Block-structured distance to infeasibility

Javier Peña
Carnegie Mellon

IMA
Minneapolis
March 18, 2003

Preamble: distance to singularity

Theorem 1 (Eckart and Young) *Assume $A \in \mathbb{R}^{n \times n}$ is non-singular. Then*

$$\text{dist}(A, \text{Sing}) = \frac{1}{\|A^{-1}\|} = \sigma_{\text{min}}(A),$$

where

$$\text{dist}(A, \text{Sing}) := \inf\{\|\Delta A\| : A + \Delta A \text{ is singular}\},$$

and

$$\sigma_{\text{min}}(A) := \sup\{\delta : \delta\mathbb{B}_{\mathbb{R}^m} \subseteq A\mathbb{B}_{\mathbb{R}^n}\}$$

($\mathbb{B}_{\mathbb{R}^d} := \{x \in \mathbb{R}^d : \|x\| \le 1\}$.)

Square matrices and linear equations

Let $A \in \mathbb{R}^{n\times n}$. Notice:

$$
\begin{aligned}
A \text{ non-singular} &\Leftrightarrow A \text{ surjective} \\
&\Leftrightarrow \{Ax : x \in \mathbb{R}^n\} = \mathbb{R}^n \\
&\Leftrightarrow Ax = b \text{ feasible for all } b \in \mathbb{R}^n.
\end{aligned}
$$

THEME:

Extend the above to *rectangular* matrices defining a *conic system*

$$\begin{array}{l} Ax = b \\ x \in C, \end{array}$$

and for perturbations restricted to some *block structure*.

Rectangular matrices and conic systems

Assume $C \subseteq \mathbb{R}^n$ is a closed convex cone.

Definition Let $A \in \mathbb{R}^{m \times n}$.

$$A \in \mathcal{F} \;\Leftrightarrow\; \{Ax : x \in C\} = \mathbb{R}^m$$

$$\Leftrightarrow \quad \begin{array}{l} Ax = b \\ x \in C \end{array} \quad \text{feasible for all } b \in \mathbb{R}^m.$$

(Put $\mathcal{I} := \mathcal{F}^c$.)

Distance to infeasibility

E-Y like identity for conic systems:

Theorem 2 (Renegar) *Assume* $A \in \mathcal{F}$. *Then*

$$\operatorname{dist}(A, \mathcal{I}) = \rho(A),$$

where

$$\operatorname{dist}(A, \mathcal{I}) := \inf\{\|\Delta A\| : \ A + \Delta A \notin \mathcal{F}\},$$

and

$$\rho(A) := \sup\{\delta : \delta \mathbb{B}_{\mathbb{R}^m} \subseteq A(\mathbb{B}_{\mathbb{R}^n} \cap C)\}.$$

Proof of Renegar's Theorem:

Key: rank-one perturbations

Assume $A \in \mathcal{F}$.

Given $v \in \mathbb{R}^m$, let

$$\begin{array}{rlll} \phi(A,v) & := & \min\limits_x & \|x\| \\ & & \text{s.t.} & Ax = v \\ & & & x \in C. \end{array}$$

Notice:

$$\rho(A) = \frac{1}{\max_{\|v\|=1} \phi(A,v)}.$$

Given $u \in \mathbb{R}^n \setminus \{0\}$, let

$$\begin{array}{rlll} \psi(A,u) & := & \sup\limits_y & \dfrac{\|y\|}{\|u\|} \\ & & \text{s.t.} & A^\top y + u \in C^*. \end{array}$$

Proposition 1 *Assume $A \in \mathcal{F}$. Then*

$$\max_{\|v\|=1} \phi(A, v) = \max_{u \neq 0} \psi(A, u).$$

(The above max is a kind of "$\|A^{-1}\|$".)

