Operator splitting, Douglas-Rachford splitting

1:01 PM

Page1 Under suitable conditions, $F(x) \ni 0$ can be solved using $R(x)=x$ where $R=(I+\lambda F)^{-1}$ is the resolvent.

0 \in F(x) <=> R(x)=x

Problem হলো in cases, • R বের করা কঠিন ক্ষেত্রবিশেষ
• parallelization কঠিন হয়ে পড়ে. $\Big\}\to$ operator splitting হয় প্রযোজ্য তখন

[Second reading]

Finding zero of a monotone operator that admits splitting into two or three monotone operators, i.e., we want to find an x such that 0 ∈ (A+B)x or 0 ∈ (A+B+C)x, where A,B and C are maximal monotone operator.

The key idea: Transform the problem into a fixed point equation. In the fixed point equation, we are interested in finding the fixed point of a resolvent and/or Cayley operator of the maximal monotone operators. Though in theory such iterations converges to the fixed point, it is useful only when computing the resolvent and/or Cayley operator is efficient.

Forward-Backward Splitting:

We want to find the an x such that [Forward-backward splitting]

$$(A+B)x \ni 0 : (A,B: \text{maximal monotone})$$

A: single valued /* i.e., $A(x)=\{\cdot\}$ */ and $\alpha \geq 0$, then we have:

$$(A+B)x \ni 0 \Leftrightarrow A(x)+B(x) \ni 0 \Leftrightarrow \exists_{\eta_1,\eta_2} : A(x)=\{\eta_1\},\ B(x)\ni \eta_2 \quad \eta_1+\eta_2=0 \Leftrightarrow$$

$$\# \ \alpha\eta_1+\alpha\eta_2=0$$

$$\Leftrightarrow \alpha\eta_1+\alpha\eta_2+x-x=0 \quad /* \text{ as } A:\mathbb{R}^n\to\mathbb{R}^n,\ B:\mathbb{R}^n\to\mathbb{R}^n \ */$$

$$\Leftrightarrow \alpha A(x)+\alpha B(x)+Ix-Ix \ni 0$$

$$(I+\alpha B)(x) \qquad -(I-\alpha A)(x)$$

$$\Leftrightarrow (I+\alpha B)(x)-\underbrace{(I-\alpha A)(x)}_{\#\text{ singleton}} \ni 0 \Leftrightarrow \exists_{\eta\in(I+\alpha B)(x)} \underbrace{\eta-(I-\alpha A)(x)=0}_{\eta=(I-\alpha A)(x)}$$

$$\Leftrightarrow (I-\alpha A)(x) \in (I+\alpha B)(x) \Leftrightarrow x \in \boxed{(I+\alpha B)}\big((I-\alpha A)(x)\big)$$

$$\Leftrightarrow \big(x,(I-\alpha A)(x)\big)\in(I+\alpha B) \Leftrightarrow \big((I-\alpha A)(x),x\big)\in(I+\alpha B)^{-1}$$

$$\Leftrightarrow \underbrace{(I+\alpha B)^{-1}}\underbrace{\big((I-\alpha A)(x)\big)}_{\text{singleton}} \ni x \Leftrightarrow x=\underbrace{(I+\alpha B)^{-1}}_{R_B}(I-\alpha A)(x)=R_B(I-\alpha A)(x)$$

resolvent of (B: maximal monotone) with positive coefficient is a function

thm: for positive coefficient and monotone operator resolvent is a

the resultant fixed point iteration:

$$x^{k+1}=R_B(I-\alpha A)(x^k)=R_B(x^k-\alpha Ax^k)$$

How do we know that this iteration will converge?

Ans:

Suppose,

- A is a subdifferential operator with Lipschitz parameter L, and α ∈ (0, 2/L], or
- A:strongly monotone and Lipschitz with parameter m, L with α ∈ (0,2m/L²)

In these cases it can be shown that $(I+\alpha A)$ is an averaged operator # More explanation needed.

Then
forward step (I-αA):averaged
backward step (I + α B)^{-1} averaged # Explanation: note that as B is maximal monotone, R_B is nonexpansive, so

So $\big((1-\theta)I+\theta R_B : (\theta=1)\big)=R_B$ is an averaged operator by definition.

