Master 2 – Probabilités et Finance
Sorbonne Université et Ecole Polytechnique

Convexity, Optimization and Stochastic Control

Exercise 1 : Carathéodory's theorem

Let E be a vector space s.t. $\dim(E) = d \in \mathbb{N}^\star$ and $A \subset E$. Let $x \in E$, $m > d+1$ and $(x_1, \cdots, x_m) \in A^m$ s.t. $x = \sum_{i=1}^m \alpha_i x_i$ with $\alpha_i \geq 0$ and $\sum_{i=1}^m \alpha_i = 1$.

1. Justify that there exists $(\kappa_i)_{1 \leq i \leq m-1} \subset \mathbb{R}$ s.t. $\sum_{i=1}^{m-1} \kappa_i v_i = 0$ where $v_i = x_i - x_m$.
2. Deduce that, for all $\lambda \in \mathbb{R}$, $\sum_{i=1}^m \alpha_i^\lambda x_i = x$, $\sum_{i=1}^m \alpha_i^\lambda = 1$ where $\alpha_i^\lambda = \alpha_i - \lambda \mu_i$, $i \leq m$ for some μ_i to be determined.
3. Justify why there exists $\lambda^\star$ s.t. $\alpha_i^{\lambda^\star} \geq 0$ and there exists $i_\star \in \{1, \cdots, m\}$ with $\alpha_{i_\star}^{\lambda^\star} = 0$.
4. Conclude that x can be written as the convex combinaison of $m-1$ points of A.

Exercise 2 : Differentiability almost everywhere

Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be a convex map et $I = \mathrm{Int}(\mathrm{dom}(f))$. Let $x_0 \in I$, we recall that, whenever they exist,

$$f'_d(x_0) := \lim_{\varepsilon \to 0,\, \varepsilon > 0} \frac{f(x_0 + \varepsilon) - f(x_0)}{\varepsilon} \text{ and } f'_g(x_0) := \lim_{\varepsilon \to 0,\, \varepsilon > 0} \frac{f(x_0) - f(x_0 - \varepsilon)}{\varepsilon}$$

1. Let $\varepsilon > 0$ small enough s.t. $x_0 + \varepsilon$ and $x_0 - \varepsilon$ belong to I, by notice that $x_0 = \frac{x_0+\varepsilon}{2} + \frac{x_0-\varepsilon}{2}$, show that $f'_g(x_0) \leq f'_d(x_0)$.
2. Verify that $\lim_{x \to x_0,\, x \leq x_0} f'_d(x) \leq f'_g(x_0)$.
3. By using the fact that $x \mapsto f'_d(x)$ is non–decreasing, deduce that f is almost surely differentiable on the interior on its domain.

Exercise 3 : Optimization

Let $\mathcal{P}(\mathbb{R})$ be the set of probability measures on $\mathbb{R}$. This set is equipped with the weak convergence topology i.e. a sequence $(m_n)_{n \geq 1} \subset \mathcal{P}(\mathbb{R})$ converges to $m \in \mathcal{P}(\mathbb{R})$ if

$\lim_{n\to\infty} \int_{\mathbb{R}} g(x) m_n(\mathrm{d}x) = \int_{\mathbb{R}} g(x) m(\mathrm{d}x)$ for any continuous bounded map $g : \mathbb{R} \to \mathbb{R}$.

Let $\lambda \in \mathcal{P}(\mathbb{R})$ be the Gaussian probability i.e. $\lambda(\mathrm{d}x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathrm{d}x$. We introduce the relative entropy w.r.t. λ by: for any $m \in \mathcal{P}(\mathbb{R})$,

$$H(m|\lambda) := \int_{\mathbb{R}} \log\left(\frac{\mathrm{d}m}{\mathrm{d}\lambda}(x)\right) m(\mathrm{d}x) \;\text{ whenever }\; \frac{\mathrm{d}m}{\mathrm{d}\lambda} \text{ exists, and } H(m|\lambda) = \infty \;\text{ otherwise,}$$

where $\frac{\mathrm{d}m}{\mathrm{d}\lambda}$ is the density in the sense of Radon–Nikodym of m w.r.t. λ.

Let $\sigma > 0$ and $F : \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ be a continuous bounded map. We consider the problem

$$\inf_{m \in \mathcal{P}(\mathbb{R})} V^\sigma(m) \text{ where } V^\sigma(m) := F(m) + \frac{\sigma^2}{2} H(m|\lambda).$$

We recall the Donsker–Varadhan variational formula:

$$H(m|\lambda) = \sup_{\psi} \left\{ \int_{\mathbb{R}} \psi(x) m(\mathrm{d}x) - \log \int_{\mathbb{R}} e^{-\psi(x)} \lambda(\mathrm{d}x) \right\}$$

where the supremum is taken over the set of bounded Borel measurable maps.

1. By using the Donsker–Varadhan variational formula, verify that the map $m \mapsto H(m|\lambda)$ is convex is lower semi–continuous.
2. Let $\varepsilon > 0$ and a compact $K \subset \mathbb{R}$ s.t. $\lambda(\mathbb{R} \setminus K) \leq \varepsilon$. Let $M > 0$, by using the Donsker–Varadhan variational formula again, show that there exists $\delta > 0$ s.t. for any $m \in \mathcal{P}(\mathbb{R})$ verifying $H(m|\lambda) \leq M$, $m(\mathbb{R} \setminus K) \leq \delta$.

   This result guarantees that for each $M > 0$, the set $\{m : \; H(m|\lambda) \leq M\}$ is compact for the weak convergence topology.
3. By using the lower semi–continuity of H, show that there exists $m^{\sigma,\star} \in \mathcal{P}(\mathbb{R})$ s.t. $V^\sigma(m^{\sigma,\star}) = \inf_m V^\sigma(m)$.
4. Verify that $\liminf_{\sigma \to 0} V^\sigma(m^{\sigma,\star}) \geq \inf_m F(m)$.
5. Let $m \in \mathcal{P}(\mathbb{R})$. We assume that there exists $(m^\sigma)_{\sigma>0}$ converges to m with $\sup_\sigma H(m^\sigma|\lambda) < \infty$. Show that $\lim_{\sigma \to 0} V^\sigma(m^{\sigma,\star}) = \inf_m F(m)$.

