DYSCO course on low-rank approximation and its applications

Computational tools

Ivan Markovsky

Vrije Universiteit Brussel

Plan

1. Introduction
2. Computational tools
3. Behavioral approach
4. System identification
5. Subspace methods
6. Generalizations

Outline

QR decomposition

SVD decomposition

Least squares and least norm problems

Weighted and regularized least squares problems

Exercise

Total least squares problems

Outline

QR decomposition

SVD decomposition

Least squares and least norm problems

Weighted and regularized least squares problems

Exercise

Total least squares problems

Orthonormal set of vectors

- consider a finite set of vectors $\mathscr{Q} := \{ q_1, \ldots, q_k \} \subset \mathbb{R}^n$
- $\mathscr{Q}$ is orthogonal $:\iff \langle q_i, q_j \rangle := q_i^\top q_j = 0$, for all $i \neq j$
- $\mathscr{Q}$ is normalized $:\iff \|q_i\|_2^2 := \langle q_i, q_i \rangle = 1$, $i = 1, \ldots, k$
- $\mathscr{Q}$ is orthonormal $:\iff \mathscr{Q}$ is orthogonal $+$ normalized
- $Q := \begin{bmatrix} q_1 & \cdots & q_k \end{bmatrix}$ orthonormal $\iff Q^\top Q = I_k$
- properties:
  - orthonormal vectors are independent
  - multiplication preserves inner product and norm

$$\langle Qz, Qy \rangle = z^\top Q^\top Q y = z^\top y = \langle z, y \rangle$$

Orthogonal projectors

- consider an orthonormal set $\mathscr{Q} := \{ q_1, \ldots, q_k \}$
- $\mathscr{Q}$ is an orthonormal basis for $\mathscr{L} := \mathrm{span}(\mathscr{Q}) \subseteq \mathbb{R}^n$
- $Q^\top Q = I_k$, however, for $k < n$, $QQ^\top \neq I_n$
- $\Pi_{\mathrm{span}(\mathscr{Q})} := QQ^\top$ is orthogonal projector on span($\mathscr{Q}$)

$$\Pi_{\mathscr{L}} x = \arg\min_y \|x - y\|_2 \quad \text{subject to} \quad y \in \mathscr{L}$$

- Properties:
  - $\Pi = \Pi^2$, $\Pi = \Pi^\top$ (necessary and sufficient conditions)
  - $\Pi^\perp := (I - \Pi)$ is orthogonal projector on

$$(\mathrm{span}(\Pi))^\perp \subseteq \mathbb{R}^n \text{— orth. complement of span}(\Pi)$$

Orthonormal basis for $\mathbb{R}^n$

- orthonormal set $\mathscr{Q} := \{ q_1, \ldots, q_n \} \subset \mathbb{R}^n$ of n vectors
- $Q := \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix}$ is orthogonal and $Q^\top Q = I_n$
- it follows that $Q^{-1} = Q^\top$ and

$$QQ^\top = \sum_{i=1}^{n} q_i q_i^\top = I_n$$

- expansion in orthonormal basis $x = QQ^\top x$
  - $\widetilde{x} := Q^\top x$ coordinates of x in the basis $\mathscr{Q}$
  - $x = Q\widetilde{x}$ reconstruct x from the coordinates a
- geometrically multiplication by Q (and $Q^\top$) is rotation

Gram-Schmidt (G-S) procedure

- given independent set $\{ a_1, \ldots, a_k \} \subset \mathbb{R}^n$
- G-S produces orthonormal set $\{ q_1, \ldots, q_k \} \subset \mathbb{R}^n$

$$\operatorname{span}(a_1, \ldots, a_r) = \operatorname{span}(q_1, \ldots, q_r), \quad \text{for all } r \leq k$$

- G-S procedure: Let $q_1 := a_1 / \|a_1\|_2$. For $i = 2, \ldots, k$
  1. orthogonalized a_i w.r.t. $q_1, \ldots, q_{i-1}$:

$$v_i := \underbrace{(I - \Pi_{\operatorname{span}(q_1, \ldots, q_{i-1})}) a_i}_{\text{projection of } a_i \text{ on } (\operatorname{span}(q_1, \ldots, q_{i-1}))^{\perp}}$$

  2. normalize the result: $q_i := v_i / \|v_i\|_2$

QR decomposition

G-S gives as a byproduct scalars r_{ji}, $j \le i$, $i = 1, \dots, k$

$$\begin{aligned} a_i &= (q_1^\top a_i) q_1 + \dots + (q_{i-1}^\top a_i) q_{i-1} + \|v_i\|_2 q_i \\ &= r_{1i} q_1 + \dots + r_{ii} q_i \end{aligned}$$

in a matrix form G-S produces the matrix decomposition

$$\underbrace{\begin{bmatrix} a_1 & a_2 & \cdots & a_k \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} q_1 & q_1 & \cdots & q_k \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1k} \\ 0 & r_{22} & \cdots & r_{2k} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{kk} \end{bmatrix}}_{R}$$

with orthonormal $Q \in \mathbb{R}^{n \times k}$ and upper triangular $R \in \mathbb{R}^{k \times k}$

- If $\{ a_1, \ldots, a_k \}$ are dependent

$$v_i := (I - \Pi_{\mathrm{span}(q_1,\ldots,q_{i-1})}) a_i = 0 \quad \text{for some } i$$

- conversely, if $v_i = 0$ for some i, a_i is linearly dependent on $\{ a_1, \ldots, a_{i-1} \}$

- Modified G-S procedure: when $v_i = 0$, skip to a_{i+1} $\implies$ *R is in upper staircase form,* *e.g.*,

$$\begin{bmatrix} \times & \times & \times & \times & \times & \times & \times \\ & \times & \times & \times & \times & \times & \times \\ & & & \times & \times & \times & \times \\ & & & & & & \times \\ & & & & & & \end{bmatrix}$$

(empty elements are zeros)

Full QR

$$A = \underbrace{\begin{bmatrix} Q_1 & Q_2 \end{bmatrix}}_{\text{orthogonal}} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \qquad \begin{array}{rcl} \operatorname{colspan}(A) & = & \operatorname{colspan}(Q_1) \\ \left(\operatorname{colspan}(A)\right)^{\perp} & = & \operatorname{colspan}(Q_2) \end{array}$$

