Available online at www.sciencedirect.com
ScienceDirect
Applied Mathematics Letters 20 (2007) 82–87
Applied Mathematics Letters
www.elsevier.com/locate/aml

Tikhonov regularization for weighted total least squares problems[☆]

Yimin Wei[a,b,*], Naimin Zhang[c], Michael K. Ng[d], Wei Xu[e]

[a] School of Mathematical Sciences, Fudan University, Shanghai 200433, PR China
[b] Key Laboratory of Mathematics for Nonlinear Sciences (Fudan University), Ministry of Education, PR China
[c] Mathematics and Information Science College, Wenzhou University, Wenzhou 325035, PR China
[d] Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong
[e] Department of Computing and Software, McMaster University, Hamilton, Ont., Canada L8S 4L7

Received 15 August 2003; received in revised form 15 January 2006; accepted 1 March 2006

Abstract

In this work, we study and analyze the regularized weighted total least squares (RWTLS) formulation. Our regularization of the weighted total least squares problem is based on the Tikhonov regularization. Numerical examples are presented to demonstrate the effectiveness of the RWTLS method.

Keywords: Tikhonov regularization; Total least squares; Weighted regularized total least squares; Matrix differentiation; Lagrange multiplier

1. Introduction

In this work, we study the regularized weighted total least squares (RWTLS) formulation. Our regularization of the weighted total least squares problem is based on the Tikhonov regularization [1].

For the total least squares (TLS) problem [2], the truncation approach has already been studied by Fierro et al. [3]. In [4], Golub et al. has considered the Tikhonov regularization approach for TLS problems. They derived a new regularization method in which stabilization enters the formulation in a natural way, and that is able to produce regularized solutions with superior properties for certain problems in which the perturbations are large. In the present work, we focus on RWTLS problems. We show that the RWTLS solution is closely related to the Tikhonov solution to the weighted least squares solution.

Our work is organized as follows. In Section 2, we introduce the RWTLS formulation and study its regularizing properties. Computational methods are described in Section 3. In Section 4, numerical examples are presented to demonstrate the RWTLS method.

☆ Research supported in part by the National Natural Science Foundation of China under grant 10471027 and Shanghai Education Committee, and Hong Kong RGC Grants Nos 7046/03P, 7035/04P and 7035/05P, and HKBU FRGs.
* Corresponding author at: School of Mathematical Sciences, Fudan University, Shanghai 200433, PR China.
E-mail addresses: ymwei@fudan.edu.cn, ymwei_cn@yahoo.com (Y. Wei), naiminzhang@yahoo.com.cn (N. Zhang), mng@math.hkbu.edu.hk (M.K. Ng).

0893-9659/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aml.2006.03.004

2. The regularized weighted total least squares

A general version of Tikhonov's formulation for the linear weighted total least squares (WTLS) problem takes the form [5]

$$\min_x \|U[(A,b)-(\tilde{A},\tilde{b})]V\|_F \quad \text{subject to} \quad \tilde{b}=\tilde{A}x, \quad \|Dx\|_S \le \delta, \tag{1}$$

where D is the regularization matrix, $V=\operatorname{diag}(W,\gamma)$ with γ being a non-zero constant, U and W are nonsingular matrices, S is a symmetric positive definite matrix with $\|y\|_S^2 = y^{\mathrm{T}}Sy$, and δ is a positive constant. By using the Lagrange multiplier formulation, this problem can be rewritten as follows:

$$\mathcal{L}(\tilde{A},x,\mu)=\|U[(A,b)-(\tilde{A},\tilde{b})]V\|_F^2+\mu(\|Dx\|_S^2-\delta^2), \quad \text{subject to} \quad \tilde{b}=\tilde{A}x, \tag{2}$$

where μ is the Lagrange multiplier, and μ is equal to zero if the inequality constraint becomes equality. The solution $\bar{x}_\delta$ to this problem is different from the solution x_{WTLS} to

$$\min_x \|U[(A,b)-(\tilde{A},\tilde{b})]V\|_F \quad \text{subject to} \quad \tilde{b}=\tilde{A}x, \tag{3}$$

for δ less than $\|Dx_{\mathrm{WTLS}}\|_2$.

Before we show the properties of the solution to (2), we have the following results about the matrix differentiation for the matrices A, $\tilde{A}$, W and U.

