hw5 Optimization

Guillaume Jarry

October 2022

1 Problem 1.2

The application is a linear form in finite dimension, thus it is continuous. Therefore the reverse image of the set $[y, +\infty[$, is a closed halfspace, even if it is empty (because this interval is closed, and the application is linear). The subgradient being the intersection of all the closed halspaces $\cap_{yR^n} f^{-1}([y, +\infty[)$, it is also a closed halfspace (even if it is empty), which closes the proof.

2 problem 1.3

Let $g \in \partial f(Ax)$. Because A is a matrix whose image is in R^m and the subgradent inequality holds for all $z \in mathbbR^m$, It also applies to the image of A. It follows:

$$\forall y \in mathbbR^n f(Ay) - f(Ax) \geq g^T(A(y - x))$$

We then use the rules of transposition of the products to reveal:

$$\forall y \in mathbbR^n g(y) - g(x) \geq (A^T g)^T (y - x)$$

And here we recognze that $A^T g$ is a subgradient of g, which proves the inclusion and closes the proof.

Let $h \in \partial f(x) + \partial g(x)$. Then there exists $f \in \partial f(x)$ and $g \in \partial g(x)$ such that:

$$\forall y \in domf, \quad f(y) - f(x) \leq \langle f | y - x \rangle$$

and

$$\forall y \in domg, \quad g(y) - g(x) \leq \langle g | y - x \rangle$$

Therefore by summing the two inequalities, and linearity of the scalar product, we get:

$$\forall y \in domg \cap domf, \quad f(y) + g(y) - f(x) - g(x) \leq \langle g + f | y - x \rangle$$

Thus:

$$f + g = h \in\in \partial(f(x) + g(x))$$

Which shows the inclusion.

3 Problem 1.9

(i) Using the fact that $N_{R^n_+}(x) = \{0\}$ when $x \in intR^n_+ = R^n_{+*}$ and $N_{R^n_+}(x) = \emptyset$ for $x \notin R^n_+$ we get:

$$N_{R^n_+}(x) = \begin{cases} \{0\} \text{ if } x > 0 \\ R^n_- \text{ if } x = 0 \\ \emptyset \text{ if } x \notin R^n_+ \end{cases}$$

(ii) if x is a minimizer, then its gradient is equal to zero and thus the forward implication holds. If $-\nabla f(x) \in N_{R^n_+}$, then according to the definition:

$$\forall y \in R^n_+, -\nabla f(x)^T(y - x) \leq 0$$

And thus we get the chain of inequality (resulting from the convexity of f) :

$$\forall y \in R^n_+ f(y) - f(x) \geq \nabla f(x)^T(y - x) \geq 0$$

Which proves that x is a minimizer ($\forall y \in R^n_+ f(y) \geq f(x)$). Thus we have proven the equivalence relationship.

4 Problem 2.1

Let us remember that throughout this exercice and operator and its graph are used interchangeably. Thus:

$$T^{-1} = \{(y, x)|(x, y) \in GraT\}$$

Thus, if T^{-1} is singled valued, for every $(y, x) \in T^{-1}$ there only exist one tuple $(x, y = Tx) \in T$. (Otherwise we would have multiple values). And thus T itself is a bijective application on $range(T)$ and we have:

$$T^{-1}Tx = x$$

5 Problem 2.3

To prove the monotonicity of ∂L, we will ahve to show that for all $(x_1, u_1), (x_2, u_2)$:

$$\begin{bmatrix} \partial_x L(x_1, u_1) - \partial_x L(x_2, u_2) \\ \partial_u(-L(x_1, u_1)) - \partial_u(-L(x_1, u_1)) \end{bmatrix}^T \begin{bmatrix} x_1 - x_2 \\ u_1 - u_2 \end{bmatrix} \geq 0$$

Let us add the four following subgradient inequality (careful, those are set inequalities and not element inequalities):

$$(1) \quad \langle \partial_x L(x_1, u_1) | x_2 - x_1 \rangle \leq L(x_2, u_1) - L(x_1, u_1)$$

And

$$(2) \quad \langle \partial_u (-L(x_1, u_1)) | u_2 - u_1 \rangle \leq -L(x_1, u_2) + L(x_1, u_1)$$

$$(3) \quad \langle \partial_x (L(x_2, u_2)) | x_1 - x_2 \rangle \leq L(x_1, u_2) - L(x_2, u_2)$$

$$(4) \quad \langle \partial_u (-L(x_2, u_2)) | u_1 - u_2 \rangle \leq -L(x_2, u_1) + L(x_2, u_2)$$

Then, by ading the opposite of line (1) and line (3), (the opposite of (1) is to put the scalar product in the right order) we get:

$$\begin{gathered} -(1) - (2) \langle \partial_x L(x_1, u_1) - \partial_x (L(x_2, u_2)) | x_1 - x_2 \rangle \geq \\ L(x_2, u_2) - L(x_2, u_1) + L(x_1, u_1) - L(x_1, u_2) \end{gathered}$$

And then, using the same reasoning:

$$\begin{gathered} -(2) - (4) \quad \langle \partial_u (-L(x_1, u_1)) - \partial_u (-L(x_2, u_2)) | u_1 - u_2 \rangle \geq \\ -L(x_2, u_2) + L(x_2, u_1) - L(x_1, u_1) + L(x_1, u_2) \end{gathered}$$

And we clearly see that by adding those two lines, we get the expected result:

$$\begin{bmatrix} \partial_x L(x_1, u_1) - \partial_x L(x_2, u_2) \\ \partial_u (-L(x_1, u_1)) - \partial_u (-L(x_1, u_1)) \end{bmatrix}^T \begin{bmatrix} x_1 - x_2 \\ u_1 - u_2 \end{bmatrix} \geq 0$$

Which proves the monotonicity of the subgradient of the Lagrangian.

6 Problem 2.5

Because of the identity $(\partial f)^{-1} = \partial f^*$, and the subgradient of a CCP function exist at every point of the domain, this guarantees the existence of the gradient of f^*. Then, let us suppose there exist a point x for which there exist two distinct $f \in \partial f^*(x)$ and $g \in \partial f^*(x)$. Then because of the equality $(\partial f)^{-1} = \partial f^*$, there exist $k, l \in \partial f(f) \times \partial f(g)$. Then:

$$\forall y \in dom f, f(y) - f(f) \geq \langle k | y - f \rangle \text{ and } f(y) - f(g) \geq \langle l | y - g \rangle$$

There, by choosing $y = g$ on the left and $y = f$ on the right, ge get:

$$f(g) - f(f) \geq \langle k | g - f \rangle \text{ and } f(f) - f(g) \geq \langle l | f - g \rangle$$

Then summing the two inequalities contradicts the monotonicity of the subgradient operator which is true because f is CCP. Therefore we have proven the fact that ∂f^* is singled-valued (i).

Then, proving (ii) comes naturally. Because ∂f^* is singled value, then for all u, f^* is going to be differentiable because its subgradient is a singleton. Thus, f^* is differentiable on $intdom f$ and we have proven the property.