Lecture Notes in Mathematics

Edited by A. Dold, B. Eckmann and F. Takens

1405

S. Dolecki (Ed.)

Optimization

Proceedings, Varetz 1988

Springer-Verlag

Lecture Notes in Mathematics

Edited by A. Dold, B. Eckmann and F. Takens

1405

S. Dolecki (Ed.)

Optimization

Proceedings of the Fifth French-German Conference
held in Castel-Novel (Varetz), France, Oct. 3–8, 1988

Springer-Verlag
Berlin Heidelberg New York London Paris Tokyo Hong Kong

Editor

Szymon Dolecki
Départment de Mathématiques, Faculté des Sciences
123 rue A. Thomas, 87060 Limoges, France

Mathematics Subject Classification (1980): 49–06

ISBN 3-540-51970-X Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-51970-X Springer-Verlag New York Berlin Heidelberg

Printed in Germany

Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
2146/3140-543210 – Printed on acid-free paper

INTRODUCTION

I am pleased to present the proceedings of the Fifth French-German Conference on Optimization held in the castle Castel-Novel in Varetz near Brive from the 3rd to 8th of October 1988. Its aim was to review the work carried out by various research groups, to intensify the exchange of ideas and to evaluate the state of arts and the trends in the area of optimization.

As a consequence of the spectacular growth in speed of computation, one witnesses an increasing role of discrete optimization and of computation complexity questions. In order to reflect these trends a survey talk on discrete optimization was invited and a special session was dedicated to projective methods in linear programming.

During the meeting, state-of-art talks were given in selected topics: identification – H. Bock (*Heidelberg*), nonsmooth optimization – A. D. Ioffe (*Haifa*), discrete optimization – B. Korte (*Bonn*), sensitivity analysis – K. Malanowski (*Warszawa*), projective methods – J.-Ph. Vial (*Genève*).

The contents of the volume differ slightly from the program of the conference (the latter is recalled on pages V and VI). In fact, the results published elsewhere do not appear here, as require the rules of the Lecture Notes. On the other hand, some authors unable to attend the meeting wished nevertheless to contribute to the proceedings. I am very grateful to the referees of this volume for the excellent work that they have done.

Previous French-German conferences were organized:

– first in *Oberwolfach* (16-24 March 1980) by A. Auslender, W. Oettli and J. Stoer;

– second in *Confolant* (16-20 March 1981) by J.-B. Hiriart-Urruty;

– third in *Luminy* (2-6 July 1984) by C. Lemaréchal;

– fourth in *Irsee* (21-26 April 1986) by K.-H. Hoffmann, J. Zowe, J.-B. Hiriart-Urruty and C. Lemaréchal.

The *Equipe d'Analyse non linéaire et d'Optimisation* did its best to prepare this meeting. On the other hand, I had received valuable assistance from several members of the Scientific Committee; I owe them my gratitude.

Let me express high appreciation to the institutions that supported this conference which would not have taken place without their generosity. I gratefully acknowledge the cordial reception by the City of Brive who offered also a recital (of Thérèse Dussaut and Jean Barthe) and by the Conseil Général de la Dordogne during the excursion to Sarlat and the Grottes de Lascaux.

Limoges, the 5th of June 1989

Szymon Dolecki

Scientific Committee

S. Dolecki (Limoges), **J.-B. Hiriart-Urruty** (Toulouse), **K.-H. Hoffmann** (Augsburg), **W. Krabs** (Darmstadt), **C. Lemaréchal** (Rocquencourt), **J.-P. Penot** (Pau), **J. Stoer** (Würzburg), **J. Zowe** (Bayreuth) .

Organizing Committee

S. Dolecki, J. Guillerme, C. Malivert, M. Théra (Equipe d'Analyse non linéaire et Optimisation de Limoges) and **V. Piqueras** (clerical assistant).

