Resolvent Averages of Monotone Operators: Dominant and Recessive Properties

Xianfu Wang

Department of Mathematics
University of British Columbia Okanagan

Terryfest 2015
International Conference on Variational Analysis, Optimization and Quantitative Finance
Université de Limoges, France: May 18-22, 2015

Joint work with Sedi Bartz, Heinz Bauschke and Sarah Moffat

Abstract

Monotone operators play important roles in optimization and convex analysis. We define a new average of monotone operators by using resolvents. The new average enjoys self-duality and inherits many nice features of given monotone operators. When the monotone operators are positive definite matrices, the new average lies between the harmonic average and arithmetic average. Appropriate limits of resolvent average lead to both harmonic average and arithmetic average. Consequences on matrix functions are also given.

Outline of Topics

1. Resolvent averages
2. Dominant properties of $\mathcal{R}_\mu(A, \lambda)$
   - At most single-valued or strict monotonicity
   - Uniform monotonicity
3. Recessive properties of $\mathcal{R}_\mu(A, \lambda)$
   - Paramonotonicity and rectangularity
   - Nonexpansive monotonicity and displacement mapping
4. Graphical limits of resolvent averages
5. Extensions and relationships

1. Resolvent averages
2. Dominant properties of $\mathcal{R}_\mu(A,\lambda)$
   - At most single-valued or strict monotonicity
   - Uniform monotonicity
3. Recessive properties of $\mathcal{R}_\mu(A,\lambda)$
   - Paramonotonicity and rectangularity
   - Nonexpansive monotonicity and displacement mapping
4. Graphical limits of resolvent averages
5. Extensions and relationships

What is a resolvent average?

$\mathcal{H}$: a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$.
For a set-valued operator $A : \mathcal{H} \rightrightarrows \mathcal{H}$,

$$\operatorname{dom} A = \{x |\ Ax \neq \emptyset\}, \quad \operatorname{ran} A = \bigcup_{x \in \operatorname{dom} A} Ax,$$

The set-valued inverse A^{-1} of A:

$$(y, x) \in \operatorname{gra} A^{-1} \ \Leftrightarrow \ (x, y) \in \operatorname{gra} A.$$

The operator A is called monotone if $\forall (x_i, x_i^*) \in \operatorname{gra} A,\ i = 1, 2,$

$$\langle x_2^* - x_1^*, x_2 - x_1 \rangle \geq 0,$$

and strictly monotone if this inequality is strict whenever $x_1 \neq x_2$.

$A : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone if the monotone set gra A is not properly contained in any other monotone set.

$\mathrm{Id} : \mathcal{H} \to \mathcal{H}$ denotes the identity mapping.

For $\lambda > 0$,

$$J_A = (\mathrm{Id} + A)^{-1} : \quad \text{resolvent of } A,$$

$$^{\lambda}A = \lambda^{-1}(\mathrm{Id} - J_{\lambda A}) : \quad \text{Yosida } \lambda\text{-regularization},$$

$$S_A = \mathrm{Id} - 2(\mathrm{Id} + A)^{-1} : \quad \text{Rockafellar-Wets regularization of } A.$$

Resolvent average

For monotone operators $A_i, i = 1, \ldots, n$ and $\lambda_i > 0$ with $\sum_{i=1}^n \lambda_i = 1$, define

$$A = (A_1, A_2, \ldots, A_n),$$

$$\lambda = (\lambda_1, \ldots, \lambda_n).$$

Resolvent average of $A_i, i = 1, \ldots, n$

$$\mathcal{R}_\mu(A, \lambda) = \left[\lambda_1(A_1 + \mu^{-1}\mathrm{Id})^{-1} + \cdots + \lambda_n(A_n + \mu^{-1}\mathrm{Id})^{-1}\right]^{-1} - \mu^{-1}\mathrm{Id},$$

Question: What does it mean?

$$\left(\mathcal{R}_\mu(A,\lambda)+\mu^{-1}\mathrm{Id}\right)^{-1}=\lambda_1\left(A_1+\mu^{-1}\mathrm{Id}\right)^{-1}+\cdots+\lambda_n\left(A_n+\mu^{-1}\mathrm{Id}\right)^{-1},$$

which is equivalent to

$$J_{\mu\mathcal{R}_\mu(A,\lambda)}=\lambda_1 J_{\mu A_1}+\cdots+\lambda_n J_{\mu A_n}. \tag{1}$$

Harmonic average and arithmetic average

Well-known harmonic average and arithmetic average are

$$\mathcal{H}(A,\lambda) = (\lambda_1 A_1^{-1} + \cdots + \lambda_n A_n^{-1})^{-1},$$

$$\mathcal{A}(A,\lambda) = \lambda_1 A_1 + \cdots + \lambda_n A_n.$$

- If $\bigcap_{i\in I} \operatorname{ran} A_i = \varnothing$, then $\mathcal{H}(A,\lambda)$ is empty-valued.
- If $\bigcap_{i\in I} \operatorname{dom} A_i = \varnothing$, then $\mathcal{A}(A,\lambda)$ is empty-valued.

Why study resolvent averages?

- $\mathcal{R}_\mu(A, \lambda)$ provides a novel method to obtain set-valued operators from known operators A_i's.
- It is very interesting to ask what properties $\mathcal{R}_\mu(A, \lambda)$ has and inherits from those A_i's.
- What are the relationships among $\mathcal{R}_\mu(A, \lambda)$, $\mathcal{H}(A, \lambda)$ and $\mathcal{A}(A, \lambda)$?

Proximal average of functions: BGLW'08

The proximal average of convex functions is defined by

$$\mathcal{P}_\mu(f,\lambda) = \left[\lambda_1(f_1 + \mu^{-1}j)^* + \cdots + \lambda_n(f_n + \mu^{-1}j)^*\right]^* - \mu^{-1}j, \qquad (2)$$

for $f = (f_1, \ldots, f_n)$ with $(\forall i)$ f_i being convex functions.

