Resolvent and Cayley Operator

9:37 PM

Page1

$R_F =$

$\text{resolvent}(F_{\{\text{relation}\}}) = (I\square + \lambda F(\square))^{-1} = (I + \lambda F)^{-1}$

$\lambda \in \mathbb{R}$

Resolvent_and_Cayley...

- $(\lambda \geq 0,\ F\{\text{monotone}\}) \Rightarrow R\{\text{nonexpansive}\}$ # as a result, R is a function. Lipschitz mapping, (e.g., nonexpansive and contraction mapping

def: monotone operators definitions and related

- $(\lambda \geq 0,\ F\{\text{maximal monotone}\}) \Rightarrow \text{dom } R = \mathbb{R}^n$

- $\text{Cayley-operator}(F_{\{\text{relation}\}}) = C_F(\square) = (2R\square - I\square) = 2(I\square + \lambda F\square)^{-1} - I\square$

- $(\lambda \geq 0,\ F\{\text{monotone}\}) \Rightarrow C_F\{\text{nonexpansive}\}$ proof: (\lambda>=0, F monotone) =

$*\ (\lambda \geq 0, F\{\text{monotone}\}) \Rightarrow (R \text{ is a function and } R \text{ is nonexpansive})$

thm: for positive coefficient and monotone operator resolvent is a nonexpansive function

Proof: Want to show

$$\forall_{(x,u)\in R}\ \forall_{(y,v)\in R}\quad \|u-v\| \leq \|x-y\|$$

now:

$$\forall_{(x,u)}\ \forall_{(y,v)}$$

$$xRu \ \wedge\ yRv$$

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

$$\leftrightarrow (I+\lambda F)u \ni x \ \wedge\ (I+\lambda F)v \ni y$$

$$\leftrightarrow u + \lambda F(u) \ni x \ \wedge\ v + \lambda F(v) \ni y$$

$$\leftrightarrow \exists_{z\in F(u)}\ u + \lambda z = x \ \wedge\ \exists_{w\in F(v)}\ v + \lambda w = y$$

$$\leftrightarrow \exists_{z\in F(u)}\ z = \frac{1}{\lambda}(x-u) \ \wedge\ \exists_{w\in F(v)}\ w = \frac{1}{\lambda}(y-v)$$

$$\therefore\quad z - w = \frac{1}{\lambda}\left((x-y) - (u-v)\right)$$

Again by given F: monotone $\Rightarrow (z-w)^T(u-v) \geq 0$

$$\Rightarrow \frac{1}{\lambda}\left((x-y)-(u-v)\right)^T(u-v) = \frac{1}{\lambda}\left((x-y)^T(u-v) - \|u-v\|_2^2\right) \geq 0$$

$$\rightarrow (x-y)^T(u-v) \geq \|u-v\|_2^2 \quad [\because \lambda > 0]$$

$$\therefore\quad 0 \leq \|u-v\|_2^2 \leq (x-y)^T(u-v) \leq |(x-y)^T(u-v)| \leq \|x-y\|_2\|u-v\|_2$$

$$\rightarrow \boxed{0 \leq \|u-v\|_2^2 \leq \|x-y\|_2\|u-v\|_2}$$

When, $u \neq v \Rightarrow 0 \leq \|u-v\|_2 \leq 1\|x-y\|$

$u = v$ 인데 trivially $0 \leq \|u-v\|_2 = 0 \leq 1\|x-y\|$

$$\therefore\ \forall_{(x,u)\in R}\ \forall_{(y,v)\in R}\quad \|u-v\|_2 \leq 1\|x-y\|$$

Alternative proof
Old proof , I suggest skip it.

