Verified Numerical Computations for Ill-Posed Optimization Problems

Viktor Härter
viktor.haerter@tuhh.de

Institute for Reliable Computing
Hamburg University of Technology

joint work with
Christian Jansson

8th GAMM-Workshop on
Applied and Numerical Linear Algebra
September, 11 - 12, 2008

Outline

1. Introduction
2. Conic programming universal form of convex optimization
3. Condition numbers in optimization
4. Verified error bounds

Outline

1. Introduction
2. Conic programming universal form of convex optimization
3. Condition numbers in optimization
4. Verified error bounds

Outline

1. Introduction
2. Conic programming universal form of convex optimization
3. Condition numbers in optimization
4. Verified error bounds

Outline

1. Introduction
2. Conic programming universal form of convex optimization
3. Condition numbers in optimization
4. Verified error bounds

Are state-of-the-art solvers reliable in optimization?

Example (by A. Neumaier and O. Shcherbina [4])

$$
\begin{array}{lll}
\min & -x_{20} & \\
\text{s.t.} & 7x_1 - x_2 \geq 5 & \\
 & -6x_{i-1} + 7x_i - x_{i+1} \geq (-1)^i 7 & \text{for } i = 2:19 \\
 & -6x_{18} - 17x_{19} + 3x_{20} \geq -23 & \\
 & 0 \leq x_i \leq 10 & \text{for } i = 1:13 \\
 & 0 \leq x_i \leq 10^6 & \text{for } i = 14:20 \\
 & \text{all } x_i \text{ integers} &
\end{array}
$$

solution: $x = (1, 2, 1, 2, \ldots, 1, 2)^T$

- **CPLEX**, **MINLP**, **GLPK**, **BONSAIG**, **XPRESS**: 'integer infeasible'
- Only **Fort MP** solved the problem correctly

Reason: ill-conditioned relaxations must be solved

Are state-of-the-art solvers reliable in optimization?

Example (by A. Neumaier and O. Shcherbina [4])

$$
\begin{array}{lll}
\min & -x_{20} & \\
\text{s.t.} & 7x_1 - x_2 \geq 5 & \\
 & -6x_{i-1} + 7x_i - x_{i+1} \geq (-1)^i 7 & \text{for } i = 2:19 \\
 & -6x_{18} - 17x_{19} + 3x_{20} \geq -23 & \\
 & 0 \leq x_i \leq 10 & \text{for } i = 1:13 \\
 & 0 \leq x_i \leq 10^6 & \text{for } i = 14:20 \\
 & \text{all } x_i \text{ integers} &
\end{array}
$$

solution: $x = (1, 2, 1, 2, \dots, 1, 2)^T$

- **CPLEX**, **MINLP**, **GLPK**, **BONSAIG**, **XPRESS**: ‘integer infeasible’
- Only **Fort MP** solved the problem correctly

Reason: ill-conditioned relaxations must be solved

Ill-posedness of real world applications

Are ill-conditioned or ill-posed problems rare in optimization?

- Ordóñez and Freund (2003):
  71% of **NETLIB lp** problems are ill-posed
- Freund, Ordóñez and Toh (2006):
  32 out of 80 problems of the **SDPLIB** are ill-posed

TUHH Technische Universität Hamburg-Harburg

Meaning of verified results

Definition (Verified results)

Verified, or sometimes also called rigorous results, means that the computed bounds are claimed to be valid with mathematical certainty even in the presence of rounding errors due to floating point arithmetic.

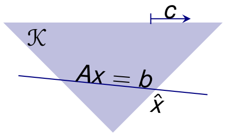
Review of Definition of a conic program

Definition (Conic Program (CP))

A linear conic program in standard form is defined as:

$$\hat{f} = \min\langle c, x\rangle \quad \text{s.t.} \quad Ax = b\,,\ x \in \mathcal{K}$$

- $A: \ \mathcal{X} \rightarrow \mathcal{Y}$ linear continuous operator
- $\mathcal{X}, \mathcal{Y}$ (infinite-dimensional) real normed vector spaces
- $\mathcal{K} \subseteq \mathcal{X}$ is a convex cone
- $c \in \mathcal{X}^*$ and $\langle c, x\rangle := c(x)$
- $\mathcal{X}^*$ the dual space of $\mathcal{X}$



Important problem classes of CP

In fact conic programming is an universal form of convex optimization.

