The Regularized Total Least Squares Problem: Theoretical Properties and Three Globally Convergent Algorithms

Amir Beck

Faculty of Industrial Engineering and Management

Technion, Haifa, Israel

4th International Workshop on TLS and Errors-in-Variables Modeling

Leuven, Belguim, August 21-23 2006

Bibliography

The lecture is based on the three papers:

D. Sima, S. Van Huffel and G. H. Golub. **”Regularized Total Least Squares Based on Quadratic Eigenvalue Problem Solvers”** *BIT* 44(4):793–812, 2004.

A. Beck, A. Ben-Tal and and M.Teboulle. **”Finding a Global Optimal Solution for a Quadratically Constrained Fractional Quadratic Problem with Applications to the Regularized Total Least Squares”***SIMAX* 28(2):425-445, 2006.

A. Beck and M.Teboulle. **”A Convex Optimization Approach for Minimizing the Ratio of Indefinite Quadratic Functions over an Ellipsoid”***Submitted for publication*.

Total Least Squares - Review

$$\mathbf{A}\mathbf{x} \approx \mathbf{b}$$

A fixed - LS

$$\min_{\mathbf{w},\mathbf{x}} \|\mathbf{w}\|^2$$

s.t.

$$\mathbf{A}\mathbf{x} = \mathbf{b} + \mathbf{w}$$

minimal perturbation to rhs which makes this linear system consistent

Total Least Squares - Review

$$\mathbf{A}\mathbf{x} \approx \mathbf{b}$$

A fixed - LS

$$\min_{\mathbf{w},\mathbf{x}} \|\mathbf{w}\|^2$$
s.t.
$$\mathbf{A}\mathbf{x} = \mathbf{b} + \mathbf{w}$$

minimal perturbation to rhs which makes this linear system consistent

A uncertain - TLS (TLS)

$$\min_{\mathbf{w},\mathbf{E},\mathbf{x}} \|\mathbf{E}\|^2 + \|\mathbf{w}\|^2$$
s.t.
$$(\mathbf{A} + \mathbf{E})\mathbf{x} = \mathbf{b} + \mathbf{w}$$

minimal perturbation to both rhs and lhs matrix which makes the system consistent (Golub, Van Loan (80))

Another Formulation of the TLS

(Golub, Van Loan, 80)

$$(TLS) \quad \min_{\mathbf{x},\mathbf{E},\mathbf{w}} \{\|\mathbf{E}\|^2 + \|\mathbf{w}\|^2 : \mathbf{b} + \mathbf{w} = (\mathbf{A} + \mathbf{E})\mathbf{x}\} =$$

Another Formulation of the TLS

(Golub, Van Loan, 80)

$$(TLS) \quad \min_{\mathbf{x},\mathbf{E},\mathbf{w}} \{\|\mathbf{E}\|^2 + \|\mathbf{w}\|^2 : \mathbf{b} + \mathbf{w} = (\mathbf{A} + \mathbf{E})\mathbf{x}\} =$$

$$= \underbrace{\min_{\mathbf{x}} \frac{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2}{\|\mathbf{x}\|^2 + 1}}_{\text{A nonconvex optimization problem}}$$

Another Formulation of the TLS

(Golub, Van Loan, 80)

$$(TLS) \quad \min_{\mathbf{x},\mathbf{E},\mathbf{w}} \{\|\mathbf{E}\|^2 + \|\mathbf{w}\|^2 : \mathbf{b} + \mathbf{w} = (\mathbf{A} + \mathbf{E})\mathbf{x}\} =$$

$$= \underbrace{\min_{\mathbf{x}} \frac{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2}{\|\mathbf{x}\|^2 + 1}}_{\text{A nonconvex optimization problem}}$$

The solution is expressed by the singular value decomposition of the augmented matrix $(\mathbf{A}, \mathbf{b})$

Regularization of the TLS solution

- Regularization of the TLS solution is required in the case where $\mathbf{A}$ is nearly rank deficient.
- **Applications:** discretization of ill posed problems, image deblurring , medical applications, signal restoration...In these problems, TLS solution can be physically meaningless, hence regularization is needed to stabilize solution.

Regularization of the TLS solution

- Regularization of the TLS solution is required in the case where **A** is nearly rank deficient.
- **Applications:** discretization of ill posed problems, image deblurring , medical applications, signal restoration...In these problems, TLS solution can be physically meaningless, hence regularization is needed to stabilize solution.

Regularization Methods

- Addition of a quadratic constraint. *([Golub, Hansen & O'leary, 1999], [Guo & Renaut, 2002, 2005], [Sima, Van Huffel & Golub, 2004], [Beck, Ben-Tal & Teboulle 2006], [Beck & Teboulle 2006])*
- Addition of a quadratic penalty to the objective function [Beck & Ben-Tal 2005].
- Truncation methods. *([Fierro, Golub Hansen & O'leary, 1997], [Hansen, 1994])*.

The RTLS problem

$$\text{(RTLS):}\ \min\left\{\frac{\|\mathbf{Ax}-\mathbf{b}\|^2}{1+\|\mathbf{x}\|^2} : \|\mathbf{Lx}\|^2 \le \rho\right\}$$

- $\mathbf{L} \in \mathbb{R}^{r\times n}(r \le n)$ has full row rank.

