Optimization
Reminder and exercises

Barbara Pascal, **Nelly Pustelnik**

CNRS, Laboratoire de Physique de l'ENS de Lyon, Univ. Lyon 1
nelly.pustelnik@ens-lyon.fr

barbara.pascal@ens-lyon.fr

Reminder of the context

$$\widehat{x} \in \operatorname*{Argmin}_{x\in\mathcal{H}} f(x), \quad \mathcal{H} \quad \text{a Hilbert space}$$

with possibility to consider $\widetilde{f}(x) = \begin{cases} f(x) & \text{if } x \in D \\ \infty & \text{otherwise} \end{cases}$

D *domain* of the function $\operatorname{dom} \widetilde{f} \equiv \{x \in \mathcal{H} \mid \widetilde{f}(x) < \infty\}$

Reminder of the context

$$\widehat{x} \in \operatorname*{Argmin}_{x \in \mathcal{H}} f(x), \quad \mathcal{H} \quad \text{a Hilbert space}$$

with possibility to consider $\widetilde{f}(x) = \begin{cases} f(x) & \text{if } x \in D \\ \infty & \text{otherwise} \end{cases}$

D *domain* of the function $\operatorname{dom} \widetilde{f} \equiv \{x \in \mathcal{H} \mid \widetilde{f}(x) < \infty\}$

$f \in \Gamma_0(\mathcal{H})$

- proper: $\operatorname{dom} f \neq \emptyset$,

Reminder of the context

$$\widehat{x} \in \operatorname*{Argmin}_{x \in \mathcal{H}} f(x), \quad \mathcal{H} \quad \text{a Hilbert space}$$

with possibility to consider $\widetilde{f}(x) = \begin{cases} f(x) & \text{if } x \in D \\ \infty & \text{otherwise} \end{cases}$

D *domain* of the function $\operatorname{dom} \widetilde{f} \equiv \{x \in \mathcal{H} \mid \widetilde{f}(x) < \infty\}$

$f \in \Gamma_0(\mathcal{H})$

- proper: $\operatorname{dom} f \neq \emptyset$,
- lower semi-continuous: closed *epigraph*

$$\operatorname{epi} f = \{(x, \zeta) \in \operatorname{dom} f \times \mathbb{R} \mid f(x) \leq \zeta\},$$

Reminder of the context

$$\widehat{x} \in \operatorname*{Argmin}_{x \in \mathcal{H}} f(x), \quad \mathcal{H} \quad \text{a Hilbert space}$$

with possibility to consider $\widetilde{f}(x) = \begin{cases} f(x) & \text{if } x \in D \\ \infty & \text{otherwise} \end{cases}$

D *domain* of the function $\operatorname{dom} \widetilde{f} \equiv \{x \in \mathcal{H} \mid \widetilde{f}(x) < \infty\}$

$f \in \Gamma_0(\mathcal{H})$

- proper: $\operatorname{dom} f \neq \emptyset$,
- lower semi-continuous: closed *epigraph*

$$\operatorname{epi} f = \{(x, \zeta) \in \operatorname{dom} f \times \mathbb{R} \mid f(x) \leq \zeta\},$$

- convex: convex epigraph.

Reminder of the context

Questions:

- Existence and uniqueness of $\widehat{x}$
  $\longrightarrow$ coercivity or compactness *(existence)*
  $\longrightarrow$ strict convexity *(uniqueness)*

- Characterization of $\widehat{x}$

  $\longrightarrow$ $\nabla f(\widehat{x}) = 0$ if f Gâteaux-differentiable
  $\longrightarrow$ $0 \in \partial f(\widehat{x})$ is f non-smooth

$$\partial f : \begin{cases} \mathcal{H} & \to & 2^{\mathcal{H}} \\ x & \mapsto & \{u \in \mathcal{H} \mid (\forall y \in \mathcal{H}), \quad \langle y - x \mid u \rangle + f(x) \leq f(y)\} \end{cases}$$