$(I+\alpha B)^{-1}(I-\alpha A)$: averaged # as, F: averaged, $\tilde{F}$: averaged $\to F\tilde{F}$: averaged

মনে রাখি: $F(x)=x$: (F: averaged) হলে $x^{k+1}=F(x^k)$ converges

So to find out the fixed point of $x=(I+\alpha B)^{-1}(I-\alpha A)(x)$ এর fixed point বের করা যাবে using $x^{k+1}=(I+\alpha B)^{-1}(I-\alpha A)(x^k)$

Example: Proximal gradient method is an example of Forward-Backward splitting. For details see [Forward-Backward Version of Proximal Gradient Method]

*** Forward-backward-forward splitting**

We again consider:

find $x : (A+B)x \ni 0$ /* A: maximal monotone, Lipschitz with parameter $L \Rightarrow$ single-valued function
B: maximal monotone */

A: function $\to (I-\alpha A)$: one to one mapping # i.e. $x\neq y \Rightarrow (I-\alpha A)x \neq (I-\alpha A)y$

• Proof:

Let, $\forall_{x,y}\ x\neq y$

$$\|(I-\alpha A)x-(I-\alpha A)y\|_2 = \|x-\alpha Ax-y+\alpha Ay\|_2$$

$$=\|(x-y)-\alpha(Ax-Ay)\|_2 \geq \|x-y\|_2-\alpha\|Ax-Ay\|_2$$

using reverse triangle inequality: $\|x-y\|_2 \geq \|x\|_2-\|y\|_2$
proof
$\|x-y\|_2+\|y\|_2 \geq \|x-y+y\|_2=\|x\|_2$
$\Leftrightarrow \|x-y\|_2 \geq \|x\|_2-\|y\|_2$

/* Again by definition. A: Lipschitz
$\|Ax-Ay\|_2 \leq L\|x-y\|_2$
$\Leftrightarrow -\alpha\|Ax-Ay\|_2 \geq -\alpha L\|x-y\|_2$ */

$$\geq \|x-y\|_2-\alpha L\|x-y\|_2=\underbrace{(1-\alpha L)}_{>0}\underbrace{\|x-y\|_2}_{>0}$$

$\alpha\in(0,\frac{1}{L})\to\alpha<\frac{1}{L}\Leftrightarrow(1-\alpha L)>0$ and $x\neq y$ $\therefore \|x-y\|_2>0$

$$>0$$

$$\therefore \|(I-\alpha A)x-(I-\alpha A)y\|_2>0$$

so if we take $(I-\alpha A)x=\eta_x$, $(I-\alpha A)y=\eta_y$ then $\forall_{x\neq y}\big(\|\eta_x-\eta_y\|>0 \Leftrightarrow \eta_x\neq\eta_y \Leftrightarrow (I-\alpha A)x\neq(I-\alpha A)y\big)$

$\therefore I-\alpha A$: one-to-one mapping

* We have already shown in forward backward splitting that

$$(A+B)(x)\ni 0 \Leftrightarrow x=R_B(I-\alpha A)x$$

$$\Leftrightarrow (I-\alpha A)x=(I-\alpha A)R_B(I-\alpha A)(x)$$ /* As both the sides are multiplied by the one-to-one operator (I - α A), information will be preserved */

$$\Leftrightarrow x-\alpha A(x)=(I-\alpha A)R_B(I-\alpha A)(x)$$

$$\Leftrightarrow x=(I-\alpha A)R_B(I-\alpha A)x+\alpha A(x)$$

$$=\big((I-\alpha A)R_B(I-\alpha A)\square+\alpha A\square\big)(x)$$

$$\Leftrightarrow x=\big((I-\alpha A)R_B(I-\alpha A)+\alpha A\big)(x)$$

$$\therefore (A+B)x\ni 0 \Leftrightarrow x=\text{fixed point}\big((I-\alpha A)R_B(I-\alpha A)+\alpha A\big)$$

Fixed point iteration scheme:

$$x^{k+\frac{1}{2}}=R_B(I-\alpha A)(x^k)=R_B(x^k-\alpha Ax^k)$$
$$x^{k+1}=x^{k+\frac{1}{2}}-\alpha(Ax^{k+\frac{1}{2}}-Ax^k)$$

Developed by Paul Ptseng

[Forward-backward-forward splitting]

Caution: (I-αA)R_B(I-αA) **not** nonexpansive ⇒ So the averaged operator convergence proof where we show that the damped iteration will converge, will not work.