The RTLS problem

$$\text{(RTLS):}\ \min\left\{\frac{\|\mathbf{Ax}-\mathbf{b}\|^2}{1+\|\mathbf{x}\|^2} : \|\mathbf{Lx}\|^2 \le \rho\right\}$$

- $\mathbf{L} \in \mathbb{R}^{r\times n}(r \le n)$ has full row rank.
- The feasible set $\{\mathbf{x} : \|\mathbf{Lx}\|^2 \le \rho\}$ represents a (possibly degenerate) ellipsoid.

The RTLS problem

$$\text{(RTLS):}\ \min\left\{\frac{\|\mathbf{A}\mathbf{x}-\mathbf{b}\|^2}{1+\|\mathbf{x}\|^2} : \|\mathbf{L}\mathbf{x}\|^2 \le \rho\right\}$$

- $\mathbf{L} \in \mathbb{R}^{r\times n}(r \le n)$ has full row rank.
- The feasible set $\{\mathbf{x} : \|\mathbf{L}\mathbf{x}\|^2 \le \rho\}$ represents a (possibly degenerate) ellipsoid.
- Popular choices for $\mathbf{L}$: identity matrix, an approximation of first or second order derivative.

The RTLS problem

$$\text{(RTLS):}\ \min\left\{\frac{\|\mathbf{Ax}-\mathbf{b}\|^2}{1+\|\mathbf{x}\|^2} : \|\mathbf{Lx}\|^2 \le \rho\right\}$$

- $\mathbf{L} \in \mathbb{R}^{r\times n}(r \le n)$ has full row rank.
- The feasible set $\{\mathbf{x} : \|\mathbf{Lx}\|^2 \le \rho\}$ represents a (possibly degenerate) ellipsoid.
- Popular choices for $\mathbf{L}$: identity matrix, an approximation of first or second order derivative.
- A nonconvex optimization problem (the objective function is nonconvex).

Main Problem

Minimization of a ratio of indefinite quadratic functions over an Ellipsoid

$$(RQ) \quad \min\left\{\frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} : \|\mathbf{L}\mathbf{x}\|^2 \le \rho\right\}$$

$$f_i(\mathbf{x}) = \mathbf{x}^T\mathbf{A}_i\mathbf{x} + 2\mathbf{b}_i^T\mathbf{x} + c_i, i = 1, 2$$

$$\mathbf{A}_i = \mathbf{A}_i^T \in \mathbb{R}^{n\times n}, \mathbf{b}_i \in \mathbb{R}^n, c_i \in \mathbb{R}, \mathbf{L} \in \mathbb{R}^{r\times n}$$

Main Problem

Minimization of a ratio of indefinite quadratic functions over an Ellipsoid

$$(RQ) \quad \min\left\{\frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} : \|\mathbf{L}\mathbf{x}\|^2 \le \rho\right\}$$

$$f_i(\mathbf{x}) = \mathbf{x}^T\mathbf{A}_i\mathbf{x} + 2\mathbf{b}_i^T\mathbf{x} + c_i, i = 1, 2$$

$$\mathbf{A}_i = \mathbf{A}_i^T \in \mathbb{R}^{n\times n}, \mathbf{b}_i \in \mathbb{R}^n, c_i \in \mathbb{R}, \mathbf{L} \in \mathbb{R}^{r\times n}$$

Assumption: the problem is well defined, i.e., $f_2(\mathbf{x}) > 0$ for every $\mathbf{x}$ such that $\|\mathbf{L}\mathbf{x}\|^2 \le \rho$

First Subclass: GTRS Problems

Generalized Trust Region Subproblem (GTRS): $f_2(\mathbf{x}) \equiv 1$

$$(GTRS)\,\min\{\mathbf{x}^T\mathbf{A}_1\mathbf{x} + 2\mathbf{b}_1^T\mathbf{x} + c_1 : \|\mathbf{L}\mathbf{x}\|^2 \le \rho\}$$

- A nonconvex problem.

First Subclass: GTRS Problems

Generalized Trust Region Subproblem (GTRS): $f_2(\mathbf{x}) \equiv 1$

$$(GTRS) \min\{\mathbf{x}^T\mathbf{A}_1\mathbf{x} + 2\mathbf{b}_1^T\mathbf{x} + c_1 : \|\mathbf{L}\mathbf{x}\|^2 \leq \rho\}$$

- A nonconvex problem.
- Nonetheless, can be efficiently solved for large-scale problems (Moré & Sorensen 83, Moré 93, Stern & Wolkowicz 95, Ben-Tal & Teboulle 96 Fortin & Wolkowicz 04).

First Subclass: GTRS Problems

Generalized Trust Region Subproblem (GTRS): $f_2(\mathbf{x}) \equiv 1$

$$(GTRS) \min\{\mathbf{x}^T\mathbf{A}_1\mathbf{x} + 2\mathbf{b}_1^T\mathbf{x} + c_1 : \|\mathbf{L}\mathbf{x}\|^2 \leq \rho\}$$

- A nonconvex problem.
- Nonetheless, can be efficiently solved for large-scale problems (Moré & Sorensen 83, Moré 93, Stern & Wolkowicz 95, Ben-Tal & Teboulle 96 Fortin & Wolkowicz 04).
- A key subproblem in Trust Region Algorithms for unconstrained minimization problems

$$\min\{f(x) : \mathbf{x} \in \mathbb{R}^n\} \Rightarrow \mathbf{x}^{k+1} \in \underset{\|\mathbf{x}-\mathbf{x}^k\|^2 \leq \Delta}{\text{argmin}} g^k(\mathbf{x})$$

Second Subclass: RTLS Problems

Regularized Total Least Squares Problem (RTLS):

$f_1(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|^2, f_2(\mathbf{x}) = \|\mathbf{x}\|^2 + 1$

$$(RTLS) \min \left\{ \frac{\|\mathbf{Ax} - \mathbf{b}\|^2}{\|\mathbf{x}\|^2 + 1} : \|\mathbf{Lx}\|^2 \leq \rho \right\}$$

- A nonconvex problem (although both the denominator and nominator are convex functions).

