Penalty and Smoothing Methods For Convex Semi-Infinite Programming

Alfred Auslender

Institut Camille Jordan

University of Lyon I, France

Talk based on a joint work with Miguel A.Goberna, and Marco A.Lopez

Introduction

Min-Max Convex Semi-Infinite Programming (CSIP)

•

Introduction

Min-Max Convex Semi-Infinite Programming (CSIP)
$T_1,\ T_2$ compact metric spaces, $Q \subset \mathbb{R}^n$, convex closed.

Introduction

Min-Max Convex Semi-Infinite Programming (CSIP)
$T_1,\ T_2$ compact metric spaces, $Q \subset \mathbb{R}^n$, convex closed.
$f : T_1 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \quad g : T_2 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$

Introduction

Min-Max Convex Semi-Infinite Programming (CSIP)
$T_1, \; T_2$ compact metric spaces, $Q \subset \mathbb{R}^n$, convex closed.
$f : T_1 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \quad g : T_2 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$
finite, continuous on $T_1 \times Q, \; T_2 \times Q$,

Introduction

Min-Max Convex Semi-Infinite Programming (CSIP)
$T_1,\ T_2$ compact metric spaces, $Q \subset \mathbb{R}^n$, convex closed.
$f : T_1 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \quad g : T_2 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$
finite, continuous on $T_1 \times Q,\ T_2 \times Q$,
$\forall t \ \ f(t, \cdot),\ g(t, \cdot)$: lsc, convex on $\mathbb{R}^n$.

Introduction

Min-Max Convex Semi-Infinite Programming (CSIP)
$T_1,\ T_2$ compact metric spaces, $Q \subset \mathbb{R}^n$, convex closed.
$f : T_1 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \quad g : T_2 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$
finite, continuous on $T_1 \times Q,\ T_2 \times Q$,
$\forall t \ \ f(t,\cdot),\ g(t,\cdot)$: lsc, convex on $\mathbb{R}^n$.

$$(P) \quad m = \inf\{F(x) |\ x \in C\},$$

Introduction

Min-Max Convex Semi-Infinite Programming (CSIP)
$T_1,\ T_2$ compact metric spaces, $Q \subset \mathbb{R}^n$, convex closed.
$f : T_1 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \quad g : T_2 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$
finite, continuous on $T_1 \times Q,\ T_2 \times Q$,
$\forall t \ \ f(t, \cdot),\ g(t, \cdot)$: lsc, convex on $\mathbb{R}^n$.

$$(P) \quad m = \inf\{F(x)|\ x \in C\},$$

$$F(x) = \sup\{f(t,x)|\ t \in T_1\}, \quad G(x) = \sup\{g(t,x)|\ t \in T_2\},$$

Introduction

Min-Max Convex Semi-Infinite Programming (CSIP)
$T_1,\ T_2$ compact metric spaces, $Q \subset \mathbb{R}^n$, convex closed.
$f: T_1 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \quad g: T_2 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$
finite, continuous on $T_1 \times Q,\ T_2 \times Q$,
$\forall t \ \ f(t,\cdot),\ g(t,\cdot)$: lsc, convex on $\mathbb{R}^n$.

$$(P) \quad m = \inf\{F(x) |\ x \in C\},$$

$F(x) = \sup\{f(t,x) |\ t \in T_1\}, \quad G(x) = \sup\{g(t,x) |\ t \in T_2\}$,
$C = Q \cap D$, and $D = \{x : G(x) \leq 0\}$.

Introduction

Min-Max Convex Semi-Infinite Programming (CSIP)
$T_1,\ T_2$ compact metric spaces, $Q \subset \mathbb{R}^n$, convex closed.
$f : T_1 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}, \quad g : T_2 \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$
finite, continuous on $T_1 \times Q,\ T_2 \times Q$,
$\forall t \ \ f(t,\cdot),\ g(t,\cdot)$: lsc, convex on $\mathbb{R}^n$.

$$(P) \quad m = \inf\{F(x) |\ x \in C\},$$

$F(x) = \sup\{f(t,x) |\ t \in T_1\}, \quad G(x) = \sup\{g(t,x) |\ t \in T_2\}$,
$C = Q \cap D$, and $D = \{x : G(x) \leq 0\}$.
Q is "simple: $Q = \mathbb{R}^n$, a box, a ball, positive orthant...

Remez-type methods:outer methods

T_1^k and T_2^k: finite subsets of T_1 and T_2

Remez-type methods:outer methods

T_1^k and T_2^k: finite subsets of T_1 and T_2

$F^k(x) = \sup\{f(t,x)|\ t \in T_1^k\}, \quad D^k = \{x : g(t,x) \leq 0\ \forall t \in T_2^k\}$

Remez-type methods:outer methods

T_1^k and T_2^k: finite subsets of T_1 and T_2

$F^k(x) = \sup\{f(t,x)|\ t \in T_1^k\}, \quad D^k = \{x : g(t,x) \leq 0\ \forall t \in T_2^k\}$

$C^k = Q \cap D^k$

Remez-type methods:outer methods

T_1^k and T_2^k: finite subsets of T_1 and T_2

$F^k(x) = \sup\{f(t,x)|\ t \in T_1^k\}, \quad D^k = \{x : g(t,x) \leq 0\ \forall t \in T_2^k\}$

$C^k = Q \cap D^k$

Initialization: Set $k = 0$ and start with $T_1^0,\ T_2^0$.

Remez-type methods:outer methods

T_1^k and T_2^k: finite subsets of T_1 and T_2

$F^k(x) = \sup\{f(t,x)|\ t \in T_1^k\}, \quad D^k = \{x : g(t,x) \leq 0\ \forall t \in T_2^k\}$

$C^k = Q \cap D^k$

Initialization: Set $k = 0$ and start with $T_1^0,\ T_2^0$.

Step 1 Compute $x^k \in \text{argmin}\{F^k(x)|\ x \in C^k\}$.

