Linear Algebra and its Applications 461 (2014) 18–41

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Two projection methods for Regularized Total Least Squares approximation

Geunseop Lee *, Jesse L. Barlow

Department of Computer Science and Engineering, The Pennsylvania State University, University Park, PA 16802, United States

A R T I C L E I N F O

Article history:
Received 18 March 2013
Accepted 9 July 2014
Available online 23 August 2014
Submitted by Z. Strakoš

MSC:
65F10
65F22

Keywords:
Projection method
Regularized Total Least Squares
Truncated GKL bidiagonalization
Generalized Krylov subspace

A B S T R A C T

Regularized Total Least Squares is a useful approach for solving ill-posed overdetermined systems of equations when both the model matrix and the observed data are contaminated by noise. A Newton-based Regularized Total Least Squares method was proposed by Lee et al. (2013) [16], but may not be efficient for large scale problems. Here we consider two projection-based algorithms applied to this method for the solution of the large scale problem. The first fixes the underlying subspace dimension, while the second expands the subspace dynamically during the iterations by employing a generalized Krylov subspace expansion. Experimental results demonstrate that the two projection-based algorithms can be successfully applied for the solution of the large scale Regularized Total Least Squares problems.

1. Introduction

We are considering the overdetermined linear system

$$X\mathbf{y} \approx \mathbf{b}, \tag{1}$$

* Corresponding author.
E-mail addresses: gxl174@cse.psu.edu (G. Lee), barlow@cse.psu.edu (J.L. Barlow).

http://dx.doi.org/10.1016/j.laa.2014.07.045

with the matrix $X \in \mathbf{R}^{m\times n}$, where $m \geq n$, the noise-contaminated observed data $\mathbf{b} \in \mathbf{R}^m$, and the solution $\mathbf{y} \in \mathbf{R}^n$. In practice, many situations occur in which errors exist in both the observed data $\mathbf{b}$ and the data matrix X. In this case, it is appropriate to adopt the Total Least Squares (TLS) model (cf. [6, §6.3], [27, Chapters 1–3]). The TLS solution minimizes the sum of squared norms given by

$$\min\left(\|E\|_F^2 + \|\mathbf{r}\|_2^2\right), \tag{2}$$

such that for some $\mathbf{y} \in \mathbf{R}^n$,

$$(X + E)\mathbf{y} = \mathbf{b} + \mathbf{r}.$$

In our context, the matrix X is ill-conditioned, thus the TLS solution

$$\mathbf{y}_{TLS} = \arg\min_{\mathbf{y}}\left(\left\|\mathbf{b} - (X + E)\mathbf{y}\right\|_2^2 + \|E\|_F^2\right) \tag{3}$$

of the linear system (1) is dominated by errors in X and $\mathbf{b}$, possibly making it meaningless [27, Chapter 7]. A recent characterization of the condition of this problem is given by Gratton et al. [26]. To stabilize the solution $\mathbf{y}_{TLS}$, we employ the Tikhonov regularization which chooses a linear operator L and a parameter δ so that the solution $\mathbf{y}$ satisfies the condition $\|L\mathbf{y}\|_2 \leq \delta$. Common choices for L include the identity, a first-order derivative, and a second-order derivative operator. The regularization parameter δ comes from knowledge of the underlying physical model. Therefore, using Tikhonov regularization, we reformulate the problem of (2) as

$$\mathbf{y}_{TLS}(\delta) = \arg\min_{\mathbf{y}}\left(\left\|\mathbf{b} - (X + E)\mathbf{y}\right\|_2^2 + \|E\|_F^2\right) \quad \text{subject to } \|L\mathbf{y}\|_2 \leq \delta. \tag{4}$$

Thus the Lagrangian from Eq. (4), as used by Golub, Hansen, and O'Leary [4], is

$$\mathcal{L}(E, \mathbf{y}, \beta) = \left\|\mathbf{b} - (X + E)\mathbf{y}\right\|_2^2 + \|E\|_F^2 + \beta\left(\|L\mathbf{y}\|_2^2 - \delta^2\right). \tag{5}$$

The Lagrange multiplier $\beta > 0$, not known *a priori*, is chosen to control the size of the solution $\mathbf{y}$. A more general model of constrained total least squares is given by Lu et al. [17].

In [4], Golub et al. show that, if the constraint $\|L\mathbf{y}\|_2 \leq \delta$ is active, the solution $\mathbf{y}_{TLS}(\delta)$ of (4), with a properly chosen parameter β, is equivalent to the solution $\mathbf{y}(\lambda, \mu)$ of the linear system

$$M(\lambda, \mu)\mathbf{y}(\lambda, \mu) = X^T\mathbf{b}, \tag{6}$$

where

$$M(\lambda, \mu) = X^TX + \lambda L^TL - \mu I, \tag{7}$$