The iterative scheme of Sima, Van Huffel and Golub

Optimality conditions: $\mathbf{x}^*$ is a global optimal solution if and only if

$$\mathbf{x}^* \in \mathrm{argmin}\{f_2(\mathbf{x})(f(\mathbf{x}) - f(\mathbf{x}^*)) : \|\mathbf{L}\mathbf{x}\|^2 \leq \rho\}, \left(f(\mathbf{x}) \equiv \frac{f_1(\mathbf{x})}{f_2(\mathbf{x})}\right)$$

The iterative scheme of Sima, Van Huffel and Golub

Optimality conditions: $\mathbf{x}^*$ is a global optimal solution if and only if

$$\mathbf{x}^* \in \operatorname{argmin}\{f_2(\mathbf{x})(f(\mathbf{x}) - f(\mathbf{x}^*)) : \|\mathbf{L}\mathbf{x}\|^2 \le \rho\}, \left(f(\mathbf{x}) \equiv \frac{f_1(\mathbf{x})}{f_2(\mathbf{x})}\right)$$

Fixed Point Iterations:

$$\mathbf{x}^{k+1} \in \operatorname{argmin}\{f_2(\mathbf{x})(f(\mathbf{x}) - f(\mathbf{x}^k)) : \|\mathbf{L}\mathbf{x}\|^2 \le \rho\}$$

Equivalently:

$$\mathbf{x}^{k+1} \in \operatorname{argmin}\{f_1(\mathbf{x}) - f(\mathbf{x}^k) f_2(\mathbf{x}) : \|\mathbf{L}\mathbf{x}\|^2 \le \rho\}$$

Each iteration involves the solution of a nonconvex GTRS.

Solving the GTRS problem

$$(P): \min\{\mathbf{x}^T\mathbf{B}\mathbf{x} - 2\mathbf{d}^T\mathbf{x} : \|\mathbf{L}\mathbf{x}\|^2 = \rho\}$$

Two solution approaches:

- Formulation as a Quadratic Eigenvalue problem:

$$(\lambda^2\mathbf{I} + 2\lambda\mathbf{W} + \mathbf{W}^2 - \rho\mathbf{h}\mathbf{h}^T)\mathbf{u} = 0,$$

where $\mathbf{W} = \mathbf{L}^{-T}\mathbf{B}\mathbf{L}^{-1}, \mathbf{h} = \mathbf{L}^{-T}\mathbf{d}$.

Solving the GTRS problem

$$(P): \min\{\mathbf{x}^T\mathbf{B}\mathbf{x} - 2\mathbf{d}^T\mathbf{x} : \|\mathbf{L}\mathbf{x}\|^2 = \rho\}$$

Two solution approaches:

- Formulation as a Quadratic Eigenvalue problem:

  $$(\lambda^2\mathbf{I} + 2\lambda\mathbf{W} + \mathbf{W}^2 - \rho\mathbf{h}\mathbf{h}^T)\mathbf{u} = 0,$$

  where $\mathbf{W} = \mathbf{L}^{-T}\mathbf{B}\mathbf{L}^{-1}, \mathbf{h} = \mathbf{L}^{-T}\mathbf{d}$.

- A dual approach:
  The dual problem:

  $$(D) \quad \max\{-\mathbf{d}^T(\mathbf{B} + \lambda\mathbf{L}^T\mathbf{L})^{-1}\mathbf{d} - \lambda\rho : \lambda \geq -\lambda_{\min}(\mathbf{L}^{-T}\mathbf{B}\mathbf{L})\}$$

  Strong duality: val(P)=val(D).

The iterative scheme of Sima, Van Huffel and Golub

- Designed to solve (RTLS) and not (RQ).

The iterative scheme of Sima, Van Huffel and Golub

- Designed to solve (RTLS) and not (RQ).
- In the case $r < n$, the initial vector is carefully chosen.

The iterative scheme of Sima, Van Huffel and Golub

- Designed to solve (RTLS) and not (RQ).
- In the case $r < n$, the initial vector is carefully chosen.
- Proof that any limit point of the generated sequence satisfies first order optimality conditions. No proof of global convergence.

The iterative scheme of Sima, Van Huffel and Golub

- Designed to solve (RTLS) and not (RQ).
- In the case $r < n$, the initial vector is carefully chosen.
- Proof that any limit point of the generated sequence satisfies first order optimality conditions. No proof of global convergence.
- Numerical experiments: convergence in at most 5 iterations to a high accuracy vector.

