k = \frac{r_k^T r_k}{r_{k-1}^T r_{k-1}},$

$p_{k+1} = r_k + \beta_k p_k.$

end for
First order approximation for the forward error

First order approximation for the forward error can be obtained as

\[
\frac{\|x - \hat{x}\|}{\|x\|} \sim \frac{\kappa_{SYS} \|[A, b, c]\|_F}{\|x\|} u, \quad u \text{ machine precision}
\]

We define the following error estimates:

\[
\hat{E}_{\text{CGLSI}} := \frac{\sqrt{\|M\| \|[A, b, c]\|_F}}{\|x\|} u,
\]

\[
\hat{E}_{\text{CGLS}_\epsilon} := \epsilon^2 \frac{c \, w}{1 + \epsilon^2 c^T w} + \frac{\sqrt{\|M_\epsilon\| \|[A, b, c]\|_F}}{\|x\|} u \left\|l_n - \frac{\epsilon^2 wc^T}{1 + \epsilon^2 c^T w}\right\|,
\]

\(u\) being the machine precision.

- **CGLS\(\epsilon\):** the error on the computed solution \(\hat{x}_\epsilon\) depends on two terms:

\[
\frac{\|x - \hat{x}_\epsilon\|}{\|x\|} \leq \frac{\|x - x_\epsilon\|}{\|x\|} + \frac{\|x_\epsilon - \hat{x}_\epsilon\|}{\|x\|} = \frac{\|x - x_\epsilon\|}{\|x\|} + \frac{\|x_\epsilon - \hat{x}_\epsilon\|}{\|x_\epsilon\|} \frac{\|x_\epsilon\|}{\|x\|}.
\]
NUMERICAL TESTS
Numerical tests: setting

- All the numerical methods have been implemented in Matlab

- $A \in \mathbb{R}^{m \times n}$, $A = U \Sigma V^T$, where $U$ and $V$ from `gallery('orthog',m/n,j), j = 1, \ldots, 6.$

- $C_1 : \Sigma_{ii} = a^{-i}$, for $a > 0$,

- $C_2 : \Sigma_{ii} = u_i$, $u = \text{linspace}(dw, up, n)$, with $dw, up > 0$, for $i = 1, \ldots, n$.

- Matrix dimensions: $m = 40$ and $n = 20$ for the tests and $m = 100$, $n = 50$ for performance profiles

- Performance profiles: 40 matrices, with condition number between 1 and $10^{10}$. The optimality measure is $\frac{\|x - \hat{x}\|}{\|x\|}$, with $x$ the exact solution ($x = (n - 1 : -1 : 0)$). A simulation is considered unsuccessful if the relative solution accuracy is larger than $10^{-2}$. 
How to choose $\epsilon$?

$$\frac{\|X - \hat{X}_\epsilon\|}{\|X\|} \leq \frac{\|X - X_\epsilon\|}{\|X\|} + \frac{\|X_\epsilon - \hat{X}_\epsilon\|}{\|X\|} = \frac{\|X - X_\epsilon\|}{\|X\|} + \frac{\|X_\epsilon - \hat{X}_\epsilon\|}{\|X_\epsilon\|} \frac{\|X_\epsilon\|}{\|X\|}.$$
Figure: Left: right hand side of small norm, Right: right hand side of large norm
Comparison with CG

Figure: Left: $\kappa(A) = 10^5$, $\kappa(\hat{A}) = 10^5$. Right: $\kappa(A) = 10^7$, $\kappa(\hat{A}) = 10^{10}$.