$\forall (x,u)\in R$, $\forall (y,v)\in R$ // $R=(I\boxminus+\lambda F\boxminus)^{-1} \rightarrow R^{-1}=(I\boxminus+\lambda F\boxminus)$

$\leftrightarrow (u,x)\in R^{-1}$, $(v,y)\in R^{-1}$

$\leftrightarrow (u,x)\in (I\boxminus+\lambda F\boxminus)$, $(v,y)\in (I\boxminus+\lambda F\boxminus)$

$\leftrightarrow (I\boxminus+\lambda F\boxminus)(u)\ni x$, $(I\boxminus+\lambda F\boxminus)(v)\ni y$ $\qquad \begin{bmatrix} I_x & 0_y \\ 0_x & I_y \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$

$\leftrightarrow u+\lambda F(u)\ni x$, $v+\lambda F(v)\ni y$

now $(-I\circ(I\boxminus+\lambda F\boxminus))\circ\boxminus = -I\left((I\boxminus+\lambda F\boxminus)(\boxminus)\right) = -I\left(I\boxminus+\lambda F(\boxminus)\right) = -I\boxminus-\lambda F\boxminus$ //same operation except //every output element //has been multiplied by -1

$v+\lambda F(v)\ni y \Rightarrow -v-\lambda F(v)\ni -y$

Now design two relations:

this is a $n\times 2n$ matrix that selects the first n elements of the input 2n vector

1) $\overbrace{\left((I\boxminus+\lambda F\boxminus)\circ\left(\begin{bmatrix} I_{\boxminus[1:n]} & 0_{\boxminus[n+1:2n]}\end{bmatrix}\boxminus\right)\right)}^{f_{\Delta_1}} : R^{2n}\rightarrow R^{n}$

$\therefore \left((I\boxminus+\lambda F\boxminus)\circ\left(I_{\boxminus[1:n]} \quad 0_{\boxminus[n+1:2n]}\right)\boxminus\right)\begin{bmatrix} u \\ v \end{bmatrix} = (I\boxminus+\lambda F\boxminus)(u) = Iu+\lambda F(u)\ni x$

and

2) $\overbrace{\left((-I\boxminus-\lambda F\boxminus)\circ\left(\begin{bmatrix} 0_{\boxminus[1:n]} & 1_{\boxminus[n+1:2n]}\end{bmatrix}\boxminus\right)\right)}^{f_{\Delta_2}} : R^{2n}\rightarrow R^{n}$

$\left((-I\boxminus-\lambda F\boxminus)\circ\left(0_{\boxminus[1:n]} \quad 1_{\boxminus[n+1:2n]}\right)\boxminus\right)\begin{bmatrix} u \\ v \end{bmatrix} = (-I\boxminus-\lambda F\boxminus)(v) = -v-\lambda F(v)\ni -y$

Then

$\underbrace{(f_{\Delta_1}+f_{\Delta_2})\begin{bmatrix} u \\ v \end{bmatrix}}_{=\, f_{\Delta_1}\left(\begin{bmatrix} u \\ v \end{bmatrix}\right)+f_{\Delta_2}\left(\begin{bmatrix} u \\ v \end{bmatrix}\right)} \ni x-y$ #As: if $(u\in F_1(x),\ v\in F_2(x)) \Rightarrow u+v\in (F_1+F_2)(x)$

statement: overloaded sum operator for relations has additivity

$\leftrightarrow$

$u+\lambda F(u)-v-\lambda F(v)\ni x-y$

$\leftrightarrow (u-v)+\lambda(F(u)-F(v))\ni x-y$ // $\lambda(F_1\boxminus+F_2\boxminus)=\lambda F_1\boxminus+\lambda F_2\boxminus$ because of overloaded nature of the sum operator

$\leftrightarrow \lambda(F(u)-F(v))\ni (x-y)-(u-v)$ // Because $(u-v)$ is a singleton so:

// $(u-v)+\underbrace{\begin{bmatrix}\boxminus \\ \vdots \\ \boxminus\end{bmatrix}}_{\lambda(F(u)-F(v))} = \begin{bmatrix}(\boxminus+u-v)=\theta \\ \boxminus+(u-v)=x-y \\ (\boxminus+u-v)=\theta\end{bmatrix} \leftrightarrow \underbrace{\begin{bmatrix}\boxminus \\ \vdots \\ \boxminus\end{bmatrix}}_{\lambda(F(u)-F(v))} = \begin{bmatrix}\theta-(u-v) \\ \vdots \\ x-y \\ \vdots \\ \theta-(u-v)\end{bmatrix}$ $\therefore \lambda(F(u)-F(v))\ni x-y$