Remez-type methods:outer methods

T_1^k and T_2^k: finite subsets of T_1 and T_2
$F^k(x) = \sup\{f(t,x)|\ t \in T_1^k\}, \quad D^k = \{x : g(t,x) \leq 0\ \forall t \in T_2^k\}$
$C^k = Q \cap D^k$
Initialization: Set $k = 0$ and start with $T_1^0,\ T_2^0$.
Step 1 Compute $x^k \in \text{argmin}\{F^k(x)|\ x \in C^k\}$.
Step 2 Compute

$$t_1^{k+1} \in \text{argmax}\{f(t,x^k)|t \in T_1\},\ t_2^{k+1} \in \text{argmax}\{g(t,x^k)|t \in T_2\}$$

Remez-type methods:outer methods

T_1^k and T_2^k: finite subsets of T_1 and T_2
$F^k(x) = \sup\{f(t,x)|\ t \in T_1^k\}, \quad D^k = \{x : g(t,x) \leq 0\ \forall t \in T_2^k\}$
$C^k = Q \cap D^k$
Initialization: Set $k = 0$ and start with $T_1^0,\ T_2^0$.
Step 1 Compute $x^k \in \text{argmin}\{F^k(x)|\ x \in C^k\}$.
Step 2 Compute

$$t_1^{k+1} \in \text{argmax}\{f(t,x^k)|t \in T_1\},\ t_2^{k+1} \in \text{argmax}\{g(t,x^k)|t \in T_2\}$$

Step 3 Choose, for $i = 1, 2,\ T_i^{k+1} \subset T_i :\ t_i^{k+1} \in T_i^{k+1}$.

Remez-type methods:outer methods

T_1^k and T_2^k: finite subsets of T_1 and T_2

$F^k(x) = \sup\{f(t,x)|\ t \in T_1^k\}, \quad D^k = \{x : g(t,x) \leq 0\ \forall t \in T_2^k\}$

$C^k = Q \cap D^k$

Initialization: Set $k = 0$ and start with $T_1^0,\ T_2^0$.

Step 1 Compute $x^k \in$ argmin$\{F^k(x)|\ x \in C^k\}$.

Step 2 Compute

$$t_1^{k+1} \in \text{argmax}\{f(t,x^k)|t \in T_1\},\ t_2^{k+1} \in \text{argmax}\{g(t,x^k)|t \in T_2\}$$

Step 3 Choose, for $i = 1,2,\ T_i^{k+1} \subset T_i :\ t_i^{k+1} \in T_i^{k+1}$.

Set $k \leftarrow k+1$; go to Step 1.

Step1

- How to choose a "good" approximation $(\tilde{P}^k)$ of

$$(P^k) \quad m_k = \inf\{F^k(x) |\, x \in D^k \cap Q\}$$

- Q polyhedral, $f(t,\cdot),\ g(t,\cdot) \Rightarrow (P^k)$ linear subproblem $\Rightarrow$the simplex dual method.
- Drawbacks: cardinality $|T_i^k|$ becomes to big
- $\Rightarrow$ dropping rules: large literature (survey of Reemtsen-Gorner (1998)).
- Cutting plane methods of Cheney-Goldstein , Kelley , Veinott or Elzinga-Moore,...
- Applied to LSIP: quasi similar properties and drawbacks.
- $\Rightarrow$ dropping rules under uniform strict convexity assumptions.

Another type of approximation for (P^k)

- $$(\tilde{P}^k) \quad \widetilde{m}_k = \inf\{\tilde{F}^k(x) + \tilde{G}^k(x) |\, x \in Q\}.$$
  $\tilde{F}^k$ smoothing F^k, $\tilde{G}^k$ penalization of D^k,
- $\Rightarrow$ data which define $(\tilde{P}^k)$ are $\mathcal{C}^1$.
- Many ways to smooth F^k: the most important:
- based on the smoothing of $\max\{\lambda_i : i = 1 \ldots m\}$.
- $$\tilde{F}^k(x) := \frac{\log \sum_{t \in T_1^k} \exp\left(f(t,x)p\right)}{p} \quad \text{with } p = [\log \left|T_1^k\right|]^2$$
- Finite minimax problem :Bertsekas , Ben-Tal and Teboulle Alvarez, Nesterov, ...
- CSIP: Polak, Royset and Womersley , by Sheu-Wu , Sheu-Lin, Fang-Wu..

Penalization of the set D^k

- Classical case with a finite number of inequalities: Auslender-Cominetti-Haddou(MOR 1997),
- $\theta : \mathbb{R} \to \mathbb{R}_+$ $\mathcal{C}^1$, convex, nondecreasing, + some additional properties.
-

$$\tilde{G}^k(x) := \frac{\gamma_k}{|T_2^k|} \frac{\sum_{t \in T_2^k} \theta(g(t,x)\delta_k)}{\delta_k}$$

  with appropriated sequences of positive scalars $\{\gamma_k\}, \ \{\delta_k\}$
- 3 others contexts in SIP with penalty functions:
- based local reduction methods: Reemsten and Gorner'survey,
- coupled with adaptive grid methods (ex:Kaplan-Tikhachke (1992))

Integral methods

- Advantage: Avoid global optimization in Step 2, via integral which convexify the approximate problem.
- At each step k compute a solution of

$$(P^k_{ips}) \quad m^k_{ips} = \inf\{\tilde{F}^k(x) + \tilde{G}^k(x) \; | x \in Q\}$$

,

- $$\tilde{F}^k(x) = \frac{1}{p_k}\log(\int_{T_1} \exp(f(t,x)p_k)dt), \quad \tilde{G}^k(x) = \gamma_k \int_{T_2} \frac{\theta(g(t,x)\delta_k)}{\delta_k} dt,$$

- whereγ_k, δ_k, p_k are adjusted for obtaining convergence.
- Auslender(1970), Conn-Gould (1987), Teo-Goh (1987), Teboulle (1990) , Teo-Rehbock-Jennings (1993), , Polak, Higgins and Mayne (1992), Schattler (1996), Lin-Fang-Wu (1998) ...

