12

More exercises

I believed myself to be a mathematician.
In these days I discovered that I am not even an amateur.
(R. Queneau, "Odile")

In this section we collect some more exercises, related to the whole content of the book.

Exercise 12.1 (About polar cones.) Let X be a reflexive Banach space, let $C \subset X$ be a closed convex cone. Then $C^{oo} = C$.

Hint. It is obvious that $C \subset C^{oo}$. Suppose now there is $x \in C^{oo} \setminus C$. Then there are $0^* \neq y^*$ and $a \in \mathbb{R}$ such that

$$\langle y^*, x \rangle > a \geq \langle y^*, c \rangle, \tag{12.1}$$

for all $c \in C$. Show that we can assume $a = 0$ in (12.1). It follows that $y^* \in C^o$ and thus, since $x \in C^{oo}$, we have that $\langle y^*, x \rangle \leq 0$.

Exercise 12.2 Let

$$f(x) = \begin{cases} -\sqrt{x} & \text{if } x \geq 0, \\ \infty & \text{elsewhere.} \end{cases}$$

Evaluate $f_k = f \nabla k \| \cdot \|$ for all k. Let $g(x) = f(-x)$. Find $\inf(f+g)$, $\inf(f_k + g_k)$ and their minimizers. Compare with the result of next exercise.

Exercise 12.3 With the notation of the previous exercise, suppose $f, g \in \Gamma(\mathbb{R}^n)$ and

$$\operatorname{ri}\operatorname{dom} f \cap \operatorname{ri}\operatorname{dom} g \neq \emptyset.$$

Then, for all large k, we have

$$\inf(f+g) = \inf(f_k + g_k)$$

and

$$\operatorname{Min}(f+g) = \operatorname{Min}(f_k + g_k).$$

Hint. Prove that $\inf(f+g) \le \inf(f_k+g_k)$. There is $y \in \mathbb{R}^n$ such that

$$-\inf(f+g) = f^*(y) + g^*(-y).$$

Take $k > \|y\|$. Then

$$\begin{aligned}-\inf(f+g) &= f^*(y) + g^*(-y) = (f^* + I_{kB})(y) + (g^* + I_{kB})(-y)\\ &= (f_k)^*(y) + (g_k)^*(-y) \ge \inf_{z\in\mathbb{R}^n}((f_k)^*(z) + (g_k)^*(-z))\\ &= -\inf(f_k+g_k) \ge -\inf(f+g).\end{aligned}$$

Observe that the above calculation also shows that y as above is optimal for the problem of minimizing $(f_k)^*(\cdot) + (g_k)^*(-\,\cdot)$ on $\mathbb{R}^n$.

Now, using $k > \|y\|$,

$$\begin{aligned}x \in \operatorname{Min}(f+g) &\Leftrightarrow f(x) + g(x) = -f^*(y) - g^*(-y)\\ &\Leftrightarrow x \in \partial f^*(y) \cap \partial g^*(-y)\\ &\Leftrightarrow x \in \partial(f^* + I_{kB})(y) \cap \partial(g^* + I_{kB})(-y)\\ &\Leftrightarrow x \in \partial(f_k)^*(y) \cap \partial(g_k)^*(-y)\\ &\Leftrightarrow x \in \operatorname{Min}(f_k+g_k).\end{aligned}$$

Exercise 12.4 Let $\{x_n^i\}$, $i = 1,\dots,k$ be k sequences in a Euclidean space, and suppose $x_n^i \to x^i$ for all i. Prove that $\operatorname{co}\bigcup x_n^i$ converges in the Hausdorff sense to $\operatorname{co}\bigcup x^i$.

Exercise 12.5 Let X be a Banach space and suppose $f, g \in \Gamma(X)$, $f \ge -g$, $f(0) = -g(0)$. Then

$$\{y^* : f^*(y^*) + g^*(-y^*) \le 0\} = \partial f(0) \cap -\partial g(0).$$

Exercise 12.6 Let X be a Banach space, let $f \in \Gamma(X)$ be Fréchet differentiable, and let $\sigma > 0$. Set

$$S_\sigma := \{x \in X : f(x) \le f(y) + \sigma\|y - x\|, \forall y \in X\},$$

and

$$T_\sigma := \{x \in X : \|\nabla f(x)\|_* \le \sigma\}.$$

Prove that $S_\sigma = T_\sigma$ are closed sets. Which relation holds between the two sets if f is not assumed to be convex?