The proximal mapping of $\mathcal{P}_\mu(f,\lambda)$ is the average of proximal mappings of f_i's, namely

$$\operatorname{Prox}_\mu \mathcal{R}_\mu(A,\lambda) = \lambda_1 \operatorname{Prox}_\mu f_1 + \cdots + \lambda_n \operatorname{Prox}_\mu f_n,$$

with the proximal mapping $\operatorname{Prox}_\mu f_i = (\mu\partial f_i + \mathrm{Id})^{-1}$.

Reformulations of $\mathcal{R}_\mu(A, \lambda)$

Proposition 1

We have

$$ {}^{\mu}[\mathcal{R}_\mu(A,\lambda)] = \lambda_1 \, {}^{\mu}A_1 + \cdots + \lambda_n \, {}^{\mu}A_n. \tag{3} $$

$$ S_{\mu\mathcal{R}_\mu(A,\lambda)} = \lambda_1 S_{\mu A_1} + \cdots + \lambda_n S_{\mu A_n}. \tag{4} $$

In terms of Yosida μ-regularization of A_i's, we have

Theorem 2

$$ \mathcal{R}_\mu(A,\lambda) = - \, {}^{\mu}\big[- (\lambda_1 \, {}^{\mu}A_1 + \cdots + \lambda_n \, {}^{\mu}A_n) \big]. $$

Proposition 3

$x^* \in \mathcal{R}_\mu(A,\lambda)(x)$ if and only if $(\forall i)\ \exists\ x_i \in \operatorname{dom} A_i$ such that

$$\begin{cases} x = \lambda_1 x_1 + \cdots + \lambda_n x_n \\ x^* \in \bigcap_{i=1}^n (A_i + \mu^{-1}\mathrm{Id})(x_i) - \mu^{-1}x. \end{cases} \tag{5}$$

Consequently, $\forall x \in \mathbb{R}^N$,

$$\mathcal{R}_\mu(A,\lambda)(x) = \bigcup \Big\{ \bigcap_{i=1}^n (A_i + \mu^{-1}\mathrm{Id})(x_i) - \mu^{-1}x : \sum_{i=1}^n \lambda_i x_i = x \Big\}, \tag{6}$$

$$\operatorname{dom} \mathcal{R}_\mu(A,\lambda) \subset \lambda_1 \operatorname{dom} A_1 + \cdots + \lambda_n \operatorname{dom} A_n. \tag{7}$$

Furthermore, $\forall x \in \mathbb{R}^N, \forall \mu > 0$,

$$\bigcap_{i=1}^{n} A_i(x) \subset \mathcal{R}_\mu(A, \lambda)(x). \tag{8}$$

Proposition 4

Let $z^*, z \in \mathbb{R}^N$. Then

$$\mathcal{R}_\mu((A_1 - z^*, \ldots, A_n - z^*), \lambda) = \mathcal{R}_\mu(A, \lambda) - z^*, \tag{9}$$

$$\mathcal{R}_\mu((A_1(\cdot - z), \ldots, A_n(\cdot - z)), \lambda) = \mathcal{R}_\mu(A, \lambda)(\cdot - z). \tag{10}$$

Proposition 5

Let $\alpha > 0$. Then

$$\mathcal{R}_\mu(\alpha A, \lambda) = \alpha \mathcal{R}_{\alpha\mu}(A, \lambda). \tag{11}$$

In particular,

$$\mathcal{R}_\mu(A, \lambda) = \mu^{-1}\mathcal{R}_1(\mu A, \lambda). \tag{12}$$

Example 6

Let $A_i = N_{C_i}$ $\forall i$. Then

$$\mathcal{R}_\mu(A,\lambda) = \mu^{-1}\big[(\lambda_1 P_{C_1} + \cdots + \lambda_n P_{C_n})^{-1} - \mathrm{Id}\big], \tag{13}$$

Inverse formula

Fact 7 (Poliquin-Rockafellar'96)

Every mapping $A : \mathcal{H} \rightrightarrows \mathcal{H}$ *obeys the identity*

$$\mathrm{Id} - (\mathrm{Id} + A)^{-1} = (\mathrm{Id} + A^{-1})^{-1}. \quad (14)$$

Indeed, the Yosida regularizations of A *are related to the resolvents of* A *by*

$${}^{\mu}A = (\mu\mathrm{Id} + A^{-1})^{-1} = \mu^{-1}[\mathrm{Id} - (\mathrm{Id} + \mu A)^{-1}] \quad \forall \mu > 0. \quad (15)$$

Theorem 8 (inverse formula: self duality)

Let $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ be any set-valued mapping and $\mu > 0$. Assume that $\sum_{i=1}^n \lambda_i = 1$ with $\lambda_i > 0$. Then

$$\left[\mathcal{R}_\mu(A,\lambda)\right]^{-1} = \mathcal{R}_{\mu^{-1}}(A^{-1},\lambda), \text{ i.e.,} \tag{16}$$

$$\left[\left(\lambda_1(A_1+\mu^{-1}\mathrm{Id})^{-1}+\cdots+\lambda_n(A_n+\mu^{-1}\mathrm{Id})^{-1}\right)^{-1}-\mu^{-1}\mathrm{Id}\right]^{-1} =$$
$$\left(\lambda_1(A_1^{-1}+\mu\mathrm{Id})^{-1}+\cdots+\lambda_n(A_n^{-1}+\mu\mathrm{Id})^{-1}\right)^{-1}-\mu\mathrm{Id}.$$

Who cares about resolvent averages?

Theorem 9 (common solutions to monotone inclusions)

Suppose that for each $i \in I$, $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone. Let x and u be points in $\mathcal{H}$. If $\bigcap_{i\in I} A_i(x) \neq \varnothing$, then

$$\mathcal{R}_\mu(A,\lambda)(x) = \bigcap_{i\in I} A_i(x). \tag{17}$$

If $\bigcap_{i\in I} A_i^{-1}(u) \neq \varnothing$, then

$$\mathcal{R}_\mu(A,\lambda)^{-1}(u) = \bigcap_{i\in I} A_i^{-1}(u). \tag{18}$$

Example 10 (convex feasibility problem)

Let $C_i \subset \mathbb{R}^N$ be non-empty closed convex, and $A_i = N_{C_i}$. If $\bigcap_{i=1}^n C_i \neq \emptyset$, then

$$\mathcal{R}_\mu(A, \lambda)^{-1}(0) = \bigcap_{i=1}^{n} C_i.$$

Who cares about resolvent averages?

