Core reduction: Necessary and sufficient information in linear approximation problems

Martin Plešinger

Department of Mathematics, TU Liberec

In collaboration with: Iveta Hnětynková, Jana Žáková, and many others

8ECM 2020, Portorož — June 21, 2021

Outline

Linear approximation problems and the TLS

Matrix right-hand side TLS minimization

Generalizations of the matrix case

Final comments and summary

Linear approximation problem

Let $A \in \mathscr{L}(\mathscr{U}, \mathscr{V})$ be a linear mapping from $\mathscr{U}$ to $\mathscr{V}$, finite-dimensional linear vector spaces, and the range $\mathscr{R}(A)$ is a proper subspace of $\mathscr{V}$.

We are interested in a linear approximation problem

$$\boxed{A(x) \approx b, \qquad x \in \mathscr{U}, \qquad b \in \mathscr{V} \setminus \mathscr{R}(A)}$$

where the inconsistency is due to errors in both A and b.

By introducing norms in $\mathscr{U}$, $\mathscr{V}$, and $\mathscr{L}(\mathscr{U}, \mathscr{V})$, we replace the original problem by some minimization, with the goal to approximate x:

☞ Least squares techniques ☜

Total least squares minimization (TLS)

The simplest case of linear approximation problem

$$A\,x \approx b\,, \qquad A \in \mathbb{R}^{m\times n}, \quad x \in \mathbb{R}^{n}, \quad b \in \mathbb{R}^{m},$$

can be replaced, e.g., by the total least squares minimization

$$\boxed{\min_{e,G} \, \| [\; e \;|\; G \;] \|_{\mathsf{F}} \qquad \text{subject to} \qquad b + e \in \mathscr{R}(A+G)}$$

If such e and G exist, then we use the (minimum norm) TLS solution

$$x \;\approx\; x_{\mathsf{TLS}} \equiv (A+G)^{\dagger}(b+e)$$

Any trouble? The minimal correction may not exist.

See [Golub, Van Loan, 80], [Van Huffel, Vandewalle, 91].

Orthogonal invariance of TLS formulation

Since the Frobenius norm $\|\cdot\|_\mathsf{F}$ used in TLS minimization is orthogonally invariant, then for any two orthogonal matrices of suitable size,

$$\forall (P, Q) \in \mathbb{O}(m) \times \mathbb{O}(n), \qquad \text{i.e.,} \qquad P^{-1} = P^\mathsf{T},\ Q^{-1} = Q^\mathsf{T}$$

we get

$$\boxed{\begin{array}{ccc} A\,x \approx b & \longleftrightarrow & \widetilde{A}\,\widetilde{x} \equiv (P^\mathsf{T} A Q)(Q^\mathsf{T} x) \approx (P^\mathsf{T} b) \equiv \widetilde{b} \\ \downarrow & & \downarrow \\ x_{\mathsf{TLS}} & \longleftrightarrow & \widetilde{x}_{\mathsf{TLS}} = Q^\mathsf{T} x_{\mathsf{TLS}} \end{array}}$$

provided the TLS solution(s) exist.

Looking for a useful representative — the Core Problem

There exist orthogonal matrices $P_\star$ and $Q_\star$, such that

$$\left[\, \widetilde{b} \,\middle|\, \widetilde{A} \,\right] = P_\star^T \left[\, b \,\middle|\, A \,\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_\star \end{array}\right] = \left[\begin{array}{c|cc} b_1 & A_{11} & 0 \\ 0 & 0 & A_{22} \end{array}\right],$$