Exercise 12.7 In the setting of Exercise 12.6, prove that f is Tykhonov well-posed if and only if $S_\sigma \ne \emptyset$ for all $\sigma > 0$ and $\operatorname{diam} S_\sigma \to 0$ as $\sigma \to 0$. Deduce an equivalence when f is also Fréchet differentiable. Is convexity needed in both implications? Give an example when the equivalence fails if f is not convex.

Hint. Suppose f is Tykhonov well-posed. Clearly, $S_\sigma \neq \emptyset$ for all σ. Without loss of generality, suppose $f(0) = 0 = \inf f$. Suppose $\operatorname{diam} S_\sigma \geq 2a$, for some $a > 0$, and let $0 < m = \inf_{\|x\|=a} f(x)$. There is $x_n \in S_{\frac{1}{n}}$ such that $\|x_n\| \geq a$. Show that this leads to a contradiction. Conversely, show that $\bigcap_{\sigma>0}$ is a singleton and the set of the minimizers of f. From the Ekeland variational principle deduce that

$$f^{\inf f + a^2} \subset B_a(S_a)$$

and use the Furi Vignoli characterization of Tykhonov well-posedness. As an example, consider $f(x) = \arctan x^2$.

Variational convergences are expressed in terms of set convergences of epigraphs. On the other hand, not only is the behavior of the epigraphs important. How the level sets move under convergence of epigraphs is an important issue. Thus, the next exercises provide gap and excess calculus with level sets and epigraphs. In the space $X \times \mathbb{R}$ we shall consider the box norm.

Exercise 12.8 Let X be a metric space, let $f\colon X \to (-\infty, \infty]$ be lower semicontinuous, let $C \in c(X)$ Prove that

(i) $D(C, f^a) = d$ implies $D(C \times \{a - d\}, \operatorname{epi} f) = d$.
(ii) $\forall b \in \mathbb{R}$ and $\forall a \geq b$ such that $f^a \neq \emptyset$,

$$D(C \times \{b\}, \operatorname{epi} f) \geq \min\{D(C, f^a), a - b\}.$$

(iii) $\forall b \in \mathbb{R}$ and $\forall a \geq b$ such that $f^a \neq \emptyset$, $D(C \times \{b\}, \operatorname{epi} f) = d$ implies $b + d \geq \inf f$.
(iv) $D(C \times \{b\}, \operatorname{epi} f) = d$ implies $D(C, f^{b+d+\varepsilon}) \leq d$, for all $\varepsilon > 0$.
(v) $D(C \times \{b\}, \operatorname{epi} f) = d$ implies $D(C, f^{b+d-\varepsilon}) \geq d$, for all $\varepsilon > 0$.

Exercise 12.9 Let X be a metric space, let $f\colon X \to (-\infty, \infty]$ be lower semicontinuous, let $C \in c(X)$ Prove that

(i) $e(C, f^a) = d$ implies $e(C \times \{a - d\}, \operatorname{epi} f) = d$.
(ii) $\forall b \in \mathbb{R}$ and $\forall a \geq b$ such that $f^a \neq \emptyset$,

$$e(C \times \{b\}, \operatorname{epi} f) \leq \max\{e(C, f^a), a - b\}.$$

(iii) $\forall b \in \mathbb{R}$ and $\forall a \geq b$ such that $f^a \neq \emptyset$, $e(C \times \{b\}, \operatorname{epi} f) = d$ implies $b + d \geq \inf f$.
(iv) $e(C \times \{b\}, \operatorname{epi} f) = d$ implies $e(C, f^{b+d+\varepsilon}) \leq d$, for all $\varepsilon > 0$.
(v) $e(C \times \{b\}, \operatorname{epi} f) = d$ implies $e(C, f^{b+d-\varepsilon}) \geq d$, for all $\varepsilon > 0$.

