Core problems in $A\,x \;\approx\; b$

– analysis of TLS revisited

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

Householder Symposium 2005, Seven Springs, Pa

MOTIVATION

Motivation

Modeling of Errors in Variables, Linear Parameter Estimation,
Linear Regression (Orthogonal Regression) . . .

In the language of computational linear algebra:

Least Squares, (Scaled) Total Least Squares, Data Least Squares.

(We only consider orthogonally invariant measures).

Main tools for analysis and computation:

Orthogonal bidiagonalization,
Singular value decomposition (SVD).

Approximation problem

$\tilde{A}$ nonzero n by k matrix, $\tilde{b}$ nonzero n-vector. With no loss of generality $n > k$ (add zero rows if necessary). Consider

$$\tilde{A}\,\tilde{x} \;\approx\; \tilde{b}, \qquad (\tilde{A}^T\tilde{b} \neq 0 \text{ for simplicity}),$$

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, of $\tilde{A}$ in DLS)
represents the distance to the nearest compatible system.

Motivation

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

$$\tilde{A}\,\tilde{x} \;=\; \tilde{b} + \tilde{r}\,, \quad \min \|\tilde{r}\|_2\,;$$

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

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

for a given scaling parameter γ;

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

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

Motivation

The data $\tilde{A}$, $\tilde{b}$ can suffer from

- multiplicities – the solution may not be unique;
- conceptual difficulties – when there are stronger colinearities among the columns of $\tilde{A}$ than between the columnspace of $\tilde{A}$ and the right hand side $\tilde{b}$, the TLS solution does not exist.

Extreme example: $\tilde{A}$ not full column rank, but $\tilde{b} \notin \mathbf{R}(\tilde{A})$.

It would be ideal to separate the information necessary and sufficient for solving the problem from the information which is irrelevant or not needed.

Motivation

Extracting the necessary and sufficient information:

In order to minimize possible numerical difficulties, it should be done at the earliest possible stage of the solution process.

We prove that this important separation step can always be achieved via some orthogonal transformations.

The resulting block structure reveals the structure of information which is present, though in most cases invisible, in the original untransformed data. In this sense, any (scaled) TLS problem can be considered structured.

Motivation

For simplicity of exposition, the presentation is mostly restricted to (unscaled) TLS.

Except for very few exceptions specified below,
this presentation assumes exact arithmetic.

Content

1. Golub and Van Loan analysis

2. Extension of Van Huffel and Vandewalle

3. Conceptual difficulty – another look

4. Core problem within $\tilde{A}\,\tilde{x} = \tilde{b}$

5. Techniques, if time permits

6. Numerical issues, regularization of ill-posed problems

1. GOLUB AND VAN LOAN ANALYSIS

1. Golub and Van Loan analysis

Compatibility condition $(\tilde{A} + \tilde{E})\,\tilde{x} \;=\; \tilde{b} + \tilde{r}$ is equivalent to

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

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

1. Golub and Van Loan analysis

Theorem

If $\sigma_{\min}(\tilde{A}) > \sigma_{\min}([\tilde{b}\,,\ \tilde{A}])$, then the Algorithm GVL gives the unique solution,

$$[\tilde{b}\,,\ \tilde{A}] = \tilde{U}\,\tilde{\Sigma}\,\tilde{V}^T = \sum_{i=1}^{k+1} \tilde{u}_i\,\tilde{\sigma}_i\,\tilde{v}_i^T\,, \qquad \tilde{v}_{k+1} = \begin{bmatrix} \nu \\ w \end{bmatrix},$$

$$\tilde{x} = -\frac{1}{\nu}\,w\,, \qquad [\tilde{r}\,,\ \tilde{E}] = -\,\tilde{u}_{k+1}\,\tilde{\sigma}_{k+1}\,\tilde{v}_{k+1}^T\,.$$

[Golub - 73], [Golub, Van Loan - 80], (see also
[Golub, Hoffman, Stewart - 87]) contain much more, in particular,

1. Golub and Van Loan analysis

- Scaling of columns and weighting of rows;
- Minimum 2-norm solution;
- Scaled TLS solution $\longrightarrow$ LS solution as $\gamma \searrow 0$;
- TLS sensitivity analysis;
- Enlightening comments on possible numerical difficulties.

