The extended normal equations: conditioning and iterative solution

E. Riccietti (ENS Lyon)
http://perso.ens-lyon.fr/elisa.riccietti/

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

Communications in NLA

28th September, 2020

Context

Given $A \in \mathbb{R}^{m \times n}$, $m \geq n$ with $\mathrm{rank}(A) = n$, $b \in \mathbb{R}^m$ and $x, c \in \mathbb{R}^n$, we consider the *extended least squares problem*

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|^2 - c^T x, \tag{ELS}$$

whose solution satisfies the *extended normal equations*

$$A^T A x = A^T b + c. \tag{ENE}$$

→ This is a generalization of the least squares problem (case $c = 0$)

Motivating applications

- Multilevel Levenberg-Marquardt method

  Calandra, H., Gratton, S., Riccietti, E., Vasseur, X., *On the approximation of the solution of partial differential equations by artificial neural networks trained by a multilevel Levenberg-Marquardt method*, OMS, 2020

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

- Penalty function method

  Fletcher, R., *A class of methods for nonlinear programming: III. Rates of convergence*, Numerical Methods for Nonlinear Optimization, 1973

  Estrin, R. and Orban, D. and Saunders, M. A., *LNLQ: An iterative method for least-norm problems with an error minimization property*, SIMAX, 2019

$$\min_{x} f(x)$$
$$\text{s.t. } g(x) = 0.$$

Our questions

1. Practical aspects:
   - How to numerically solve (ENE) by a stable iterative method?

2. Theoretical aspects:
   - How to build a good bound for the forward error on the computed solution by such method?
     - What is the conditioning of (ENE)?
     - What is the backward error of (ENE)?

NUMERICAL SOLUTION OF THE SYSTEM

Exploit the structure of the problem

1. **Case** $c = 0$
   - Forming matrix $A^T A$ leads to a loss of accuracy
   - Practical solution methods do not form this product:

$$A^T Ax - A^T b = A^T (Ax - b)$$

   - Direct methods: employ a factorization of A rather than of $A^T A$
   - Iterative methods: perform matrix-vector multiplications Ax and $A^T y$.

2. **Case** $c \neq 0$ **?**

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

Paige, C. C. and Saunders, M. A., *LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares,* ACM Trans. Math. Softw., 1982

Björck, A. and Elfving, T. and Strakos, Z. , *Stability of conjugate gradient and Lanczos methods for linear least squares problems,* SIMAX, 1998

Stable method for solving (ENE): CGLSc

- Extend the successful algorithmic procedures to the case $c \neq 0$

Algorithm 3 CG for $A^T Ax = A^T b + c$

Input: A, b, c, x_0.
Define $r_0 = A^T(b - Ax_0) + c$, $p_1 = r_0$.
for $k = 1, 2, \ldots$ **do**
$\alpha_k = \|r_{k-1}\|^2 / \|Ap_k\|^2$,
$x_k = x_{k-1} + \alpha_k p_k$,
$r_k = r_{k-1} - \alpha_k A^T(Ap_k)$,
$\beta_k = \|r_k\|^2 / \|r_{k-1}\|^2$,
$p_{k+1} = r_k + \beta_k p_k$.
end for

Algorithm 4 CGLSc for $A^T Ax = A^T b + c$

Input: A, b, x_0
Define $r_0 = b - Ax_0$, $s_0 = A^T r_0 + c$, $p_1 = s_0$.
for $k = 1, 2, \ldots$ **do**
$t_k = Ap_k$
$\alpha_k = \|s_{k-1}\|^2 / \|t_k\|^2$
$x_k = x_{k-1} + \alpha_k p_k$
$r_k = r_{k-1} - \alpha_k t_k$
$s_k = A^T r_k + c$
$\beta_k = \|s_k\|^2 / \|s_{k-1}\|^2$
$p_{k+1} = r_k + \beta_k p_k$
end for

Numerical tests: setting

- All the numerical methods have been implemented in Matlab
- Matrix of dimensions $m = 100$, $n = 50$ with known singular values distribution
- Performance profiles: 55 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}.