The iterative scheme of Sima, Van Huffel and Golub

- Designed to solve (RTLS) and not (RQ).
- In the case $r < n$, the initial vector is carefully chosen.
- Proof that any limit point of the generated sequence satisfies first order optimality conditions. No proof of global convergence.
- Numerical experiments: convergence in at most 5 iterations to a high accuracy vector.
- The numerical experiments suggest that the algorithm converges to a global optimum

The iterative scheme of Sima, Van Huffel and Golub

- Designed to solve (RTLS) and not (RQ).
- In the case $r < n$, the initial vector is carefully chosen.
- Proof that any limit point of the generated sequence satisfies first order optimality conditions. No proof of global convergence.
- Numerical experiments: convergence in at most 5 iterations to a high accuracy vector.
- The numerical experiments suggest that the algorithm converges to a global optimum
- Question 1: Does the algorithm converge to a global optimum for (RTLS)? (RQ)?

The iterative scheme of Sima, Van Huffel and Golub

- Designed to solve (RTLS) and not (RQ).
- In the case $r < n$, the initial vector is carefully chosen.
- Proof that any limit point of the generated sequence satisfies first order optimality conditions. No proof of global convergence.
- Numerical experiments: convergence in at most 5 iterations to a high accuracy vector.
- The numerical experiments suggest that the algorithm converges to a global optimum
- Question 1: Does the algorithm converge to a global optimum for (RTLS)? (RQ)?
- Question 2: What is the reason for the small number of iterations?

A Globally Convergent Algorithm

Dinkelbach's principal for fractional programming (67)

$$F(\alpha) = \min\{f_1(\mathbf{x}) - \alpha f_2(\mathbf{x}) : \|\mathbf{L}\mathbf{x}\|^2 \le \rho\}$$

- F is a decreasing function of α.

A Globally Convergent Algorithm

Dinkelbach's principal for fractional programming (67)

$$F(\alpha) = \min\{f_1(\mathbf{x}) - \alpha f_2(\mathbf{x}) : \|\mathbf{L}\mathbf{x}\|^2 \leq \rho\}$$

- F is a decreasing function of α.
- α^* is the optimal value if and only if $F(\alpha^*) = 0$.

A Globally Convergent Algorithm

Dinkelbach's principal for fractional programming (67)

$$F(\alpha) = \min\{f_1(\mathbf{x}) - \alpha f_2(\mathbf{x}) : \|\mathbf{Lx}\|^2 \le \rho\}$$

- F is a decreasing function of α.
- α^* is the optimal value if and only if $F(\alpha^*) = 0$.

Outer Bisection Algorithm (Beck, Ben-Tal, Teboulle, 06)

Initialization: α_l, α_u - lower and upper bounds on α^*.
while $\alpha_u - \alpha_l > \epsilon$ repeat
$\alpha_h = \frac{\alpha_u+\alpha_l}{2}$
If $F(\alpha_h) > 0$ then $\alpha_u = \alpha_h$, else $\alpha_l = \alpha_h$

Analysis of the Outer Bisection Algorithm

- The algorithm and analysis relate to the (RQ) problem

Analysis of the Outer Bisection Algorithm

- The algorithm and analysis relate to the (RQ) problem
- Each iteration requires the solution of a GTRS problem.

Analysis of the Outer Bisection Algorithm

- The algorithm and analysis relate to the (RQ) problem
- Each iteration requires the solution of a GTRS problem.
- An ϵ-global optimal solution is obtained after $O(\log(1/\epsilon))$ iterations.

Analysis of the Outer Bisection Algorithm

- The algorithm and analysis relate to the (RQ) problem
- Each iteration requires the solution of a GTRS problem.
- An ϵ-global optimal solution is obtained after $O(\log(1/\epsilon))$ iterations.
- Acceleration of the algorithm is made by using the following simple fact:

for each feasible $\tilde{\mathbf{x}}$ one has $\alpha^* < f(\tilde{\mathbf{x}})$

Analysis of the Outer Bisection Algorithm

- The algorithm and analysis relate to the (RQ) problem
- Each iteration requires the solution of a GTRS problem.
- An ϵ-global optimal solution is obtained after $O(\log(1/\epsilon))$ iterations.
- Acceleration of the algorithm is made by using the following simple fact:

for each feasible $\tilde{\mathbf{x}}$ one has $\alpha^* < f(\tilde{\mathbf{x}})$

- Extension: Nonconvex feasible set $\{m \leq \|\mathbf{L}\mathbf{x}\|^2 \leq M\}$. Usage of the hidden convexity property of problems of the form:

$$\min\{\mathbf{x}^T\mathbf{B}\mathbf{x} - 2\mathbf{d}^T\mathbf{x} : m \leq \|\mathbf{L}\mathbf{x}\|^2 \leq M\}$$

Image Deblurring Example

- **Problem: estimate a 32×32 two dimensional image obtained from the sum of three harmonic oscillations:**

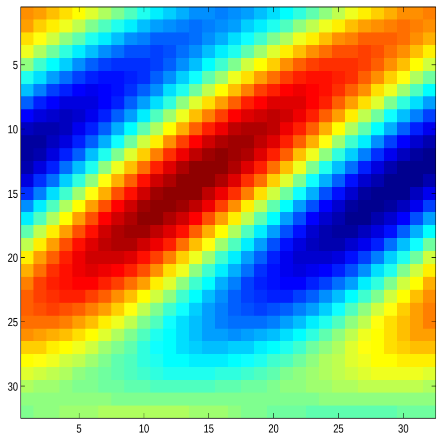
$$\mathbf{x}(\mathbf{z_1}, \mathbf{z_2}) = \sum_{\mathbf{l}=1}^{3} \mathbf{a_i} \cos(\mathbf{w_{l,1}} \mathbf{z_1} + \mathbf{w_{l,2}} \mathbf{z_2} + \phi_{\mathbf{l}}), \quad \left(\mathbf{w_{l,i}} = \frac{2\pi \mathbf{k_{l,i}}}{\mathbf{n}} \right),$$

where $1 \le \mathbf{z_1}, \mathbf{z_2} \le \mathbf{32}$, $\mathbf{k_{l,i}} \in \mathbb{Z}^2$, and $\mathbf{a_i}, \phi_{\mathbf{l}}$ given parameters.

