Monotone Operators: Introduction, Relations

5:33 AM

Introduction: Structure: $\boxed{2}$ Relations

$\boxed{3}$ Lipschitz constants, nonexpansive and contractive operators

$\boxed{4}$ Monotone operator {generalizatied idea of monotone increasing function} (e.g. subdifferential mapping)

$\boxed{5}$ Fixed point iteration algorithm {finds the fixed points of Lipschitz monotone operators}

$\boxed{6}$ Basic results for resolvent and Cayley operations.

$\boxed{7}$ Proximal point method {finds zeros of general monotone operators}

$\boxed{8}$ Operator splitting methods {better than proximal point method}

$\boxed{2}$

* Relation other names: point-to-set mapping,
  set-valued mapping,
  multi-valued function

R: Relation on $\mathbb{R}^n$ is subset on $\mathbb{R}^n \times \mathbb{R}^n$

$R = \{(x,y) : y \in R(x)\}$ // set of all pairs such that $y \in R(x)$

$R(x)$ {overloaded function notation} $= \{y \mid (x,y) \in R\}$

$R(x) = \{\text{singleton} \vee \phi\} \Rightarrow R$ {function} // in such case $y = R(x)$ is written

also called operator // though technically correct is $R(x) = \{y\}$

{Set-image notation for functions to relations} $R(S) = \bigcup_{s \in S} R(s)$

S {a set of n dimensional points, $\subseteq \mathbb{R}^n$}

অর্থাৎ কোন relation এর input হিসেবে যদি set, i.e., ∃ একাধিক গুলি element প্রবেশ করালে, then the resultant union of all such output sets for each of those points

Example:
- Empty relation, $R = \emptyset$
- Full relation, $R = \mathbb{R}^n \times \mathbb{R}^n$
- Zero relation {a function too}, $R = 0 = \{(x,0) \mid x \in \mathbb{R}^n\}$ (overloaded)
- Identity relation, $I = \{(x,x) \mid x \in \mathbb{R}^n\}$ // this is a function too
- $f\{\mathcal{D}, : C \to \mathbb{R}, C \subseteq \mathbb{R}^n\}$
  ∂f {subdifferential} $= \{(x, g_x) \in C \times C \mid \forall_{z \in C}\ f(z) \geq f(x) + g_x^T(z-x)\}$ // technically $x \in \operatorname{int} \operatorname{dom} f$ for $\partial f \neq \emptyset$

  অর্থাৎ $(x,g) \in \partial f \leftrightarrow x \in C \wedge g \in C \wedge \forall_{z \in C}\ f(z) \geq f(x) + g_x^T(z-x)$ (g_x: subgradient)

* Operations on relations: Say $R: A \to B$

- $\operatorname{dom} R = \{x \mid R(x) \neq \emptyset\} = \{a \in A \mid \exists_{b \in B}\ (a,b) \in R\}$, $\operatorname{range}(R) = \{b \in B \mid \exists_{a \in A}\ (a,b) \in R\}$, $\operatorname{inv}(R) = R^{-1} = \{(b,a) \in B \times A \mid (a,b) \in R\}$ (always exists)

Note: $\emptyset + R = \emptyset$ অর্থাৎ empty set annhilates any other set during (overloaded) set addition

- composition
  {R {relation}, S {relation}}

$R \circ S = RS$ {composition} $= \{(x,y) \mid \exists_z\ xSz \wedge zRy\}$ // Velleman notation: $xRy \overset{\text{def}}{=} (x,y) \in R$

Similar to matrix multiplication: the relation to be applied first is innermost, the second relation to be applied is the left of the first relation and so on...

by sum of relation law:

$$\text{রি}_1 + \text{রি}_2 = \{(\text{ই}, \text{আ}_1 + \text{আ}_2) \mid (\text{ই}, \text{আ}_1) \in \text{রি}_1, (\text{ই}, \text{আ}_2) \in \text{রি}_2\}$$

so,

$$\text{আ} \in (\text{রি}_1 + \text{রি}_2)(\text{ই}) \leftrightarrow (\text{ই}, \text{আ}) \in (\text{রি}_1 + \text{রি}_2) \Leftrightarrow \exists_{\text{আ}_1, \text{আ}_2}\ \underbrace{\text{আ}_1 \in \text{রি}_1(\text{ই}), \text{আ}_2 \in \text{রি}_2(\text{ই})}, \text{আ} = \text{আ}_1 + \text{আ}_2$$

$$\Leftrightarrow \exists_{\text{আ}_1, \text{আ}_2}\ \text{আ} = \text{আ}_1 + \text{আ}_2 \in \text{রি}_1(\text{ই}) + \text{রি}_2(\text{ই})$$ # By overloaded set addition