Remark

(ENE) is equivalent to the augmented system

$$\begin{bmatrix} \xi I_m & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \begin{bmatrix} b \\ -c/\xi \end{bmatrix}, \quad r = \xi y = b - Ax, \qquad \text{(AUG)}$$

Comparison with iterative methods

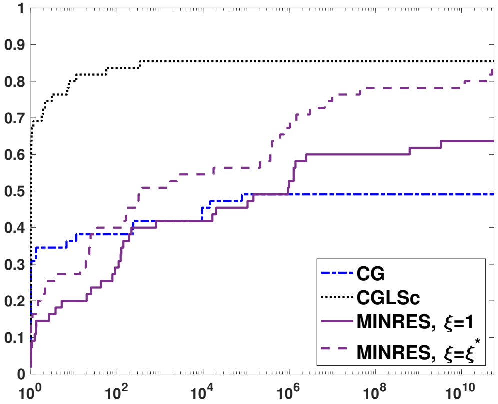
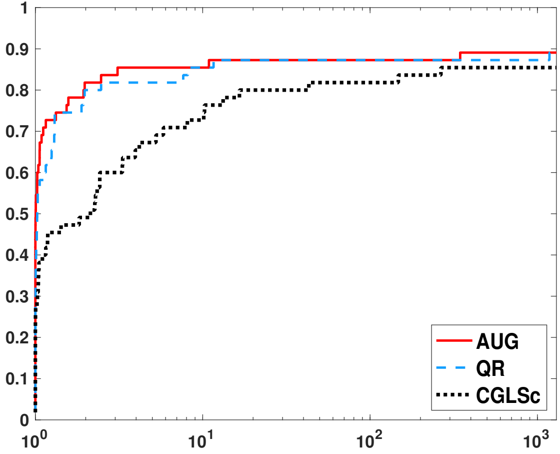


Figure: Performance profile in logarithmic scale. The optimality measure considered is the relative solution accuracy $\|x - \hat{x}\|/\|x\|$.

Comparison with direct methods



- QR: solves (AUG) with $\xi = 1$, employing the QR factorization of $[A, b]$.
- AUG: solves (AUG) with $\xi = \xi^*$ using an LBL^T factorization (Matlab ldl).

→ CGLS*c* can compare with direct methods in terms of solution accuracy

THEORETICAL RESULTS:
Error bounds

Why can't we use existing theory?

Can we use standard linear systems theory?

This gives underwhelming results already for normal equations.

Let x and $\hat{x}$ be an exact and a perturbed solution of (LS), $\delta x = x - \hat{x}$, u the machine precision, $r = b - Ax$ the residual.

Forward error bound

Linear systems' theory:

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

Least squares theory:

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{m}{1 - mu} \kappa(A) \left(1 + \frac{\|A^\dagger\| \|r\|}{\|x\|}\right) u$$

Underwhelming result!

The conditioning of the problem depends on $\kappa(A)^2$ only if $\|r\|$ is large! The bound from linear systems' theory is pessimistic.

Why such underwhelming results?

Standard linear systems theory:

- Based on the assumption that the matrix $A^T A$ is formed explicitly.
- Practical solution methods do not form this product:

$$A^T Ax - A^T b = A^T(Ax - b)$$

  - Direct methods: employ a factorization of A rather than of $A^T A$
  - Iterative methods: perform matrix-vector multiplications Ax and $A^T y$.
- We should consider perturbations on matrix A rather than on matrix $A^T A$: we need a structured analysis to obtain condition number and backward error
- Better error bounds:

$$FE := \frac{\|x - \hat{x}\|}{\|x\|} \sim \text{relative condition number} \times \text{backward error}$$

Condition number

Definition

If F is a continuously differentiable function

$$\begin{aligned} F : \mathcal{X} &\to \mathcal{Y} \\ x &\longmapsto F(x), \end{aligned}$$

the absolute condition number of F at x is the scalar

$$\|F'(x)\|_{\mathrm{op}} := \sup_{\|v\|_{\mathcal{X}}=1} \|F'(x)v\|_{\mathcal{Y}},$$

where $F'(x)$ is the Fréchet derivative of F at x.
The relative condition number of F at x is

$$\frac{\|F'(x)\|_{\mathrm{op}} \, \|x\|_{\mathcal{X}}}{\|F(x)\|_{\mathcal{Y}}}.$$

J . R . Rice, *A theory of condition*, SIAM J . Numer . Anal ., 1966

Conditioning, normal equations ($c = 0$)

Definition of F

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

$$F : \mathbb{R}^{m\times n} \times \mathbb{R}^m \to \mathbb{R}^n$$
$$(A, b) \longmapsto F(A, b) = A^\dagger b.$$