Exercise 12.10 Let X be an E-space and $f \in \Gamma(X)$. Then, setting $f_n(x) = f(x) + \frac{1}{n}\|x\|^2$, prove that $f_n \to f$ for the Attouch–Wets convergence and that $f_n(\cdot) - \langle p, \cdot \rangle$ is Tykhonov well-posed for all n and for all $p \in X^*$.

Exercise 12.11 Let X be a reflexive Banach space, and $f \in \Gamma(X)$. Find a sequence $\{f_n\}$ such that $f_n \in \Gamma(X)$ are Tykhonov well-posed, everywhere Fréchet differentiable, and $f_n \to f$ for the Attouch–Wets convergence.

Hint. Take an equivalent norm $\|\cdot\|$ in X such that both X and X^* are now E-spaces. From Exercise 12.10 we know that $f^* + \frac{1}{n}\|\cdot\|_*^2 - \langle p, \cdot\rangle$ is Tykhonov well-posed for all n. Thus $(f^*+\frac{1}{n}\|x\|_*^2)^*$ is everywhere Fréchet differentiable for all n. It follows that $g_n(x) = (f^*+\frac{1}{n}\|\cdot\|_*^2)^*(x)+\frac{1}{n}\|x\|^2$ is Fréchet differentiable and Tykhonov well-posed for all n. Prove that $g_n \to f$ for the Attouch–Wets convergence.

Exercise 12.12 Consider the following game. Rosa and Alex must say, at the same time, a number between 1 and 4 (inclusive). The one saying the highest number gets from the other what was said. There is one exception for otherwise the game is silly. If Alex says n and Rosa $n-1$, then Rosa wins n, and conversely. Write down the matrix associated with the game, and find its value and its saddle points.

Hint. Observe that it is a fair game, and use Exercise 7.2.6.

We make one comment on the previous exercise. The proposed game (or maybe an equivalent variant of it) was invented by a rather famous person, with the intention of creating a computer program able to *learn* from the behavior of an opponent, in order to be able to understand its psychology and to beat it after several repetitions of the game. Unfortunately, he had a student with some knowledge of game theory, proposing to him the use of the optimal strategy, whose existence is guaranteed by the theorem of von Neumann. Thus, when telling the computer to play this strategy over and over, no clever idea could do better than a tie (on average) with resulting great disappointment for the famous person. I like this story, since it shows well how challenging game theory can be from the point of view of psychology.

Exercise 12.13 Consider the following game. Emanuele and Alberto must show each other one or two fingers and say a number, at the same time. If both are right or wrong, they get zero. If one is wrong and the other one is right, the one who is right gets the number he said. Determine what they should play, knowing that both are very smart. Do the same if the winner always gets 1, instead of the number he said.

Hint. The following matrix should tell you something.

$$\begin{pmatrix} 0 & 2 & -3 & 0 \\ -2 & 0 & 0 & 3 \\ 3 & 0 & 0 & -4 \\ 0 & -3 & 4 & 0 \end{pmatrix}.$$

Ask yourself if the result of Exercise 7.2.6 can be used. My answer (but you should check) is that they always say "three" and play 1 with probability x, 2 with probability $1-x$, where $\frac{4}{7} \le x \le \frac{3}{5}$.

Exercise 12.14 Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be continuous convex, and suppose $\lim_{|x|\to\infty} f(x, mx) = \infty$ for all $m \in \mathbb{R}$. Prove that f is Tykhonov well-posed in the generalized sense. Does the same hold in infinite dimensions?

Hint. Consider a separable Hilbert space with basis $\{e_n : n \in \mathbb{N}\}$, and the function

$$f(x) = \sum_{n=1}^{\infty} \frac{\langle x, e_n \rangle^2}{n^2} - \langle x^*, x \rangle,$$

where $x^* = \sum \frac{1}{n} e_n$. Then show f is not even lower bounded.

