Extension of the Total Least Square Problem Using General Unitarily Invariant Norms

Chi-Kwong Li
Department of Mathematics
College of William and Mary
Williamsburg, Virginia 23187-8795

R.C. Thompson Matrix Meeting, San Jose, 2004.

Joint work with:
Xin-Guo Liu and Xue-Feng Wang
Department of Mathematics, Ocean University of China.

Least square problem

Let A be $m \times n$, b be $m \times 1$, $Ax = b$ has no solution.

Determine f with smallest $\ell_2(f)$ such that $Ax = b - f$ is solvable. Equivalently, we find x so that $f = b - Ax$ has minimum ℓ_2-norm.

Example 1

Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $b = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Then $Ax = b$ has no solution, and $Ax = 0$ is the least square solution. [Here $f = b$.]

A different approach - Total least square problem

Change (A, b) to $(A-E, b-f)$ with minimum $\ell_2([E|f])$ so that $(A-E)x = b-f$ is solvable.

In Example 1, we can take $f = 0$, $E = \begin{pmatrix} 0 & 0 \\ 0 & \varepsilon \\ 0 & \varepsilon \end{pmatrix}$ for any $\varepsilon \neq 0$.

Some definitions

Let $m, n, p \in \mathbb{N}$ with $m \geq n+p$, A be $m \times n$, B be $m \times p$, and

$$\mathcal{P}(A,B) = \{(E,F) : (A-E)X = B-F \text{ is solvable}\}.$$

Clearly, $(E,F) \in \mathcal{P}(A,B)$ if and only if

$$\operatorname{rank}([A-E|B-F]) = \operatorname{rank}(A-E).$$

For a given unitarily invariant (u.i.) norm $\|\cdot\|$, define

$$\rho(A,B) = \inf\{\|[E|F]\| : (E,F) \in \mathcal{P}(A,B)\}.$$

A norm $\|\cdot\|$ on rectangular matrices is u.i. if

$$\|UZV\| = \|Z\| \text{ for all } A \text{ and unitary } U \text{ and } V.$$

So, $\|Z\|$ depends only on the singular values of Z.

Goals

a. Determine $\rho(A, B)$.

b. Check the existence of $(E, F) \in \mathcal{P}(A, B)$ attaining $\rho(A, B)$.
Construct/find such (E, F) if exists.

c. Obtain the results for arbitrary unitarily invariant norms.

Example 1 (again) Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $b = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Then

$(A - E)x = b$ is solvable for any $E = \begin{pmatrix} 0 & 0 \\ 0 & \varepsilon \\ 0 & \varepsilon \end{pmatrix}$ with $\varepsilon \neq 0$.

Then $\rho(A, b) = 0$ and there is no $(E, f) \in \mathcal{P}(A, b)$ attaining $\rho(A, b)$ for any u.i. norm.

Example 2 Let $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$.

Then

$$(E, F) = \left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \in \mathcal{P}(A, B)$$

satisfies $\|[E|F]\| = \rho(A, B)$ for any u.i. norm.

Example 3 Let $A = \begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$.

a. Then for any u.i. norm,

$$(E_m, F) = \left(\begin{pmatrix} 0 & 0 \\ 0 & 1/m \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \right) \in \mathcal{P}(A, B),$$

and $\|[E_m|F]\| \to \|[0|F]\| = \rho(A, B)$.

b. For the operator norm, $(0, B) \in \mathcal{P}(A, B)$ attains $\rho(A, B)$. For the Frobenius norm, there is no matrix in $\mathcal{P}(A, B)$ attaining $\rho(A, B)$.

Results

Theorem *Let* $\|\cdot\|$ *be a unitarily invariant norm. Suppose* A *is* $m \times n$ *and* B *is* $m \times p$*. If the* $m \times (n+p)$ *matrix* $[A|B]$ *has singular values* $s_1 \geq \cdots \geq s_{n+p}$*, then*

$$\rho(A,B) = \left\| \sum_{j=n+1}^{n+p} s_j E_{jj} \right\| .$$

Proposition *Let A be $m \times n$ and B be $m \times p$. Then $(E, F) \in \mathcal{P}(A, B)$ if and only if any one of the following holds.*