Some basic tools: Asymptotic functions

- Definition (Steinitz 1913)Let $Q \subset \mathbb{R}^n$, its asymptotic cone:

$$Q_\infty = \{y \ : \exists t_k \to +\infty,\ x_k \in Q \quad \text{with } y = \lim_{k\to\infty} \frac{x_k}{t_k}\}.$$

Some basic tools: Asymptotic functions

- Definition (Steinitz 1913)Let $Q \subset \mathbb{R}^n$, its asymptotic cone:

$$Q_\infty = \{y \ : \exists t_k \to +\infty,\ x_k \in Q \quad \text{with } y = \lim_{k\to\infty} \frac{x_k}{t_k}\}.$$

- When Q is closed and convex, Q_∞ coincides with the recession cone $0^+(Q)$ defined by

$$0^+(Q) := \{y \ : x + \lambda y \in Q\ \forall \lambda > 0,\ \forall x \in Q\}.$$

Some basic tools: Asymptotic functions

- Definition (Steinitz 1913)Let $Q \subset \mathbb{R}^n$, its asymptotic cone:

$$Q_\infty = \{y \ : \exists t_k \to +\infty, \ x_k \in Q \quad \text{with } y = \lim_{k\to\infty} \frac{x_k}{t_k}\}.$$

- When Q is closed and convex, Q_∞ coincides with the recession cone $0^+(Q)$ defined by

$$0^+(Q) := \{y \ : x + \lambda y \in Q \ \forall \lambda > 0, \ \forall x \in Q\}.$$

- Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ lsc , proper, then the asymptotic function f_∞ is defined by

$$\text{epi} f_\infty = (\text{epi} f)_\infty.$$

Some basic tools: Asymptotic functions

- Definition (Steinitz 1913)Let $Q \subset \mathbb{R}^n$, its asymptotic cone:

$$Q_\infty = \{y \ : \exists t_k \to +\infty,\ x_k \in Q \quad \text{with } y = \lim_{k\to\infty} \frac{x_k}{t_k}\}.$$

- When Q is closed and convex, Q_∞ coincides with the recession cone $0^+(Q)$ defined by

$$0^+(Q) := \{y \ : x + \lambda y \in Q \ \forall \lambda > 0,\ \forall x \in Q\}.$$

- Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ lsc , proper, then the asymptotic function f_∞ is defined by

$$\text{epi} f_\infty = (\text{epi} f)_\infty.$$

- Theorem (Dedieu 1977)

$$f_\infty(y) = \inf\{\liminf_{k\to+\infty} \frac{f(x_k t_k)}{t_k} :\ t_k \to +\infty\ x_k \to y\},$$

where $\{t_k\}$ and $\{x_k\}$ are sequences in $\mathbb{R}$ and in $\mathbb{R}^n$.

Some basic tools: Asymptotic functions

- Definition (Steinitz 1913)Let $Q \subset \mathbb{R}^n$, its asymptotic cone:

$$Q_\infty = \{y \ : \exists t_k \to +\infty, \ x_k \in Q \quad \text{with } y = \lim_{k\to\infty} \frac{x_k}{t_k}\}.$$

- When Q is closed and convex, Q_∞ coincides with the recession cone $0^+(Q)$ defined by

$$0^+(Q) := \{y \ : x + \lambda y \in Q \ \forall \lambda > 0, \ \forall x \in Q\}.$$

- Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ lsc , proper, then the asymptotic function f_∞ is defined by

$$\text{epi} f_\infty = (\text{epi} f)_\infty.$$

- Theorem (Dedieu 1977)

$$f_\infty(y) = \inf\{\liminf_{k\to+\infty} \frac{f(x_k t_k)}{t_k} : \ t_k \to +\infty \ x_k \to y\},$$

where $\{t_k\}$ and $\{x_k\}$ are sequences in $\mathbb{R}$ and in $\mathbb{R}^n$.

- For a lsc convex proper function it coincides with the recession

Asymptotic function of a composite function

- $$\mathcal{F} = \left\{ \begin{array}{c} \theta : \mathbb{R} \to \mathbb{R}_+ \text{ convex, nondecreasing, nonconstant,} \\ \mathcal{C}^1, \text{ and such that } \lim_{u \to -\infty} \theta(u) = 0 \end{array} \right\}.$$

Asymptotic function of a composite function

- $$\mathcal{F} = \left\{ \begin{array}{c} \theta : \mathbb{R} \longrightarrow \mathbb{R}_+ \text{ convex, nondecreasing, nonconstant,} \\ \mathcal{C}^1, \text{ and such that } \lim_{u \to -\infty} \theta(u) = 0 \end{array} \right\}.$$

- Proposition (Auslender, Cominetti, Haddou (SIOPT, 1999)) Let $\theta \in \mathcal{F}$, and let f be a proper lsc convex function, and consider the composite function

$$g(x) = \begin{cases} \theta(f(x)), & \text{if } x \in dom\, f, \\ +\infty, & \text{otherwise.} \end{cases}$$

Then g is a proper lsc convex function, and we have

$$g_\infty(d) = \begin{cases} \theta_\infty(f_\infty(d)), & \text{if } d \in dom\, f_\infty, \\ +\infty, & \text{otherwise.} \end{cases}$$

The penalty functions

- We will consider two subsets of $\mathcal{F}$:

$$\mathcal{F}_1 = \{\theta \in \mathcal{F} : \ \theta_\infty(1) < +\infty\} \text{ and } \mathcal{F}_2 = \{\theta \in \mathcal{F} : \ \theta_\infty(1) = +\infty\}.$$

