The Core Problem Story 2004—2024

Martin Plešinger
Dept. of Mathematics, TU Liberec

& Iveta Hnětynková, Jana Žáková
& Zdeněk Strakoš, Diana Maria Sima, Sabine Van Huffel

PANM22, Hejnice, June 23—28, 2024

Outline:

Introduction, recapitulation, & motivation

- Linear approximation problem & total least squares
- The vector right-hand side TLS
- The core problem theory
- The matrix right-hand side TLS
- Could the CP theory help again?
- Key questions and how to approach them

Look outside

- A few concepts from tensor computations
- Tensor right-hand side TLS
- Bi-linear matrix right-hand side TLS
- k-linear tensor right-hand side TLS

Look inside

- (De-)composing of CPs
- Irreducible representation w.r.t. decomposing
- Tuples of matrices — A general framework
- Search for irreducible representation

Introduction, recapitulation, & motivation

Linear approximation problem:

$$\mathcal{A}(\xi) \approx \beta, \qquad \text{where} \qquad \beta \notin \mathrm{Im}(\mathcal{A})$$

$\xi \in \mathscr{V}, \quad \beta \in \mathscr{W}, \quad$ and $\quad \mathcal{A} \in \mathscr{L}(\mathscr{V}, \mathscr{W})$

(finite-dimensional; over the same field — $\mathbb{R}$ or $\mathbb{C}$)

Least squares (LS):

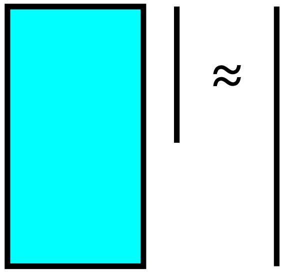
$$\min_{\gamma \in \mathscr{W}} \|\gamma\|_{\mathscr{W}} \quad \text{s.t.} \quad (\beta + \gamma) \in \mathrm{Im}(\mathcal{A})$$

Total least squares (TLS):

$$\min_{\substack{\gamma \in \mathscr{W} \\ \mathcal{E} \in \mathscr{S} \subseteq \mathscr{L}(\mathscr{V}, \mathscr{W})}} \big\| \big[\, \|\gamma\|_{\mathscr{W}}, \|\mathcal{E}\|_{\mathscr{L}\ldots} \,\big] \big\|_{\mathbb{R}^2} \quad \text{s.t.} \quad (\beta + \gamma) \in \mathrm{Im}(\mathcal{A} + \mathcal{E})$$

$$\iff \quad \exists\, \xi_{\mathrm{TLS}} : \quad (\mathcal{A} + \mathcal{E})(\xi_{\mathrm{TLS}}) = (\beta + \gamma)$$

Vector (or single) right-hand side (RHS) TLS:



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

$$\min_{\substack{g \in \mathbb{R}^{m} \\ E \in \mathbb{R}^{m \times n}}} \quad \|[g, E]\|_{\mathsf{F}} \quad \text{s.t.} \quad (b + g) \in \mathscr{R}(A + E)$$

How to get the TLS solution?

$$\begin{aligned} Ax &\approx b \\ -b + Ax &\approx 0_m \\ [b, A]\begin{bmatrix} -1 \\ x \end{bmatrix} &\approx 0_m \end{aligned} \qquad \Bigg| \qquad \begin{aligned} (A + E)x_{\mathsf{TLS}} &= (b + g) \\ -(b + g) + (A + E)x_{\mathsf{TLS}} &= 0_m \\ \big([b, A] + [g, E]\big)\begin{bmatrix} -1 \\ x_{\mathsf{TLS}} \end{bmatrix} &= 0_m \end{aligned}$$

Singular value decomposition (SVD) of a matrix:

Let $M \in \mathbb{R}^{s \times t}$ be a matrix of rank $r = \operatorname{rank}(M)$. Then

$$M = U\Sigma V^\top = \sum_{j=1}^{r} u_j \sigma_j v_j^\top$$

where

$$U = [u_1, u_2, \ldots, u_s] \in \mathbb{O}_s \quad \text{i.e.} \quad U \in \mathbb{R}^{s \times s} \text{ and } U^\top = U^{-1},$$

$$V = [v_1, v_2, \ldots, v_t] \in \mathbb{O}_t \quad \text{i.e.} \quad V \in \mathbb{R}^{t \times t} \text{ and } V^\top = V^{-1},$$

$$\Sigma = \begin{bmatrix} \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r) & 0_{r,t-r} \\ 0_{s-r,r} & 0_{s-r,t-r} \end{bmatrix} \in \mathbb{R}^{s \times t}$$

and

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Eckart—Young—Mirsky (EYM) theorem:

Let $M \in \mathbb{R}^{s\times t}$, $\mathrm{rank}(M) = r$, i.e. $M = \sum_{j=1}^{r} u_j \sigma_j v_j^\top$. Then

$$\|M - N_{\star}\|_{\mathsf{F}} = \min_{\substack{N \in \mathbb{R}^{s\times t} \\ \mathrm{rank}(N) < r}} \|M - N\|_{\mathsf{F}} = \sigma_r$$

where

$$N_{\star} = \sum_{j=1}^{r-1} u_j \sigma_j v_j^\top .$$

$N_{\star}$ is the optimal low-rank approximation of M. Since the smallest singular value is q-tuple, where $q \geq 1$, i.e.,

$$\sigma_{r-q} > \sigma_{r-q+1} = \cdots = \sigma_r > 0,$$

the minimizer $N_{\star}$ may not be unique.

So, how to get the TLS solution?

Let for simplicity assume $\text{rank}(A) = n$. Then $\text{rank}([b, A]) = n + 1$ and

$$[b, A] = \sum_{j=1}^{n+1} u_j \sigma_j v_j^{\mathsf{T}},$$

EYM theorem gives us the optimal correction

$$[g, E] = -u_{n+1}\sigma_{n+1}v_{n+1}^{\mathsf{T}} \qquad \text{and} \qquad [b, A] + [g, E] = \sum_{j=1}^{n} u_j \sigma_j v_j^{\mathsf{T}}.$$

Moreover

$$\Big([b, A] + [g, E]\Big) v_{n+1} = 0_m.$$

Now it remains to scale v_{n+1} to

$$v_{n+1} \cdot \left(\frac{-1}{v_{1,n+1}}\right) = \begin{bmatrix} -1 \\ x_{\mathsf{TLS}} \end{bmatrix} \ldots \qquad \text{Wait!}$$

Since the smallest singular value is $(q+1)$-tuple, we can play the same game with other right singular vectors. If the first row of

$$[v_{n-q+1}, \dots, v_{n+1}]$$

is nonzero, we have the TLS solution (for $q \geq 1$ non-unique).

And what if it is zero? Then our construction fails (and in fact the TLS solution does not exist). Technically, we can employ larger and larger singular values and play similar game, yielding so-called non-generic solution (see Van Huffel), but the optimality is lost (it can also be reformulated as a constrained TLS).

But what it means? Why there is no solution (on the contrary to the standard LS)?

Example #1 (Golub—Van Loan):

$$Ax = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \approx \begin{bmatrix} 1 \\ 1 \end{bmatrix} = b.$$

Let E and g be arbitrary corrections that make the system compatible. Denote $\|[g, E]\|_{\mathsf{F}} = \varepsilon$. Then

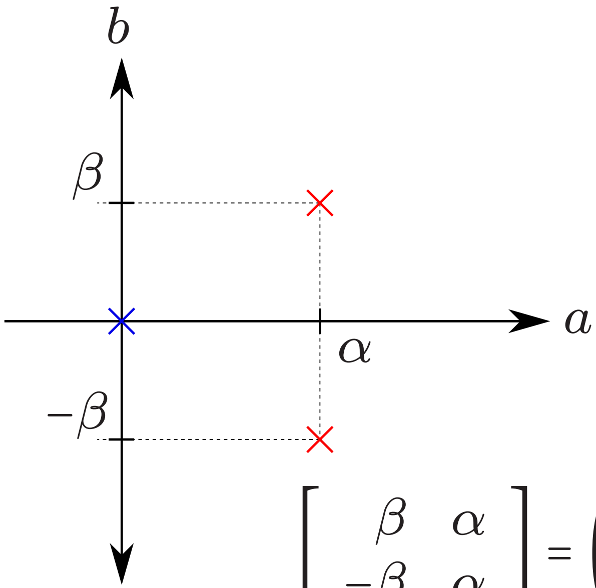
$$E' = \begin{bmatrix} 0 & 0 \\ 0 & \varepsilon/2 \end{bmatrix}, \qquad g' = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

make the system compatible as well but with $\|[g, E]\|_{\mathsf{F}} = \varepsilon/2$.

