Arithmetic Duality for Norms

Yao-Liang Yu
Department of Computing Science
University of Alberta
Edmonton, AB T6G 2E8, Canada
yaoliang@cs.ualberta.ca

June 11, 2012

Abstract

Some relatively easy but occasionally useful results about norm duality are recorded.

1 Norms

Recall that norms, defined on some vector space $\mathcal{X}$ (with scalar field $\mathbb{K}$), are functions $\|\cdot\| : \mathcal{X} \mapsto \mathbb{R}_+$ that satisfy the following three properties:

1. (positive definiteness) $\|x\| = 0 \iff x = 0$;
2. (homogeneity) $\|\alpha x\| = |\alpha|\|x\|, \forall \alpha \in \mathbb{K}$;
3. (triangle inequality) $\|x + y\| \leq \|x\| + \|y\|$.

It follows immediately from the last two properties that (pseudo-)norms are actually convex functions, and Lipschitz continuous on the whole space $\mathcal{X}$. Note that norms are not differentiable at the origin (including those induced by inner products, if you did not forget to take the square root ☺).

It follows easily from the definition that the (open) norm ball $B := \{x \in \mathcal{X} : \|x\| < 1\}$ is convex, balanced, absorbing and bounded. In fact, the other direction is also true. Let us put aside the positive definiteness condition temporarily (*i.e.* consider pseudo-norms). Any convex, balanced and absorbing set, say B, induces a pseudo-norm on the underlying vector space through the Minkowski functional (known also as the gauge function):

$$p(x) := \inf\{t \in \mathbb{R} : x \in tB\}.$$

It is an easy exercise to show that the (open) pseudo-norm ball $\{x \in \mathcal{X} : p(x) < 1\}$ in this case is B (provided that B is open, whatever that means ☺). Finally, we mention that $p(\cdot)$ is a norm iff its ball B does not contain a line.

2 Norm Duality

Given some norm $\|\cdot\|$ on the vector space X, we define its *dual* norm on the (topological) dual space $\mathcal{X}^*$ as:

$$\|x^*\|^* := \sup_{x \in B} \langle x; x^* \rangle,$$

where B is the norm ball of $\|\cdot\|$, and the notation $\langle \cdot; \cdot \rangle$ signifies the dual paring. From the definition we have immediately the "Cauchy-Schwarz inequality"

$$\langle x; x^* \rangle \leq \|x\| \|x^*\|^*.$$

It is clear from the definition that the dual norm is nothing but the support function of B [1], whence the two balls B and $B^* := \{x^* \in \mathcal{X}^* : \|x^*\|^* < 1\}$ are polar to each other. It is an amazing fact that "duality",

[1]hence is the Fenchel conjugate of the indicator function $\delta_B(x)$ whose value is 0 on B and infinity otherwise.

in many disguises, reflects (and can be characterized (!) by) "order" relations. For sets, in particular norm balls, we can roughly say that big ones are dual to small ones, and vice versa (*i.e.* if we order sets by their "sizes", then duality reverses this order).

It can be shown that the bidual norm $\|\cdot\|^{**}$, defined on $\mathcal{X}^{**}{}_{|\mathcal{X}}$ (whatever this notation means ☺), is nothing but the original norm $\|\cdot\|$.

3 Norm Arithmetics

We now consider some elementary "arithmetic" operations on norms. Take two norms $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$, where the underlying vector space is always fixed to be $\mathcal{X}$. Let B_1 and B_2 be the respective norm balls.

The first operation we consider is the summation. Define

$$h(x) := \|x\|_{(1)} + \|x\|_{(2)}.$$

It should be straightforward to verify that $h(\cdot)$ is indeed a norm (on $\mathcal{X}$). The norm ball of h turns out to be the inverse summation[2] of B_1 and B_2:

$$B_h := \{x \in \mathcal{X} : h(x) < 1\} = \bigcup_{\lambda \in [0,1]} \{\lambda B_1 \cap (1-\lambda) B_2\} =: B_1 \sharp B_2.$$

Naturally, the next operation to consider, is the inverse summation. Define

$$k(x) := \sup_{\lambda \in [0,1]} \inf_{x = x_1 + x_2} \lambda \|x_1\|_{(1)} + (1-\lambda)\|x_2\|_{(2)} = \inf_{x = x_1 + x_2} \|x_1\|_{(1)} \vee \|x_2\|_{(2)} =: \|x\|_{(1)} \sharp \|x\|_{(2)}.$$