Reminder of the context

Explicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_n)$$

Reminder of the context

Explicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_n)$$

Implicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_{n+1})$$

Reminder of the context

Explicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_n)$$

Implicit (sub)gradient descent

$$x_{n+1} + \gamma \partial f(x_{n+1}) = x_n$$

Reminder of the context

Explicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_n)$$

Implicit (sub)gradient descent

$$(\mathrm{Id} + \gamma \partial f)\, x_{n+1} = x_n$$

Reminder of the context

Explicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_n)$$

Implicit (sub)gradient descent

$$(\mathrm{Id} + \gamma \partial f)\, x_{n+1} = x_n$$

Proximal operator

$$\mathrm{prox}_{\gamma f}(x) = \operatorname*{arg\,min}_{y \in \mathcal{H}} \frac{1}{2}\|y - x\|^2 + \gamma f(y) = (\mathrm{Id} + \gamma \partial f)^{-1}(x)$$

Reminder of the context

Explicit (sub)gradient descent

$$x_{n+1} = x_n - \gamma \partial f(x_n)$$

Implicit (sub)gradient descent

$$(\mathrm{Id} + \gamma \partial f)\, x_{n+1} = x_n$$

Proximal operator

$$\mathrm{prox}_{\gamma f}(x) = \underset{y \in \mathcal{H}}{\arg\min}\, \frac{1}{2}\|y - x\|^2 + \gamma f(y) = (\mathrm{Id} + \gamma \partial f)^{-1}(x)$$

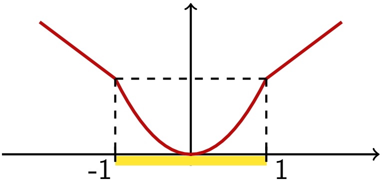
Another way to say that $p = \mathrm{prox}_f(x) \quad \Leftrightarrow \quad x - p \in \partial f(p)$.

Exercise 1

Provide an example of a function $f : \mathbb{R} \to \mathbb{R}$ and a nonempty set $C \subset \mathbb{R}$ such that

- f is nonconvex
- C is convex
- $f + \iota_C$ is convex.

$C = [-1, 1]$



$$f(x) = \begin{cases} x^2 & \text{if } |x| \leq 1 \\ \frac{3}{4}|x| + \frac{1}{4} & \text{otherwise} \end{cases}$$

Exercise 2

1. Let $f : \mathcal{H} \to]-\infty, +\infty]$ be a convex function.
   Prove that for every $\zeta \in \mathbb{R}$, the lower level set

   $$\operatorname{lev}_{\leq \zeta} f = \left\{ x \in \mathcal{H} \;\middle|\; f(x) \leq \zeta \right\}$$

   is convex.
   Let $x, y \in \operatorname{lev}_{\leq \zeta} f$, and $\alpha \in [0,1]$, then

   $$\begin{aligned} f((1-\alpha)x + \alpha y) &\underset{f \text{ convex}}{\leq} (1-\alpha)f(x) + \alpha f(y) \\ &\underset{x,y \in \operatorname{lev}_{\leq \zeta} f}{\leq} (1-\alpha)\zeta + \alpha\zeta = \zeta \end{aligned}$$

2. Show that the converse is false by providing an example of a nonconvex function the lower level sets of which are all convex.

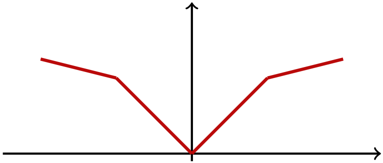
Exercise 2

1. Let $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ be a convex function.
   Prove that for every $\zeta \in \mathbb{R}$, the lower level set

$$\operatorname{lev}_{\leq \zeta} f = \{x \in \mathcal{H} \mid f(x) \leq \zeta\}$$

   is convex.

2. Show that the converse is false by providing an example of a nonconvex function the lower level sets of which are all convex.



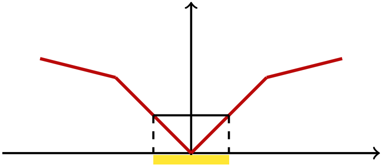
Exercise 2

1. Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ be a convex function.
   Prove that for every $\zeta \in \mathbb{R}$, the lower level set

$$\operatorname{lev}_{\leq \zeta} f = \left\{ x \in \mathcal{H} \;\middle|\; f(x) \leq \zeta \right\}$$

   is convex.