$\leftrightarrow F(u)-F(v)\ni \frac{1}{\lambda}\left((x-y)-(u-v)\right)$ // Because $\underbrace{\lambda}_{>0} I\boxminus$ is one to one function with inverse function $\frac{1}{\lambda}$

// $(x-y)-(u-v) \leftarrow \underbrace{\begin{bmatrix}\boxminus \\ \vdots \\ \boxminus\end{bmatrix}}_{\lambda(F(u)-F(v))} = \lambda \underbrace{\begin{bmatrix}\frac{1}{\lambda}\boxminus \\ \vdots \\ \frac{1}{\lambda}\boxminus\end{bmatrix}}_{F(u)-F(v)} \rightarrow \frac{(x-y)-(u-v)}{\lambda} \leftrightarrow \underbrace{\begin{bmatrix}\boxminus \\ \vdots \\ \boxminus\end{bmatrix}}_{F(u)-F(v)} = \begin{bmatrix}\frac{1}{\lambda}\boxminus \\ \vdots \\ \frac{(x-y)-(u-v)}{\lambda} \\ \vdots \\ \frac{1}{\lambda}\boxminus\end{bmatrix}$ $\therefore F(u)-F(v)\ni \frac{(x-y)-(u-v)}{\lambda}$

Now, F is monotone:

$(F(u)-F(v))^T(u-v)\geq 0$ //elementwise

$\forall_{\bar{u}-\bar{v}\in F(u)-F(v)}$ $(\bar{u}-\bar{v})^T(u-v)\geq 0$

so,

$\left(\frac{1}{\lambda}\left((x-y)-(u-v)\right)\right)^T(u-v)\geq 0$

// As a mnemonic (unverified), in a composite relation, the constituent relation gulor majhe jara function, tader jonno ukto equ ation e functional operation chalano jabe

$\rightarrow \frac{1}{\lambda}\left((x-y)^T-(u-v)^T\right)(u-v)\geq 0$

$\rightarrow (x-y)^T(u-v)-\|u-v\|_2^2\geq 0$ $[\because \lambda>0]$

$\rightarrow \underbrace{0\leq \|u-v\|_2^2}_{\text{norm } \geq 0}\leq (x-y)^T(u-v)$

if $x=y$ then $\|u-v\|_2^2=0 \leftrightarrow u=v$

But, $u\in R(x)$, $v\in R(y)$

so when $x=y$, $u=v=R(x)=R(y)$

$\therefore R$ is a function (part 1 done)

Showing R is nonexpansive:

We have already shown:

$\forall_{(x,u)\in R}$, $\forall_{(y,v)\in R}$ $\|u-v\|_2^2\leq (x-y)^T(u-v)\leq |(x-y)^T(u-v)|\leq \|x-y\|_2\,\|u-v\|_2$

$\rightarrow \|u-v\|_2^2\leq \|x-y\|_2^2\,\|u-v\|_2$

$\left.\begin{array}{l}\text{when } u\neq v, \\ \quad \|u-v\|_2\leq \|x-y\|_2 \\ \quad \leftrightarrow \|R(x)-R(y)\|_2\leq 1\,\|x-y\|_2 \\ \text{when } u=v \\ \quad \|u-v\|_2=0\leq 1\,\|x-y\|_2 \text{ trivially true}\end{array}\right\}$ $\therefore \forall_{x,y\in \mathrm{dom} R}$ $\|R(x)-R(y)\|_2\leq 1\,\|x-y\|_2$

