

Duality in Splitting Methods

Ernest K. Ryu

Large-Scale Convex Optimization:
Algorithms and Analyses via Monotone Operators

Stanford University
EE392F Spring 2023

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

Attouch–Théra duality

Duality in splitting methods

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

Fenchel duality

Attouch–Théra duality

Duality in splitting methods

Attouch–Théra duality

Consider

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

where \mathbf{A} and \mathbf{B} are maximal monotone.

Define $\mathbf{A}^{-\circlearrowleft}(u) = -\mathbf{A}^{-1}(-u)$.

Attouch–Théra dual monotone inclusion problem is

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

Attouch–Théra duality

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

$$\text{Zer}(\mathbf{A} + \mathbf{B}) \neq \emptyset \quad \Leftrightarrow \quad \text{Zer}(\mathbf{A}^{-\circledast} + \mathbf{B}^{-1}) \neq \emptyset,$$

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

Proof.

$$\begin{aligned} \exists x [0 \in (\mathbf{A} + \mathbf{B})x] &\Leftrightarrow \exists x, u [-u \in \mathbf{A}x, u \in \mathbf{B}x] \\ &\Leftrightarrow \exists x, u [-x \in \mathbf{A}^{-\circledast}u, x \in \mathbf{B}^{-1}u] \\ &\Leftrightarrow \exists u [0 \in (\mathbf{A}^{-\circledast} + \mathbf{B}^{-1})u]. \end{aligned}$$

□

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

Attouch–Théra vs. Fenchel duality

Attouch–Théra generalizes 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^*, u^*) satisfying

$$-u^* \in \mathbb{A}x^* \text{ and } u^* \in \mathbb{B}x^* \quad (1)$$

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

If only a primal solution x^* is provided, we must verify $0 \in \mathbb{A}x^* + \mathbb{B}x^*$.
How do we compute the Minkowski sum $\mathbb{A}x^* + \mathbb{B}x^*$?

Outline

Fenchel duality

Attouch–Théra duality

Duality in splitting methods

FBS

The FPI with FBS

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

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

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

Note $u^{k+1} \in \mathbf{B}x^{k+1}$. If $x^k \rightarrow x^*$, then

$$u^{k+1/2} \rightarrow u^*, \quad u^{k+1} \rightarrow u^*, \quad u^* \in \text{Zer}(\mathbf{A}^{-\odot} + \mathbf{B}^{-1}).$$

Characterization of fixed points of DRS

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

$$\text{Fix}(\mathbf{R}_{\alpha\mathbf{A}}\mathbf{R}_{\alpha\mathbf{B}}) \subseteq \text{Zer}(\mathbf{A} + \mathbf{B}) + \alpha\text{Zer}(\mathbf{A}^{-\heartsuit} + \mathbf{B}^{-1})$$

Proof.

$$z = \mathbf{R}_{\alpha\mathbf{A}}\mathbf{R}_{\alpha\mathbf{B}}z$$

$$\Leftrightarrow z + 2\mathbf{J}_{\alpha\mathbf{A}}(2\mathbf{J}_{\alpha\mathbf{B}} - \mathbf{I})z - 2\mathbf{J}_{\alpha\mathbf{B}}z = z, x = \mathbf{J}_{\alpha\mathbf{B}}z$$

$$\Leftrightarrow \mathbf{J}_{\alpha\mathbf{A}}(x - \alpha u) = x, z = x + \alpha u, u \in \mathbf{B}x$$

$$\Leftrightarrow x - \alpha u = x + \alpha v, v \in \mathbf{A}x, z = x + \alpha u, u \in \mathbf{B}x$$

$$\Leftrightarrow v = -u, v \in \mathbf{A}x, u \in \mathbf{B}x, z = x + \alpha u$$

$$\Leftrightarrow -u \in \mathbf{A}x, u \in \mathbf{B}x, z = x + \alpha u$$

$$\Leftrightarrow -u \in \mathbf{A}x, u \in \mathbf{B}x, -x \in \mathbf{A}^{-\heartsuit}u, x \in \mathbf{B}^{-1}u, z = x + \alpha u$$

$$\Rightarrow 0 \in (\mathbf{A} + \mathbf{B})x, 0 \in (\mathbf{A}^{-\heartsuit} + \mathbf{B}^{-1})u, z = x + \alpha u.$$