Under some regularity over F, the optimum of $V^{\sigma,\star}$ is unique and can be explicitly computed. Let us see an example. We now consider that

$$F(m) := \int_{\mathbb{R}} f(x) m(\mathrm{d}x) \text{ for some bounded continuous map } f.$$

- Verify that $m^{\sigma,\star}(\mathrm{d}x) = \left(\int_{\mathbb{R}} e^{-\frac{2}{\sigma^2} f(x')} \lambda(\mathrm{d}x')\right)^{-1} e^{-\frac{2}{\sigma^2} f(x)} \lambda(\mathrm{d}x)$.
- Deduce that if the sequence $(m^{\sigma,\star})_{\sigma>0}$ converges to a dirac measure, the limit is necessary $m^\star(\mathrm{d}x) = \delta_{a^\star}(\mathrm{d}x)$ where $a^\star \in \arg\min f$.

Exercise 4 : Fenchel–Rockafellar Theorem and applications

Part I : Fenchel–Rockafellar Theorem

The goal here is to prove that: for two proper convex functions $(f, g) : \mathbb{R}^d \to \mathbb{R} \times \mathbb{R}$ s.t. $\mathrm{Dom}(f) \cap \mathrm{Dom}(g)$ is non–empty, we have for any $x_0 \in \mathbb{R}^d$

$$(f+g)^\star(x_0) = \inf_{z \in \mathbb{R}^d} \{f^\star(z) + g^\star(x_0 - z)\}. \tag{1}$$

1. Verify that for any $x \in \mathbb{R}^d$ we have $(f+g)^\star(x) \leq \inf_{z \in \mathbb{R}^d}\{f^\star(z) + g^\star(x-z)\}$.
2. Justify why the map $h(u) := \inf_{x \in \mathbb{R}^d} f(x) + g(x+u)$ is convex and its domain is non–empty ?
3. Let $x_0 = 0$.
   (a) Show that there exists $\overline{x} \in \mathbb{R}^d$ s.t.
   $$h(0) + \langle \overline{x}, u \rangle \leq h(u), \quad \text{for any } u \in \mathbb{R}^d.$$
   (b) Verify that for any x and u,
   $$f(x) + g(x+u) - \langle \overline{x}, u \rangle \geq h(0).$$
   (c) Deduce that (1) is true for $x_0 = 0$.
4. Let $z \in \mathbb{R}^d$ and we set $\ell(u) := g(u) - \langle z, u \rangle$.
   (a) Verify that $\ell^\star(x) = g^\star(x+z)$ and $(f+\ell)^\star(0) = (f+g)^\star(z)$.
   (b) Deduce (1) for the general case.

Part II – Application: Monge Kantorovich duality

Let $d, \ell \geq 1$. We give ourselves two probability distributions $\nu \in \mathcal{P}(\{1, \cdots, d\})$ and $\mu \in \mathcal{P}(\{1, \cdots, \ell\})$. We identify ν and μ as vectors of $\mathbb{R}^d$ and $\mathbb{R}^\ell$ respectively. We denote by $\Gamma(d, \ell) \subset \mathbb{R}^d \times \mathbb{R}^\ell$ the set $(\gamma(i,j))_{1 \leq i \leq d, 1 \leq j \leq \ell}$ s.t. $\gamma \geq 0$, $\sum_{j=1}^{\ell} \gamma(i,j) = \nu(j)$ and $\sum_{i=1}^{d} \gamma(i,j) = \mu(i)$. We consider $(c(i,j))_{1 \leq i \leq d, 1 \leq j \leq \ell} \in \mathbb{R}^d \times \mathbb{R}^\ell$. We want to prove that

$$\begin{aligned} &\inf_{\gamma \in \Gamma(d,\ell)} \sum_{i=1}^{d} \sum_{j=1}^{\ell} \gamma(i,j) c(i,j) \\ &= \sup \Big\{ \sum_{i=1}^{d} \phi(i)\nu(i) + \sum_{j=1}^{\ell} \psi(j)\mu(j) : \quad \phi(i) + \psi(j) \leq c(i,j), \quad \text{for any } (i,j) \Big\}. \end{aligned}$$

We introduce

$$f(\gamma) = \begin{cases} \sum_{i=1}^d \sum_{j=1}^\ell \gamma(i,j)c(i,j), & \text{if } \gamma \geq 0 \\ \\ +\infty, & \text{otherwise} \end{cases}$$

and

$$g(\gamma) = \begin{cases} 0, \;\text{if } \sum_{j=1}^\ell \gamma(i,j) = \nu(i) \text{ and } \sum_{i=1}^d \gamma(i,j) = \mu(j) \\ \\ +\infty, \;\text{otherwise} \end{cases}$$

1. Show that $f^\star(\pi) = 0$ if $\pi(i,j) \leq c(i,j)$ and $f^\star(\pi) = +\infty$ otherwise.
2. Compute that $g^\star(\pi)$ for any $\pi \in \mathbb{R}^d \times \mathbb{R}^\ell$.
3. Let $i_0 \neq i_1$ s.t. $\exists j_0 \neq j_1$, $\pi(i_0,j_0) - \pi(i_1,j_0) \neq \pi(i_0,j_1) - \pi(i_1,j_1)$. We suppose that $\pi(i_0,j_0) - \pi(i_1,j_0) > \pi(i_0,j_1) - \pi(i_1,j_1)$

   (a) Let $\eta \in \mathbb{R}^{d+\ell}$ with $\eta(i_0,j_0) = \eta(i_1,j_1) = 1$ and $\eta(i_0,j_1) = \eta(i_1,j_0) = -1$. Show that, for all t, $\sum_{j=1}^\ell (\gamma + t\eta)(i,j) = \nu(i)$ and $\sum_{i=1}^d (\gamma + t\eta)(i,j) = \mu(j)$ and

   $$\begin{aligned} &\sum_{i,j} \pi(i,j)(\gamma + t\eta)(i,j) \\ &= \sum_{i,j} \pi(i,j)\gamma(i,j) + t(\pi(i_0,j_0) + \pi(i_1,j_1) - \pi(i_1,j_0) - \pi(i_0,j_1)) \end{aligned}$$

   (b) Deduce that $g^\star(\pi) = \sum_i \phi(i)\nu(i) + \sum_j \psi(j)\mu(j)$ for $\pi(i,j) = \phi(i) + \psi(j)$ and $g^\star(\pi) = +\infty$ otherwise.

   (c) Apply the Fenchel–Rockafellar Theorem and deduce the result

Part III – Application: Super replication in complete market

The general version of the Monge Kantorovich duality is

$$\begin{aligned} &\inf_{(X,Y)\in\Theta} \mathbb{E}^{\mathbb{P}}[c(X,Y)] \\ &= \sup\left\{ \int_{\mathbb{R}^d} \phi(e)\nu(\mathrm{d}e) + \int_{\mathbb{R}^\ell} \psi(u)\mu(\mathrm{d}u) : \;\; \phi(e) + \psi(u) \leq c(e,u), \;\text{ for any } (e,u) \right\} \end{aligned}$$

where $(X,Y) \in \Theta$ if $\mathcal{L}^{\mathbb{P}}(X) = \nu$ and $\mathcal{L}^{\mathbb{P}}(Y) = \mu$. The distributions ν and μ admit a finite first moment and c satisfies $|c(x,y)| \leq 1 + |x| + |y|$.