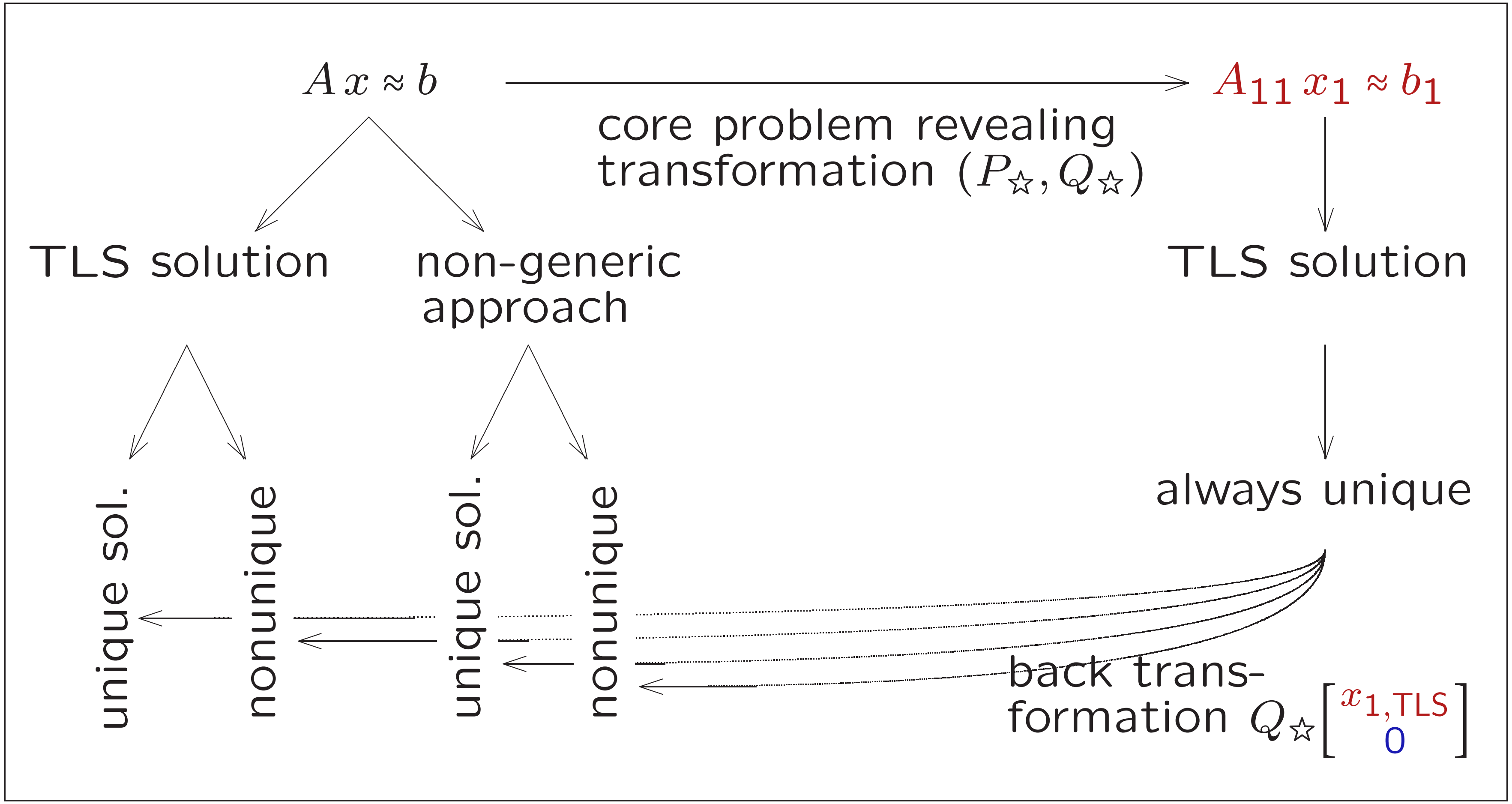
i.e.,

$$A\,x \approx b \qquad \longleftrightarrow \qquad A_{11}\,x_1 \approx b_1 \quad \& \quad A_{22}\,x_2 \approx 0,$$

and $A_{11}\,x_1 \approx b_1$, called the *core problem* (CP) has minimal dimensions over all $(P,Q) \in \mathbb{O}(m) \times \mathbb{O}(n)$.

The CP has a lot of interesting properties. In particular, for any (A,b) the CP within always has the unique TLS solution $x_{1,\mathrm{TLS}}$; see [Paige, Strakoš, 06], [Hnětynková, Strakoš, 07].

Insight to the TLS via the core problem



For the non-generic approach see [Van Huffel, Vandewalle, 91].

