Advances in Applied Mathematical Analysis.
ISSN 0973-5313 Volume 2 Number 1 (2007), pp. 51–55

http://www.ripublication.com/aama.htm

Data-Ordinary-Total Least Norm (Squares) Problems

Faranges Kyanfar
Department of Mathematics, Shahid Bahonar University of Kerman, 7619614111, Kerman, Iran
E-mail: kyanfar@mail.uk.ac.ir

Abstract

This paper presents an analysis of the errors of the Data-Ordinary-Total least norm (squares) problem $AX \approx B$, where $A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{m\times p}$ by using arbitrary norm (Frobenius norm). Also we state some conditions which the minimum errors cannot be attained.

AMS subject classification: 65F20, 65F35.
Keywords: Total least square, data least square, least norm.

1. Introduction

Let $A \in \mathbb{C}^{m\times n}$ be a data matrix and $B \in \mathbb{C}^{m\times p}$ be an observation matrix. An over determined set of linear equations $AX \approx B$ arises in a broad class of scientific problems, where usually, $R(B) \nsubseteq R(A)$ and hence this set does not have an exact solution. Throughout this paper $\|.\|_F$ denotes the Frobenius norm on $\mathbb{C}^{m\times(n+p)}$ and $\|.\|$ denotes a fixed arbitrary norm on $\mathbb{C}^{m\times(n+p)}$.

In ordinary least squares approach the measurements A are assumed to be free of error and all errors are confined to data matrix B. Then we are looking to find

$$\inf_{\Delta B\in\mathbb{C}^{m\times p}} \|\Delta B\|_F, \quad AX = (B - \Delta B) \quad \text{has a solution.} \tag{1.1}$$

In underlying assumption the matrix A is exactly known, see [2, 4, 5].

In data least squares approach observation matrix B are assumed to be free of error and all errors are confined to data matrix A, We are looking to find

$$\inf_{\Delta A\in\mathbb{C}^{m\times n}} \|\Delta A\|_F, \quad (A - \Delta A)X = B \quad \text{has a solution.} \tag{1.2}$$

In underlying assumption the matrix B is exactly known see [3].

In the above assumptions the data matrix A or observation matrix B are free of error, but these assumptions are frequently unrealistic: sampling errors, human errors, modeling errors, and instrument errors may imply inaccuracies in data matrix A and observation matrix B. In total least squares approach random errors occur in both the observation matrix B and data matrix A,

$$\inf_{[\Delta A|\Delta B]\in\mathbb{C}^{m\times(n+p)}} \|[\Delta A|\Delta B]\|_F, \quad (A-\Delta A)X=(B-\Delta B) \quad \text{has a solution.} \tag{1.3}$$

In the above equation if the infimum is attained i.e. there exist ΔA and ΔB such that $(A-\Delta A)X=(B-\Delta B)$ has a solution, then X is called the total least square solution. The following example shows that in some cases there is no total or data least square solution.

Example 1.1. Let $A=\begin{pmatrix}1&0\\0&0\end{pmatrix}$ and $B=\begin{pmatrix}1\\1\end{pmatrix}$. Let ε be an arbitrary non zero real number. Define $\Delta A=\begin{pmatrix}0&0\\0&\varepsilon\end{pmatrix}$ and $\Delta B=\begin{pmatrix}0\\0\end{pmatrix}$. Then $(A-\Delta A)X=(B-\Delta B)$ has a solution for all non zero $\varepsilon\in\mathbb{R}$ and hence $\inf\|[\Delta A|\Delta B]\|=0$. Thus, there is no total or data least square solution.

Although the name "total least squares" has appeared only recently in the literatures, this method of fitting is certainly not new and has a long history in the statistical literature where the method is known as orthogonal regression or errors-in-variables regression. In the field of numerical analysis, this problem was first studied by Golub and Van Loan [5]. Their analysis, as well as their algorithm, is strongly based on the Singular Value Decomposition (SVD). Some descriptions of TLS are given in [6, 7, 10, 11].