Rephrase Theorem 2:

Theorem 3 *Assume $A \in \mathcal{F}$. Then*

$$\text{dist}(A, \mathcal{I}) = \frac{1}{\max_{u \neq 0} \psi(A, u)} = \frac{1}{\max_{\|v\|=1} \phi(A, v)}.$$

Notice:

$$A + \Delta A \notin \mathcal{F} \Leftrightarrow \exists y \neq 0 \text{ s.t. } (A + \Delta A)^{\mathsf{T}} y \in C^*,$$

and given $\theta \geq 0$

$$\theta < \psi(A, u) \Leftrightarrow$$
$$\exists y \text{ s.t. } A^{\mathsf{T}} y + u \in C^*, \ \frac{\|y\|}{\|u\|} > \theta.$$

Proof of Thm 3.
"$\leq$": Given $0 \leq \theta < \psi(A, u)$, let y be as above. Then

$$A + \frac{1}{\|y\|^2} y u^{\mathsf{T}} \notin \mathcal{F}$$

with

$$\left\| \frac{1}{\|y\|^2} y u^{\mathsf{T}} \right\| = \frac{\|u\|}{\|y\|} < \frac{1}{\theta}.$$

"$\geq$": Proceed in reverse order.

Block-structured distance to infeasibility

Simple block:

Suppose $\mathbb{R}^n = X_1 \times X_2$, and only allow perturbations in the set

$$\Delta := \left\{ \begin{bmatrix} B_1 & 0 \end{bmatrix} \,\middle|\, B_1 : X_1 \to \mathbb{R}^m \right\}.$$

Obvious modification of proof above shows:

Theorem 4 (P.) *Assume* $A \in \mathcal{F}$. *Then*

$$\mathrm{dist}_\Delta(A,\mathcal{I}) = \frac{1}{\max\limits_{\|v\|=1} \varphi(A,v)} = \frac{1}{\max\limits_{u_1 \neq 0} \psi(A,u)}.$$

Where

$$\mathrm{dist}_\Delta(A,\mathcal{I}) := \inf\{\|B\| : B \in \Delta,\ A+B \notin \mathcal{F}\},$$

$$\begin{array}{rll} \varphi(A,v) := & \min\limits_{x} & \|x_1\| \\ & \text{s.t.} & Ax = v \\ & & x \in C, \end{array}$$

and

$$\begin{array}{rll} \psi(A,u) := & \sup\limits_{y} & \dfrac{\|y\|}{\|u_1\|} \\ & \text{s.t.} & A^\top y + u \in C^*. \end{array}$$

Multiple blocks:

Suppose $\mathbb{R}^n = X_1 \times X_2$, $Y_1, Y_2 \subseteq \mathbb{R}^m$, and only allow perturbations in the set

$$\Delta := \left\{ \begin{bmatrix} B_1 & B_2 \end{bmatrix} \,\middle|\, B_j : X_j \to Y_j,\ j = 1, 2 \right\}.$$

For $B = \begin{bmatrix} B_1 & B_2 \end{bmatrix} \in \Delta$, define

$$\|B\|_\Delta := \max\{\|B_1\|, \|B_2\|\},$$

and

$$\text{dist}_\Delta(A, \mathcal{I}) := \inf\{\|B\|_\Delta : B \in \Delta,\ A + B \notin \mathcal{F}\}.$$

E-Y like characterization for $\text{dist}_\Delta(A, \mathcal{I})$:

Given $u \in \mathbb{R}^n \setminus \{0\}$, define

$$\begin{array}{rcl} \psi(A,u) := \sup\limits_{y} & \min & \left\{ \dfrac{\|y_1\|}{\|u_1\|}, \dfrac{\|y_2\|}{\|u_2\|} \right\} \\ & \text{s.t.} & A^\top y + u \in C^*. \end{array}$$

Theorem 5 (P.) *Assume $A \in \mathcal{F}$. Then*

$$\text{dist}_\Delta(A, \mathcal{I}) = \frac{1}{\max\limits_{u \neq 0} \psi(A,u)}.$$

Proof. (Similar to the unstructured case.)
For a given $\theta \geq 0$

$$\theta < \psi(A,u) \Leftrightarrow \exists y \text{ s.t. } A^\top y + u \in C^*,\ \frac{\|y_j\|}{\|u_j\|} > \theta.$$

"$\leq$": Given $u \neq 0$ and $0 \leq \theta < \psi(A,u)$, let y be as above. Then

$$\left[A_1 + \tfrac{1}{\|y_1\|^2} y_1 u_1^\top \quad A_2 + \tfrac{1}{\|y_2\|^2} y_2 u_2^\top\right] \notin \mathcal{F}$$

with

$$\|\frac{1}{\|y_j\|^2} y_j u_j^\top\| = \frac{\|u_j\|}{\|y_j\|} \leq \frac{1}{\theta}.$$

"$\geq$": Proceed in reverse order.