Exercise 12.15 This is a cute example taken from T. Rockafellar's book *Convex Analysis*, i.e., the example of a function $f : \mathbb{R}^2 \to \mathbb{R}$ continuous convex, assuming a minimum on each line, and not assuming a minimum on $\mathbb{R}^2$. Let C be the epigraph of the function $g(x) = x^2$ and consider the function $f(x, y) = d^2[(x, y), C] - x$. Prove that f fulfills the above property, and prove also that f is $C^1(\mathbb{R}^2)$.

Exercise 12.16 Let $f \in \Gamma(\mathbb{R}^n)$. The following are equivalent:

- f is lower bounded and $\operatorname{Min} f = \emptyset$;
- $0 \in \operatorname{dom} f^*$ and there is y such that $(f^*)'(0, ; y) = -\infty$.

Hint. Remember that $f^*(0) = -\inf f$ and that $\operatorname{Min} f = \partial f^*(0)$. Prove that $\partial f(x) = \emptyset$ if and only if there exists a direction y such that $f'(x; y) = -\infty$ (remember that $f'(x; \cdot)$ is sublinear).

Exercise 12.17 Prove that $\operatorname{cl} \operatorname{cone} \operatorname{dom} f = \big(0^+((f^*)^a)\big)^\circ$, for $a > -f(0)$.

Hint. Observe that $(f^*)^a \neq \emptyset$. $(f^*)^a = \{x^* : \langle x^*, x \rangle - f(x) \leq a, \forall x \in \operatorname{dom} f\}$. Thus $z^* \in (0^+((f^*)^a))^\circ$ if and only if $\langle z^*, x \rangle \leq 0$ for all $x \in \operatorname{dom} f$, if and only if $\langle z^*, y \rangle \leq 0$ for all $y \in \operatorname{cl} \operatorname{cone} \operatorname{dom} f$.

Exercise 12.18 This is much more than an exercise. Here I want to introduce the idea of "minimizing" a function which is not real valued, but rather takes values in a Euclidean space. This subject is known under the name of *vector optimization* (also Pareto optimization, multicriteria optimization) and it is a very important aspect of the general field of optimization. Minimizing a function often has the meaning of having to minimize some cost. However, it can happen that one must take into account several cost functions at the same time, not just one. Thus it is important to give a meaning to the idea of minimizing a function $f = (f_1, \dots, f_n)$, where each f_i is a scalar function. And this can be generalized by assuming that f takes values on a general space, ordered in some way (to give a meaning to the idea of minimizing). Here I want to talk a little about this. I will consider very special cases, in order to avoid any technicalities. What I will say can be deeply generalized. The interested reader could consult the book by Luc [Luc] to get a more complete idea of the subject.

So, let $P \subset \mathbb{R}^l$ be a pointed (i.e., $P \cap -P = \{0\}$) closed and convex cone with nonempty interior. The cone P induces on $\mathbb{R}^l$ the order relation $\leq_P$ defined as follows: for every $y_1, y_2 \in \mathbb{R}^l$,

$$y_1 \leq_P y_2 \overset{\text{def}}{\Longleftrightarrow} y_2 \in y_1 + P.$$

Here are some examples of cones: in $\mathbb{R}^n$, $P = \{x = (x_1, \dots, x_n) : x_i \geq 0, \forall i\}$; in $\mathbb{R}^2$, $P = \{x = (x, y)$: either $x \geq 0$ or $x = 0$ and $y \geq 0\}$: this cone, which is *not* closed, induces the so called *lexicographic order*. In l^2, let $P = \{x = (x_1, \dots, x_n, \dots) : x_i \geq 0,\ \forall i\}$: this cone has empty interior, in l^∞ let $P = \{x = (x_1, \dots, x_n, \dots) : x_i \geq 0,\ \forall i\}$: this cone has nonempty interior.

Given C, a nonempty subset of $\mathbb{R}^l$, we denote by $\operatorname{Min} C$ the set

$$\operatorname{Min} C \overset{\text{def}}{=} \{y \in C : C \cap (y - P) = \{y\}\}.$$

The elements of the set $\operatorname{Min} C$ are called the *minimal points* of C (with respect to the order induced by the cone P).

