Absolute minimizer in convex programming by exponential penalty*

F. Alvarez†

Abstract

We consider a nonlinear convex program. Under some general hypotheses, we prove that approximate solutions obtained by exponential penalty converge toward a particular solution of the original convex program as the penalty parameter goes to zero. This particular solution is called the absolute minimizer and is characterized as the unique solution of a hierarchical scheme of minimax problems.

Keywords. Convexity, minimax problems, penalty methods, nonuniqueness, optimal trajectory, convergence.

AMS 1991 subject classifications. 90C25, 90C31.

1 Introduction

Let us consider a mathematical program of the type:

$$(P) \qquad \min_{x\in I\!R^n} \{f_0(x) \mid \ f_i(x) \le 0,\ i=1,...,m\},$$

where for each $i=0,...,m$, f_i is a convex function. The exponential penalty method consists in solving for $r>0$ small enough the unconstrained problem

$$(P_r) \qquad \min_{x\in I\!R^n} \left\{ f_0(x) + r\sum_{i=1}^{m} \exp[f_i(x)/r] \right\}.$$

We denote by $x(r)$ an optimal solution of (P_r) and we regard it as an approximate solution of the original problem (P). Exponential penalty methods

*Partially supported by FONDECYT 1990884 and FONDAP Matemáticas Aplicadas.

†Depto. Ingeniería Matemática, Universidad de Chile, Casilla 170/3 Correo 3, Santiago, Chile. Email: falvarez@dim.uchile.cl.

have been widely applied since the pioneering work of Motzkin [12]. Generally speaking, the convergence as $r \to 0^+$ of $x(r)$ is well determined when (P) has a unique optimal solution. For constrained problems, it was introduced in [10] an *exponential multiplier* method (see also [4]); under uniqueness and second order sufficiency conditions, the convergence of this kind of algorithms can be established for several penalty functions (see [5, 14, 15]). Although it is possible to obtain dual convergence without these restrictive hypotheses (see [16]), primal convergence is not well understood in the case of multiple optimal solutions. For an implementable algorithm for solving convex programs by applying an exponential penalty technique, we refer the reader to [9]. A different approach is given in [13], where convergence is forced by combining the exponential penalty with a proximal regularization.

We are interested in the convergence of the whole *optimal path* $\{x(r) : r \to 0^+\}$ when (P) admits a multiplicity of optimal solutions. In this direction, it is proved in [6] that for linear programs, $x(r)$ converges toward the *centroid*, a sort of analytic center of the optimal polytope. A similar situation occurs for linear-quadratic minimax problems (see [1]). The aim of this paper is to extend the convergence results of [1, 6] to a more general nonlinear setting. Under a regularity condition on all the f_i's, we prove that the approximate solution $x(r)$ converges to a "distinguished" solution of (P), which is called the *absolute minimizer* and is characterized as the unique solution of a recursive hierarchy of reduced minimax problems. A similar selection of a particular solution appears in the L^p approximation of L^∞ problems (see [2]). See [3, 7, 11] for analogous convergence results for other path-following methods in linear and convex programming.

2 Absolute minimizer for convex minimax.

In this section we give a construction of a particular solution of a given minimax problem: the absolute minimizer. From now on, $I = \{1, ..., m\}$ and $f_i : I\!R^n \to I\!R$ is convex for each $i \in I$. Set $\bar{f}(x) := \max_{i \in I}\{f_i(x)\}$ and consider the minimax problem

$$(\widehat{P}) \qquad \mu^* := \min_{x \in F} \bar{f}(x).$$

where $F \subset I\!R^n$ is a closed convex set. We denote by $S(\widehat{P})$ the set of optimal solutions of $(\widehat{P})$, and we assume that $S(\widehat{P})$ is nonempty and compact. The question is how to distinguish a best optimal solution among all the elements in $S(\widehat{P})$. To this end, we define the set of optimal active indices by

$$I_0 := \{i \in I \mid \forall x \in S(\widehat{P}),\ f_i(x) = \mu^*\}.$$

It is easy to see that I_0 is nonempty. For each $i \in I_0$ the corresponding f_i is constant on $S(\widehat{P})$ and for such an f_i all the solutions are in some sense equivalent. If $I_0 = I$ there is nothing else to do. Otherwise, there exist $i_0 \in I \backslash I_0$ and an optimal solution $\widehat{x}$ such that $f_{i_0}(\widehat{x}) < \mu^*$; we consider