Supporting Institutions

Conseil Général de la Corrèze
Conseil Régional du Limousin
Département de Mathématiques de Limoges
Direction des Recherches Etudes et Techniques
Electricité de France
European Research Office
Ministère des Affaires Etrangères
Ministère de l'Education Nationale
Ministère de la Recherche et de la Technologie
Université de Limoges
Ville de Brive
Excursion partially supported by
Conseil Général de la Dordogne
Technical support
Macorbur

Sponsors

Centre National de la Recherche Scientifique
Faculté des Sciences de Limoges
Société Mathématique de France

TABLE OF CONTENTS

Optimization – Fifth French-German Conference
Castel Novel 1988, Lecture Notes in Mathematics 1405
Springer Verlag 1989, pp. 1-11

A SMOOTHING TECHNIQUE FOR NONDIFFERENTIABLE OPTIMIZATION PROBLEMS

A. BEN-TAL[1] AND M. TEBOULLE[2]

Abstract . We introduce a smoothing technique for nondifferentiable optimization problems. The approach is to replace the original problem by an approximate one which is controlled by a smoothing parameter. The recession function is instrumental in the construction of the approximate problem. An a priori bound on the difference between the optimal values of the original problem and the approximate one is explicitly derived in term of the smoothing parameter. The relationships between the primal approximated problem and its corresponding dual are investigated.

1. INTRODUCTION

In this paper we introduce a smoothing mechanism for nondifferentiable optimization problems. The idea underlying our approach is to approximate the original nondifferentiable problem by a perturbed problem. The basic tool to generate such an approximate problem is through the use of recession functions. The resulting approximate problem is a smooth optimization problem which contains a smoothing parameter. This parameter controls the accuracy of the approximation. When the parameter approaches zero, the original problem is recovered. In Section 2 we recall some basic properties of recession functions and present the framework for smoothing nondifferentiable optimization problems. Our approach is general enough to cover many interesting problems, including ℓ_1-norm minimization and min-max optimization. This is illustrated via examples. In Section 3, we derive an a priori error bound on the difference between the optimal values of the original problem and the approximate one. The duality correspondence existing between the recession function and the support function leads naturally to explore the relationships between the primal perturbed problem and its corresponding dual. These duality results are derived in Section 4 and some applications are given.

We will frequently referred to results in Rockafellar book [3]. The notations and definitions used here are the same as in that book. Recall that the recession function of the function g is denoted by $g0^+$, the domain of g is denoted by dom g. The conjugate function of g is denoted by g^* and is defined as

$$g^*(z) = \sup\{ \langle x, z \rangle - g(x) : x \in \text{dom } g\}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^m$. The support function of a set S is given by $\delta^*(z|S) = \sup\{ \langle x, z \rangle : x \in S\}$ and the relative interior of S is denoted by ri S.

---

[1] Faculty of Industrial Engineering and Management, Technion, Israel Institute of Technology, Haifa 32000, Israel.
Supported by NSF Grant ECS-8801240.

[2] Department of Mathematics and Statistics, University of Maryland, Baltimore County, Baltimore. Supported by AFOSR Grant 0218-88 and NSF Grant ECS-8802239.

2. The smoothing method

We consider the following optimization problem :

$$\text{(P)} \qquad \inf_{x \in R^n} \{G(x) := F(f_1(x),..., f_m(x))\}$$

and where we assume that $\{f_i(x), i = 1,..., m\}$ are real functions over $\mathbb{R}^n$, and F is the recession function of some proper convex function g,

$$\text{(2.1)} \qquad F(y) = g0^+(y) = \sup\{g(x + y) - g(x) : x \in \text{dom } g\}$$

The recession function is a positively homogeneous convex function (see [3, Theorem 8.5, p.66]). In the sequel we assume that $g : \mathbb{R}^m \to \mathbb{R}$ is a closed proper convex function and that $0 \in \text{dom } g$. Let us define

$$F_\varepsilon(y) = \varepsilon\, g(\frac{y}{\varepsilon})$$

Then from [3, Corollary 8.5.2, p.67] we have

$$F(y) = \lim_{\varepsilon \to 0} F_\varepsilon(y) \quad \forall y \in \mathbb{R}^m$$

Thus, for small $\varepsilon > 0$, we can approximate F by F_ε. Hence, an approximated (perturbed) problem for (P) is :

$$(P_\varepsilon) \qquad \inf_{x \in \mathbf{R}^n} \left\{ G_\varepsilon(x) := \varepsilon\, g(\frac{f_1(x)}{\varepsilon},..., \frac{f_m(x)}{\varepsilon}) \right\}$$

The usefulness of the approximation lies in the fact that frequently g (and hence F_ε) is a smooth function, while F is not. Here $\varepsilon > 0$ plays the role of a smoothing parameter.