Let $(S^1_t)_{t\le T}$ and $(S^2_t)_{t\le T}$ be two assets martingale under the risk neutral measures $\mathbb{P}^1$ and $\mathbb{P}^2$ respectively (the interest rate $r=0$). The price p_0 of super replication of an European option with payoff $c(S^1_T, S^2_T)$ is given by

$$p_0 := \inf\left\{\mathbb{E}^{\mathbb{P}^1}[\lambda^1(S^1_T)] + \mathbb{E}^{\mathbb{P}^2}[\lambda^2(S^2_T)] :\ \ c(x,y) \le \lambda^1(x) + \lambda^2(y) \ \text{ for all } (x,y)\right\}.$$

Apply the Monge Kantorovich duality and comment.

Bonus exercise : Maximum Principle (deterministic case)

Let $T>0$ and $A\subset\mathbb{R}$. We give ourselves the maps $[0,T]\times\mathbb{R}\times A \ni (t,x,a)\mapsto (b,f)(t,x,a)\in\mathbb{R}\times\mathbb{R}$ and $\mathbb{R}\ni x\mapsto g(x)\in\mathbb{R}$ differentiable in (x,a) with bounded derivatives and there exists $C>0$ s.t. for each (t,x,a)

$$|b(t,x,a)| + |f(t,x,a)| + |g(x)| \le C(1+|x|).$$

We denote by $\mathcal{A}$ the set of all predictable process $(\alpha_t)_{t\in[0,T]}$ satisfying $\int_0^T \alpha_t^2 \mathrm{d}t < \infty$. Wet set $x\in\mathbb{R}$. For any $\alpha\in\mathcal{A}$, let X^α be a process satisfying:

$$X^\alpha_t = x + \int_0^t b(r, X^\alpha_r, \alpha_r)\,\mathrm{d}r, \quad \text{for each } t\le T.$$

The goal of this exercise is to solve the following optimization problem

$$R := \inf_{\alpha\in\mathcal{A}} J(\alpha) \ \text{ with } \ J(\alpha) := \int_0^T f(t, X^\alpha_t, \alpha_t)\mathrm{d}t + g(X^\alpha_T).$$

We introduce for any α and $\beta\in\mathcal{A}$, the process $V^{\alpha,\beta} := V$ verifying:

$$V_t = \int_0^t V_r\partial_x b(r, X^\alpha_r, \alpha_r) + \beta_r \partial_a b(r, X^\alpha_r, \alpha_r)\,\mathrm{d}r, \quad \text{for each } t\le T.$$

1. Let $\alpha,\beta\in\mathcal{A}$. We assume that for any sufficiently small $\varepsilon>0$, $\alpha^\varepsilon := \alpha+\varepsilon\beta\in\mathcal{A}$. Show that

$$\lim_{\varepsilon\to 0}\sup_{t\in[0,T]}\left|\frac{X^{\alpha^\varepsilon}_t - X^\alpha_t}{\varepsilon} - V^{\alpha,\beta}_t\right| = 0.$$

2. Deduce that

$$\lim_{\varepsilon\to 0}\frac{J(\alpha+\varepsilon\beta) - J(\alpha)}{\varepsilon} = \int_0^T V^{\alpha,\beta}_t \partial_x f(t, X^\alpha_t, \alpha_t) + \beta_t\partial_a f(t, X^\alpha_t, \alpha_t)\,\mathrm{d}t + V^{\alpha,\beta}_T \partial_x g(X^\alpha_T).$$

We introduce the map $H : [0,T] \times \mathbb{R} \times \mathbb{R} \times A \to \mathbb{R}$ defined by

$$H(t,x,y,a) := b(t,x,a)\; y + f(t,x,a).$$

The map H is called the Hamiltonian. For $\alpha \in \mathcal{A}$, we denote Y^α the process verifying

$$Y_t = \partial_x g(X_T^\alpha) - \int_t^T \partial_x H(r, X_r^\alpha, Y_r^\alpha, \alpha_r)\; \mathrm{d}r, \quad \text{for all } t \leq T.$$

3. Show that

$$Y_T^\alpha V_T^{\alpha,\beta} = \int_0^T Y_t^\alpha \partial_a b(t, X_t^\alpha, \alpha_t)\beta_t - \partial_x f(t, X_t^\alpha, \alpha_t) V_t^{\alpha,\beta}\; \mathrm{d}t.$$

4. Deduce that

$$\lim_{\varepsilon \to 0} \frac{J(\alpha + \varepsilon\beta) - J(\alpha)}{\varepsilon} = \int_0^T \partial_a H(t, X_t^\alpha, Y_t^\alpha, \alpha_t)\beta_t\; \mathrm{d}t.$$

5. We assume that H is convex in a, show that if α is an optimal control then for each $a \in A$,

$$H(t, X_t^\alpha, Y_t^\alpha, a) \geq H(t, X_t^\alpha, Y_t^\alpha, \alpha_t).$$

6. Let us assume that g is convex, H is convex in (x,a) and $\alpha \in \mathcal{A}$ s.t.

$$H(t, X^\alpha, Y_t^\alpha, \alpha_t) = \inf_{a \in A} H(t, X_t^\alpha, Y_t^\alpha, a).$$

Show that $\alpha \in \mathcal{A}$ is an optimal control i.e. $R = J(\alpha)$.

Exercice 1 : Soit E un ev tq $\dim(E) = d \in \mathbb{N}^*$ et $A \subset E$.

Soit $x \in E$, $m > d+1$ et $(x_1, \dots, x_m) \in A^m$ tq $x = \sum_{i=1}^{m} \alpha_i x_i$ avec $\alpha_i \geq 0$ et $\sum_{i=1}^{m} \alpha_i = 1$

1) Justifier qu'$\exists (k_i)_{1 \leq i \leq m-1} \subset \mathbb{R}$ tq $\sum_{i=1}^{m-1} k_i v_i = 0$ où $v_i = x_i - x_m$

Comme $m-1 > d$ ($= \dim(E)$) les points $v_1, \dots, v_{m-1}$ sont liés donc $\exists (k_i)_{1 \leq i \leq m-1}$ tq $\sum_{i=1}^{m-1} k_i v_i = 0$

2) Déduire que, $\forall \lambda \in \mathbb{R}$, $\sum_{i=1}^{m} \alpha_i^\lambda x_i = x$, $\sum_{i=1}^{m} \alpha_i^\lambda = 1$ où $\alpha_i^\lambda = \alpha_i - \lambda \mu_i$ $i \leq m$ pour des μ_i à déterminer.