Consequently, there is no optimal correction. Moreover the norm of the solution of $(A + E')x = (b + g')$ grows to infinity as $1/\varepsilon$.

(Note, it does not satisfy our extra assumption — A is not of full column rank.)

Example #2 (geometric): Given two measurements (red points × in (a,b)-plane), goal is to find a linear model of the data — by TLS:



$$Ax = \begin{bmatrix} \alpha \\ \alpha \end{bmatrix} [\xi] \approx \begin{bmatrix} \beta \\ -\beta \end{bmatrix} = b$$

$$[b, A] = \begin{bmatrix} \beta & \alpha \\ -\beta & \alpha \end{bmatrix}$$ is of full column rank for $\beta \neq 0$, $\alpha \neq 0$

with the SVD

$$\begin{bmatrix} \beta & \alpha \\ -\beta & \alpha \end{bmatrix} = \left(\frac{\sqrt{2}}{2} \begin{bmatrix} \mathrm{sgn}(\beta) & \mathrm{sgn}(\alpha) \\ -\mathrm{sgn}(\beta) & \mathrm{sgn}(\alpha) \end{bmatrix} \right) \begin{bmatrix} \sqrt{2}|\beta| & 0 \\ 0 & \sqrt{2}|\alpha| \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{\mathsf{T}}$$

If $|\beta| < |\alpha|$, then there is unique TLS solution $\beta = 0 \cdot \alpha$.
If $|\beta| = |\alpha|$, then there are infinitely many TLS solutions.
If $|\beta| > |\alpha|$, then our construction gives (and there is) no TLS solution.

Example #3 (Paige—Strakoš): Let

$$Ax = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \approx \begin{bmatrix} b_1 \\ 0 \end{bmatrix} = b$$

then

$$[b, A] \begin{bmatrix} -1 \\ x \end{bmatrix} = \begin{bmatrix} b_1 & A_{11} & 0 \\ 0 & 0 & A_{22} \end{bmatrix} \begin{bmatrix} -1 \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} [b_1, A_{11}] \begin{bmatrix} -1 \\ x_1 \end{bmatrix} \\ A_{22} x_2 \end{bmatrix} \approx 0,$$

where the red sub-problem needs to be solved, the blue one has obvious (minimum norm) solution

$$[b_1, A_{11}] \begin{bmatrix} -1 \\ x_1 \end{bmatrix} \approx 0, \qquad A_{22} x_2 \approx 0, \qquad x_2 = 0.$$

Assuming A_{11} (and thus also $[b_1, A_{11}]$) is of full column rank, we have g_1, E_{11} with $\|[g_1, E_{11}]\|_F = \sigma_{\min}([b_1, A_{11}])$, and possibly also $x_{1,\mathrm{TLS}}$.

Example #3 (Paige—Strakoš): Thus

$$\left(\begin{bmatrix} b_1 & A_{11} & 0 \\ 0 & 0 & A_{22} \end{bmatrix} + \begin{bmatrix} g_1 & E_{11} & 0 \\ 0 & 0 & 0 \end{bmatrix}\right)\begin{bmatrix} -1 \\ x_{1,\mathrm{TLS}} \\ 0 \end{bmatrix} = 0.$$

If $\sigma_{\min}([b_1, A_{11}]) > \sigma_{\min}(A_{22})$: Take corresponding singular vecs u, v, take an arbitrary vector z, $r = b_1 - A_{11}z$, and number $\varepsilon > 0$, then

$$\left(\begin{bmatrix} b_1 & A_{11} & 0 \\ 0 & 0 & A_{22} \end{bmatrix} + \begin{bmatrix} 0 & 0 & r\varepsilon v^T \\ 0 & 0 & -u\sigma_{\min}v^T \end{bmatrix}\right)\begin{bmatrix} -1 \\ z \\ v\varepsilon^{-1} \end{bmatrix} = 0,$$

with the norm of the correction $(\|r\|_2^2\,\varepsilon^2 + \sigma_{\min}{}^2)^{\frac{1}{2}}$.

With sufficiently small ε we get smaller correction, larger solution, moreover with arbitrarily chosen component.

The core problem (CP) theory (Paige—Strakoš):

We have $Ax \approx b$, $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ $(b \notin \mathscr{R}(A))$, and $[b, A] = U\Sigma V^\top$. For any

$$(P, Q) \in \mathbb{O}_m \times \mathbb{O}_n$$

we define a transformed problem $\widetilde{A}\widetilde{x} = (P^\top AQ)(Q^\top x) \approx (P^\top b) = \widetilde{b}$,

$$[\widetilde{b}, \widetilde{A}]\begin{bmatrix} -1 \\ \widetilde{x} \end{bmatrix} = \left(P^\top[b, A]\begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}\right)\begin{bmatrix} -1 \\ Q^\top x \end{bmatrix} \approx 0_m,$$

$$[\widetilde{b}, \widetilde{A}] = (P^\top U)\,\Sigma\left(\begin{bmatrix} 1 & 0 \\ 0 & Q^\top \end{bmatrix} V\right)^\top .$$

Corrections E, g and $\widetilde{E} = P^\top EQ$, $\widetilde{g} = P^\top g$ have the same norm

$$\|[\widetilde{g}, \widetilde{E}]\|_F = \left\| P^\top[g, E]\begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}\right\|_F = \|[g, E]\|_F.$$

The CP theorem:

$$\forall\ (A, b) \in \mathbb{R}^{m\times n} \times \mathbb{R}^{m} \quad \exists\ (P_{\star}, Q_{\star}) \in \mathbb{O}_m \times \mathbb{O}_n$$

such that

$$[\widetilde{b}, \widetilde{A}] = P_{\star}^{\top}[b, A]\begin{bmatrix} 1 & 0 \\ 0 & Q_{\star} \end{bmatrix} = \begin{bmatrix} b_1 & A_{11} & 0 \\ 0 & 0 & A_{22} \end{bmatrix},$$

where

(CP1) $A_{11} \in \mathbb{R}^{\bar{m}\times\bar{n}}$ is of *full column rank* $\bar{n}$.

(CP2) $b_1 \in \mathbb{R}^{\bar{m}}$ is *nonzero*.

(CP3) $u_i^{\top} b_1$ are *nonzero*; u_i are left singular vecs of A_{11}, $i = 1, \ldots \bar{m}$.

$\Longrightarrow$ (CP4) $[b_1, A_{11}]$ is of *full row rank* $\bar{m}$.

(CP5) $e_1^{\top} v_\ell$ are *nonzero*; v_ℓ are right sing vecs of $[b_1, A_{11}]$, $\ell = 1, \ldots \bar{n}{+}1$.

(CP6) Singular values of A_{11} are *simple*.

(CP7) Singular values of $[b_1, A_{11}]$ are *simple*.

(CP8) $A_{11} x_1 \approx b_1$, the *core problem*, always has *unique TLS solution*.

Note: How the CP looks like and how to get it?