General block structure:

Suppose $X_j \subseteq \mathbb{R}^n$, $Y_j \subseteq \mathbb{R}^m$, $j = 1, \ldots, k$ are given.

Let

$$\Delta := \left\{ \left[B_j \right] \,\middle|\, B_j : X_j \to Y_j,\ j = 1, \ldots, k \right\}.$$

(Notation: $\left[B_j \right] x := \sum_{j=1}^{k} B_j x_j$.)

For $B = \left[B_j \right] \in \Delta$, let

$$\|B\|_\Delta := \max\{\|B_j\|, j = 1, \ldots, k\}.$$

Given $u_j \in X_j$ not all zero, put $U := \begin{bmatrix} u_1 \ldots u_k \end{bmatrix}$ and let

$$\psi(A,U) := \begin{array}{ll} \sup\limits_{y} & \min\limits_{j} \dfrac{\|y_j\|}{\|u_j\|} \\ \text{s.t.} & A^\mathsf{T} y + Ue \in C^*. \end{array}$$

Theorem 6 (P.) *Assume $A \in \mathcal{F}$. Then*

$$\text{dist}_\triangle(A,\mathcal{I}) = \frac{1}{\max\limits_{u_j \in X_j,\, U \neq 0} \psi(A,U)}.$$

Can also consider the dual counterpart:

Given $v_j \in Y_j$, put $V := \begin{bmatrix} v_1 & \ldots & v_k \end{bmatrix}$ and let

$$\begin{array}{rll} \varphi(A,V) := & \inf\limits_{x,z} & \max\limits_j \dfrac{\|x_j\|}{z_j} \\ & \text{s.t.} & Ax = Vz \\ & & x \in C \\ & & z > 0. \end{array}$$

Proposition 2 *Assume $A \in \mathcal{F}$. Then*

$$\max_{v_j \in Y_j, \|v_j\|=1} \varphi(A,V) = \max_{u_j \in X_j, U \neq 0} \psi(A,U).$$

(Again, these max are kind of "$\|A^{-1}\|_\triangle$".)

Thus, have a "dual" version of Thm 6:

Theorem 7 *Assume $A \in \mathcal{F}$. Then*

$$\text{dist}_\triangle(A,\mathcal{I}) = \frac{1}{\max\limits_{v_j \in Y_j, \|v_j\|=1} \varphi(A,V)}.$$

<u>Special cases:</u>

Componentwise distance to infeasibility.

Suppose $E \subseteq \{0,1\}^{m\times n}$ determines some sparsity structure, e.g.,

$$\begin{bmatrix} \mathsf{X} & \mathsf{X} & 0 & 0 & \mathsf{X} \\ 0 & \mathsf{X} & \mathsf{X} & 0 & 0 \\ 0 & 0 & \mathsf{X} & \mathsf{X} & 0 \\ \mathsf{X} & 0 & 0 & \mathsf{X} & 0 \end{bmatrix}.$$

Consider the 1×1 blocks defined by E.
In this case $\text{dist}_\Delta(\cdot,\mathcal{I}) = \text{dist}_E(\cdot,\mathcal{I})$, where

$$\text{dist}_E(A,\mathcal{I}) := \inf\{\delta : \exists \Delta A \text{ with } |\Delta A| \leq \delta\, E \\ \text{s.t. } A + \Delta A \notin \mathcal{F}\}.$$

In this case Thm 7 can be sharpened:

Given $B \in \mathbb{R}^{m\times n}$, let

$$\begin{aligned} \Phi(A,B) := \inf_{x,z} \quad & \max_{j=1\ldots n} \frac{|x_j|}{z_j} \\ \text{s.t.} \quad & Ax = Bz \\ & x \in C \\ & z > 0. \end{aligned}$$

Theorem 8 (P.) *Assume* $A \in \mathcal{F}$. *Then*

$$\begin{aligned} \text{dist}_E(A,\mathcal{I}) &= \frac{1}{\max_{|B|=E} \Phi(A,B)} \\ &= \frac{1}{\max_S \Phi(A,SE)}, \end{aligned}$$

max taken over signature *matrices:*
S *is a signature matrix iff* $|S| = I$.