2. Show that the converse is false by providing an example of a nonconvex function the lower level sets of which are all convex.



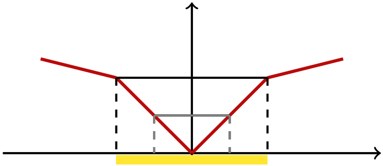
Exercise 2

1. Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ be a convex function.
   Prove that for every $\zeta \in \mathbb{R}$, the lower level set

$$\operatorname{lev}_{\leq \zeta} f = \left\{ x \in \mathcal{H} \;\middle|\; f(x) \leq \zeta \right\}$$

   is convex.

2. Show that the converse is false by providing an example of a nonconvex function the lower level sets of which are all convex.



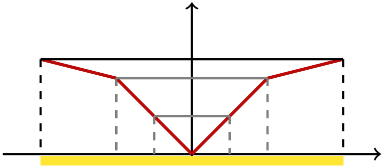
Exercise 2

1. Let $f\colon \mathcal{H} \to\ ]-\infty, +\infty]$ be a convex function.
   Prove that for every $\zeta \in \mathbb{R}$, the lower level set

$$\operatorname{lev}_{\leq \zeta} f = \left\{ x \in \mathcal{H} \;\middle|\; f(x) \leq \zeta \right\}$$

   is convex.

2. Show that the converse is false by providing an example of a nonconvex function the lower level sets of which are all convex.



Exercise 3

Let $y \in \mathbb{R}$. Show that

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto \log(1 + \exp(-yx))$$

is convex. When is it strictly convex ?

Exercise 3

Let $y \in \mathbb{R}$. Show that

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto \log(1 + \exp(-yx))$$

is convex. When is it strictly convex ?

$$f'(x) = \frac{-y \exp(-yx)}{1 + \exp(-yx)} = \frac{-y}{\exp(yx) + 1}$$

Exercise 3

Let $y \in \mathbb{R}$. Show that

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto \log(1 + \exp(-yx))$$

is convex. When is it strictly convex ?

$$f'(x) = \frac{-y\exp(-yx)}{1+\exp(-yx)} = \frac{-y}{\exp(yx)+1}$$

$$f''(x) = \frac{-(-y) \times y\exp(yx)}{(\exp(yx)+1)^2} = \frac{y^2\exp(yx)}{(\exp(yx)+1)^2} \geq 0$$

Exercise 3

Let $y \in \mathbb{R}$. Show that

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto \log(1 + \exp(-yx))$$

is convex. When is it strictly convex ?

$$f'(x) = \frac{-y\exp(-yx)}{1+\exp(-yx)} = \frac{-y}{\exp(yx)+1}$$

$$f''(x) = \frac{-(-y) \times y\exp(yx)}{(\exp(yx)+1)^2} = \frac{y^2\exp(yx)}{(\exp(yx)+1)^2} \geq 0$$

f is strictly convex iff $y \neq 0$.

Exercise 4

Let $A \in \mathbb{R}^{M \times N}$ and $z \in \mathbb{R}^M$. Let $g\colon \mathbb{R}^N \to \mathbb{R}\colon x \mapsto \|Ax - z\|^2$.

1. Prove that g is convex.
2. Give a necessary and sufficient condition on A for g to be strictly convex.
3. Find the minimizers of g.

Exercise 4

Let $A \in \mathbb{R}^{M\times N}$ and $z \in \mathbb{R}^M$. Let $g\colon \mathbb{R}^N \to \mathbb{R}\colon x \mapsto \|Ax - z\|^2$.

1. Prove that g is convex.
2. Give a necessary and sufficient condition on A for g to be strictly convex.
3. Find the minimizers of g.

1. g is twice differentiable and

$$\nabla g(x) = 2A^\top (Ax - z),$$

Exercise 4

Let $A \in \mathbb{R}^{M\times N}$ and $z \in \mathbb{R}^{M}$. Let $g \colon \mathbb{R}^N \to \mathbb{R} \colon x \mapsto \|Ax - z\|^2$.

1. Prove that g is convex.
2. Give a necessary and sufficient condition on A for g to be strictly convex.
3. Find the minimizers of g.