***Condition of convergence:** Forward-backward-forward splitting will converge when A: (Lipschitz with parameter L, α ∈ (0,1/L))
Note that in Forward-backward we required strongly monotone A and Lipschitz, so in FBF we have better assumption.

Convergence Proof: May be later...

Example: Extragradient method:

Consider finding the zero of a maximal monotone function A, i.e., A(x)=0, we can write this as A x + 0 x = 0 so,
We have B=0 so resolvent is $R_B=(I+\alpha\, 0)^{-1}= I$

So by using Ptseng's method we have:

$x^{k+\frac{1}{2}} = R_B\,(x^k - \alpha A x^k) = I\,(x^k - \alpha A x^k) = x^k - \alpha A x^k$

$x^{k+1}=x^{k+1} - \alpha\,(A x^{k+\frac{1}{2}} - A x^k)$

Which converges when A is Lipschitz with parameter L, and α∈ (0,1/L)

Operator Splitting : why maximality is needed?

• We want to solve $F(x) \ni 0$ {F{maximal monotone}}

idea • $\exists_{A,B} \left(F = A+B \wedge \{A,B\}.\{\text{maximal monotone}\}\right)$

• $R_A = (I+\lambda A)^{-1}$, $R_B = (I+\lambda B)^{-1}$ # R_A, R_B determination, R_F এর চাইতে সহজতর, you can hit the #jackpot

$C_A = 2R_A - I$, $C_B = 2R_B - I$

$(\{A,B\}.\{\text{maximal monotone}\}, \lambda \geq 0) \Rightarrow \underbrace{\{C_A, C_B\}.\{\text{nonexpansive}\}}$ # as $(F\{\text{monotone}\}, \lambda \geq 0) \Rightarrow C\{\text{nonexpansive}\}$

$\Downarrow$

$C_A C_B$ {nonexpansive} # composition of nonexpansive (contraction) mapping is nonexpansive (contraction)

$$\boxed{\left\{\exists_{\{A,B\}.\{\text{maximal monotone}\}} F = A+B\right\} \; F(x) = A(x)+B(x) \ni 0 \;\leftrightarrow\; \left(C_A C_B(z) = z,\; x = R_B(z)\right) \leftrightarrow \left(x = R_B(z),\; \tilde{z} = 2x - z,\; \tilde{x} = R_A(\tilde{z}),\; z = 2\tilde{x} - \tilde{z}\right)}$$

thm: main theorem behind operator splitting

অর্থাৎ, $F(x) \ni 0$ এর solution $C_A C_B(\square)$ এর fixed point থেকে বের করা যাবে

$(L.H.S \leftarrow M.H.S) \wedge (M.H.S \leftrightarrow R.H.S)$

Proof Strategy: At first we will prove R.H.S ⇔ M.H.S, then we will prove M.H.S ⇒ L.H.S. This is the computationally important part, as it tells us that if we solve $C_A C_B$ (z)=z, x=R_B(z), then the x will be a zero of the maximal montobe relation F=A+B i.e, F(x)=A(x)+B(x) ∋ 0.

