Technische Universität München
Department of Mathematics
Chair of Mathematical Optimization, M1
Prof. Dr. Michael Ulbrich
Dennis Klöckner, M.Sc. Winter Semester 2013/14

First Order Primal-Dual Optimization Methods

Exercise Sheet 1

Exercise 1.1 (Dual Norms):
Let $v \in \mathbb{R}^n, 1 < p < \infty$, and q such that $\frac{1}{p} + \frac{1}{q} = 1$. Show the following identities:

a) $\|v\|_1 = \max\limits_{\|y\|_\infty \leq 1} \langle y, v \rangle$,

b) $\|v\|_p = \max\limits_{\|y\|_q \leq 1} \langle y, v \rangle$,

c) $\|v\|_\infty = \max\limits_{\|y\|_1 \leq 1} \langle y, v \rangle$,

where $\langle y, v \rangle = y^T v$ denotes the euclidean scalar product, $\|v\|_r = \left(\sum\limits_{i=1}^n |v_i|^r \right)^{1/r}$ for $1 \leq r < \infty$, and $\|v\|_\infty = \max\limits_{i=1,\dots,n} |v_i|$.

Hint: For b) you can either use Hölder's inequality and then find a suitable y such that the maximum is attained, or alternatively consider the KKT-conditions of the minimization problem

$$\min_{y \in \mathbb{R}^n} -y^T v \quad \text{s.t} \quad \sum_{i=1}^n |y_i|^q \leq 1.$$

Exercise 1.2 (Image Processing Model):
Consider the image processing denoising model with box constraints

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^n \|D_i x\|_2 + \|Bx - u\|_1 \quad \text{s.t.} \quad x \in [0,1]^n, \tag{1}$$

where $x \in \mathbb{R}^n$ is the reconstruced image consisting of n pixels that are suitably arranged in the vector, $B : \mathbb{R}^n \to \mathbb{R}^n$ denotes a transformation, $u \in \mathbb{R}^n$ is the corrupted image data (in the transformed domain), and $D_i : \mathbb{R}^n \to \mathbb{R}^2$ represent discrete difference operators .

a) Show that (1) can be written in the form

$$\min_{x \in C} \max_{y \in K} y^T A x - h^T y \tag{2}$$

for suitable $h \in \mathbb{R}^{3n}, A \in \mathbb{R}^{3n \times n}$, and $C \subset \mathbb{R}^n, K \subset \mathbb{R}^{3n}$.

b) Compute the primal-dual algorithm from the lecture for problem (1), using reformulation (2).

Exercise 1.3 (Proximity Operator):
Let $f : \mathbb{R}^n \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ be a convex function, $x \in \mathbb{R}^n$, and $\gamma \in (0, \infty)$. We consider the problem

$$\inf_{y \in \mathbb{R}^n} f(y) + \frac{1}{2\gamma} \|x - y\|^2. \tag{3}$$

a) Show that the minimum of (3) is uniquely attained, if

(i) $f : \mathbb{R}^n \to \mathbb{R}$ is additionally continuously differentiable or

(ii) f is the convex indicator function $\mathcal{I}_C : \mathbb{R}^n \to \overline{\mathbb{R}}$ of a non-empty, closed and convex set $C \subset \mathbb{R}^n$, given by

$$\mathcal{I}_C(x) = \begin{cases} 0 & \text{, if } x \in C, \\ +\infty & \text{, else.} \end{cases}$$

One can generalize these results to proper, convex and lower semicontinuous functions f. This motivates the definition of the so-called **proximity operator** $\text{prox}_f : \mathbb{R}^n \to \mathbb{R}^n$, defined as

$$\text{prox}_f(x) = \arg\min_{y \in \mathbb{R}^n} f(y) + \frac{1}{2} \|x - y\|^2.$$

b) Let now, according to the lecture, $C \subset \mathbb{R}^n, K \subset \mathbb{R}^m$ be non-empty, closed and convex sets. We define the functions $G : \mathbb{R}^n \to \mathbb{R},\ G(x) = g^T x$ for $g \in \mathbb{R}^n$, $H : \mathbb{R}^m \to \mathbb{R},\ H(y) = h^T y$ for $h \in \mathbb{R}^m$. The convex indicator functions $\mathcal{I}_C$ and $\mathcal{I}_K$ are defined as explained above, and $\sigma, \tau > 0$ are step sizes. Show that for $w \in \mathbb{R}^m$

$$\text{prox}_{\sigma H + \mathcal{I}_K}(w) = P_K(w - \sigma h),$$

and for $v \in \mathbb{R}^n$

$$\text{prox}_{\tau G + \mathcal{I}_C}(v) = P_C(v - \tau g)$$

and infer that this implies

$$\text{prox}_{\sigma H + \mathcal{I}_K}(y + \sigma A\tilde{x}) = P_K(y + \sigma(A\tilde{x} - h)),$$
$$\text{prox}_{\tau G + \mathcal{I}_C}(x - \tau A^T y) = P_C(x - \tau(A^T y + g)).$$

This shows that the first primal-dual algorithm from the lecture is a special case of a more general algorithm, which we will encounter later in this course.