1. g is twice differentiable and

$$\nabla g(x) = 2A^\top (Ax - z), \quad \nabla^2 g(x) = 2A^\top A$$

Exercise 4

Let $A \in \mathbb{R}^{M \times N}$ and $z \in \mathbb{R}^M$. Let $g \colon \mathbb{R}^N \to \mathbb{R} \colon x \mapsto \|Ax - z\|^2$.

1. Prove that g is convex.
2. Give a necessary and sufficient condition on A for g to be strictly convex.
3. Find the minimizers of g.

1. g is twice differentiable and

$$\nabla g(x) = 2A^\top (Ax - z), \quad \nabla^2 g(x) = 2A^\top A$$

The Hessian of g is *positive*, hence g is convex.

Exercise 4

Let $A \in \mathbb{R}^{M\times N}$ and $z \in \mathbb{R}^M$. Let $g\colon \mathbb{R}^N \to \mathbb{R}\colon x \mapsto \|Ax - z\|^2$.

1. Prove that g is convex.
2. Give a necessary and sufficient condition on A for g to be strictly convex.
3. Find the minimizers of g.

1. g is twice differentiable and

$$\nabla g(x) = 2A^\top (Ax - z)\,, \quad \nabla^2 g(x) = 2A^\top A$$

   The Hessian of g is *positive*, hence g is convex.

2. If $A^\top A$ is *definite* positive, then g is strictly convex.
   Reciprocally if $A^\top A$ is not *definite*, then $\exists v \in \mathbb{R}^N$ s.t. $Av = 0$ and thus $g(x+v) = g(x)$, that is g has a “flat” direction (along eigenvector v).

Exercise 4

Let $A \in \mathbb{R}^{M \times N}$ and $z \in \mathbb{R}^M$. Let $g \colon \mathbb{R}^N \to \mathbb{R} \colon x \mapsto \|Ax - z\|^2$.

1. Prove that g is convex.
2. Give a necessary and sufficient condition on A for g to be strictly convex.
3. Find the minimizers of g.

3. Since g is convex and coercive, it has (at least) one minimizer and

$$\widehat{x} \text{ is a minimizer of } g \text{ iff } \nabla g(\widehat{x}) = 0,$$

Exercise 4

Let $A \in \mathbb{R}^{M\times N}$ and $z \in \mathbb{R}^M$. Let $g\colon \mathbb{R}^N \to \mathbb{R}\colon x \mapsto \|Ax - z\|^2$.

1. Prove that g is convex.
2. Give a necessary and sufficient condition on A for g to be strictly convex.
3. Find the minimizers of g.

3. Since g is convex and coercive, it has (at least) one minimizer and

$$\widehat{x} \text{ is a minimizer of } g \text{ iff } \nabla g(\widehat{x}) = 0, \text{ that is iff}$$

$$2A^\top(A\widehat{x} - z) = 0 \quad \Leftrightarrow \quad A^\top A\widehat{x} - A\top z = 0$$

Exercise 4

Let $A \in \mathbb{R}^{M\times N}$ and $z \in \mathbb{R}^M$. Let $g\colon \mathbb{R}^N \to \mathbb{R}\colon x \mapsto \|Ax - z\|^2$.

1. Prove that g is convex.
2. Give a necessary and sufficient condition on A for g to be strictly convex.
3. Find the minimizers of g.

3. Since g is convex and coercive, it has (at least) one minimizer and

$$\widehat{x} \text{ is a minimizer of } g \text{ iff } \nabla g(\widehat{x}) = 0, \text{ that is iff}$$

$$\begin{aligned} 2A^\top (A\widehat{x} - z) = 0 \quad &\Leftrightarrow \quad A^\top A\widehat{x} - A\top z = 0 \\ &\Leftrightarrow \quad \widehat{x} = \left(A^\top A\right)^{-1} A^\top z \end{aligned}$$

if $A^\top A$ is *definite* positive.