Proof :

#Proof of R.H.S ⇔ M.H.S

$\tilde{z} = 2x - \tilde{z}$, $x = R_A(\tilde{z})$, $2\tilde{x} = \tilde{z} + z$ assignment গুলোকে alternatively এভাবে justify করা যায়:

$(C_A C_B(z) = z, x = R_B(z)) \Leftrightarrow (x = R_B(z), \tilde{z} = 2x - z, \tilde{x} = R_A(\tilde{z}), z = 2\tilde{x} - \tilde{z})$

Proof: $(x = R_B(z) \wedge z = C_A C_B(z))$ # $P \wedge Q \overset{k}{=} (P \wedge P) \wedge Q \overset{k}{=} P \wedge (P \wedge Q)$

$(x = R_B(z) \wedge z = C_A C_B(z) = C_A(2R_B \boxminus - I\boxminus)(z) = C_A(2R_B(z) - z) = C_A(2x - z) = (2R_A \boxminus - I\boxminus)\underbrace{(2x - z)}_{\tilde{z}}$

$= (2R_A \boxminus - I \boxminus)(\tilde{z})$ # $\tilde{z} = 2x - z$ # $= (2\underbrace{R_A(\tilde{z})}_{\tilde{x}} - \tilde{z})$ # $\tilde{z} = 2x - z$ #

$= (2\tilde{x} - \tilde{z})$ # $\tilde{z} = 2x - z$, $\tilde{x} = R_A(\tilde{z})$ # $)$

$\leftrightarrow (x = R_B(z) \wedge \tilde{z} = 2x - z \wedge \tilde{x} = R_A(\tilde{z}) \wedge z = 2\tilde{x} - \tilde{z})$ # written in the order of the arrival

$\therefore (C_A C_B(z) = z, x = R_B(z)) \leftrightarrow (x = R_B(z), \tilde{z} = 2x - z, \tilde{x} = R_A(\tilde{z}), z = 2\tilde{x} - \tilde{z})$ (M.H.S=R.H.S)

$\tilde{z} = C_B(z)$ না লিখে $\tilde{z} = 2x - z$ form এ কেন লিখা হলো? কারণ কি R_A, R_B [illegible]

A: Because R_A, R_B (resolvent) is often easier to calculate than $C_A, C_B = (2R_B(\boxminus) - I\boxminus)$

Now let's prove $(x = R_B(z), \tilde{z} = R_B(z), \tilde{x} = R_A(\tilde{z}), z = 2\tilde{x} - \tilde{z}) \Rightarrow F(x) = A(x) + B(x) \ni 0$

now $x = R_B(z) = (I\boxminus + \lambda B(\boxminus))^{-1} z$

$\leftrightarrow (z, x) \in (I + \lambda B)^{-1}$

$\leftrightarrow (x, z) \in (I + \lambda B)$

$\leftrightarrow z \in (I\boxminus + \lambda B(\boxminus))(x) = x + \lambda B(x)$

$x - \boxed{I\boxminus + \lambda A(\boxminus)} - z$

and

$\tilde{x} = R_A(\tilde{z}) = (I\boxminus + \lambda A \boxminus)^{-1}$

$\leftrightarrow (\tilde{z}, \tilde{x}) \in (I\boxminus + \lambda A \boxminus)^{-1}$

$\leftrightarrow (\tilde{x}, \tilde{z}) \in (I\boxminus + \lambda A(\boxminus))$

$\leftrightarrow \tilde{z} \in (I\boxminus + \lambda A(\boxminus))(\tilde{x}) = \tilde{x} + \lambda A(\tilde{x})$

$\leftrightarrow \tilde{z} \in (I\boxminus + \lambda A \boxminus)(x)$ [∵ $x = \tilde{x}$]

(statement: overloaded sum operator for relations has additivity + [eg: Sum of relations]) এর মানে: $(y \in R(x), z \in S(x)) \Leftrightarrow y + z \in (R+S)(x) = R(x) + S(x)$

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

$\rightarrow (z \in (I\boxminus + \lambda B(\boxminus))(x), \tilde{z} \in (I\boxminus + \lambda A(\boxminus))(\tilde{x})) \leftrightarrow \underbrace{z + \tilde{z}}_{2x} \in (I\boxminus + \lambda A(\boxminus))(x) + (I\boxminus + \lambda B(\boxminus))(x) = x + \lambda A(x) + x + \lambda B(x) = 2x + \lambda(A(x) + B(x))$

$\leftrightarrow 2x \in 2x + \lambda(A(x) + B(x)) = (2I\boxminus + \lambda(A+B)(\boxminus))(x)$

