Inverse Problems — Example Sheet 2

Lucas Riedstra

16 November 2020

Question 1. *Let $\mathcal{U}$ be a Banach space and $J \colon \mathcal{U} \to \overline{\mathbb{R}}$ a functional. We define the* subdifferential *of J at any $v \in \mathcal{U}$ as*

$$\partial \mathcal{J}(v) \coloneqq \{p \in \mathcal{U}^* \mid J(u) \geq J(v) + \langle p, u - v\rangle \textit{ for all } u \in \mathcal{U}\}.$$

Characterise the subdifferential for the

(a) absolute value function: $\mathcal{U} = \mathbb{R}$, $J(v) = |v|$,

(b) ℓ^1-norm: $\mathcal{U} = \ell^2$,

$$J(u) = \|u\|_{\ell^1} \coloneqq \begin{cases} \sum_{j=1}^\infty |u_j| & \textit{if } u \in \ell^1; \\ \infty & \textit{else.} \end{cases}$$

(c) characteristic function of the unit ball in $\mathbb{R}$: $\mathcal{U} = \mathbb{R}$, $J(u) = \chi_C(u)$, $C \coloneqq \{u \in \mathbb{R} : |u| \leq 1\}$.

(d) Total Variation $\mathrm{TV} \colon L^1(\Omega) \to \overline{\mathbb{R}}$, where $\Omega \subset \mathbb{R}^n$ is bounded Lipschitz

$$\mathrm{TV}(u) = \sup_{\varphi \in \mathcal{D}} \langle u, \boldsymbol{\nabla} \cdot \varphi\rangle, \quad \mathcal{D} = \{\varphi \in \mathcal{C}_0^\infty(\Omega; \mathbb{R}^n) : \|\varphi(x)\|_2 \leq 1 \ \forall x \in \Omega\}.$$

Solution. Note: the spaces $\mathcal{U}$ in parts (a) to (c) are Hilbert spaces, which means we can identify $\mathcal{U}^*$ with $\mathcal{U}$ (since any functional in $\mathcal{U}^*$ is of the form $\langle u, \cdot\rangle$ for some $u \in \mathcal{U}$).

(a) Let $v \in \mathbb{R}$. We know that $|\cdot|$ is differentiable at $v \neq 0$, so

$$v > 0 \implies \partial J(v) = \{1\} \quad \text{and} \quad v < 0 \implies \partial J(v) = \{-1\}.$$

For $v = 0$ we have

$$\begin{aligned} p \in \partial J(v) &\iff |u| \geq p \cdot u \text{ for all } u \in \mathbb{R} \\ &\iff p \in [-1, 1], \end{aligned}$$

so $\partial J(0) = [-1, 1]$.

(b) Let $v \in \ell^2$. Firstly, if $v \notin \ell^1 = \mathrm{dom}(J)$, then we have $\partial J(v) = \varnothing$. Assume now that $v \in \ell_1 \cap \ell_2$. Then we have, for $p \in \ell^2$, that

$$\begin{aligned} p \in \partial J(v) &\iff \|u\|_{\ell^1} \geq \|v\|_{\ell^1} + \langle p, u - v\rangle && \text{for all } u \in \ell^2 \\ &\iff \|u\|_{\ell^1} - \|v\|_{\ell^1} - \langle p, u - v\rangle \geq 0 && \text{for all } u \in \ell^2 \\ &\iff \sum_{j=1}^\infty |u_i| - |v_i| - p_i(u_i - v_i) \geq 0 && \text{for all } u \in \ell^2 \qquad (1) \\ &\overset{\star}{\iff} |x| - |v_i| - p_i(x - v_i) \geq 0 && \text{for all } i \in \mathbb{N} \text{ and } x \in \mathbb{R}. \qquad (2) \end{aligned}$$

We first prove the bi-implication $\star$. If (2) holds, it is clear that (1) holds. If (2) does not hold, then we can find x, i such that $|x| - |v_i| - p_i(x - v_i) < 0$. By now letting $u = xe_i$ in (1) we find that (1) does not hold.