Example 11 (homotopy transform)

Let $A_1, A_2 : \mathbb{R}^N \rightrightarrows \mathbb{R}^N$ be maximal monotone operators. The mapping $(\forall \lambda \in [0,1])\ h_\lambda : \mathbb{R}^N \rightrightarrows \mathbb{R}^N$ given by

$$\mathbb{R}^N \ni x \mapsto \begin{cases} A_1 x & \text{if } \lambda = 0; \\ A_2 x & \text{if } \lambda = 1; \\ \mathcal{R}_1(A,\lambda)x & \text{if } 0 < \lambda < 1. \end{cases} \tag{19}$$

is a homotopy in the graphical convergence topology. More precisely, $\lambda \mapsto h_\lambda$ is continuous on $[0,1]$ in the graphical convergence topology.

Example 1

Define $A_1(x) = x$ and $A_2(x) = \begin{cases} -1 & \text{if } x < 0 \\ [-1,1] & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$.

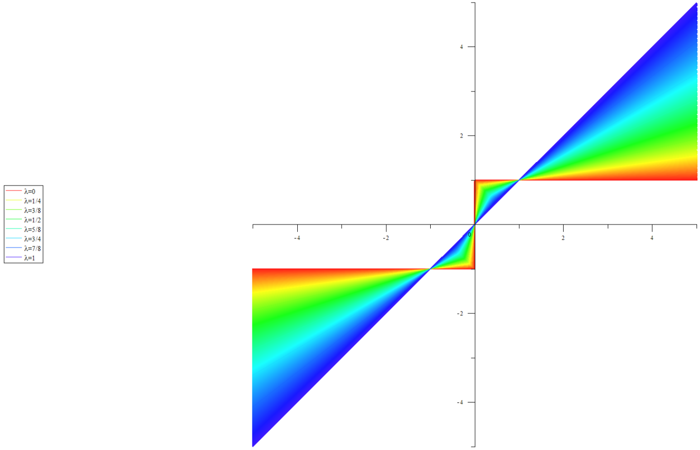
Then,

$$\lambda J_{A_1}(x) = \frac{\lambda}{2}x \text{ and } (1-\lambda)J_{A_2}(x) = \begin{cases} (1-\lambda)(x+1) & \text{if } x < -1 \\ 0 & \text{if } 1 \leq x \leq 1 \\ (1-\lambda)(x-1) & \text{if } x > 1 \end{cases}$$

Now to see the graph of $\mathcal{R}_1(A, \lambda)$, we use the Minty Parameterization,

$$(J_{\mathcal{R}(A,\lambda)}, x - J_{\mathcal{R}(A,\lambda)})$$

Example 1 cont.



Example 2

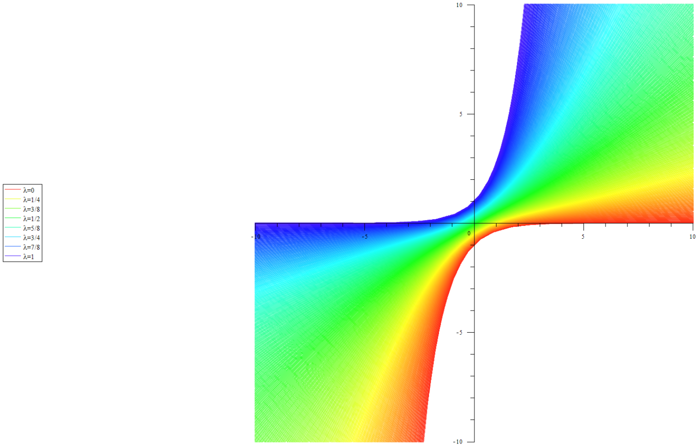
Define $A_1(x) = e^x$ and $A_2(x) = -e^{-x}$. Solving for J_{A_1}, J_{A_2},

$$J_{A_1}(x) = -\mathrm{W}(e^x) + x \text{ and } J_{A_2}(x) = \mathrm{W}(\frac{1}{e^x}) + x$$

Where W is the Lambert W function. Now to see the graph of $\mathcal{R}_1(A, \lambda)$ we use the Minty Parameterizations,

$$(J_{\mathcal{R}(A,\lambda)}, x - J_{\mathcal{R}(A,\lambda)})$$

Example 2 cont.



Inheritance of properties

Definition 12

We say that a property (p) is

1. **dominant** if the existence of $i_0 \in I$ such that A_{i_0} has property (p) implies that $\mathcal{R}_\mu(A, \lambda)$ has property (p);
2. **recessive** if (p) is not dominant, and for all $i \in I$, A_i having property (p) implies that $\mathcal{R}_\mu(A, \lambda)$ has property (p).
3. **indeterminate** if (p) is neither dominant nor recessive.

1. Resolvent averages

2. Dominant properties of $\mathcal{R}_\mu(A,\lambda)$
   - At most single-valued or strict monotonicity
   - Uniform monotonicity

3. Recessive properties of $\mathcal{R}_\mu(A,\lambda)$
   - Paramonotonicity and rectangularity
   - Nonexpansive monotonicity and displacement mapping

4. Graphical limits of resolvent averages

5. Extensions and relationships

What have we gained?

Theorem 13

Let $A_1, \cdots, A_n : \mathbb{R}^N \rightrightarrows \mathbb{R}^N$ be monotone. Then $\mathcal{R}_\mu(A, \lambda)$ is monotone. Moreover,

$$\operatorname{dom} J_{\mu\mathcal{R}_\mu(A,\lambda)} = \operatorname{dom} J_{\mu A_1} \cap \cdots \cap \operatorname{dom} J_{\mu A_n}, \quad \textit{i.e.}, \tag{20}$$

Consequently, $\mathcal{R}_\mu(A, \lambda)$ is maximal monotone if and only if $(\forall i)$ A_i is maximal monotone.