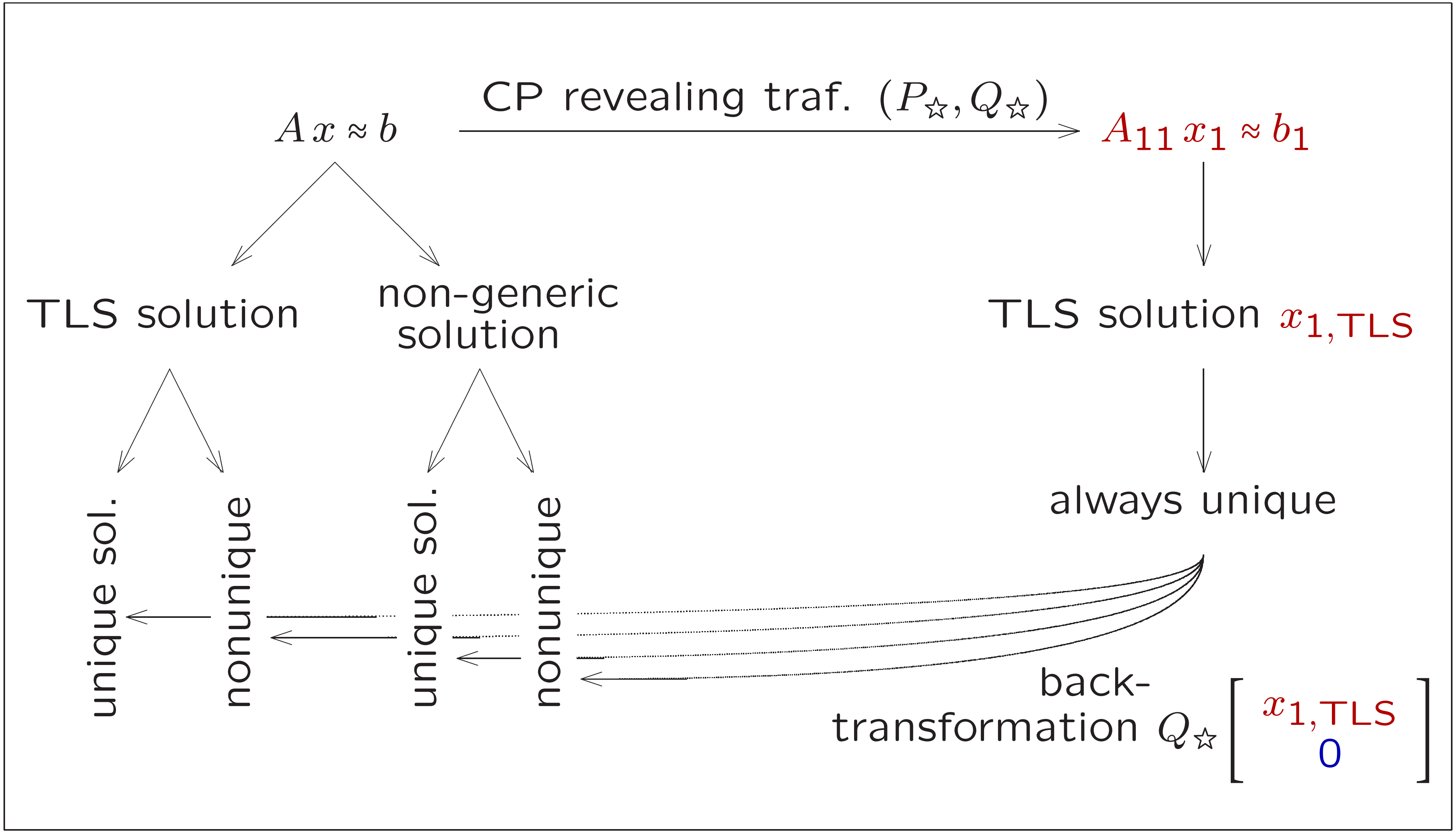
By using the SVD of A:

$$[b_1, A_{11}] = \left[\begin{array}{c|cccc} \mu_1 & \sigma_1 & & & \\ \mu_2 & & \sigma_2 & & \\ \vdots & & & \ddots & \\ \mu_{\bar{n}} & & & & \sigma_{\bar{n}} \\ \mu_{\bar{m}} & 0 & 0 & 0 & 0 \end{array}\right], \qquad \begin{array}{l} \mu_i \neq 0, \quad i = 1, \dots, \bar{m}, \\ \sigma_1 > \sigma_2 > \cdots > \sigma_{\bar{n}} > 0. \end{array}$$

By using the Golub—Kahan bidiagonalization:

$$[b_1, A_{11}] = \left[\begin{array}{c|cccc} \gamma_1 & \delta_1 & & & \\ & \gamma_2 & \delta_2 & & \\ & & \ddots & \ddots & \\ & & & \gamma_{\bar{n}} & \delta_{\bar{n}} \\ & & & & \gamma_{\bar{m}} \end{array}\right], \qquad \begin{array}{l} \gamma_i > 0, \quad i = 1, \dots, \bar{m}, \\ \delta_j > 0, \quad j = 1, \dots, \bar{n}. \end{array}$$

Long story short:



Outline:

Introduction, recapitulation, & motivation

- Linear approximation problem & total least squares
- The vector right-hand side TLS
- The core problem theory
- **The matrix right-hand side TLS**
- Could the CP theory help again?
- Key questions and how to approach them

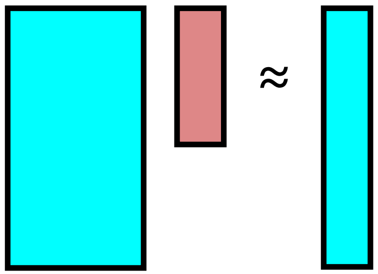
Look outside

- A few concepts from tensor computations
- Tensor right-hand side TLS
- Bi-linear matrix right-hand side TLS
- k-linear tensor right-hand side TLS

Look inside

- (De-)composing of CPs
- Irreducible representation w.r.t. decomposing
- Tuples of matrices — A general framework
- Search for irreducible representation

Matrix (or multiple) right-hand side TLS:



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

$$\min_{\substack{G \in \mathbb{R}^{m \times d} \\ E \in \mathbb{R}^{m \times n}}} \|[G, E]\|_{\mathsf{F}} \quad \text{s.t.} \quad \mathscr{R}(B + G) \in \mathscr{R}(A + E)$$

How to get the TLS solution?

$$\begin{aligned} AX &\approx B \\ -B + AX &\approx 0_{m,d} \\ [B, A]\begin{bmatrix} -I_d \\ X \end{bmatrix} &\approx 0_{m,d} \end{aligned}$$

$$\begin{aligned} (A + E)X_{\mathsf{TLS}} &= (B + G) \\ -(B + G) + (A + E)X_{\mathsf{TLS}} &= 0_{m,d} \\ \big([B, A] + [G, E]\big)\begin{bmatrix} -I_d \\ X_{\mathsf{TLS}} \end{bmatrix} &= 0_{m,d} \end{aligned}$$

In the context of the general setting

$$\mathcal{A}(X) \approx B, \qquad \mathcal{A} : \mathbb{R}^{n\times d} \longrightarrow \mathbb{R}^{m\times d},$$

is mapping from (nd)- into (md)-dimensional space.

After reshaping the problem

$$AX \approx B \qquad \Longleftrightarrow \qquad (I_d \otimes A)\text{vec}(X) \approx \text{vec}(B),$$

we see that the correction $\mathcal{E}$ is sought in proper subset — *subspace*

$$\mathscr{S} = \left\{ (I_d \otimes E) \;:\; E \in \mathbb{R}^{m\times n} \right\} \subsetneq \mathbb{R}^{md\times nd} \cong \mathscr{L}\left(\mathbb{R}^{n\times d}, \mathbb{R}^{m\times d}\right).$$

This is the origin of several new substantial difficulties.

Here $(I \otimes A) = \text{diag}(A, \ldots, A)$ is the Kronecker product,
and $\text{vec}([x_1, \ldots, x_d]) = [x_1^{\mathsf{T}}, \ldots, x_d^{\mathsf{T}}]^{\mathsf{T}}$.