However, if we define $H(x) \coloneqq |x|$, we see that eq. (2) is equivalent to $p_i \in \partial H(v_i)$ for all i. Therefore, by (a) we have

$$\partial J(v) = \big\{p \in \ell^2 \mid p_i = \operatorname{sign}(v_i) \text{ if } v_i \neq 0 \text{ and } p_i \in [-1, 1] \text{ for all } i\big\}.$$

(c) Clearly, if $|v| < 1$, then χ_C is differentiable with derivative 0 so $\partial J(v) = \{0\}$. If $|v| > 1$, then $v \notin \operatorname{dom}(J)$, and therefore $\partial J(v) = \varnothing$.

Consider the point $v = 1$, then we have

$$p \in \partial J(\chi_C) \iff \chi_C(u) \geq p \cdot (u - 1) \ \forall u.$$

For $u > 1$, this equation is satisfied regardless of p. Therefore, the above equation is equivalent to

$$p \cdot (u - 1) \leq 0 \ \forall u \leq 1,$$

which is satisfied for all $p \geq 0$, so we conclude $\partial J(1) = [0, \infty)$. Analogously, we find $\partial J(-1) = (-\infty, 0]$. We conclude that

$$\partial J(v) = \begin{cases} \varnothing & \text{if } |v| > 1; \\ (-\infty, 0] & \text{if } v = -1; \\ \{0\} & \text{if } v \in (-1, 1); \\ [0, \infty) & \text{if } v = 1. \end{cases}$$

(d) Let $f \in L^1(\Omega) \setminus \mathrm{BV}(\Omega)$, then clearly $\partial\,\mathrm{TV}(f) = \varnothing$. Now suppose $f \in \mathrm{BV}(\Omega)$. It is known that the dual of $L^1(\Omega)$ is $L^\infty(\Omega)$. Therefore, we have for $p \in L^\infty(\Omega)$ that

$$p \in \partial\,\mathrm{TV}(f) \iff \mathrm{TV}(g) \geq \mathrm{TV}(f) + \int_\Omega p(x)(g - f)(x)\,\mathrm{d}x \ \ \forall g \in L^1(\Omega)$$

I do not know how to continue from here.

Question 2. *Let $\mathcal{U}$ be a Banach space and let $E \colon \mathcal{U} \to \overline{\mathbb{R}}$ be proper, lower semi-continuous and convex. Then the* Fenchel conjugate *or* convex conjugate *of E is defined to be the mapping $E^* \colon \mathcal{U}^* \to \mathbb{R}$ with*

$$E^*(v) \coloneqq \sup_{u \in \mathcal{U}} \{\langle v, u\rangle - E(u)\}.$$

(a) Compute the convex conjugates of the following functionals.

(i) $E(u) = \|u\|_{\mathcal{U}}$ for a Banach space $\mathcal{U}$,

(ii) $E(u \mid f) = \sum_{i=1}^n u_i \log\left(\frac{u_i}{f_i}\right)$, where $f \in \mathbb{R}^n_{>0}$ is a positive vector and $u \in \mathbb{R}^n$. What is the effective domain of E? (here we define $\log(x) = -\infty$ for $x < 0$).

(b) Let $E \colon \mathcal{U} \to \overline{\mathbb{R}}$ be a proper, lower semi-continuous and convex functional. Show that

$$p \in \partial E(u) \iff u \in \partial E^*(p)$$

for all $u, p \in \mathcal{U}$.

Hint: *You may exploit the fact that under the stated assumptions $E = E^{**}$ holds true.*

Solution. (a) (i) We have

$$E^*(v) = \sup_{u\in\mathcal{U}} (\langle v, u\rangle - \|u\|).$$

Suppose $\langle v, u^*\rangle - \|u^*\| = \xi > 0$ for some u^*. Then we have for $\alpha > 0$ that

$$\langle v, \alpha u^*\rangle - \|\alpha u^*\| = \alpha(\langle v, u^*\rangle - \|u^*\|) = \alpha\xi,$$

and therefore clearly $E^*(v) = \infty$.
On the other hand, if $\langle v, u\rangle - \|u\| \le 0$ for all u, then the supremum is attained in $u = 0$ with value 0, and therefore $E^*(v) = 0$.
We see that

$$\begin{aligned}
&\langle v, u^*\rangle - \|u^*\| > 0 && \text{for some } u^* \in \mathcal{U};\\
\iff\ &\langle v, u^*\rangle > \|u^*\| && \text{for some } u^* \in \mathcal{U};\\
\iff\ &\|v\|_{\mathcal{U}^*} > 1.
\end{aligned}$$