Theorem 14 (domain and range of $\mathcal{R}_\mu$)

Suppose that for each $i \in I$, $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone. Then

$$\operatorname{ran} \mathcal{R}_\mu(A, \lambda) \simeq \sum_{i \in I} \lambda_i \operatorname{ran} A_i, \quad \operatorname{dom} \mathcal{R}_\mu(A, \lambda) \simeq \sum_{i \in I} \lambda_i \operatorname{dom} A_i. \tag{21}$$

Theorem 15 (nonempty interior of the domain, fullness of the domain and surjectivity are dominant)

Suppose that for each $i \in I$, $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone.

1. *If there exists $i_0 \in I$ such that $\operatorname{int}\operatorname{dom} A_{i_0} \neq \varnothing$, then $\operatorname{int}\operatorname{dom} \mathcal{R}_\mu(A, \lambda) \neq \varnothing$.*
2. *If there exists $i_0 \in I$ such that $\operatorname{dom} A_{i_0} = \mathcal{H}$, then $\operatorname{dom} \mathcal{R}_\mu(A, \lambda) = \mathcal{H}$.*
3. *If there exists $i_0 \in I$ such that A_{i_0} is surjective, then $\mathcal{R}_\mu(A, \lambda)$ is surjective.*

Convex combinations of NE or FNE

Lemma 16

Suppose that for each $i \in I$, $N_i : \mathcal{H} \to \mathcal{H}$ is nonexpansive, $T_i : \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive and set $N = \sum_{i\in I} \lambda_i N_i$ and $T = \sum_{i\in I} \lambda_i T_i$. Let x and y be points in $\mathcal{H}$ such that $\|Tx - Ty\|^2 = \langle x - y, Tx - Ty\rangle$. Then $T_i x - T_i y = Tx - Ty$ for every $i \in I$. As a consequence, the following assertions hold:

1. *If there exits $i_0 \in I$ such that T_{i_0} is injective, then T is injective.*
2. *If x and y are points in $\mathcal{H}$ such that $\|Nx - Ny\| = \|x - y\|$, then $N_i x - N_i y = Nx - Ny$ for every $i \in I$.*
3. *(Reich' 83) If $\bigcap_{i\in I} \operatorname{Fix} N_i \neq \varnothing$, then $\operatorname{Fix} N = \bigcap_{i\in I} \operatorname{Fix} N_i$.*

- T_{i_0} being injective: $T_{i_0}(x) = T_{i_0} y \quad \Rightarrow \quad x = y$.

Lemma 17

Suppose that for each $i \in I$, $T_i : \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive and set $T = \sum_{i\in I} \lambda_i T_i$. If there exits $i_0 \in I$ such that

$$T_{i_0}x \neq T_{i_0}y \quad \Rightarrow \quad \|T_{i_0}x - T_{i_0}y\|^2 < \langle x - y, T_{i_0}x - T_{i_0}y\rangle, \tag{22}$$

then T has property (22) as well.

Outline

1. Resolvent averages
2. Dominant properties of $\mathcal{R}_\mu(A, \lambda)$
   - At most single-valued or strict monotonicity
   - Uniform monotonicity
3. Recessive properties of $\mathcal{R}_\mu(A, \lambda)$
   - Paramonotonicity and rectangularity
   - Nonexpansive monotonicity and displacement mapping
4. Graphical limits of resolvent averages
5. Extensions and relationships

Lemma 16 gives:

Theorem 18

1. *Assume that some A_i is at most single-valued. Then $\mathcal{R}_\mu(A, \lambda)$ is also at most single-valued.*
2. *Assume that some A_{i_0} is strictly monotone. Then $\mathcal{R}_\mu(A, \lambda)$ is also strictly monotone.*

- A mapping $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is at most single-valued if for every $x \in \mathcal{H}$, Ax is either empty or a singleton.
- A mapping $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is said to be strictly monotone if whenever $u \in Ax$ and $v \in Ay$ are such that $x \neq y$, then $0 < \langle u - v, x - y \rangle$.

Outline

1. Resolvent averages
2. Dominant properties of $\mathcal{R}_\mu(A, \lambda)$
   - At most single-valued or strict monotonicity
   - Uniform monotonicity
3. Recessive properties of $\mathcal{R}_\mu(A, \lambda)$
   - Paramonotonicity and rectangularity
   - Nonexpansive monotonicity and displacement mapping
4. Graphical limits of resolvent averages
5. Extensions and relationships

Uniform monotonicity and uniform FNE

Definition 19

A mapping $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is monotone with modulus $\phi : [0, \infty[\to [0, \infty]$ if for every two points (x, u) and (y, v) in gra A,

$$\phi(\|x - y\|) \leq \langle u - v, x - y \rangle.$$

The mapping A is said to be uniformly monotone with modulus ϕ if $\phi(t) = 0 \Leftrightarrow t = 0$.

Definition 20

A mapping $T : \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive with modulus $\phi : [0, \infty[\to [0, \infty]$ if for every pair of points x and y in $\mathcal{H}$,

$$\|Tx - Ty\|^2 + \phi(\|Tx - Ty\|) \leq \langle Tx - Ty, x - y \rangle.$$

The mapping $T : \mathcal{H} \to \mathcal{H}$ is said to be uniformly firmly nonexpansive with modulus ϕ if $\phi(t) = 0 \Leftrightarrow t = 0$.

- A is φ-monotone $\Leftrightarrow$ J_A is φ-FNE.

Proposition 21

Suppose that for each $i \in I$, $T_i : \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive with modulus ϕ_i which is lower semicontinuous and convex and set $T = \sum_{i \in I} \lambda_i T_i$. Then T is firmly nonexpansive with modulus $\phi = p_{\frac{1}{2}}(\phi, \boldsymbol{\lambda})$ which is proper, lower semicontinuous and convex. In particular, if there exists $i_0 \in I$ such that T_{i_0} is ϕ_{i_0}-uniformly firmly nonexpansive, then T is ϕ-uniformly firmly nonexpansive.