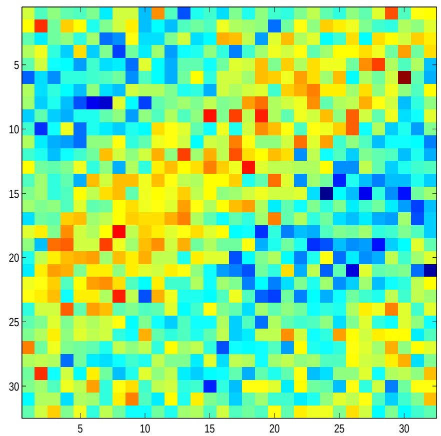
- The image is blurred by atmospheric turbulence blur which results with a highly noisy image (see Fig. B).
- We ran algorithms and show the results for:
  • RLS with standard regularization ($\mathbf{L} = \mathbf{I}$).
  • RLS with $\mathbf{L}$ as a discrete approximation of the Laplace operator, which is standard in image processing .
  • TTLS and our algorithm RTLSC.

Results for Regularization Solvers: RLS

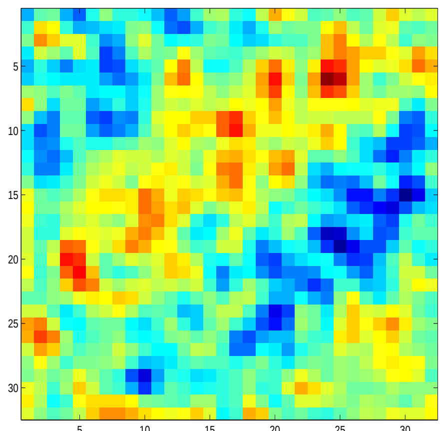
(A) True Image



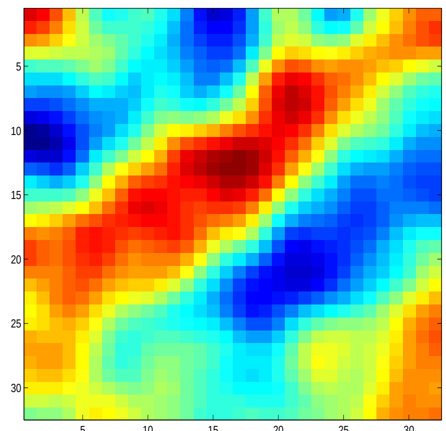
(B) Observation



(C) RLS with $\mathbf{L} = \mathbf{I}$

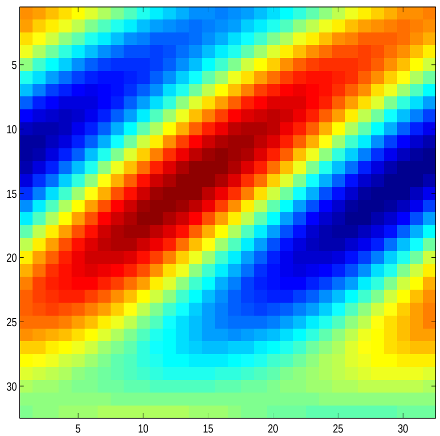


(D) RLS with Laplace operator

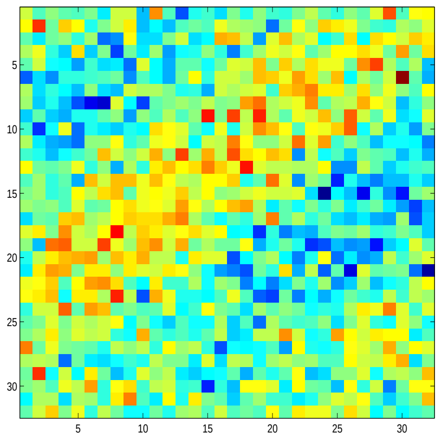


Results for Regularization Solvers: TTLS and RTLSC

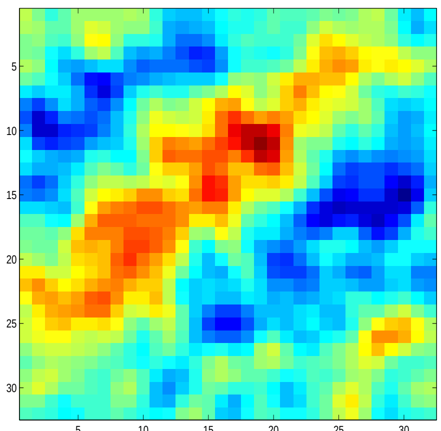
(A) True Image



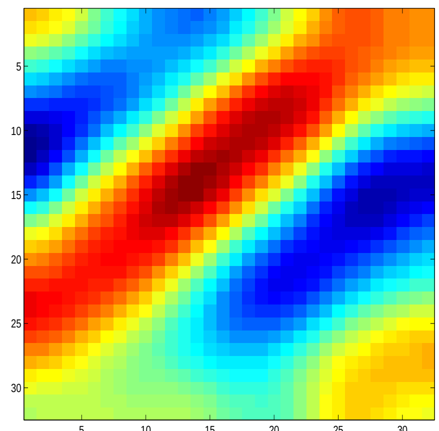
(B) Observation



(E) TTLS

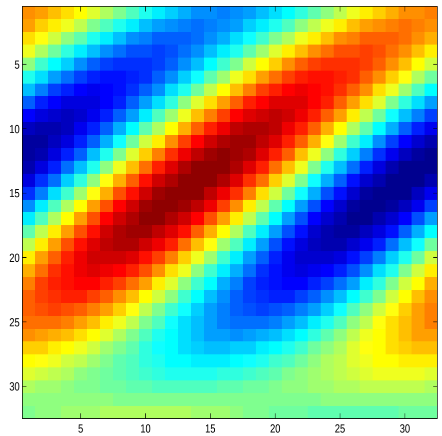


(F) Our Algorithm: RTLSC

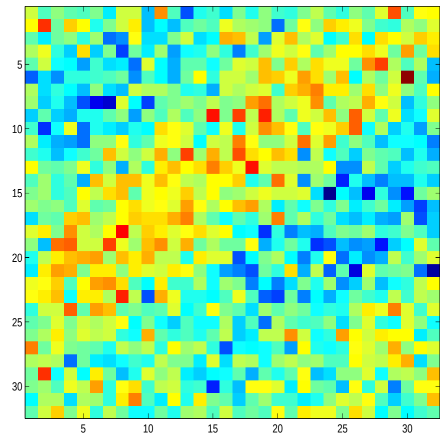


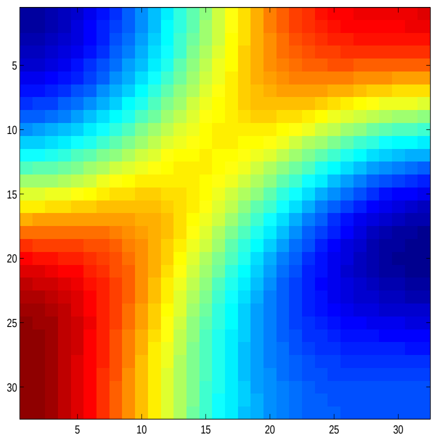
First three iterations of algorithm RTLSC

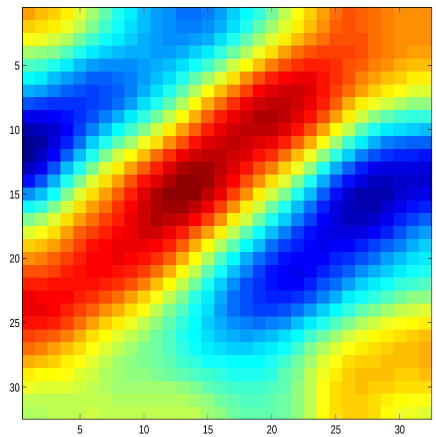
(A) True Image

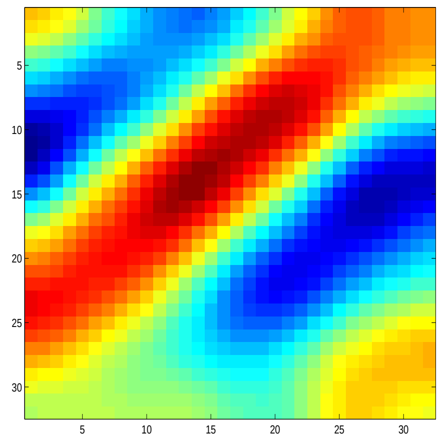


(B) Observation









An Empirical Observation and Some Questions...

Thousands of simulations suggest that both methods converge very quickly to a global minimum

- Can the iterative scheme of Sima et al. be proven to converge to the global minimum of RTLS? RQ?

An Empirical Observation and Some Questions...

Thousands of simulations suggest that both methods converge very quickly to a global minimum

- Can the iterative scheme of Sima et al. be proven to converge to the global minimum of RTLS? RQ?
- What is the theoretical rate of convergence of the iterative scheme?

An Empirical Observation and Some Questions...

Thousands of simulations suggest that both methods converge very quickly to a global minimum

- Can the iterative scheme of Sima et al. be proven to converge to the global minimum of RTLS? RQ?
- What is the theoretical rate of convergence of the iterative scheme?
- Does there exists a more general/unifying theory behind such algorithms and their good performance?

An Empirical Observation and Some Questions...

Thousands of simulations suggest that both methods converge very quickly to a global minimum

- Can the iterative scheme of Sima et al. be proven to converge to the global minimum of RTLS? RQ?
- What is the theoretical rate of convergence of the iterative scheme?
- Does there exists a more general/unifying theory behind such algorithms and their good performance?
- Hidden convexity...