Outline

Linear approximation problems and the TLS

Matrix right-hand side TLS minimization

Generalizations of the matrix case

Final comments and summary

Matrix right-hand side (RHS) case

Similarly to the vector (or single) right-hand side problem,

$$Ax \approx b\,, \qquad A \in \mathbb{R}^{m\times n}, \quad x \in \mathbb{R}^{n}, \quad b \in \mathbb{R}^{m},$$

one can consider the matrix (or multiple) right-hand side problem,

$$AX \approx B\,, \qquad A \in \mathbb{R}^{m\times n}, \quad X \in \mathbb{R}^{n\times d}, \quad B \in \mathbb{R}^{m\times d},$$

with the corresponding TLS minimization

$$\boxed{\min_{E,G} \|[\, E \,|\, G \,]\|_{\mathsf{F}} \qquad \text{subject to} \qquad \mathscr{R}(B+E) \subseteq \mathscr{R}(A+G)}$$

see [Van Huffel, Vandewalle, 91].

Again, the TLS solution not exist, but the overall behavior is even more tangled; see [Hnětynková, P, Sima, Strakoš, Van Huffel, 11].

Similarly, for any three orthogonal matrices of suitable size, we get

$$\forall (P, Q, R) \in \mathbb{O}(m) \times \mathbb{O}(n) \times \mathbb{O}(d)$$

$$\boxed{\begin{array}{ccc} AX \approx B & \longleftrightarrow & \widetilde{A}\widetilde{x} \equiv (P^{\mathsf{T}}AQ)(Q^{\mathsf{T}}XR) \approx (P^{\mathsf{T}}BR) \equiv \widetilde{b} \\ \downarrow & & \downarrow \\ X_{\mathsf{TLS}} & \longleftrightarrow & \widetilde{X}_{\mathsf{TLS}} = Q^{\mathsf{T}}X_{\mathsf{TLS}}R \end{array}}$$

provided the TLS solution(s) exist.

Moreover, there exists triplet $(P_{\star}, Q_{\star}, R_{\star})$ so that

$$[\ \widetilde{B} \mid \widetilde{A}\ ] = P_{\star}^{\mathsf{T}} [\ B \mid A\ ] \left[\begin{array}{c|c} R_{\star} & 0 \\ \hline 0 & Q_{\star} \end{array} \right] = \left[\begin{array}{cc|cc} B_1 & 0 & A_{11} & 0 \\ 0 & 0 & 0 & A_{22} \end{array} \right],$$

and $A_{11}X_{11} \approx B_1$, called the core problem has minimal dimensions over all $(P, Q, R) \in \mathbb{O}(m) \times \mathbb{O}(n) \times \mathbb{O}(d)$; see [Hnětynková, P, Strakoš, 13, 15].

Properties of and troubles with CPs

Minimality of $A_{11}X_{11} \approx B_1$ in the matrix RHS case (and of $A_{11}x_1 \approx b_1$ in the vector RHS case) guarantees a lot of interesting properties:

- A_{11} is of full column rank.
- B_1 is of full column rank (b_1 is nonzero).
- $[B_1, A_{11}]$ ($[b_1, A_{11}]$) is of full row rank.
- Singular values of A_{11} and $[B_1, A_{11}]$ ($[b_1, A_{11}]$) are of mutiplicities at most $\mathsf{rank}(B_1)$ (are simple).
- Matrices* $U_j^\top B_1$ are of full row rank (b_1 has nonzero components in all left singular vectors sub-spaces of A_{11}).

*Columns of U_j form an orthomormal basis of left singular vector sub-space of A_{11} corresponding to the jth largest singular value of A_{11}.

The core problems can be:

- obtained by emploiyng the SVD of A or
- *computed* by the (band -or- block generalized) Golub–Kahan iterative bidiagonalization of A with column(s) of B (b) as starting vector(s).

Unfortunatelly, although the CP sligtly *simplifies* the behavior of the matrix RHS TLS minimization, the core problem itself

$$A_{11} X_{11} \approx B_1$$

still may not have a TLS solution contrary to the vector RHS case; see [Hnětynková, P, Sima, 16].

☞ CP does not proivde the insight to the matrix RHS TLS ☜

Two possible directions of further analysis

The matrix right-hand side
core problem $A_{11}\,X_{11} \approx B_1$
may not have a TLS solution

↙ ↘

Get closer: Look inside the core problem and explore its internal structure, if there is any.