$\therefore (\text{রি}_1 + \text{রি}_2)(\text{ই}) = \text{রি}_1(\text{ই}) + \text{রি}_2(\text{ই})$ [So the overloaded sum operator for relation has additivity]

statement: overloaded sum operator for relations has additivity

$R + S$ {sum of relations} $= \{(x, y+z) \mid (x,y) \in R, (x,z) \in S\}$ (this is an overloaded '+')
$= \{(x, y+z) \mid xRy, xSz\}$ [eq: Sum of relations]

// e.g. for a function: $u = F(x)$, $v = \bar{F}(x)$ হলে $(F+\bar{F})(x) = F(x) + \bar{F}(x) = u+v$ $\therefore F + \bar{F} = \{(x, u+v) : u = F(x), v = \bar{F}(x)\}$

// $(y \in R(x), z \in S(x)) \Leftrightarrow y + z \in (R+S)(x) = R(x) + S(x)$

Zero of a relation: $0 \in R(x) \overset{\text{def}}{=} x$ is a zero of R

// extension to zero of a function: $0 = f(x) \overset{\text{def}}{=} x$ is a zero of f
this is

zero set of a relation: $R^{-1}(\{0\}) = \{x \mid (x,0) \in R\}$ অর্থাৎ set of all zeros of a relation R

// i.e., $x \in R^{-1}(\{0\}) \leftrightarrow 0 \in R(x)$

$\lambda F = \{(x, \lambda y) \mid (x,y) \in F\}$

$\lambda \bar{F} = \{(x, \lambda z) \mid (x,z) \in F\}$

$F + \bar{F} = \{(x, y+z) \mid (x,y) \in F, (x,z) \in F\}$

$\lambda F + \lambda \bar{F} = \{(x, \lambda y + \lambda z) \mid (x, \lambda y) \in \lambda F, (x, \lambda z) \in \lambda \bar{F}\}$

$\lambda(F + \bar{F}) = \{(x, \lambda(y+z)) \mid (x,y) \in F, (x,z) \in F\}$

$= \{(x, \lambda y + \lambda z) \mid (x,y) \in F, (x,z) \in F\}$

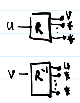
$\Big\}\ \lambda(F + \bar{F}) = \lambda F + \lambda \bar{F}$ distributive property of relations

* Example Resolvent: # Resolvent of Cayley operator is going to play an important role later on, so let's try to get used to it gradually

$R = (I + \lambda F)^{-1}$ let's find out what is this (λ: nonzero)

// Here I = identity matrix
// F is another relation / point to set mapping

$(u,v) \in R = (I + \lambda F)^{-1} \leftrightarrow (v,u) \in R^{-1} = (I + \lambda F)$



$\leftrightarrow u \in (I + \lambda F)v$ {$\because (x,y) \in R \overset{\text{def}}{=} xRy$}

$= Iv + \lambda F(v)$

$= v + \lambda F(v)$

$= v + \lambda\{\pi : vF\pi\}$

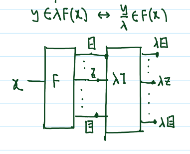
$= \{v + \lambda \pi : vF\pi\}$

$\leftrightarrow \exists_\pi (\underbrace{u = v + \lambda \pi}_{\pi = \frac{1}{\lambda}(u-v)} \wedge vF\pi)$

$\leftrightarrow vF\left(\frac{1}{\lambda}(u-v)\right)$ // $\exists_\pi$ আর দরকার নাই

$\lambda \neq 0$

$y \in \lambda F(x) \leftrightarrow \frac{y}{\lambda} \in F(x)$



y \in \lambda F(x) <=> y/\lambda \in F(x)

$\Leftrightarrow \quad \frac{1}{\lambda}(u-v) \in F(v)$ // this shows how $(u,v) \in R$ is related to the set valued mapping in F

So for any output ∈ (operator//resolvent⊡)(input) hole a scaled version of the input output difference will belong to operator(output)
% Let's make up a story suppose aliens are sending us some signal x but while it hits earth it changes into y by going throug h resolvent of some space-time operator (alien technology). All we get to see is y, now we want to find out what is the original x, now suppose constru cting F is cheap, so we take that y input it through F and find out that one element (we are in luck and F is a function) is (1/λ)(x-y), from which we can reconstruct original alien message y.