Explicit formula for the conditioning

The absolute condition number of the normal equations, with Euclidean norm on the solution and Frobenius norm on the data[a], is given by

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

Gratton, S., *On the condition number of linear least squares problems in a weighted Frobenius norm*, BIT Numerical Mathematics, 1996

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

Backward error

Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and $\tilde{x}$ a perturbed solution of the normal equations. Find the smallest perturbation (E, f) of (A, b) such that the vector $\tilde{x}$ exactly solves

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

i.e. given

$$\mathcal{G} := \{(E, f) \in \mathbb{R}^{m\times n+1} : (A + E)^T (A + E)\tilde{x} = (A + E)^T (b + f)\},$$

we want to compute the quantity:

$$\eta(\tilde{x}) = \min_{(E,f)\in\mathcal{G}} \|[E, f]\|_F.$$

Well studied problem → explicit formula for $\eta(\tilde{x})$

Why can't we use standard least squares theory?

Presence of c:

- Conditioning: different mapping from data to solution.
- Backward error: different set of admissible perturbations.

Conditioning for (ENE)

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

$$F : \mathbb{R}^{m\times n} \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$$
$$(A, b, c) \longmapsto F(A, b, c) = A^\dagger b + A^\dagger (A^\dagger)^T c.$$

Lemma

The absolute condition number of the problem ENE is given by

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

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

Case $c = 0$

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

An explicit formula for the condition number, $c \neq 0$

Theorem

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

$$\bar{M} = (1 + \|r\|^2)(A^T A)^{-2} + (1 + \|x\|^2)(A^T A)^{-1} - 2\,\mathrm{sym}(B),$$

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

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

Remark

The structured relative condition number is

$$\kappa_S = \frac{\sqrt{\|\bar{M}\|}\,\|A, b, c\|_F}{\|x\|}$$

There are problems in which κ_S can be as large as a quantity of order $\kappa(A)^2$, while in others it can be as low as $\kappa(A)$.

Backward error for (ENE)

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 (ENE). Find the smallest perturbation (E, f, g) of (A, b, c) such that the vector $\tilde{x}$ exactly solves

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

i.e. given

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

we want to compute the quantity:

$$\eta(\tilde{x}) = \min_{(E,f,g)\in\mathcal{G}} \|(E, f, g)\|_F := \sqrt{\|E\|_F^2 + \|f\|^2 + \|g\|^2}$$

Difficult to solve → we use a linearized estimate $\bar{\eta}$

First order approximation for the forward error

- *Classical analysis*:

$$\Delta_C = \kappa(A)^2 \frac{\|A^T A\hat{x} - A^T b - c\|}{\|A\|^2 \|\hat{x}\|}$$

- *Structured analysis*:

$$\Delta_S = \frac{\sqrt{\|\bar{M}\|}\,\|(A, b, c)\|_F}{\|\hat{x}\|} \bar{\eta}_r(\hat{x}).$$

This is valid only if matrix $A^T A$ is not explicitly formed.

Validation of the structured error bound

Table: $\kappa(A)$: condition number, κ_S: structured condition number, FE: $\|x - \hat{x}\|/\|x\|$ forward error, Δ_C: standard bound, Δ_S: structured estimate.

				CGLSc	
Pb.	$\kappa(A)$	κ_S	FE	Δ_C	Δ_S
1	$9 \cdot 10^2$	$1 \cdot 10^6$	$5 \cdot 10^{-13}$	$2 \cdot 10^{-10}$	$1 \cdot 10^{-11}$
2	$2 \cdot 10^3$	$4 \cdot 10^3$	$7 \cdot 10^{-15}$	$3 \cdot 10^{-10}$	$3 \cdot 10^{-13}$
3	$5 \cdot 10^5$	$6 \cdot 10^5$	$1 \cdot 10^{-12}$	$3 \cdot 10^{-5}$	$5 \cdot 10^{-11}$
4	$4 \cdot 10^7$	$4 \cdot 10^7$	$4 \cdot 10^{-11}$	$6 \cdot 10^{-2}$	$4 \cdot 10^{-9}$
5	$1 \cdot 10^9$	$5 \cdot 10^8$	$3 \cdot 10^{-8}$	$7 \cdot 10^2$	$3 \cdot 10^{-7}$
6	$1 \cdot 10^5$	$3 \cdot 10^{10}$	$2 \cdot 10^{-8}$	$3 \cdot 10^{-6}$	$1 \cdot 10^{-7}$
7	$1 \cdot 10^4$	$5 \cdot 10^5$	$6 \cdot 10^{-13}$	$2 \cdot 10^{-8}$	$2 \cdot 10^{-12}$
8	$1 \cdot 10^4$	$8 \cdot 10^9$	$9 \cdot 10^{-10}$	$8 \cdot 10^{-8}$	$7 \cdot 10^{-8}$
9	$1 \cdot 10^4$	$3 \cdot 10^7$	$5 \cdot 10^{-11}$	$2 \cdot 10^{-8}$	$1 \cdot 10^{-10}$
10	$1 \cdot 10^7$	$3 \cdot 10^{10}$	$3 \cdot 10^{-8}$	$3 \cdot 10^{-2}$	$1 \cdot 10^{-7}$