The penalty functions

- We will consider two subsets of $\mathcal{F}$:

$$\mathcal{F}_1 = \{\theta \in \mathcal{F} : \ \theta_\infty(1) < +\infty\} \text{ and } \mathcal{F}_2 = \{\theta \in \mathcal{F} : \ \theta_\infty(1) = +\infty\}.$$

- Chen and Mangasarian provided a systematic way to generate functions which belong to $\mathcal{F}_1$. These are smooth approximations of the function $u^+ = max(u, 0)$. Specific cases of interest

$$\theta_1(u) = \log(1 + \exp(u)), \quad \theta_2(u) = 2^{-1}(u + \sqrt{u^2 + 4}),$$

$$\theta_3(u) = \begin{cases} 0, & u \leq -\frac{1}{2}, \\ \frac{1}{2}\left(u + \frac{1}{2}\right)^2, & -\frac{1}{2} < u < \frac{1}{2}, \\ u, & u \geq \frac{1}{2}, \end{cases}$$

The penalty functions

- We will consider two subsets of $\mathcal{F}$:

$$\mathcal{F}_1 = \{\theta \in \mathcal{F} : \ \theta_\infty(1) < +\infty\} \text{ and } \mathcal{F}_2 = \{\theta \in \mathcal{F} : \ \theta_\infty(1) = +\infty\}.$$

- Chen and Mangasarian provided a systematic way to generate functions which belong to $\mathcal{F}_1$. These are smooth approximations of the function $u^+ = max(u, 0)$. Specific cases of interest

$$\theta_1(u) = \log(1 + \exp(u)), \quad \theta_2(u) = 2^{-1}(u + \sqrt{u^2 + 4}),$$

$$\theta_3(u) = \begin{cases} 0, & u \leq -\frac{1}{2}, \\ \frac{1}{2}\left(u + \frac{1}{2}\right)^2, & -\frac{1}{2} < u < \frac{1}{2}, \\ u, & u \geq \frac{1}{2}, \end{cases}$$

- the following classical belong to $\mathcal{F}_2$

$$\theta_4(u) = \frac{1}{2}(u^+)^2, \ \theta_5(u) = (u^+)^3, \ \theta_6(u) = \exp(u).$$

The penalty functions

- We will consider two subsets of $\mathcal{F}$:

$$\mathcal{F}_1 = \{\theta \in \mathcal{F} : \ \theta_\infty(1) < +\infty\} \text{ and } \mathcal{F}_2 = \{\theta \in \mathcal{F} : \ \theta_\infty(1) = +\infty\}.$$

- Chen and Mangasarian provided a systematic way to generate functions which belong to $\mathcal{F}_1$. These are smooth approximations of the function $u^+ = max(u, 0)$. Specific cases of interest

$$\theta_1(u) = \log(1 + \exp(u)), \quad \theta_2(u) = 2^{-1}(u + \sqrt{u^2 + 4}),$$

$$\theta_3(u) = \begin{cases} 0, & u \leq -\frac{1}{2}, \\ \frac{1}{2}\left(u + \frac{1}{2}\right)^2, & -\frac{1}{2} < u < \frac{1}{2}, \\ u, & u \geq \frac{1}{2}, \end{cases}$$

- the following classical belong to $\mathcal{F}_2$

$$\theta_4(u) = \frac{1}{2}(u^+)^2, \ \theta_5(u) = (u^+)^3, \ \theta_6(u) = \exp(u).$$

Remez-type method with penalty methods

The basic assumption (A_1) : F is level bounded on C.

Remez-type method with penalty methods

The basic assumption (A_1) : F is level bounded on C.

$F^k(x) = \sup\{f(t,x)|\ t \in T_1^k\},\ C^k = \{x \in Q : g(t,x) \leq 0\ \forall t \in T_2^k\}$

Remez-type method with penalty methods

The basic assumption $(A_1): \; F$ is level bounded on C.

$F^k(x) = \sup\{f(t,x) |\; t \in T_1^k\}, \; C^k = \{x \in Q : g(t,x) \leq 0 \; \forall t \in T_2^k\}$

Lemma: Assume that (A_1) holds. Then, $\exists$ finite non-empty subsets $T_1^0 \subset T_1, \; T_2^0 \subset T_2 : \; F^0$ is level bounded on C^0.

Remez-type method with penalty methods

The basic assumption $(A_1):\ F$ is level bounded on C.
$F^k(x) = \sup\{f(t,x) |\ t \in T_1^k\},\ C^k = \{x \in Q : g(t,x) \le 0\ \forall t \in T_2^k\}$
Lemma: Assume that (A_1) holds. Then, $\exists$ finite non-empty subsets $T_1^0 \subset T_1,\ T_2^0 \subset T_2:\ F^0$ is level bounded on C^0.

$$\tilde{F}^k(x) := \frac{\log \sum_{t\in T_1^k} \exp\left(f(t,x)p_k\right)}{p_k},\ \text{with}\ p_k = [\log(|T_1^0| + k)]^2$$

Remez-type method with penalty methods

The basic assumption $(A_1):\ F$ is level bounded on C.
$F^k(x) = \sup\{f(t,x)|\ t \in T_1^k\},\ C^k = \{x \in Q : g(t,x) \leq 0\ \forall t \in T_2^k\}$
Lemma: Assume that (A_1) holds. Then, $\exists$ finite non-empty subsets $T_1^0 \subset T_1,\ T_2^0 \subset T_2:\ F^0$ is level bounded on C^0.

$$\tilde{F}^k(x) := \frac{\log \sum_{t \in T_1^k} \exp\left(f(t,x)p_k\right)}{p_k},\ \text{with } p_k = [\log(|T_1^0| + k)]^2$$

Let $\theta \in \mathcal{F},\ \delta_k > 0,\ \gamma_k > 0,\ \epsilon_k > 0:\ \lim_{k\to\infty} \epsilon_k = 0.$