Analysis of the solvability:

We again for simplicity assume $\text{rank}(A) = n$, $\text{rank}(B) = d$, and at least one $b_j \notin \mathscr{R}(A)$. Then $\text{rank}([B, A]) \geq n + 1$ and

$$[B, A] = U\Sigma V^{\mathsf{T}},$$

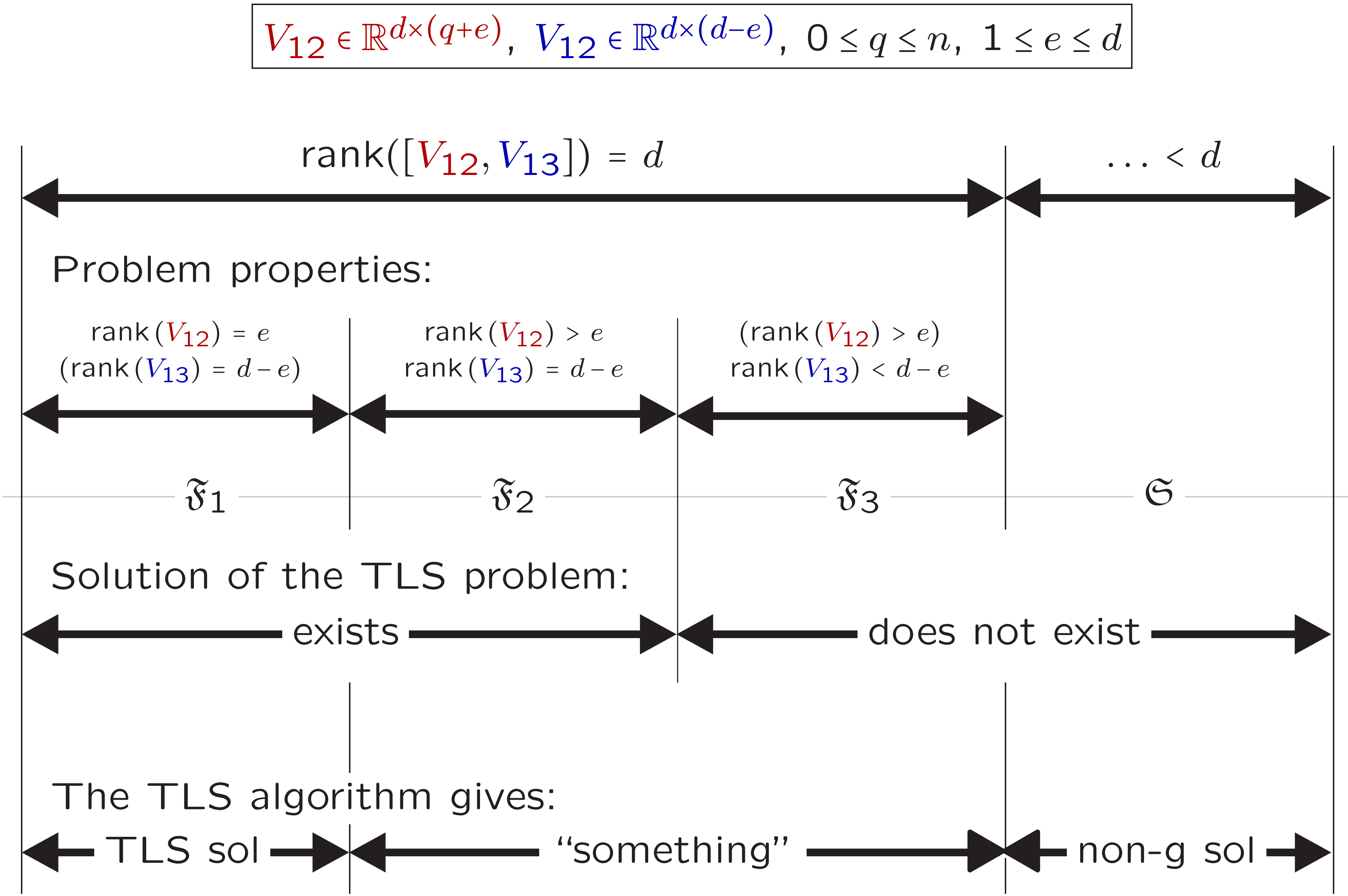
where for some q $(0 \leq q \leq n)$ and e $(1 \leq e \leq d)$

$$\sigma_{n-q} > \sigma_{n-q+1} = \cdots = \sigma_{n+1} = \cdots = \sigma_{n+e} > \sigma_{n+e+1}$$

and

$$V = \begin{array}{c} \left[\begin{array}{ccc} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \end{array}\right] \begin{array}{l} \}d \\ \}n \end{array} \\ \begin{array}{ccc} \underbrace{\qquad\quad}_{n-q} & \underbrace{\qquad\quad}_{q+e} & \underbrace{\qquad\quad}_{d-e} \end{array} \end{array} .$$

The goal is $\left[\begin{smallmatrix} V_{12} & V_{13} \\ V_{22} & V_{23} \end{smallmatrix}\right]$ transform to $\left[\begin{smallmatrix} -I_d \\ X_{\mathsf{TLS}} \end{smallmatrix}\right]$.



$V_{12} \in \mathbb{R}^{d\times(q+e)}$, $V_{12} \in \mathbb{R}^{d\times(d-e)}$, $0 \le q \le n$, $1 \le e \le d$
rank$([V_{12}, V_{13}]) = d$
$\ldots < d$
Problem properties:
rank$(V_{12}) = e$
(rank$(V_{13}) = d - e$)
rank$(V_{12}) > e$
rank$(V_{13}) = d - e$
(rank$(V_{12}) > e$)
rank$(V_{13}) < d - e$
$\mathfrak{F}_1$
$\mathfrak{F}_2$
$\mathfrak{F}_3$
$\mathfrak{S}$
Solution of the TLS problem:
exists
does not exist
The TLS algorithm gives:
TLS sol
“something”
non-g sol

Could the CP theory help again?

Similarly as before, the problem is orthogonally invariant and

$$\forall\,(A,B)\in\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times d}\qquad \exists\,(P_{\star},Q_{\star},R_{\star})\in\mathbb{O}_m\times\mathbb{O}_n\times\mathbb{O}_d$$

such that

$$P_{\star}^{\top}[B,A]\begin{bmatrix} R_{\star} & 0\\ 0 & Q_{\star}\end{bmatrix}=\left[\begin{array}{cc|cc} B_1 & 0 & A_{11} & 0\\ 0 & 0 & 0 & A_{22}\end{array}\right],$$

or, equivalently,

$$(P_{\star}^{\top}AQ_{\star})(Q_{\star}^{\top}XR_{\star})=\begin{bmatrix} A_{11} & 0\\ 0 & A_{22}\end{bmatrix}\begin{bmatrix} X_{11} & X_{12}\\ X_{21} & X_{22}\end{bmatrix}\approx\begin{bmatrix} B_1 & 0\\ 0 & 0\end{bmatrix}=(P_{\star}^{\top}BR_{\star}),$$

so the original problem splits into four sub-problems

$$A_{11}X_{11}\approx B_1,\qquad A_{11}X_{12}\approx 0,\quad A_{22}X_{21}\approx 0,\quad A_{22}X_{22}\approx 0,$$

where (...)

where the first one $A_{11}X_{11} \approx B_1$, the *core problem* satisfies
(CP1) $A_{11} \in \mathbb{R}^{\bar{m}\times\bar{n}}$ is of *full column rank* $\bar{n}$.
(CP2) $B_1 \in \mathbb{R}^{\bar{m}\times\bar{d}}$ is of *full column rank* $\bar{d}$.
(CP3) $U_i^\top B_1 \in \mathbb{R}^{\bar{\mu}_i\times\bar{d}}$ are of *full row rank* $\bar{\mu}_i$; U_i are bases of left singular vec subspaces of A_{11}, $i = 1, \ldots \bar{\xi}, \bar{\xi}+1$.

$\Longrightarrow$ (CP4), (CP5), (CP6), (CP7) can be reasonably generalized.

In fact (CP8) can also be geralized, but the result are not very helpful: $A_{11}X_{11} \approx B_1$ can still belong into any of the four classes

$$\mathfrak{F}_1, \quad \mathfrak{F}_2, \quad \mathfrak{F}_3, \quad \mathfrak{S},$$

there are explicitely examples, i.e., it may not have a TLS solution.
Hovever, from (CP5) it follows that:

If matrix RHS CP belongs to $\mathfrak{F}_1$, then it has unique TLS solution.

Note: How the CP looks like and how to get it?

By using a rank-revealing decomposition of B together with *either* the SVD of A *or* the block/band generalization of the Golub—Kahan bidiagonalization, e.g.:

$$
[B_1, A_{11}] = \left[\begin{array}{ccc|ccccccc}
\gamma_{1,1} & \kappa_{1,2} & \kappa_{1,3} & \delta_{1,4} & & & & & & \\
 & \gamma_{2,2} & \kappa_{2,3} & \kappa_{2,4} & \delta_{2,5} & & & & & \\
 & & \gamma_{3,3} & \kappa_{3,4} & \kappa_{3,5} & \delta_{3,6} & & & & \\
 & & & \gamma_{4,4} & \kappa_{4,5} & \kappa_{4,6} & \delta_{4,7} & & & \\
 & & & & 0 & \gamma_{5,6} & \kappa_{5,7} & \delta_{5,8} & & \\
 & & & & & & \gamma_{6,7} & \kappa_{6,8} & 0 & \\
 & & & & & & & \gamma_{7,8} & \delta_{7,9} & \\
 & & & & & & & & \gamma_{8,9} & \delta_{8,10}
\end{array}\right],
$$

where $\quad \gamma_{i,\pi(i)} > 0,\ i = 1, \ldots, \bar{m}, \quad \delta_{\psi(j),\bar{d}+j} > 0,\ j = 1, \ldots, \bar{n}.$

Key questions and how to approach them:

We would like to understand:

Why the matrix RHS core problem does not have the TLS solution...
What is "wrong" with the data...
How to make more obvious (how to visualize) this quality...