1. Golub and Van Loan analysis

The condition $\sigma_{\min}(\tilde{A}) > \sigma_{\min}([\tilde{b}, \tilde{A}])$ is sufficient,

but not necessary: If $\sigma_{\min}(\tilde{A}) = \sigma_{\min}([\tilde{b}, \tilde{A}])$,

then there might be a solution, or it *can* happen that

$$\tilde{v}_{k+1} = \begin{bmatrix} 0 \\ w \end{bmatrix}$$

and the TLS formulation does not have a solution.

1. Golub and Van Loan analysis

The minimum norm solution: (Remember $[\tilde{b}\,,\ \tilde{A}] \;=\; \tilde{U}\,\tilde{\Sigma}\,\tilde{V}^T$)

$$\tilde{\sigma}_j \;>\; \tilde{\sigma}_{j+1} \;=\; \ldots \;=\; \tilde{\sigma}_{k+1}\,, \quad \begin{aligned} V' &= [\tilde{v}_{j+1}\,,\ \ldots\ ,\ \tilde{v}_{k+1}]\,, \\ U' &= [\tilde{u}_{j+1}\,,\ \ldots\ ,\ \tilde{u}_{k+1}]\,. \end{aligned}$$

If $e_1^T\,V' \;\neq\; 0$, then take Q', $Q'^T Q' \;=\; Q' Q'^T \;=\; I$ such that

$(e_1^T\,V')\,Q' \;=\; \nu\,e_1^T$; set $\tilde{v} \;=\; (V'\,Q')\,e_1 \;=\; \begin{bmatrix} \nu \\ w \end{bmatrix}$, $\tilde{u} \;=\; U'\,Q'\,e_1$.

The solution is given by

$$x \;=\; -\,\frac{1}{\nu}\,w\,, \qquad [\tilde{r}\,,\ \tilde{E}] \;=\; -\,\tilde{u}\;\tilde{\sigma}_{k+1}\,\tilde{v}^T\,.$$

2. EXTENSION OF VAN HUFFEL AND VANDEWALLE

2. Extension of Van Huffel and Vandewalle

If $e_1^T V' = 0$, i.e. no column of V' has a nonzero first component, then the corresponding directions in the columnspace of $\tilde{A}$ bear no information whatsoever about the “observation” or “response” $\tilde{b}$. In other words, the correlations between the columns of $\tilde{A}$ are stronger than the correlations between the columnspace of $\tilde{A}$ and the vector $\tilde{b}$.

[Van Huffel, Vandewalle – 91]:

Eliminate some unwanted directions in the columnspace of $\tilde{A}$ (nonpredictive colinearities) uncorrelated with the vector $\tilde{b}$.

2. Extension of Van Huffel and Vandewalle

Consider the splitting

$$[\tilde{b}\,,\ \tilde{A}] \;=\; \sum_{i=1}^{q} \tilde{u}_i\,\tilde{\sigma}_i\,\tilde{v}_i^T \;+\; \sum_{i=q+1}^{k+1} \tilde{u}_i\,\tilde{\sigma}_i\,\tilde{v}_i^T\,,$$

where q is the maximal value of i such that $e_1^T\,\tilde{v}_i \neq 0$.

The *nongeneric* TLS formulation uses the additional restriction:

$$(\tilde{A} + \tilde{E})\,\tilde{x} \;=\; \tilde{b} + \tilde{r}\,, \quad \min \|\,[\tilde{r}\,,\ \tilde{E}]\,\|_F \quad \text{subject to}$$

$$[\tilde{r}\,,\ \tilde{E}]\,[\tilde{v}_{q+1}\,,\ \ldots\,,\ \tilde{v}_{k+1}] \;=\; 0\,.$$

Theorem

The nongeneric TLS solution always exists,
the minimum norm nongeneric TLS solution is unique.

We call the nongeneric extension of Van Huffel and Vandewalle EVHV.

Any decision as to whether the problem is generic or nongeneric can be made only after completing the SVD of $[\tilde{b}\,,\ \tilde{A}]$.