- procedure for finding Q_2

  complete A to full rank matrix, e.g.,
  $A_m := \begin{bmatrix} A & I \end{bmatrix}$, *and apply G-S on* A_m

- application:
  complete an orthonormal matrix $Q_1 \in \mathbb{R}^{n \times k}$
  to an orthogonal matrix $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \in \mathbb{R}^{n \times n}$
  (by computing the full QR of $\begin{bmatrix} Q_1 & I \end{bmatrix}$)

Outline

QR decomposition

SVD decomposition

Least squares and least norm problems

Weighted and regularized least squares problems

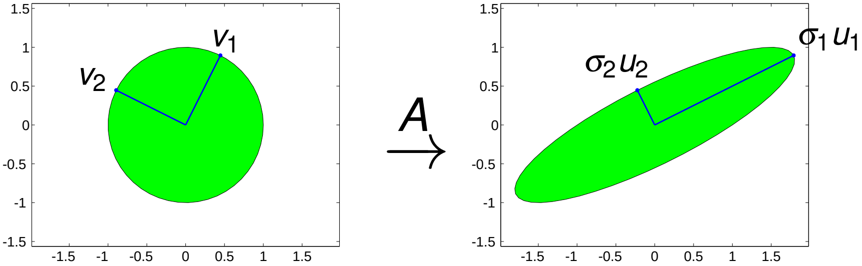
Exercise

Total least squares problems

Geometric fact motivating the SVD

The image of a unit ball under linear map is a hyperellips.

$$\underbrace{\begin{bmatrix} 1.00 & 1.50 \\ 0 & 1.00 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 0.89 & -0.45 \\ 0.45 & 0.89 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 2.00 & 0 \\ 0 & 0.50 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} 0.45 & -0.89 \\ 0.89 & 0.45 \end{bmatrix}}_{V^\top}$$



Singular value decomposition

any $m \times n$ matrix A of rank r has a reduced SVD

$$A = \underbrace{\begin{bmatrix} u_1 & \cdots & u_r \end{bmatrix}}_{U_1} \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}}_{\Sigma_1} \underbrace{\begin{bmatrix} v_1 & \cdots & v_r \end{bmatrix}^\top}_{V_1^\top}$$

with U_1 and V_1 orthonormal

- $\sigma_1 \geq \cdots \geq \sigma_r$ are called singular values
- $u_1, \ldots, u_r$ are called left singular vectors
- $v_1, \ldots, v_r$ are called right singular vectors

The SVD is both computational and analytical tool

Full SVD $\quad A = U\Sigma V^\top$

where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal and

$$\Sigma = \begin{array}{cc} \begin{array}{cc} r & n-r \end{array} & \\ \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} & \begin{array}{c} r \\ m-r \end{array} \end{array} \qquad \text{where} \qquad \Sigma_1 = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$$

the singular values of A are

$$\sigma(A) := \left(\sigma_1, \ldots, \sigma_r, \quad \underbrace{0, \ldots, 0}_{\min(n-r, m-r)} \quad\right)$$

- $\sigma_{\min}(A)$ — smallest singular value of A
- $\sigma_{\max}(A)$ — largest singular value of A

Proof of existence of an SVD

- constructive, based on induction, assume $m \geq n$
- end of induction: vector $A \in \mathbb{R}^{m \times 1}$ has reduced SVD

$$A = U \Sigma V^\top, \quad \text{with} \quad U := A / \|A\|_2, \quad \Sigma := \|A\|_2, \quad V := 1$$

- inductive step: let $\sigma_i := \|A_i\|_2$, $\exists\ u_i \in \mathbb{R}^m$ and $v_i \in \mathbb{R}^n$

$$A_i v_i =: \sigma_i u_i, \quad \text{where} \quad \|u_i\|_2 = 1, \quad \text{with} \quad \|v_i\|_2 = 1$$

- complete u_i and v_i to orthogonal matrices (QR)

$$U_i := \begin{bmatrix} u_i & \star \end{bmatrix} \qquad \text{and} \qquad V_i := \begin{bmatrix} v_i & \star \end{bmatrix}$$

- for certain $w \in \mathbb{R}^{n-1}$ and $A_{i+1} \in \mathbb{R}^{(n-1)\times(n-1)}$

$$U_i^\top A_i V_i = \begin{bmatrix} \sigma_i & w^\top \\ 0 & A_{i+1} \end{bmatrix}$$

- next we show that $w = 0$

$$\sigma_i^2 = \|A_i\|_2^2 = \max_v \frac{\|A_i v\|_2^2}{\|v\|_2^2} \geq \frac{\|A_i \left[\begin{smallmatrix} \sigma_i \\ w \end{smallmatrix}\right]\|_2^2}{\|\left[\begin{smallmatrix} \sigma_i \\ w \end{smallmatrix}\right]\|_2^2}$$

$$= \frac{1}{\sigma_i^2 + w^\top w} \left\| \begin{bmatrix} \sigma_i^2 + w^\top w \\ A_{i+1} w \end{bmatrix} \right\|_2^2$$

$$\geq \frac{1}{\sigma_i^2 + w^\top w} (\sigma_i^2 + w^\top w)^2 = \sigma_i^2 + w^\top w$$

- $\sigma_i^2 \geq \sigma_i^2 + w^\top w \implies w = 0$

Low-rank approximation

given

- a matrix $A \in \mathbb{R}^{m \times n}$, $m \geq n$, and
- an integer r, $0 < r < n$,

find

$$\widehat{A} := \arg\min_{\widehat{A}} \|A - \widehat{A}\| \quad \text{subject to} \quad \operatorname{rank}(\widehat{A}) \leq r$$

- Interpretation: $\widehat{A}^*$ is optimal rank-r approx. of A w.r.t.

$$\|A\|_{\mathrm{F}}^2 := \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \qquad \text{or} \qquad \|A\|_2 := \max_x \frac{\|Ax\|_2}{\|x\|_2}$$

- $\widehat{A}^*$ is optimal in any unitarily invariant norm

Solution via truncated SVD

$$\widehat{A}^* := \arg\min_{\widehat{A}} \|A - \widehat{A}\|_F \quad \text{subject to} \quad \operatorname{rank}(\widehat{A}) \leq r \quad \text{(LRA)}$$

