Core problems in linear parameter estimation

Christopher C. Paige
School of Computer Science, McGill University, Montreal, Canada

Zdeněk Strakoš
Institute of Computer Science, Academy of Sciences, Prague, Czech Republic

http://www.cs.cas.cz/~strakos

COMPSTAT 2004, August 2004, Prague

1. Introduction

Modelling of Errors in Variables
Linear Parameter Estimation
Linear Regression (Orthogonal regression)

In the language of computational linear algebra:

Least Squares, Total Least Squares, Data Least Squares

Main tool for analysis and computation:
Singular value decomposition (SVD)

$\tilde{A}$ a nonzero n by k matrix, $\tilde{b}$ a nonzero n-vector

$$\tilde{A}\,\tilde{x} \;\approx\; \tilde{b}$$

where $\approx$ typically means using data corrections of the prescribed type in order to get the nearest compatible system

The size of the required minimal data correction (of $\tilde{b}$ in LS, of $\tilde{b}$ and $\tilde{A}$ in (Scaled) TLS, or of $\tilde{A}$ in DLS) represents the distance to the nearest compatible system

- when errors are confined to $\tilde{b}$: **LS**

$$\tilde{A}\,\tilde{x} \;=\; \tilde{b} + \tilde{r}\,, \qquad \min\,\|\tilde{r}\|$$

- when errors are contained in both $\tilde{A}$ and $\tilde{b}$: (Scaled) **TLS**

$$(\tilde{A} \;+\; \tilde{E})\,\tilde{x}\,\gamma \;=\; \tilde{b}\,\gamma \;+\; \tilde{r}\,, \qquad \min\,\|[\tilde{r},\;\tilde{E}]\|_F,$$

for a given scaling parameter γ

- when errors are restricted to $\tilde{A}$: **DLS**

$$(\tilde{A} \;+\; \tilde{E})\,\tilde{x} \;=\; \tilde{b}\,, \qquad \min\,\|\tilde{E}\|_F$$

2. Description of the difficulty

Suppose
$$\left[\begin{array}{c|c} \tilde{b} & \tilde{A} \end{array}\right] = \left[\begin{array}{c||c|c} b_1 & A_{11} & 0 \\ \hline 0 & 0 & A_{22} \end{array}\right],$$

so that the problem can be rewritten as two independent approximation problems

$$\begin{aligned} A_{11}\, x_1 &\approx b_1\,, \\ A_{22}\, x_2 &\approx 0\,, \end{aligned}$$

with the solution $\tilde{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

It seems that $A_{22}\ x_2 \ \approx \ 0$ has a meaningful solution $x_2 = 0$ and only $A_{11}\ x_1 \ \approx \ b_1$ need be solved.

However, the situation of TLS, DLS is not so simple. Assume, e.g., that for the TLS solution of $A_{11}\ x_1 \ \approx \ b_1$ we get the compatible system $(A_{11} + E_{11})\ x_1 \ = \ b_1 + r_1$, and

$$\sigma_{\min}\,(A_{22}) \ < \ \sigma_{11} \ \equiv \ \min \ \|[r_1\,,\ E_{11}]\|_F\,.$$

Then the above TLS solution x_1, with $x_2 = 0$, does not give a TLS solution of the whole $\tilde{A}\tilde{x} \ \approx \ \tilde{b}$, and x_1, x_2 are not computed by the basic algorithm suggested by Golub and Van Loan [1980], see also Van Huffel and Vandewalle [1991]:

- There exist $\hat{x}_1$, $\hat{x}_2 \neq 0$, possibly not optimal, such that

$$
\begin{aligned}
(A_{11} + \hat{E}_{11})\, \hat{x}_1 \;+\; \hat{E}_{12}\, \hat{x}_2 &= b_1 + \hat{r}_1 \\
\hat{E}_{21}\, \hat{x}_1 \;+\; (A_{22} + \hat{E}_{22})\, \hat{x}_2 &= \hat{r}_2
\end{aligned}
$$

and

$$
\left\| \left[\begin{array}{c|c|c} \hat{r}_1 & \hat{E}_{11} & \hat{E}_{12} \\ \hline \hat{r}_2 & \hat{E}_{21} & \hat{E}_{22} \end{array} \right] \right\|_F < \sigma_{11} \equiv \min \, \| [r_1 \, , \, E_{11}] \|_F
$$

- Golub, Van Loan algorithm: Compatibility condition

$$(\tilde{A} + \tilde{E})\,\tilde{x} \;=\; \tilde{b} + \tilde{r}$$

means

$$\left([\tilde{b}, \tilde{A}] \;+\; [\tilde{r}, \tilde{E}]\right) \begin{bmatrix} -1 \\ \tilde{x} \end{bmatrix} \;=\; 0\,.$$

Look for the smallest perturbation $[\tilde{r}, \tilde{E}]$ of $[\tilde{b}, \tilde{A}]$ which makes it rank deficient. If the right singular vector corresponding to the smallest singular value of $[\tilde{b}, \tilde{A}]$ has a non-zero first component, then scaling it so that the first component is -1 gives the basic TLS solution.