THANK YOU FOR YOUR ATTENTION

Calandra, H., Gratton, S., Riccietti, E., Vasseur, X., On iterative solution of the extended normal equations, SIMAX, 2020
http://perso.ens-lyon.fr/elisa.riccietti/doc/linear.pdf

QR method

- Solves the augmented system:

$$\begin{bmatrix} \xi I_m & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = \begin{bmatrix} b \\ -c/\xi \end{bmatrix}, \quad r = \xi y = b - Ax,$$

with $\xi = 1$, employing the QR factorization of $[A, b]$, as described in theorem below.

Theorem

Let $A \in \mathbb{R}^{m \times n}$, $m \geq n$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$. Assume that $\mathrm{rank}(A) = n$ and let

$$[A, b] = Q \begin{bmatrix} R & d_1 \\ 0 & d_2 \end{bmatrix}.$$

For any $\xi \neq 0$, the solution to the augmented system can be computed from

$$R^T z = -c, \quad Rx = (d_1 - z), \quad r = Q \begin{bmatrix} z \\ d_2 \end{bmatrix}.$$

Remark

- (ENE) and (AUG) also give the first-order optimality conditions for the problems

$$\min_{x,r} \frac{1}{2}\|r\|^2 - c^T x \quad \text{subject to} \quad Ax + r = b, \tag{ELS-primal}$$

and

$$\min_{r} \frac{1}{2}\|r\|^2 - b^T r \quad \text{subject to} \quad A^T r = -c. \tag{ELS-dual}$$

Motivating applications (I)

- Multilevel Levenberg-Marquardt method

Calandra, H., Gratton, S., Riccietti, E., Vasseur, X., *On the approximation of the solution of partial differential equations by artificial neural networks trained by a multilevel Levenberg-Marquardt method*, OMS, 2020

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

We have at disposal an approximation to the objective function:

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

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

Fletcher, R., *A class of methods for nonlinear programming: III. Rates of convergence*, Numerical Methods for Nonlinear Optimization, 1973

Estrin, R. and Orban, D. and Saunders, M. A., *LNLQ: An iterative method for least-norm problems with an error minimization property*, SIMAX, 2019

$$\min_x f(x)$$
$$\text{s.t. } g(x) = 0,$$

Penalty function :

$$\Phi_\sigma(x) = f(x) - g(x)^T y_\sigma(x),$$

where $y_\sigma(x) \in \mathbb{R}^m$ is the solution of

$$\min_y \|A(x)^T y - \nabla f(x)\|^2 + \sigma g(x)^T y,$$

with $A(x)$ the Jacobian matrix of $g(x)$ at x and $\sigma > 0$, a given real-valued penalty parameter.

Theorem

The absolute condition number of problem (ENE), with Euclidean norm on the solution and Frobenius norm (parameterized by α, β, γ) on the data, is $\sqrt{\|\bar{M}\|}$, with $\bar{M} \in \mathbb{R}^{n \times n}$ given by

$$\bar{M} = \left(\frac{1}{\gamma^2} + \frac{\|r\|^2}{\alpha^2}\right)(A^T A)^{-2} + \left(\frac{1}{\beta^2} + \frac{\|x\|^2}{\alpha^2}\right)(A^T A)^{-1} - \frac{2}{\alpha^2}\,\mathrm{sym}(B), \qquad (1)$$

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

The structured conditioning of the normal equations is

$$\|F'(A,b)\| = \|A^\dagger\| \sqrt{\frac{1}{\beta^2} + \frac{\|x\|^2 + \|A^\dagger\|^2 \|r\|^2}{\alpha^2}}.$$

If $c = 0$ and $\gamma \to \infty$, the known result for least squares problems is recovered (note that in this case $B = 0$ as $A^T r = 0$).
Taking large values of γ allows us to perturb A and b only, and to include the case $c = 0$. This is because the condition $\gamma \to \infty$ implies $g \to 0$, from the constraint $\alpha^2 \|E\|_F^2 + \beta^2 \|f\|^2 + \gamma^2 \|g\|^2 = 1$ in the definition of the condition number.