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:

$$x^{k+1/2} = \mathbb{J}_{\alpha\mathbf{B}}(z^k)$$

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

$$x^{k+1} = \mathbb{J}_{\alpha\mathbf{A}}(2x^{k+1/2} - z^k)$$

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

$$z^{k+1} = z^k + x^{k+1} - x^{k+1/2}.$$

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

$$\begin{aligned}x^{k+1/2} &\rightarrow x^*, & x^{k+1} &\rightarrow x^*, & x^* &\in \text{Zer}(\mathbf{A} + \mathbf{B}) \\u^{k+1/2} &\rightarrow u^*, & u^{k+1} &\rightarrow u^*, & u^* &\in \text{Zer}(\mathbf{A}^{-\odot} + \mathbf{B}^{-1}) \\z^k &\rightarrow x^* + \alpha u^*.\end{aligned}$$

Self-dual property of DRS

PRS and DRS are self-dual:

$$\mathbf{R}_A \mathbf{R}_B = \mathbf{R}_{A-\ominus} \mathbf{R}_{B^{-1}}$$

Follows from $\mathbf{J}_{A-\ominus} = \mathbf{I} + \mathbf{J}_A(-\mathbf{I})$ and $\mathbf{J}_{B^{-1}} = \mathbf{I} - \mathbf{J}_B$:

$$\begin{aligned}(2\mathbf{J}_{A-\ominus} - \mathbf{I})(2\mathbf{J}_{B^{-1}} - \mathbf{I}) &= (2\mathbf{J}_A(-\mathbf{I}) + \mathbf{I})(\mathbf{I} - 2\mathbf{J}_B) \\ &= (2\mathbf{J}_A(-\mathbf{I}) + \mathbf{I})(-\mathbf{I})(2\mathbf{J}_B - \mathbf{I}) \\ &= (2\mathbf{J}_A - \mathbf{I})(2\mathbf{J}_B - \mathbf{I})\end{aligned}$$

Self-dual property of DRS

When $\alpha = 1$, DRS is:

$$x^{k+1/2} = \mathbf{J}_B(z^k)$$

$$u^{k+1/2} = \mathbf{J}_{B^{-1}}(z^k) = z^k - x^{k+1/2}$$

$$x^{k+1} = \mathbf{J}_A(2x^{k+1/2} - z^k)$$

$$u^{k+1} = \mathbf{J}_{A^{-1}}(2u^{k+1/2} - z^k) = x^{k+1} - x^{k+1/2} + u^{k+1/2}$$

$$z^{k+1} = z^k + x^{k+1} - x^{k+1/2} = z^k + u^{k+1} - u^{k+1/2}$$

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

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

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

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

Fixed points of DYS:

$$\begin{aligned} \text{Fix} (\mathbf{I} - \mathbf{J}_{\alpha\mathbf{B}} + \mathbf{J}_{\alpha\mathbf{A}}(\mathbf{R}_{\alpha\mathbf{B}} - \alpha\mathbf{C}\mathbf{J}_{\alpha\mathbf{B}})) \\ \subseteq \text{Zer} (\mathbf{A} + \mathbf{B} + \mathbf{C}) + \alpha\text{Zer} ((\mathbf{A} + \mathbf{C})^{-\circledast} + \mathbf{B}^{-1}). \end{aligned}$$

Primal-dual DYS

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

$$x^{k+1/2} = \mathbf{J}_{\alpha\mathbf{B}}(z^k)$$

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

$$x^{k+1} = \mathbf{J}_{\alpha\mathbf{A}}(2x^{k+1/2} - z^k - \alpha\mathbf{C}x^{k+1/2})$$

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

$$z^{k+1} = z^k + x^{k+1} - x^{k+1/2}.$$

Note $u^{k+1/2} \in \mathbf{B}x^{k+1/2}$, $-u^{k+1} \in \mathbf{A}x^{k+1} + \mathbf{C}x^{k+1/2}$. If $z^k \rightarrow z^*$, then

$$x^{k+1/2} \rightarrow x^*, \quad x^{k+1} \rightarrow x^*, \quad x^* \in \text{Zer}(\mathbf{A} + \mathbf{B} + \mathbf{C})$$

$$u^{k+1/2} \rightarrow u^*, \quad u^{k+1} \rightarrow u^*, \quad u^* \in \text{Zer}((\mathbf{A} + \mathbf{C})^{-\textcircled{V}} + \mathbf{B}^{-1})$$

$$z^k \rightarrow x^* + \alpha u^*.$$

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