2. Extension of Van Huffel and Vandewalle

The nongeneric approach completes the definition of (Scaled) TLS. It always leads to a meaningful well justified solution. The computation, however, does not remove all directions in the column space of $\tilde{A}$ uncorrelated with the vector $\tilde{b}$, nor all redundant data.

The basic condition is $\sigma_{\min}(\tilde{A}) > \sigma_{\min}([\tilde{b}, \tilde{A}])$.

When $\sigma_{\min}(\tilde{A}) = \sigma_{\min}([\tilde{b}, \tilde{A}])$ there might still be a solution, and this can be extended to the minimum norm solution in the case of nonuniqueness. The theory was then advanced by the nongeneric extension EVHV. The fact that the basic condition is sufficient but not necessary complicates the whole theory and computations.

3. CONCEPTUAL DIFFICULTY – ANOTHER LOOK

3. Conceptual difficulty – another look

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

so that the problem $Ax \approx b$ 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 $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

3. Conceptual difficulty – another look

But $A_{22}x_2 \approx 0$ says x_2 lies approximately in the null space of A_{22}, and no more.

Thus unless there is a reason not to, we can set $x_2 = 0$.

Now since we have obtained b with the intent to estimate x,
and since x_2 does not contribute to b in any way —

the best we can do is estimate x_1 from $A_{11}\,x_1 \approx b_1$.

3. Conceptual difficulty – another look

We need only consider the case where $Ax \approx b$ is incompatible.
Then $A_{11}x_1 \approx b_1$ is also incompatible.
We will show later that we can get:

- A_{11} is a $(p+1) \times p$ matrix with no zero or multiple singular values,
- b_1 has nonzero components in all left singular vector subspaces of A_{11}. That is if $A_{11} = U_{11}\Sigma_1 V_{11}^T$, then $U_{11}^T b_1$ has no zero entry.

As a consequence we will have the desired basic condition:

- $\sigma_{\min}(A_{11}) > \sigma_{\min}([b_1, A_{11}])$.

3. Conceptual difficulty – another look

What will the standard approaches give?

The SVD of $[b, \; A]$ is the direct sum of the SVDs of $[b_1, \; A_{11}]$ and A_{22}. Indeed,

$$\left[\begin{array}{c||c|c} b_1 & A_{11} & 0 \\ \hline 0 & 0 & A_{22} \end{array} \right] = \left[\begin{array}{c|c} U_1 \, \Sigma_1 \, V_1^T & 0 \\ \hline 0 & U_2 \, \Sigma_2 \, V_2^T \end{array} \right] ,$$

then extend the singular vectors by zeros.

3. Conceptual difficulty – another look

Since $\sigma_{\min}(A_{11}) > \sigma_{\min}([b_1, A_{11}])$,

- $\sigma_{\min}(A_{22}) > \sigma_{\min}([b_1, A_{11}])$ implies $\sigma_{\min}(A) > \sigma_{\min}([b, A])$ and the algorithm of Golub-Van Loan (AGVL) finds the unique solution.

- $\sigma_{\min}(A_{22}) = \sigma_{\min}([b_1, A_{11}])$ implies $\sigma_{\min}(A) = \sigma_{\min}([b, A])$; $\sigma_{\min}([b, A])$ is multiple, but $e_1^T V' \neq 0$. Consequently, AGVL finds the unique minimum norm solution.

- $\sigma_{\min}(A_{22}) < \sigma_{\min}([b_1, A_{11}])$ implies $\sigma_{\min}(A) = \sigma_{\min}([b, A])$ and $e_1^T V' = 0$. The problem is considered by AGVL unsolvable. The nongeneric extension EVHV has to be applied.

3. Conceptual difficulty – another look

The EVHV projects out (by imposing the additional condition) “the part of the block” A_{22} with singular values below $\sigma_{\min}([b_1, A_{11}])$. Then it solves the projected problem using the standard (minimum norm solution) approach.

The situation is illustrated on a simple example.