Get more far: Look at the core problem in wider context and explore its surroundings. ✓

Outline

Linear approximation problems and the TLS

Matrix right-hand side TLS minimization

Generalizations of the matrix case

Final comments and summary

Generalizations: I. Tensor right-hand side

The very first and very natural way is a further generalization to tensor right-hand side linear approximation problems:

Vector RHS:	$Ax \approx b$,	$A \in \mathbb{R}^{m \times n}$,	$x \in \mathbb{R}^{n}$,	$b \in \mathbb{R}^{m}$
	$\downarrow$			
Matrix RHS:	$AX \approx B$,	$A \in \mathbb{R}^{m \times n}$,	$X \in \mathbb{R}^{n \times d}$,	$B \in \mathbb{R}^{m \times d}$
	$\downarrow$			
Tensor RHS:	$\boxed{A \times_1 \mathcal{X} \approx \mathcal{B}}$	$A \in \mathbb{R}^{m \times n}$,	$\mathcal{X} \in \mathbb{R}^{n \times d_2 \times \cdots \times d_k}$,	$\mathcal{B} \in \mathbb{R}^{m \times d_2 \times \cdots \times d_k}$

Here "$\times_\ell$" denotes the ℓ-mode matrix-tensor (pre-)mutlipliciation

$$(Z \times_\ell \mathcal{T})_{j_1,\ldots,j_{\ell-1},i,j_{\ell+1},\ldots,j_k} = \sum_{\alpha=1}^{n_\ell} z_{i,\alpha} \cdot t_{j_1,\ldots,j_{\ell-1},\alpha,j_{\ell+1},\ldots,j_k}$$

I. Tensor right-hand side: TLS point of view

Suitable applications:

- Collections of snapshots of time-dependent matrix RHS case:
  $A\,X(t) \approx B(t)$, $t \in \{t_1, t_2, \ldots\}$ $(k = 3)$.
- Problem depending on several parameters:
  $A\,x(p_1, \ldots, p_k) \approx b(p_1, \ldots, p_k)$, $p_\ell \in \{p_{\ell,1}, p_{\ell,2}, \ldots\}$, $\ell = 1, \ldots, k$.

The corresponding TLS minimization:

$$\boxed{\min_{\mathcal{E}, G} \left(\|\mathcal{E}\|_2^2 + \|G\|_F^2 \right)^{\frac{1}{2}} \qquad \text{subject to} \qquad \mathscr{R}\Big((\mathcal{B} + \mathcal{E})^{\{1\}} \Big) \subseteq \mathscr{R}(A + G)}$$

$\mathcal{T}^{\{\ell\}}$ is the ℓ-mode unfolding of $\mathcal{T}$, a matrix of ℓ-mode fibres as col's.
The TLS minimization is <u>essentially the same</u> as in the matrix case.

I. Tensor right-hand side: CP point of view

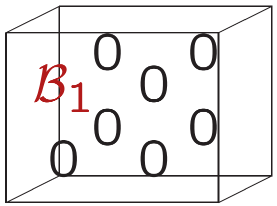
On the other hand, the CP theory can be extended to the tensor RHS case. Moreover, it can be shown, there exist

$$(P_\star, Q_\star, R_{2,\star}, \ldots, R_{k,\star}) \in \mathbb{O}(m) \times \mathbb{O}(n) \times \mathbb{O}(d_2) \times \cdots \times \mathbb{O}(d_k),$$

so that

$$\widetilde{A} = P_\star^{\mathsf{T}} A \, Q_\star = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix},$$

$$\widetilde{\mathcal{B}} = P_\star^{\mathsf{T}} \times_1 (R_{2,\star}^{\mathsf{T}} \times_2 (\cdots \times_{k-1} (R_{k,\star}^{\mathsf{T}} \times_k \mathcal{B}) \cdots)) = \quad ,$$



and $A_{11} \times_1 \mathcal{X}_{11\ldots1} \approx \mathcal{B}_1$, called the core problem has minimal dimensions; see [Hnětynková, P, Žáková, 18].

I. Tensor right-hand side: CP point of view

• Note that the CPs obtained from the tensor problem and its matricized counter-part are not the same

$$\begin{array}{ccc} A \times_1 \mathcal{X} \approx \mathcal{B} & \longleftrightarrow & A\,X \equiv A\,\mathcal{X}^{\{1\}} \approx \mathcal{B}^{\{1\}} \equiv B \\ \downarrow & & \downarrow \\ A_{11} \times_1 \mathcal{X}_{11\cdots 1} \approx \mathcal{B}_1 & \not\longleftrightarrow & A_{11}\,X_{11} \approx B_1 \end{array}$$

i.e., $\mathcal{B}_1^{\{1\}} \neq B_1$, unless the transformation matrix $R_\star \in \mathbb{O}(d)$ $(d = \prod_{j=2}^{k} d_j)$ has a Kronecker product structure w.r.t. the tensor dimensions, i.e.,

$$R_{k,\star} \otimes \cdots \otimes R_{2,\star} = R_\star$$

• The CP transformation in the tensor case is based on the Tucker decomposition of $\mathcal{B}$.