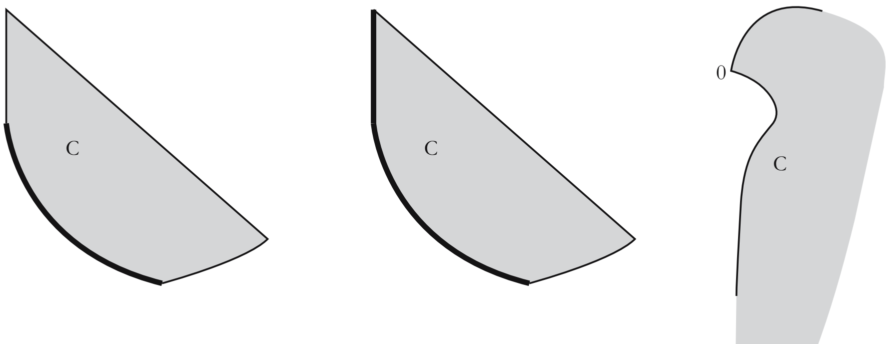
This is not the only notion of minimality one can think of. For instance, the above notion of minimality can be strengthened by introducing the notion of *proper minimality*. A point $y \in C$ is a *properly minimal point* of C if there exists a convex cone P_0 such that $P \setminus \{0\} \subset \operatorname{int} P_0$ and y is a minimal point of C with respect to the order given by the cone P_0. We denote the set of the properly minimal points of C by $\operatorname{Pr} \operatorname{Min} C$.

The concept of minimal point can also be weakened. Define the set

$$\operatorname{Wmin} C \overset{\text{def}}{=} \{y \in C : C \cap (y - \operatorname{int} P) = \emptyset\}$$

of the weakly minimal points of the set C. Clearly

$$\operatorname{Pr} \operatorname{Min} C \subset \operatorname{Min} C \subset W \operatorname{Min} C.$$



Efficient points of C.

Weakly efficient points of C.

0 is not properly efficient for C.

Figure 12.1.

Let us now consider a function $f\colon \mathbb{R}^k \to \mathbb{R}^l$. Let A be a subset of $\mathbb{R}^k$. The set of the efficient points of A is

$$\operatorname{Eff}(A, f) \stackrel{\text{def}}{=} \{x \in A : f(x) \in \operatorname{Min} f(A)\}.$$

In the same way we can introduce the sets $\operatorname{WEff}(A, f)$ and $\operatorname{PrEff}(A, f)$.

And it is clearly possible and interesting to define a notion of convexity for vector valued functions. Here it is.

Let $A \subset \mathbb{R}^k$ be a convex set, and $f\colon A \subset \mathbb{R}^k \to \mathbb{R}^l$. Then f is said to be a *P-convex* (or simply convex, when it is clear which is the cone P inducing the order relation) function on A if for every $x_1, x_2 \in A$ and for every $\lambda \in [0, 1]$,

$$f(\lambda x_1 + (1-\lambda)x_2) \in \lambda f(x_1) + (1-\lambda) f(x_2) - P,$$

and it is said to be a *strictly P-convex* function if for every $x_1, x_2 \in A$, $x_1 \neq x_2$ and for every $\lambda \in (0, 1)$,

$$f(\lambda x_1 + (1-\lambda)x_2 \in \lambda f(x_1) + (1-\lambda) f(x_2) \setminus \operatorname{int} P.$$

Now I only suggest some results, focusing essentially on some aspects of convexity, and stability. I leave the proofs as exercises, and sometimes outline the main ideas of the proofs. The first is an existence result, which is stated in a very particular case.

Proposition 12.19 *Under the setting previously described, let $A \subset \mathbb{R}^k$ be nonempty, closed and such that there exists $x \in \mathbb{R}^k$ such that $A \subset x + P$. Then $\operatorname{Min} A$ is nonempty.*

Proof. (Outline) Without loss of generality, suppose $x = 0$. Prove that there exists $x^* \in \mathbb{R}^k$ such that $\langle x^*, p\rangle > 0$ for all $p \in P$, $p \neq 0$ (the origin can be separated from $\operatorname{co}(A \cap \partial B)$, since the cone P is pointed). Prove that $\lim_{c \in C, \|c\| \to \infty} \langle x^*, c\rangle = \infty$ (arguing by contradiction). Then $g(a) = \langle x^*, a\rangle$ assumes minimum on A. Prove that if $\bar{x}$ minimizes g on A then $\bar{x} \in \operatorname{Min} A$.