Three important classes of conic programs:

- **Linear Programming LP**

$$\min c^T x \quad s.t. \quad Ax = b,\ x \in \mathcal{K} = \mathbb{R}^n_+$$

- **Second Order Cone Programming SOCP**

$$\mathcal{K} = \{x \in \mathbb{R}^n :\ x_1 \geq \|x_{2:n}\|\}$$

- **Semidefinite Programming SDP**

$$\mathcal{K} = \{X \in \mathbb{R}^{n\times n} :\ X \text{ positive semidefinite }\}$$

Important problem classes of CP

In fact conic programming is an universal form of convex optimization.

Three important classes of conic programs:

- **Linear Programming LP**

$$\min c^T x \quad s.t. \quad Ax = b\,,\ x \in \mathcal{K} = \mathbb{R}^n_+$$

- **Second Order Cone Programming SOCP**

$$\mathcal{K} = \{x \in \mathbb{R}^n :\ x_1 \geq \|x_{2:n}\|\}$$

- **Semidefinite Programming SDP**

$$\mathcal{K} = \{X \in \mathbb{R}^{n \times n} :\ X \ \text{positive semidefinite}\ \}$$

Important problem classes of CP

In fact conic programming is an universal form of convex optimization.

Three important classes of conic programs:

- **Linear Programming LP**

$$\min c^T x \quad s.t. \quad Ax = b\,,\; x \in \mathcal{K} = \mathbb{R}^n_+$$

- **Second Order Cone Programming SOCP**

$$\mathcal{K} = \{x \in \mathbb{R}^n : \; x_1 \geq \|x_{2:n}\|\}$$

- **Semidefinite Programming SDP**

$$\mathcal{K} = \{X \in \mathbb{R}^{n \times n} : \; X \;\text{ positive semidefinite }\}$$

TUHH Technische Universität Hamburg-Harburg

Important problem classes of CP

In fact conic programming is an universal form of convex optimization.

Three important classes of conic programs:

- **Linear Programming LP**

$$\min c^T x \quad s.t. \quad Ax = b\,,\ x \in \mathcal{K} = \mathbb{R}^n_+$$

- **Second Order Cone Programming SOCP**

$$\mathcal{K} = \{x \in \mathbb{R}^n :\ x_1 \geq \|x_{2:n}\|\}$$

- **Semidefinite Programming SDP**

$$\mathcal{K} = \{X \in \mathbb{R}^{n \times n} :\ X \text{ positive semidefinite }\}$$

Condition number for matrices

- relative distance to the next singular matrix

$$\varrho_I := \min\{\frac{\| \Delta A \|}{\| A \|} : A + \Delta A \text{ singular}\}$$

- condition for matrices: reciprocal of the relative distance to singularity

$$\kappa(A) := \|A\| \|A^{-1}\| = \frac{1}{\varrho_I}$$

TUHH
Technische Universität Hamburg-Harburg

Distances to infeasibility

Renegar ([5]) adopted above-mentioned understanding of condition measures to optimization problems.

Definition (Distance to primal/dual infeasibility)

Let $p := (A, b, c)$ denote the problem data of a conic program.

$\varrho_P(p)$: relative distance to primal infeasibility

$$\varrho_P(p) := \inf \left\{ \frac{\|\Delta p\|}{\|p\|} : \text{ problem } p + \Delta p \text{ is primal infeasible} \right\}$$

$\varrho_D(p)$: relative distance to dual infeasibility

$$\varrho_D(p) := \inf \left\{ \frac{\|\Delta p\|}{\|p\|} : \text{ problem } p + \Delta p \text{ is dual infeasible} \right\}$$

Renegar's condition number for CP

Definition (Condition number for CP)

Condition number of the problem instance p is the scale-invariant reciprocal of the smallest data perturbation Δp that will render the perturbed data instance $p + \Delta p$ either primal or dual infeasible

$$\kappa(p) := \frac{1}{\min\{\varrho_P(p), \varrho_D(p)\}}$$

Sensitivity and stability issues

- p is called ill-posed if $\min\{\varrho_P(p), \varrho_D(p)\} = 0$
- Theorem by Renegar [5]: the optimal value depends cubically on the condition number of the problem
- $\Rightarrow$ for very ill-conditioned problems very small data perturbations may yield arbitrarily large changes of the optimal value, but the optimality conditions (termination criteria) change only little, Reliability?