2.1. Example . Consider the ℓ_1-norm minimization problem :

$$\text{(P1)} \qquad \min_{x \in R^n} \sum_{i=1}^{m} | f_i(x) |$$

Here, $F(y) = \sum_{i=1}^{m} | y_i |$. It is easily verified that F(y) is the recession function of

$$g(y) = \sum_{i=1}^{m} \sqrt{1 + y_i^2}\, ,$$

hence the resulting approximate problem is the smoothed problem :

$$(P1)_\varepsilon \qquad \min_{x \in R^n} \sum_{i=1}^{m} \sqrt{f_i(x)^2 + \varepsilon^2}$$

This is precisely the nonlinear approximation problem suggested by El-Attar et al. [2] for solving the ℓ_1-norm optimization problem.

2.2. EXAMPLE . Consider the continuous-discrete min-max optimization problem :

$$(P2) \qquad \min_{x \in R^n} \max_{1 \le i \le m} f_i(x)$$

Here, $F(y) = \max_{1 \le i \le m} y_i$, which is the recession function of

$$g(y) = \log \sum_{i=1}^{m} e^{y_i}$$

Hence, the resulting approximation is the smoothed problem :

$$(P2)_\varepsilon \qquad \min_{x \in R^n} \varepsilon \log \sum_{i=1}^{m} \exp\left(\frac{f_i(x)}{\varepsilon}\right)$$

This approximate function here is similar to the penalty function used in Bertsekas [1, Section 5.1.3].

3. ERROR ANALYSIS

In this section we show that the difference between the optimal values of the original function and the approximate one, is bounded by a term which is proportional to the smoothing parameter ε. We first recall the following result.

3.1. LEMMA . [3, Theorem 13.3, p. 116] *Let* g *be a proper convex function. Then*

$$(g^*0^+)(x^*) = \delta^*(x^* \mid \operatorname{dom} g) = \sup_{x \in \operatorname{dom} g} \langle x, x^* \rangle$$

If g is closed, then

$$(g0^+)(x) = \delta^*(x \mid \operatorname{dom} g) = \sup_{x^* \in \operatorname{dom} g} \langle x, x^* \rangle$$

We denote by x^* and x^*_ε the minimizers of $G(x)$ and $G_\varepsilon(x)$ respectively. We may now establish the main result of this section.

3.2. THEOREM . *Let* g *be a closed proper convex function and suppose that the following assumption holds :*

$$(A) \qquad \forall \varepsilon > 0 \qquad \varepsilon\, g(z/\varepsilon) \ge (g0^+)(z) \qquad \forall z$$

Then for every $\varepsilon > 0$

$$0 \le G(x^*_\varepsilon) - G(x^*) \le \varepsilon g(0)$$

Proof . Since g is closed we have

$$g^{**}(z) = g(z) = \sup_{t \in \mathrm{dom}\, g^*} \{ \langle t, z \rangle - g^*(t) \}$$
$$\le \sup_{t \in \mathrm{dom}\, g^*} \langle t, z \rangle + \sup_{t \in \mathrm{dom}\, g^*} (-g^*(t))$$

but $\sup_{t \in \mathrm{dom}\, g^*} (-g^*(t)) = g(0)$ and, from Lemma 3.1 $\sup_{t \in \mathrm{dom}\, g^*} \langle t, z \rangle = (g0^+)(z)$. Hence,

$$g(z) - (g0^+)(z) \le g(0) \text{ for every } z$$

Applying the above inequality at $z = f(x) / \varepsilon = (f_1(x) / \varepsilon, \ldots, f_m(x) / \varepsilon)^T$ and using the fact that $(g0^+)$ is positively homogeneous, it follows that