This gives

Theorem 22 (uniform monotonicity is dominant)

Suppose that for each $i \in I$, $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone with modulus ϕ_i which is lower semicontinuous and convex. Then $\mathcal{R}_\mu(A, \lambda)$ is monotone with modulus $\phi = p_{\frac{\mu}{2}}(\phi, \boldsymbol{\lambda})$ which is lower semicontinuous and convex. In particular, if there exists $i_0 \in I$ such that A_{i_0} is ϕ_{i_0}-uniformly monotone, then $\mathcal{R}_\mu(A, \lambda)$ is ϕ-uniformly monotone.

Definition 23

A mapping $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is ϵ-monotone, where $\epsilon \geq 0$, if $A - \epsilon \mathrm{Id}$ is monotone, that is, if for any two points (x, u) and (y, v) in gra A,

$$\epsilon \|x - y\|^2 \leq \langle v - u, x - y \rangle.$$

We also say that A^{-1} is ε-cocoercive.

Suppose that for each $i \in I$, $0 \leq \alpha_i \leq \infty$ and set $\boldsymbol{\alpha} = (\alpha_1 \cdots, \alpha_n)$. Define

$$r_\mu(\boldsymbol{\alpha}, \boldsymbol{\lambda}) = \Big[\sum_{i \in I} \lambda_i (\alpha_i + \mu^{-1})^{-1} \Big]^{-1} - \mu^{-1} \quad \text{and} \quad r(\boldsymbol{\alpha}, \boldsymbol{\lambda}) = r_1(\boldsymbol{\alpha}, \boldsymbol{\lambda}). \tag{23}$$

Theorem 24 (strong monotonicity is dominant)

Suppose that for each $i \in I$, $\epsilon_i \geq 0$, $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone and ϵ_i-monotone. Then $\mathcal{R}_\mu(A, \lambda)$ is ϵ-monotone where $\epsilon = r_\mu(\boldsymbol{\epsilon}, \boldsymbol{\lambda})$. In particular, if there exists $i_0 \in I$, such that A_{i_0} is ϵ_{i_0}-strongly monotone, then $\mathcal{R}_\mu(A, \lambda)$ is ϵ-strongly monotone.

Corollary 25 (cocoerciveness is dominant)

Suppose that for each $i \in I$, $\epsilon_i \geq 0$, $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone and A_i^{-1} is ϵ_i-monotone. Then $(\mathcal{R}_\mu(A, \lambda))^{-1}$ is ϵ-monotone where $\epsilon = r_{\mu^{-1}}(\boldsymbol{\epsilon}, \boldsymbol{\lambda})$. In particular, if there exists $i_0 \in I$ such that A_{i_0} is ϵ_{i_0}-cocoercive, then $\mathcal{R}_1(A, \lambda)$ is ϵ-cocoercive.

1. Resolvent averages

2. Dominant properties of $\mathcal{R}_\mu(A, \lambda)$
   - At most single-valued or strict monotonicity
   - Uniform monotonicity

3. Recessive properties of $\mathcal{R}_\mu(A, \lambda)$
   - Paramonotonicity and rectangularity
   - Nonexpansive monotonicity and displacement mapping

4. Graphical limits of resolvent averages

5. Extensions and relationships

Example 26

Let $f = \|\cdot\|$, $A_1 = \partial f$ and $A_2 = \mathbf{0}$. Then

$$J_{A_1}x = \begin{cases} \left(1 - \frac{1}{\|x\|}\right)x, & \text{if } \|x\| > 1; \\ 0, & \text{if } \|x\| \leq 1 \end{cases}$$

and $J_{A_2} = \mathrm{Id}$. J_{A_1} is not an affine relation and J_{A_2} is linear. However, for $0 < \lambda < 1$, $\lambda_1 = \lambda$,

$$J_{\mathcal{R}_1(A,\lambda)}x = \lambda J_{A_1}x + (1-\lambda)J_{A_2}x = \begin{cases} \left(1 - \lambda\frac{1}{\|x\|}\right)x, & \text{if } \|x\| > 1; \\ (1-\lambda)x, & \text{if } \|x\| \leq 1, \end{cases}$$

is not an affine relation. Thus, $\mathcal{R}_1(A,\lambda)$ is not an affine relation.

- Example 26 says: linearity and affinity are not dominant properties w.r.t. $\mathcal{R}_\mu(A,\lambda)$.

Theorem 27 (Linearity and affinity are recessive)

Suppose that for each $i \in I$, $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is a maximally monotone linear (resp. affine) relation. Then $\mathcal{R}_\mu(A, \lambda)$ is a maximally monotone linear (resp. affine) relation.

Paramonotone and rectangularity

Definition 28 (rectangular and paramonotone mappings)

The monotone mapping $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is said to be

1. rectangular (also known as 3* monotone) if for every $x \in \operatorname{dom} A$ and every $v \in \operatorname{ran} A$ we have

$$\inf_{(z,w)\in \operatorname{gra} A} \langle v - w, x - z\rangle > -\infty, \tag{24}$$

equivalently, if

$$\operatorname{dom} A \times \operatorname{ran} A \subseteq \operatorname{dom} F_A. \tag{25}$$

2. paramonotone if whenever we have a pair of points (x, v) and (y, u) in $\operatorname{gra} A$ such that $\langle x - y, v - u\rangle = 0$, then (x, u) and (y, v) are also in $\operatorname{gra} A$.

