Duality in Splitting Methods

Ernest K. Ryu
Seoul National University

Mathematical and Numerical Optimization
Fall 2020

Last edited: 11/20/2020

Attouch–Théra duality and splitting methods

We present Attouch–Théra duality, which is analogous to, but *simpler* than, convex duality, and explore its connection to base splitting methods.

Outline

Fenchel duality

Fenchel duality: primal

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) + g(x),$$

and dual

$$\underset{u \in \mathbb{R}^n}{\text{maximize}} \quad -f^*(-u) - g^*(u)$$

generated by

$$\mathbf{L}(x, u) = f(x) + \langle x, u \rangle - g^*(u).$$

Total duality is subtle.

One Interpretation of Fenchel duality

For simplicity, assume total duality and f, g, f^*, and g^* differentiable.

Primal is to find point x such that ∇f and ∇g at x sum to 0:

$$\underset{x\in\mathbb{R}^n}{\text{find}} \quad 0 = \nabla f(x) + \nabla g(x),$$

Dual is to find gradient u such that ∇f produces $-u$ and ∇g produces u at the same point:

$$\underset{u\in\mathbb{R}^n}{\text{find}} \quad (\nabla f)^{-1}(-u) = (\nabla g)^{-1}(u)$$

This is one of the many viewpoints of convex duality.

Outline

Attouch–Théra duality

Consider

$$\underset{x \in \mathbb{R}^n}{\text{find}} \quad 0 \in (\mathbb{A} + \mathbb{B})x,$$

where $\mathbb{A}$ and $\mathbb{B}$ are maximal monotone.

Define $\mathbb{A}^{-\varovee}(u) = -\mathbb{A}^{-1}(-u)$.

Attouch–Théra dual monotone inclusion problem is

$$\underset{u \in \mathbb{R}^n}{\text{find}} \quad 0 \in (\mathbb{A}^{-\varovee} + \mathbb{B}^{-1})u.$$

Attouch–Théra duality

Attouch–Théra duality is, in a sense, easier than Fenchel duality since

$$\operatorname{Zer}(\mathbb{A}+\mathbb{B}) \neq \emptyset \quad \Leftrightarrow \quad \operatorname{Zer}(\mathbb{A}^{-\circledvee}+\mathbb{B}^{-1}) \neq \emptyset,$$

i.e., a primal solution exists if and only if a dual solution exists.

Proof.

$$\begin{aligned} \exists\, x\, [0 \in (\mathbb{A}+\mathbb{B})x] \quad &\Leftrightarrow \quad \exists\, x, u\, [-u \in \mathbb{A}x,\ u \in \mathbb{B}x] \\ &\Leftrightarrow \quad \exists\, x, u \left[-x \in \mathbb{A}^{-\circledvee}u,\ x \in \mathbb{B}^{-1}u\right] \\ &\Leftrightarrow \quad \exists\, u \left[0 \in (\mathbb{A}^{-\circledvee}+\mathbb{B}^{-1})u\right]. \end{aligned}$$

□

(No notion of strong duality, since no function values.)

Attouch–Théra vs. Fenchel duality

Attouch–Théra generalized Fenchel duality in the following sense:

$$\partial(\text{proper convex function}) \subset \text{monotone operators}$$

However, Attouch–Théra fails to capture the subtleties of Fenchel duality.

In Fenchel duality, strong duality may fail, a primal solution may exist while a dual solution does not, or vice versa. No analogous pathologies in Attouch–Théra duality.

Dual solutions as certificates

It is desirable for a method to produce both primal and dual solutions as the dual solution can certify correctness of the primal solution.

If a primal-dual solution $(x^\star, u^\star)$ satisfying

$$-u^\star \in Ax^\star \text{ and } u^\star \in Bx^\star \tag{1}$$

is provided, verifying (1) certifies correctness of the solutions.

If only a primal solution $x^\star$ is provided, we must verify $0 \in Ax^\star + Bx^\star$. How do we compute the Minkowski sum $Ax^\star + Bx^\star$?

Outline

FBS

The FPI with FBS

$$\begin{aligned} x^{k+1/2} &= x^k - \alpha \mathbb{A} x^k \\ x^{k+1} &= \mathbb{J}_{\alpha\mathbb{B}} x^{k+1/2} \end{aligned}$$

often not considered a primal-dual method. We can make it primal-dual:

$$\begin{aligned} x^{k+1/2} &= x^k - \alpha \mathbb{A} x^k \\ u^{k+1/2} &= -\mathbb{A} x^k \\ x^{k+1} &= \mathbb{J}_{\alpha\mathbb{B}} x^{k+1/2} \\ u^{k+1} &= \alpha^{-1}(x^{k+1/2} - x^{k+1}). \end{aligned}$$