Theorem Let $A = U\Sigma V^\top$ be the SVD of A and define

$$U =: \begin{array}{cc} r & r-n \\ \left[U_1 \right. & \left. U_2 \right] \end{array} \; n\,, \quad \Sigma =: \begin{array}{cc} \begin{array}{cc} r & r-n \end{array} & \\ \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} & \begin{array}{c} r \\ r-n \end{array} \end{array}, \quad V =: \begin{array}{cc} r & r-n \\ \left[V_1 \right. & \left. V_2 \right] \end{array} \; n$$

A solution to (LRA) is

$$\widehat{A}^* = U_1 \Sigma_1 V_1^\top$$

It is unique if and only if $\sigma_r \neq \sigma_{r+1}$

Numerical rank

- distance of A to the manifold of rank-r matrices

$$\sqrt{\textstyle\sum_{i=r+1}^{n} \sigma_i^2} = \min_{\widehat{A}} \|A - \widehat{A}\|_{\mathsf{F}} \quad \text{subject to} \quad \operatorname{rank}(\widehat{A}) \le r$$

$$\sigma_{r+1} = \min_{\widehat{A}} \|A - \widehat{A}\|_2 \quad \text{subject to} \quad \operatorname{rank}(\widehat{A}) \le r$$

- $\sigma_{\min}(A)$ is the distance of A to rank deficiency

- numerical rank: $\operatorname{rank}(A, \varepsilon) :=$ # of singular values $> \varepsilon$

- $\operatorname{rank}(A, \varepsilon)$ depends on an a priori given tolerance ε

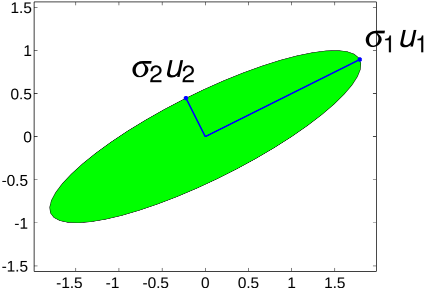
Pseudo-inverse $A^+ := V_1 \Sigma_1^{-1} U_1^\top \in \mathbb{R}^{n \times m}$

$$\begin{aligned}
\mathrm{rank}(A) = n = m \quad &\Longrightarrow \quad A^+ = A^{-1} \\
\mathrm{rank}(A) = n \quad &\Longrightarrow \quad A^+ = (A^\top A)^{-1} A^\top \\
\mathrm{rank}(A) = m \quad &\Longrightarrow \quad A^+ = A^\top (A A^\top)^{-1}
\end{aligned}$$

- $A^+ y$ is least squares-least norm solution of $Ax = y$
- the pseudo-inverse depends on the rank of A
- in practice, the numerical rank $\mathrm{rank}(A, \varepsilon)$ is used
- the SVD, gives reliable way of solving $Ax = y$

Condition number $\kappa(A) := \sigma_{\max}(A)/\sigma_{\min}(A)$

- $\kappa(A)$ is eccentricity of hyperellipsoid $A\{\, x \mid \|x\|_2 = 1 \,\}$



- $\kappa(A)$ — sensitivity of $A^+ y$ to perturbations in y, A
- for large $\kappa(A)$ (≥ 1000) A is called ill-conditioned

Outline

QR decomposition

SVD decomposition

Least squares and least norm problems

Weighted and regularized least squares problems

Exercise

Total least squares problems

Least squares

- overdetermined system of linear equations $Ax = b$
- given $A \in \mathbb{R}^{m \times n}$, $m > n$ and $b \in \mathbb{R}^m$, find $x \in \mathbb{R}^n$
- for "most" A and b, there is no solution x
- Least squares approximation:

  choose x that minimizes 2-norm of the residual

$$e(x) := b - Ax$$

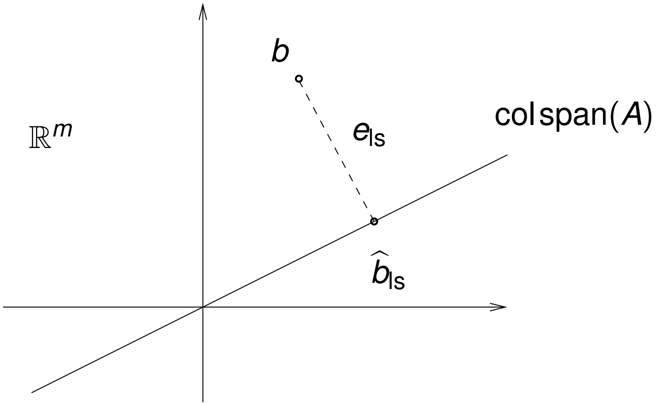
- least squares approximate solution

$$\widehat{x}_{\mathrm{ls}} := \arg\min_x \| \underbrace{b - Ax}_{e(x)} \|_2$$

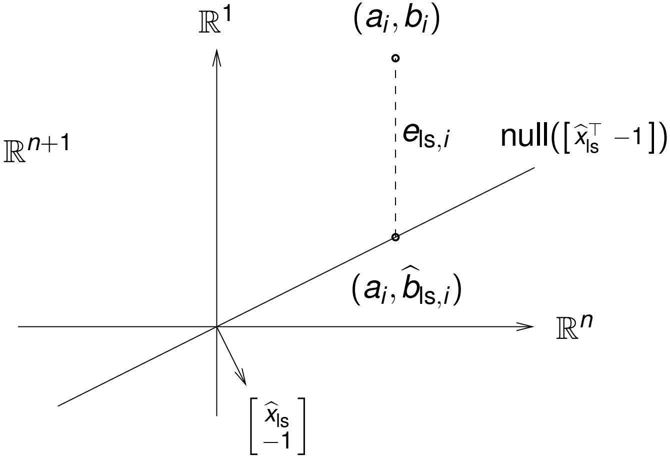
Geometric interpretation: project b onto the image of A