It might take some effort to convince oneself that $k(\cdot)$ is indeed a norm (but should not be too hard ☺). Note that we have deliberately defined $k(\cdot)$ in a way that resembles the $\sharp$ operation for sets, while the last (cleaner) equality implies immediately that the norm ball

$$B_k := \{x \in \mathcal{X} : k(x) < 1\} = B_1 + B_2.$$

Next we consider the maximum. Define

$$f(x) := \|x\|_{(1)} \vee \|x\|_{(2)}.$$

Easily verified that $f(\cdot)$ is indeed a norm. Its norm ball is the intersection:

$$B_f := \{x \in \mathcal{X} : f(x) < 1\} = B_1 \cap B_2.$$

It is again natural to consider the "dual" notion, this time, the "minimum". Well, not quite, since the minimum of two norms need not be a norm (although positive definiteness and homogeneity still hold). Note that we are not saying that the minimum of two norms can never be a norm. A trivial example is when one norm "dominates" the other. Perhaps a little bit surprisingly (is it? one cannot "bend" convex functions!), this is the only possibility[3]. So we need to bring in convexity explicitly. Define:

$$g(x) := \operatorname{conv}\{\|\cdot\|_{(1)} \wedge \|\cdot\|_{(2)}\} = \inf_{x = x_1 + x_2} \|x_1\|_{(1)} + \|x_2\|_{(2)} =: \|\cdot\|_{(1)} \Box \|\cdot\|_{(2)},$$

where $\Box$ signifies the infimal convolution and conv denotes the (closed) convex hull (*i.e.* the largest closed convex minorant)[4]. One is encouraged to verify that the norm ball satisfies:

$$B_g := \{x \in \mathcal{X} : g(x) < 1\} = \operatorname{conv}\{B_1 \cup B_2\}.$$

Duality is apparent in the last two operations, if one interprets (as usual) $\vee$ as $\cup$, $\wedge$ as $\cap$ and vice versa (they look similar, don't they? ☺). We have also a nice example that illustrates that small balls (intersection)

[2]For a justification of the name "inverse sum", refer to T. Rockafellar's classic book: Convex Analysis.

[3]Suppose $\exists x, y \in \mathcal{X}$ such that $\|x\|_{(1)} > \|x\|_{(2)}, \|y\|_{(1)} < \|y\|_{(2)}$ and yet $\|\cdot\|_{(1)} \wedge \|\cdot\|_{(2)}$ is a norm. Restricting all relevant norms into the two dimensional subspace spanned by x and y. Since all norms are equivalent in finite dimensional spaces, we easily conclude that $\|\cdot\|_{(1)} = \|\cdot\|_{(2)}$ on that two dimensional subspace, an immediate contradiction.

[4]In general, the infimal convolution differs from the convex hull; they coincide on positively homogeneous and subadditive functions (*i.e.* pseudo-norms if $\mathbb{K} = \mathbb{R}$).

are dual to big balls (union). It is probably not hard to guess that the first two operations are dual to each other too.

It is tempting to generalize the above results to arbitrary (but finite, since we do not want to mess with limits ☺) number of norms. For this, we need to convince ourselves that the inverse sum and the infimal convolution are both commutative and associative. Well, yes, we also need to show sum and maximum are commutative and associative......

Instead of repeating the arguments, we summarize directly the results in Table 1.

operation	formula	ball	dual	dual formula
$+$	$+_i \|\cdot\|_{(i)}$	$\sharp_i B_i$	$\sharp$	$\sharp_i \|\cdot\|_{(i)}^*$
$\sharp$	$\sharp_i \|\cdot\|_{(i)}$	$+_i B_i$	$+$	$+_i \|\cdot\|_{(i)}^*$
$\vee$	$\vee_i \|\cdot\|_{(i)}$	$\cap_i B_i$	$\square$	$\square_i \|\cdot\|_{(i)}^*$
$\square$	$\square_i \|\cdot\|_{(i)}$	$\operatorname{conv} \cup_i B_i$	$\vee$	$\vee_i \|\cdot\|_{(i)}^*$

Table 1: Duality relations among elementary arithmetic operations on norms.

We end this section with a theorem that summarizes and generalizes the above four arithmetics. (In retrospect, this theorem is what needs to be kept in mind.)

Theorem 1 *The two norms,* $\left(\sum_{i=1}^n \|x\|_{(i)}^p\right)^{1/p}$ *and* $\inf_{x=\sum_{i=1}^n x_i} \left(\sum_{i=1}^n \|x_i\|_{(i)}^{*q}\right)^{1/q}$, *are dual to each other, where* $1/p + 1/q = 1, p \geq 1$.