Underlying Assumption

Assumption:

$$\exists \eta \geq 0 : \begin{pmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^T & c_2 \end{pmatrix} + \eta \begin{pmatrix} \mathbf{L}^T\mathbf{L} & \mathbf{0} \\ \mathbf{0} & -\rho \end{pmatrix} \succ \mathbf{0}. \quad (1)$$

Underlying Assumption

Assumption:

$$\exists \eta \geq 0 : \begin{pmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^T & c_2 \end{pmatrix} + \eta \begin{pmatrix} \mathbf{L}^T\mathbf{L} & \mathbf{0} \\ \mathbf{0} & -\rho \end{pmatrix} \succ \mathbf{0}. \quad (2)$$

- Implies that the problem is well-defined ($f_2(\mathbf{x}) > 0$ for every $\mathbf{x}$ such that $\|\mathbf{L}\mathbf{x}\|^2 \leq \rho$).
- Automatically satisfied for the RTLS problem ($\eta = 0$).
- Satisfied for the GTRS problem if $r = n$.

$$\mathbf{A} \succeq \mathbf{B} \Leftrightarrow \mathbf{A} - \mathbf{B} \textbf{ PSD}$$

$$\mathbf{A} \succ \mathbf{B} \Leftrightarrow \mathbf{A} - \mathbf{B} \textbf{ PD}$$

Attainability of the minimum

The minimum is not always attained. For example,

$$\min_{x_1,x_2}\left\{f(x_1,x_2)=\frac{5-4x_1+2x_1^2+x_2^2+x_1x_2}{1+x_1^2+x_2^2+x_1x_2}:x_1^2\leq 1\right\}.$$

Attainability of the minimum

The minimum is not always attained. For example,

$$\min_{x_1,x_2} \left\{ f(x_1,x_2) = \frac{5 - 4x_1 + 2x_1^2 + x_2^2 + x_1x_2}{1 + x_1^2 + x_2^2 + x_1x_2} : x_1^2 \le 1 \right\}.$$

The infimum is 1.

Attainability of the minimum

The minimum is not always attained. For example,

$$\min_{x_1,x_2}\left\{f(x_1,x_2)=\frac{5-4x_1+2x_1^2+x_2^2+x_1x_2}{1+x_1^2+x_2^2+x_1x_2} : x_1^2\leq 1\right\}.$$

The infimum is 1.

The infimum not attained since

$$f(x_1,x_2)=1+\frac{(x_1-2)^2}{1+x_1^2+x_2^2+x_1x_2}>1.$$

Attainability of the minimum

Attainability Condition: Either the feasible set if compact or

$$\lambda_{\min}(\mathbf{M}_1, \mathbf{M}_2) < \lambda_{\min}(\mathbf{F}^T\mathbf{A}_1\mathbf{F}, \mathbf{F}^T\mathbf{A}_2\mathbf{F}),$$

where

$$\mathbf{M}_1 = \begin{pmatrix} \mathbf{F}^T\mathbf{A}_1\mathbf{F} & \mathbf{F}^T\mathbf{b}_1 \\ \mathbf{b}_1^T\mathbf{F} & c_1 \end{pmatrix}, \mathbf{M}_2 = \begin{pmatrix} \mathbf{F}^T\mathbf{A}_2\mathbf{F} & \mathbf{F}^T\mathbf{b}_2 \\ \mathbf{b}_2^T\mathbf{F} & c_2 \end{pmatrix}$$

and **F** is an $n \times (n-r)$ matrix whose columns form an orthonormal basis for the null space of L.

- Weak inequality is always satisfied.
- (B-, T-, 06) The minimum is attained under the above assumption. Mathematical tools: recession function and sets.
- A generalization of the attainability condition for the unconstrained TLS problem : $\sigma_{\min}(\mathbf{A}, \mathbf{b}) < \sigma_{\min}(\mathbf{A})$.

Reformulation as a Nonconvex Quadratic Problem

Under the attainability condition, (RQ) can be homogenized:

$$\min_{\mathbf{z}\in\mathbb{R}^n, s\in\mathbb{R}} \left\{\varphi_1(\mathbf{z},s) : \varphi_2(\mathbf{z},s) = 1, \varphi_3(\mathbf{z},s) \leq 0\right\},$$

where

$$\begin{aligned}
\varphi_i(\mathbf{z},s) &= \mathbf{z}^T\mathbf{A}_i\mathbf{z} + 2\mathbf{b}_i^T\mathbf{z}s + c_i s^2, \quad i = 1,2,\\
\varphi_3(\mathbf{z},s) &= \|\mathbf{L}\mathbf{z}\|^2 - \rho s^2.
\end{aligned}$$

Reformulation as a Nonconvex Quadratic Problem

Under the attainability condition, (RQ) can be homogenized:

$$\min_{\mathbf{z}\in\mathbb{R}^n, s\in\mathbb{R}} \left\{\varphi_1(\mathbf{z},s) : \varphi_2(\mathbf{z},s)=1, \varphi_3(\mathbf{z},s)\le 0\right\},$$

where

$$\begin{aligned}\varphi_i(\mathbf{z},s) &= \mathbf{z}^T\mathbf{A}_i\mathbf{z}+2\mathbf{b}_i^T\mathbf{z}s+c_is^2, \quad i=1,2,\\ \varphi_3(\mathbf{z},s) &= \|\mathbf{L}\mathbf{z}\|^2-\rho s^2.\end{aligned}$$

S-Lemma of Polyak (98): under some mild conditions the following are equivalent for three symmetric matrices $\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3$:

(i) $\mathbf{y}^T\mathbf{R}_2\mathbf{y}=a_2, \mathbf{y}^T\mathbf{R}_3\mathbf{y}\le a_3 \Rightarrow \mathbf{y}^T\mathbf{R}_1\mathbf{y}\ge a_1.$

(ii) $\exists \alpha\in\mathbb{R}, \beta\in\mathbb{R}_+ : \mathbf{R}_1\succeq \alpha\mathbf{R}_2-\beta\mathbf{R}_3, \quad \alpha a_2\ge a_1+\beta a_3$