In this case, a best estimate of X is found by fitting a best subspace S to $[A;B]$ such that rank $(\hat{A})=\operatorname{rank}([\hat{A};\hat{B}])$, where $[\hat{A};\hat{B}]$ is the projection of $[A;B]$ into S. By solving this adjusted set $\hat{A}X=\hat{B}$, X is obtained and is called the total least square solution.

2. Data-Ordinary-Total Least Norm

In some situations, the use of a norm other than the Frobenius norm may be preferable. For example, if the data contains outliers, an L_1 norm might be more suitable [1].

Let $A\in\mathbb{C}^{m\times n}$ is a data matrix and $B\in\mathbb{C}^{m\times p}$ is an observation matrix. Define

$$P_t^s(A,B)=\{[tE;sF]\in\mathbb{C}^{m\times n}\times\mathbb{C}^{m\times p}:(A-tE)X=B-sF \text{ has a solution}\},$$

$$\rho_t^s(A,B)=\inf_{[tE;sF]\in P_t^s(A,B)}\|[tE;sF]\|. \tag{2.1}$$

where $\|.\|$ denotes an arbitrary norm on $\mathbb{C}^{m\times(n+p)}$ and $t,s\in\{0,1\}$.

If we consider the Frobenius norm, the above notations have explained ordinary least square problem whenever $t=0$, $s=1$, data least square problem whenever $t=1$, $s=0$ and total least square problem whenever $t=1$, $s=1$.

Also, if there exists an $[E_0; F_0] \in \mathbb{C}^{m\times n}\times\mathbb{C}^{m\times p}$ such that $\rho_t^s(A, B) = |[tE_0; sF_0]\|_F$, and $[tE_0; sF_0] \in P_t^s(A, B)$. Then any solution of $(A - tE_0)X = B - sF_0$ is called the ordinary least squares solution whenever $t = 0, s = 1$, data least squares solution whenever $t = 1, s = 0$ and total least squares solution whenever $t = 1, s = 1$.

Lemma 2.1. Let $(A, B) \in \mathbb{C}^{m\times n}\times\mathbb{C}^{m\times p}$ and let $\rho_t^s(A, B)$ be as in (2.1), then $\rho_1^1(A, B) \leq \min\{\rho_1^0(A, B), \rho_0^1(A, B)\}$.

Proof. Since $P_1^1(A, B) \subseteq P_1^0(A, B)$ and $P_1^1(A, B) \subseteq P_0^1(A, B)$, then $\rho_1^1(A, B) \leq \rho_0^1(A, B)$ and $\rho_1^1(A, B) \leq \rho_1^0(A, B)$. ■

Now, in the following examples we compare $\rho_1^0(A, B)$ and $\rho_0^1(A, B)$.

Example 2.2. Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\|.\|$ be an arbitrary norm on $\mathbb{C}^2$. Then, $0 = \rho_1^0(A, B) < \rho_0^1(A, B) = \left\|\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\|$.

Example 2.3. Let $A = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $B = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}$ and $\|.\|$ be an arbitrary permutation invariant norm on $\mathbb{C}^2$. Then, $\left\|\begin{pmatrix} 1/2 \\ 0 \end{pmatrix}\right\| = \rho_0^1(A, B) < \rho_1^0(A, B) = \left\|\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\|$.

Now, we state a necessary and sufficient condition $\rho_t^s(A, B) = 0$. The following Theorem cover and extend Proposition 3.1 [6].

Theorem 2.4. Let $(A, B) \in \mathbb{C}^{m\times n}\times\mathbb{C}^{m\times p}$ and let $\rho_t^s(A, B)$ be as in (2.1). The following are equivalent,

(i) $\operatorname{rank}([A|B]) \leq n$.

(ii) $\rho_1^0(A, B) = 0$.

(iii) $\rho_1^1(A, B) = 0$.