Note $u^{k+1} \in \mathbb{B} x^{k+1}$. If $x^k \to x^\star$, then

$$u^{k+1/2} \to u^\star, \quad u^{k+1} \to u^\star, \quad u^\star \in \operatorname{Zer}(\mathbb{A}^{-\circledvee} + \mathbb{B}^{-1}).$$

Characterization of fixed points of DRS

With Attouch–Théra dual, characterize fixed points of PRS and DRS:

$$\mathrm{Fix}\,(\mathbb{R}_{\alpha\mathbb{A}}\mathbb{R}_{\alpha\mathbb{B}}) \subseteq \mathrm{Zer}\,(\mathbb{A}+\mathbb{B}) + \alpha\mathrm{Zer}\,(\mathbb{A}^{-\circledvee}+\mathbb{B}^{-1})$$

Proof.

$$\begin{aligned}
z =& \mathbb{R}_{\alpha\mathbb{A}}\mathbb{R}_{\alpha\mathbb{B}}z \\
&\Leftrightarrow \quad z + 2\mathbb{J}_{\alpha\mathbb{A}}(2\mathbb{J}_{\alpha\mathbb{B}} - I)z - 2\mathbb{J}_{\alpha\mathbb{B}}z = z,\ x = \mathbb{J}_{\alpha\mathbb{B}}z \\
&\Leftrightarrow \quad \mathbb{J}_{\alpha\mathbb{A}}(x-\alpha u) = x,\ z = x+\alpha u,\ u\in\mathbb{B}x \\
&\Leftrightarrow \quad x-\alpha u = x+\alpha\mu,\ \mu\in\mathbb{A}x,\ z=x+\alpha u,\ u\in\mathbb{B}x \\
&\Leftrightarrow \quad \mu=-u,\ \mu\in\mathbb{A}x,\ u\in\mathbb{B}x,\ z=x+\alpha u \\
&\Leftrightarrow \quad -u\in\mathbb{A}x,\ u\in\mathbb{B}x,\ z=x+\alpha u \\
&\Leftrightarrow \quad -u\in\mathbb{A}x,\ u\in\mathbb{B}x,\ -x\in\mathbb{A}^{-\circledvee}u,\ x\in\mathbb{B}^{-1}u,\ z=x+\alpha u \\
&\Rightarrow \quad 0\in(\mathbb{A}+\mathbb{B})x,\ 0\in(\mathbb{A}^{-\circledvee}+\mathbb{B}^{-1})u,\ z=x+\alpha u.
\end{aligned}$$

Last step is not an equivalence, so characterization with $\subseteq$, not $=$.

Primal-dual DRS

We can make the FPI with DRS more explicitly primal-dual:

$$\begin{aligned}
x^{k+1/2} &= \mathbb{J}_{\alpha\mathbb{B}}(z^k) \\
u^{k+1/2} &= \frac{1}{\alpha}(z^k - x^{k+1/2}) \\
x^{k+1} &= \mathbb{J}_{\alpha\mathbb{A}}(2x^{k+1/2} - z^k) \\
u^{k+1} &= \frac{1}{\alpha}(x^{k+1} - x^{k+1/2} + \alpha u^{k+1/2}) \\
z^{k+1} &= z^k + x^{k+1} - x^{k+1/2}.
\end{aligned}$$

Note $u^{k+1/2} \in \mathbb{B}x^{k+1/2}$, $-u^{k+1} \in \mathbb{A}x^{k+1}$. If $\mathrm{Zer}\,(\mathbb{A} + \mathbb{B}) \neq \emptyset$, then

$$x^{k+1/2} \to x^\star, \quad x^{k+1} \to x^\star, \quad x^\star \in \mathrm{Zer}\,(\mathbb{A} + \mathbb{B})$$
$$u^{k+1/2} \to u^\star, \quad u^{k+1} \to u^\star, \quad u^\star \in \mathrm{Zer}\,(\mathbb{A}^{-\circledvee} + \mathbb{B}^{-1})$$
$$z^k \to x^\star + \alpha u^\star.$$

Self-dual property of DRS

PRS and DRS are self-dual:

$$\mathbb{R}_{\mathbb{A}}\mathbb{R}_{\mathbb{B}} = \mathbb{R}_{\mathbb{A}^{-\circledvee}}\mathbb{R}_{\mathbb{B}^{-1}}$$

