

The Concept of Duality in Convexity Theory

S. Artstein-Avidan and V.D. Milman

Tel Aviv University

Cortona, 2007

A. Let $Cvx(\mathbb{R}^n)$ denote the family of all *lower-semi-continuous* convex functions $\{f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\pm\infty\}\}$.

Define on \mathbb{R}^n a scalar product (x, y) .

Then the Legendre transform is defined by

$$(\mathcal{L}f)(x) = \sup_{y \in \mathbb{R}^n} \{(x, y) - f(y)\}$$

and usually associated with the notion of “duality” on this class of functions.

The transform $\mathcal{L} : Cvx(\mathbb{R}^n) \rightarrow Cvx(\mathbb{R}^n)$ is onto and 1–1. It has many remarkable properties.

Some of them:

(i) $\mathcal{L}\mathcal{L}f = f$ (involution);

(ii) If $f \leq g$ then $\mathcal{L}f \geq \mathcal{L}g$;

(iii) Let $\widehat{\inf}(f, g) = \sup\{h : h \in Cvx(\mathbb{R}^n) \text{ and } h \leq \inf(f, g)\}$; then

$\mathcal{L}[\widehat{\inf}(f, g)] = \sup(\mathcal{L}f, \mathcal{L}g)$ and

$\mathcal{L}\sup(f, g) = \widehat{\inf}(\mathcal{L}f, \mathcal{L}g)$;

(iv) $\mathcal{L}(f+g)(x) := \inf_{x=x_1+x_2} \{(\mathcal{L}f)(x_1) + (\mathcal{L}g)(x_2)\}$,

i.e. sum is transferred to convolution and vice versa.

B. Convex sets containing 0

We are very familiar with the notion of polarity for convex sets although two cases should be distinguished here:

1. \mathcal{K}_o – the family of all compact convex bodies with 0 in their interiors;
2. \mathcal{K} – the family of all convex closed bodies containing 0 (no boundedness, and 0 may not be in their interiors).

However, the duality formula looks the same for both cases: if $K \in \mathcal{K}$ (or $K \in \mathcal{K}_\circ$)

$$K^\circ := \{x \in \mathbb{R}^n \mid (x, y) \leq 1 \text{ for } \forall y \in K\}.$$

Again, a very concrete formula and many remarkable properties.

To unify our discussion, we will identify $K \in \mathcal{K}$ with the function $\mathbf{1}_K(x) = \begin{cases} 1 & x \in K \\ 0 & x \notin K \end{cases}$.

C. We may also think about the family of norms on \mathbb{R}^n and consider a bilinear (positive definition) form to define the duality for this class.

More examples:

D. Let $LC(\mathbb{R}^n)$ be the class of (all) upper semicontinuous non-negative functions s.t. their logs are concave (in short, “log-concave” functions.)

What should be a duality transform for this class (and does it exist in some natural sense)?

In Artstein-Klartag-Milman [AKM], we introduced a duality by formula: for $f \in LC(\mathbb{R}^n)$,

$$f^\circ(x) = \inf_{y \in \mathbb{R}^n} \frac{e^{-(x,y)}}{f(y)}.$$

(One may see that $-\log f^\circ = \mathcal{L}(-\log f)$.)

E. s -concave functions (for $s > 0$):

$Conc_s^+(\mathbb{R}^n)$ is the family of all bounded s -concave functions $f : \mathbb{R}^n \rightarrow \mathbb{R}^+$ (i.e. $f^{1/s}$ is concave on its convex support) and s.t. $f(0) > 0$.

Again we define, in [AKM], a duality for this class by

$$\mathcal{L}_s f = \inf_{y: f(y) > 0} \frac{(1 - (x, y))_+^s}{f(y)}.$$

Nobody (for decades or centuries) has asked justification for the definitions and notions of duality for classes A,B, and C.

However, the dualities for classes D and E are new notions and must be justified

Theorem (Arstein-Avidan, Milman). *Let*
 $T : Cvx(\mathbb{R}^n) \rightarrow Cvx(\mathbb{R}^n)$ *(1-1 and onto)*
satisfying

1. $T \cdot T\varphi = \varphi$ *(for any* $\varphi \in Cvx(\mathbb{R}^n)$ *);*
2. $\varphi \leq \psi$ *implies* $T\varphi \geq T\psi$.