With a little more effort one could prove that under the previous assumptions $\operatorname{Pr}\operatorname{Min} A$ is actually nonempty. □

We now see some properties of the convex functions.

Proposition 12.20 *Let $A \subset \mathbb{R}^k$ be a convex set and let $f\colon \mathbb{R}^k \to \mathbb{R}^l$ be a P-convex function. Then*

(i) *$f(A) + P$ is a convex subset of $\mathbb{R}^l$.*
(ii) *f is continuous.*
(iii) *If f is strictly $P-$ convex then* $\operatorname{WEff}(A, f) = \operatorname{Eff}(A, f)$.
(iv) *Defining in the obvious way the level sets of f, prove that, for all $a, b \in \mathbb{R}^k$ such that $f^a \neq \emptyset$, $f^b \neq \emptyset$, it holds $0^+(f^a) = 0^+(f^b)$.*
(v) *Calling H the common recession cone of the level sets of f, show that, if $0^+(A) \cap H = \{0\}$, then $f(A) + P$ is closed.*

We turn now our attention to convergence issues. Prove the following.

Proposition 12.21 *Let C_n be closed convex subsets of $\mathbb{R}^l$. Suppose $C_n \xrightarrow{K} C$. Then*

(i) $\operatorname{Li}\operatorname{Min} C_n \supset \operatorname{Min} C$;
(ii) $\operatorname{Li} Pr \operatorname{Min} C_n \supset Pr \operatorname{Min} C$;
(iii) $\operatorname{Ls}\operatorname{Wmin} C_n \subset \operatorname{Min} C$.

Proof. (Outline) For (i), it is enough to prove that for every $c \in C$ and for every $\varepsilon > 0$ there exists $y_n \in \operatorname{Min} C_n$ such that $d(y_n, c) < \varepsilon$. There exists a sequence $\{c_n\}$ such that $c_n \in C_n$ for all n and $c_n \to c$. Show that $D_n := (c_n - P) \cap C_n \subset B(c; \varepsilon)$ eventually. Since $\operatorname{Min} D_n$ is nonempty and $\operatorname{Min} D_n \subset \operatorname{Min} C_n$, the conclusion of (i) follows. The proof of (ii) relies on the fact that the proper minimal points are, under our assumptions, a dense subset of the minimal points. The proof of (iii) is straightforward. □

Thus the minimal and properly minimal sets enjoy a property of lower convergence, while the weakly minimal sets enjoy a property of upper convergence. Easy examples show that opposite relations do not hold in general. However it should be noticed that, if $\operatorname{Min} A = W \operatorname{Min} A$, then actually from (i) and (iii) above we can trivially conclude that $\operatorname{Min} C_n$ converges to $\operatorname{Min} C$ in Kuratowski sense.

Theorem 12.22 *Let $A_n \subset \mathbb{R}^k$ be closed convex sets, let f_n and f be P-convex functions. Suppose*

(i) $0^+(A) \cap H_f = \{0\}$;
(ii) $A_n \xrightarrow{K} A$;
(iii) *$f_n \to f$ with respect to the continuous convergence (i.e., $x_n \to x$ implies $f_n(x_n) \to f(x)$).*

Then

$$f_n(A_n) + P \xrightarrow{K} f(A) + P.$$

Theorem 12.23 *Under the same assumptions as the previous theorem we have*

(i) $\operatorname{Min} f(A) \subset \operatorname{Li}\operatorname{Min} f_n(A_n)$.
(ii) *If moreover f is strictly convex,*

$$\operatorname{Min} f_n(A_n) \xrightarrow{K} \operatorname{Min} f(A) \quad \text{and} \quad \operatorname{Eff}(A_n, f_n) \xrightarrow{K} \operatorname{Eff}(A, f).$$

If anyone is really interested in having the proofs of the previous exercises, he can send me an e-mail and I will send back the paper.