Multi-parameter regularization and regularized total least squares in Hilbert spaces

Ulrich Tautenhahn
University of Applied Sciences Zittau/Görlitz, Germany
http://www.hs-zigr.de/matnat/MATH/tautenhahn/index.html

Talk presented at the Workshop IP-TA 2010
Warsaw, February 9 – 12

Joint work with Sergei Pereverzev and Shuai Lu, RICAM Linz

Content

1. Introduction
2. What is
   - Total Least Squares (TLS)?
   - Regularized Total Least Squares (RTLS)?
   - Dual Regularized Total Least Squares (DRTLS)?
3. What it has to do with multi-parameter regularization?
4. Are there any error bounds for RTLS and DRTLS?
5. How to compute the RTLS and the DRTLS solution?
6. Conclusions

1 Introduction

Problem formulation, notational issues

Consider linear ill-conditioned systems

$$\boxed{A_0 x = y_0}$$

$x^\dagger$ – *unknown (generalized) solution*

$A_0 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ *(generally $m \geq n$)*

(A_0, y_0) – *error-free data and*

1. y_δ – *given noisy right hand side with* $\|y_0 - y_\delta\|_2 \leq \delta$
2. A_h – *given noisy system matrix with* $\|A_0 - A_h\|_F \leq h$

Ill-conditioning:
$$\left. \begin{array}{l} \|y_0 - y_\delta\|_2 \leq \delta \\ \|A_0 - A_h\|_F \leq h \end{array} \right\} \not\Rightarrow \|x^\dagger - x_{\delta,h}\|_2 \leq \varepsilon$$

2 TLS, RTLS and DRTLS

Least Squares for $A_h x \approx y_\delta$

Recall that in the LS problem we solve $\min_x \|A_h x - y_\delta\|_2$.

Alternatively, we look for x, y such that $y = A_h x$:

$$\min_{x,y} \|y - y_\delta\|_2 \quad \textit{subject to} \quad y = A_h x$$

Total Least Squares for $A_h x \approx y_\delta$

TLS takes care for perturbations in A_h:

$$\min_{x,y,A} \left\{ \|A - A_h\|_F^2 + \|y - y_\delta\|_2^2 \right\} \quad \textit{subject to} \quad y = Ax$$

Hence, we look for x, y, A to make the system compatible.

How to compute the TLS solution x_{TLS}?

1. *Compute the SVD of* $\quad (A_h|y_\delta)_{m,n+1} = \sum_{i=1}^{n+1} v_i\,\sigma_i\,u_i^T$
2. *Partition* $U = (u_1\,u_2\,...\,|u_{n+1})$ *as*

$$U = \begin{pmatrix} U_{11} & | U_{12} \\ U_{21} & | U_{22} \end{pmatrix} \begin{matrix} n \\ 1 \end{matrix}$$

3. *Then, if* $U_{22} \neq 0$,

$$x_{TLS} = -\frac{1}{U_{22}} U_{12}.$$

Alternatively, if $\sigma_{n+1} \notin \sigma(A_h^T A_h)$,

$$x_{TLS} = (A_h^T A_h - \sigma_{n+1} I)^{-1} A_h^T y_\delta.$$

Tikhonov regularization

Remember different formulations:

(i) $$\min_x \left\{ \|A_h x - y_\delta\|_2^2 + \alpha \|Bx\|_2^2 \right\}$$

(ii) $$\left(A_h^T A_h + \alpha B^T B \right) x = A_h^T y_\delta$$

. . . .

(iii) $$\min_x \|A_h x - y_\delta\|_2 \quad \textit{subject to} \quad \|Bx\|_2 \le R$$

(iv) $$\min_x \|Bx\|_2 \quad \textit{subject to} \quad \|A_h x - y_\delta\|_2 \le \delta$$

Question: How to introduce TLS in the Tikhonov setting?

