Exercise sheet 3

You can return your solutions to questions 5 and 8 to get them marked. If so, please upload them on Moodle before Tuesday 7/3 at 5pm.

1. Compute the proximal operator of the following functions:
   (i) $f(x) = (1/2)x^T Ax$ where A is positive definite
   (ii) $f(x) = -\sum_{i=1}^n \log x_i$
   (iii) $f(x) = \|x\|_2$

2. Let $\lambda > 0$ and let $f(x) = \lambda g(x/\lambda)$ where g is a convex function. Express the proximal operator of f in terms of that of g.

3. (*Bregman subgradient method*) Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a smooth and strictly convex function, and let D_ϕ be its Bregman divergence. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a potentially nonsmooth convex function, and consider the following Bregman subgradient method:

$$x_{k+1} = \operatorname*{argmin}_{x\in\mathbb{R}^n} \{t_k \langle g_k, x - x_k\rangle + D_\phi(x|x_k)\}$$

   where $g_k \in \partial f(x_k)$.
   (a) Show that for $\phi(x) = \|x\|_2^2/2$ we recover the usual subgradient method.
   (b) Let $\|\cdot\|$ be an arbitrary norm on $\mathbb{R}^n$. We assume that ϕ is 1-strongly convex with respect to $\|\cdot\|$. Show that the iterates of the Bregman subgradient method satisfy:

$$D_\phi(x^*|x_{k+1}) \le D_\phi(x^*|x_k) + \frac{1}{2}\|t_k g_k\|_*^2 + t_k(f(x^*) - f(x_k))$$

   where $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$. Deduce:

$$f_{\text{best},k} - f^* \le \frac{D_\phi(x^*\|x_0)}{\sum_{i=0}^k t_i} + \frac{\sum_{i=0}^k t_i^2\|g_i\|_*^2}{2\sum_{i=0}^k t_i}.$$

   where $f_{\text{best},k} = \min\{f(x_0), \dots, f(x_k)\}$.

4. Examples: Compute the Fenchel conjugates of the following functions
   (a) $f(x) = \frac{1}{2}x^T Ax + b^T x$
   (b) $f(x) = \|x\|$ for some norm $\|\cdot\|$.

5. (*) (*Moreau's identity*) Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex and lower semicontinuous. Prove Moreau's identity: $\mathbf{prox}_{f^*}(x) = x - \mathbf{prox}_f(x)$.

6. Compute the Lagrangian dual of the following optimization problems
   (a) $\min_{x\in\mathbb{R}^n} \langle c, x\rangle$ subject to $x \ge 0, Ax = b$, where $A \in \mathbb{R}^{m\times n}$
   (b) $\min_{X\in\mathbf{S}^n} \operatorname{tr}(CX)$ subject to $X \succeq 0, A(X) = b$, where $A : \mathbf{S}^n \to \mathbb{R}^m$ is a linear map.

(c) $\min \|x\|_1$ subject to $Ax = b$

7. (*Smoothing via conjugate functions*)

   (a) Assume f is a convex function given as $f(x) = h^*(Ax + b)$ where h is convex lower-semicontinuous, defined on a *compact domain* D, i.e.,

   $$f(x) = \max_{y \in D} \left\{ y^T(Ax + b) - h(y) \right\}.$$

   Let d be a nonnegative convex function defined on D which is 1-strongly convex with respect to the Euclidean norm, and consider for $\mu > 0$ the function

   $$f_\mu(x) = (h + \mu d)^*(Ax + b).$$

   Show that f_μ is smooth, with smoothness parameter (with respect to Euclidean norm) $L = \|A\|^2/\mu$ where $\|A\|$ is the operator norm of A. Further, show that

   $$f - \mu R \leq f_\mu \leq f$$

   where $R = \max_{d \in D} d(x)$.

   (b) Examples: (i) let $f(x) = \|Ax + b\|_1$ which we can write as $f(x) = h^*(Ax + b)$ where h is the indicator function of the unit ℓ_∞ ball. Compute $f_\mu(x)$ explicitly for $d(y) = \|y\|_2^2/2$, and for $d(y) = \sum_i 1 - \sqrt{1 - y_i^2}$ (check that both functions are 1-strongly convex).

8. (*) Consider the following optimization problem for denoising a one-dimensional signal $b \in \mathbb{R}^n$:

   $$\min_{x \in \mathbb{R}^n} \|x - b\|_2^2 + \gamma \|Dx\|_1$$

   where $\gamma > 0$, and $D \in \mathbb{R}^{(n-1) \times n}$ is the finite-difference operator $Dx = [x_{i+1} - x_i]_{1 \leq i \leq n-1}$.

   After introducing the new variable $y = Dx$, compute the Lagrangian and the dual problem, and discuss algorithms to solve the dual problem as well as their convergence properties. Compare with the subgradient method applied to the original problem. Extra: implement the algorithms with b a piecewise constant signal corrupted by some Gaussian noise.