Performance of CGLSI and CGLS$\epsilon$ is comparable but

- CGLSI is parameter free
- CGLSI is less sensible to the right hand side

Much better performance than CG
Numerical tests

Validation of error bounds

<table>
<thead>
<tr>
<th>Problem</th>
<th>$\kappa(A)^2 u$</th>
<th>$E_{CGLSI}$</th>
<th>$\hat{E}_{CGLSI}$</th>
<th>$E_{CGLS_\epsilon}$</th>
<th>$\hat{E}<em>{CGLS</em>\epsilon}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a = 2.0$</td>
<td>$10^{-10}$</td>
<td>$10^{-11}$</td>
<td>$10^{-8}$</td>
<td>$10^{-9}$</td>
<td>$10^{-9}$</td>
</tr>
<tr>
<td>$a = 2.5$</td>
<td>$10^{-2}$</td>
<td>$10^{-9}$</td>
<td>$10^{-6}$</td>
<td>$10^{-9}$</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>$a = 1.5$</td>
<td>$10^{-10}$</td>
<td>$10^{-13}$</td>
<td>$10^{-11}$</td>
<td>$10^{-13}$</td>
<td>$10^{-11}$</td>
</tr>
<tr>
<td>$a = 1.3$</td>
<td>$10^{-12}$</td>
<td>$10^{-13}$</td>
<td>$10^{-13}$</td>
<td>$10^{-14}$</td>
<td>$10^{-13}$</td>
</tr>
<tr>
<td>$a = 1.1$</td>
<td>$10^{-14}$</td>
<td>$10^{-14}$</td>
<td>$10^{-13}$</td>
<td>$10^{-14}$</td>
<td>$10^{-14}$</td>
</tr>
<tr>
<td>$a = 0.7$</td>
<td>$10^{-10}$</td>
<td>$10^{-12}$</td>
<td>$10^{-12}$</td>
<td>$10^{-12}$</td>
<td>$10^{-12}$</td>
</tr>
<tr>
<td>up = 1</td>
<td>$10^{-2}$</td>
<td>$10^{-9}$</td>
<td>$10^{-7}$</td>
<td>$10^{-8}$</td>
<td>$10^{-8}$</td>
</tr>
<tr>
<td>up = 1</td>
<td>$10^{-2}$</td>
<td>$10^{-8}$</td>
<td>$10^{-4}$</td>
<td>$10^{-4}$</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>$a = 1.5$</td>
<td>$10^{-10}$</td>
<td>$10^{-13}$</td>
<td>$10^{-12}$</td>
<td>$10^{-10}$</td>
<td>$10^{-10}$</td>
</tr>
</tbody>
</table>

- Better performance than standard CG, both in terms of accuracy and of rate of convergence.
- The error bounds much better predict forward errors than classical bounds.
Propose method can compare with direct methods in terms of solution accuracy.
THANK YOU FOR YOUR ATTENTION

Calandra, H., Gratton, S., Riccietti, E., Vasseur, X., On the solution of systems of the form $A^T Ax = A^T b + c$, In preparation
Effect of large right-hand sides

Let us assume to apply CG to SYS and CGLS to SYS\(_{\epsilon}\). We would respectively compute:

\[
\alpha_1 = \frac{\|r_0\|^2}{p_1^T A^T A p_1} = \frac{\|A^T b + c\|^2}{\|A(A^T b + c)\|^2}, \quad x_1 = \alpha_1(A^T b + c) = \alpha_1 p_1,
\]

and

\[
\alpha_1(\epsilon) = \frac{\|A^T b + c\|^2}{\|A(A^T b + c)\|^2 + \epsilon \|c^T (A^T b + c)\|^2}, \quad x_1(\epsilon) = \alpha_1(\epsilon) p_1(\epsilon) = \alpha_1(\epsilon) p_1.
\]

Notice that if \(\epsilon\) tends to zero, so does the term \(\epsilon \|c^T (A^T b + c)\|\) in the denominator of \(\alpha_1(\epsilon)\). Consequently \(\alpha_1(\epsilon)\) tends toward \(\alpha_1\) and \(x_1(\epsilon)\) tends toward \(x_1\). If \(\epsilon\) has to be fixed, its value should be small enough to let \(\epsilon \|c^T (A^T b + c)\|\) be small compared to \(\|A(A^T b + c)\|^2\), otherwise the found approximation will be close to a solution of SYS\(_{\epsilon}\) rather than to one of SYS. This choice is then particularly difficult when \(\|A^T b + c\|\) is large.