Componentwise distance to singularity:

When $m = n$ and $C = \mathbb{R}^n$:

$$\begin{aligned} \text{dist}_E(A, \mathcal{I}) &= \text{dist}_E(A, \text{Sing}) \\ &:= \inf\{\delta : \exists \Delta A \text{ with } |\Delta A| \le \delta\, E \\ &\qquad\qquad \text{s.t. } A + \Delta A \in \text{Sing}\}. \end{aligned}$$

On the other hand, have a connection between distance to singularity and eigenvalues:

Example. Assume $A \in \mathbb{R}^{n\times n}$ is non-singular and $B \in \mathbb{R}^{n\times n}$. Then

$$\inf\{|\delta| : A + \delta B \in \text{ Sing}\} = \frac{1}{\rho_0(A^{-1}B)}.$$

$\rho_0(\cdot)$ is the *real spectral radius:*

$\rho_0(M) := \max\{|\lambda| : \lambda$ is a *real* eigenvalue of $M\}$.

(If M has no real eigenvalues, $\rho_0(M) := 0$.)

Is there a connection between $\text{dist}_E(A, \text{Sing})$ and eigenvalues?

Theorem 9 (Rohn) *Assume $A \in \mathbb{R}^{n\times n}$ is non-singular. Then*

$$\text{dist}_E(A, \text{Sing}) = \frac{1}{\max_{S_1,S_2} \rho_0(A^{-1}S_1ES_2)},$$

max taken over signature matrices.

Can recover Thm 9 from Thm 8.

Key step:

Theorem 10 (Rump) *Assume* $M \in \mathbb{R}^{n \times n}$. *Then*

$$\begin{aligned} \max_S \rho_0(MS) &= \max_{x \neq 0} \min_{x_i \neq 0} \frac{|(Mx)_i|}{|x_i|} \\ &= \max_S \inf_{z > 0} \max_i \frac{|(MSz)_i|}{z_i} \\ &= \max_S \Phi(I, MS). \end{aligned}$$

(Can be shown via LP duality.)

Thus,

$$\begin{aligned} \max_S \Phi(A, SE) &= \max_{S_1, S_2} \Phi(I, A^{-1} S_1 E S_2) \\ &= \max_{S_1, S_2} \rho_0(A^{-1} S_1 E S_2). \end{aligned}$$

Hence Thm 9 follows from Thm 8.

The structured singular value

Assume $n = m$, $\mathbb{R}^n = X_1 \times \cdots \times X_k$, and $C = \mathbb{R}^n$.

Let

$$\Delta := \left\{ \left[B_j \right] \middle| B_j : X_j \to X_j \right\}.$$

(Δ : a diagonal block-structure with $Y_j = X_j$.)

Definition (Doyle)

$$\mu_\Delta(M) := \frac{1}{\inf\{\|B\| : B \in \Delta, \det(I - MB) = 0\}}$$

μ_Δ: important parameter in robust control (Doyle et al.)

Notice: For M non-singular

$$\frac{1}{\mu_\Delta(M)} = \inf\{\|B\|_\Delta : B \in \Delta, M^{-1} - B \in \text{Sing}\}$$
$$= \text{dist}_\Delta(M^{-1}, \mathcal{I}).$$

Hence by Thm 7,

$$\mu_\Delta(M) = \max\{\varphi(M^{-1}, V) : v_j \in X_j, \|v_j\| = 1\}$$
$$= \max\{\varphi(I, MV) : v_j \in X_j, \|v_j\| = 1\}.$$

Connection with real-spectral radius:

$$\mu_\Delta(M) = \max\{\rho_0(MB) : B \in \Delta, \|B\| = 1\}$$
$$= \max\{\varphi(I, MV) : v_j \in X_j, \|v_j\| = 1\}.$$

(This holds for all $M \in \mathbb{R}^{n\times n}$.)

A bit more generality

Weighted blocks

Suppose X_j, Y_j and $\alpha_j > 0$ are given.
For $B \in \Delta$ let the α-weighted norm be

$$\|B\|_{\Delta,\alpha} := \max\{\|B_j\|/\alpha_j\}.$$

For this norm have:

$$\mathrm{dist}_{\Delta,\alpha}(A, \mathcal{I}) = \frac{1}{\max\limits_{v_j \in Y_j,\, \|v_j\| = \alpha_j} \varphi(A, V)}.$$

General conic systems

Suppose $C_X \subseteq \mathbb{R}^n$, $C_Y \subseteq \mathbb{R}^m$. Consider

$$\begin{aligned} & b - Ax \in C_Y \\ & x \in C_X. \end{aligned}$$

Just rewrite as

$$\begin{aligned} & \begin{bmatrix} A & I \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} = b \\ & \begin{bmatrix} x \\ s \end{bmatrix} \in C_X \times C_Y, \end{aligned}$$

and declare identity block not subject to perturbations.