Exercise 5

Let $\mathcal{H}$ be a Hilbert space and let $f\colon \mathcal{H} \to \,]-\infty,+\infty]$ be a convex function. Let g be the perspective function of f defined as

$$(\forall (x,t) \in \mathcal{H} \times \mathbb{R}) \qquad g(x,t) = \begin{cases} t\, f(x/t) & \text{if } t > 0 \\ +\infty & \text{otherwise.} \end{cases}$$

1. How is the epigraph of g related to the epigraph of f ?
2. Deduce that g is a convex function.
3. As a consequence of this result, show that the Kullback-Leibler divergence defined as

$$\big(\forall x = (x^{(i)})_{1\le i\le N} \in \mathbb{R}^N\big)\big(\forall y = (y^{(i)})_{1\le i\le N} \in \mathbb{R}^N\big)$$
$$h(x,y) = \begin{cases} \sum_{i=1}^N x^{(i)} \ln(x^{(i)}/y^{(i)}) & \text{if } (x,y) \in (\,]0,+\infty[^N)^2 \\ +\infty & \text{otherwise,} \end{cases}$$

is convex.

Exercise 5

Let $\mathcal{H}$ be a Hilbert space and let $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ be a convex function. Let g be the perspective function of f defined as

$$(\forall (x,t) \in \mathcal{H} \times \mathbb{R}) \qquad g(x,t) = \begin{cases} t\, f(x/t) & \text{if } t > 0 \\ +\infty & \text{otherwise.} \end{cases}$$

1. Link between epigraphs

$$\begin{aligned} \operatorname{epi} g &= \{(y,\zeta) \in \operatorname{dom} g \times \mathbb{R} \mid g(y) \le \zeta\} \\ &= \{(x,t,\zeta) \in \operatorname{dom} f \times \mathbb{R}_+^* \times \mathbb{R} \mid tf(x/t) \le \zeta\} \\ &= \{(x,t,\zeta) \in \operatorname{dom} f \times \mathbb{R}_+^* \times \mathbb{R} \mid f(x/t) \le \zeta/t\} \\ &= \{(x,t,\zeta),\ t \in \mathbb{R}_+^*, (x/t, \zeta/t) \in \operatorname{epi} f\} \\ &= \{(tx', t\zeta', t),\ t \in \mathbb{R}_+^*, (x', \zeta') \in \operatorname{epi} f\} \end{aligned}$$

setting $x' = x/t, \zeta' = \zeta/t$.

The epigraph of g is the perspective of the epigraph of f.

Exercise 5

Let $\mathcal{H}$ be a Hilbert space and let $f\colon \mathcal{H} \to \left]-\infty, +\infty\right]$ be a convex function. Let g be the perspective function of f defined as

$$(\forall (x,t) \in \mathcal{H} \times \mathbb{R}) \qquad g(x,t) = \begin{cases} t\, f(x/t) & \text{if } t > 0 \\ +\infty & \text{otherwise.} \end{cases}$$

2. f convex $\Leftrightarrow \operatorname{epi} f$ convex $\underset{\text{Property of perspective}}{\Rightarrow} \operatorname{epi} g$ convex $\Leftrightarrow$ g convex
3. One remarks that the function $g(x^{(i)}, y^{(i)}) =$

$$\begin{cases} x^{(i)} \ln(x^{(i)}/y^{(i)}) = x^{(i)} \left(-\ln(y^{(i)}/x^{(i)})\right) & \text{if } x^{(i)},\ y^{(i)} > 0 \\ +\infty & \text{otherwise} \end{cases}$$

is the perspective function of the **convex** function $-\ln$, and thus it is convex. h being the sum of convex functions it is convex.

Exercise 6: Huber function

Let $\rho > 0$ and set

$$f : \mathbb{R} \to \mathbb{R} : \mapsto \begin{cases} \frac{x^2}{2}, & \text{if } |x| \leq \rho \\ \rho|x| - \frac{\rho^2}{2}, & \text{otherwise} \end{cases}$$

1. What is the domain of f ?
2. Is f differentiable ? twice-differentiable ?
3. Prove that f is convex.

1. $(\forall x \in \mathbb{R}) \quad f(x) < \infty$, thus $\operatorname{dom} f = \mathbb{R}$.
2. f is differentiable on $\mathbb{R} \setminus \{\pm\rho\}$. Further

$$\lim_{x \to \rho^-} \underbrace{f'(x)}_{=x} = \rho = \lim_{x \to \rho^+} \underbrace{f'(x)}_{=\rho}$$

Thus f is differentiable at $x = \rho$ and by symmetry, f is also differentiable at $-\rho$. Finally f is differentiable on $\mathbb{R}$.