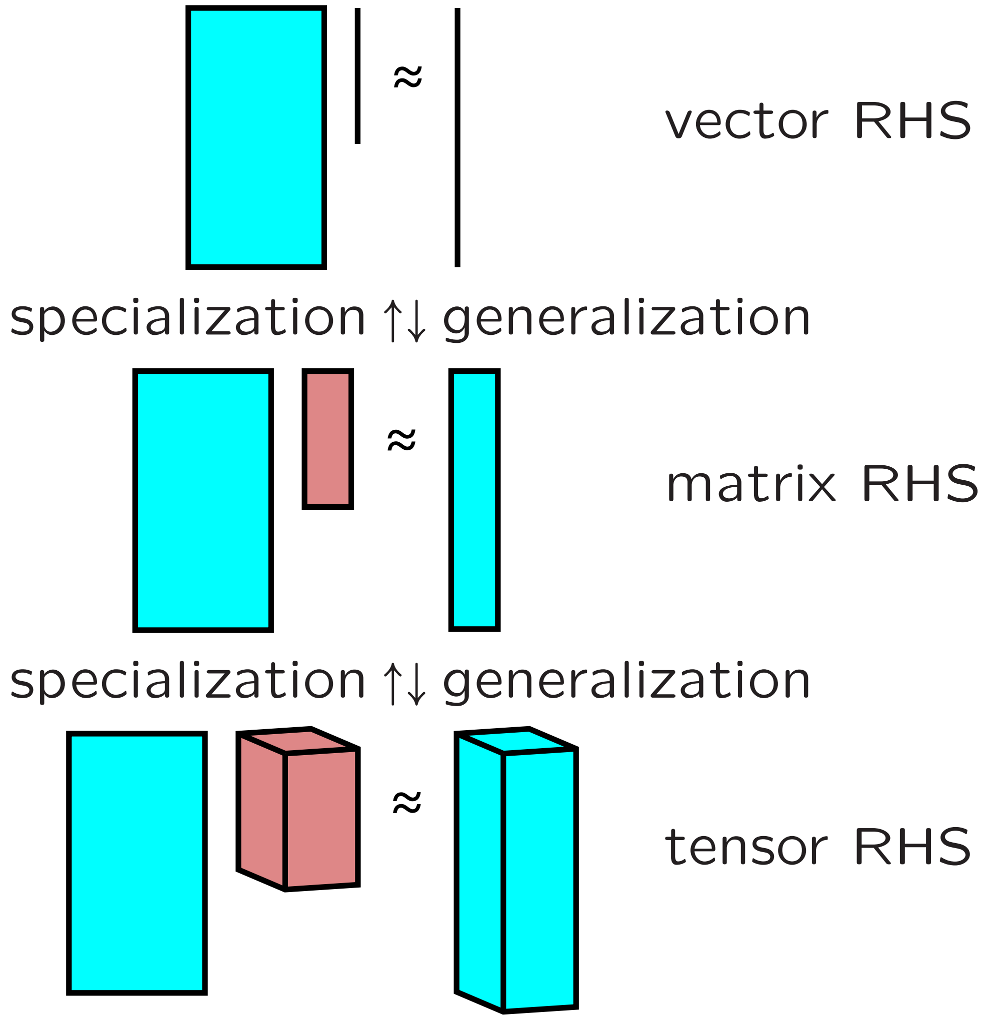
We try to follow two possible paths:

'Look outside' — try to generalize even more...
'Look inside' — explore the internal structure of core problems...

Look outside

Generalize even more...

There is one very natural way of generalization:



A few concepts from tensor computations (Kolda—Bader):

By *tensor* of order k we understand a k-way array of N numbers

$$\mathcal{T} = (t_{j_1,j_2,\ldots,j_k}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k}, \qquad N = \prod_{\lambda=1}^{k} n_\lambda$$

Matrix-tensor product (MT) in ℓth mode ($1 \le \ell \le k$)

$$Q \in \mathbb{R}^{m \times n_\ell}, \qquad \mathcal{S} = Q \times_\ell \mathcal{T} \in \mathbb{R}^{n_1 \times \cdots \times n_{\ell-1} \times m \times n_{\ell+1} \times \cdots \times n_k}$$

$$s_{j_1,\ldots,j_{\ell-1},i,j_{\ell+1},\ldots,j_k} = \sum_{j_\ell=1}^{n_\ell} q_{i,j_\ell} \cdot t_{j_1,\ldots,j_{\ell-1},j_\ell,j_{\ell+1},\ldots,j_k}$$

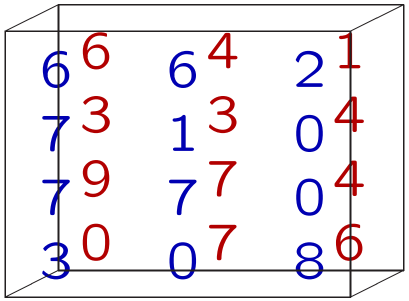
(Einstein notation does not help us too much.)

Short exercise (all 8 products of two square matrices):

$$\begin{array}{llll} AB = A \times_1 B, & AB^\top = A \times_2 B, & A^\top B = A^\top \times_1 B, & A^\top B^\top = A^\top \times_2 B, \\ BA = B \times_1 A, & BA^\top = B \times_2 A, & B^\top A = B^\top \times_1 A, & B^\top A^\top = B^\top \times_2 A, \end{array}$$

A few concepts from tensor computations (part II):

ℓ-mode *matricization* of a tensor $\qquad \mathcal{T} \longmapsto \mathcal{T}^{\{\ell\}} \in \mathbb{R}^{n_\ell \times (N/n_\ell)}$,

$\mathcal{T} =$ 

,

$$\mathcal{T}^{\{1\}} = \begin{array}{c} \begin{bmatrix} 6 & 6 & 2 & 6 & 4 & 1 \\ 7 & 1 & 0 & 3 & 3 & 4 \\ 7 & 7 & 0 & 9 & 7 & 4 \\ 3 & 0 & 8 & 0 & 7 & 6 \end{bmatrix} \\ \begin{array}{cccccc} (*,1,1) & (*,2,1) & (*,3,1) & (*,1,2) & (*,2,2) & (*,3,2) \end{array} \end{array},$$

$$\mathcal{T}^{\{2\}} = \begin{array}{c} \begin{bmatrix} 6 & 7 & 7 & 3 & 6 & 3 & 9 & 0 \\ 6 & 1 & 7 & 0 & 4 & 3 & 7 & 7 \\ 2 & 0 & 0 & 8 & 1 & 4 & 4 & 6 \end{bmatrix} \\ \begin{array}{cccccccc} (1,*,1) & (2,*,1) & (3,*,1) & (4,*,1) & (1,*,2) & (2,*,2) & (3,*,2) & (4,*,2) \end{array} \end{array},$$

$$\mathcal{T}^{\{3\}} = \begin{array}{c} \begin{bmatrix} 6 & 7 & 7 & 3 & 6 & 1 & 7 & 0 & 2 & 0 & 0 & 8 \\ 6 & 3 & 9 & 0 & 4 & 3 & 7 & 7 & 1 & 4 & 4 & 6 \end{bmatrix} \\ \begin{array}{cccccccccccc} (1,1,*) & (2,1,*) & (3,1,*) & (4,1,*) & (1,2,*) & (2,2,*) & (3,2,*) & (4,2,*) & (1,3,*) & (2,3,*) & (3,3,*) & (4,3,*) \end{array} \end{array},$$

with *fibres* in the *inverse lexicographical ordering* w.r.t. *multi-indices*

A few concepts from tensor computations (part III):