$\leftrightarrow R$ is nonexpansive

$\rightarrow \|u-v\|_2^2 \le \|x-y\|_2 \cdot \|u-v\|_2$

when $u \neq v$,

$\|u-v\|_2 \le \|x-y\|_2$

$\Leftrightarrow \|R(x)-R(y)\|_2 \le 1\,\|x-y\|_2$

when $u=v$

$\|u-v\|_2 = 0 \le 1\,\|x-y\|_2$ trivially true

$\therefore \forall_{x,y \in \operatorname{dom} R}\ \|R(x)-R(y)\|_2 \le 1\,\|x-y\|_2$

$\Leftrightarrow$ R is nonexpansive

- $(\lambda \ge 0,\ F$ monotone operator$) \Rightarrow C = 2R\square - I\square$ is a nonexpansive function

(\lambda>=0, F monotone) => C [*nonexpansive function*]

Proof: (continued from previous proof:) R is a nonexpansive function, let, $R(x)=u$, $R(y)=v$

$C = 2\overset{\text{function}}{R}\square - \overset{\text{function}}{I}\square \Rightarrow C$ is a function

$C(x) = 2R(x) - Ix = 2R(x) - x$

$C(y) = 2\underset{\text{function}}{R(y)} - y$

$C(x) - C(y) = 2\underbrace{R(x)}_{u} - x - 2\underbrace{R(y)}_{v} + y = 2(u-v) - (x-y)$

$\therefore \|C(x)-C(y)\|_2^2 = \|2(u-v)-(x-y)\|_2^2 = \left(2(u-v)^T - (x-y)^T\right)\left(2(u-v)-(x-y)\right)$

$= 4\|u-v\|_2^2 - 4(x-y)^T(u-v) + \|x-y\|_2^2$

$= \underbrace{4\left(\underbrace{\|u-v\|_2^2 - (x-y)^T(u-v)}_{\le 0}\right)}_{\text{nonpositive number so}} + \|x-y\|_2^2$

[illegible]

$\le \|x-y\|_2^2$

$\therefore \|C(x)-C(y)\|_2^2 \le 1\,\|x-y\|_2^2$

$\therefore C$ is a nonexpansive function.

Page2 Example:

* Subdifferential mapping 의 resolvent

thm: proximal operator is the resolvent of subdifferential operator

thm: proximal operator is the resolvent of subdifferential operator

[illegible]:

$\| \ (I+\lambda\partial f)^{-1}\square = \operatorname{prox}_{\lambda f}(\square) = \underset{\square}{\operatorname{argmin}}\left[\lambda f(\square) + \frac{1}{2}\|\square - \square\|_2^2\right]$

Page3 * Normal cone operator 의 resolvent

normal cone operator

$R_{N_C}(x) = (I+\lambda \underbrace{N_C(x)}_{\partial I_C(x)})^{-1}(x) = (I+\lambda\partial I_C(x))^{-1}(x) = \underset{\square}{\operatorname{argmin}}\underbrace{\left[\lambda I_C(\square) + \frac{1}{2}\|\square - x\|_2^2\right]}_{\text{optimization problem}} = \Pi_C(x)$

$\boxed{\therefore R_{N_C}(x) = \Pi_C(x)}$ # Resolvent of normal cone operator

$\left(\underset{\square}{\forall}\ \lambda I_C(\square) + \frac{1}{2}\|\square - x\|_2^2\right) =$

$\begin{pmatrix} \underset{\square}{\forall}\ \frac{1}{2}\|\square - x\|_2^2 \\ \text{ㄱ}\quad \square \in C \end{pmatrix} \xrightarrow{\text{argmin}} \Pi_C(x)$

def: multiplier to residual mapping

* Multiplier to residual mapping operator 의 resolvent (MRM) # resolvent for multiplier to residual mapping

By definition,

$F(\square) = b - A\,\underset{x}{\operatorname{argmin}}\, L(x,\square) = b - A\left(\partial_{x^*\square} f\right)^{-1}(-A^T\square)$ # Alternative definition of MRM

$L(x,v) = f(x) + v^T(Ax-b)$

$x^*(\square) \in \underset{\bar{x}}{\operatorname{argmin}}\, L(\bar{x}, \square)$

