hw12 Optimization

Guillaume Jarry

December 2022

1 Problem 2.34

a) First let us remember that:

$$(\delta_V)^*(c) = \underset{x \in V}{sup}\{\langle x|c\rangle - \delta_V(x)\}$$

ie, because $\delta_V(x) = 0$ for $x \in V$

$$(\delta_V)^*(c) = \underset{x \in V}{sup}\{\langle x|c\rangle\}$$

Then, because V is a subspace, if $c \notin V^\perp$, there exist $u \in V$ such that: $\langle u|c\rangle \geq 0$. And thus, by homogeneity of the subspace and the scalar product, $\langle \lambda u|c\rangle \to \infty$ as $\lambda \to \infty$. When $c \in V^\perp$, then for all $x \in V$, by definition $\langle v|x\rangle = 0$. To summarize, we have:

$$(\delta_V)^*(c) = \begin{cases} 0 \text{ if } c \in V^\perp \\ \infty \text{ else, ie } c \notin V^\perp \end{cases}$$

And this is clearly $\delta_{V^\perp}$. Thus:

$$(\delta_V)^*(c) = \delta_{V^\perp}$$

b) Then, let $v \in V$. Because V is a vector space, for all $z \in V$, we have the belonging $v - x \in V$.

Once that is noticed, clearly $V^\perp \subset \mathbb{N}_V(v)$ because for all $y \in V^\perp, u \in V, \langle u|y\rangle = 0$. But this inclusion is also an equality because if we have $y \in \mathbb{N}_V(v)$, such then there exist $z \in V$ so that $\langle y|z - v < 0$, then because V is a vector space, we also have the inequality $\langle y|v - z > 0$, which is a contradiction. Therefore we need to admit:

$$\forall y \in \mathbb{N}_V(v), \forall z \in V, \langle y|z - v = 0$$

And that is clearly $V^\perp$. This way we have shown: $\mathbb{N}_V(v) = V^*$

c) Using Moreau's identity for $\alpha = 1$, applied to δ_V, we get:

$$Prox_{\delta_V} + Prox_{(\delta_V)^*} = \mathbb{I}$$

Then, using what we have shown in question a), we get:

$$Prox_{\delta_V} + Prox_{\delta_{V^\perp}} = \mathbb{I}$$

$$\Pi_V + \Pi_{V^\perp} = \mathbb{I}$$

Which is what we had to show.

2 Problem 2.36

Let us show the hint (which is also exercise 11.8). Let $y \in C^\perp$, and let us denotes its column by $y_1, \dots, y_n$. Let $x \in C$ We have:

$$\langle x|y\rangle = \sum_{i=1}^{n} \langle x_i|y_i\rangle = 0$$

Then because x is part of the consensus set, we can factorize the expression and we get:

$$\langle x|y\rangle = \langle x|y_1 + \cdots + y_n\rangle = 0$$

Because the above equality stands for all $x \in C$, which is isomoprhic to $\mathbb{R}^{\ltimes}$, and that $y_1 + \cdots + y_n \in mathbbR^n$, then $y_1 + \cdots + y_n = 0$, because only 0 is orthogonal to the entire vector space of which it is the nul element. Therefore we have shown the first inclusion $C^\perp \subset \{v \in \mathbb{R}^{\ltimes |}, v_1 + \cdots + v_n = 0\}$. The equality simply stems from the very same factorization we have exposed above. Let $y \in$, $x \in C$:

$$\langle x|y\rangle = \sum_{i=1}^{n} \langle x_i|y_i\rangle = \langle x|y_1 + \cdots + y_n\rangle = \langle x|0\rangle = 0$$

Therefore we have shown the equality : $C^\perp = \{v \in \mathbb{R}^{\ltimes |}, v_1 + \cdots + v_n = 0\}$.

Once that is established, because of the problem 2.34, we can identify that the two problems are equivalent, because the operator $\mathbb{N}_C(x_1, \dots, x_n)$ enforces the constraints that $x_1, \dots, x_n$ have to be in the consensus set. Indeed, if $x_1, \dots, x_n \in C, \mathbb{N}_C(x_1, \dots, x_n) = C^\perp$, which includes 0. If $x_1, \dots, x_n \notin C$, then $\mathbb{N}_C(x_1, \dots, x_n) =$ and therefore does not include 0, so the solution has to be in the consensus set. Then, let x^* be a zero of $\sum A_i x$. Either x^* is a zero common to all the A_i, and is thus a solution of the second problem because the zeros of $\begin{bmatrix} A_1(x_1) \\ \vdots \\ A_n(x) \end{bmatrix}$, are the zeros common to all the A_i, and that $0 \in C^\perp$.