$*$ Inverse of subdifferential: $(\partial f)^{-1}$

We are going to show something amazing: we will show that if a pair belong to inverse of a subdifferential, then those pairs are tight in Younge's inequality with the input being the argument of the conjugate function and the output being the argument of the function itself.

$(u,v) \in (\partial f)^{-1} \Leftrightarrow v \in (\partial f)^{-1} u$

$\Leftrightarrow (v,u) \in \partial f$

v is some point $\quad$ u is the subgradient of f at v

$\Leftrightarrow \forall_x \; \underbrace{f(x) \geq f(v) + u^T(x-v)}$

$f(x) - u^T x \geq f(v) - u^T v$ // so v minimizes $f(x) - u^T x$ over all x

$\Leftrightarrow \quad v \in \underset{x}{\operatorname{argmin}} \, (f(x) - u^T x)$

$\therefore \; v \in (\partial f)^{-1}(u) \Leftrightarrow v \in \underset{x}{\operatorname{argmin}} \, (f(x) - u^T x)$

$(\partial f)^{-1}(u) = \underset{x}{\operatorname{argmin}} \, (f(x) - u^T x) = \underset{x}{\operatorname{argmax}} \, (u^T x - f(x))$ // $\because \; \underset{x}{\vee} f(x) = -\underset{x}{\wedge} -f(x) \; \therefore \; \underset{x}{\operatorname{argmin}} \, f(x) = \underset{x}{\operatorname{argmax}} -f(x)$

We can relate this to conjugate function f^*, where $f^*(u) = \sup_x (u^T x - f(x))$ # From the defn we can immediately obtain Younge's inequality:
$\forall_{u,x} \; f^*(u) = \sup_x (u^T x - f(x)) \geq u^T x - f(x) \Rightarrow \forall_{u,x} \; f^*(u) + f(x) \geq u^T x$

$= [u^T \boxdot - f(\boxdot)]_{\boxdot \in \underset{x}{\operatorname{argmax}} (u^T x - f(x))}$

// as $v \in \underset{x}{\operatorname{argmax}} \, (u^T x - f(x))$

$= u^T v - f(v)$

$\Leftrightarrow \; f^*(u) + f(v) = u^T v$

$\therefore \; (u,v) \in (\partial f)^{-1} \Leftrightarrow f^*(u) + f(v) = u^T v$

$(u,v) \in (\partial f)^{-1} \Leftrightarrow$ u,v are tight in Younge's inequality

def: CCP

$*$ Another important result: For a CCP{closed, convex, proper} function f, $(\partial f)^{-1} = \partial f^*$

{ f ∈ CCP = {closed, convex, proper} } $f^{**} = f$ # So, applying conjugate operator twice on a CCP function is the function itself.

$\text{epi } f = \{(x,t) \in \mathbb{R}^{n+1} \mid x \in \operatorname{dom} f, \; f(x) \leq t\}$ {[]}

$\operatorname{dom} f = \neg \phi$

now for $f \in \{CCP\}$

$(u,v) \in (\partial f)^{-1} \qquad \phi(u) \qquad \phi(v)$

$\Leftrightarrow \; f^*(u) + \underbrace{f(v)}_{f^{**}(v) \; [\because f \in \{CCP\}]} = u^T v \Leftrightarrow (f^*(u)) + (f^*(v))^* = \phi(u) + \phi^*(v) = u^T v$

$\Leftrightarrow \; \phi^*(v) + \phi(u) = v^T u$ # Remember conjugate function is always convex

$\Leftrightarrow \; (v,u) \in (\partial \phi)^{-1} = (\partial f^*)^{-1}$ $\quad$ conjugate_function(input)+normal_function(output)=input^{T}output ⇔ (input,output) ∈ (∂ function)$^{-1}$

$\Leftrightarrow \; (u,v) \in \partial f^*$

$\therefore \; (u,v) \in (\partial f)^{-1} \Leftrightarrow (u,v) \in \partial f^*$

$\overset{k}{=} \; (\partial f)^{-1} = (\partial f^*)$