Duality in Splitting Methods

Ernest K. Ryu and Wotao Yin

Large-Scale Convex Optimization via Monotone Operators
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

\[
\begin{align*}
\text{minimize} \quad & f(x) + g(x), \\
\text{subject to} \quad & x \in \mathbb{R}^n
\end{align*}
\]

and dual

\[
\begin{align*}
\text{maximize} \quad & -f^*(-u) - g^*(u) \\
\text{subject to} \quad & u \in \mathbb{R}^n
\end{align*}
\]

generated by

\[
L(x, u) = f(x) + \langle x, u \rangle - g^*(u).
\]

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 $\nabla f$ and $\nabla g$ at $x$ sum to 0:

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

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

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

This is one of the many viewpoints of convex duality.
Outline

Fenchel duality

Attouch–Théra duality

Duality in splitting methods
Consider
\[
\begin{align*}
\text{find } & \quad 0 \in (A + B)x, \\
\text{where } & \quad A \text{ and } B \text{ are maximal monotone.}
\end{align*}
\]
Define \( A^{-\nabla}(u) = -A^{-1}(-u) \).

Attouch–Théra dual monotone inclusion problem is
\[
\begin{align*}
\text{find } & \quad 0 \in (A^{-\nabla} + B^{-1})u. \\
\end{align*}
\]
Attouch–Théra duality

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

\[ \text{Zer}(A + B) \neq \emptyset \iff \text{Zer}(A^\ominus + B^{-1}) \neq \emptyset, \]

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

**Proof.**

\[ \exists x \left[ 0 \in (A + B)x \right] \iff \exists x, u \left[ -u \in Ax, u \in Bx \right] \]
\[ \iff \exists x, u \left[ -x \in A^\ominus u, x \in B^{-1}u \right] \]
\[ \iff \exists u \left[ 0 \in (A^\ominus + B^{-1})u \right]. \]

(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

\[-u^* \in Ax^* \text{ and } u^* \in Bx^*\]  

(1)

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

If only a primal solution \(x^*\) is provided, we must verify \(0 \in Ax^* + Bx^*\).

How do we compute the Minkowski sum \(Ax^* + Bx^*\)?
Outline

Fenchel duality

Attouch–Théra duality

Duality in splitting methods
The FPI with FBS

\[
x^{k+1/2} = x^k - \alpha A x^k
\]
\[
x^{k+1} = J_{\alpha B} x^{k+1/2}
\]

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

\[
x^{k+1/2} = x^k - \alpha A x^k
\]
\[
u^{k+1/2} = -A x^k
\]
\[
x^{k+1} = J_{\alpha B} x^{k+1/2}
\]
\[
u^{k+1} = \alpha^{-1} (x^{k+1/2} - x^{k+1}).
\]

Note \(u^{k+1} \in B x^{k+1}\). If \(x^k \to x^*\), then

\[
u^{k+1/2} \to u^*, \quad u^{k+1} \to u^*, \quad u^* \in \text{Zer} (A^{-\ominus} + B^{-1}).
\]
Characterization of fixed points of DRS

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

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

Proof.

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

$$\iff z + 2J_{\alpha A} (2J_{\alpha B} - I) z - 2J_{\alpha B} z = z, x = J_{\alpha B} z$$

$$\iff J_{\alpha A} (x - \alpha u) = x, z = x + \alpha u, u \in Bx$$

$$\iff x - \alpha u = x + \alpha v, v \in A x, z = x + \alpha u, u \in Bx$$

$$\iff v = -u, v \in A x, u \in Bx, z = x + \alpha u$$

$$\iff -u \in A x, u \in Bx, z = x + \alpha u$$

$$\iff -u \in A x, u \in Bx, -x \in A^{-\nabla} u, x \in B^{-1} u, z = x + \alpha u$$