3. Conceptual difficulty – another look

$$[b\,,\ A] = \left[\begin{array}{c||c|c} b_1 & A_{11} & 0 \\ \hline 0 & 0 & A_{22} \end{array}\right] = \left[\begin{array}{c||c|c} 1 & 1 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & \omega \end{array}\right]$$

$$\text{SVD of} \quad [b_1\,,\ A_{11}] =$$

$$\begin{bmatrix} 0.8507 & -0.5257 \\ 0.5257 & 0.8507 \end{bmatrix} \begin{bmatrix} 1.618 & 0 \\ 0 & 0.618 \end{bmatrix} \begin{bmatrix} 0.5257 & -0.8507 \\ 0.8507 & 0.5257 \end{bmatrix}^T$$

- If $\omega \geq \sigma_{\min}([b_1, A_{11}]) = 0.618$, then all is fine.
- If $\omega < \sigma_{\min}([b_1, A_{11}]) = 0.618$, then we see the trouble:

3. Conceptual difficulty – another look

Take any z, define $r_1 = b_1 - A_{11} z$.

Then for any $\theta > 0$, (denoting v_2 , u_2 the singular vectors corresponding to $\sigma_{\min}(A_{22}) \equiv \sigma_2$, here $v_2 = 1$, $u_2 = 1$, $\sigma_{\min}(A_{22}) = \omega$)

$$\left[\begin{array}{c||c|c} b_1 & A_{11} & r_1 \theta^{-1} v_2^T \\ \hline 0 & 0 & A_{22} - u_2 \sigma_2 v_2^T \end{array} \right] \begin{bmatrix} -1 \\ z \\ v_2 \theta \end{bmatrix} \equiv$$

$$\left[\begin{array}{c||c|c} b_1 & A_{11} & r_1 \theta^{-1} \\ \hline 0 & 0 & 0 \end{array} \right] \begin{bmatrix} -1 \\ z \\ \theta \end{bmatrix} = 0 ,$$

3. Conceptual difficulty – another look

For large θ we have $\| [r\,,\ E] \|_F \;\longrightarrow\; \sigma_{\min}\left([A_{22}]\right) \;=\; \omega$ and "close to optimal solution vector"

$$\begin{bmatrix} z \\ v_2\,\theta \end{bmatrix} \equiv \begin{bmatrix} z \\ \theta \end{bmatrix}$$

which is absolutely meaningless, since it couples the blocks and reflects no useful information whatsoever.

The nongeneric EVHV imposes condition $[r\,,\ E]\ [0\,,\ 0\,,\ 1]^T \;=\; 0$,
and constructs the unique nongeneric solution from the block $[b_1\,,\ A_{11}]$.

Motivation for the next step

In this section, the problem was structured so that the difficulty was clearly revealed and the solution was transparent.

We claim and show that analogous structure, fully determined by the multiplicities and irelevant information in the data $\tilde{b}\,,\ \tilde{A}$ can always be found via proper orthogonal transformations.

The solution can then be found by ignoring all multiplicities and irelevant information (i.e. block A_{22}).

4. CORE PROBLEM

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

4. Core problem within $\tilde{A}\,\tilde{x} \;\approx\; \tilde{b}$

Our suggestion is to find an orthogonal transformation

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

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

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

4. Core problem within $\tilde{A}\,\tilde{x} \;\approx\; \tilde{b}$

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

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

if $\beta_{p+1} = 0$ or $p = n\,,$ (where $\tilde{A}$ is $n \times k$), or

4. Core problem within $\tilde{A}\,\tilde{x} \;\approx\; \tilde{b}$

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

if $\alpha_{p+1} = 0$ or $p = k$ (where $\tilde{A}$ is $n \times k$).

In both cases: $[b_1, A_{11}]$ has full row rank and A_{11} has full column rank.

Technique: Householder reflections or Lanczos-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) All components of b_1 in the left singular vector subspaces of A_{11} are nonzero. Consequently, 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. Core problem within $\tilde{A}\,\tilde{x} \approx \tilde{b}$

The core problem approach consists of three steps:

1. Orthogonal transformation $[b,\, A] = P^T\,[\tilde{b}\,,\, \tilde{A}\,Q]\,,$ where the upper bidiagonal block $[b_1,\, A_{11}]$ is as above and A_{22} is not bidiagonalized. All irrelevant and multiple information is filtered out to A_{22} .