Remez-type method with penalty methods

The basic assumption (A_1) : F is level bounded on C.
$F^k(x) = \sup\{f(t,x) | \, t \in T_1^k\}, \; C^k = \{x \in Q : g(t,x) \le 0 \; \forall t \in T_2^k\}$
Lemma: Assume that (A_1) holds. Then, $\exists$ finite non-empty subsets $T_1^0 \subset T_1, \; T_2^0 \subset T_2$: F^0 is level bounded on C^0.

$$\tilde{F}^k(x) := \frac{\log \sum_{t \in T_1^k} \exp\left(f(t,x) p_k\right)}{p_k}, \text{ with } p_k = [\log(|T_1^0| + k)]^2$$

Let $\theta \in \mathcal{F}, \; \delta_k > 0, \; \gamma_k > 0, \; \epsilon_k > 0$: $\lim_{k \to \infty} \epsilon_k = 0$.

$$\tilde{G}^k(x) := \frac{\gamma_k}{|T_2^k|} \sum_{t \in T_2^k} \frac{\theta(g(t,x)\delta_k)}{\delta_k}, \qquad \tilde{H}^k(x) := \tilde{F}^k(x) + \tilde{G}^k(x) + \epsilon_k ||x||^2$$

Remez-type method with penalty methods

The basic assumption (A_1) : F is level bounded on C.
$F^k(x) = \sup\{f(t,x) |\ t \in T_1^k\},\ C^k = \{x \in Q : g(t,x) \leq 0\ \forall t \in T_2^k\}$
Lemma: Assume that (A_1) holds. Then, $\exists$ finite non-empty subsets $T_1^0 \subset T_1,\ T_2^0 \subset T_2$: F^0 is level bounded on C^0.

$$\tilde{F}^k(x) := \frac{\log \sum_{t \in T_1^k} \exp\left(f(t,x) p_k\right)}{p_k}, \text{ with } p_k = [\log(|T_1^0| + k)]^2$$

Let $\theta \in \mathcal{F},\ \delta_k > 0,\ \gamma_k > 0,\ \epsilon_k > 0$: $\lim_{k \to \infty} \epsilon_k = 0$.

$$\tilde{G}^k(x) := \frac{\gamma_k}{|T_2^k|} \sum_{t \in T_2^k} \frac{\theta(g(t,x)\delta_k)}{\delta_k}, \qquad \tilde{H}^k(x) := \tilde{F}^k(x) + \tilde{G}^k(x) + \epsilon_k ||x||^2$$

$(\tilde{P}^k)$: Find an ϵ_k minimizer of $\tilde{H}^k$ over Q.

Remez penalty smoothing algorithm-RPSALG

Initialization: Let $\mu_k \geq 0 \ \ \forall k$ and $\lim_{k \to +\infty} \mu_k = 0$. Set $k = 0$ start with T_i^0 defined in Lemma 1.

Remez penalty smoothing algorithm-RPSALG

Initialization: Let $\mu_k \geq 0 \;\; \forall k$ and $\lim_{k \to +\infty} \mu_k = 0$. Set $k = 0$ start with T_i^0 defined in Lemma 1.

Step 1 Compute $x^k \in Q : \;\; \tilde{H}^k(x^k) \leq \min\{\tilde{H}^k(x) | x \in Q\} + \epsilon_k$

Remez penalty smoothing algorithm-RPSALG

Initialization: Let $\mu_k \geq 0 \ \forall k$ and $\lim_{k\rightarrow+\infty} \mu_k = 0$. Set $k = 0$ start with T_i^0 defined in Lemma 1.

Step 1 Compute $x^k \in Q : \ \tilde{H}^k(x^k) \leq \min\{\tilde{H}^k(x) | x \in Q\} + \epsilon_k$

Step 2 Compute $t_1^{k+1} \in T_1$ and $t_2^{k+1} \in T_2$:

$$f(t_1^{k+1}, x^k) \geq \max\{f(t, x^k) \mid t \in T_1\} - \mu_k,$$

$$g(t_2^{k+1}, x^k) \geq \max\{g(t, x^k) \mid t \in T_2\} - \mu_k,$$

Remez penalty smoothing algorithm-RPSALG

Initialization: Let $\mu_k \geq 0 \;\; \forall k$ and $\lim_{k\to+\infty}\mu_k = 0$. Set $k = 0$ start with T_i^0 defined in Lemma 1.

Step 1 Compute $x^k \in Q : \; \tilde{H}^k(x^k) \leq \min\{\tilde{H}^k(x) | x \in Q\} + \epsilon_k$

Step 2 Compute $t_1^{k+1} \in T_1$ and $t_2^{k+1} \in T_2$:

$$f(t_1^{k+1}, x^k) \geq \max\{f(t, x^k) \mid t \in T_1\} - \mu_k,$$

$$g(t_2^{k+1}, x^k) \geq \max\{g(t, x^k) \mid t \in T_2\} - \mu_k,$$

Step 3 Set $T_i^{k+1} = T_i^k \cup \left\{t_i^{k+1}\right\}, \; i = 1, 2.$

Remez penalty smoothing algorithm-RPSALG

Initialization: Let $\mu_k \geq 0 \;\; \forall k$ and $\lim_{k \to +\infty} \mu_k = 0$. Set $k = 0$ start with T_i^0 defined in Lemma 1.

Step 1 Compute $x^k \in Q : \;\; \tilde{H}^k(x^k) \leq \min\{\tilde{H}^k(x) | x \in Q\} + \epsilon_k$

Step 2 Compute $t_1^{k+1} \in T_1$ and $t_2^{k+1} \in T_2$:

$$f(t_1^{k+1}, x^k) \geq \max\{f(t, x^k) \mid t \in T_1\} - \mu_k,$$

$$g(t_2^{k+1}, x^k) \geq \max\{g(t, x^k) \mid t \in T_2\} - \mu_k,$$

Step 3 Set $T_i^{k+1} = T_i^k \cup \left\{ t_i^{k+1} \right\}, \;\; i = 1, 2.$

Set $k \leftarrow k + 1$; go to Step 1.