$\leftrightarrow 2x \in 2x + \lambda(A+B)(x)$

$\leftrightarrow \exists_{\eta \in (A+B)(x)} (\lambda\eta + 2x = 2x \leftrightarrow \lambda\eta = 0)$

$\rightarrow \exists_{\eta \in (A+B)(x)} \eta = 0$ [∵ $\lambda > 0$]

$\leftrightarrow (A+B)(x) \ni 0$ # y \in \lambda F(x) <=> y/\lambda \in F(x)

$\therefore (C_A C_B(z) = z, x = R_A(z)) \Rightarrow (0 \in A(x) + B(x))$

$\downarrow$

অর্থাৎ এর solution বের করলে $F(x) = A(x) + B(x) \ni 0$ এর সমাধান পেয়ে যাবো. $\therefore$ L.H.S ← M.H.S ∎

✱ Douglas Rachford splitting

So, দেখা গেলো.

$\left\{\exists_{\{A,B\}.\{\text{maximal monotone}\}} F = A+B\right\} \left(F(x) = A(x) + B(x) \ni 0 \leftrightarrow (C_A C_B(z) = z, x = R_B(z)) \leftrightarrow (x = R_B(z), \tilde{z} = 2x - z, \tilde{x} = R_A(\tilde{z}), z = 2\tilde{x} - \tilde{z})\right)$

intermediate variable, $\tilde{z} = C_B(z)$

$z = C_A(\tilde{z})$

intermediate but 2,4th থেকে: $\tilde{z} = 2x - z$, $z = 2\tilde{x} - \tilde{z}$ (-) $\overline{\tilde{z} - z = 2(x - \tilde{x}) - z + \tilde{z}} \rightarrow \boxed{x = \tilde{x}}$.

now, to solve $x = R_B(z)$, $z = \underbrace{C_A C_B}_{\text{nonexpansive}}(z)$ so iterative way তে z বের করতে হবে আমাদের damped iteration apply করতে হবে, এই damped iteration ই Douglas-Rachford splitting.

(however Peaceman-Rachford splitting contraction algorithm apply হয় which might not necessarily converge)

≠ damped iteration

$z^{k+1} = (1-\theta) z^k + \theta C_A C_B(z^k)$ # $\theta \in (0,1)$

$= \frac{1}{2} z^k + \frac{1}{2} C_A \underbrace{(2R_B \boxminus - I\boxminus)(z^k)}$ # $\theta = \frac{1}{2}$ হলে

$2\underbrace{R_B(z^k)} - z^k$

$x^k = R_B(z^k)$ 1st intermediate iteration impose করতে পারি, as $x = R_B(z)$, note that z^k kth iteration এ বের করছি, so after that $x^k = R_B(z^k)$ ও বের করতে পারি.

$= \frac{1}{2} z^k + \frac{1}{2} C_A(2x^k - z^k)$

let, $\tilde{z}^k = 2x^k - z^k$, 2nd intermediate iteration note that x^k মানে $x^k = R_B(z^k)$ থেকে যা 1st intermediate action থেকে পাই, আর z^k is the kth master iteration

$= \frac{1}{2} z^k + \frac{1}{2} C_A(\tilde{z}^k) = \frac{1}{2} z^k + \frac{1}{2}(2R_A \boxminus - I\boxminus)(\tilde{z}^k)$

$= \frac{1}{2} z^k + \frac{1}{2}(2R_A(\tilde{z}^k) - \tilde{z}^k)$ # let $\tilde{x}^k = R_A(\tilde{z}^k)$, 3rd intermediate iteration

$= \frac{1}{2} z^k + \frac{1}{2}(2\tilde{x}^k - \tilde{z}^k)$

$= \frac{1}{2} z^k + \frac{1}{2}(2\tilde{x}^k - (2x^k - z^k))$

$= \frac{1}{2} z^k + \tilde{x}^k - x^k + \frac{1}{2} z^k$

$= z^k + \tilde{x}^k - x^k$

So, in the order of execution Douglas-Rachford splitting iterations are:

$$\boxed{\begin{array}{ll} x^k := R_B(z^k) & \text{\# 1st intermediate iteration} \\ \tilde{z}^k := 2x^k - z^k & \text{\# 2nd intermediate iteration} \\ \tilde{x}^k := R_A(\tilde{z}^k) & \text{\# 3rd intermediate iteration} \\ z^{k+1} := z^k + \underbrace{\tilde{x}^k - x^k}_{\text{this is a residual}} & \text{\# master iteration} \end{array}}$$