Si on prend $\mu_i = k_i$ pour $i \in [\![1, m-1]\!]$ et $\mu_m = -\sum_{i=1}^{m-1} k_i$

Alors, $\sum_{i=1}^{m} \alpha_i^\lambda x_i = \sum_{i=1}^{m} (\alpha_i - \lambda \mu_i) x_i = \sum_{i=1}^{m-1} (\alpha_i - \lambda k_i) x_i + (\alpha_m + \lambda \sum_{i=1}^{m-1} k_i) x_m$

$$= \underbrace{\sum_{i=1}^{m} \alpha_i x_i}_{=x} - \lambda \sum_{i=1}^{m-1} k_i \underbrace{(x_i - x_m)}_{v_i} \quad \text{or } \sum_{i=1}^{m-1} k_i v_i = 0$$

$$= x$$

3) Justifier qu'$\exists \lambda^*$ tq $\alpha_i^{\lambda^*} \geq 0$ et $\exists i_* \in \{1, \dots, m\}$ avec $\alpha_{i_*}^{\lambda^*} = 0$

Si λ^* existe $\Rightarrow \lambda^* \in V := \{\lambda \in \mathbb{R}, \alpha_i - \lambda \mu_i \geq 0, i \leq m\}$

Pour $\lambda^* = \min_{i \leq m} \left(\frac{\alpha_i}{\mu_i}\right)$ cela est vérifié et comme $\left(\frac{\alpha_i}{\mu_i}\right)_{1 \leq i \leq m}$ est fini, $\exists i_*$ tq $\lambda^* = \frac{\alpha_{i_*}}{\mu_{i_*}} \Rightarrow \alpha_{i_*}^{\lambda^*} = \alpha_{i_*} - \lambda^* \mu_{i_*} = \alpha_{i_*} - \frac{\alpha_i^*}{\mu_{i_*}} \mu_{i_*} = 0$

4) Conclure que x peut être écrit comme une combinaison cvx de $m-1$ pts de A.

On a $\alpha_i^{\lambda^*} \geq 0$ pour $i \leq m$ et $\alpha_{i_*}^{\lambda^*} = 0$ alors $\sum_{\substack{i=1 \\ i \neq i_*}}^{m} \alpha_i^{\lambda^*} x_i = x$

Théorème de Carathéodory :

En dim d, $\forall A \subset \mathbb{R}^d$ et $\forall x \in \mathrm{Conv}(A)$ alors x s'écrit comme combinaison convexe de $d+1$ éléments de A.

Exercice 2 : $f : \mathbb{R} \to \overline{\mathbb{R}}$ cvx et $I = \mathrm{int}(\mathrm{dom}(f))$. Soit $x_0 \in I$,

$$f_d'(x_0) := \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} \frac{f(x_0+\varepsilon) - f(x_0)}{\varepsilon} \quad \text{et} \quad f_g'(x_0) := \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} \frac{f(x_0) - f(x_0-\varepsilon)}{\varepsilon}$$

1) Soit $\varepsilon > 0$ suffisamment petit tq $x_0 + \varepsilon$ et $x_0 - \varepsilon \in I$, en remarquant que $x_0 = \frac{x_0+\varepsilon}{2} + \frac{x_0 - \varepsilon}{2}$, mq $f_g'(x_0) \leq f_d'(x_0)$

f est cvx $\Rightarrow f(x_0) \leq \frac{1}{2} f(x_0+\varepsilon) + \frac{1}{2} f(x_0 - \varepsilon) \Rightarrow 0 \leq \frac{1}{2}(f(x_0+\varepsilon) - f(x_0)) + \frac{1}{2}(f(x_0-\varepsilon) - f(x_0))$

$\Rightarrow 0 \leq \frac{1}{\varepsilon}(f(x_0+\varepsilon) - f(x_0)) + \frac{1}{\varepsilon}(f(x_0-\varepsilon) - f(x_0))$

$\Rightarrow \frac{1}{\varepsilon}(f(x_0) - f(x_0-\varepsilon)) \leq \frac{1}{\varepsilon}(f(x_0+\varepsilon) - f(x_0))$

$\Rightarrow$ lorsque $\varepsilon \to 0$, $f_g'(x_0) \leq f_d'(x_0)$

2) Vérifier que $\lim_{\substack{x \to x_0 \\ x \leq x_0}} f_d'(x) \leq f_g'(x_0)$

$$f_d'(x_0) = \lim_{y \to x_0^+} \frac{f(y) - f(x_0)}{y - x_0} \quad \text{et} \quad f_d'(x_0) = \lim_{y \to x_0^-} \frac{f(x_0) - f(y)}{x_0 - y}$$

On sait aussi que $y \mapsto \frac{f(y) - f(x_0)}{y - x_0} \nearrow$

$\Rightarrow$ Soit $x \leq y < x_0$, $\frac{f(y) - f(x)}{y - x} \leq \frac{f(x_0) - f(x)}{x_0 - x}$

Pour $y \underset{y > x}{\to} x$ $\quad f_d'(x) \leq \frac{f(x_0) - f(x)}{x_0 - x} \Rightarrow$ Pour $x \underset{x < x_0}{\to} x_0$ $\quad f_d'(x) \leq f_g'(x_0)$

3) En utilisant que $x \mapsto f_d'(x)$ est non décroissante, déduire que f est pas différentiable à l'intérieur de son domaine.

Rappel : Si $f : E \to \overline{\mathbb{R}}$, $\mathrm{dom}(f) = \{x \in E, f(x) < \infty\}$

Si f est dérivable en $x_0 \Rightarrow f_d'(x_0) = f_g'(x_0)$.

Par 1) et 2), $\lim_{\substack{x \to x_0 \\ x < x_0}} f_d'(x_0) \leq f_g'(x_0) \leq f_d'(x_0)$

Version prof

Si $x \mapsto f'_d(x)$ est continue en x_0, $f'_d(x_0) = \lim_{x \to x_0^-} f'_d(x_0) \leq f'_g(x_0) \leq f'_d(x_0)$
$\Rightarrow f'_d(x_0) = f'_g(x_0) \Rightarrow f$ dérivable en x_0

Comme $x \mapsto f'_d(x)$ est $\nearrow$, elle est discontinue sur un ens. dénombrable $\Rightarrow f$ est non dérivable sur un ens dénombrable
$\Rightarrow f$ dérivable presque partout.