- Current techniques look for some $\hat{x}_1, \hat{x}_2$ by changing the problem (applying additional constraints) so that the idea of the basic GVL algorithm could be used. They need SVD of both $\tilde{A}$ and $[\tilde{b}, \tilde{A}]$.

Van Huffel, Vandewalle:

The concept of a nongeneric solution.

Observation:

Since the norms $\|\cdot\|$, $\|\cdot\|_F$ are orthogonally invariant, the trouble exists for all $\tilde{A}$, $\tilde{b}$ which can be orthogonally transformed to the given form,

$$P^T\,[\tilde{b}\,,\;\tilde{A}\,Q] \;=\; \left[\begin{array}{c||c|c} b_1 & A_{11} & 0 \\ \hline 0 & 0 & A_{22} \end{array}\right]$$

P and Q orthogonal.

3. Core problem within $\tilde{A}\tilde{x} \approx \tilde{b}$

Our suggestion is to find an orthogonal transformation

$$P^T\,[\tilde{b}\,,\ \tilde{A}\,Q] = \left[\begin{array}{c||c|c} b_1 & A_{11} & 0 \\ \hline 0 & 0 & A_{22} \end{array}\right], \quad P^{-1} = P^T\,, \quad Q^{-}$$

so that A_{11} has minimal dimensions and $A_{11}x_1 \approx b_1$ can be solved by the algorithm given by Golub and Van Loan. Then solve $A_{11}x_1 \approx b_1$, and take the original problem solution to be

$$\tilde{x} = Q \begin{bmatrix} x_1 \\ 0 \end{bmatrix}.$$

Such orthogonal transformation is given (assuming $\tilde{b} \neq 0$) by reducing $[\tilde{b}, \tilde{A}]$ to an upper bidiagonal matrix. In fact, A_{22} need not be bidiagonalized while upper bidiagonal $[b_1, A_{11}] = P_1^T [\tilde{b},\ \tilde{A}\, Q_1]$ has nonzero bidiagonal elements and is either

$$[b_1 \mid A_{11}] = \left[\begin{array}{c|cccc} \beta_1 & \alpha_1 & & & \\ & \beta_2 & \alpha_2 & & \\ & & \cdot & \cdot & \\ & & & \beta_p & \alpha_p \end{array}\right], \qquad \beta_i \alpha_i \neq 0, \quad i = 1, \ldots, p$$

if $\beta_{p+1} = 0$ or $p = n$, or

$$[b_1 \mid A_{11}] = \left[\begin{array}{c|ccccc} \beta_1 & \alpha_1 & & & & \\ & \beta_2 & \alpha_2 & & & \\ & & \cdot & \cdot & & \\ & & & \beta_p & \alpha_p & \\ & & & & \beta_{p+1} & \end{array}\right], \quad \beta_i \alpha_i \neq 0\,,\ i = 1, \ldots,$$

if $\alpha_{p+1} = 0$ or $p = k$.

From the construction, $[b_1, A_{11}]$ has full row rank and A_{11} has full column rank.

Technique: Householder reflections and
Golub-Kahan bidiagonalization.

Theorem

(a) A_{11} has no zero or multiple singular values, so any zero singular values or repeats that $\tilde{A}$ has must appear in A_{22};

(b) A_{11} has minimal dimensions, and A_{22} maximal dimensions, over all orthogonal transformations of the form given above;

(c) The solution of the TLS problem $A_{11}x_1 \approx b_1$ can be obtained by the algorithm of Golub and Van Loan.

4. Concluding remarks

- In theory, the core problem approach differs from the standard Van Huffel and Vandewalle approach for "non-basic" TLS problems.
- In practice, the suggested bidiagonalization (leading to the core problem) is an ideal first step in solving the total least squares, scaled total least squares or data least squares problems with single right hand sides.
- Unlike in the approach of Van Huffel and Vandewalle, the extension to multiple right hand side problems is not obvious.
- The paper will be submitted to CSDA, published papers are referenced in the abstract.

THANK YOU!