↗ solves $F(x) = A(x) + B(x) \ni 0 \Leftrightarrow z = C_A C_B(z),\ x = R_B(z)$ # thm: main theorem behind operator splitting

[# eq: Douglas-Rachford splitting (my notation)]

as $z^k \hookrightarrow z^* = C_A C_B(z^*)$ so, $\tilde{x}^k - x^k$ is sort of a residual and will approach zero : $\tilde{x}^k - x^k \hookrightarrow 0 \Rightarrow \lim_{k\to\infty} \tilde{x}^k = x^k$

z^k is the running sum of residuals

, inspired by this observation we can write D-R splitting in Ernest's notation

$$\boxed{\begin{array}{l} x^{k+\frac{1}{2}} := R_B(z^k) \\ \tilde{z}^{k+\frac{1}{2}} := 2x^{k+\frac{1}{2}} - z^k \\ x^{k+1} := R_A(\tilde{z}^{k+\frac{1}{2}}) \\ z^{k+1} := z^k + x^{k+1} - x^{k+\frac{1}{2}} \end{array}}$$

↗ solves $F(x) = A(x) + B(x) \ni 0 \Leftrightarrow z = C_A C_B(z),\ x = R_B(z)$

[# eq: Douglas-Rachford splitting (Ernest's notation)]

Note that, here $k+\frac{1}{2}$ has a special meaning, it signifies an intermediate piece of information that uses information from the kth iterate, and $\square^{(k+1/2)}$ itself will be used to compute the next iterate x^{k+1}.
Also from my notation → Ernest's notation, the intermediate iteration গুলো subscript change করা হয়েছে, which is upto us, as numerically everything is still same.

Douglas-Rachford splitting:

- নানা way তে rearrange করা যায়
- equivalent to many other algorithm, but not obvious
- কেবলমাত্র A, B maximally monotone হলেই solution exists
- A, B are handled seperately (by R_A, R_B); they are 'uncoupled'.

• Alternating direction method of multipliers: # ADMM by Douglas-Rachford splitting

suppose we want:

$$\min f(x) + g(x),\ \text{so the minimizer } x^* \square\ \partial(f(x^*) + g(x^*)) = \underbrace{\partial f(x^*)}_{B} + \underbrace{\partial g(x^*)}_{A} \ni 0$$

কোনটা A, কোনটা B তা matter করে না.

remember using indicator function $\begin{pmatrix} \min f(x) \\ \text{s.t. } x \in C \end{pmatrix} = \left(\min f(x) + I_C(x) \right)$

thm: proximal operator is the resolvent of subdifferential operator

- the resolvent of a subdifferential operator w.r.t function ψ is the proximal map, i.e.

$$R_{\partial\psi}(\square) = (I + \lambda\psi)^{-1}\square = \mathrm{prox}_{\lambda\psi}(\square) = \underset{\square}{\mathrm{argmin}} \left(\lambda\psi(\square) + \tfrac{1}{2}\|\square - \square\|_2^2 \right) = \underset{\square}{\mathrm{argmin}} \left(\psi(\square) + \tfrac{1}{2\lambda}\|\square - \square\|_2^2 \right)$$

by definition $\mathrm{prox}_{\psi}(\square) = \underset{\square}{\mathrm{argmin}} \left(\psi(\square) + \tfrac{1}{2}\|\square - \square\|_2^2 \right)$

so Douglas-Rachford splitting will become: # [# eq: Douglas-Rachford splitting (Ernest's notation