$$(3.1) \qquad G_\varepsilon(x) - G(x) \le \varepsilon \cdot g(0)$$

Now, since $x^* = \operatorname{argmin} G(x)$, then $G(x^*) \le G(x)$ for every x and thus in particular

$$G(x^*_\varepsilon) - G(x^*) \ge 0.$$

From assumption (A) we have $G_\varepsilon(x) \ge G(x)$ for every x. Hence

$$G(x^*_\varepsilon) - G(x^*) \le G_\varepsilon(x^*_\varepsilon) - G(x^*)$$

But, $x^*_\varepsilon = \operatorname{argmin} G_\varepsilon(x)$, then $G_\varepsilon(x^*_\varepsilon) \le G_\varepsilon(x)$ for every x and thus in particular $G_\varepsilon(x^*_\varepsilon) \le G_\varepsilon(x^*)$. Combining with the above inequality it follows that

$$G(x^*_\varepsilon) - G(x^*) \le G_\varepsilon(x^*_\varepsilon) - G(x^*) \le G_\varepsilon(x^*) - G(x^*) \le \varepsilon \cdot g(0)$$

The last inequality following from (3.1) □

Note that assumption (A) holds trivially for Examples 2.1 and 2.2

3.3. EXAMPLE . From Example 2.1 we have $g(0) = m$, thus for the ℓ_1-norm minimization problem the following holds :

$$0 \le G(x^*_\varepsilon) - G(x^*) \le \varepsilon \cdot m$$

3.4. EXAMPLE . From Example 2.2 we have $g(0) = \log m$, hence for the min-max problem

$$0 \le G(x^*_\varepsilon) - G(x^*) \le \varepsilon \cdot \log m$$

4. DUAL PROBLEMS

The dual correspondence existing between the recession function and the support function as given in Lemma 3.1, leads naturally to explore the relationships between the primal problem and the induced dual. In the sequel we assume that the functions $f_1(x),\dots, f_m(x)$ are convex and we denote $f(x) = (f_1(x), \dots, f_m(x))^T$.

Recall that the primal problem is

(P) $$\inf_{x\in R^n} \{G(x) = (g0^+)(f_1(x),\dots, f_m(x))\}$$

By Lemma 3.1, we have

$$(g0^+)(y) = \delta^*(y \mid \mathrm{dom}\, g^*)$$

Thus (P) can be rewritten as the following min-max problem :

(P) $$\inf_{x\in R^n} \delta^*(f(x) \mid \mathrm{dom} g^*) = \inf_{x\in R^n} \sup_y \{ <y, f(x)> : y \in \mathrm{dom} g^*\}$$

Therefore a natural dual problem for (P) is

(D) $$\sup\{H(y) : y \in \mathrm{dom}\, g^*\}$$

where the dual objective function is

$$H(y) = \inf\{ <y, f(x)> : x \in \mathbb{R}^n\}$$

4.1. THEOREM . *Under one of the following two conditions, a strong duality result holds for the pair of problems* (P) – (D) *i.e.* inf(P) = max(D).

(a) The functions f_i *are affine and* dom g^* *is closed*

(b) $\mathrm{dom}\, g^* \subset \mathbb{R}^m_+$ *and* $\exists \bar{x} \in \mathbb{R}^n$ $f(\bar{x}) < 0$

Proof . (a) Problem (P) can be rewritten equivalently as a linearly convex problem

(P) $$\inf_{x,z} \{\delta^*(z \mid \mathrm{dom}\, g^*) : f(x) = z\}$$

The corresponding Lagrangian is

$$L(x, z) = \delta^*(z \mid \mathrm{dom}\, g^*) + <y, f(x)> - <y, z>, \quad y \in \mathbb{R}^m$$

and thus the dual objective is

$$\begin{aligned}\inf_{x,z} L(x, z) &= \inf_x <y, f(x)> + \inf_z \{\delta^*(z \mid \mathrm{dom}\, g^*) - <y, z>\}\\ &= H(y) - \sup\{ <y, z> - \delta^*(z \mid \mathrm{dom}\, g^*)\}\\ &= H(y) - \delta^{**}(y \mid \mathrm{dom}\, g^*)\\ &= H(y) - \delta(y \mid \mathrm{dom}\, g^*) \text{ since dom g is closed}\end{aligned}$$