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

$$
\begin{aligned}
\mathcal{E} =& \{ E \in \mathbb{R}^{m\times n} : \ (A+E)^T (b - (A+E)\tilde{x}) = -c \}, \\
\mathcal{M} =& \{ v \left(\alpha c^T - v^\dagger A \right) + (I_m - vv^\dagger)(\tilde{r}\tilde{x}^\dagger + Z(I_n - \tilde{x}\tilde{x}^\dagger)) \ : \\
& v \in \mathbb{R}^m, \ Z \in \mathbb{R}^{m\times n}, \alpha \in \mathbb{R}, \ s.t. \ \alpha \|v\|^2 (v^\dagger b - \alpha c^T \tilde{x}) = -1 \}.
\end{aligned}
$$

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

Case $c = 0$

$$
\begin{aligned}
\mathcal{E} =& \{ E \in \mathbb{R}^{m\times n} : \ (A+E)^T (b - (A+E)\tilde{x}) = 0 \}, \\
\mathcal{M} =& \{ -vv^\dagger A + (I_m - vv^\dagger)(\tilde{r}\tilde{x}^\dagger + Z(I_n - \tilde{x}\tilde{x}^\dagger)) \ : v \in \mathbb{R}^m, \ Z \in \mathbb{R}^{m\times n} \}.
\end{aligned}
$$

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

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

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 $c = 0$

$$\min_{\mathcal{E}} \|E\|_F^2 = \frac{\|\tilde{r}\|^2}{\|\tilde{x}\|^2} + \min\{\lambda_*, 0\}, \quad \lambda_* = \lambda_{\min}\left(AA^T - \frac{\tilde{r}\tilde{r}^T}{\|\tilde{x}\|^2}\right).$$

Linearization estimate of $\eta(\tilde{x})$

Given $h(A, b, c, x) = A^T(b - Ax) + c$, find (E, f, g) such that

$$\bar{\eta}(\tilde{x}) = \min \|[E, f, g]\|_F \quad \text{s.t.} \quad h(A, b, c, \tilde{x}) + [J_A, J_b, J_c] \begin{bmatrix} \mathrm{vec}(E) \\ f \\ g \end{bmatrix} = 0,$$

where J_A, J_b and J_c are the Jacobian matrices of h with respect to $\mathrm{vec}(A)$, b, c.

Lemma

$$\bar{\eta}(\tilde{x}) = \left\| \begin{bmatrix} vec(E) \\ f \\ g \end{bmatrix} \right\| = \|J^{\dagger} h(A, b, c, \tilde{x})\|, \quad J := [I_n \otimes \tilde{r}^T - A^T(\tilde{x} \otimes I_m), A^T, I_n].$$

Moreover, assume that $\tilde{r} \neq 0$. If $4\sqrt{2 + \|\tilde{x}\|^2}\|J^{\dagger}\|\eta(\tilde{x}) \leq 1$, then

$$\frac{2}{1 + \sqrt{2}} \bar{\eta}(\tilde{x}) \leq \eta(\tilde{x}) \leq 2\, \bar{\eta}(\tilde{x}),$$

Comparison with CG: solution accuracy

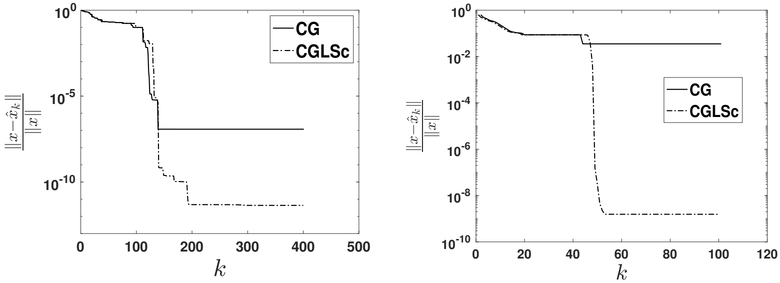


Figure: Left: $\kappa(A) = 10^5$. Right: $\kappa(A) = 5 \times 10^7$.