$F(x) = A(x) + B(x) = \partial f(x) + \partial g(x) \ni 0$ # $A = \partial g$, $B = \partial f$

$$\boxed{\begin{array}{l} x^{k+\frac{1}{2}} := R_B(z^k) = R_{\partial g}(z^k) = \underset{\square}{\mathrm{argmin}} \left(f(\square) + \frac{1}{2\lambda}\|\square - z^k\|_2^2 \right) \\ \tilde{z}^{k+\frac{1}{2}} := 2x^{k+\frac{1}{2}} - z^k \\ x^{k+1} := R_A(\tilde{z}^{k+\frac{1}{2}}) = R_{\partial f}(\tilde{z}^{k+\frac{1}{2}}) = \underset{\square}{\mathrm{argmin}} \left(f(\square) + \frac{1}{2\lambda}\|\square - \tilde{z}^{k+\frac{1}{2}}\|_2^2 \right) \\ z^{k+1} := z^k + x^{k+1} - x^{k+\frac{1}{2}} \end{array}}$$

↝ This is a special case of alternating direction method of multipliers.

page4 Constrined convex optimization problem:

$$\begin{pmatrix} \min f(x) \\ \text{s.t. } x \in C \end{pmatrix} = \min \left(f(x) + I_C(x) \right)$$

so the minimizer $x^* \square\ \partial(f(x) + I_C(x)) = \partial f(x) + \underbrace{\partial I_C(x)}_{= N_C(x)} = \underbrace{\partial f(x)}_{B} + \underbrace{N_C(x)}_{A}$

$N_C(x)$ # normal cone operator

now

- resolvent of the subdifferential operator of any function ψ

$$R_{\partial\psi}(\square) = \underset{\square}{\mathrm{argmin}} \left(\psi(\square) + \tfrac{1}{2\lambda}\|\square - \square\|_2^2 \right)$$

- resolvent of the normal cone operator of the normal cone operator of a set C, N_C: # Resolvent of normal cone operator

$$R_{N_C}(\square) = \Pi_C(\square)$$

So, for this case Douglas-Rachford splitting will become:

$A(x) + B(x) \ni 0$ where $A(x) = N_C(x)$, $B(x) = \partial f(x)$

$x^{k+\frac{1}{2}} := R_B(z^k) = R_{\partial f}(z^k) = \underset{\square}{\mathrm{argmin}} \left(f(\square) + \frac{1}{2\lambda}\|\square - z^k\|_2^2 \right)$ # note that this step can be parallelized, if f is seperable

$\tilde{z}^{k+\frac{1}{2}} := 2x^{k+\frac{1}{2}} - z^k$

$x^{k+1} := R_A(\tilde{z}^{k+\frac{1}{2}}) = \Pi_C(\tilde{z}^{k+\frac{1}{2}})$ # if C is a direct product of simpler convex sets (see # List of projections on simple convex sets

$z^{k+1} := z^k + x^{k+1} - x^{k+\frac{1}{2}}$

✱ Dykstra's alternating projections for finding a point in the intersection of convex sets C, D

$x \in C \cap D \Leftrightarrow N_C(x) + N_D(x) \ni 0$ // as $I_{set}(\square)$ is a function

proof:

($\rightarrow$) Want to prove: $x \in C \cap D \rightarrow N_C(x) + N_D(x) \ni 0$

from normal cone operator we know: $N_C(x) = \begin{cases} \emptyset, & x \notin C \\ \{g_x \mid \forall_{y \in C}\ g_x^T(y - x) \leq 0\}, & x \in C \end{cases}$

as $x \in C \cap D \Leftrightarrow x \in C \wedge x \in D$

$x \in C \Rightarrow N_C(x) = \{g_x \mid \forall_{y \in C}\ g_x^T(y - x) \leq 0\}$

clearly $g_x = 0 \in N_C(x)$ as $\forall_{y \in C}\ 0^T(y - x) = 0 \leq 0$

$\left.\begin{array}{l} \therefore x \in C \Rightarrow N_C(x) \ni 0 \Leftrightarrow (x, 0) \in N_C \\ \text{similarly } x \in D \Rightarrow N_D(x) \ni 0 \Leftrightarrow (x, 0) \in N_D \end{array}\right\} \therefore (x, 0+0) \in N_C + N_D \Leftrightarrow (N_C + N_D)(x) = N_C(x) + N_D(x) \ni 0$