Hence, the dual of (P) is

(D) $$\sup\{H(y) : y \in \operatorname{dom} g^*\}$$

and since (P) is a convex linearly constrained problem the result follows from standard duality arguments, see e.g. [3].
(b) Let $K(x, y) = \langle y, f(x) \rangle$. Then, problem (P) is simply

(P) $$\inf_{x \in \mathbb{R}^n} \sup_{y \in \operatorname{dom} g^*} K(x, y)$$

Since the f_i are convex and $\operatorname{dom} g^* \subset \mathbb{R}^m_+$, then $K(\cdot, y)$ are convex for all $y \in \operatorname{dom} g^*$, and $K(x, \cdot)$ are concave (linear) for all $x \in \mathbb{R}^n$. By a result in [4], a sufficient condition for the validity of a strong duality result for a general convex-concave saddle function $K(x, y)$ is :

There is no $d \in \operatorname{dom} g^*$ such that $\langle d, \nabla_y K(x,y)\rangle \geq 0 \ \forall x \in \mathbb{R}^n, \forall y \in \operatorname{ri} \operatorname{dom} g^*$.

Here we have $\nabla_y K(x, y) = f(x)$. Hence the above means

There is no $d \in \operatorname{dom} g^*$ such that $\langle d, f(x)\rangle \geq 0 \ \forall x \in \mathbb{R}^n, \forall y \in \operatorname{ri} \operatorname{dom} g^*$

But the latter is certainly satisfied if there exists $\bar{x}$ such that $f(\bar{x}) < 0$ and thus the result follows.

4.2. REMARK . The above theorem remains true for the constrained nonsmooth program i.e. (P) is $\inf\{G(x) : x \in S\}$ where S is a given non empty convex subset of $\mathbb{R}^n$. In that case problem (P) can be written as :

(P) $$\inf_{x \in \mathbb{R}^n} \delta^*(f(x) \mid \operatorname{dom} g^*) + \delta(x \mid S)$$

and the dual problem for (P) is

(D) $$\sup\{H(y) : y \in \operatorname{dom} g^*\}$$

where the dual objective function is

$$H(y) = \inf\{\langle y, f(x) \rangle + \delta(x \mid S) : x \in \mathbb{R}^n\}$$

When f(x) are affine, say $f(x) = b - Ax$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, then it is easy to verify that the dual problem is given by

$$\text{(D)} \quad \sup\{y^T b - \delta^*(A^T y \mid S) : \ y \in \operatorname{dom} g^*\}$$

This is illustrated below in Example 4.4.

4.3. REMARK . It is possible to derive an alternative representation for the dual objective function H(y) by using the infimal convolution property, see [3, Theorem 16.1 and 16.4, pp. 140-145]. We have

$$H(y) = \inf_x \langle y, f(x) \rangle = -\left(\sum_{i=1}^m y_i f_i(.)\right)^*(0) = -\inf_{z_1,\dots,z_m} \left\{\sum_{i=1}^m (f_i^* y_i)(z_i) : \sum_{i=1}^m z_i = 0\right\}$$

where

$$(f_i^* y_i)(z_i) = \begin{cases} y_i f_i^*(\frac{z_i}{y_i}) & \text{if } y_i \in \text{int dom } g^* \\ \delta(z_i \mid 0) & \text{if } y_i = 0 \end{cases}$$

Hence an alternative dual problem to (D) is

$$(\bar{D}) \qquad \sup_{y \in \text{dom} g^*} \sup_{z_1,\dots,z_m} \left\{ -\sum_{i=1}^{m} (f_i^* y_i)(z_i) \ : \ \sum_{i=1}^{m} z_i = 0 \right\}$$