(a) *There is an $n \times p$ matrix X such that*

$$[E|F]\begin{bmatrix} -X \\ I_p \end{bmatrix} = [A|B]\begin{bmatrix} -X \\ I_p \end{bmatrix}.$$

(b) *There is an $(n+p) \times p$ matrix $Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ such that Y_2 is a $p \times p$ invertible matrix, $Y^*Y = I_p$, and $[E|F]Y = [A|B]Y$.*

Theorem *Let* $\|\cdot\|$ *be a u.i. norm,* A *be* $m \times n$*,* B *be* $m \times p$*. Then there is* $(E,F) \in \mathcal{P}(A,B)$ *attaining* $\rho(A,B)$ *if and only if there is an* $(n+p) \times p$ *matrix* $Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ *such that* Y_2 *is a* $p \times p$ *invertible matrix,* $Y^*Y = I_p$*, and*

$$\|[A|B][Y|0_{n+p,n}]\| = \left\| \sum_{j=n+1}^{n+p} s_j([A|B])E_{jj} \right\|.$$

Computable criteria

Suppose $[A|B]$ has rank $\leq n$. Then $\rho(A, B) = 0$ is attainable by $(E, F) = (O, O) \in \mathcal{P}(A, B)$ if and only if $[A|B]$ and A have the same rank.

Theorem *Let* $\| \cdot \|$ *be a u.i. norm,* A *be* $m \times n$ *and* B *be* $m \times p$ *such that* $[A|B]$ *has singular values* $s_1 \geq \cdots \geq s_{n+p}$ *with* $s_n > 0$*. Suppose*

$$\rho(A,B) = \left\| \sum_{j=1}^{r} s_{n+j} E_{jj} \right\| > \left\| \sum_{j=1}^{r-1} s_{n+j} E_{jj} \right\|.$$

Let $b \leq n < d \leq n+t$ *satisfy*

$$s_b > s_{b+1} = \cdots = s_d > s_{d+1} \text{ and } s_{n+r} = \cdots = s_{n+t} > s_{n+t+1},$$

$$W = \begin{array}{cc} \begin{array}{c} n \\ p \end{array} & \begin{bmatrix} W_{11} & W_{12} & W_{13} & W_{14} \\ W_{21} & W_{22} & W_{23} & W_{24} \end{bmatrix} \\ & \begin{array}{cccc} b & d-b & n+t-d & p-t \end{array} \end{array}$$

be unitary such that

$$W[A|B]^*[A|B]W^* = \operatorname{diag}(s_1^2, \ldots, s_{n+p}^2).$$

Then there is $(E,F) \in \mathcal{P}(A,B)$ *attaining* $\rho(A,B)$ *if and only if* W_{23} *has rank* $n+t-d$ *and* $[W_{22}|W_{23}]$ *has rank at least* t*.*

Suppose $t = p$. (This happens $\|\cdot\|$ is the Frobenius norm or any strictly convex u.i. norm.) Then W_{14} and W_{24} are vacuous, and condition (b) reduces to the condition that W_{23} has rank $n + p - d$ and $[W_{22}|W_{23}]$ has rank p.

Corollary *Let* $t = p$. *If* $(E, F) \in \mathcal{P}(A, B)$ *attains* $\rho(A, B)$, *then* $\rho(A, B)$ *is attained by* (E, F) *for any other u.i. norm.*

Suppose $\| \cdot \|$ is the spectral norm. Then W_{13} and W_{23} are vacuous, and condition (b) reduces to the condition that W_{22} has rank at least $t = d - n$.

Corollary *If $(E, F) \in \mathcal{P}(A, B)$ satisfies $\|[E|F]\| = \rho(A, B)$ for a given u.i. norm $\| \cdot \|$, then (E, F) also attains $\rho(A, B)$ for the spectral norm.*

Further research

Applications?

Efficient computation algorithms?

Obtain results for other norms.

Proposition *Suppose* $\|\cdot\|$ *is a (any!) norm,* A *is* $m \times n$*,* B *is* $m \times p$*. Then* $\rho(A, B) = 0$ *is attainable* (by $(E, F) = (0, 0)$) *if and only if* $\operatorname{rank}([A|B]) = \operatorname{rank}(A)$.

THANK YOU FOR YOUR ATTENTION!