Recall

Definition 29 (Fitzpatrick function)

With the mapping $A : \mathcal{H} \rightrightarrows \mathcal{H}$ we associate the Fitzpatrick function $F_A : \mathcal{H} \times \mathcal{H} \to \left]-\infty, +\infty\right]$, defined by

$$F_A(x, v) = \sup_{(z,w)\in \operatorname{gra} A} \big(\langle w, x\rangle + \langle v, z\rangle - \langle w, z\rangle\big), \qquad (x, v) \in \mathcal{H} \times \mathcal{H}. \tag{26}$$

Fact 30

Let $A \in \mathbb{R}^{N\times N}$ be monotone and set $A_+ = \frac{1}{2}A + \frac{1}{2}A^\intercal$. Then the following assertions are equivalent:

1. *A is paramonotone;*
2. *A is rectangular;*
3. $\operatorname{rank} A = \operatorname{rank} A_+$;
4. $\operatorname{ran} A = \operatorname{ran} A_+$.

Example 31

In $\mathbb{R}^2$, let $A_1 = N_{\mathbb{R}\times\{0\}}$. Then J_{A_1} is the projection on $\mathbb{R}\times\{0\}$. Since A_1 is a subdifferential, it is rectangular and paramonotone. Let $A_2 : \mathbb{R}^2 \to \mathbb{R}^2$ be the counterclockwise rotation by $\pi/2$. Then

$$J_{A_1} = P_{\mathbb{R}\times\{0\}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad J_{A_2} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Since

$$A_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad A_{2+} = \frac{1}{2}(A_2 + A_2^{\intercal}) = 0,$$

A_2 is neither rectangular nor paramonotone.

Letting $\lambda_1 = \lambda_2 = \frac{1}{2}$, we obtain

$$\mathcal{R}(\mathbf{A}) = (\frac{1}{2}J_{A_1} + \frac{1}{2}J_{A_2})^{-1} - \mathrm{Id} = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} \quad \text{and}$$

$$\mathcal{R}(\mathbf{A})_+ = \frac{1}{2}\big(\mathcal{R}(\mathbf{A}) + \mathcal{R}(\mathbf{A})^\intercal\big) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

By employing Fact 30 we see that $\mathcal{R}_1(A, \lambda)$ is neither rectangular nor paramonotone.

Outline

1. Resolvent averages
2. Dominant properties of $\mathcal{R}_\mu(A, \lambda)$
   - At most single-valued or strict monotonicity
   - Uniform monotonicity
3. Recessive properties of $\mathcal{R}_\mu(A, \lambda)$
   - Paramonotonicity and rectangularity
   - Nonexpansive monotonicity and displacement mapping
4. Graphical limits of resolvent averages
5. Extensions and relationships

Lemma 32 (Fitzpatrick function of $\mathcal{R}_\mu$)

Suppose that for each $i \in I$, $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone. Then

$$F_{\mu\mathcal{R}_\mu(A,\lambda)} \leq \sum_{i\in I} \lambda_i F_{\mu A_i} \quad \textit{in particular,} \quad F_{\mathcal{R}_1(A,\lambda)} \leq \sum_{i\in I} \lambda_i F_{A_i} \tag{27}$$

and

$$\sum_{i\in I} \lambda_i \operatorname{dom} F_{\mu A_i} \subseteq \operatorname{dom} F_{\mu\mathcal{R}_\mu(A,\lambda)} \quad \textit{in particular,} \quad \sum_{i\in I} \lambda_i \operatorname{dom} F_{A_i} \subseteq \operatorname{dom} F_{\mathcal{R}_1(A,\lambda)}. \tag{28}$$

Theorem 33 (rectangularity is recessive)

Suppose that for each $i \in I$, $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is rectangular and maximally monotone. Then $\mathcal{R}_\mu(A,\lambda)$ is rectangular.

To study paramonotonicity, we need:

Proposition 34

Suppose that for each $i \in I$, $T_i : \mathcal{H} \to \mathcal{H}$ is firmly nonexpansive and set $T = \sum_{i \in I} \lambda_i T_i$. Then:

1. *If for each $i \in I$, given points x and y in $\mathcal{H}$,*

$$\|T_i x - T_i y\|^2 = \langle x - y, T_i x - T_i y\rangle \quad \Rightarrow \quad \begin{cases} T_i x = T_i(T_i x + y - T_i y) \\ T_i y = T_i(T_i y + x - T_i x), \end{cases} \tag{29}$$

   then T also has property (29).

2. *If there exists $i_0 \in I$ such that T_{i_0} has property (29) and is injective, then T has property (29) and is injective.*

Theorem 35 (paramonotonicity is recessive)

Suppose that for each $i \in I$, $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximally monotone and paramonotone. Then $\mathcal{R}_\mu(A, \lambda)$ is paramonotone.

Outline

1. Resolvent averages
2. Dominant properties of $\mathcal{R}_\mu(A, \lambda)$
   - At most single-valued or strict monotonicity
   - Uniform monotonicity
3. Recessive properties of $\mathcal{R}_\mu(A, \lambda)$
   - Paramonotonicity and rectangularity
   - Nonexpansive monotonicity and displacement mapping
4. Graphical limits of resolvent averages
5. Extensions and relationships

Definition 36

The mapping $D : \mathcal{H} \to \mathcal{H}$ is said to be a displacement mapping if there exists a nonexpansive mapping $N : \mathcal{H} \to \mathcal{H}$ such that $D = \mathrm{Id} - N$, in which case D is maximally monotone.

Lemma 37

The maximally monotone mapping $A : \mathcal{H} \rightrightarrows \mathcal{H}$ is $\frac{1}{2}$-strongly monotone if and only if A^{-1} is a displacement mapping.

Theorem 38 (being a displacement mapping is recessive)

Suppose that for each $i \in I$, $A_i : \mathcal{H} \rightrightarrows \mathcal{H}$ is a displacement mapping. Then $\mathcal{R}_1(A, \lambda)$ is a displacement mapping.

Lemma 39

The maximally monotone mapping $N : \mathcal{H} \to \mathcal{H}$ is nonexpansive if and only if $N = 2J_B - \mathrm{Id}$ for a maximally monotone and nonexpansive mapping B.