The above representation is particularly useful when g^* is separable, i.e. $g^*(y) = \Sigma_{i=1}^m g_i^*(y_i)$. Indeed, in that case, dom $g^* = \bigcap_{i=1}^m$ dom g_i^* and therefore problem $(\bar{D})$ can be written as

$$(\bar{D}) \qquad \sup \left\{ \sum_{i=1}^{m} k_i(z_i) \ : \ \sum_{i=1}^{m} z_i = 0 \right\}$$

where $k_i(z_i) = \sup_{y_i} - (f_i^* y_i)(z_i)$ is usually easy to compute. As an example take $f_i(x) = \frac{1}{2} x^T Q_i x - b_i$ where the Q_i are $n \times n$ positive definite matrices. Then, $k_i(z_i) = -(b_i^* Q_i^{-1} b_i)^{1/2} (z_i^T Q_i^{-1} z_i)^{1/2} - b_i^T Q_i^{-1} z_i$.

4.4. EXAMPLE . Consider the convex constrained nonsmooth problem

$$(P) \qquad \min \left\{ \sum_{i}^{m} | a_i^T x - b_i | \ : \ x_j^2 + x_{j+1}^2 \le 1,\ j = 1, 3, \dots, n-1 \right\}$$

where n is assumed even. Then the dual objective function is

$$H(y) = - b^T y + \inf \left\{ \sum_{i}^{m} a_i^T x : \ x_j^2 + x_{j+1}^2 \le 1,\ j \in J \right\}$$

where $J = \{1, 3, \dots, n-1\}$. But the above can be rewritten as :

$$\begin{aligned} H(y) &= - b^T y + \sum_{j \in J} \inf_{(x_j, x_{j+1})} \{ x_j y^T a^j + x_{j+1} y^T a^{j+1} : x_j^2 + x_{j+1}^2 \le 1, j \in J \} \\ &= - b^T y - \sum_{j \in J} \delta^* \left((y^T a^j, y^T a^{j+1}) \mid x_j^2 + x_{j+1}^2 \le 1 \right) \\ &= - b^T y - \sum_{j \in J} \sqrt{(y^T a^j)^2 + (y^T a^{j+1})^2} \end{aligned}$$

where $a^j \in \mathbb{R}^m$ is the j-th column of the matrix $A = (a_1, \dots, a_m)^T$. Also, from Example 2.1 we know that $g(z) = \Sigma_{i=1}^m \sqrt{1 + z_i^2}$. Then, the conjugate is

$$g^*(y) = \sum_{i=1}^{m} \sqrt{1 - y_i^2},\ \forall y \in \text{dom } g^* = \{ y \in R^m : -1 \le y \le 1 \}$$

and hence, the dual problem is

(P) $$\sup\{-b^Ty - \sum_{j\in J}\sqrt{(y^Ta^j)^2 + (y^Ta^{j+1})^2} : -1 \le y \le 1\}$$

We may now establish a similar duality result for the approximate smoothed primal problem :

(P_ε) $$\inf_{x\in \mathbb{R}^n} \varepsilon g\left(\frac{f_1(x)}{\varepsilon},\dots \frac{f_m(x)}{\varepsilon}\right)$$

Using the definition of a conjugate function and the fact that g is closed, problem (P_ε) can be written as

$$\inf_{x\in \mathbb{R}^n} \sup_{y\in \operatorname{dom} g^*} \{<y, f(x)> - \varepsilon g^*(y)\}$$

Thus again, a natural dual is obtained by replacing min-max with max-min. This yields the following dual approximate problem

(D_ε) $$\sup\{H_\varepsilon(y) : y \in \operatorname{dom} g^*\}$$

where

$$H_\varepsilon(y) = H(y) - \varepsilon\, g^*(y)$$

and H(y) is the dual objective of the original problem (P).