($\widehat{b}_{\text{ls}} := A\widehat{x}_{\text{ls}}$ is the projection)

$$e_{\text{ls}} := \widehat{b}_{\text{ls}} - A\widehat{x}_{\text{ls}}$$



Another geometric interpretation of the LS approximation:



$$
\begin{aligned}
A\widehat{x}_{\mathrm{ls}} = \widehat{b}_{\mathrm{ls}} \quad &\iff \quad \begin{bmatrix} A & \widehat{b}_{\mathrm{ls}} \end{bmatrix} \begin{bmatrix} \widehat{x}_{\mathrm{ls}} \\ -1 \end{bmatrix} = 0 \\
&\iff \quad \begin{bmatrix} a_i & \widehat{b}_{\mathrm{ls},i} \end{bmatrix} \begin{bmatrix} \widehat{x}_{\mathrm{ls}} \\ -1 \end{bmatrix} = 0, \quad \text{for } i = 1, \ldots, m
\end{aligned}
$$

(a_i is the ith row of A)

- $\begin{bmatrix} a_i \\ \widehat{b}_{\mathrm{ls},i} \end{bmatrix}$ lies on subspace perpendicular to span$\left(\begin{bmatrix} \widehat{x}_{\mathrm{ls}} \\ -1 \end{bmatrix}\right)$
- "data point" $\begin{bmatrix} a_i \\ b_i \end{bmatrix} = \begin{bmatrix} a_i \\ \widehat{b}_{\mathrm{ls},i} \end{bmatrix} + \begin{bmatrix} 0 \\ e_{\mathrm{ls},i} \end{bmatrix}$
- approx. error $\begin{bmatrix} 0 \\ e_{\mathrm{ls},i} \end{bmatrix}$ is the vertical distance

Notes

- assuming $m \geq n = \text{rank}(A)$, *i.e.*, A is full column rank,

$$\widehat{x}_{\text{ls}} = (A^\top A)^{-1} A^\top b$$

  is the unique least squares approximate solution

- $\widehat{x}_{\text{ls}}$ is a linear function of b
- if A is square, $\widehat{x}_{\text{ls}} = A^{-1} b$
- $\widehat{x}_{\text{ls}}$ is an exact solution if $Ax = b$ has an exact solution
- $\widehat{b}_{\text{ls}} := A\widehat{x}_{\text{ls}} = A(A^\top A)^{-1} A^\top b$ is LS approx. of b

Projector onto the span of A

- the $m \times m$ matrix

$$\Pi_{\mathrm{col\,span}(A)} := A(A^\top A)^{-1}A^\top$$

  is the orthogonal projector onto $\mathscr{L} := \mathrm{col\,span}(A)$

- the columns of A are an arbitrary basis for $\mathscr{L}$
- if the columns of Q form an orthonormal basis for $\mathscr{L}$

$$\Pi_{\mathrm{col\,span}(Q)} := QQ^\top$$

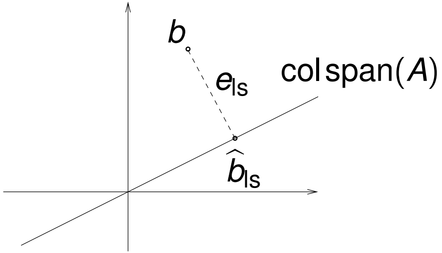
Orthogonality principle

the least squares residual vector

$$e_{\text{ls}} := b - A\widehat{x}_{\text{ls}} = \underbrace{\left(I_m - A(A^\top A)^{-1}A^\top\right)}_{\Pi_{(\text{col span}(A))^\perp}} b$$

is orthogonal to col span(A)

$$\langle e_{\text{ls}}, A\widehat{x}_{\text{ls}}\rangle = b^\top\left(I_m - A(A^\top A)^{-1}A^\top\right)A\widehat{x}_{\text{ls}} = 0, \quad \text{for all } b \in \mathbb{R}^m$$



Least squares via QR decomposition

Let $A = QR$ be the reduced QR decomposition of A.

$$
\begin{aligned}
(A^\top A)^{-1} A^\top &= (R^\top Q^\top Q R)^{-1} R^\top Q^\top \\
&= (R^\top Q^\top Q R)^{-1} R^\top Q^\top = R^{-1} Q^\top
\end{aligned}
$$

$$\widehat{x}_{\mathrm{ls}} = R^{-1} Q^\top b \quad \text{and} \quad \widehat{b}_{\mathrm{ls}} := A x_{\mathrm{ls}} = Q Q^\top b$$

we have a sequence of LS problems ($A =: \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$)

$$A^i x^i = b, \quad \text{where } A^i := \begin{bmatrix} a_1 & \cdots & a_i \end{bmatrix}, \quad \text{for } i = 1, \dots, n$$

R_i — leading $i \times i$ submatrix of R and $Q_i := \begin{bmatrix} q_1 & \cdots & q_i \end{bmatrix}$

$$\widehat{x}_{\mathrm{ls}}^i = R_i^{-1} Q_i^\top b$$

Least norm solution

underdetermined system $Ax = b$, with full rank $A \in \mathbb{R}^{m \times n}$

The set of solutions is

$$\{\, x \in \mathbb{R}^n \mid Ax = b \,\} = \{\, x_{\mathrm{p}} + z \mid z \in \mathrm{null}(A) \,\}$$

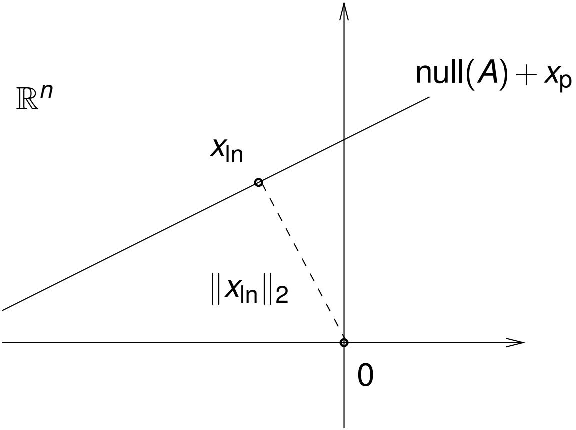
where x_{p} is a particular solution, *i.e.*, $Ax_{\mathrm{p}} = b$.