Proof. Let $\|.\|$ be an arbitrary norm on $\mathbb{C}^{m\times(n+p)}$. First we show that $(i) \Rightarrow (ii)$. Let $A = [A_1|\cdots|A_n]$ and $B = [B_1|\cdots|B_p]$ such that $\{A_{i_1}, \ldots, A_{i_k}\}$ is a basis for the column space of A and $\{A_{i_1}, \ldots, A_{i_k}, B_{j_1}, \ldots, B_{j_l}\}$ is a basis for the column space of $[A; B]$. Assume that $\operatorname{rank}([A|B]) \leq n$, then $k + l \leq n$. Since $k + l \leq n$, we define an $m \times n$ matrix $E = [E_1|\cdots|E_n]$ such that $E_1 = A_{i_1}, \ldots, E_{i_k} = A_{i_k}, E_{i_{k+1}} = B_{j_1}, \ldots, E_{i_{k+l}} = B_{j_l}$, and the rest of the columns of E be zero. The system of equations $(A - \delta E)X = B$ has a solution, because $\operatorname{rank}(A - \delta E) = \operatorname{rank}([A - \delta E|B]) = k + l$. Then $[\delta E|0] \in \mathcal{P}_1^0(A, B)$ for all $\delta > 0$. Let $\varepsilon > 0$ be given, then there exists a $\delta > 0$ such that $[\delta E|0] \in \mathcal{P}_1^0(A, B)$ and $\|[\delta E|0]\| \leq \varepsilon$. Therefore, $\rho_1^0(A, B) = 0$. By Lemma 2.1, $(ii) \Rightarrow (iii)$. It is enough to show that $(iii) \Rightarrow (i)$. Assume that $\rho_1^1(A, B) = 0$. Then, there exist $[E_k|F_k] \in P_1^1(A, B)$ such that $\|[E_k|F_k]\| \to 0$ as $k \to 0$. Since all norms in a finite dimensional spaces are equivalent. Then, there exists$\alpha > 0$ such that $\|.\|_F \leq \alpha\|.\|$. Hence $\|[E_k|F_k]\|_F \to 0$. By [8, 6], we known that for any $k = 1, 2, \ldots$, $\|[E_k|F_k]\|_F \geq \sigma_{n+1}$, where $\sigma_1 \geq \cdots \geq \sigma_{n+p}$ are the singular values of the matrix $[A; B]$. Then $\sigma_{n+1} = \cdots = \sigma_{n+p} = 0$, and hence $\operatorname{rank}[A; B] \leq n$. ■

Now, we state a necessary and sufficient condition that $\rho_0^1(A, B) = 0$

Proposition 2.5. Let $(A, B) \in \mathbb{C}^{m\times n} \times \mathbb{C}^{m\times p}$, $\rho_t^s(A, B)$ be as in (2.1)and $\|.\|$ be an arbitrary norm on $\mathbb{C}^{m\times p}$. Then $\rho_0^1(A, B) = 0$ if and only if col space$(B) \subseteq$ col space(A), where col space(A) denotes the column space of A.

Proof. Assume that col space$(B) \subseteq$ col space(A), then $[0; 0] \in \mathcal{P}_0^1(A, B)$ and hence $\rho_0^1(A, B) = 0$. Conversely, assume $\rho_0^1(A, B) = 0$. then, there exist $[0|F_k] \in P_0^1(A, B)$ such that $\|F_k\| \to 0$. This means that col space$(B - F_k) \subseteq$ col space(A) for all $k = 1, 2, \ldots$. Since col space(A) is a closed subspace and $\|B - F_k\| \to \|B\|$. Thus, col space$(B) \subseteq$ col space(A). ■

3. Free Structured Data-Ordinary-Total Least Norm Problem

Let $A \in \mathbb{C}^{m\times n}$ be a data matrix and $B \in \mathbb{C}^{m\times p}$ be an observation matrix. An alternative formulation of the total least squares problem as a matrix low rank approximation problem was introduced by S.Van Huffel [7] as, free total least square problem,

$$[\hat{A}, \hat{B}] := \arg\min_{[\hat{A},\hat{B}]} \|[A - \hat{A}, B - \hat{B}]\|_F \quad \text{subject to } \operatorname{rank}([\hat{A}, \hat{B}]) \leq n. \tag{3.1}$$

Example 1.1 shows that a total least squares problem may has no solution. But it is easy to see that a free total least squares problem always has a solution.