Regularized TLS (RTLS)

Introduce TLS in the Tikhonov setting (iii) as follows:

$$\min_{x,y,A} \left\{ \|A - A_h\|_F^2 + \|y - y_\delta\|_2^2 \right\} \quad \textit{subject to} \quad \begin{aligned} & y = Ax, \\ & \|Bx\|_2 \le R \end{aligned}$$

Dual RTLS

Introduce TLS in the Tikhonov setting (iv) as follows:

$$\min_{x,y,A} \|Bx\|_2^2 \quad \textit{subject to} \quad \begin{aligned} & y = Ax, \\ & \|y - y_\delta\|_2 \le \delta, \\ & \|A - A_h\|_F \le h \end{aligned}$$

Some references on TLS and RTLS

Huffel, S. V. and Vanderwalle, J. (1991)
The TLS Problem: Computational Aspects and Analysis
Philadelphia: SIAM

Golub, G. H., Hansen, P. C. and O'Leary, D. P. (1999)
Tikhonov regularization and total least squares
SIAM J. Matrix Anal. Appl. 21, 185 - 194

Further references on RTLS:

Sima, D., Huffel, S. V. and Golub, G. H. (2004), Renaut, R. A. and Guo, H. (2005), Beck, A. and Ben-Tal, A. (2006), Beck, A. and Ben-Tal, A. and Teboulle, M. (2006), Sima, D. (2006), Lu, S. and Pereverzev, S. V. and Tautenhahn, U. (2007), Lampe, J. and Voss, H. (2007,2009)

Some references on DRTLS

- Lu, S. and Pereverzev, S. V. and Tautenhahn, U. (2009)
  Regularized total least squares: computational aspects and error bounds
  SIAM J Matrix Anal Appl **31**, *918 – 941*
- Lu, S. and Pereverzev, S. V. and Tautenhahn, U. (2008)
  Dual regularized total least squares and multi-parameter regularization
  Comput. Meth. Appl. Math. 8, 253–262
- Lu, S. and Pereverzev, S. V. and Tautenhahn, U. (2008)
  A model function method in total least squares
  RICAM-Preprint 2008-18

3 RTLS, DRTLS and Multi-parameter regularization

Theorem 1 (RTLS and multi-parameter regularization)

Consider the RTLS problem

$$\min_{x,y,A} \left\{ \|A - A_h\|_F^2 + \|y - y_\delta\|_2^2 \right\} \quad \text{subject to} \quad y = Ax, \quad \|Bx\|_2 \le R.$$

If the constraint is active, then the solution $x_{RTLS} = x_{\alpha,\beta}^R$ *can be obtained by* *multi-parameter regularization*

$$\|A_h x - y_\delta\|_2^2 + \alpha \|Bx\|_2^2 + \beta \|x\|_2^2 \to \min$$

where (α, β) *obey*

$$\|Bx_{\alpha,\beta}^R\|_2 = R \quad \text{and} \quad \beta = -\frac{\|y_\delta - A_h x_{\alpha,\beta}^R\|_2^2}{1 + \|x_{\alpha,\beta}^R\|_2^2}.$$

Properties in the general case $B \neq I$

Let $x = x_{RTLS}$ *be given by*

$$(A_h^T A_h + \alpha B^T B + \beta I)x = A_h^T y_\delta$$

Depending on the choice of R, *the solutions are related by*

R	*solutions*	α	β
$R < \|Bx_{TLS}\|_2$	$x_{RTLS} \neq x_{TLS}$	$\alpha > 0$	$\beta < 0,\ \partial\beta/\partial R > 0$
$R \geq \|Bx_{TLS}\|_2$	$x_{RTLS} = x_{TLS}$	$\alpha = 0$	$\beta = -\sigma_{n+1}^2$

- No equivalent interpretation with Tikhonov setting (iii)
- Multi-parameter regularization with $\beta < 0$
- De-regularization due to the negativity of β

Theorem 2 (DRTLS and multi-parameter regularization)