Exercise 6: Huber function

Let $\rho > 0$ and set

$$f : \mathbb{R} \to \mathbb{R} : \mapsto \begin{cases} \frac{x^2}{2}, & \text{if } |x| \leq \rho \\ \rho|x| - \frac{\rho^2}{2}, & \text{otherwise} \end{cases}$$

1. What is the domain of f ?
2. Is f differentiable ? twice-differentiable ?
3. Prove that f is convex.
2. f is twice-differentiable on $\mathbb{R} \setminus \{\pm\rho\}$ and

$$f''(x) = \begin{cases} 1 & \text{if } |x| < \rho \\ 0 & \text{if } |x| > \rho \end{cases}$$

thus it is not twice-differentiable.

Exercise 6: Huber function

Let $\rho > 0$ and set

$$f : \mathbb{R} \to \mathbb{R} : \mapsto \begin{cases} \frac{x^2}{2}, & \text{if } |x| \leq \rho \\ \rho|x| - \frac{\rho^2}{2}, & \text{otherwise} \end{cases}$$

1. What is the domain of f ?
2. Is f differentiable ? twice-differentiable ?
3. Prove that f is convex.
3. f is differentiable on $\mathbb{R}$ and

$$f'(x) = \begin{cases} -\rho & \text{if } x < -\rho \\ x & \text{if } |x| < \rho \\ \rho & \text{otherwise} \end{cases}$$

which is increasing. Thus f is convex.

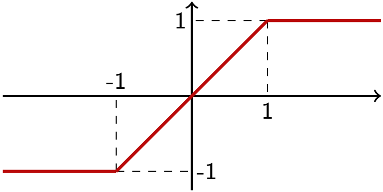
Exercise 1: Huber function

Let $\rho > 0$ and set

$$f : \mathbb{R} \to \mathbb{R} : \mapsto \begin{cases} \frac{x^2}{2}, & \text{if } |x| \leq \rho \\ \rho|x| - \frac{\rho^2}{2}, & \text{otherwise.} \end{cases}$$

1. What is the domain of f ?
2. Plot the subdifferential of f.
3. Is f differentiable ? Prove that f is convex.

2. (See the computation of $f'(x)$ done above.) For $\rho = 1$



Exercise 2

Let $\mathcal{H}$ be a Hilbert space. Let $f\colon \mathcal{H} \to \,]-\infty, +\infty]$ and let $C \subset \mathcal{H}$ such that $\operatorname{dom} f \cap C \neq \varnothing$.

- Give a sufficient condition for $x \in \mathcal{H}$ to be a global minimizer of $f + \iota_C$.
- *Assume that $f \in \Gamma_0(\mathcal{H})$ and that C is a closed convex set.*

Then, from the properties of C, $\iota_C \in \Gamma_0(\mathcal{H})$.
From Fermat's rule,

$$\widehat{x} \text{ is a minimizer of } f + \iota_C \text{ iff } 0 \in \partial(f + \iota_C)(\widehat{x}).$$

Since $\operatorname{dom} f \cap C \neq \varnothing$, then $\partial(f + \iota_C) = \partial f + \partial \iota_C$.
Moreover $(\forall x \in \mathcal{H}) \quad \partial \iota_C(x) = N_C(x)$, the normal cone of C at x.

$$\text{Thus, } \widehat{x} \text{ is a minimizer of } f + \iota_C \text{ iff } 0 \in \partial f(\widehat{x}) + N_C(\widehat{x}).$$

That is *if the normal cone of C at $\widehat{x}$ contains a subgradient of f at $\widehat{x}$.*

Féjer monotonicity

Definition

A sequence $(x_n)_{n\in\mathbb{N}}$ is said to be Féjer monotone with respect to a set C if

$$(\forall c \in C)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - c\| \leq \|x_n - c\|.$$