Least norm problem

$$x_{\mathrm{ln}} := \arg\min_{x} \|x\|_2 \quad \text{subject to} \quad Ax = b$$

Geometric interpretation:

- x_{ln} is the projection of 0 onto the solution set
- orthogonality principle $x_{\text{ln}} \perp \text{null}(A)$



Derivation via Lagrange multipliers

consider the least norm problem with A full rank

$$\min_x \|x\|_2^2 \quad \text{subject to} \quad Ax = b$$

introduce Lagrange multipliers $\lambda \in \mathbb{R}^m$

$$L(x, \lambda) = xx^\top + \lambda^\top (Ax - b)$$

the optimality conditions are

$$\nabla_x L(x, \lambda) = 2x + A^\top \lambda = 0$$
$$\nabla_\lambda L(x, \lambda) = Ax - b = 0$$

substituting $x = -A^\top \lambda / 2$ into the second eqn.

$$\lambda = -2(AA^\top)^{-1} b \quad \Longrightarrow \quad x_{\text{ln}} = A^\top (AA^\top)^{-1} b$$

Solution via QR decomposition

Let $A^\top = QR$ be the reduced QR decomposition of $A^\top$.

$$A^\top (AA^\top)^{-1} = QR(R^\top Q^\top QR)^{-1} = Q(R^\top)^{-1}$$

is a right inverse of A. Then

$$x_{\text{ln}} = Q(R^\top)^{-1} b$$

Outline

Weighted least squares

- weighted 2-norm, defined by $W \in \mathbb{R}^{m \times m}$, $W > 0$

$$\|e\|_W^2 := e^\top W e$$

- weighted least squares approximation problem

$$\widehat{x}_{W,\text{ls}} := \arg\min_x \|b - Ax\|_W$$

- orthogonality principle holds with inner product

$$\langle e, b \rangle_W := e^\top W b$$

- solution

$$\widehat{x}_{W,\text{ls}} = (A^\top W A)^{-1} A^\top W b$$

Recursive least squares

- let $a_i^\top$ be the ith row of A

$$A = \begin{bmatrix} — & a_1^\top & — \\ & \vdots & \\ — & a_m^\top & — \end{bmatrix}$$

$$\|b - Ax\|_2^2 = \sum_{i=1}^{m} (b_i - a_i^\top x)^2$$

$$\hat{x}_{\text{ls}} = \hat{x}_{\text{ls}}(m) := \left(\sum_{i=1}^{m} a_i a_i^\top \right)^{-1} \left(\sum_{i=1}^{m} a_i b_i \right)$$

- (a_i, b_i) correspond to a measurement
- often the (a_i, b_i)'s come sequentially (*e.g.*, in time)

Recursive comput. of $\widehat{x}_{\text{ls}}(m) = \left(\sum_{i=1}^{m} a_i a_i^\top\right)^{-1} \left(\sum_{i=1}^{m} a_i b_i\right)$

- $P(0) = 0 \in \mathbb{R}^{n \times n}$, $q(0) = 0 \in \mathbb{R}^n$

- For $m = 0, 1, \ldots$
  - $P(m+1) := P(m) + a_{m+1} a_{m+1}^\top$
    $q(m+1) := q(m) + a_{m+1} b_{m+1}$
  - $x_{\text{ls}}(m) = P^{-1}(m) q(m)$

Notes:

- the algorithm requires inversion of an $n \times n$ matrix

- $P(m)$ invertible $\implies$ $P(m')$ invertible, for all $m' > m$

Rank-1 update formula:

$$(P + aa^\top)^{-1} = P^{-1} - \frac{1}{1 + a^\top P^{-1} a}(P^{-1}a)(P^{-1}a)^\top$$

Notes:

- $O(n^2)$ method for computing $P^{-1}(m+1)$ from $P^{-1}(m)$
- standard methods based on dense LU, QR, or SVD for computing $P^{-1}(m+1)$ require $O(n^3)$ operations

Multiobjective least squares

- least squares minimizes $J_1(x) := \|b - Ax\|_2^2$
- consider second cost function $J_2(x) := \|z - Bx\|_2^2$
- usually $\min_x J_1(x)$ and $\min_x J_2(x)$ are competing
- common example: $J_2(x) := \|x\|_2^2$ — small x
- feasible objectives:

  $$\{\, (\alpha, \beta) \in \mathbb{R}^2 \mid \exists\, x \in \mathbb{R}^n \text{ subject to } J_1(x) = \alpha,\ J_2(x) = \beta \,\}$$

- trade-off curve: boundary of the feasible objectives
- the corresponding x is called Pareto optimal

Set of Pareto optimal solutions

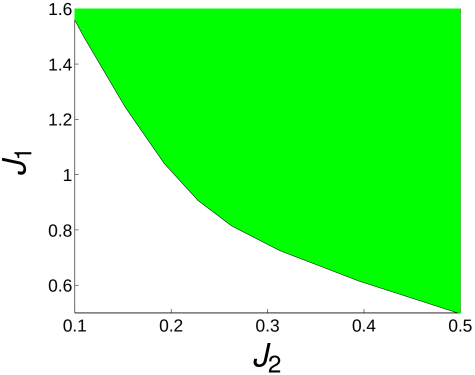
Example:

green area — feasible

white area — infeasible

black line — marginally feasible

Pareto optimal solutions $\leftrightarrow$ points on the line



$\widehat{x}(\mu) = \arg\min_x J_1(x) + \mu J_2(x)$ is Pareto optimal.

varying $\mu \in [0, \infty)$, $\widehat{x}(\mu)$ sweeps the Pareto solutions

Regularized least squares

- Tychonov regularization

$$\widehat{x}_{\text{tych}}(\mu) = \arg\min_x \|b - Ax\|_2^2 + \mu \|x\|_2^2$$

- solution

$$\widehat{x}_{\text{tych}}(\mu) = (A^\top A + \mu I_n)^{-1} A^\top b$$

- exists for any $\mu > 0$, independent of size / rank of A

- trade-off between
  - fitting accuracy $J_1(x) = \|b - Ax\|_2$, and
  - solution size $J_2(x) = \|x\|_2$