$\left(\forall f(x) \in \mathcal{D}_{\text{strictly}}\ \text{ㄱ}\ Ax=b\right)$

We want to find the resolvent of MRM operator which is: $R = (I+\lambda F)^{-1}$, (F is monotone (def: multiplier to residual mapping), $\lambda \ge 0$) $\Rightarrow$ R is a nonexpansive function thm: for positive coefficient and monotone operator resolvent is a nonexpansive function

$R(y) = z$

$\Leftrightarrow (I+\lambda F)^{-1}(y) = z$ [function [illegible]]

$\Leftrightarrow (y,z) \in (I+\lambda F)^{-1} \Leftrightarrow (z,y) \in (I+\lambda F)$

$\Leftrightarrow y \in (I+\lambda F)(z) = z + \lambda F(z) = z + \lambda\left(b - A\underbrace{\underset{x}{\operatorname{argmin}}\, L(x,z)}_{\text{multivalued part arises from this}}\right)$

$\rightarrow \exists_{x^*(z) \in \underset{x}{\operatorname{argmin}}\, L(x,z)}\ y = z + \lambda\left(b - Ax^*(z)\right)$

$\Leftrightarrow \exists_{x^*}\ y = z + \lambda(b - Ax^*) \wedge \underbrace{x^* \in \underset{x}{\operatorname{argmin}}\, L(x,z) = f(x) + (A^Tz)^Tx - b^Tz}$ ✓

$\Updownarrow$

$\partial f(x^*) + A^Tz \ni 0$ # remember $\partial \equiv \partial_x$ here as
z is a constant

$\Leftrightarrow \exists_{x^*}\ z = y - \lambda(b - Ax^*) \wedge \partial f(x^*) + A^Tz \ni 0$ [# Note, we want to find out z=R(y), but to find out z here we also need to know x*, so this is what we are going to do, solve for x* first, and then set the value in z=y-\lambda (b- Ax*) #]

$\Leftrightarrow \exists_{x^*}\ \partial f(x^*) + A^T(y - \lambda(b - Ax^*)) \ni 0$

$\Leftrightarrow \exists_{x^*} \; z = y - \lambda(b - Ax^*) \wedge \partial f(x^*) + A^T z \ni 0$ [# Note, we want to find out z=R(y), but to find out z here we also need to know x*, so this is what we are going to do, solve for x* first, and then set the value in z=y-\lambda (b- Ax*) #]

$\Leftrightarrow \exists_{x^*} \; \partial f(x^*) + A^T(y - \lambda(b - Ax^*)) \ni 0$

#

$$\partial f(x^*) + A^T y + \lambda A^T(Ax^* - b)$$

$$= \Big[\partial_w f(w) + \nabla_w \big(\underbrace{(A^T y)^T w}_{y^T A w} - \underbrace{y^T b}_{\text{added}} \big) + \lambda \nabla_w \tfrac{1}{2} \|Aw - b\|_2^2 \Big]_{w = x^*}$$

$$= \Big(\partial_w \big[f(w) + y^T(A^T w - b) + \tfrac{\lambda}{2} \|Aw - b\|_2^2 \big] \Big)_{w = x^*}$$

$$\Leftrightarrow \exists_{x^*} \; \Big(\partial_w \big[\underbrace{\underbrace{f(w)}_{\mathcal{D}\ \text{strictly}} + \underbrace{y^T(A^T w - b)}_{\text{affine}} + \underbrace{\tfrac{\lambda}{2} \|Aw - b\|_2^2}_{\mathcal{D}\ \text{strongly}}}_{\mathcal{D}} \big] \Big)_{w = x^*} \ni 0$$

$$\Leftrightarrow x^* \in \underset{w}{\operatorname{argmin}} \left(f(w) + y^T(A^T w - b) + \tfrac{\lambda}{2} \|Aw - b\|_2^2 \right) \quad \left[\because \{ f(x) \in \mathcal{D} \} \; x^* \in \underset{x}{\operatorname{argmin}} f(x) \Leftrightarrow [\partial f(x)]_{x = x^*} \ni 0 \right]$$

