1 Dual norm

Definition 1. Let $\| \cdot \|$ be a norm on $\mathbb{R}^n$. Its dual norm is

$$\|u\|_* = \sup_x \{ u^T x \mid \|x\| \leq 1 \}$$

Problem 1. Verify that $\| \cdot \|_*$ is a norm.

Problem 2. Let $1/p + 1/q = 1$, where $p, q \in [1, \infty]$. Show that $\| \cdot \|_q$ is dual to $\| \cdot \|_p$. In particular, $\ell_2$-norm is self-dual, and $\| \cdot \|_\infty$ is dual to $\| \cdot \|_1$. (Hint: Use Hölder’s inequality: $\sum |u_i v_i| \leq \|u\|_p \|v\|_q$)

2 Fenchel Conjugate

Definition 2. The Fenchel conjugate of a function $f$ is

$$f^*(z) = \sup_{x \in \text{dom } f} \{ x^T z - f(x) \}$$

Note 1. Some remarks:

- $f^*$ is the pointwise sup over $x$ of affine functions of $z$, so it is convex.
- Fenchel’s inequality: $f(x) + f^*(y) \geq x^T y$, $\forall x, y$.
- If $f$ is closed and convex, then $f^{**} = f$.

Problem 3. Find $f^*$ when

- $f(x) = -\log(x)$, $x > 0$
- $f(x) = \frac{1}{2} x^T Q x$, $Q > 0$, $x \in \mathbb{R}^n$

Definition 3. The indicator function of a set $C$ is

$$\mathbb{I}_C(x) = \begin{cases} 0 & \text{if } x \in C \\ \infty & \text{otherwise} \end{cases}$$

Problem 4. Prove that the conjugate of $f = \| \cdot \|$ is the indicator function of dual unit norm ball, that is $f^* = \mathbb{I}_{\| \cdot \|_* \leq 1}$. 

3 Subgradient

Definition 4. A vector \( g \in \mathbb{R}^n \) is called a subgradient at point \( y \), if for all \( x \in \text{dom} f \), it holds that
\[
f(x) \geq f(y) + \langle g, x - y \rangle
\]

Definition 5. The set \( \partial f(y) \) of all subgradients at \( y \) is called the subdifferential of \( f \) at \( y \).

Note 2. Subgradient gives affine global underestimator of \( f \). Some properties:
- \( 0 \in \partial f(y) \iff f(x) \geq f(y) \) for all \( x \in \text{dom} f \)
- If \( f \) differentiable at \( x \), then \( \partial f(x) = \{ \nabla f(x) \} \)
- If \( \partial f(x) = \{ g \} \), then \( f \) is differentiable and \( g = \nabla f(x) \)
- Scaling: \( \forall \alpha > 0, \partial (\alpha f)(x) = \alpha \partial f(x) = \{ \alpha g | g \in \partial f(x) \} \)
- Addition: \( \partial (f + g)(x) = \partial f(x) + \partial g(x) = \{ u + v | u \in \partial f(x), v \in \partial g(x) \} \)
- Maximum function: If \( f(x) := \max_{i=1,\ldots,m} f_i(x) \), \( \partial f(x) = \text{conv} (\bigcup_i \partial f_i(x) | f_i(x) = f(x)) \)

Problem 5. Find \( \partial f(x) \) when
- \( f = | \cdot | \) for all \( x \in \mathbb{R} \)
- \( f = \| \cdot \|_2 \) for all \( x \in \mathbb{R}^n \)
- \( f = \| \cdot \|_\infty \) for all \( x \in \mathbb{R}^n \)

Problem 6. For \( f \) closed and convex, prove that \( \forall x, y \)
\[
y \in \partial f(x) \iff x \in \partial f^*(y) \iff x^T y = f(x) + f^*(y)
\]

4 Application: proximal mapping

Definition 6. For closed convex \( h \), the proximal mapping of \( h \) is
\[
\text{prox}_h(x) = \arg\min_u \{ h(u) + \frac{1}{2} \| u - x \|^2 \}
\]

Definition 7. In the proximal gradient algorithm, the update step with step size \( t > 0 \) is
\[
x^+ = \text{prox}_{th}(x - t \nabla f(x))
\]

Problem 7. Find \( \text{prox}_h \) for \( h = 0, h = \mathbb{1}_{\mathcal{C}} \). What are the associated proximal gradient methods?

Problem 8. Show: \( u = \text{prox}_h(x) \iff x - u \in \partial h(u) \)

Problem 9. Prove Moreau decomposition: \( x = \text{prox}_h(x) + \text{prox}_{h^*}(x), \forall x \)

Problem 10. Suppose you know how to compute projection onto the unit norm ball for any norm, how can you leverage that to compute the prox operator for \( h = \| \cdot \| \)?