We conclude that

$$E^*(v) = \chi_{\{\|v\|\le 1\}} = \begin{cases} 0 & \text{if } \|v\| \le 1,\\ \infty & \text{else.}\end{cases}$$

(ii) Suppose first that $p \in \partial E(u)$. Then we have

$$\begin{aligned}
&p \in \partial E(u)\\
\implies\ &E(v) \ge E(u) + \langle p, v-u\rangle && \text{for all } v\\
\implies\ &\langle p, u\rangle - E(u) \ge \langle p, v\rangle - E(v) && \text{for all } v\\
\implies\ &\langle p, u\rangle - E(u) \ge \sup_v (\langle p, v\rangle - E(v)) = E^*(p)\\
\implies\ &\langle p, u\rangle - E(u) + \langle q-p, u\rangle \ge E^*(p) + \langle q-p, u\rangle && \text{for all } q\\
\implies\ &\langle q, u\rangle - E(u) \ge E^*(p) + \langle q-p, u\rangle && \text{for all } q\\
\implies\ &\sup_v (\langle q, v\rangle - E(v)) \ge E^*(p) + \langle q-p, u\rangle && \text{for all } q\\
\implies\ &E^*(q) \ge E^*(p) + \langle q-p, u\rangle && \text{for all } q\\
\implies\ &u \in \partial E^*(p).
\end{aligned}$$

Now, for the reverse implication, note that by what we just proved we have

$$u \in \partial E^*(p) \implies p \in \partial E^{**}(u) \iff p \in \partial E(u),$$

which proves the claim.

Question 3. *Let $u, v \in \mathcal{U}$ and $p \in \partial J(v)$. Recall that the Bregman distance of J at u, v is defined as*

$$D_J^p(u, v) := J(u) - J(v) - \langle p, u-v\rangle.$$

(a) Show that Bregman distances are non-negative.

(b) Show that Bregman distances may not be symmetric, i.e., there exists a J and $u, v \in \mathcal{U}$ with $p \in \partial J(v), q \in \partial J(u)$ such that $D_J^p(u, v) \neq D_J^q(v, u)$.

(c) Show that a vanishing Bregman distance may not imply that the two arguments are the same. What if J is strictly convex?

Proof. (a) Since $p \in \partial J(v)$, we have $J(u) \geq J(v) + \langle p, u-v \rangle$, or equivalently $J(u) - J(v) - \langle p, u-v \rangle \geq 0$.

(b) Let $\mathcal{U} = \mathbb{R}$, $J(x) = |x|$, and choose $u = 0, v = 1$ and $p = 1, q = 0$. Then

$$D_J^p(u,v) = -1 - (-1) = 0 \quad \text{and} \quad D_J^q(v,u) = 1.$$

(c) In the previous part we had an example $J(x) = |x|, u = 0, v = 1, p = 1$, where $D_J^p(u,v) = 0$ while $u \neq v$.

Suppose that J is strictly convex and that $u \neq v$ but $D_J^p(u,v) = 0$, so $J(u) = J(v) + \langle p, u-v \rangle$. Then we have for all $t \in (0,1)$ that

$$\begin{aligned} J(v) + \langle p, (1-t)(u-v) \rangle &= J(v) + \langle p, (tv + (1-t)u) - v \rangle \\ &\overset{\star}{\leq} J(tv + (1-t)u) \\ &< tJ(v) + (1-t)J(u) \\ &= tJ(v) + (1-t)(J(v) + (1-t)\langle p, u-v \rangle) \\ &= J(v) + \langle p, (1-t)(u-v) \rangle, \end{aligned}$$

a contradiction (here $\star$ follows from $p \in \partial J(v)$). We conclude that, if J is strictly convex, the Bregman distance does satisfy $u \neq v \implies D_J^p(u,v) > 0$.