Follows from $\mathbb{J}_{\mathbb{A}^{-\circledvee}} = \mathbb{I} + \mathbb{J}_{\mathbb{A}}(-\mathbb{I})$ and $\mathbb{J}_{\mathbb{B}^{-1}} = \mathbb{I} - \mathbb{J}_{\mathbb{B}}$:

$$\begin{aligned}(2\mathbb{J}_{\mathbb{A}^{-\circledvee}} - \mathbb{I})(2\mathbb{J}_{\mathbb{B}^{-1}} - \mathbb{I}) &= (2\mathbb{J}_{\mathbb{A}}(-\mathbb{I}) + \mathbb{I})(\mathbb{I} - 2\mathbb{J}_{\mathbb{B}}) \\ &= (2\mathbb{J}_{\mathbb{A}}(-\mathbb{I}) + \mathbb{I})(-\mathbb{I})(2\mathbb{J}_{\mathbb{B}} - \mathbb{I}) \\ &= (2\mathbb{J}_{\mathbb{A}} - \mathbb{I})(2\mathbb{J}_{\mathbb{B}} - \mathbb{I})\end{aligned}$$

Self-dual property of DRS

When $\alpha = 1$, DRS is:

$$
\begin{aligned}
x^{k+1/2} &= \mathbb{J}_{\mathbb{B}}(z^k) \\
u^{k+1/2} &= \mathbb{J}_{\mathbb{B}^{-1}}(z^k) = z^k - x^{k+1/2} \\
x^{k+1} &= \mathbb{J}_{\mathbb{A}}(2x^{k+1/2} - z^k) \\
u^{k+1} &= \mathbb{J}_{\mathbb{A}^{-\circledcirc}}(2u^{k+1/2} - z^k) = x^{k+1} - x^{k+1/2} + u^{k+1/2} \\
z^{k+1} &= z^k + x^{k+1} - x^{k+1/2} = z^k + u^{k+1} - u^{k+1/2}
\end{aligned}
$$

Nicely reveals the symmetry. (Algorithmically no need to use both the x and u.) When $\alpha \neq 1$, similar but less elegant self-dual form.

(This self-dual property explains why the infimal postcomposition technique and the dualization technique yield the same ADMM.)

DYS

For

$$\underset{x\in\mathbb{R}^n}{\text{find}} \quad 0 \in (\mathbb{A}+\mathbb{B}+\mathbb{C})x,$$

where $\mathbb{A}$, $\mathbb{B}$, and $\mathbb{C}$ are maximal monotone and $\mathbb{C}$ is single-valued, Attouch–Théra dual is

$$\underset{u\in\mathbb{R}^n}{\text{find}} \quad 0 \in ((\mathbb{A}+\mathbb{C})^{-\circledcirc} + \mathbb{B}^{-1})u.$$

Fixed points of DYS:

$$\begin{aligned}\operatorname{Fix}\,(\mathbb{I} - \mathbb{J}_{\alpha\mathbb{B}} + \mathbb{J}_{\alpha\mathbb{A}}(\mathbb{R}_{\alpha\mathbb{B}} - \alpha\mathbb{C}\mathbb{J}_{\alpha\mathbb{B}}))&\\ \subseteq \operatorname{Zer}\,(\mathbb{A}+\mathbb{B}+\mathbb{C}) + \alpha\operatorname{Zer}\,((\mathbb{A}+\mathbb{C})^{-\circledcirc} + \mathbb{B}^{-1}).&\end{aligned}$$

Primal-dual DYS

We can make the FPI with DYS more explicitly primal-dual:

$$\begin{aligned} x^{k+1/2} &= \mathbb{J}_{\alpha\mathbb{B}}(z^k) \\ u^{k+1/2} &= \frac{1}{\alpha}(z^k - x^{k+1/2}) \\ x^{k+1} &= \mathbb{J}_{\alpha A}(2x^{k+1/2} - z^k - \alpha\mathbb{C}x^{k+1/2}) \\ u^{k+1} &= \frac{1}{\alpha}(x^{k+1} - x^{k+1/2} + \alpha u^{k+1/2}) \\ z^{k+1} &= z^k + x^{k+1} - x^{k+1/2}. \end{aligned}$$

Note $u^{k+1/2} \in \mathbb{B}x^{k+1/2}$, $-u^{k+1} \in \mathbb{A}x^{k+1} + \mathbb{C}x^{k+1/2}$. If $z^k \to z^\star$, then

$$x^{k+1/2} \to x^\star, \quad x^{k+1} \to x^\star, \quad x^\star \in \operatorname{Zer}(\mathbb{A} + \mathbb{B} + \mathbb{C})$$
$$u^{k+1/2} \to u^\star, \quad u^{k+1} \to u^\star, \quad u^\star \in \operatorname{Zer}((\mathbb{A} + \mathbb{C})^{-\circledvee} + \mathbb{B}^{-1})$$
$$z^k \to x^\star + \alpha u^\star.$$

DYS is not self-dual as it uses an evaluation of $\mathbb{C}$, a primal operation.