TUHH
Technische Universität Hamburg-Harburg

Sensitivity and stability issues

- p is called ill-posed if $\min\{\varrho_P(p), \varrho_D(p)\} = 0$
- Theorem by Renegar [5]: the optimal value depends cubically on the condition number of the problem
- $\Rightarrow$ for very ill-conditioned problems very small data perturbations may yield arbitrarily large changes of the optimal value, but the optimality conditions (termination criteria) change only little, Reliability?

Sensitivity and stability issues

- p is called ill-posed if $\min\{\varrho_P(p), \varrho_D(p)\} = 0$
- Theorem by Renegar [5]: the optimal value depends cubically on the condition number of the problem
- $\Rightarrow$ for very ill-conditioned problems very small data perturbations may yield arbitrarily large changes of the optimal value, but the optimality conditions (termination criteria) change only little, Reliability?

Notation and preliminaries

Each convex cone $\mathcal{K} \subseteq \mathcal{X}$ induces a partial ordering on $\mathcal{X}$:

$$a \preccurlyeq b \iff b - a \in \mathcal{K}$$

For $\{a, b\} \subseteq \mathcal{X}$ we define a lower bound $\underline{d}$ by:

$$\underline{d} \preceq \{a, b\} \iff \underline{d} \preccurlyeq a\,,\ \underline{d} \preccurlyeq b$$

We define the negative part of a vector d by:

$$\underline{d}^- \preccurlyeq \{d, 0\}$$

Example LP:

$$\underline{d}_i^- := \min\{d_i, 0\}$$

TUHH Technische Universität Hamburg-Harburg

Verified error bounds: lower bound

Theorem (Jansson (2007)*)

Given a Conic Program $p := (A, b, c)$. Let $\tilde{y} \in \mathcal{Y}^*$ and $d := -A^*\tilde{y} + c$. Suppose further that

$$\underline{d}^- \preccurlyeq \{d, 0\} \text{ (includes dual violations)}$$

If there is an optimal solution $\hat{x}$ for which an upper bound $\hat{x} \preccurlyeq \bar{x}$ can be found (primal boundedness qualification), then a lower bound for the primal optimal value is given by:

$$\hat{f} \geq \langle \tilde{y}, b \rangle + \langle \underline{d}^-, \overline{x} \rangle =: \underline{f}$$

* More general results can be found in [1].

Example LP: rigorous lower bound

Consider the linear program in standard form:

$$\hat{f} = \min c^T x \quad s.t. \quad Ax = b\,,\ x \in \mathbb{R}^n_+\,,\ \hat{x} \leq \overline{x}$$

- determine with a linear solver an approximate solution $\tilde{y}$
- compute the rigorous lower bound using directed rounding:

$$\begin{aligned} \hat{f} &\geqslant \tilde{y}^T b + (\underline{d}^-)^T \overline{x} && \text{with} \\ \underline{d}_j^- &= \min\{(-A^T\tilde{y} + b)_j, 0\} && j = 1 \dots n \end{aligned}$$

- asymptotic complexity for LP lower bound is $O(mn)$, compared with $O(m^2 n)$ operations for LP solvers

Example LP: rigorous lower bound

Consider the linear program in standard form:

$$\hat{f} = \min c^T x \quad s.t. \quad Ax = b\,,\; x \in \mathbb{R}^n_+\,,\; \hat{x} \leq \overline{x}$$

- determine with a linear solver an approximate solution $\tilde{y}$
- compute the rigorous lower bound using directed rounding:

$$\begin{array}{rcl} \hat{f} \geqslant \tilde{y}^T b + (\underline{d}^-)^T \overline{x} & & \text{with} \\ \underline{d}_j^- = \min\{(-A^T \tilde{y} + b)_j, 0\} & & j = 1 \dots n \end{array}$$

- asymptotic complexity for LP lower bound is $O(mn)$, compared with $O(m^2 n)$ operations for LP solvers

TUHH Technische Universität Hamburg-Harburg

Example LP: rigorous lower bound

Consider the linear program in standard form:

$$\hat{f} = \min c^T x \quad s.t. \quad Ax = b\,,\; x \in \mathbb{R}^n_+\,,\; \hat{x} \leq \overline{x}$$