Generalizations: II. Bi-linear problem

Another possibility, originated in the observation that the matrix $A \in \mathbb{R}^{m\times d}$ cannot represent a general linear mapping in $\mathscr{L}(\mathbb{R}^{n\times d}, \mathbb{R}^{m\times d})$:

Vector RHS:	$Ax \approx b,$	$A \in \mathbb{R}^{m\times n},$	$x \in \mathbb{R}^{n},$	$b \in \mathbb{R}^{m}$
	$\downarrow$			
Matrix RHS:	$AX \approx B,$	$A \in \mathbb{R}^{m\times n},$	$X \in \mathbb{R}^{n\times d},$	$B \in \mathbb{R}^{m\times d}$
	$\downarrow$			
Bi-linear MRHS:	$\boxed{A_{\mathfrak{L}} X A_{\mathfrak{R}}^{\mathsf{T}} \approx B}$	$A_{\mathfrak{L}} \in \mathbb{R}^{m\times n}, A_{\mathfrak{R}} \in \mathbb{R}^{c\times d},$	$X \in \mathbb{R}^{n\times d},$	$B \in \mathbb{R}^{m\times c}$

The corresponding **TLS formulation** exists and it has some **applications**; see [Kukush, Makrovsky, Van Huffel, 02, 03].

Existence and uniqueness of TLS solution?

II. Bi-linear problem: CP point of view

Again, it can be shown, there exist four orthogonal matrices

$$(P_{\star}, Q_{\star}, R_{\star}, K_{\star}) \in \mathbb{O}(m) \times \mathbb{O}(n) \times \mathbb{O}(d) \times \mathbb{O}(c),$$

so that

$$\widetilde{A}_{\mathfrak{L}} \widetilde{X} \widetilde{A}_{\mathfrak{R}}^{\mathsf{T}} \equiv (P_{\star}^{\mathsf{T}} A_{\mathfrak{L}} Q_{\star})(Q_{\star}^{\mathsf{T}} X R_{\star})(K_{\star}^{\mathsf{T}} A_{\mathfrak{R}} R_{\star})^{\mathsf{T}}$$

$$= \begin{bmatrix} A_{\mathfrak{L},11} & 0 \\ 0 & A_{\mathfrak{L},22} \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} A_{\mathfrak{R},11} & 0 \\ 0 & A_{\mathfrak{R},22} \end{bmatrix}^{\mathsf{T}} \approx \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$= (P_{\star}^{\mathsf{T}} B K_{\star}) \equiv \widetilde{B},$$

and $A_{\mathfrak{L},11} X_{11} A_{\mathfrak{R},11}^{\mathsf{T}} \approx B_1$, called the core problem has minimal dimensions; see [Hnětynková, P, Žáková, 19].

Properties of and troubles with CPs

In both generalizations, CPs in some way inherit all the important properties. In particular:

- A_{11}, $A_{\mathfrak{L},11}$, $A_{\mathfrak{R},11}$ are of full column rank.
- $\mathcal{B}_1^{\{\ell\}}$ are of full row rank for $\ell = 2, \ldots, k$.
- Matrices† $U_j^\top \mathcal{B}_1^{\{1\}}$ are of full row rank, $j = 1, 2, \ldots$
- Matrices† $U_{\mathfrak{L},j}^\top B_1$ are of full row rank, $j = 1, 2, \ldots$
- Matrices† $B_1 U_{\mathfrak{R},j}$ are of full column rank, $j = 1, 2, \ldots$

†Columns of U_j, $U_{\mathfrak{L},j}$, and $U_{\mathfrak{R},j}$ form an orthomormal bases of left singular vector sub-spaces of A_{11}, $A_{\mathfrak{L},j}$, and $A_{\mathfrak{R},j}$ corresponding to the jth largest singular values of A_{11}, $A_{\mathfrak{L},j}$, and $A_{\mathfrak{R},j}$, respectively.