Consider the DRTLS problem

$$\min_{x,y,A} \|Bx\|_2^2 \quad \text{subject to} \quad \begin{aligned} & y = Ax, \\ & \|y - y_\delta\|_2 \le \delta, \\ & \|A - A_h\|_F \le h. \end{aligned}$$

If the constraints are active, then the solution $x_{DRTLS} = x_{\alpha,\beta}^{\delta,h}$ can be obtained by multi-parameter regularization

$$\|A_h x - y_\delta\|_2^2 + \alpha\|Bx\|_2^2 + \beta\|x\|_2^2 \to \min$$

where the parameters α and β obey

$$\|A_h x_{\alpha,\beta}^{\delta,h} - y_\delta\|_2 = \delta + h\|x_{\alpha,\beta}^{\delta,h}\|_2 \quad \text{and} \quad \beta = -h^2 - \frac{h\delta}{\|x_{\alpha,\beta}^{\delta,h}\|_2}.$$

Ideas of proof

1. *Use classical Lagrange multiplier formulation:*

$$\mathcal{L}(x, A, \mu, \nu) = \|Bx\|_2^2 + \mu(\|Ax - y_\delta\|_2^2 - \delta^2) + \nu(\|A - A_h\|_F^2 - h^2)$$

2. *Optimality conditions:*

$$\begin{aligned} \mathcal{L}_x &= 2B^T Bx + 2\mu A^T (Ax - y_\delta) = 0 \\ \mathcal{L}_A &= 2\mu(Ax - y_\delta)x^T + 2\nu(A - A_h) = 0 \\ \mathcal{L}_\mu &= \|Ax - y_\delta\|_2^2 - \delta^2 = 0 \\ \mathcal{L}_\nu &= \|A - A_h\|_F^2 - h^2 = 0 \end{aligned}$$

3. *Manipulation of these equations gives*

$$A = A_h - \frac{h}{\|(A_h x - y_\delta)x^T\|_F}(A_h x - y_\delta)x^T$$

4. *Further manipulation gives our characterization result*

6 Error bounds

- $A_0 \in \mathcal{L}(X,Y)$ with non-closed range $\mathcal{R}(A_0)$
- X, Y – Hilbert spaces
- B – strictly pos. self-adjoint (unbounded) operator in X

Assumption A1 (Link condition between A_0 and B^{-1})

$$m\|B^{-a}x\| \le \|A_0 x\| \quad \textit{for some } a > 0, m > 0$$

Assumption A2 (Solution smoothness)

$$x^\dagger \in M_{B,E} = \left\{x \in X \,:\, \|B^p x\| \le E\right\} \quad \textit{for some } p > 0$$

Theorem 1 (Order optimality for the RTLS solution)

$$\left.\begin{array}{l} A1, A2 \\ p \in [1, 2+a] \\ R = \|Bx^\dagger\| \end{array}\right\} \Rightarrow \|x_{RTLS} - x^\dagger\| = O\left((\delta + h)^{\frac{p}{p+a}}\right)$$

Extensions:

- More general link conditions
- More general conditions for solution smoothness
- Case $0 < p < 1$?

Problem: Exact magnitude of $\|Bx^\dagger\|$ necessary!

Theorem 2 (Order optimality for the DRTLS solution)

$$\left.\begin{array}{l} A1, A2 \\ p \in [1, 2+a] \end{array}\right\} \Rightarrow \|x_{DRTLS} - x^\dagger\| = O\left((\delta + h)^{\frac{p}{p+a}}\right)$$

Extensions:

- More general link conditions
- More general conditions for solution smoothness
- Case $0 < p < 1$?