- determine with a linear solver an approximate solution $\tilde{y}$
- compute the rigorous lower bound using directed rounding:

$$\begin{aligned} \hat{f} &\geqslant \tilde{y}^T b + (\underline{d}^-)^T \overline{x} && \text{with} \\ \underline{d}_j^- &= \min\{(-A^T \tilde{y} + b)_j, 0\} && j = 1 \dots n \end{aligned}$$

- asymptotic complexity for LP lower bound is $O(mn)$, compared with $O(m^2 n)$ operations for LP solvers

Example LP: rigorous lower bound

Consider the linear program in standard form:

$$\hat{f} = \min c^T x \quad s.t. \quad Ax = b\,,\ x \in \mathbb{R}^n_+\,,\ \hat{x} \leq \overline{x}$$

- determine with a linear solver an approximate solution $\tilde{y}$
- compute the rigorous lower bound using directed rounding:

$$\begin{aligned} \hat{f} &\geqslant \tilde{y}^T b + (\underline{d}^-)^T \overline{x} && \text{with} \\ \underline{d}_j^- &= \min\{(-A^T \tilde{y} + b)_j, 0\} && j = 1 \ldots n \end{aligned}$$

- asymptotic complexity for LP lower bound is $O(mn)$, compared with $O(m^2 n)$ operations for LP solvers

TUHH Technische Universität Hamburg-Harburg

Verification of regularization methods

Consider the convex constrained least squares problem: [1]

$$\hat{f} = \min \|Ax - b\| \quad s.t. \quad \|Lx\| \leq \Delta$$

- this least squares problem can be easy reformulated as a SOCP
- lower bound: (Theorem)
- upper bound: $\overline{f} = \|A\tilde{x}_p - b\|$

$$\Rightarrow \ 0 \leq \| A\tilde{x}_p - b \| - \hat{f} \leq \ \|A\tilde{x}_p - b\| - \underline{f}$$

[1] many regularization methods can be put into this framework,

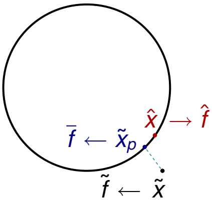
Verification of regularization methods

Consider the convex constrained least squares problem: [1]

$$\hat{f} = \min \|Ax - b\| \quad s.t. \quad \|Lx\| \leq \Delta$$

Let $\tilde{x}$ be an approximate solution.

- this least squares problem can be easy reformulated as a SOCP
- lower bound: (Theorem)
- upper bound: $\bar{f} = \|A\tilde{x}_p - b\|$



$$\Rightarrow \; 0 \leq \| A\tilde{x}_p - b \| - \hat{f} \leq \|A\tilde{x}_p - b\| - \underline{f}$$

[1] many regularization methods can be put into this framework,

Verification of regularization methods

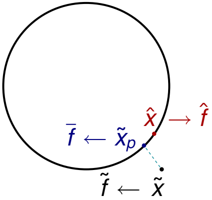
Consider the convex constrained least squares problem: [1]

$$\hat{f} = \min \|Ax - b\| \quad s.t. \quad \|Lx\| \leq \Delta$$

Actually there is no guarantee about the accuracy of $\tilde{x}$:

$$\|A\tilde{x} - b\| - \hat{f} \leq ?$$

- this least squares problem can be easy reformulated as a SOCP
- lower bound: (Theorem)
- upper bound: $\bar{f} = \|A\tilde{x}_p - b\|$



$$\Rightarrow \; 0 \leq \| A\tilde{x}_p - b \| - \hat{f} \leq \|A\tilde{x}_p - b\| - \underline{f}$$

[1] many regularization methods can be put into this framework,

Verification of regularization methods

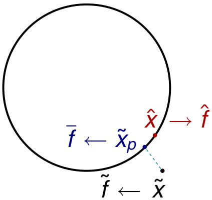
Consider the convex constrained least squares problem: [1]

$$\hat{f} = \min \|Ax - b\| \quad s.t. \quad \|Lx\| \leq \Delta$$

Actually there is no guarantee about the accuracy of $\tilde{x}$:

$$\|A\tilde{x} - b\| - \hat{f} \leq ?$$

- this least squares problem can be easy reformulated as a SOCP
- lower bound: (Theorem)
- upper bound: $\bar{f} = \|A\tilde{x}_p - b\|$



$$\Rightarrow \; 0 \leq \| A\tilde{x}_p - b \| - \hat{f} \leq \|A\tilde{x}_p - b\| - \underline{f}$$