4.5. THEOREM . *Under conditions* (a) *or* (b) *of Theorem 4.1 we have*

$$\inf(P_\varepsilon) = \max(D_\varepsilon)$$

Moreover if g *is differentiable, then if* x^*_ε *solves* (P_ε), *the optimal solution* y^*_ε *is given by*

$$y^*_\varepsilon = \nabla g\left(\frac{f_1(x^*_\varepsilon)}{\varepsilon},\dots,\frac{f_m(x^*_\varepsilon)}{\varepsilon}\right)$$

Proof . (a) Following step by step the proof of Theorem 4.1 (a), the result follows. Also, since $g^*(y) = \sup_z\{<y, z/\varepsilon> - g(z/\varepsilon)\}$, by simple calculus the optimality condition gives the primal-dual relationships $y^* = \nabla g(f(x^*)/\varepsilon)$.
(b) Let $K(x, y) = <y, f(x)> - \varepsilon g^*(y)$. Since f_i are convex and dom $g^* \subset \mathbb{R}^m_+$, then $K(\cdot,y)$ are convex for all $y \in$ dom g^*, and $K(x, \cdot)$ are concave (since g^* convex and $\varepsilon > 0$) for all $x \in \mathbb{R}^n$. By [3, Theorem 37.3, p. 392] a sufficient condition for the validity of a strong duality result for a general convex-concave saddle function K is that the convex functions $-K(x,.)$ have no common directions of recession. By [3, Theorem 37.1, and Corollary 37.21 pp. 391-392] this means that, for all $0 \neq w \in$ dom g^* ,

$$\sup_{x\in \mathbf{R}^n} \sup_{y\in \mathrm{ri\ dom}\ g^*} \{K(x, y) - K(x, y + w)\} > 0$$

Substituting the values of K this means that we have to show that

$$\varepsilon \sup_{y\in \mathrm{ri\ dom}\ g^*} \{g^*(y + w) - g^*(y)\} > \inf_{x\in \mathbf{R}^n} < w, f(x) >$$

But from (2.1), the left hand side of the above inequality is exactly $(g^*0^+)(w)$ and from Lemma 3.1

$$(g^*0^+)(w) = \delta^*(w \mid \mathrm{dom}\ g) = (\delta^*(w \mid \mathrm{ri\ dom}\ g))$$

Thus we have to show that $\forall w \neq 0 \in \mathrm{dom}\ g^*$

$$\varepsilon . \sup_{t\in \mathrm{dom} g} < t, w > > \inf_{x\in R^n} < w, f(x) >$$

or equivalently that

$$\sup\{ < \varepsilon t - f(x), w > : t \in \mathrm{dom}\ g, x \in \mathbb{R}^n\} > 0$$

But the later is always satisfied since $0 \in \mathrm{dom}\ g$ and we assumed the existence of a point $\overline{x}$ such that $f(\overline{x}) < 0$. □

4.6. EXAMPLE . Consider problem (P1) given in Example 2.1 with f_i given convex. The function $g(z) = \Sigma_{i=1}^m \sqrt{1 + z_i^2}$. Then, the conjugate is

$$g^*(y) = \sum_{i=1}^m \sqrt{1 - y_i^2},\ \forall y \in \mathrm{dom}\ g^* = \{ y \in R^m : -1 \le y \le 1\}$$

Then a perturbed dual is given by

$$(D1_\varepsilon) \qquad \sup\left\{H(y) - \varepsilon \sum_{i=1}^m \sqrt{1 - y_i^2} : -1 \le y \le 1\right\}$$

Clearly the conditions (b) of Theorem 4.5 are satisfied if there is a Slater point for the function f_i. The optimal solution of the dual is given by

$$(y_\varepsilon)_i = \frac{f_i(x_\varepsilon)}{\sqrt{f_i^2(x_\varepsilon) + \varepsilon^2}} \qquad i = 1,..., m$$

where x_ε is the optimal solution of the primal. Note that the perturbed solution y_ε is feasible for the original dual problem (D).