Remez penalty smoothing algorithm-RPSALG

Initialization: Let $\mu_k \geq 0 \ \forall k$ and $\lim_{k \to +\infty} \mu_k = 0$. Set $k = 0$ start with T_i^0 defined in Lemma 1.

Step 1 Compute $x^k \in Q: \ \tilde{H}^k(x^k) \leq \min\{\tilde{H}^k(x) | x \in Q\} + \epsilon_k$

Step 2 Compute $t_1^{k+1} \in T_1$ and $t_2^{k+1} \in T_2$:

$$f(t_1^{k+1}, x^k) \geq \max\{f(t, x^k) \mid t \in T_1\} - \mu_k,$$

$$g(t_2^{k+1}, x^k) \geq \max\{g(t, x^k) \mid t \in T_2\} - \mu_k,$$

Step 3 Set $T_i^{k+1} = T_i^k \cup \left\{t_i^{k+1}\right\}, \ i = 1, 2.$

Set $k \leftarrow k + 1$; go to Step 1.

Implementable stopping rule. Ex when $Q = \mathbb{R}^n$:

$$||\nabla \tilde{H}^k(x^k)|| \leq \sqrt{2}\epsilon_k.$$

Convergence Theorem

- 3 types of conditions on $(\theta, \gamma_k, \delta_k,)$ $(s_k := |T_2^0| + k)$
- (a) $\theta \in \mathcal{F}_1$, $\lim_{k\to\infty} \frac{\gamma_k}{\delta_k} = 0$, and $\lim_{k\to\infty} \frac{\gamma_k}{s_k} = +\infty$.
- (b) $\theta \in \mathcal{F}_2$, $\lim_{k\to\infty} \frac{\gamma_k}{\delta_k} = 0$, and $\frac{\gamma_k}{s_k} > \varepsilon \, \forall k$, for a certain $\varepsilon > 0$.
- (c) $\theta \in \mathcal{F}_2$, $\lim_{k\to\infty} \delta_k = +\infty$, $\frac{\gamma_k}{s_k} > \varepsilon \, \forall k$ for a certain $\varepsilon > 0$, $\left\{ \frac{\gamma_k}{\delta_k} \right\}$ is bounded, and either $\theta(0) = 0$ or Slater condition holds.
- Theorem 1 Assume that (A_1) holds. If $(\theta, \{\gamma_k\}, \{\delta_k\})$ satisfies at least one of the conditions (a), (b), (c) then the sequence built by RPSALG is bounded and each limit point of this sequence is an optimal solution of (P) .

Duality results

- Assumptions: $Q = \mathbb{R}^n,\ |T_1| = 1,\ \nabla_x g(\cdot,\cdot)$ continuous on $T_2 \times \mathbb{R}^n$.
- $\mathcal{C}(T_2)$: Banach space of real-valued continuous functions on T_2, with $\|h\| = \max\{|h(t)| : t \in T_2\}$.
- $\mathcal{M}(T_2)$ its topological dual: the finite signed Borel measures on T_2, embedded with the total variation norm,
- $\langle h, \sigma \rangle := \int_{T_2} h(t)\sigma(dt)\ \forall \sigma \in M(T_2),\ \forall h \in C(T_2)$.
- $M_+(T_2)$: the positive cone of $M(T_2)$, i.e. the finite Borel measures on T_2. For $\sigma \in M_+(T_2)$ we have $||\sigma|| = \int_{T_2} \sigma(dt)$.
- $L(x, \sigma)$: the usual Lagrangian function associated to (P)

$$L(x,\sigma) := F(x) + \langle g(.,x), \sigma \rangle = F(x) + \int_{T_2} g(t,x)\sigma(dt).$$

Dual multipliers

- The dual functional: $\psi(\sigma) := \inf\{L(x,\sigma) | \; x \in \mathbb{R}^n\}$ and the dual problem
- $(D) \quad v(D) = \sup\{\psi(\sigma) | \; \sigma \in M_+(T_2)\}$
- Assumptions: (A_1)+Slater condition $\Rightarrow v(P) = v(D)$ and the optimal set of (D) is not empty.
- Stopping rule in:**RPSALG** compute $x^k \in \mathbb{R}^n$:

$$\left\| \nabla F(x^k) + \frac{\gamma_k}{|T_2^k|} \sum_{t \in T_2^k} \theta'(g(t,x^k)\delta_k) \nabla_x g(t,x^k) + 2\epsilon_k x^k \right\| \leq \sqrt{2}\epsilon_k.$$

- Multipliers: discrete measures $\{\sigma^k\}$ associated with $\{x^k\}$

$$\sigma^k := \frac{\gamma_k}{|T_2^k|} \sum_{t \in T_2^k} \theta'(g(t,x^k)\delta_k)\alpha_t,$$

where α_t is the Dirac distribution concentrated at t.

Dual convergence

- Theorem: The sequence $\{\sigma^k\}$ is strongly bounded. There exists at least a $weak^*-$limit point of this sequence, and each $weak^*-$limit point of this sequence belongs to S_D.
- Definition: a sequence $\{\sigma^k\} \in M(T_2)$ converges $weakly^*$ to $\sigma \in M(T_2)$ if

$$\lim_{k\to\infty} \left\langle h, \sigma^k \right\rangle = \langle h, \sigma \rangle, \quad \forall h \in \mathcal{C}(T_2).$$

Comparisons with related papers

For $T_2 = \emptyset$: Sheu-Lin (2004, JOTA) with Q compact.