Then T *is essentially the Legendre transform* \mathcal{L} :
 $\exists c_0 \in \mathbb{R}, v_0 \in \mathbb{R}^n$, *symmetric linear* $B \in GL_n$ *s.t.*

$$(T\varphi)(x) = (\mathcal{L}\varphi)(Bx + v_0) + (x, v_0) + C_0.$$

(discuss)

Corollary (for our case D). *Let*

$T : LC(\mathbb{R}^n) \rightarrow LC(\mathbb{R}^n)$ *satisfy, $\forall f \in LC(\mathbb{R}^n)$*

1. $T \cdot Tf = f;$

2. $f \leq g$ *implies* $Tf \geq Tg.$

Then $\exists 0 < c_0 \in \mathbb{R}^n, v_0 \in \mathbb{R}^n$ and a symmetric operator $B \in GL_n$ s.t. T is defined by

$$(Tf)(x) = c_0 e^{-(v_0, x)} \inf_y \frac{e^{-(Bx + v_0, y)}}{f(y)},$$

i.e. $(Tf)(x) = c_0 e^{-(v_0, x)} f^\circ(Bx + v_0).$

The Concept of Duality

Definition. Let S be a set of functions defined on \mathbb{R}^n . We say that a transform $T : S \rightarrow S$ (onto, 1–1) generates a duality if

1. $\forall f \in S, T \cdot Tf = f$;
2. $\forall f, g \in S, f \leq g$ implies $Tf \geq Tg$.

The above theorem and corollary describe all existing dualities on classes A and D.

A recent result by Böröczky–Schneider implies the same kind of characterization for the case B.1 (i.e. the standard duality is, essentially, the only existing duality for the class of compact convex bodies).

Our method gives another proof of this and also provides the same result for the class B.2. The duality C is also characterized in the same way, which follows from joining Gruber’s (1992) and Böröczky–Schneider’s results (the whole weight of the proof lies within Gruber’s paper, but it was written and developed for a different problem).

The result for class E. We recall

$$\text{Conc}_s^+(\mathbb{R}^n) = \{f \in \text{Conc}_s(\mathbb{R}^n) : f(0) > 0 \\ \text{and } \exists C \text{ s.t. } f \leq C\}.$$

Theorem (Artstein–Milman). *Let*

$T : \text{Conc}_s^+(\mathbb{R}^n) \rightarrow \text{Conc}_s^+(\mathbb{R}^n)$, *1–1, onto*

1. $T \cdot Tf = f$;
2. $f \leq g$ *implies* $Tf \geq Tg$.

Then, there exists a constant $C_0 \in \mathbb{R}$ and an invertible symmetric linear transformation

$B \in \text{GL}_n$ *such that*

$$(Tf)(x) = C_0 \inf_{(y:f(By)>0)} \frac{(1 - \langle x, y \rangle)_+^s}{f(By)}.$$

Consequences, remarks

1. The condition of involution is not needed.

Instead, we need that

1. T is onto, 1-1; and

2a. $\varphi \leq \psi \Rightarrow T\varphi \geq T\psi$;

2b. $T\varphi \leq T\psi \Rightarrow \varphi \geq \psi$

(and the result is not true without the “onto” condition).

2. Consider now the order preserving 1–1 and onto maps from $Cvx(\mathbb{R}^n)$ to itself.

Corollary (Artstein–Milman). *Let*

$$\mathcal{F} : Cvx(\mathbb{R}^n) \longrightarrow Cvx(\mathbb{R}^n), \text{ 1–1, onto.}$$

Let

1. $\varphi \leq \psi \Rightarrow \mathcal{F}\varphi \leq \mathcal{F}\psi;$

2. $\mathcal{F}\varphi \leq \mathcal{F}\psi \Rightarrow \varphi \leq \psi.$

Then $\exists C_0 \in \mathbb{R}, C_1 > 0, v_0, v_1 \in \mathbb{R}^n$ and $B \in GL^n$, s.t.

$$(\mathcal{F}\varphi)(x) = C_0 + (v_1, x) + C_1\varphi(Bx + v_0).$$

(to compare with P. Gruber’s result and approach of Böröczky–Schneider)