Lemma 1.

$$\begin{array}{llllll}
\text{(i)} & \dfrac{\partial\, \mathrm{tr}(W^{\mathrm{T}}A^{\mathrm{T}}U^{\mathrm{T}}U\tilde{A}W)}{\partial\tilde{A}} = U^{\mathrm{T}}UAWW^{\mathrm{T}} & \text{(ii)} & \dfrac{\partial\, \mathrm{tr}(W^{\mathrm{T}}\tilde{A}^{\mathrm{T}}U^{\mathrm{T}}UAW)}{\partial\tilde{A}} = U^{\mathrm{T}}UAWW^{\mathrm{T}} \\
\text{(iii)} & \dfrac{\partial\, \mathrm{tr}(W^{\mathrm{T}}\tilde{A}^{\mathrm{T}}U^{\mathrm{T}}U\tilde{A}W)}{\partial\tilde{A}} = 2U^{\mathrm{T}}U\tilde{A}WW^{\mathrm{T}} & \text{(iv)} & \dfrac{\partial\, (b^{\mathrm{T}}U^{\mathrm{T}}U\tilde{A}x)}{\partial\tilde{A}} = U^{\mathrm{T}}Ubx^{\mathrm{T}} \\
\text{(v)} & \dfrac{\partial\, (x^{\mathrm{T}}\tilde{A}^{\mathrm{T}}U^{\mathrm{T}}Ub)}{\partial\tilde{A}} = U^{\mathrm{T}}Ubx^{\mathrm{T}} & \text{(vi)} & \dfrac{\partial\, (x^{\mathrm{T}}\tilde{A}^{\mathrm{T}}U^{\mathrm{T}}U\tilde{A}x)}{\partial\tilde{A}} = 2U^{\mathrm{T}}U\tilde{A}xx^{\mathrm{T}}.
\end{array}$$

Proof. Since (i) is equivalent to (ii), (iv) is equivalent to (v), and (vi) is a special case of (iii), we only give the proofs of (i) and (iii).

Let Z be a $p\times q$ matrix of differentiable functions of the $m\times n$ matrix X. If

$$\frac{\partial Z}{\partial x_{ij}} = GE_{ij}^{(mn)}H + C(E_{ij}^{(mn)})^{\mathrm{T}}F, \quad i=1,\ldots,m,\ j=1,\ldots,n$$

then

$$\frac{\partial z_{ij}}{\partial X} = G^{\mathrm{T}}E_{ij}^{(pq)}H^{\mathrm{T}} + F(E_{ij}^{(pq)})^{\mathrm{T}}C, \quad i=1,\ldots,p,\ j=1,\ldots,q,$$

and the converse is also true (see p. 57, Theorem 7.1 in [6]), where $G=(g_{ij})$ is a $p\times m$ matrix, $H=(h_{ij})$ is an $n\times q$ matrix, $C=(c_{ij})$ is a $p\times n$ matrix, $F=(f_{ij})$ is an $m\times q$ matrix $E_{ij}^{(kl)}$ is a k-by-l zero matrix except the (i,j)-entry being equal to one.

For (i), we consider $Y=W^{\mathrm{T}}A^{\mathrm{T}}U^{\mathrm{T}}U$ and we have $\frac{\partial\, \mathrm{tr}(Y\tilde{A}W)}{\partial\tilde{A}} = \frac{\partial}{\partial\tilde{A}}\left(\sum_i (Y\tilde{A}W)_{ii}\right) = \sum_i \frac{\partial (Y\tilde{A}W)_{ii}}{\partial\tilde{A}}$. Since $\frac{\partial(Y\tilde{A}W)}{\partial\tilde{A}_{ij}} = YE_{ij}W$ and $\frac{\partial(Y\tilde{A}W)_{ij}}{\partial\tilde{A}} = Y^{\mathrm{T}}E_{ij}W^{\mathrm{T}}$, we obtain $\frac{\partial\, \mathrm{tr}(Y\tilde{A}W)}{\partial\tilde{A}} = \sum_i Y^{\mathrm{T}}E_{ii}W^{\mathrm{T}} = Y^{\mathrm{T}}W^{\mathrm{T}}$. The result follows.

For (iii), we find that $\frac{\partial[(U\tilde{A}W)^{\mathrm{T}}(U\tilde{A}W)]}{\partial\tilde{A}_{ij}} = W^{\mathrm{T}}E_{ij}^{\mathrm{T}}U^{\mathrm{T}}U\tilde{A}W + (U\tilde{A}W)^{\mathrm{T}}UE_{ij}W$, and therefore we have $\frac{\partial[(U\tilde{A}W)^{\mathrm{T}}(U\tilde{A}W)]_{ii}}{\partial\tilde{A}} = U^{\mathrm{T}}U\tilde{A}WE_{ii}^{\mathrm{T}}W^{\mathrm{T}} + U^{\mathrm{T}}U\tilde{A}WE_{ii}W^{\mathrm{T}}$. It follows that