Quadratically constrained least squares

- consider biobjective LS problem $\min_x J_1(x)$ and $J_2(x)$
- scalarization approach:

$$\widehat{x}_{\text{tych}}(\mu) = \operatorname*{arg\,min}_x \; J_1(x) + \mu J_2(x)$$

where μ is trade-off parameter

- constrained optimization approach:

$$\widehat{x}_{\text{constr}}(\gamma) = \operatorname*{arg\,min}_x \; J_1(x) \quad \text{subject to} \quad J_2(x) \leq \gamma$$

where γ is upper bound on the J_2 objective

Regularized least squares

- Tychonov regularization is scalarization with
  - fitting accuracy $J_1(x) = \|b - Ax\|_2$, and
  - solution size $J_2(x) = \|x\|_2$
- the constrained optimization approach leads to

$$\widehat{x}_{\text{constr}}(\gamma) = \arg\min_x \|b - Ax\|_2^2 \quad \text{subject to} \quad \|x\|_2^2 \leq \gamma^2$$

- least squares minimization over the ball*

$$\mathscr{U}_{\gamma^2} := \{\, x \mid \|x\|_2^2 \leq \gamma^2 \,\}$$

- solution involves scalar nonlinear equation

Secular equation

- if $\|A^+ b\|_2^2 \leq \gamma^2$, then $\widehat{x}_{\text{constr}}(\gamma) = \|A^+ b\|_2^2$
- if $\|A^+ b\|_2^2 > \gamma^2$, then $\widehat{x}_{\text{constr}}(\gamma) \in \mathscr{U}_{\gamma^2}$
- the Lagrangian of

$$\text{minimize}_x \quad \|b - Ax\|_2^2 \quad \text{subject to} \quad \|x\|_2^2 = \gamma^2$$

  is $\|b - Ax\|_2^2 + \mu(\|x\|_2^2 - \gamma^2)$, μ — Lagrange multiplier

- necessary and sufficient optimality condition

$$x_{\text{tych}}^\top(\mu) x_{\text{tych}}(\mu) = \gamma^2, \quad \text{where} \quad x_{\text{tych}}(\mu) := (A^\top A + \mu I)^{-1} b$$

- secular equation (nonlinear equation in μ)

$$b^\top (A^\top A + \mu I)^{-2} b = \gamma^2$$

- has unique positive solution because
  - $\|x_{\text{tych}}(\mu)\|$ is monotonically decreasing on $\mu \in [0, \infty)$ (by assumption $\|x_{\text{tych}}(0)\|_2^2 > \gamma^2$)
  - $\|x_{\text{tych}}(\infty)\|_2^2 = 0$

Outline

Outline

Total least squares (TLS)

- LS minimizes 2-norm of the eqn. error $e(x) := b - Ax$

$$\min_{x,e} \|e\|_2 \quad \text{subject to} \quad Ax = b - e$$

- alternatively, e can be viewed as a correction on b

- the TLS method is motivated by the asymmetry

  both A and b are given data, but only b is corrected

- TLS problem:

$$\min_{x,\Delta A,\Delta b} \left\| \begin{bmatrix} \Delta A & \Delta b \end{bmatrix} \right\|_F \quad \text{subject to} \quad (A + \Delta A)x = b + \Delta b$$

  - ΔA — correction on A, Δb — correction on b
  - Frobenius matrix norm: $\|C\|_F := \sqrt{\sum_{i=1}^m \sum_{j=1}^n c_{ij}^2}$

Geometric interpretation of the TLS criterion

- with $n = 1$, $x \in \mathbb{R}$, $a = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$, $b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$

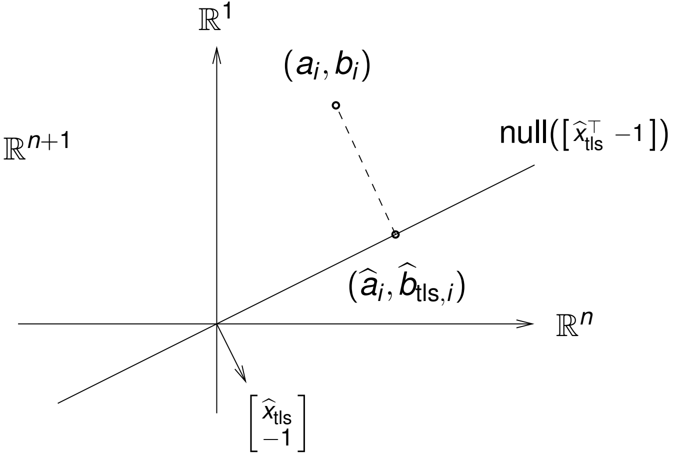
Geometric interpretation:

fit a line $\mathscr{L}(x)$ passing through 0 to the points

$$\begin{bmatrix} a_1 \\ b_1 \end{bmatrix}, \ldots, \begin{bmatrix} a_m \\ b_m \end{bmatrix}$$

- LS minimizes $\sum$ vertical distances^2 from $\begin{bmatrix} a_i \\ b_i \end{bmatrix}$ to $\mathscr{L}(x)$
- TLS minimizes $\sum$ orth. distances^2 from $\begin{bmatrix} a_i \\ b_i \end{bmatrix}$ to $\mathscr{L}(x)$

Geometric interpretation of the TLS criterion



Solution of the TLS problem

Let $\begin{bmatrix} A & b \end{bmatrix} = U\Sigma V^\top$ be the reduced SVD of $\begin{bmatrix} A & b \end{bmatrix}$ and

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_{n+1} \end{bmatrix}, \quad U = \begin{bmatrix} u_1 & \cdots & u_{n+1} \end{bmatrix}, \quad V = \begin{bmatrix} v_1 & \cdots & v_{n+1} \end{bmatrix}$$

TLS solution of $Ax = b$ exists iff $v_{n+1,n+1} \neq 0$ and is unique iff $\sigma_n \neq \sigma_{n+1}$.