4.7. EXAMPLE . Consider problem (P2) given in Example 2.2 with f_i given convex. The function $g(z) = \log \sum_{i=1}^{m} e^{z_i}$. Then, the conjugate is

$$g^*(y) = \sum_{i=1}^{m} y_i \log y_i, \quad \forall y \in \text{dom } g^* = \{y \in R^m : y \geq 0, \sum_{i=1}^{m} y_i = 1\}$$

Then a perturbed dual is given by

$$(D2)_\varepsilon \qquad \sup \{H(y) - \varepsilon \sum_{i=1}^{m} y_i \log y_i : y \geq 0, \sum_{i=1}^{m} y_i = 1\}$$

Clearly the conditions (b) of Theorem 4.5 are satisfied if there is a Slater point for the functions f_i. The optimal solution of the dual is given by

$$(y_\varepsilon)_i = \frac{e^{f_i(x_\varepsilon)/\varepsilon}}{\sum_{j=1}^{m} e^{f_j(x_\varepsilon)/\varepsilon}}, \quad i = 1, \ldots, m$$

where x_ε is the optimal solution of the primal. Note that the perturbed solution y_ε is feasible for the original dual problem.
In the linear case ($f_i(x) = a_i^T x - b_i$), conditions (a) of Theorem 4.5 are satisfied and for both examples we obtain

$$H(y) = \begin{cases} -b^T y & \text{if } A^T y = 0 \\ -\infty & \text{otherwise} \end{cases}$$

where $A^T = (a_1, \ldots, a_m)$. Problems $(D1)_\varepsilon$ and $(D2)_\varepsilon$ are thus producing two different type of perturbed linear programs.

4.8. EXAMPLE . We consider a generalization of Example 2.1. We define $F(z) = \sum_{i=1}^{m} F_i(z_i)$ where

$$F_i(z_i) = u_i \max(0, z_i) - l_i \inf(0, z_i)$$

where u_i, l_i are given non negative real numbers. A direct computation shows that F(z) is the recession function of $g(y) = \sum_{i=1}^{m} g_i(y_i)$ where

$$g_i(z_i) = (1 + l_i^2 \min^2(0, y_i) + u_i^2 \max^2(0, y_i))^{1/2}$$

and the resulting approximate smoothed problem is

$$(P3)_\varepsilon \qquad \inf_{x \in R^n} \sum_{i=1}^{m} \varepsilon\, g_i\left(\frac{f_i(x)}{\varepsilon}\right)$$

To construct the dual of $(P3_\varepsilon)$ we compute the conjugate of g_i :

$$g_i^*(t_i) = \left(1 - \frac{1}{l_i^2}\min^2(0, t_i) - \frac{1}{u_i^2}\max{}^2(0, t_i)\right)^{1/2} ..$$

then the dual objective function is given by

$$H_\varepsilon(y) = H(y) - \varepsilon \sum_{i=1}^{m} g_i^*(y_i)$$

and the dual problem $(D3_\varepsilon)$ is

$(D3_\varepsilon)$ $$\max\{H_\varepsilon(y) \ : \ y \in \text{dom } g^* = \{y : -l_i \le y_i \le u_i \ \forall \ i\}\}$$

In the linear case $(f_i(x) = a_i^T x - b_i)$ we thus obtain the perturbed dual linear program

$$\max \{ -b^t y - \varepsilon \sum_{i=1}^{m} \left(1 - \frac{1}{l_i^2}\min^2(0, y_i) - \frac{1}{u_i^2}\max{}^2(0, t_i)\right)^{1/2} \}$$

such that

$$A^T y = 0 \ -l_i \le y_i \le u_i, \quad i = 1,..., m.$$

REFERENCES

[1] **Bertsekas, D.P.**, Constrained Optimization and Lagrangian Multiplier Methods, Academic Press, New York 1982.

[2] **El-Attar, R.A., Vidyasagar, M., Dutta, S.R.K.**, An algorithm for ℓ_1-norm minimization with application to nonlinear ℓ_1-approximation, SIAM J. Numer. Anal., **16** (1979), 70-86.

[3] **Rockafellar, R.T.**, Convex Analysis, Princeton University Press, New Jersey, 1970.

[4] **Rockafellar, R.T.**, Minimax theorems and conjugate saddle functions, Math. Scand., **14** (1964), 151-173.