2. Solving the minimally dimensioned $A_{11}\,x_1 \approx b_1$ by AGVL.

3. Setting $\tilde{x} = Q\,x \equiv Q \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$, (if we take $x_2 = 0$).

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

The core problem approach does not need to complete the SVD of all of $[\tilde{b}\,,\; \tilde{A}]$. When the bidiagonalization stops, we use only the necessary (and sufficient) information for computing the solution.

The approximation problems for the original data $[\tilde{b},\, \tilde{A}]$ and the orthogonally transformed data $[b,\, A]$ are equivalent. Consequently the core problem approach always gives meaningful solutions by setting $x_2 = 0$.

4. Core problem within $\tilde{A}\,\tilde{x} \;\approx\; \tilde{b}$

Theorem

The core problem approach gives in exact arithmetic the minimum norm (Scaled) TLS solution of $\tilde{A}\tilde{x} \approx \tilde{b}$ determined by the algorithm of Golub and Van Loan, if it exists. If such a solution does not exist, then the core problem approach gives the nongeneric minimum norm (Scaled) TLS solution determined by the algorithm of Van Huffel and Vandewalle.

5. TECHNIQUES, IF TIME PERMITS

5. Techniques, if time permits

5.1. Understanding core problems. Start with the SVD of A:

$$[\tilde{b}, \tilde{A}] = \left[\tilde{b} \;\middle|\; U \left[\begin{array}{c|c} S & 0 \\ \hline 0 & 0 \end{array}\right] V^T \right] = U \left[\begin{array}{c||c|c} \tilde{c} & S & 0 \\ \hline d & 0 & 0 \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & V^T \end{array}\right]$$

Use orthogonal transformations from the left and right in order to

- transform nonzero d to δe_1;
- create as many zeros in $\tilde{c}$ as possible;
- move out all zeros in $\tilde{c}$,
- and so move out all multiplicities and unneeded elements in S.

5. Techniques, if time permits

Result (with new U, V):

$$U^T[\tilde{b}, \| \tilde{A}V] = \left[\begin{array}{c||c|c} b_1 & A_{11} & 0 \\ \hline 0 & 0 & A_{22} \end{array}\right] = \left[\begin{array}{c||c|c} c & S_1 & 0 \\ \delta & 0 & 0 \\ \hline 0 & 0 & S_2 \end{array}\right]$$

δ is nonzero (and the corresponding row exists) if and only if the system is incompatible. Size of the core problem ($p \times p$ or $(p+1) \times p$) is given by the number of the left singular subspaces of $\tilde{A}$, corresponding to distinct nonzero singular values, in which $\tilde{b}$ has a nonzero component. (c has all its components nonzero, singular values in S_1 are distinct and nonzero).

5.2. Obtaining this structure from the bidiagonalization

Upper bidiagonalization of $[\tilde{b}, \tilde{A}]$. Then, using $A_{11} = U_{11} S_1 V_{11}^T$, (obtaining $A_{22} = U_{22} S_2 V_{22}^T$ is unnecessary),

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

$$\left[\begin{array}{cc|c} U_{11} & r_1 & 0 \\ \hline 0 & 0 & U_{22} \end{array}\right] \left[\begin{array}{c||c|c} c & S_1 & 0 \\ \delta & 0 & 0 \\ \hline 0 & 0 & S_2 \end{array}\right] \left[\begin{array}{c|c|c} 1 & 0 & 0 \\ 0 & V_{11}^T & 0 \\ \hline 0 & 0 & V_{22}^T \end{array}\right]$$

where $c \equiv U_{11}^T b_1$, $\delta \equiv \|w\| \equiv \|b_1 - U_{11} c\|$, and, if $\delta \neq 0$, $r_1 \equiv w/\delta$.

5.3. Equivalence with the minimum norm TLS

Orthogonal transformations do not change the problem. Therefore, consider the (partial) upper bidiagonal form in the incompatible case (the compatible case is obvious).

$$[b, A] = \left[\begin{array}{c||c|c} b_1 & A_{11} & 0 \\ \hline 0 & 0 & A_{22} \end{array}\right] = \left[\begin{array}{c|cccc|c} \beta_1 & \alpha_1 & & & & \\ & \beta_2 & \ddots & & & 0 \\ & & & \ddots & \alpha_p & \\ & & & & \beta_{p+1} & \\ \hline 0 & & & 0 & & A_{22} \end{array}\right]$$

5. Techniques, if time permits

Case 1: $\sigma_{\min}(A) > \sigma_{\min}([b, A]) > 0$.