Comparisons with related papers

For $T_2 = \emptyset$: Sheu-Lin (2004, JOTA) with Q compact.
For $|T_1| = 1$: Martinet (1972, thesis)

$$\tilde{G}^k(x) = \alpha_k \sum_{t \in T_2^k} \theta(g(t,x)), \quad \text{with } \alpha_k \geq 1.$$

Comparisons with related papers

For $T_2 = \emptyset$: Sheu-Lin (2004, JOTA) with Q compact.
For $|T_1| = 1$: Martinet (1972, thesis)

$$\tilde{G}^k(x) = \alpha_k \sum_{t \in T_2^k} \theta(g(t,x)), \quad \text{with } \alpha_k \geq 1.$$

Non-convex case but stronger conditions

Comparisons with related papers

For $T_2 = \emptyset$: Sheu-Lin (2004, JOTA) with Q compact.
For $|T_1| = 1$: Martinet (1972, thesis)

$$\tilde{G}^k(x) = \alpha_k \sum_{t \in T_2^k} \theta(g(t,x)), \quad \text{with } \alpha_k \geq 1.$$

Non-convex case but stronger conditions
a) F is level bounded on Q instead on the feasible set,

Comparisons with related papers

For $T_2 = \emptyset$: Sheu-Lin (2004, JOTA) with Q compact.
For $|T_1| = 1$: Martinet (1972, thesis)

$$\tilde{G}^k(x) = \alpha_k \sum_{t \in T_2^k} \theta(g(t,x)), \quad \text{with } \alpha_k \geq 1.$$

Non-convex case but stronger conditions
a) F is level bounded on Q instead on the feasible set,
b) $\theta(t) = 0, \ \forall t \leq 0$ excluding important penalty functions : $\exp t$ and $\theta_1, \theta_2, \theta_3$..,

Comparisons with related papers

For $T_2 = \emptyset$: Sheu-Lin (2004, JOTA) with Q compact.
For $|T_1| = 1$: Martinet (1972, thesis)

$$\tilde{G}^k(x) = \alpha_k \sum_{t \in T_2^k} \theta(g(t,x)), \quad \text{with } \alpha_k \geq 1.$$

Non-convex case but stronger conditions
a) F is level bounded on Q instead on the feasible set,
b) $\theta(t) = 0,\ \forall t \leq 0$ excluding important penalty functions : $\exp t$ and $\theta_1, \theta_2, \theta_3$..,
c) concerns essentially functions as $(t_+)^2, (t_+)^3$ and in this case the two models coincide.

Integral methods

- Assumptions: $int(T_i)$ nonempty in $\mathbb{R}^m$ and either Q is compact either

Integral methods

- Assumptions: $int(T_i)$ nonempty in $\mathbb{R}^m$ and either Q is compact either
-

$$f(t,x) = h(x) + \langle a(t), x\rangle - b(t) \; \forall t \in T_1, \quad g(t,x) = \langle a(t), x\rangle - b(t) \; \forall t \in T_2,$$

where $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, convex, and $\mathcal{C}^1$ on Q.

Integral methods

- Assumptions: $int(T_i)$ nonempty in $\mathbb{R}^m$ and either Q is compact either
-

$$f(t,x) = h(x) + \langle a(t), x\rangle - b(t)\ \forall t \in T_1, \quad g(t,x) = \langle a(t), x\rangle - b(t)\ \forall t \in T_2,$$

where $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, convex, and $\mathcal{C}^1$ on Q.

- For $\epsilon_k > 0,\ \delta_k > 0,\ p_k > 0,\ \gamma_k > 0$, and $\theta \in \mathcal{F}$ we set

$$I_k(x) = \frac{1}{p_k} \log\left(\int_{T_1} \exp[(f(t,x)p_k]dt\right),\ E_k(x) = \gamma_k \int_{T_2} \frac{\theta(g(t,x)\delta_k)}{\delta_k} dt,$$

where dt is the Lebesgue measure.

Integral methods

- Assumptions: $int(T_i)$ nonempty in $\mathbb{R}^m$ and either Q is compact either
-

$$f(t,x) = h(x) + \langle a(t), x\rangle - b(t)\ \forall t \in T_1, \quad g(t,x) = \langle a(t), x\rangle - b(t)\ \forall t \in T_2,$$

where $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, convex, and $\mathcal{C}^1$ on Q.

- For $\epsilon_k > 0,\ \delta_k > 0,\ p_k > 0,\ \gamma_k > 0$, and $\theta \in \mathcal{F}$ we set

$$I_k(x) = \frac{1}{p_k} \log\left(\int_{T_1} \exp[(f(t,x)p_k]dt\right),\ E_k(x) = \gamma_k \int_{T_2} \frac{\theta(g(t,x)\delta_k)}{\delta_k} dt,$$

where dt is the Lebesgue measure.

- $R_k(x) := I_k(x) + E_k(x) + \epsilon_k ||x||^2,\ \lim_{k\to\infty} \epsilon_k = 0,$

Integral methods

- Assumptions: $int(T_i)$ nonempty in $\mathbb{R}^m$ and either Q is compact either
-

$$f(t,x) = h(x) + \langle a(t), x\rangle - b(t) \ \forall t \in T_1, \quad g(t,x) = \langle a(t), x\rangle - b(t) \ \forall t \in T_2,$$

  where $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, convex, and $\mathcal{C}^1$ on Q.
- For $\epsilon_k > 0, \ \delta_k > 0, \ p_k > 0, \ \gamma_k > 0$, and $\theta \in \mathcal{F}$ we set

$$I_k(x) = \frac{1}{p_k} \log \left(\int_{T_1} \exp[(f(t,x)p_k]dt \right), \ E_k(x) = \gamma_k \int_{T_2} \frac{\theta(g(t,x)\delta_k)}{\delta_k} dt,$$

  where dt is the Lebesgue measure.
- $R_k(x) := I_k(x) + E_k(x) + \epsilon_k ||x||^2, \ \lim_{k\to\infty} \epsilon_k = 0,$
- $(P^k_{ips}) \quad v(P^k_{ips}) = \min\{R_k | \ x \in Q\}.$