Version perso:

f cvx $\Rightarrow \forall x_0 \in I$, $f'_g(x_0) \leq f'_d(x_0)$
f'_d est non-décroissante sur I car f cvx
$\forall x_0$, $f(x)$ est différentiable ssi $f'_g(x_0) = f'_d(x_0)$

Les discontinuités potentielles de f'_d apparaissent uniquement si $f'_g(x_0) \neq f'_d(x_0)$.
Ces discontinuités correspondent aux pts où f n'est pas différentiable.
Cependant:

* f'_d, en tant que fct° $\nearrow$, peut être discontinue uniquement sur un ens. dénombrable
* $\Rightarrow f$ n'est pas différentiable uniquement sur un ens. dénombrable.

L'ens des pts où f n'est pas diff. est exactement l'ens des discontinuités de f'_d, qui est dénombrable.
Par conséquent:

* f est diff sur tout intervalle sauf un ens. dénomb.
* f est diff presque partout sur $I = \mathrm{int}(\mathrm{dom}(f))$

Exercice 3: Soit $\mathcal{P}(\mathbb{R})$ l'ens. des mesures de proba sur $\mathbb{R}$. Cet ens est muni de la topologie de la cv faible, c-à-d qu'une suite $(m_n)_{n \geq 1} \subset \mathcal{P}(\mathbb{R})$ CV vers $m \in \mathcal{P}(\mathbb{R})$ si $\lim_{n \to \infty} \int_{\mathbb{R}} g(x)\, m_n(dx) = \int_{\mathbb{R}} g(x)\, m(dx)$ $\forall g: \mathbb{R} \to \mathbb{R}$ bornée.
Soit $\lambda \in \mathcal{P}(\mathbb{R})$, une proba gaussienne, i.e $\lambda(dx) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$
On introduit l'entropie relative pr^t à λ comme suit:

$$\forall m \in \mathcal{P}(\mathbb{R}),\ H(m|\lambda) := \int_{\mathbb{R}} \log\left(\frac{dm}{d\lambda}(x)\right) m(dx)$$ lorsque $\frac{dm}{d\lambda}$ (la densité au sens de Radon-Nikodym de m pr^t à λ) existe, et $H(m|\lambda) = \infty$ sinon

Soit $\sigma > 0$ et $F: \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ une fct° continue et bornée.
On considère le pb suivant:

$$\inf_{m \in \mathcal{P}(\mathbb{R})} V^\sigma(m) \quad \text{où} \quad V^\sigma(m) := F(m) + \frac{\sigma^2}{2} H(m|\lambda)$$

Nous rappelons la formule variationnelle de Donsker-Varadhan:

$$H(m|\lambda) = \sup_{\varphi} \left\{ \int_{\mathbb{R}} \varphi(x)\, m(dx) - \log\left(\int_{\mathbb{R}} e^{\varphi(x)} \lambda(dx)\right) \right\},$$

où le supremum est pris sur l'ens des applications boréliennes bornées.

1) En utilisant la formule variationnelle de Donsker-Varadhan, mq $m \mapsto H(m|\lambda)$ est cvx et semi-continue inférieurement.

Par la formule ___ de $H(m|\lambda)$, on a $H(m|\lambda) = \sup_{\varphi} \{ \int \varphi\, dm - \log(\int e^{\varphi} d\lambda) \}$

pour chaque fct° borélienne bornée φ, $\mathcal{L}^\varphi(m): m \mapsto \int \varphi\, dm - \log(\int e^{\varphi} d\lambda)$ est cvx (car 1er terme linéaire en m et le 2^e $\perp$ de m)
Comme le sup de fct° cvx est cvx, alors $m \mapsto H(m|\lambda) = \sup_{\varphi} \mathcal{L}^\varphi(m)$ est cvx

Rappel (admis): $(f^a)_{a \in A}$ famille de fct° avec $A \neq \emptyset$. Si $\forall a \in A$, $y \in E \mapsto f^a(y)$ est continue et cvx où E espace topo, alors $\sup_a f^a$ est sci et cvx.

On observe $\sup_{\varphi \in \mathcal{B}} \{ \int \varphi\, dm - \log(\int e^{\varphi} d\lambda) \} = \sup_{\varphi \in C_b} \{ \int \varphi\, dm - \log(\int e^{\varphi} d\lambda) \}$ $\quad$ C_b un ss ens dense dans $\mathcal{B}$, ici l'espace des fct° continues et bornées

Si $\varphi \in C_b$, alors $m \mapsto \mathcal{L}^\varphi(m)$ est continue pour la topo de la cv faible et $m \mapsto H(m|\lambda) = \sup_{\varphi} \mathcal{L}^\varphi(m)$ est sci

ψ borel bornée $\Rightarrow \exists (\psi_m)_{m\geq 0} \in \mathcal{C}_0^\infty,\ \psi_m \xrightarrow[m\to\infty]{} \psi$ p.p.

2) Soit $\varepsilon>0$ et un compact $K\subset\mathbb{R}$ tq $\lambda(\mathbb{R}\setminus K)\leq\varepsilon$.

Soit $M>0$, en utilisant la formule - - -, mq $\exists\delta>0$ tq $\forall m\in\mathcal{P}(\mathbb{R})$ vérifiant $H(m|\lambda)\leq M$, alors $m(\mathbb{R}\setminus K)\leq\delta$.

Soit m tq $H(m|\lambda)\leq M$, $\forall\psi\in\mathcal{B}$, $\int\psi\,dm-\log(\int e^{\psi}d\lambda)\leq H(m|\lambda)\leq M$

$\Rightarrow \int\psi\,dm\leq M+\log(\int e^{\psi}d\lambda)$

$*_1 : \int e^{\psi}\mathbb{1}_K d\lambda = \int_K e^0 d\lambda = \lambda(K)$

$*_2 : \int e^{\psi}\mathbb{1}_{K^c} d\lambda = \int_{K^c} e^{c}d\lambda = e^{c}\lambda(K^c)$

$(m(K^c) = \int_{\mathbb{R}}\mathbb{1}_{K^c}(x)m(dx))$

Soit $\psi(x) = \underset{>0}{\underline{c}}\,\mathbb{1}_{K^c}(x) \Rightarrow c\,m(K^c)\leq M+\log\Big(\underbrace{\int e^{\psi}\mathbb{1}_K d\lambda}_{\lambda(K)}{}^{*_1}+\underbrace{\int e^{\psi}\mathbb{1}_{K^c}d\lambda}_{e^{c}\lambda(K^c)}{}^{*_2}\Big)$

$$\Rightarrow c\,m(K^c)\leq M+\log\big(\underbrace{1-\lambda(K^c)+e^{c}\lambda(K^c)}_{\lambda(K^c)(e^{c}-1)+1}\big)$$

$$\Rightarrow m(K^c)\leq\frac{M+\log(\varepsilon(e^{c}-1)+1)}{c}$$

On veut choisir c (en fct° de ε) tq si $\varepsilon\to 0$, $c\to\infty$

Il suffit de prendre $c=\log(1+1/\varepsilon) \Rightarrow \delta=\frac{M+\log(2)}{c}$

famille tendue

A_M compact $\Rightarrow$ Si K est tq $\lambda(K^c)\leq c$ alors $\forall m$ tq $H(m|\lambda)\leq M$, $\exists\delta(\varepsilon)$ vérifiant $\lim_{\varepsilon\to 0}\delta(\varepsilon)=0$ et $m(K^c)\leq\delta(\varepsilon)$) Thm de Prokhorov