If x^* is not a zero not common to all the A_i, then there exist an element $v = v_1, \ldots, v_n \in A_1 x \times A_n x$ such that $\sum_{i=1}^n v_i = 0$, ie $v \in C^\perp$. But then, its opposite is also in $^\perp$ and thus 0 is contained in the set $\begin{bmatrix} A_1(x_1) \\ \vdots \\ A_n(x) \end{bmatrix} + \mathbb{N}_C(x_1, \ldots, x_n)$. And this way we have shown that the two problems are indeed equivalent.

PS: We must clarify that a zero here is a element x such that $\sum A_i x$ includes 0 !

3 Problem 11.1

Using the hint, we first of all notice that, by definition of the infimum:

$$\forall z \in Y,\ \forall x \in X,\ f(x,z) - h(z) = f(x,z) - \underset{x \in X}{inf} f(x,z) \geq 0$$

And, furthermore, by definition of $x^*(y)$, ie $f(x^*(y), y) = \underset{x \in X}{inf} f(x,z)$ we have:

$$\forall y \in Y, f(x^*(y), y) - h(y) = 0$$

We conclude from the two above definition that y is a minimizer of the function (for the given $x^*(y) \in X$), $z \to f(x^*, z) - h(z)$. By differentiabilty of the two function in Y for $x^*(y)$ given, the above sum function is differentiable in y and because it is a minimizer, the gradient of the sum of function will be zero, ie:

$$\nabla_y f(x^*(y), y) - \nabla_y h(y) = 0$$

And therefore:

$$\nabla_y f(x^*(y), y) = \nabla_y h(y)$$

4 Problem 11.2

First let us show that h is convex. Let $y, z \in Y$ and $t \in [0, 1]$. For the sake of conveniance, we will write: $w = tz + (1-t)y$, and will reuse the notation of the above exercise $y \in Y, x^{ast}(y) = argmin_{x \in X} f(x, y)$.

$$h(w) = inf_{x \in X} f(x, w)$$

By definition of $h(w)$:

$$h(w) \leq f(tx^*(z) + (1-t)x^*(y), w)$$

And then, by joint convexity of f, f is convex on the segment $[(x^*(z), z), (x^*(y), y)]$, and therefore:

$$h(w) \leq f(tx^*(z) + (1-t)x^*(y), tz + (1-t)y \leq tf(x^*(z), z) + (1-t)f(x^*(y), y)$$

This way we have shown that h is convex.

b) Then, we reuse the trick of problem 11.1:

$$\forall (x, y) \in X \times Y,\ f(x, y) - h(y) \geq 0$$

And $x^*(y)$ is the minimizer of the function $y \to f(x^*, y) - h(y)$. Then, because f is convex, this implies that the subgradient of the difference function in y is zero.

$$\partial_y f(x^*, y) - \partial_y h(y) = 0$$

5 Problem 11.4

Let us write $\phi^*(z)$:

$$\phi^*(z) = \underset{y \in Y}{sup}\{z^T y - \underset{x \in X}{inf} f(x, y)\}$$

But because of the minus sign before the function f, maximizing the quantity in the curlybrackets will also be findng the x that minimizes the function f (as it is retrieved from the quantity. And therefore the x that minimizes the function maximizes the expression in the brackets. And because x does not intervene elsewhere, we have:

$$\phi^*(z) = \underset{y \in Y, x \in X}{sup}\{z^T y - f(x, y)\} = \psi(x)(v)$$

Which is exactly what we had to show.

6 Problem 11.8

This was already proven in exercice 2.36, and therefore we will just refer to the proof of exercice 2.36 to justify that this exercise was done.

7 Problem 10.1

a) Let $(x, u) \in A^{-1}, (y, v) \in A^{-1}$, Because A is maximal monotone, we get :

$$\langle x - y | u - v \rangle \geq 0$$

Then, obviously, we can reverse the order of the term the scalar product to make the characterization of the monotony of A^{-1} appear. Then, suppose their

exist a pair (z, w) such that $A^{-1} \cup \{(z, w)\}$ is monotone. If that was the case, then because of the commutativity of the scalar product, the reverse pair (w, z) would make $A \cup \{(w, z)\}$ monotone. Or A is already maximal monotone, so such a pair does not exist. Therefore A^{-1} is maximal.
We have show the maximal monotonicity of the inverse of A.

b) Let (x, u) and (y, v) belong to $M^T AM$, and let $s = AMx, t = AMy$. Then:

$$\langle (u - v) | x - y \rangle = \langle M^T (s - t) | (x - y) \rangle$$

And by monotonicity of A, we have:

$$\langle (s - t) | (x - y) \rangle 0$$

c) For the reverse implication, if (y, v) is in A, then by monotonicity of A, we have the result $\langle x - y | u - v \rangle \geq 0$ for all pair (x, u) in A. The direct implication stems from the maximality of A. Suppose we have a pair $(y, v) \notin A$ such that $\forall (x, u) \in A, \langle x - y | u - v \rangle \geq 0$. Then $A \cup \{(y, v)\}$ would be monotone. But by maximality of A, that is impossible. Therefore $(y, v) \in A$.

d)