□

Question 4. *Recall that a function $J : \mathcal{U} \to \overline{\mathbb{R}}$ is called absolutely one-homogeneous if $J(\lambda u) = |\lambda| J(u)$ for all $\lambda \in \mathbb{R}, u \in \mathcal{U}$. Let J be convex, proper, l.s.c. and absolutely one-homogeneous.*

(a) Show that $p \in \partial J(v)$ if and only if $p \in \partial J(0)$ and $J(v) = \langle p, v \rangle$. Therefore,

$$D_J^p(u,v) = J(u) - \langle p, u \rangle.$$

Show that

$$\partial J(0) = \bigcup_{u \in \mathcal{U}} \partial J(u).$$

(b) Show that the Bregman distances associated with absolutely one-homogeneous functionals fulfill the triangle inequality in the first argument, i.e., for all $u, v, w \in \mathcal{U}$ and $p \in \partial J(w)$ there is

$$D_J^p(u+v, w) \leq D_J^p(u,w) + D_J^p(v,w).$$

(c) Show that the convex conjugate $J^*(\cdot)$ is the characteristic function of the convex set $\partial J(0)$. Compare this to the results of Exercise 2(a)(i).

Proof. It is clear that $J(0) = 0$.

(a) Suppose $p \in \partial J(v)$. Then we have $J(u) \geq J(v) + \langle p, u-v \rangle$ for all u, which we can rewrite as $J(u) - \langle p, u \rangle \geq J(v) - \langle p, v \rangle$. Plugging in $u = 0$ we obtain $J(v) - \langle p, v \rangle \leq 0$, but plugging in $u = 2v$ we obtain

$$2(J(v) - \langle p, v \rangle) = J(2v) - \langle p, 2v \rangle \geq J(v) - \langle p, v \rangle \implies J(v) - \langle p, v \rangle \geq 0,$$

so we conclude $J(v) - \langle p, v \rangle = 0$ or $J(v) = \langle p, v \rangle$. This also implies that

$$J(u) \geq \langle p, u \rangle \text{ for all } u \implies p \in \partial J(0).$$

Conversely, if $p \in \partial J(0)$ and $J(v) = \langle p, v \rangle$, then for all u we have

$$J(u) \geq \langle p, u \rangle + (J(v) - \langle p, v \rangle) \implies p \in \partial J(v).$$

This concludes the first claim.

From this claim, it follows that $\partial J(u) \subseteq \partial J(0)$ for all $u \in \mathcal{U}$, and therefore trivially $\partial J(0) = \cup_u \partial J(u)$.

(b) Note that we have

$$J(u+v) = 2J\left(\frac{1}{2}u + \frac{1}{2}v\right) \leq 2\left(\frac{1}{2}J(u) + \frac{1}{2}J(v)\right) = J(u) + J(v),$$

and therefore

$$D_J^p(u+v,w) = J(u+v) - \langle p, u+v\rangle \leq J(u) + J(v) - \langle p, u\rangle - \langle p, v\rangle = D_J^p(u,w) + D_J^p(v,w).$$

(c) We can reason analogously to 2(a)(i): we have

$$J^*(v) = \sup_{u\in U}(\langle v, u\rangle - J(u)).$$

Suppose that $v \notin \partial J(0)$, i.e., $\langle v, u^*\rangle - J(u) = \xi > 0$ for some u^*. Then we have for all $\lambda > 0$ that

$$\langle v, \lambda u^*\rangle - J(\lambda u^*) = \lambda\xi,$$

and letting $\lambda \to \infty$ shows $J^*(v) = \infty$.

On the other hand, suppose that $v \in \partial J(0)$, i.e., $\langle v, u\rangle - J(u) \leq 0$ for all u. Then the supremum is attained in $u = 0$ and therefore we have $J^*(v) = 0$.

It follows that $J^*(v) = \partial J(0)$, which is indeed also what we saw in 2(a)(i), since the subdifferential of the norm at 0 is exactly $\{v \in \mathcal{U}^* : \|v\|_{\mathcal{U}^*} \leq 1\}$.

$\square$