$\Rightarrow \forall M>0$, l'ens $\{m : H(m|\lambda)\leq M\} = A^M$ est compact pour topo de la CV faible i.e $\forall(m_m)_{m\geq 0}\subset A_M$, $\exists\bar m$, $(m_k)_{k\geq 1}\subseteq\mathbb{N}$ tq $m_k \xrightarrow[k\to\infty]{}\infty$ et $m_{m_k}\xrightarrow[k\to\infty]{}\bar m$

Finir jusqu'au 5

28/11 suite de l'exo 3 (quest° après 5)

$F(m) := \int_{\mathbb{R}} f(x)\,m(dx)$, f donnée

$*$ Mq $m^{\sigma,*}(dx) = c\,\underbrace{e^{-\frac{2}{\sigma^2}f(x)}}_{>0\ \forall x\in\mathbb{R}}{}^{*}\lambda(dx)$ avec $c^{-1}=\int_{\mathbb{R}}e^{-\frac{2}{\sigma^2}f(x')}\lambda(dx')$

On veut mq $\underbrace{V^\sigma(m)}_{=F(m)+\frac{\sigma^2}{2}H(m|\lambda)}\geq V^\sigma(m^{\sigma,*})$ $\quad\forall m\in\mathcal{P}(\mathbb{R})$

Remarquons que : $H(m|\lambda) = \int_{\mathbb{R}}\log(\frac{dm}{d\lambda}(x))\,m(dx)$ $\quad m\in\mathcal{P}(\mathbb{R})$ dominée par λ. $m^{\sigma,*}$ est aussi dominée par λ. car $*$

$$= \int_{\mathbb{R}}\log\Big(\frac{dm}{dm^{\sigma,*}}(x)\,\frac{dm^{\sigma,*}}{d\lambda}(x)\Big)\,m(dx)$$

$$= \underbrace{\int_{\mathbb{R}}\log\Big(\frac{dm}{dm^{\sigma,*}}(x)\Big)m(dx)}_{H(m|m^{\sigma,*})} + \int_{\mathbb{R}}\log\Big(\frac{dm^{\sigma,*}}{d\lambda}(x)\Big)m(dx)$$

$$= H(m|m^{\sigma,*}) + \int_{\mathbb{R}}\log\big(c\,e^{-\frac{2}{\sigma^2}f(x)}\big)\,m(dx)$$

$$= H(m|m^{\sigma,*}) - \log(c^{-1}) - \frac{2}{\sigma^2}\underbrace{\int_{\mathbb{R}}f(x)\,m(dx)}_{F(m)}$$

Donc $F(m)+\frac{\sigma^2}{2}H(m|\lambda) = \frac{\sigma^2}{2}H(m|m^{\sigma,*}) - \frac{\sigma^2}{2}\log(c^{-1})$

$$= \frac{\sigma^2}{2}\underbrace{H(m|m^{\sigma,*})}_{\geq 0} - \frac{\sigma^2}{2}\log\Big(\int_{\mathbb{R}}e^{-\frac{2}{\sigma^2}f(x')}\lambda(dx')\Big)$$

On obtient que $V^\sigma(m)\geq -\frac{\sigma^2}{2}\log(c^{-1})$ $\quad\forall m\in\mathcal{P}(\mathbb{R})$ dominée par λ

On a aussi que $V^\sigma(m^{\sigma,*}) = \frac{\sigma^2}{2}\underbrace{H(m^{\sigma,*}|m^{\sigma,*})}_{=0} - \frac{\sigma^2}{2}\log(c^{-1})$

$$= -\frac{\sigma^2}{2}\log(c^{-1})$$

$\Rightarrow \inf_{m\in\mathcal{P}(\mathbb{R})}V^\sigma(m)\geq V^\sigma(m^{\sigma,*})$ d'où le résultat car $\Rightarrow m^{\sigma,*}$ est le min de $V^\sigma(m)$

$*$ Si $(m^{\sigma,*})_{\sigma>0}$ CV vers une dirac, alors $\lim m^{\sigma,*} = \delta_{a^*}$ où $a^* = \arg\min f$?

On sait que $\inf_m V^\sigma(m) = V^\sigma(m^{\sigma,*}) \xrightarrow[\sigma\to 0]{} \inf_m F(m)$ $\quad \inf_a f(a)$

Mais $\inf_m F(m) = \inf_m\big(\int_{\mathbb{R}}f(x)\,m(dx)\big)$

Soit $m\in\mathcal{P}(\mathbb{R})$, $\int_{\mathbb{R}}f(x)\,m(dx)\geq\inf_{a\in\mathbb{R}}(f(a))$

Soit $a\in\mathbb{R}$, $\delta_a\in\mathcal{P}(\mathbb{R})$, $f(a)\geq\inf_{m\in\mathcal{P}(\mathbb{R})}\big(\int f(x)\,m(dx)\big)$ $\quad\Rightarrow \inf_m\big(\int_{\mathbb{R}}f(x)\,m(dx)\big) = \inf_a(f(a))$

Pour conclure, $f(a^\infty) = \inf_a f(a)$ où $\delta_{a^\infty}(dx) = \lim m^{\sigma,*}(dx)$
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \| \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad m^\infty(dx)$

$$** \left(\underbrace{\underbrace{\int_{\mathbb{R}} f(x) m^\infty(dx)}_{F(m^\infty)}}_{\lim_{\sigma \to 0} F(m^{\sigma,*})} \overset{?}{=} \inf_m \Big(\underbrace{\int_{\mathbb{R}} f(x) m(dx)}_{F(m)} \Big) \right.$$

On sait que $F(m^{\sigma,*}) \leq V^\sigma(m^{\sigma,*})$

$$\Rightarrow \limsup_{\sigma \to 0} F(m^{\sigma,*}) \leq \lim_{\sigma \to 0} V^\sigma(m^{\sigma,*}) \underset{Q5}{=} \inf_m F(m)$$