underlying proof structure,

$$\begin{array}{l} \text{crude given} \\ \\ \text{implicitly given} \end{array} \left[\begin{array}{l} \overbrace{A \checkmark}^{\text{given}} \\ B \checkmark \\ A \rightarrow C \\ C \rightarrow D \\ (C \wedge D) \rightarrow E \end{array}\right] \xrightarrow{[\text{modus ponens}]} \overset{\text{given}}{\left[\begin{array}{l} C \\ D \\ (C \wedge D) \rightarrow E \end{array}\right]} \Rightarrow \overset{\text{given}}{\left[\begin{array}{l} (C \wedge D) \checkmark \\ (C \wedge D) \rightarrow E \checkmark \end{array}\right]} \xrightarrow{[\text{modus ponens}]} \overset{\text{goal}}{E}\ ☺$$

($\leftarrow$) Want to prove: $\left(N_C(x) + N_D(x) \ni 0 \rightarrow x \in C \wedge D \right) \overset{\text{def}}{=} \left(\neg(N_C(x) + N_D(x) \ni 0) \vee x \in C \wedge D \right)$

Per absurdum,

$\neg\left(\neg(N_C(x) + N_D(x) \ni 0) \vee x \in C \wedge D \right) = \underbrace{N_C(x) + N_D(x) \ni 0}_{P} \wedge (x \notin C \wedge D)$

$(x \notin C \wedge D) \Leftrightarrow \underbrace{x \notin C}_{Q} \vee \underbrace{x \notin D}_{R}$

$= P \wedge (Q \vee R) = (P \wedge Q) \vee (P \wedge R)$

case I: $(P \wedge Q) \overset{\text{def}}{=} N_C(x) + N_D(x) \ni 0 \wedge x \notin C$

$$\Updownarrow$$

$$\underbrace{x \notin C}_{Q} \vee \underbrace{x \notin D}_{R}$$

$= P \wedge (Q \vee R) = (P \wedge Q) \vee (P \wedge R)$

Case I: $(P \wedge Q) \triangleq N_C(x) + N_D(x) \ni 0 \wedge x \notin C$

now $x \notin C \rightarrow N_C(x) = \phi$, as $N_C(x) = \begin{cases} \phi, & x \notin C \\ \{g_x\} \ \forall_{y \in C} \ g_x^T(y-x) \leq 0, & x \in C \end{cases}$

so, $N_C(x) + N_D(x) = \phi + N_D(x) = \phi$ // overloaded set addition notation: $\phi + A = \phi$

so $N_C(x) + N_D(x) \not\ni 0$ contradiction for case 1.

Case II:

Similarly, $(N_C(x) + N_D(x) \ni 0 \wedge x \notin C) \Rightarrow N_C(x) + N_D(x) \not\ni 0$

contradiction for case II.

Combining both of them: we arrive at a contradiction for the perabsurdum statement

$\therefore N_C(x) + N_D(x) \ni 0 \rightarrow x \in C \cap D$ ∎

So now: $x \in C \cap D \leftrightarrow \underbrace{N_C(x)}_{A(x)} + \underbrace{N_D(x)}_{B(x)} \ni 0$

resolvent of the normal cone operator of a set C, N_C: [# Resolvent of normal cone operator]

$$R_{N_C}(\boxminus) = \Pi_C(\boxminus)$$

- So, for this case Douglas-Rachford splitting will become:

$$x^{k+\frac{1}{2}} := R_B(z^k) = R_{N_C}(z^k) = \Pi_C(z^k)$$
$$\tilde{z}^{k+\frac{1}{2}} := 2x^{k+\frac{1}{2}} - z^k$$
$$x^{k+1} := R_A(\tilde{z}^{k+\frac{1}{2}}) = R_{N_D}(\tilde{z}^{k+\frac{1}{2}}) = \Pi_D(\tilde{z}^{k+\frac{1}{2}})$$
$$z^{k+1} := z^k + x^{k+1} - x^{k+\frac{1}{2}}$$

solves $(A(x) + B(x) \ni 0$ where $A(x) = N_C(x)$, $B(x) = N_D(x))$