Summary of E-Y like identities

E-Y:

$$\text{dist}(A, \text{Sing}) = \frac{1}{\|A^{-1}\|} = \sigma_{\text{min}}(A),$$

Renegar:

$$\text{dist}(A, \mathcal{I}) = \rho(A) = \frac{1}{\max_{\|v\|=1} \phi(A, v)},$$

Rohn:

$$\text{dist}_E(A, \text{Sing}) = \frac{1}{\max_{S_1, S_2} \rho_0(A^{-1} S_1 E S_2)},$$

P:

$$\text{dist}_\Delta(A, \mathcal{I}) = \frac{1}{\max_{v_j \in Y_j,\, \|v_j\|=1} \varphi(A, V)}.$$

(Δ: block-structure defined by X_j, Y_j.)

E-Y identity for set-valued maps

Let X, Y be Banach spaces.

Consider a set-valued map $F : X \rightrightarrows Y$.

Define graph(F) $\subseteq X \times Y$, $F^{-1} : Y \rightrightarrows X$, and $\|F\|$ as follows

$$\text{graph}(F) := \{(x, y) : y \in F(x)\},$$

$$x \in F^{-1}(y) \Leftrightarrow y \in F(x),$$

$$\|F\| = \sup_{x \in \mathbb{B}_X} \inf_{y \in F(x)} \|y\|,$$

$\mathbb{B}_X$: unit ball in X.

A set-valued map F is *sublinear* if graph(F) is a convex cone.

(Sublinear maps a.k.a. *convex processes*.)

Theorem 11 (Lewis) *Assume F is a sublinear map with closed graph. Then*

$$\frac{1}{\|F^{-1}\|} = \inf\{\|\Delta F\| : \Delta F \in L(X,Y),\\ F + \Delta F \text{ not surjective}\}.$$

Thm 2 (Renegar): special case of Thm 11.

Dontchev, Lewis, Rockafellar: E-Y identity for the radius of metric regularity (distance to metric irregularity)

What is next?

E-Y

$A \in \mathbb{R}^{n \times n}$

$\text{dist}(A, \text{Sing})$

Rohn

$A \in \mathbb{R}^{n \times n}$

$\text{dist}_E(A, \text{Sing})$

Renegar

$A \in \mathbb{R}^{m \times n}$

$\text{dist}(A, \mathcal{I})$

Lewis

$F : X \rightrightarrows Y$

$\text{dist}(F, \text{non-Surj})$

P.

$A \in \mathbb{R}^{m \times n}$

$\text{dist}_\Delta(A, \mathcal{I})$

P.

$F : X \rightrightarrows Y$

$\text{dist}_\Delta(F, \text{non-Surj})$

??

Block-structured distance to non-surjectivity:

Suppose $X_j \subseteq X$, $Y_j \subseteq Y$, $j = 1, \ldots, k$ are given.

Let

$$\Delta := \left\{ \left[B_j \right] : B_j \in L(X_j, Y_j),\ j = 1, \ldots, k \right\}.$$

For $B = \left[B_j \right] \in \Delta$, let

$$\|B\|_\Delta := \max\{\|B_j\|, j = 1, \ldots, k\}.$$

Theorem 12 (P.) *Assume F is a sublinear map with closed graph. Then*

$$\frac{1}{\|F^{-1}\|_\Delta} = \inf\{\|B\|_\Delta : B \in \Delta, \quad F + B \text{ not surjective}\}.$$

(For suitable $\|F^{-1}\|_\Delta$.)

Concluding remarks

– Extensions of the Eckart and Young identity hold for rectangular matrices, conic systems, and block-structured perturbations.

– Proof technique: low-rank construction.

– Work in progress: block-structured
- dist. to non-surj. (sublinear maps)
- radius of metric reg. (set-valued maps)

– Future work:
- Other types of structure (e.g., Toeplitz).
- Other norms to measure perturbations (e.g., Frobenius).

– Acknowledgments:
Raphael Hauser, Oxford University.