#[remember, standard augmented Lagrangian for : $\begin{pmatrix} \forall f(w) \\ \urcorner \; Aw = b \end{pmatrix}$ is $L_{aug}(w, y) = f(w) + y^T(Aw - b) + \frac{\lambda}{2} \|Aw - b\|_2^2$]#

$\therefore z = R(y)$ can be determined from:

$$\boxed{\begin{aligned} x &:= \underset{w}{\operatorname{argmin}} \left(f(w) + y^T(Aw - b) + \left(\tfrac{\lambda}{2}\right) \|Aw - b\|_2^2 \right) \\ z &:= y + \lambda(Ax - b) \end{aligned}}$$

The resolvent determing equations for Multiplier to Residual Mapping ; $F(\square) = b - A \underset{\square}{\operatorname{argmin}} L(\square, \square)$

Lagrangian

$L(\square, \square) = f(\square) + \square^T(A\square - b)$

underlying optimization problem

$\forall f(\square)$

$\urcorner \; A\square = b$

Compact form: Resolvent of the multiplier to residual mapping

Compact form for resolvent of the MRM mapping , $R = (I + \lambda F)^{-1}$
$R(\square)$ # underlying optimization problem $\forall f(\square) \urcorner A\square = b$
can be calculated by the equation:
$\square = \underset{\square}{\operatorname{argmin}} \left(f(\square) + \square^T(A\square - b) + \frac{\lambda}{2} \|A\square - b\|_2^2 \right)$
$R(\square) = \square + \lambda(A\square - b)$

Page4:

- Why find fixed point of Cayley and resolvent of some operator?

Fixed points of Cayley and resolvent operators :

why maximality is needed?

def: maximal monotone

$\{F \in \text{maximal monotone}\}, \lambda > 0\}$ Maximality is needed because we know that, $(A \in \{\text{maximal monotone}\}, \lambda \geq 0) \Rightarrow \operatorname{dom} R_A = \operatorname{dom} C_A = \mathbb{R}^n$, now the damped iteration, or contraction mapping iteration repetatively apply contractive/nonexpansive mapping. If, $\operatorname{dom} R_A \neq \mathbb{R}^n$, or $\operatorname{dom} C_A \neq \mathbb{R}^n$ what if the iterate x^k escapes the dom C_A or dom R_A? Then the iteration cannot proceed.

$x^{k+1} = (1 - \theta) x^k + R(x^k)$ or C

$x^{k+1} = R(x^k)$

Often the argmin of an optimization problem can be written as the zero set of some operator : i.e. $(x, 0) \in F$

$\Leftrightarrow F(x) \ni 0$ # e.g. $\partial f(x) \ni 0 \Leftrightarrow x = \underset{y}{\operatorname{argmin}} f(y)$ etc

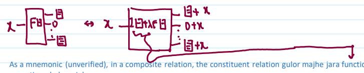
- (solutions to $F(x) \ni 0$) = (fixed points of R) = (fixed points of C)

Proof:

but this is a function → thm: for positive coefficient and monotone operator resolvent is a nonexpansive function

$F(x) \ni 0 \Leftrightarrow Ix + \lambda F(x) \ni x \Leftrightarrow (I\square + \lambda F(\square)) x \ni x \Leftrightarrow (x, x) \in (I\square + \lambda F(\square)) \Leftrightarrow (x, x) \in (I\square + \lambda F(\square))^{-1} \Leftrightarrow x = \underbrace{(I\square + \lambda F\square)^{-1}}_{R} = R(x) = \text{fixedpoint of } R$

this is the resolvent



As a mnemonic (unverified), in a composite relation, the constituent relation gulor majhe jara function, tader jonno ukto eqation e functional operation chalano jabe

$\Leftrightarrow 2x = 2R(x)$

$\Leftrightarrow x = 2R(x) - x = (2R(\square) - I\square)(x) = C(x)$ ∎