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

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

Duality in splitting methods
Primal-dual DRS

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

\[
\begin{align*}
    x^{k+1/2} &= J_{\alpha B}(z^k) \\
    u^{k+1/2} &= \frac{1}{\alpha}(z^k - x^{k+1/2}) \\
    x^{k+1} &= J_{\alpha A}(2x^{k+1/2} - z^k) \\
    u^{k+1} &= \frac{1}{\alpha}(x^{k+1} - x^{k+1/2} + \alpha u^{k+1/2}) \\
    z^{k+1} &= z^k + x^{k+1} - x^{k+1/2}.
\end{align*}
\]

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

\[
\begin{align*}
    x^{k+1/2} &\to x^*, \\
    x^{k+1} &\to x^*, \\
    x^* &\in \text{Zer}(A + B) \\
    u^{k+1/2} &\to u^*, \\
    u^{k+1} &\to u^*, \\
    u^* &\in \text{Zer}(A^{-\odot} + B^{-1}) \\
    z^k &\to x^* + \alpha u^*.
\end{align*}
\]
Self-dual property of DRS

PRS and DRS are self-dual:

$$R_A R_B = R_{A - \varnothing} R_{B - 1}$$

Follows from $J_{A - \varnothing} = I + J_A(-I)$ and $J_{B - 1} = I - J_B$:

$$(2J_{A - \varnothing} - I)(2J_{B - 1} - I) = (2J_A(-I) + I)(I - 2J_B)$$

$$= (2J_A(-I) + I)(-I)(2J_B - I)$$

$$= (2J_A - I)(2J_B - I)$$
Self-dual property of DRS

When $\alpha = 1$, DRS is:

$$
\begin{align*}
\frac{x^{k+1}}{2} &= J_B(z^k) \\
\frac{u^{k+1}}{2} &= J_{B^{-1}}(z^k) = z^k - \frac{x^{k+1}}{2} \\
x^{k+1} &= J_A(2\frac{x^{k+1}}{2} - z^k) \\
u^{k+1} &= J_{A^{-\otimes}}(2\frac{u^{k+1}}{2} - z^k) = x^{k+1} - \frac{x^{k+1}}{2} + \frac{u^{k+1}}{2} \\
z^{k+1} &= z^k + x^{k+1} - \frac{x^{k+1}}{2} = z^k + \frac{u^{k+1}}{2} - \frac{u^{k+1}}{2}
\end{align*}
$$

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.)
For
\[ \text{find } x \in \mathbb{R}^n \quad 0 \in (A + B + C)x, \]

where $A$, $B$, and $C$ are maximal monotone and $C$ is single-valued, 
Attouch–Théra dual is
\[ \text{find } u \in \mathbb{R}^n \quad 0 \in ((A + C)^{-\ominus} + B^{-1})u. \]

Fixed points of DYS:
\[ \text{Fix } (I - J_{\alpha B} + J_{\alpha A}(R_{\alpha B} - \alpha CJ_{\alpha B})) \]
\[ \subseteq \text{Zer } (A + B + C) + \alpha \text{Zer } ((A + C)^{-\ominus} + B^{-1}). \]
Primal-dual DYS

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

\[ x^{k+1/2} = J_{\alpha B}(z^k) \]
\[ u^{k+1/2} = \frac{1}{\alpha}(z^k - x^{k+1/2}) \]
\[ x^{k+1} = J_{\alpha A}(2x^{k+1/2} - z^k - \alpha Cx^{k+1/2}) \]
\[ u^{k+1} = \frac{1}{\alpha}(x^{k+1} - x^{k+1/2} + \alpha u^{k+1/2}) \]
\[ z^{k+1} = z^k + x^{k+1} - x^{k+1/2}. \]

Note \( u^{k+1/2} \in Bx^{k+1/2}, -u^{k+1} \in Ax^{k+1} + Cx^{k+1/2} \). If \( z^k \to z^* \), then

\[ x^{k+1/2} \to x^*, \quad x^{k+1} \to x^*, \quad x^* \in \text{Zer}(A + B + C) \]
\[ u^{k+1/2} \to u^*, \quad u^{k+1} \to u^*, \quad u^* \in \text{Zer}((A + C)^{-\odot} + B^{-1}) \]
\[ z^k \to x^* + \alpha u^*. \]

DYS is not self-dual as it uses an evaluation of \( C \), a primal operation.