In the case when unique TLS solution exists, it is given by

$$\widehat{x}_{\text{tls}} = -\frac{1}{v_{n+1,n+1}} v_{n+1}(1:n)$$

The TLS correction is
$$\begin{bmatrix} \Delta A_{\text{tls}} & \Delta b_{\text{tls}} \end{bmatrix} = -\sigma_{n+1} u_{n+1} v_{n+1}^\top = \begin{bmatrix} A & b \end{bmatrix} v_{n+1} v_{n+1}^\top .$$

Link to low-rank approximation

- TLS approx. $\begin{bmatrix}\widehat{A}_{\text{tls}} & \widehat{b}_{\text{tls}}\end{bmatrix} := \begin{bmatrix}A & b\end{bmatrix} - \begin{bmatrix}\Delta A_{\text{tls}} & \Delta b_{\text{tls}}\end{bmatrix}$ is optimal (in the Frobenius norm) LRA of $\begin{bmatrix}A & b\end{bmatrix}$
- TLS approx. solution of $Ax = b$, $x \in \mathbb{R}^n$ is equivalent to LRA of $D := \begin{bmatrix}A & b\end{bmatrix}$ by rank-n matrix $\widehat{D}$ with

$$\begin{bmatrix}0 & \cdots & 0 & 1\end{bmatrix} \notin \text{kernel}(\widehat{D}) \qquad (*)$$

- generically, the condition $(*)$ is satisfied
- in nongeneric cases, the TLS solution does not exist
- note that the LRA always exists

Statistical properties of TLS

- errors-in-variables (EIV) model

$$A = \overline{A} + \widetilde{A} \quad \text{and} \quad b = \overline{b} + \widetilde{b}$$

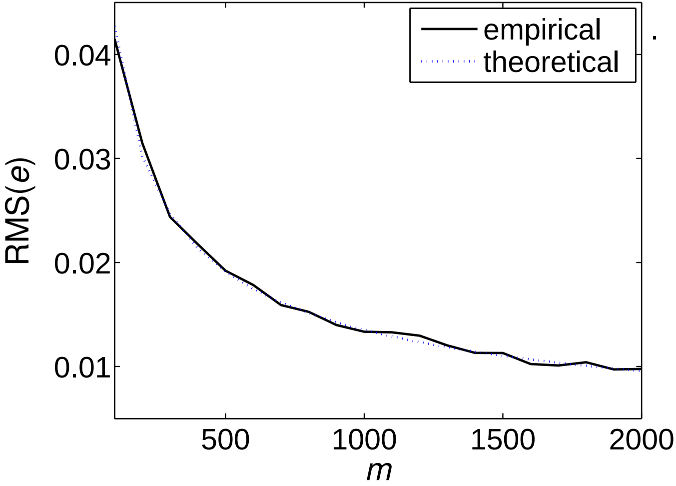
- true values $\overline{A}$, $\overline{b}$ satisfy $\overline{A}\overline{x} = \overline{b}$, for some $\overline{x} \in \mathbb{R}^n$

- perturbations $\widetilde{A}$, $\widetilde{b}$ are zero mean element-wise i.i.d.

- under additional mild assumptions the TLS approx. solution $\widehat{x}$ is a consistent estimator of the true value $\overline{x}$

- measurement errors model
  - A, b — measured data
  - $\overline{x}$ / $\widehat{x}$ — true/estimated model parameters

Estimation error $e = \overline{x} - \widehat{x}$



empirical — solid line, theoretical — dotted line

Notes

- TLS problem vs EIV model
  - TLS approx. can be used without EIV model
  - EIV model shows the correct testbed TLS approx.

- distinguish
  - corrections ΔA, Δb in the TLS problem, and
  - noise/perturbations $\widetilde{A}$, $\widetilde{b}$ in the EIV model

Confidence bounds

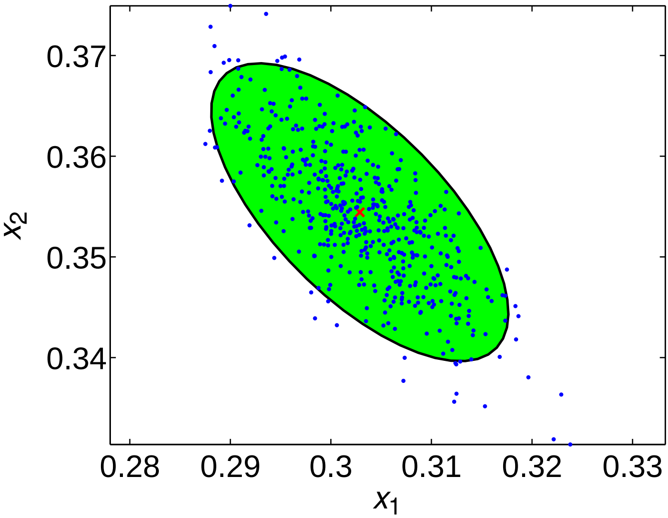
- assume that $\widetilde{A}$, $\widetilde{b}$ are i.i.d. normal with variance ξ^2
- the estimation error e is asymptorically normal
  $\rightsquigarrow$ confidence bounds for $\widehat{x}$
- the asymptotic error $e := \overline{x} - \widehat{x}$ covariance matrix is

$$V_e = \xi^2(1 + \widehat{x}^\top \widehat{x})(A^\top A - m\xi^2 I)^{-1}$$

- the noise variance ξ^2 can be estimated from the data

$$\widehat{\xi}^2 = \frac{1}{m}\sigma_{n+1}^2$$

95% confidence ellipsoid



Weighted total least squares problem

- replace the Frobenius norm by the weighted 2-norm

$$\|D\|_W := \sqrt{\operatorname{vec}^\top(D) W \operatorname{vec}(D)}$$

- $W =$ inverse noise $(\operatorname{vec}([\widetilde{A} \ \ \widetilde{b}]))$ covariance matrix
- in general, WTLS doesn't have analytic solution
- special cases $\rightsquigarrow$ structure of W
  - column/row-wise weighting
  - element-wise weighting
  - generalized TLS
  - restricted TLS

Hierarchy of WTLS problems

1. fully weighted $W \geq 0$

2. column-wise weighted

$$W = \operatorname{diag}(W_1, \ldots, W_m), \quad W_i \in \mathbb{R}_+^{(n+1)\times(n+1)}$$

3. element-wise weighted

$$W = \operatorname{diag}(w), \quad w \in \mathbb{R}_+^{m(n+1)}$$

4. column-wise GTLS: case 2, with W_i's equal

5. column-wise scaled: case 3, with W_i — diagonal

Relative error TLS

- consider the element-wise weighted case

$$\|D\|_w = \|D\|_\Sigma := \|\Sigma \odot D\|_\mathrm{F}$$

($\odot$ — element-wise product)