Theorem 40 (nonexpansiveness is recessive)

Suppose that for each $i \in I$, $A_i : \mathcal{H} \to \mathcal{H}$ is a nonexpansive and monotone mapping. Then $\mathcal{R}_1(A, \lambda)$ is nonexpansive. Furthermore, for each $i \in I$, $A_i = 2J_{B_i} - \mathrm{Id}$ where B_i is maximally monotone, nonexpansive and $\mathcal{R}_1(A, \lambda) = 2J_B - \mathrm{Id}$ where B is the maximally monotone and nonexpansive mapping given by $B = \sum_{i \in I} \lambda_i B_i$.

Theorem 41 (within the class of nonexpansive mappings, being a Banach contraction is dominant)

Suppose that for each $i \in I$, $A_i : \mathcal{H} \to \mathcal{H}$ is nonexpansive and monotone. If there exists $i_0 \in I$ such that A_{i_0} is a Banach contraction, then $\mathcal{R}_1(A, \lambda)$ is a Banach contraction.

Theorem 42 (in $MLR(\mathcal{H})$, being $BML(\mathcal{H})$ and being $BMLI(\mathcal{H})$ are dominant)

Suppose that for each $i \in I$, $A \in MLR(\mathcal{H})$ and there exists $i_0 \in I$ such that $A_{i_0} \in BML(\mathcal{H})$. Then $\mathcal{R}_\mu(A, \lambda) \in BML(\mathcal{H})$. Furthermore:

1. *If $A_{i_0} \in BMLI(\mathcal{H})$, then $\mathcal{R}_\mu(A, \lambda) \in BMLI(\mathcal{H})$;*
2. *If A_{i_0} is paramonotone, then $\mathcal{R}_\mu(A, \lambda)$ is paramonotone.*

- $MLR(\mathcal{H})$: maximally monotone and linear relation on $\mathcal{H}$.
- $BML(\mathcal{H})$: bounded monotone linear operators.
- $BMLI(\mathcal{H})$: bounded monotone linear operators with a bounded inverse.

Indeterminate properties

1. Being a projection is indeterminate;
2. Being a normal cone operator is indeterminate.

1. Resolvent averages
2. Dominant properties of $\mathcal{R}_\mu(A, \lambda)$
   - At most single-valued or strict monotonicity
   - Uniform monotonicity
3. Recessive properties of $\mathcal{R}_\mu(A, \lambda)$
   - Paramonotonicity and rectangularity
   - Nonexpansive monotonicity and displacement mapping
4. Graphical limits of resolvent averages
5. Extensions and relationships

The graphical limits by Attouch, Rockafellar and Wets are effective for analyzing the convergence of sequences of resolvent averages.

Definition 43

For a sequence of mappings $S_k : \mathbb{R}^N \rightrightarrows \mathbb{R}^N$, we say that S_k converges graphically to S, denoted by $S_k \xrightarrow{g} S$, if

$$\limsup_k \,(\operatorname{gra} S_k) = \liminf_k \,(\operatorname{gra} S_k) = \operatorname{gra} S.$$

Equivalenlty, $\forall x \in \mathbb{R}^N$ one has

$$\bigcup_{x_k \to x} \limsup_k S_k(x_k) \subset S(x) \subset \bigcup_{x_k \to x} \liminf_k S(x_k).$$

Fact 44 (Rockafellar & Wets)

(i) *If a sequence of maximal monotone mappings* $S_k : \mathbb{R}^N \rightrightarrows \mathbb{R}^N$ *converges graphically, then the limit mapping* S *must be maximal monotone.*

(ii) *For maximal monotone mappings* S_k *and* S*, for any choice of* $\mu > 0$ *one has*

$$S_k \xrightarrow{g} S \quad \Leftrightarrow \quad (\mathrm{Id} + \mu S_k)^{-1} \xrightarrow{p} (\mathrm{Id} + \mu S)^{-1}.$$

Theorem 45

Let $(A_{i,k})_{k\in\mathbb{N}}$ be sequences of maximal monotone mappings and let $(\lambda_{i,k})_{k\in\mathbb{N}}$ and $(\mu_k)_{k\in\mathbb{N}}$ be sequences in $(0,+\infty)$ such that that $(\forall i)$ $A_{i,k} \xrightarrow{g} A_i$, $\lambda_{i,k} \to \lambda_i > 0$ and $\mu_k \to \mu > 0$. Then

$$\mathcal{R}_{\mu_k}((A_{1,k},\ldots,A_{n,k}),(\lambda_{1,k},\ldots,\lambda_{n,k})) \xrightarrow{g} \mathcal{R}_{\mu}(A,\lambda), \text{ as } k \to \infty.$$

Moreover, $\mathcal{R}_{\mu}(A,\lambda)$ is maximal monotone.

Question: What happens for $\mu \downarrow 0$ or $\mu \uparrow \infty$?

Lemma 46

Let $(A_n)_{n\in\mathbb{N}}, A$ *be linear operators from* $\mathbb{R}^N$ *to* $\mathbb{R}^N$. *If* $A_n \xrightarrow{g} A$, *then there exists* $M > 0$ *such that*

$$\|A_n\| < M \;\forall n \in \mathbb{N}, \text{ and } \|A\| < M. \tag{30}$$

Consequently, for linear operators $(A_n)_{n\in\mathbb{N}}$ *and* A *on* $\mathbb{R}^N$, *the followings are equivalent:*

(i) *graphical convergence:* $A_n \xrightarrow{g} A$;

(ii) *point-wise convergence:* $A_n \xrightarrow{p} A$;

(iii) *norm convergence:* $A_n \xrightarrow{n} A$.

Theorem 47

Let $A_i, i = 1, \ldots, k$ be invertible linear monotone operators. Assume that at least one of $A_i^{-1}, i = 1, \ldots, k$ is striclty monotone. Then

$$\mathcal{R}_\mu(A, \lambda) \to (\lambda_1 A_1^{-1} + \cdots + \lambda_n A_n^{-1})^{-1},$$

when $\mu \to \infty$.

Theorem 48

Let $A_i, i = 1, \ldots, n$ be linear and monotone operators. Then

$$\mathcal{R}_\mu(A, \lambda) \to \lambda_1 A_1 + \cdots + \lambda_n A_n,$$

when $\mu \downarrow 0$.