Mais on a aussi que, $F(m^{\sigma,*}) \geq \inf_m F(m)$ par déf

$$\Rightarrow \liminf_{\sigma \to 0} F(m^{\sigma,*}) \geq \inf F(m)$$

On a majoré limsup et minoré liminf $\Rightarrow \lim_{\sigma \to 0} F(m^{\sigma,*}) = \inf_m F(m)$

D'où le résultat par **

Car $\Rightarrow f(a^\infty) = \inf_a f(a) \Rightarrow a^\infty \in \operatorname{argmin}(f)$

<u>Complément</u>: On peut mq $m^{\sigma,*}(dx) = \mathcal{L}(X_t)\ \forall t$ où $dX_t \overset{(*)}{=} -(\nabla f(X_t) + \frac{\sigma^2}{2} X_t)dt + \sigma dW_t$ sous des condit° sur f

mesure invariante

<u>Méthode de MC</u>: $X_{t^\kappa_{k+1}} = X_{t^\kappa_k} - (\nabla f(X_{t^\kappa_k}) + \frac{\sigma^2}{2} X_{t^\kappa_k})(t^\kappa_{k+1} - t^\kappa_k) + \sigma(W_{t^\kappa_{k+1}} - W_{t^\kappa_k})$

La discrétisat° de (*) est liée à la descente de gradient (stochastique) vu en cours.

Exercice 4 :

Partie I :

Le but ici est de prouver que : pour deux fct° cvx propres $(f,g) : \mathbb{R}^d \to \mathbb{R}\times\mathbb{R}$ tq $\mathrm{Dom}(f)\cap\mathrm{Dom}(g)\neq\emptyset$, on a $\forall x_0\in\mathbb{R}^d$

$$\underset{\sup_y(\langle x_0,y\rangle-(f+g)(y))}{\underbrace{(f+g)^*(x_0)}} = \inf_{z\in\mathbb{R}^d}\{f^*(z)+g^*(x_0-z)\} \quad (1)$$

$$h^*(x) := \sup_y \langle x,y\rangle - h(y)$$

1) Mq $\forall x\in\mathbb{R}^d$, on a $(f+g)^*(x) \leq \inf_{z\in\mathbb{R}^d}\{f^*(z)+g^*(x-z)\}$

$$(f+g)^*(x) = \sup_y(\langle x,y\rangle-(f+g)(y))$$

$$\forall(x,y,z),\ \langle x,y\rangle - f(y) - g(y) = \underbrace{\langle x-z,y\rangle - g(y)}_{\leq g^*(x-z)} + \underbrace{\langle z,y\rangle - f(y)}_{\leq f^*(z)}$$

$$\Rightarrow \langle x,y\rangle - (f+g)(y) \leq g^*(x-z)+f^*(z) \quad \forall x,y,z$$

$$\Rightarrow (f+g)^*(x) = \sup_y\{\langle x,y\rangle-(f+g)(y)\} \leq g^*(x-z)+f^*(z) \quad \forall z$$

$$\Rightarrow (f+g)^*(x) \leq \inf_z\{g^*(x-z)+f^*(z)\}$$

2) Mq $h(u) := \inf_{x\in\mathbb{R}^d}\{f(x)+g(x+u)\}$ est cvx et son domaine est non vide

Faux

$\forall x\in\mathbb{R}$, $J_x : u\mapsto f(x)+g(x+u)$ est cvx car g cvx donc ~~$h(\cdot) = \inf_{x\in\mathbb{R}^d}(J_x(\cdot))$ est cvx~~ à revoir !

$h(u) := \inf\{f(x)+g(x-u)\} = (f\,\square\,g)(u)$ l'inf convolut° de f et g. Ici on est dans le cas où l'inf convolut° est bien cvx.

On observe que $h(0) = \inf_x(f(x)+g(x))$

Comme $\mathrm{Dom}(f)\cap\mathrm{Dom}(g)\neq\emptyset$, alors $\exists x_0\in\mathrm{Dom}(f)\cap\mathrm{Dom}(g)$ tq $h(0)\leq f(x_0)+g(x_0)<\infty$

$$\Rightarrow 0\in\mathrm{Dom}(h)$$

3) Soit $x_0=0$

a) Mq $\exists\bar{x}\in\mathbb{R}^d$ tq $h(0)+\langle\bar{x},u\rangle\leq h(u)\quad\forall u\in\mathbb{R}^d$

Mq $\bar{x}\in\partial h(0)$

On veut donc mq $\underset{\text{sous gradient de } h \text{ en } 0}{\underbrace{\partial h(0)}}\neq\emptyset$

$\partial h(0)$ est non vide car h cvx

$0\in\mathrm{Dom}(f)$

Def : On appelle le sous-gradient de h en y l'ensemble $x\in\mathbb{R}^d$ tq

$$h(y)+\langle x,u-y\rangle\leq h(u)\quad\forall u$$

On note $\partial h(y)$

b) Vérifier que $\forall x,u$, $f(x)+g(x+u)-\langle\bar{x},u\rangle\geq h(0)$

En 3-a, On a vu que $h(0)+\langle\bar{x},u\rangle\leq h(u)$

Par déf de h, $h(u)\leq f(x)+g(x+u)\quad\forall(x,u)$

$$\Rightarrow h(0)+\langle\bar{x},u\rangle\leq f(x)+g(x+u)\quad\forall(x,u)$$

c) Déduire (1) pour $x_0=0$

Prouver (1) en $x_0=0$, en utilisant la Q1, ça revient à prouver $(f+g)^*(0)\geq\inf_z(f^*(z)+g^*(-z))$

Mais $(f+g)^*(0) = \sup_y(\langle y, \underset{0}{\underline{0}}\rangle - (f+g)(y)) = -\inf_y((f+g)(y)) = -h(0)$

← z sur $\mathbb{R}$ donc on peut prendre $z \mapsto -z$

On doit prouver que $-h(0) \geq \inf_z \{f^*(z) + g^*(-z)\}$

On sait que Q3-b, $-h(0) \geq \langle \bar{x}, u\rangle - f(x) - g(x+u)$ $\forall (x,u)$

$$\geq \langle \bar{x}, x+u\rangle - g(x+u) \underbrace{- \langle \bar{x}, x\rangle}_{+\langle -\bar{x}, x\rangle} - f(x) \qquad \forall (x,u)$$

On va faire sup sur u puis sur x.

$$\geq \underbrace{\sup_u(\langle \bar{x}, x+u\rangle - g(x+u) + \langle -\bar{x}, x\rangle - f(x))}_{\sup_{u'}(\langle \bar{x}, u'\rangle - g(u') + \langle -\bar{x}, x\rangle - f(x))}$$

$$\geq \underbrace{\sup_{u'}(\langle \bar{x}, u'\rangle - g(u'))}_{g^*(\bar{x})} + \underbrace{\sup_x(\langle -\bar{x}, x\rangle - f(x))}_{f^*(-\bar{x})}$$

$$\Rightarrow -h(0) \geq g^*(\bar{x}) + f^*(-\bar{x}) \geq \inf_z \{g^*(z) + f^*(-z)\}$$

On a bien (1) pour $x_0 = 0$.