- $\Sigma_{ij} = 1/d_{ij} \rightsquigarrow$ approximation in relative error sense

$$e_{ij} = \frac{d_{ij} - \widehat{d}_{ij}}{d_{ij}}$$

GTLS problem

- TLS approximation with criterion

$$\|D\|_{\Sigma_l,\Sigma_r} := \|\Sigma_l D \Sigma_r\|_F$$

- link to WTLS

$$\begin{aligned}\|\Sigma_l(D-\widehat{D})\Sigma_r\|_F^2 &= \left\|\operatorname{vec}(\Sigma_l(D-\widehat{D})\Sigma_r)\right\|^2 \\ &= \left\|(\Sigma_r \otimes \Sigma_l)\operatorname{vec}(D-\widehat{D})\right\|^2 \\ &= \operatorname{vec}^\top(D-\widehat{D})\big(W_r \otimes W_l\big)\operatorname{vec}(D-\widehat{D})\end{aligned}$$

where $\sqrt{W_r} = \Sigma_r$ and $\sqrt{W_l} = \Sigma_l$

- WTLS problem with weight matrix $W = W_r \otimes W_l$

Element-wise GTLS

- element-wise weighted total least squares

$$\|D\|_w = \|D\|_\Sigma := \|\Sigma \odot D\|_\mathrm{F}$$

- element-wise generalized total least squares

$$W_\mathrm{r} = \mathrm{diag}(w_\mathrm{r}) \quad \text{and} \quad W_\mathrm{l} = \mathrm{diag}(w_\mathrm{l})$$

- $\rightsquigarrow$ rank-1 matrix $\Sigma = w_\mathrm{l} w_\mathrm{r}^\top$

GTLS solution

- $\sqrt{W_\text{r}} = \Sigma_\text{r}$, w.l.o.g. we can choose Σ_r upper triangular, *e.g.*, the Cholesky factor of W_r

- modified data matrix: $D_\text{m} := \Sigma_\text{l} D \Sigma_\text{r}$

- TLS approximation of D_m: $\widehat{D}_\text{m,tls}$ and $\widehat{x}_\text{m,tls}$

- partition Σ_r as $\begin{bmatrix} \Sigma_\text{r,11} & \Sigma_\text{r,12} \\ 0 & \Sigma_\text{r,22} \end{bmatrix}$, with $\Sigma_\text{r,11} \in \mathbb{R}^{n \times n}$

- GTLS solution

$$\widehat{x}_\text{gtls} = \frac{\Sigma_\text{r,11} \widehat{x}_\text{tls} - \Sigma_\text{r,11}}{\Sigma_\text{r,22}}, \quad \widehat{D}_\text{gtls} = (\Sigma_\text{l})^{-1} \widehat{D}_\text{m,tls} (\Sigma_\text{r})^{-1}$$

Singular weight matrix

- consider the element-wise weighted case

$$\|D\|_w = \|D\|_\Sigma := \|\Sigma \odot D\|_\mathrm{F}$$

- Σ is a matrix of element-wise nonnegative weights
- $\sigma_{ij} = 0 \implies$ the solution doesn't depend on d_{ij}
- zero weights allow us to consider missing data

Restricted total least squares problem

- impose structured correction ΔD

$$\begin{array}{rl} \text{minimize} & \|E\|_{\mathsf{F}} \\ \text{subject to} & (A+\Delta b)x = b+\Delta b \\ \text{and} & \begin{bmatrix}\Delta A & \Delta b\end{bmatrix} = LER \end{array}$$

- link to WTLS: RTLS is a GTLS problem with

$$W_{\mathrm{l}} = (LL^{\top})^{+} \quad \text{and} \quad W_{\mathrm{r}} = (RR^{\top})^{+}$$

(A^{+} is the pseudo-inverse of A)

Structured total least squares

- T. Abatzoglou, J. Mendel, and G. Harada. The constrained total least squares technique and its application to harmonic superresolution. *IEEE Trans. Signal Proc.*, 39:1070–1087, 1991

$$\begin{aligned} &\text{minimize} \quad \text{over } x, \Delta A, \Delta b \quad \left\| \begin{bmatrix} \Delta A & \Delta b \end{bmatrix} \right\|_{\mathrm{F}} \\ &\text{subject to} \quad (A + \Delta A)x = b + \Delta b \text{ and} \\ &\begin{bmatrix} \Delta A & \Delta b \end{bmatrix} \text{ has the same structure as } \begin{bmatrix} A & b \end{bmatrix} \end{aligned}$$

- types of structures
  - linear: Hankel/Toeplitz, Sylvester
  - nonlinear: Vandermonde

Link to structured low-rank approximation

- STLS is equivalent to structured low-rank approx.

$$\begin{array}{ll} \text{minimize} & \text{over } \Delta D \quad \|\Delta D\|_{\mathrm{F}} \\ \text{subject to} & \operatorname{rank}(D + \Delta D) \leq r \text{ and} \\ & \color{red}{\Delta D \text{ has the same structure as } D} \end{array}$$

with $D := \begin{bmatrix} A & b \end{bmatrix}$, $r = n$, and

$$\begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix} \notin \operatorname{kernel}(\widehat{D}) \qquad (*)$$

- generically, the condition $(*)$ is satisfied
- in nongeneric cases, the STLS solution does not exist

History of the problem

- Errors-in-variables system identification

  M. Aoki and P. Yue. On a priori error estimates of some identification methods. *IEEE Trans. Automat. Control*, 15(5):541–548, 1970

- Sum-of-exponetials estimation

  Y. Bresler and A. Macovski. Exact maximum likelihood parameter estimation of superimposed exponential signals in noise. *IEEE Trans. Acust., Speech, Signal Proc.*, 34:1081–1089, 1986

  J. Cadzow. Signal enhancement—A composite property mapping algorithm. *IEEE Trans. Signal Proc.*, 36:49–62, 1988

- Rimmanian SVD algorithm

  B. De Moor. Structured total least squares and L_2 approximation problems. *Linear Algebra Appl.*, 188–189:163–207, 1993

- Structured total least norm algorithm

  J. Rosen, H. Park, and J. Glick. Total least norm formulation and solution of structured problems. *SIAM J. Matrix Anal. Appl.*, 17:110–126, 1996

- Variable projection algorithm

  I. Markovsky, S. Van Huffel, and R. Pintelon. Block-Toeplitz/Hankel structured total least squares. *SIAM J. Matrix Anal. Appl.*, 26(4):1083–1099, 2005