Now we define the free Data-Ordinary-Total Least Norm problems as follows:

$$\hat{A} := \arg\min_{\hat{A}} \|[A - \hat{A}, 0]\| \quad \text{subject to } \operatorname{rank}([\hat{A}, B]) \leq n, \quad \text{(Free Data Least Square)}. \tag{3.2}$$

$$\hat{B} := \arg\min_{\hat{B}} \|[0, B - \hat{B}]\| \quad \text{subject to } \operatorname{rank}([A, \hat{B}]) \leq n, \tag{3.3}$$
$$\text{(Free Ordinary Least Square)}.$$

$$[\hat{A}, \hat{B}] := \arg\min_{[\hat{A},\hat{B}]} \|[A - \hat{A}, B - \hat{B}]\| \quad \text{subject to } \operatorname{rank}([\hat{A}, \hat{B}]) \leq n, \tag{3.4}$$
$$\text{(Free Total Least Square)}.$$

In many applications of signal processing, system identification and system response prediction, the matrix A has a special structure, such as Toeplitz or Hankel matrix, see [1].

The important advantage of the Structured Data-Ordinary-Total Least square formulation is that it permits a known structure of the matrix A and $[A|B]$ to be preserved in $\hat{A}$ and $[\hat{A}|\hat{B}]$, respectively. It will be nice to consider the free structured Data Ordinary Total least square (norm) problems.

Acknowledgement

The research was supported by Mahani Mathematical Research Center and Linear Algebra and Optimization Center of Excellence of Shahid Bahonar University of Kerman.

References

[1] J. Ben Rosen, Haesun Park, and John Glick, Total least norm formulation and solution for structured problems, *SIAM J. Matrix Anal. Appl.*, **17**, pp. 110–126, 1996.

[2] Ake Björck, Numerical methods for least squares problems, *SIAM*, Philadelphia, 1996.

[3] R. Degroat and E. Dowling, The data least squares problem and channel equalization, *IEEE Trans. Signal Process.*, **41**, pp. 407–411, 1991.

[4] G. Golub and C. Van Loan, Matrix Computations, second edition, Johns Hopkins University Press, Baltimore, 1989.

[5] G. Golub and C. Van Loan, An analysis of the total least squares problem, *SIAM J. Numer. Anal.*, **17**, pp. 883–893, 1980.

[6] Chi-Kwong Li, Xin-Guo Liu and Xue-Feng Wang, Extension of the totl least square problem using general unitarily invariant norms, *Linear and Multilinear Algebra*, **55**, pp. 71–79, 2007.

[7] Ivan Markovsky and S. Van Huffel, Overview of total least-square methods, *Signal processing*, **87**, pp. 2283–2302, 2007.

[8] L. Mirsky, Symmetric gauge functions and unitarily invariant norms, *Quart. J. Math.*, **11**, pp. 50–59, 1960.

[9] C.C. Paige and Z. Strakos, Unifying least squares, total least squares and data least squares, in Total Least Squares and Errors-in-Variables Modeling, S. van Huffel and P. Lemmerling, eds., Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 25–34, 2002.

[10] S. Van Huffel and J. Vandewalle, The Total Least Squares Problem: Computational Aspects and Analysis, *SIAM*, Philadelphia, 1991.

[11] S. Van Huffel (Ed.), Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modeling, *SIAM*, Philadelphia, 1997.