Journal of Computational Mathematics, Vol.21, No.4, 2003, 463–472.

A SUCCESSIVE LEAST SQUARES METHOD FOR STRUCTURED TOTAL LEAST SQUARES [*1)]

Plamen Y. Yalamov
(*Center of Applied Mathematics and Informatics, University of Rousse, 7017 Rousse, Bulgaria*)

Jin-yun Yuan
(*Departamento de Matemática, Universidade Federal do Paraná, Centro Politécnico, Caixa Postal 19.081, 81531-990 Curitiba – PR, Brazil*)

Abstract

A new method for Total Least Squares (TLS) problems is presented. It differs from previous approaches and is based on the solution of successive Least Squares problems. The method is quite suitable for Structured TLS (STLS) problems. We study mostly the case of Toeplitz matrices in this paper. The numerical tests illustrate that the method converges to the solution fast for Toeplitz STLS problems. Since the method is designed for general TLS problems, other structured problems can be treated similarly.

Key words: Structure total least squares, Linear Least squares, Successive linear squares method, Toeplitz systems, Structure least squares.

1. Introduction

Total Least Squares (TLS) problems appear in many engineering applications such as signal and image processing, systems identification, and systems response prediction. A good survey of areas of application and computational methods is given in [17].

The TLS problem can be stated as follows:

$$\|E \mid r\|_F = \min, \text{ where } (A+E)x = b + r. \tag{1}$$

Here $A, E \in \mathcal{R}^{m\times n}$ (usually $m \geq n$), and $x \in \mathcal{R}^n, b, r \in \mathcal{R}^m$. The subscript F denotes the Frobenius norm. E and r are called errors in the model.

All the algorithms in [17] are based on the Singular Value Decomposition (SVD) analysis (see [9, 10]). Other approaches are taken in [3] (general matrices) and [12] (Toeplitz matrices) where methods for nonlinear equations are used to solve the problem. All these methods are suitable for general matrices, and do not take into account any structure in the matrix E. Very often in practice the matrix A has some structure, e. g., Toeplitz, or Hankel [13]. Sometimes the matrix E requires to have the same structure as A. We will call this problem a Structured TLS (STLS) problem. For this problem the SVD based methods of [17], and the methods of [3, 12] do not produce a matrix E with the desired structure.

A different approach is applied in [15] where minimization techniques are used to solve STLS problems. Toeplitz and sparsity structures are considered as an application. This method produces a matrix E with a prescribed structure. The method is suitable only for TLS problems of small size.

[*] Received March 13, 2001; final revised January 15, 2002.

[1)] This research was started when the first author visited the Federal University of Paraná, Curitiba, Brazil and finished when the second author visited Stanford University, USA. The work of the first author was also supported by Grant MM-707/97 from the National Scientific Research Fund of the Bulgarian Ministry of Education and Science. The work of the second author was partially supported by CNPq, CAPES, FINEP, Fundação Araucaria and FUNPAR, Brazil.

A number of methods for solving large Structured LS (SLS) problems ([2, 4, 5, 6, 7]) motivates us to establish some method for solving STLS problems by using SLS methods. The purpose of this paper is to propose such a method for the STLS problem in which the basic kernel is the solution of a LS problem. In this way the proposed method can be used for solving large STLS problems. We give a general framework of the method. Then, by a suitable choice of a parameter, the method is applicable to structured, or unstructured problems (We use the same idea as in [15]). We prove global convergence for any structure. In the case of Toeplitz A and E we show also that each iteration step is faster than one step of the method in [15]. While the minimization of the errors in [15] is with respect to the 1, 2, and infinity norms, here we discuss only the 2-norm. Clearly, this norm is the best choice when LS solutions are involved. In this paper, the existence of solution of the structured total least squares problem (1) always assumed. The outline of the paper is as follows. In Section 2 we present the new method and study its convergence. In Section 3 the implementation for Toeplitz STLS problems is considered. Finally, numerical experiments are give in Section 4.

2. The LS Method

Since the equation

$$(A+E)x = b + r \tag{2}$$

is nonlinear with respect to the unknowns E, x, and r, we assume that the unknowns in a nonlinear system can be split into two groups, for example, one group for x and another group for E and r. With this splitting, if the unknowns in one of the groups are constants the problem becomes linear with respect to the unknowns from the other group, and vice versa.

In such nonlinear problems we can start with some initial value for one of the groups of unknowns, and then alternatively compute approximations of the two groups of unknowns by solving linear problems according to some iteration scheme.

For the TLS problem we suggest the LS solution $x^{(0)}, Ax^{(0)} = b + r^{(0)}$, as an initial value for x. The same initial value is chosen also in [3, 15]. This choice is natural because the LS problem is just a special case of the TLS problem, and in many cases $x^{(0)}$ will be close enough to the solution of the TLS problem.

Let us also note that if x is constant, and E and r are variables then problem (1) can be rewritten as

$$\left\| \begin{array}{c} r \\ \alpha \end{array} \right\|_2 = \min, \quad Ax + X\alpha = b + r. \tag{3}$$

Here the matrix X and vector α are chosen in such a way that

$$X\alpha = Ex.$$

This choice depends on the structure of the matrix E. We present a few examples:

- E is unstructured. Then we have

$$X = \begin{pmatrix} x_1 & \cdots & x_n & & & & & & \\ & & & x_1 & \cdots & x_n & & & \\ & & & & & & \ddots & & \\ & & & & & & x_1 & \cdots & x_n \end{pmatrix} \in \mathcal{R}^{m \times mn},$$

$$\alpha =$$

$$\mathrm{Vec}(E) = (e_{11}, e_{12}, \ldots, e_{1n}, e_{21}, e_{22}, \ldots, e_{2n}, \ldots, e_{m1}, e_{m2}, \ldots, e_{mn})^T.$$

- E is general sparse. Then X and α are also sparse, and their sparsity pattern depends on the sparsity pattern of E.