Semidefinite formulation of (RQ)

Under the attainability condition:

$$\begin{aligned} \max_{\beta\geq 0,\alpha,\lambda\in\mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & \begin{pmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^T & c_1 \end{pmatrix} \succeq \alpha \begin{pmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^T & c_2 \end{pmatrix} - \beta \begin{pmatrix} \mathbf{L}^T\mathbf{L} & \mathbf{0} \\ \mathbf{0} & -\rho \end{pmatrix}, \\ & \alpha \geq \lambda. \end{aligned}$$

Semidefinite formulation of (RQ)

Under the attainability condition:

$$
\begin{aligned}
\max_{\beta \geq 0, \alpha, \lambda \in \mathbb{R}} \quad & \lambda \\
\text{s.t.} \quad & \begin{pmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^T & c_1 \end{pmatrix} \succeq \alpha \begin{pmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^T & c_2 \end{pmatrix} - \beta \begin{pmatrix} \mathbf{L}^T\mathbf{L} & \mathbf{0} \\ \mathbf{0} & -\rho \end{pmatrix}, \\
& \alpha \geq \lambda.
\end{aligned}
$$

- Under the attainability condition, problem (RQ) is equivalent to a single convex semidefinite problem.
- The SDP problem can be solved efficiently via interior point methods.
- The solution of (RQ) can be extracted from the solution of the semidefinite formulation.

Hidden Convexity

The class of problems (RQ) belongs to a small but prestigious classes of nonconvex problems that can be reformulated as convex problems.

Hidden Convexity

The class of problems (RQ) belongs to a small but prestigious classes of nonconvex problems that can be reformulated as convex problems.
Nonconvex problems that can be transformed into convex problems:

- GTRS problems: $\min\{\mathbf{x}^T\mathbf{A}_1\mathbf{x} + 2\mathbf{b}_1^T\mathbf{x} + c_1 : \|\mathbf{L}\mathbf{x}\|^2 \leq \rho\}$
- Nonconvex homogenous quadratic programming with two quadratic constraints (Polyak, 98):

$$\min\{\mathbf{x}^T\mathbf{Q}_0\mathbf{x} : \mathbf{x}^T\mathbf{Q}_1\mathbf{x} \leq \rho_1, \mathbf{x}^T\mathbf{Q}_2\mathbf{x} \leq \rho_2\}$$

- Nonconvex quadratic optimization problems with two quadratic constraints over the complex domain (Beck & Eldar, 2006):

$$\min\{f_0(\mathbf{z}) : f_1(\mathbf{z}) \leq 0, f_2(\mathbf{z}) \leq 0, \mathbf{z} \in \mathbb{C}^n\},$$

where $f_i(\mathbf{z}) = \mathbf{z}^*\mathbf{A}_i\mathbf{z} + 2\Re(\mathbf{b}_i^*\mathbf{z}) + c_i$

Back to the fixed point algorithm...

Under the attainability condition,

- The iterative scheme of Sima et al. converges to a global optimum for the general problem (RQ).

Back to the fixed point algorithm...

Under the attainability condition,

- The iterative scheme of Sima et al. converges to a global optimum for the general problem (RQ).
- Superlinear rate of convergence

Back to the fixed point algorithm...

Under the attainability condition,

- The iterative scheme of Sima et al. converges to a global optimum for the general problem (RQ).
- Superlinear rate of convergence
- An ϵ-global optimal solution is obtained after at most $O(\sqrt{\log(1/\epsilon)})$ iterations.

Back to the fixed point algorithm...

Under the attainability condition,

- The iterative scheme of Sima et al. converges to a global optimum for the general problem (RQ).
- Superlinear rate of convergence
- An ϵ-global optimal solution is obtained after at most $O(\sqrt{\log(1/\epsilon)})$ iterations.

$\Rightarrow$ three globally convergent algorithms for solving (RQ).

Summary and Future Research

- (RQ) is a wide class of problems containing the class of GTRS and RTLS problems

Summary and Future Research

- (RQ) is a wide class of problems containing the class of GTRS and RTLS problems
- Three globally convergent algorithms for solving (RQ).

Summary and Future Research

- (RQ) is a wide class of problems containing the class of GTRS and RTLS problems
- Three globally convergent algorithms for solving (RQ).
- The exists an equivalent semidefinite formulation for the class of RQ problems $\Rightarrow$ hidden convexity.

Summary and Future Research

- (RQ) is a wide class of problems containing the class of GTRS and RTLS problems
- Three globally convergent algorithms for solving (RQ).
- The exists an equivalent semidefinite formulation for the class of RQ problems $\Rightarrow$ hidden convexity.
- Potential applications: conic trust region subproblems, min-max problems involving fractional terms.

Summary and Future Research

- (RQ) is a wide class of problems containing the class of GTRS and RTLS problems
- Three globally convergent algorithms for solving (RQ).
- The exists an equivalent semidefinite formulation for the class of RQ problems $\Rightarrow$ hidden convexity.
- Potential applications: conic trust region subproblems, min-max problems involving fractional terms.

Summary and Future Research

- (RQ) is a wide class of problems containing the class of GTRS and RTLS problems
- Three globally convergent algorithms for solving (RQ).
- The exists an equivalent semidefinite formulation for the class of RQ problems $\Rightarrow$ hidden convexity.
- Potential applications: conic trust region subproblems, min-max problems involving fractional terms.

Thank you for listening!