When $A_i = \partial f_i$:

Theorem 49

Let $f_i : X \to\,]-\infty, +\infty]$, $i = 1, \dots, n$ be proper lower semi-continuous convex functions. We have

(i) *When $\mu \downarrow 0$,*

$$\mathcal{R}_\mu(\partial f, \lambda) \xrightarrow{g} \partial(\lambda_1 f_1 + \cdots + \lambda_n f_n).$$

If, in addition, $\bigcap_{i=1}^n \mathrm{ri}(\mathrm{dom}\, \mathrm{f_i}) \neq \emptyset$, then

$$\mathcal{R}_\mu(\partial f, \lambda) \xrightarrow{g} \lambda_1 \partial f_1 + \cdots + \lambda_n \partial f_n \quad \text{when } \mu \downarrow 0;$$

(ii) When $\mu \uparrow \infty$,

$$\mathcal{R}_\mu(\partial f, \lambda) \xrightarrow{g} \partial(\lambda_1 f_1^* + \cdots + \lambda_n f_n^*)^*.$$

If, in addition, $\bigcap_{i=1}^n \mathrm{ri}(\mathrm{dom}\, \mathrm{f}_\mathrm{i}^*) \neq \emptyset$, then

$$\mathcal{R}_\mu(\partial f, \lambda) \xrightarrow{g} [\lambda_1(\partial f_1)^{-1} + \cdots + \lambda_n(\partial f_n)^{-1}]^{-1} \quad \text{when } \mu \uparrow \infty.$$

When $(\forall i)$ A_i are positive definite matrice:

Let S_+^N (resp. S_{++}^N) be the set of positive semidefinite matrices (resp. positive definite matrices). For symmetric matrices X, Y, if $X - Y \in S_+^N$ we write $X \preceq Y$.

Theorem 50

Let $A_1, \ldots, A_n \in S_{++}^n$. We have

1. $$\mathcal{H}(A, \lambda) \preceq \mathcal{R}_\mu(A, \lambda) \preceq \mathcal{A}(A, \lambda); \tag{31}$$
2. $\mathcal{R}_\mu(A, \lambda) \to \mathcal{A}(A, \lambda)$ *when* $\mu \to 0$;
3. $\mathcal{R}_\mu(A, \lambda) \to \mathcal{H}(A, \lambda)$ *when* $\mu \to \infty$.

Corollary 51

Assume that $(\forall i)\ A_i \in S^N_{++}$ *and* $\sum_{i=1}^n \lambda_i = 1$ *with* $\lambda_i > 0$. *Then*

$$(\lambda_1 A_1 + \cdots + \lambda_n A_n)^{-1} \preceq \lambda_1 A_1^{-1} + \cdots + \lambda_n A_n^{-1}.$$

Consequently, the matrix function $X \mapsto X^{-1}$ *is matrix convex on* S^N_{++}.

Corollary 52

For every $\mu > 0$, *the resolvent average matrix function* $A \mapsto \mathcal{R}_\mu(A, \lambda)$ *given by*

$$\begin{aligned}(A_1, \cdots, A_n) \mapsto [\lambda_1(A_1^{-1} + \mu^{-1}\mathrm{Id})^{-1} + \cdots + \lambda_n(A_n^{-1} + \mu^{-1}\mathrm{Id})^{-1}]^{-1}\\ - \mu^{-1}\mathrm{Id}\end{aligned}$$

is matrix concave on $S^N_{++} \times \cdots \times S^N_{++}$.

For each $\lambda = (\lambda_1, \cdots, \lambda_n)$ with $\sum_{i=1}^n \lambda_i = 1$ and $\lambda_i > 0 \ \forall i$, the harmonic average matrix function

$$(A_1, \cdots, A_n) \mapsto (\lambda_1 A_1^{-1} + \cdots + \lambda_n A_n^{-1})^{-1} \text{ is matrix concave}$$

on $S_{++}^N \times \cdots \times S_{++}^N$. Consequenlty, the harmonic average function

$$(x_1, \cdots, x_n) \mapsto \frac{1}{x_1^{-1} + \cdots + x_n^{-1}} \text{ is concave} \tag{32}$$

on $\mathbb{R}_{++} \times \cdots \times \mathbb{R}_{++}$.

1. Resolvent averages
2. Dominant properties of $\mathcal{R}_\mu(A, \lambda)$
   - At most single-valued or strict monotonicity
   - Uniform monotonicity
3. Recessive properties of $\mathcal{R}_\mu(A, \lambda)$
   - Paramonotonicity and rectangularity
   - Nonexpansive monotonicity and displacement mapping
4. Graphical limits of resolvent averages
5. Extensions and relationships

How far does this take us?

- Extensions to averaged mappings by Combettes?
- Relationships to geometric averages of matrices, variational sums of monotone operators by Attouch, Baillon & Thera?
- For general monotone operators, under what conditions

$$\mathcal{R}_\mu(A,\lambda) \xrightarrow{g} \mathcal{H}(A,\lambda) \text{ when } \mu \uparrow \infty,$$

$$\mathcal{R}_\mu(A,\lambda) \xrightarrow{g} \mathcal{A}(A,\lambda) \text{ when } \mu \downarrow 0$$

?

Thank You Very Much!

Bibliography

- H. Attouch, J.-B. Baillon and M. Théra, "Variational sum of monotone operators", *J. Convex Anal.* 1, pp. 1–29, 1994.
- S. Bartz, H. H. Bauschke, S. M. Moffat and X. Wang, "The resolvent average of monotone operators: dominant and recessive properties", preprint, `http://arxiv.org/abs/1505.02718`.
- H. H. Bauschke, R. Goebel, Y. Lucet, and X. Wang, "The proximal average: basic theory", *SIAM J. Optim.* 19, pp. 766–785, 2008.
- H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, 2011.
- R. T. Rockafellar and R. J-B Wets, *Variational Analysis*, Springer-Verlag, New York, 1998.