4) Soit $z \in \mathbb{R}^d$, et $\ell(u) := g(u) - \langle z, u\rangle$

a) Vérifier que $\ell^*(x) = g^*(x+z)$ et $(f+\ell)^*(0) = (f+g)^*(z)$

$$\ell^*(x) = \sup_y(\langle x, y\rangle - \ell(y)) = \sup_y(\langle x, y\rangle - g(y) + \langle z, y\rangle)$$
$$= \sup_y(\langle x+z, y\rangle - g(y)) = g^*(x+z)$$

$$(f+\ell)^*(0) = \sup_y(\langle y, 0\rangle - (f+\ell)(y)) = \sup_y(-f(y) - g(y) + \langle z, y\rangle)$$
$$= (f+g)^*(z)$$

b) Déduire (1) dans le cas général

f et ℓ sont propres et $Dom(f) \cap Dom(\ell) \neq \emptyset$ car $Dom(f) \cap Dom(g) \subset Dom(f) \cap Dom(\ell)$

On peut appliquer le résultat à $x_0 = 0$

$$\underbrace{(f+\ell)^*(0)}_{(f+g)^*(z)} = \inf_{z'} \{f^*(z') + \underbrace{\ell^*(-z')}_{g^*(z-z')}\}$$

$$\Rightarrow (f+g)^*(z) = \inf_{z'} \{f^*(z') + g^*(z-z')\}$$
$$= (f^* \square g^*)(-z) \qquad \square.$$

Partie II :

Soit $d, \ell \geq 1$. On a deux distribut° de proba $\nu \in \mathcal{P}(\{1, \dots, d\})$ et $\mu \in \mathcal{P}(\{1, \dots, \ell\})$.

ν et μ sont des vecteurs de $\mathbb{R}^d$ et $\mathbb{R}^\ell$ resp.

On note $\Gamma(d,\ell) \subset \mathbb{R}^d \times \mathbb{R}^\ell$ l'ensemble des $(\gamma(i,j))_{1 \leq i \leq d, 1 \leq j \leq \ell}$ tq $\gamma \geq 0$, $\sum_{j=1}^{\ell} \gamma(i,j) = \nu(j)$ et $\sum_{i=1}^{d} \gamma(i,j) = \mu(i)$

On considère $(c(i,j))_{\substack{1 \leq i \leq d \\ 1 \leq j \leq \ell}} \in \mathbb{R}^d \times \mathbb{R}^\ell$

On veut mq $\inf_{\gamma \in \Gamma(d,\ell)} \{\sum_{i=1}^{d} \sum_{j=1}^{\ell} \gamma(i,j) c(i,j)\} = \sup \{\sum_{i=1}^{d} \phi(i)\nu(i) + \sum_{j=1}^{\ell} \psi(j)\mu(j) : \phi(i) + \psi(j) \leq c(i,j) \ \forall (i,j)\}$

On introduit $f(\gamma) = \begin{cases} \sum_{i=1}^{d} \sum_{j=1}^{\ell} \gamma(i,j) c(i,j) & \text{si } \gamma \geq 0 \\ \infty & \text{sinon} \end{cases}$

$$g(\gamma) = \begin{cases} 0 & \text{si } \underbrace{\sum_{j=1}^{\ell} \gamma(i,j) = \nu(i) \text{ et } \sum_{i=1}^{d} \gamma(i,j) = \mu(j)}_{**} \\ +\infty & \text{sinon} \end{cases}$$

Rq : $(f+g)^*(0) = \sup_\gamma(\langle 0, \gamma\rangle - (f+g)(\gamma)) = \sup_{\gamma \text{ vérifie } **}(-(f+\cancel{g})(\gamma)) = \sup_{\gamma **}(-f(\gamma)) = \sup_{\substack{\gamma \Rightarrow ** \\ \gamma \geq 0}}(-f(\gamma)) = -\inf_{\substack{\gamma \Rightarrow ** \\ \gamma \geq 0}} \sum_i \sum_j \gamma(i,j) c(i,j)$

1) Mq $f^*(\pi) = 0$ si $\pi(i,j) \leq c(i,j)$ et $f^*(\pi) = +\infty$ sinon

$$f^*(\pi) = \sup_{\gamma} (\langle \pi, \gamma \rangle - f(\gamma))$$
$$= \sup_{\gamma \geq 0} (\langle \pi, \gamma \rangle - \langle c, \gamma \rangle)$$
$$= \sup_{\gamma \geq 0} (\langle \pi - c, \gamma \rangle)$$
$$= \begin{cases} 0 & \text{si } \pi \leq c \\ +\infty & \text{sinon} \end{cases}$$

2) Calculer $g^*(\pi)$ $\forall \pi \in \mathbb{R}^d \times \mathbb{R}^e$

$$g^*(\pi) = \sup_{\gamma} (\langle \pi, \gamma \rangle - g(\gamma))$$
$$= \sup_{\gamma \Rightarrow xx} \langle \pi, \gamma \rangle = \sup_{\gamma \Rightarrow xx} \sum_{i,j} \pi(i,j) \gamma(i,j)$$

3-a et 3-b on Mq si $\pi(i,j) = \phi(i) + \psi(j)$ avec $\phi + \psi \leq c$ alors $g^*(\pi) = \sum_i \phi(i) \nu(i) + \sum_j \psi(j) \mu(j)$ sinon $g^*(\pi) = \infty$.

Pour conclure, il suffit d'appliquer (1) en $x_0 = 0$, $(f+g)^*(0) = \inf_{\pi} \{ f^*(\pi) + g^*(-\pi) \}$

$$= \inf_{\pi} \Big\{ 0 \, \mathbb{1}_{(\pi \leq c)} + \infty \, \mathbb{1}_{(\pi \leq c)^c} + \Big(\sum_i \phi(i)\nu(i) + \sum_j \psi(j)\mu(j)\Big) \mathbb{1}_{\substack{\phi+\psi \leq c \\ \phi+\psi = -\pi}} + \infty \, \mathbb{1}_{\text{sinon}} \Big)$$

$$= - \sup_{\pi} \Big\{ \sum_i \phi(i)\nu(i) + \sum_j \psi(j)\mu(j) \;\Big|\; \psi + \phi \leq c \Big\}$$