Proof: It follows from the definition. (Prove me wrong if you are skeptical. Oh yeah!) ■

4 Discussion

Unlike previous section, let us take two norms defined on two different vector spaces: $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$.

Connection to interpolation spaces.

5 Examples

5.1 Vector example: $\alpha\|\cdot\|_1 + \|\cdot\|_2$

Let $\mathcal{X} = \mathbb{R}^d$, and define $\|x\| := \alpha\|x\|_1 + \|x\|_2 = \alpha\sum_{i=1}^d |x_i| + \sqrt{\sum_{i=1}^d x_i^2}$, where $\alpha > 0$ is a constant. The dual norm of $\|\cdot\|$ is

$$\|y\|^* = \min_z\{\|y - z\|_2 \vee \|z/\alpha\|_\infty\} = \{\min_{z,t} t, \text{ s.t. } \|y - z\|_2 \leq t, \|z\|_\infty \leq \alpha t\}. \tag{1}$$

It is clear that we can assume $y \geq 0$ and y is arranged in decreasing order, hence w.l.o.g., we can assume $z \geq 0$ and z is arranged in decreasing order too. The key observation is that the right-hand side of (1) depends only on $s := \|z\|_\infty$, *i.e.*

$$(1) = \min_{s \leq \alpha t} t, \text{ s.t. } (y_1 - s)^2 + \sum_{i=2}^d (y_i - y_i \wedge s)^2 \leq t^2.$$

It is clear that $t \leq \|y/\alpha\|_\infty$, hence at optimum, we may have $s = \alpha t$ since the left-hand side of the above constraint is a decreasing function of s. Therefore

$$(1) = \min_t t, \text{ s.t. } \sum_{i=1}^d (y_i - y_i \wedge (\alpha t))^2 \leq t^2.$$

We realize that the left-hand side of the constraint is decreasing in t, hence this one dimensional problem can be very quickly solved. In fact, an analytic solution exists: By checking $d+1$ points $\{y_1/\alpha, \ldots, y_d/\alpha, y_{d+1} := 0\}$ we know the optimal t lies in $[y_{j+1}/\alpha, y_j/\alpha]$ (where $j \in [d]$). Solving the quadratic inequality we obtain

$$\tilde{t} = \begin{cases} \left(\alpha\|y_{1:j}\|_1 - \sqrt{\alpha^2\|y_{1:j}\|_1^2 - (j\alpha^2 - 1)\|y_{1:j}\|_2^2}\right)/(j\alpha^2 - 1), & d > 1 \\ y_1/(1+\alpha), & d = 1 \end{cases},$$

$t = (\tilde{t} \wedge y_j) \vee y_{j+1}$ and $z = t \wedge y$. Moreover,

$$\partial\|y\|^* = \frac{y - z}{\|y - z\|_2 + \alpha\|y - z\|_1}.$$

Similarly, we can derive the dual norm of $\|\cdot\| := (\alpha\|\cdot\|_1) \vee \|\cdot\|_2$. Although this case is of less interest to us, we nevertheless record the result (as before assume w.l.o.g. $y \geq 0$ and y is arranged in decreasing order):

$$\|y\|^* = \min_z \|y - z\|_2 + \|z/\alpha\|_\infty = \min_s s/\alpha + \sqrt{\sum_{i=1}^d (y_i - y_i \wedge s)^2}.$$

The right-hand side is apparently a (univariate) convex function of s, hence can be quickly solved. Again, an analytic solution can be found by checking the $d+1$ points $\{y_1, \ldots, y_d, y_{d+1} := 0\}$. The subdifferential $\partial\|y\|^* = \frac{y-z}{\|y-z\|_2}$.

5.2 Matrix example: $\alpha\|\cdot\|_{1,1} + \|\cdot\|_{\text{tr}}$

Let $\mathcal{X} = \mathbb{R}^{m\times n}$, and define $\|X\| := \alpha\|X\|_{1,1} + \|X\|_{\text{tr}} = \alpha\sum_{ij}|X_{ij}| + \sum_{i=1}^{m\wedge n}\sigma_i$, where $\alpha > 0$ is a constant, $\{\sigma_i\}$ are the singular values of X. The dual norm of $\|\cdot\|$ is

$$\|Y\|^* = \min_Z\{\|Y - Z\|_{\text{sp}} \vee \|Z/\alpha\|_{\infty,\infty}\} = \{\min_{Z,t} t, \text{ s.t. } \|Y - Z\|_{\text{sp}} \leq t, |Z_{ij}| \leq \alpha t\}. \quad (2)$$