Given $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k}$, and $Q_\ell \in \mathbb{R}^{m_\ell \times n_\ell}$,

$$\left(Q_\ell \times_\ell \mathcal{T} \right)^{\{\ell\}} = Q_\ell \mathcal{T}^{\{\ell\}}.$$

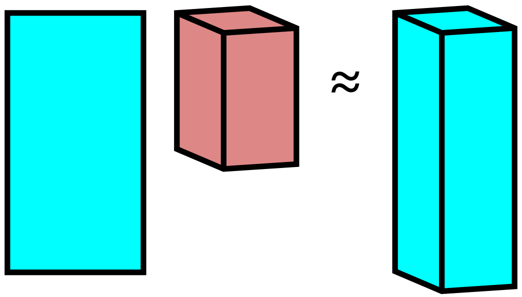
Associativity of matrix multiplication translates into MT products as

$$\ell_1 \neq \ell_2 \qquad \Longrightarrow \qquad Q_{\ell_1} \times_{\ell_1} \left(Q_{\ell_2} \times_{\ell_2} \mathcal{T} \right) = Q_{\ell_2} \times_{\ell_2} \left(Q_{\ell_1} \times_{\ell_1} \mathcal{T} \right)$$

That brings us to general linear transformation of a tensor

$$\left(Q_1, Q_2, \ldots, Q_k \middle| \mathcal{T} \right) = Q_1 \times_1 \left(Q_2 \times_2 \cdots \left(Q_k \times_k \mathcal{T} \right) \cdots \right)$$

Tensor right-hand side TLS:



$$A \times_1 \mathcal{X} \approx \mathcal{B}, \qquad A \in \mathbb{R}^{m\times n},\ \mathcal{X} \in \mathbb{R}^{n\times d_2\times\cdots\times d_k},\ \mathcal{B} \in \mathbb{R}^{m\times d_2\times\cdots\times d_k}$$

$$\min_{\substack{\mathcal{G} \in \mathbb{R}^{m\times d_2\times\cdots\times d_k} \\ E \in \mathbb{R}^{m\times n}}} \left(\|\mathcal{G}\|^2 + \|E\|_F^2\right)^{\frac{1}{2}} \text{ s.t. } \mathscr{R}((\mathcal{B}+\mathcal{G})^{\{1\}}) \in \mathscr{R}(A+E)$$

where $\|\mathcal{G}\| = \|\mathcal{G}^{\{1\}}\|_F$. Clearly

$$A \times_1 \mathcal{X} \approx \mathcal{B} \qquad \Longleftrightarrow \qquad A\mathcal{X}^{\{1\}} \approx \mathcal{B}^{\{1\}}$$

i.e. all the results about TLS solvability (existence and uniqueness of TLS solution) in the matrix case, are directly applicable here. Schematically:

$$\Big(\, \mathit{tensor}\ \mathsf{TLS}(A,\mathcal{B})\, \Big)^{\{1\}} = \Big(\, \mathit{matrix}\ \mathsf{TLS}(A,\mathcal{B}^{\{1\}})\, \Big).$$

Core problem within tensor RSH problem: Similarly as always

$$\forall\ (A, \mathcal{B}) \in \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times d_2\times\cdots\times d_k}$$
$$\exists\ (P_\star, Q_\star, R_{2,\star}, \ldots, R_{k,\star}) \in \mathbb{O}_m \times \mathbb{O}_n \times \mathbb{O}_{d_2} \times \cdots \times \mathbb{O}_{d_k}$$

($R_{j,\star}$ are from the Tucker decomposition) such that

$$P_\star^\top A Q_\star = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}, \qquad \left(P_\star^\top, R_{2,\star}^\top, \ldots, R_{k,\star}^\top \,\middle|\, \mathcal{B} \right) = \mathrm{diag}_k(\mathcal{B}_1, 0)$$

so the original problem splits into 2^k sub-problems

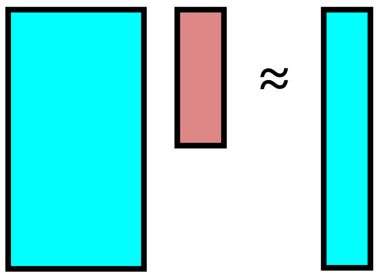
$$A_{11} \times_1 \mathcal{X}_{11\ldots1} \approx \mathcal{B}_1, \quad A_{11} X_{1 i_2 \ldots i_k} \approx 0, \quad A_{22} X_{21\ldots1} \approx 0, \quad A_{22} X_{2 i_2 \ldots i_k} \approx 0,$$

where the core problem has a lot of nice properties (CP1)—... But

$$\Big(\mathit{tensor}\ \mathrm{CP}(A, \mathcal{B}) \Big)^{\{1\}} \neq \Big(\mathit{matrix}\ \mathrm{CP}(A, \mathcal{B}^{\{1\}}) \Big).$$

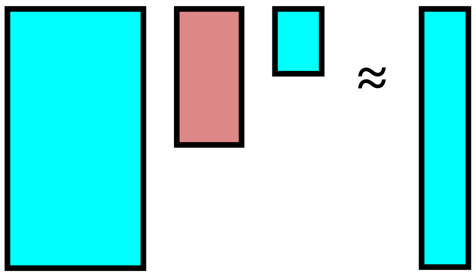
Tensor structure has to be preserved.

There is another way of generalizalization:



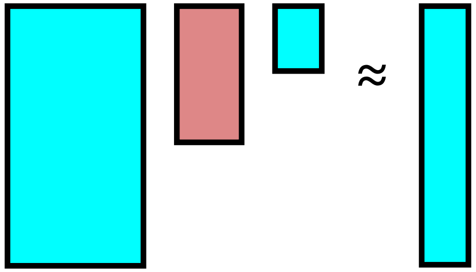
matrix RHS

specialization ↑↓ generalization



bi-linear matrix RHS

Bi-linear matrix right-hand side TLS:



$$A_{\mathsf{L}} X A_{\mathsf{R}}^{\mathsf{T}} \approx B, \qquad A_{\mathsf{L}} \in \mathbb{R}^{m\times n},\ X \in \mathbb{R}^{n\times s},\ A_{\mathsf{R}} \in \mathbb{R}^{d\times s},\ B \in \mathbb{R}^{m\times d}$$

$$\min_{\substack{G \in \mathbb{R}^{m\times d}\\ E_{\mathsf{L}} \in \mathbb{R}^{m\times n}\\ E_{\mathsf{R}} \in \mathbb{R}^{d\times s}}} \left\| \begin{bmatrix} G & E_{\mathsf{L}} \\ E_{\mathsf{R}}^{\mathsf{T}} & 0 \end{bmatrix} \right\|_{\mathsf{F}} \quad \text{s.t.} \quad \begin{array}{l} \mathscr{R}(B+G) \in \mathscr{R}(A_{\mathsf{L}} + E_{\mathsf{L}}) \\ \mathscr{R}((B+G)^{\mathsf{T}}) \in \mathscr{R}(A_{\mathsf{R}} + E_{\mathsf{R}}) \end{array}$$

- For applications and first steps of analysis (existence and uniqueness of TLS solution) see Kukush—Markovsky—Van Huffel.
- In the context of general setting $\mathcal{A}(X) \approx B$, $\mathcal{A} : \mathbb{R}^{n\times s} \longrightarrow \mathbb{R}^{m\times d}$, the correction $\mathcal{E}$ is sought in proper subset — *submanifold*

$$\mathscr{S} = \left\{ (E_{\mathsf{R}} \otimes E_{\mathsf{L}}) \ :\ E_{\mathsf{L}} \in \mathbb{R}^{m\times n},\ E_{\mathsf{R}} \in \mathbb{R}^{d\times s} \right\} \subsetneq \mathbb{R}^{md\times ns} \cong \mathscr{L}\left(\mathbb{R}^{n\times s}, \mathbb{R}^{m\times d}\right).$$

- Enforcing $E_{\mathsf{R}} = 0$ or $E_{\mathsf{L}} = 0$ yields a *joined LS-TLS minimization*.