[1] many regularization methods can be put into this framework,

Verification of regularization methods

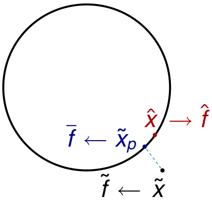
Consider the convex constrained least squares problem: [1]

$$\hat{f} = \min \|Ax - b\| \quad s.t. \quad \|Lx\| \leq \Delta$$

Actually there is no guarantee about the accuracy of $\tilde{x}$:

$$\|A\tilde{x} - b\| - \hat{f} \leq ?$$

- this least squares problem can be easy reformulated as a SOCP
- lower bound: (Theorem)
- upper bound: $\bar{f} = \|A\tilde{x}_p - b\|$



$$\Rightarrow \; 0 \leq \| A\tilde{x}_p - b \| - \hat{f} \leq \|A\tilde{x}_p - b\| - \underline{f}$$

[1] many regularization methods can be put into this framework,

Verification of regularization methods

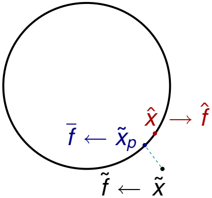
Consider the convex constrained least squares problem: [1]

$$\hat{f} = \min \|Ax - b\| \quad s.t. \quad \|Lx\| \leq \Delta$$

Actually there is no guarantee about the accuracy of $\tilde{x}$:

$$\|A\tilde{x} - b\| - \hat{f} \leq ?$$

- this least squares problem can be easy reformulated as a SOCP
- lower bound: (Theorem)
- upper bound: $\bar{f} = \|A\tilde{x}_p - b\|$



$$\Rightarrow \; 0 \leq \|A\tilde{x}_p - b\| - \hat{f} \leq \|A\tilde{x}_p - b\| - \underline{f}$$

[1] many regularization methods can be put into this framework,

Verification of regularization methods

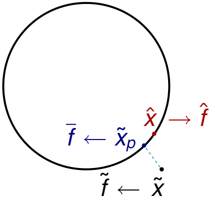
Consider the convex constrained least squares problem: [1]

$$\hat{f} = \min \|Ax - b\| \quad s.t. \quad \|Lx\| \le \Delta$$

Actually there is no guarantee about the accuracy of $\tilde{x}$:

$$\|A\tilde{x} - b\| - \hat{f} \le \, ?$$

- this least squares problem can be easy reformulated as a SOCP
- lower bound: (Theorem)
- upper bound: $\bar{f} = \|A\tilde{x}_p - b\|$



$$\Rightarrow \; 0 \le \| A\tilde{x}_p - b \| - \hat{f} \le \|A\tilde{x}_p - b\| - \underline{f}$$

[1] many regularization methods can be put into this framework,

Available software for verified computations

Two easy to use software packages are available:

- **Lurupa:** Verified Linear Programming by C. Keil ([3])
- **VSDP:** Verified Semidefinite Programming by C. Jansson ([2])

Available software for verified computations

Two easy to use software packages are available:

- **Lurupa:** Verified Linear Programming by C. Keil ([3])
- **VSDP:** Verified Semidefinite Programming by C. Jansson ([2])

Available software for verified computations

Two easy to use software packages are available:

- **Lurupa:** Verified Linear Programming by C. Keil ([3])
- **VSDP:** Verified Semidefinite Programming by C. Jansson ([2])

TUHH
Technische Universität Hamburg-Harburg

Numerical Results of Lurupa for NETLIB LP

Numerical results of Lurupa:

- summarizing: for all problems with zero distance to primal or dual infeasibility finite lower respectively upper bounds are computed
- the relative error varies between 10^{-8} and 10^{-16}
- median of time ratios: $(\underline{t}/t) = 0.5$ and $(\bar{t}/t) := 5.3$

Numerical Results of Lurupa for NETLIB LP

Numerical results of Lurupa:

- summarizing: for all problems with zero distance to primal or dual infeasibility finite lower respectively upper bounds are computed
- the relative error varies between 10^{-8} and 10^{-16}
- median of time ratios: $(\underline{t}/t)= 0.5$ and $(\bar{t}/t): = 5.3$

Numerical Results of Lurupa for NETLIB LP

Numerical results of Lurupa:

- summarizing: for all problems with zero distance to primal or dual infeasibility finite lower respectively upper bounds are computed
- the relative error varies between 10^{-8} and 10^{-16}
- median of time ratios: $(\underline{t}/t) = 0.5$ and $(\bar{t}/t): = 5.3$

Numerical Results of Lurupa for NETLIB LP

Numerical results of Lurupa:

- summarizing: for all problems with zero distance to primal or dual infeasibility finite lower respectively upper bounds are computed
- the relative error varies between 10^{-8} and 10^{-16}
- median of time ratios: $(\underline{t}/t) = 0.5$ and $(\bar{t}/t): = 5.3$

Numerical Results of VSDP for SDPLIB

Numerical results of VSDP:

- for all well-posed problems rigorous bounds could be computed
- for ill-posed problems infinite upper bounds are computed, reflecting the difficulty of these problems
- median of guaranteed relative accuracy: $4.9 \cdot 10^{-7}$
- median of time ratios: $(\underline{t}/t)= 0.045$ and $(\bar{t}/t) := 2.4$

Numerical Results of VSDP for SDPLIB

Numerical results of VSDP:

- for all well-posed problems rigorous bounds could be computed
- for ill-posed problems infinite upper bounds are computed, reflecting the difficulty of these problems
- median of guaranteed relative accuracy: $4.9 \cdot 10^{-7}$
- median of time ratios: $(\underline{t}/t) = 0.045$ and $(\bar{t}/t) := 2.4$

Numerical Results of VSDP for SDPLIB

Numerical results of VSDP:

- for all well-posed problems rigorous bounds could be computed
- for ill-posed problems infinite upper bounds are computed, reflecting the difficulty of these problems
- median of guaranteed relative accuracy: $4.9 \cdot 10^{-7}$
- median of time ratios: $(\underline{t}/t) = 0.045$ and $(\bar{t}/t) := 2.4$

Numerical Results of VSDP for SDPLIB

Numerical results of VSDP:

- for all well-posed problems rigorous bounds could be computed
- for ill-posed problems infinite upper bounds are computed, reflecting the difficulty of these problems
- median of guaranteed relative accuracy: $4.9 \cdot 10^{-7}$
- median of time ratios: $(\underline{t}/t)$= 0.045 and $(\bar{t}/t)$: = 2.4

Numerical Results of VSDP for SDPLIB

Numerical results of VSDP:

- for all well-posed problems rigorous bounds could be computed
- for ill-posed problems infinite upper bounds are computed, reflecting the difficulty of these problems
- median of guaranteed relative accuracy: $4.9 \cdot 10^{-7}$
- median of time ratios: $(\underline{t}/t) = 0.045$ and $(\bar{t}/t)\!: = 2.4$

Summary

- ill-conditioned or ill-posed optimization problems are not rare in practice
- rigorous results are provided even for very ill-conditioned or ill-posed problems
- methods and software for computing verified results are available

Summary

- ill-conditioned or ill-posed optimization problems are not rare in practice
- rigorous results are provided even for very ill-conditioned or ill-posed problems
- methods and software for computing verified results are available

Summary

- ill-conditioned or ill-posed optimization problems are not rare in practice
- rigorous results are provided even for very ill-conditioned or ill-posed problems
- methods and software for computing verified results are available

References

- C. Jansson.
  Guaranteed accuracy for conic programming problems in vector lattices, 2007, submitted.
- C. Jansson.
  VSDP: Verified SemiDefinite Programming.
  http://www.ti3.tu-harburg.de/jansson/vsdp/
- C. Keil.
  Lurupa - Rigorous Error Bounds in Linear Programming.
  http://www.ti3.tu-harburg.de/~keil/
- A. Neumaier and O. Shcherbina.
  Safe bounds in linear and mixed-integer linear programming.
  Math. Program., 99(2):283–296, 2004.
- J. Renegar.
  Some perturbation theory for linear programming.
  Math. Program., 65:73–91, 1994.

TUHH
Technische Universität Hamburg-Harburg

Any questions?

Thank You for Your Attention!!!