Case 2: $\sigma_j([b, A]) > \sigma_{j+1}([b, A]) = \ldots = \sigma_{k+1}([b, A])$,

$$V' = [\tilde{v}_{j+1}, \tilde{v}_{j+2}, \ldots, \tilde{v}_{k+1}],$$

Case 2a: $e_1^T V' \neq 0$.

Case 2b: $e_1^T V' = 0$.

6. NUMERICAL ISSUES, REGULARIZATION OF ILL-POSED PROBLEMS

6. Numerical issues and regularization

Numerically, determining b_1 , A_{11} , A_{22} will depend on some threshold criterion.

If the problem is ill-posed and the data are corrupted by noise, then determining and solving the numerical core problem should also incorporate some way of determining what we can of a meaningful solution, such as regularization.

A survey of regularization in connection with TLS is given in [Hansen, O'Leary –97], [Golub, Hansen, O'Leary – 99].
Also in computational statistics, and the Russian school inspired by Tikhonov [Zhdanov et al. – 86, 89, 90, 91].

Truncated TLS

$$(A+E)\,x \;=\; b \,+\, r, \quad \min \|\,[r, E]\,\|_F \quad \text{subject to}$$

$$(\text{rank}\;([b+r,\; A+E]) =)\;\; \text{rank}\;(A+E) = m\,.$$

Its (minimum norm nongeneric TLS) solution is constructed by considering the small singular values equal and set to zero, while preserving the singular vectors. With the restriction of the rank, the T-TLS distance is (unlike in the nongeneric TLS problem) the sqare root of the sum of squares of the neglected singular values.

Suggested in [van Huffel, Vandewalle - 91, Section 3.6.1].
Analyzed in [Fierro, Bunch – 94], [Fierro, Bunch – 96], [Wei – 92], see also [Stewart – 84], [van der Sluis, Veltkamp – 79].

Lanczos Truncated TLS

Lanczos bidiagonalization of $[\tilde{b}\,,\ \tilde{A}]$. Then compute an approximate truncated TLS solution by applying TLS to the bidiagonal system with the $(k+1) \times k$ matrix at each step k . Stopping criterion is based on the TLS solution of the $(k+1)$ by k bidiagonal problem.

[Fierro, Golub, Hansen, O'Leary – 97], [Sima, Van Huffel – 05]

Lanczos Truncated TLS "approximates" the core problem.

6. Numerical issues and regularization

An analogy for solving ill-posed LS problems?

LSQR [Paige, Saunders – 82],
[Björck – 88], [Björck, Grimme, Van Dooren – 94],
see also [O'Leary, Simmons – 81], [Hanke, Hansen – 93], [Hanke – 01],
book [Hansen – 98], . . .

Another field uses different names:

Principal component regression (Truncated SVD) [Massy – 65],
partial least squares [Wold – 75], see the explanatory paper [Elden – 04].

6. Numerical issues and regularization

In regularization of noisy ill-posed problems, interesting questions remain open. Consider, e.g., noisy ill-posed LS problems and Modified TSVD [Hansen, Sekii, Shibahaski – 92]

$$\min \; \| L\tilde{x} \|_2 \quad \text{subject to} \quad \min \; \| \tilde{A}\tilde{x} - \tilde{b} \| \, .$$

If L is a general matrix with full row rank, then one can consider $x_2 \neq 0$ for numerically determined A_{22} . This does not alter the core problem concept theoretically or computationally,

cf. [Fierro, Golub, Hansen, O'Leary – 97, Section 5].

CLOSING REMARKS

Closing remarks

The core problem approach represents a clear computationally efficient concept which in exact arithmetic gives in all cases (Scaled) TLS solutions identical to the minimum norm solutions given by AGVL resp. EVHV.

Theoretically, it simplifies and extends the previous (Scaled) TLS analysis.

Computationally, it can lead to interesting numerical questions and applications. A close connection to regularization.

It needs to be extended to problems with multiple right hand sides.

Closing remarks

THANK YOU!