Core problem within bi-linear problem: Similarly as always

$\forall\ (A_{\mathsf{L}}, A_{\mathsf{R}}, B) \in \mathbb{R}^{m\times n} \times \mathbb{R}^{d\times s} \times \mathbb{R}^{m\times d} \quad \exists\ (P_{\star}, Q_{\star}, K_{\star}, R_{\star}) \in \mathbb{O}_m \times \mathbb{O}_n \times \mathbb{O}_s \times \mathbb{O}_d$

such that

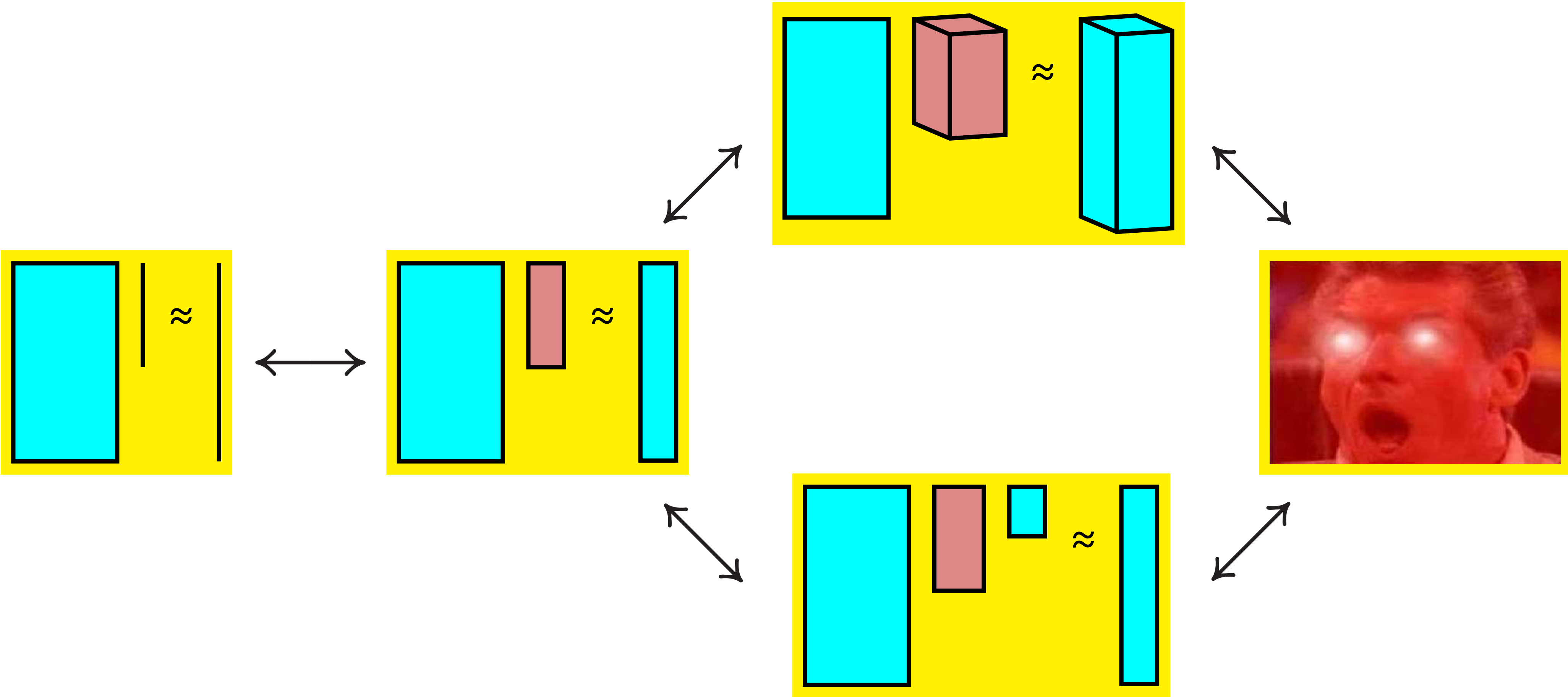
$$(P_{\star}^{\mathsf{T}} A_{\mathsf{L}} Q_{\star})\,(Q_{\star}^{\mathsf{T}} X K_{\star})\,(R_{\star}^{\mathsf{T}} A_{\mathsf{R}} K_{\star})^{\mathsf{T}} =$$
$$\begin{bmatrix} A_{\mathsf{L},11} & 0 \\ 0 & A_{\mathsf{L},22} \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} A_{\mathsf{R},11} & 0 \\ 0 & A_{\mathsf{R},22} \end{bmatrix}^{\mathsf{T}} \approx \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix}$$
$$= (P_{\star}^{\mathsf{T}} B R_{\star}),$$

so the original problem splits into four sub-problems

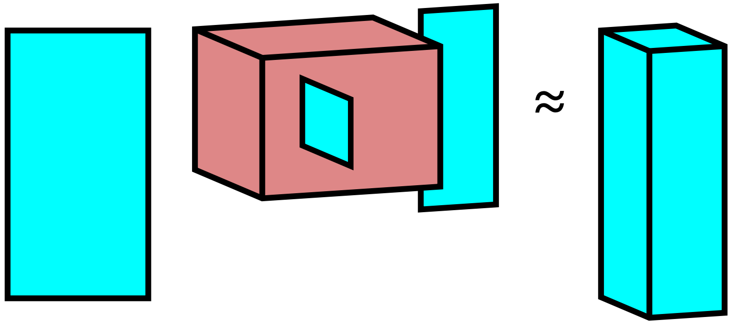
$$A_{\mathsf{L},11} X_{11} A_{\mathsf{R},11}^{\mathsf{T}} \approx B_1, \qquad A_{\mathsf{L},11} X_{12} A_{\mathsf{R},22}^{\mathsf{T}} \approx 0, \quad A_{\mathsf{L},22} X_{21} A_{\mathsf{R},11}^{\mathsf{T}} \approx 0, \quad \ldots$$

where the core problem has a lot of nice properties (CP1)—...

GUT — Grant Unification of TLS:



k-linear tensor right-hand side TLS:



$$\left(A_1, A_2, \dots, A_k \,\middle|\, \mathcal{X} \right) \approx \mathcal{B}$$

$A_\ell \in \mathbb{R}^{m_\ell \times n_\ell}, \quad \ell = 1, 2, \dots, k$

$\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k}$

$\mathcal{B} \in \mathbb{R}^{m_1 \times m_2 \times \cdots \times m_k}$

- TLS solvability is not clear.
- For some ℓ it may be $A_\ell = I$.
- For some ℓ it may be $E_\ell = 0$ (the *joined LS-TLS minimization*).
- The core problem (the minimally dimensioned sub-problem) exists.
- But...

...it does not help us to answer our *original question*...

(On the left hand-side, there may be sum of terms with the same $\mathcal{X}$, yielding *Lyapunov-like* or *Sylvester-like* problems.)

Look inside

Explore the internal structure of core problems...

(De-)composing of CPs:

(*We restrict ourselves on matrix case only.*)

Let for some $(P, Q, R) \in \mathbb{O}_{\bar{m}} \times \mathbb{O}_{\bar{n}} \times \mathbb{O}_{\bar{d}}$

$$P^{\mathsf{T}}[B_1, A_{11}] \begin{bmatrix} R & 0 \\ 0 & Q \end{bmatrix} = \left[\begin{array}{cc|cc} B_{1,\mu} & 0 & A_{11,\mu} & 0 \\ 0 & B_{1,\nu} & 0 & A_{11,\nu} \end{array}\right].$$

Then

$$A_{11}X_{11} \approx B_1 \text{ is CP} \qquad \Longleftrightarrow \qquad \begin{cases} & A_{11,\mu}X_{11,\mu} \approx B_{1,\mu} \text{ is CP} \\ \& & A_{11,\nu}X_{11,\nu} \approx B_{1,\nu} \text{ is CP} \end{cases} ;$$

where CP is any problem satisfying (CP1)—(CP3).

Schematically

$$[B_1, A_{11}] \equiv [B_{1,\mu}, A_{11,\mu}] \boxplus [B_{1,\nu}, A_{11,\nu}].$$

Our goal is to study **how**

- TLS solvability of composed problem $[B_1, A_{11}]$ **depends** on
- TLS solvabilities of individual components $[B_{1,\mu}, A_{11,\mu}]$ and $[B_{1,\nu}, A_{11,\nu}]$.