Integral Penalty Smoothing Algorithm: IPSALG

- Assumptions: $\lim_{k\to\infty} p_k = \infty$ and either

Integral Penalty Smoothing Algorithm: IPSALG

- Assumptions: $\lim_{k\to\infty} p_k = \infty$ and either
- $\theta \in \mathcal{F}_1$, $\lim_{k\to\infty} \frac{\gamma_k}{\delta_k} = 0$, and $\lim_{k\to\infty} \gamma_k = +\infty$, either

Integral Penalty Smoothing Algorithm: IPSALG

- Assumptions: $\lim_{k\to\infty} p_k = \infty$ and either
- $\theta \in \mathcal{F}_1$, $\lim_{k\to\infty} \frac{\gamma_k}{\delta_k} = 0$, and $\lim_{k\to\infty} \gamma_k = +\infty$, either
- $\theta \in \mathcal{F}_2$, $\lim_{k\to\infty} \frac{\gamma_k}{\delta_k} = 0$, and $\gamma_k > \varepsilon \, \forall k$, for a certain $\varepsilon > 0$.

Integral Penalty Smoothing Algorithm: IPSALG

- Assumptions: $\lim_{k\to\infty} p_k = \infty$ and either
- $\theta \in \mathcal{F}_1$, $\lim_{k\to\infty} \frac{\gamma_k}{\delta_k} = 0$, and $\lim_{k\to\infty} \gamma_k = +\infty$, either
- $\theta \in \mathcal{F}_2$, $\lim_{k\to\infty} \frac{\gamma_k}{\delta_k} = 0$, and $\gamma_k > \varepsilon\ \forall k$, for a certain $\varepsilon > 0$.
- IPSALG: Compute $x^k \in Q : \ R_k(x^k) \leq \min\{R_k(x) |\ x \in Q\} + \epsilon_k$.

Integral Penalty Smoothing Algorithm: IPSALG

- Assumptions: $\lim_{k\to\infty} p_k = \infty$ and either
- $\theta \in \mathcal{F}_1$, $\lim_{k\to\infty} \frac{\gamma_k}{\delta_k} = 0$, and $\lim_{k\to\infty} \gamma_k = +\infty$, either
- $\theta \in \mathcal{F}_2$, $\lim_{k\to\infty} \frac{\gamma_k}{\delta_k} = 0$, and $\gamma_k > \varepsilon \; \forall k$, for a certain $\varepsilon > 0$.
- IPSALG: Compute $x^k \in Q : \; R_k(x^k) \leq \min\{R_k(x) | \; x \in Q\} + \epsilon_k$.
- Theorem: Under Assumption (A_1) the sequence $\{x^k\}$ built by **IPSALG** is bounded and all its limit points are optimal solutions

Duality results and comparisons

- Assumptions: $Q = \mathbb{R}^n, \ |T_1| = 1$ and $x^k : \ ||\nabla R_k(x^k)|| \leq \sqrt{2}\epsilon_k$.

Duality results and comparisons

- Assumptions: $Q = \mathbb{R}^n$, $|T_1| = 1$ and $x^k : \ ||\nabla R_k(x^k)|| \leq \sqrt{2}\epsilon_k$.
- $$x^k \to \sigma^k := \gamma_k J_0(\theta'(g(\cdot, x^k)\delta_k)) \in M_+(T_2),$$

Duality results and comparisons

- Assumptions: $Q = \mathbb{R}^n,\ |T_1| = 1$ and $x^k:\ ||\nabla R_k(x^k)|| \leq \sqrt{2}\epsilon_k$.
- $$x^k \to \sigma^k := \gamma_k J_0(\theta'(g(\cdot, x^k)\delta_k)) \in M_+(T_2),$$
- where $J_0 : \mathcal{C}_+(T_2) \to M_+(T_2)$ is the linear map

$$\langle J_0 f, h\rangle := \int_{T_2} hf\,dt \quad \forall f \in \mathcal{C}_+(T_2),\ \forall h \in \mathcal{C}(T_2).$$

Duality results and comparisons

- Assumptions: $Q = \mathbb{R}^n$, $|T_1| = 1$ and $x^k : \; ||\nabla R_k(x^k)|| \leq \sqrt{2}\epsilon_k$.
- $$x^k \to \sigma^k := \gamma_k J_0(\theta'(g(\cdot, x^k)\delta_k)) \in M_+(T_2),$$
- where $J_0 : \mathcal{C}_+(T_2) \to M_+(T_2)$ is the linear map

$$\langle J_0 f, h \rangle := \int_{T_2} h f dt \quad \forall f \in \mathcal{C}_+(T_2), \; \forall h \in \mathcal{C}(T_2).$$

- Theorem: Assume (A_1) and Slater condition, then $\{\sigma^k\}$ is strongly bounded. There exists at least a $weak^*$–limit point of this sequence, and each $weak^*$–limit point of this sequence is an optimal solution of the dual (D).

Duality results and comparisons

- Assumptions: $Q = \mathbb{R}^n$, $|T_1| = 1$ and $x^k : \ ||\nabla R_k(x^k)|| \leq \sqrt{2}\epsilon_k$.
-

$$x^k \to \sigma^k := \gamma_k J_0(\theta'(g(\cdot, x^k)\delta_k)) \in M_+(T_2),$$

- where $J_0 : \mathcal{C}_+(T_2) \to M_+(T_2)$ is the linear map

$$\langle J_0 f, h\rangle := \int_{T_2} hf\,dt \quad \forall f \in \mathcal{C}_+(T_2),\ \forall h \in \mathcal{C}(T_2).$$

- Theorem: Assume (A_1) and Slater condition, then $\{\sigma^k\}$ is strongly bounded. There exists at least a $weak^*-$limit point of this sequence, and each $weak^*-$limit point of this sequence is an optimal solution of the dual (D).
- For $|T_1| = 1$ particular cases :$\theta_4, \theta_3, \theta_6$. For $T_2 = \emptyset$:Sheu-Lin (2004). In all cases weaker convergence results.