Outline

Linear approximation problems and the TLS

Matrix right-hand side TLS minimization

Generalizations of the matrix case

Final comments and summary

Further development

- Can be the overall picture further extended to general tensor problems, or higher Kronecker rank (e.g., Sylvester-like) problems, etc.

$$
\begin{array}{ccccccc}
 & & & A\,x \approx b & & & \\
 & & & \downarrow & & & \\
 & & & A\,X \approx B & & & \\
 & & \swarrow & & \searrow & & \\
 & A \times_1 \mathcal{X} \approx \mathcal{B} & & & & A_{\mathfrak{L}}\,X\,A_{\mathfrak{R}}^{\top} \approx B & \\
 \swarrow & & \searrow & & \swarrow & & \searrow \\
 \textbf{?} & & & \mathcal{A} \times_{\dots} \mathcal{X} \approx \mathcal{B} & & & \sum_i A_{\mathfrak{L},i}\,X\,A_{\mathfrak{R},i}^{\top} \approx B
\end{array}
$$

where $\mathcal{A}$, $\mathcal{X}$, and $\mathcal{B}$ are tensors of suitable orders and dimensions?

- The underlying TLS theory (existence, uniqueness of the solution)?
- The computational aspects?

Krylov subspace procedures

The generalized GK of A starting with B gives a subproblem, e.g.,

$$
\left[\; B_1 \,\middle|\, A_{11} \;\right] = \left[\begin{array}{ccc||ccc|cc|c|c}
\gamma_1 & \beta_{12} & \beta_{13} & \alpha_1 & & & & & & \\
 & \gamma_2 & \beta_{23} & \beta_{24} & \alpha_2 & & & & & \\
 & & \gamma_3 & \beta_{34} & \beta_{35} & \alpha_3 & & & & \\
\hline
 & & & \gamma_4 & \beta_{45} & \beta_{46} & \alpha_4 & & & \\
 & & & & & \gamma_5 & \beta_{57} & \alpha_5 & & \\
\hline
 & & & & & & \gamma_6 & \beta_{68} & & \\
 & & & & & & & \gamma_7 & \alpha_6 & \\
\hline
 & & & & & & & & \gamma_8 & \alpha_7 \\
\hline
 & & & & & & & & & \gamma_9
\end{array}\right],
$$

where $\alpha_j > 0$, $\gamma_\ell > 0$, proposed by Åke Björck for computation CP in the matrix right-hand side case.

References & related works

Golub, Van Loan: *An analysis of the total least squares problem*, Numer. Anal., **1980**.
Van Huffel, Vandewalle: *The TLS Problem: Computational aspects and analysis*, SIAM, **1991**.
Paige, Strakoš: *Core problem in linear algebraic systems*, SIMAX, **2006**.
Hnětynková, Strakoš: *Lanczos tridiagonalization and core problems*, LAA, **2007**.
Hnětynková, P, Sima, Strakoš, Van Huffel: *The total least squares problem in $AX \approx B$: A new classification*, SIMAX, **2011**.
Hnětynková, P, Strakoš: *The core problem within a linear approximation problem $AX \approx B$ with multiple right-hand sides*, SIMAX, **2013**
Hnětynková, P, Strakoš: *Band generalization of the Golub–Kahan bidiagonalization, generalized Jacobi matrices, and the core problem*, SIMAX, **2015**
Hnětynková, P, Sima: *Solvability of the core problem with multiple right-hand sides in the TLS sense*, SIMAX, **2016**
Hnětynková, P, Žáková: *TLS formulation and core reduction for problems with structured right-hand sides*, LAA, **2018**.
Kukush, Markovsky, Van Huffel: *On consistent estimators in linear and bilinear multivariate errors-in-variables models*, Proc. of TLS-EIV workshop (Van Huffel, Lemmerling, eds.), **2002**.
Kukush, Markovsky, Van Huffel: *Consistent estimation in the bilinear multivariate errors-in-variables model*, Metrika **2003**.
Hnětynková, P, Žáková: *On TLS formulation and core reduction for data fitting with generalized models*, LAA, **2019**.
Björck: *Bidiagonal Decomposition and Least Squares*, Talk, Canberra, **2005**.
Björck: *A Band-Lanczos Generalization of Bidiagonal Decomp.*, Talk, Stockholm, **2006**.

THANK YOU

FOR YOUR ATTENTION