For example we already know that some compositions of two components are *admissible* and some are *forbiden*:

$\boxplus$	$\mathfrak{F}_1$	$\mathfrak{F}_2$	$\mathfrak{F}_3$	$\mathfrak{S}$
$\mathfrak{F}_1$	$\mathfrak{F}_1$, $\mathfrak{F}_2$, $\mathfrak{F}_3$, $\mathfrak{S}$	—"—	—"—	—"—
$\mathfrak{F}_2$	$\mathfrak{F}_2$	$\mathfrak{F}_2$	—"—	—"—
$\mathfrak{F}_3$	$\cancel{\mathfrak{F}_1}$, $\cancel{\mathfrak{F}_2}$, $\mathfrak{F}_3$	$\cancel{\mathfrak{F}_1}$, $\cancel{\mathfrak{F}_2}$	$\cancel{\mathfrak{F}_1}$, $\cancel{\mathfrak{F}_2}$, $\mathfrak{F}_3$	—"—
$\mathfrak{S}$	$\cancel{\mathfrak{F}_1}$, $\cancel{\mathfrak{F}_2}$, $\mathfrak{F}_3$, $\mathfrak{S}$	$\cancel{\mathfrak{F}_1}$, $\cancel{\mathfrak{F}_2}$	$\cancel{\mathfrak{F}_1}$, $\cancel{\mathfrak{F}_2}$	$\cancel{\mathfrak{F}_1}$, $\cancel{\mathfrak{F}_2}$, $\cancel{\mathfrak{F}_3}$, $\mathfrak{S}$

However, it will be useful to know something more about the *decomposability* of CPs, and how the *irreducible components* looks like.

Irreducible representation w.r.t. decomposing:

Let start with a single matrix $M \in \mathbb{R}^{s \times t}$

$$P^\top M Q = \begin{bmatrix} M_\mu & 0 \\ 0 & M_\nu \end{bmatrix} \qquad \Longleftrightarrow \qquad M \equiv M_\mu \boxplus M_\nu$$

Since the SVD $M = U \Sigma V^\top = \sum_{j=1}^{r} u_j \sigma_j v_j^\top$, where $r = \mathrm{rank}(M)$, performs *orthogonal block diagonalization with smallest blocks* of M, it gives the *irreducible representation* of M in terms of *decomposing*

$$M \equiv [\sigma_1] \boxplus [\sigma_2] \boxplus \cdots \boxplus [\sigma_r] \boxplus [0_{s-r,t-r}] = \left(\bigboxplus_{j=1}^{r} [\sigma_j] \right) \boxplus [0_{s-r,t-r}].$$

How to generalize such concept on pairs (or in general tuples) of matrices?

Tuples of matrices — A general framework:

The set of all matrices, and the set of all k-tuples of all matrices

$$\mathscr{M} = \mathscr{M}(\mathbb{R}) = \bigcup_{s,t \in \mathbb{N}_0} \mathbb{R}^{s \times t}, \qquad \mathscr{M}^k,$$

form together with $\boxplus$ (applied componentwisely on tuples) monoids.

Let $k = 4$ and let

$$\big(M_1, M_2, M_3, M_4 \big) \in \mathscr{M}^4, \qquad M_\ell \in \mathbb{R}^{s_\ell \times t_\ell}, \quad \ell = 1, 2, 3, 4.$$

We introduce three objects:

- *Tuple alignment* — set of algebraic relations among dimension, e.g.,

$$\mathsf{TA} = \big\{ s_1 - s_2 = 0,\ s_1 + s_3 - s_4 = 0,\ t_1 + t_2 - t_3 = 0 \big\}$$

then $\mathscr{M}^4_{\mathsf{TA}}$ denotes the set of all TA-aligned quaruples of matrices.

Tuples of matrices — A general framework:

- *Tuple matricizaton* — mapping from $\mathscr{M}^k_{\mathsf{TA}}$ back to $\mathscr{M}$

$$\mathsf{TM} : \mathscr{M}^4_{\mathsf{TA}} \longrightarrow \mathscr{M}, \quad \text{e.g.,} \quad \mathsf{TM}\big(M_1, M_2, M_3, M_4\big) = \left[\begin{array}{c|c} \begin{array}{c|c} M_1 & M_2 \\ \hline \multicolumn{2}{c}{M_3} \end{array} & M_4 \end{array}\right]$$

- *Tuple transformation* — group of allowed orthogonal trafsformations $\mathsf{TT} = \mathsf{TT}(\mathsf{TA}, \mathsf{TM})$ that defines the congruence, e.g.,

$$\left[\begin{array}{c|c} \begin{array}{c|c} M_1 & M_2 \\ \hline \multicolumn{2}{c}{M_3} \end{array} & M_4 \end{array}\right] \equiv_{\mathsf{TT}} \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}^{\mathsf{T}} \left[\begin{array}{c|c} \begin{array}{c|c} M_1 & M_2 \\ \hline \multicolumn{2}{c}{M_3} \end{array} & M_4 \end{array}\right] \begin{bmatrix} Q_1 & 0 & 0 \\ 0 & Q_2 & 0 \\ 0 & 0 & Q_3 \end{bmatrix}$$

Search for irreducible representation:

Thus we are living in the set of *general problems* — pairs of matrices

$$\mathscr{M}^2_{\mathsf{TA}} = \left\{(A,B) : A \in \mathbb{R}^{m_A \times n},\ B \in \mathbb{R}^{m_B \times d}\right\}, \qquad \mathsf{TA} = \left\{m_A - m_B = 0\right\}$$

where $\mathsf{TM}(A,B) = [B,A]$ and $[B,A] \equiv_{\mathsf{TT}} P^{\mathsf{T}}[B,A]\,\mathrm{diag}(R,Q)$.

• The *core problem* revealing transformation realizes decomposition within $\mathscr{M}^2_{\mathsf{TA}}$ of the form

$$(A,B) \equiv (A_{11}, B_1) \boxplus (A_{22}, 0)$$

where the *core problem reduction* itself represents orthogonal projection onto a proper subset of $\mathscr{M}^2_{\mathsf{TA}}$

$$\mathsf{CPR}(A,B) = (A_{11}, B_1), \qquad \mathsf{CPR} : \mathscr{M}^2_{\mathsf{TA}} \longrightarrow \mathscr{M}^2_{\mathsf{TA,CP}}$$

with a lot of nice properties.

Search for irreducible representation:

- For example

$$\mathrm{CPR}\big((A_\mu, B_\mu) \boxplus (A_\nu, B_\nu)\big) \equiv \mathrm{CPR}(A_\mu, B_\mu) \boxplus \mathrm{CPR}(A_\nu, B_\nu)$$

- We are able to extract something we call *degenerated component* that only increases the residuum (if it is present)

$$(A, B) \equiv (A_{11,\varphi}, B_{1,\varphi}) \boxplus (0, B_{1,\psi}) \boxplus (A_{22}, 0)$$

- We know that $(A_{11,\varphi}, B_{1,\varphi})$ is uniquely decomposable into irreducible components (sub-problems).

- We know *some* irreducible components (VRHS CP, but not only). General description is not clear.


That's all Folks!

References:

Golub, Van Loan: SIAM NA 1980 (10.1137/0717073)
Van Huffel, Vandewalle: SIAM 1991 (10.1137/1.9781611971002)
Paige, Strakoš: SIMAX 2005 (10.1137/040616991)
Hnětynková, P, Sima, Strakoš, Van Huffel: SIMAX 2012 (10.1137/100813348)
Hnětynková, P, Strakoš: SIMAX 2013 (10.1137/120884237)
Hnětynková, P, Strakoš: SIMAX 2015 (10.1137/140968914)
Hnětynková, P, Sima: SIMAX 2016 (10.1137/15M1028339)

Kolda, Bader: SIREV 2009 (10.1137/07070111X)
Hnětynková, P, Žáková: LAA 2018 (10.1016/j.laa.2018.06.016)
Kukush, Markovsky, Van Huffel: TLS-EIV 2002 (10.1007/978-94-017-3552-0_14)
Kukush, Markovsky, Van Huffel: Metrika 2003 (10.1007/s001840200217)
Hnětynková, P, Žáková: LAA 2019 (10.1016/j.laa.2019.04.018)
Hnětynková, P, Žáková: SIMAX 2023 (10.1137/21M1462155)

Hnětynková, P, Žáková: AoM 2019 (10.21136/am.2019.0252-18)
Žáková: PhD Thesis, TUL 2024 (https://kma.fp.tul.cz/externi-data/soubory/zaverecne-prace/2024-OM-PHD_Zakova.pdf)