Advantage over RTLS

- No need to know $\|Bx^\dagger\|$
- Only noise-levels δ and h necessary

Error bounds under general smoothness conditions

For results with $h = 0$, see

- *Nair/Pereverzev/Tautenhahn (2005)*
  Reg. in Hilbert scales under general smoothing conditions
  Inverse Problems 21, *1851 – 1869*
- *Böttcher/Hofmann/Tautenhahn/Yamamoto (2006)*
  Convergence rates for Tikhonov regularization from different kinds of smoothness conditions
  Applicable Analysis 85, *555 – 578*
- *Mathé/Tautenhahn (2006)*
  Interpolation in variable Hilbert scales with application to inverse problems
  Inverse Problems 22, *2271 – 2297*

4 Computational aspects for RTLS and DRTLS

RTLS: Eigenvalue-eigenvector-problem for $x = x_{RTLS}$

If the constraint $\|Bx\|_2 \le R$ *is active, then*

$$\begin{pmatrix} A_h^T A_h + \alpha B^T B & A_h^T y_\delta \\ y_\delta^T A_h & -\alpha R^2 + y_\delta^T y_\delta \end{pmatrix} \begin{pmatrix} x \\ -1 \end{pmatrix} = -\beta \begin{pmatrix} x \\ -1 \end{pmatrix}$$

with parameters α *and* β *given as above.*

RTLS: Constrained minimization problem for $x = x_{RTLS}$

The RTLS solution $x = x_{RTLS}$ *is the solution of*

$$\min_x \frac{\|A_h x - y_\delta\|_2^2}{1 + \|x\|_2^2} \quad \textit{subject to} \quad \|Bx\|_2 \le R$$

DRTLS: Special case h=0

1. *If* $\|y - y_\delta\| \le \delta$ *is active, then* $x = x_{DRTLS}$ *satisfies*

$$(A_0^T A_0 + \alpha B^T B)x = A_0^T y_\delta$$

and the parameter $\alpha > 0$ *is the solution of*

$$f(\alpha) := \|A_0 x - y_\delta\|_2^2 - \delta^2 = 0.$$

2. $f : \mathbb{R}^+ \to \mathbb{R}$ *is continuous and monotonically increasing.*
3. *Newton's method for* $h(r) := f(1/r) = 0$ *converges globally and monotonically where*

$$\text{(i)} \quad h'(r) < 0 \qquad\qquad \text{(ii)} \quad h''(r) > 0.$$

4. *Rewriting Newton's method with* r *in terms of* $\alpha = 1/r$ *leads to following algorithm:*

Algorithm for the Dual RTLS problem, case h=0

Input: $\varepsilon > 0$, y_δ, A_0, B *and* δ *satisfying* $\delta < \|y_\delta\|_2$.

1: *Choose some starting value* $\alpha \geq \alpha^*$.

2: *Solve* $(A_0^T A_0 + \alpha B^T B)x = A_0^T y_\delta$.

3: *Solve* $(A_0^T A_0 + \alpha B^T B)v = B^T Bx$.

4: *Update* $\alpha_{\text{new}} := \dfrac{2\alpha^3(v, B^T Bx)}{2\alpha^2(v, B^T Bx) + \|A_0 x - y_\delta\|_2^2 - \delta^2}$.

5: **if** $|\alpha_{\text{new}} - \alpha| \geq \varepsilon|\alpha|$ **then** $\alpha := \alpha_{\text{new}}$ *and* **goto** *2*

6: **else** *solve* $(A_0^T A_0 + \alpha_{\text{new}} B^T B)x = A_0^T y_\delta$.

- Extensions to the general case $h \neq 0$ possible
- Alternative: Use model function approach

7 Conclusions

1. Review on TLS:

   ⇒ takes care for perturbations in A_h

2. Introducing TLS in Tikhonov's setting:

   ⇒ leads to RTLS and DRTLS

3. Characterization results for RTLS and DRTLS

   ⇒ leads to multi-parameter regularization

4. Order optimal error bounds:

   ⇒ $\|Bx^\dagger\|$ required for RTLS, not for DRTLS

5. Computational aspects for DRTLS:

   ⇒ Newton's method applied to transformed equations