$$(\widehat{P}^1) \qquad \mu_1^* := \min_{x \in S(\widehat{P})} \max_{i \in I \backslash I_0} \{f_i(x)\},$$

in order to select the minimizers of $\max_{i \in I \backslash I_0}\{f_i\}$ among all the minimizers of $\max_{i \in I}\{f_i\}$. Of course, $\mu_1^* < \mu^*$ and $S(\widehat{P}^1)$ is nonempty and compact. Let $A := \{i \in I \backslash I_0 \mid \forall x \in S(\widehat{P}^1),\ f_i(x) = \mu_1^*\}$, which is nonempty. If the set $I_1 := I_0 \cup A$ is not equal to I, we can proceed recursively and consider the following minimax problem

$$(\widehat{P}^2) \qquad \mu_2^* = \min_{x \in S(\widehat{P}^1)} \max_{i \in I \backslash I_1} \{f_i(x)\}.$$

We continue in this manner obtaining a sequence of problems of the type:

$$(\widehat{P}^t) \qquad \mu_t^* = \min_{x \in S(\widehat{P}^{t-1})} \max_{i \in I \backslash I_{t-1}} \{f_i(x)\}.$$

By construction, we have a strictly increasing sequence of sets $I_0 \subset I_1 \subset ....$ Therefore, $I_{p+1} = \{1, ..., m\} = I$ for some $p \leq m$. Observe that for each $\widehat{x} \in S(\widehat{P}^p)$

$$\max_{i \in I \backslash I_t} \{f_i(\widehat{x})\} = \mu_{t+1}^* = \min_{x \in S(\widehat{P}^t)} \max_{i \in I \backslash I_t} \{f_i(x)\},$$

for all $t \in \{0, ..., p\}$, where $I_t \backslash I_{t-1} = \{i \in I \mid \forall x \in S(\widehat{P}^t), f_i(x) = \mu_t^*\}$ is the set of active indices for the problem $(\widehat{P}^{t-1})$. We consider $S(\widehat{P}^p)$ as the set of best optimal solutions of the original minimax problem $(\widehat{P})$. Similar constructions can be found in [2, 3, 6].

The final optimal set $S(\widehat{P}^p)$ depends on the analytical representation of $\bar{f} = \max_{i \in I}\{f_i\}$. For instance, take $\bar{f}(x) = |x|$ if $|x| > 1$ and $\bar{f}(x) = 1$ otherwise. Then $S(\widehat{P}) = [-1, 1]$. Setting $f_1 := 1/2$ and $f_2 := \bar{f}$, we can write $\bar{f} = \max\{f_1, f_2\}$ to obtain $S(\widehat{P}^1) = [-1, 1]$. But if we set $f_1(x) := 1$ and $f_2(x) := |x|$, then $S(\widehat{P}^1) = \{0\}$. As this trivial example illustrates, in general we cannot ensure uniqueness of the solution generated by the hierarchical process defined above. Nevertheless, we may overcome this disadvantage by restricting our analysis to a suitable class of max-type representations.

Definition 2.1 [2] *A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be* quasi-analytic *if whenever f is constant on the segment $[x, y]$ with $x \neq y$, then f is constant on the whole line passing through x and y.*

Affine, quadratic and analytic functions are quasi-analytic. If we assume that for each $i \in I$, f_i is quasi-analytic then there exists a unique solution x^* of the recursive hierarchy of minimax problems; to see this, fix $x_1, x_2 \in S(\widehat{P}^p)$ and note that each f_i is constant on $[x_1, x_2]$ so that $x_2 - x_1$ is a constancy direction for every f_i and consequently $x_2 = x_1$ (recall that $S(\widehat{P})$ is bounded). We call such x^* the *absolute minimizer* of $(\widehat{P})$.

3 Convergence toward the absolute minimizer.

Let us return to the convex program

$$(P) \qquad \alpha := \min_{x \in I\!R^n} \{f_0(x) \mid \ f_i(x) \leq 0, \ i \in I\}.$$

We assume that $f_0 : I\!R^n \to I\!R \cup \{+\infty\}$ is a closed proper convex function, and $f_i : I\!R^n \to I\!R$ are all convex. We tacitly assume that α is finite and we consider the penalty approximation

$$(P_r) \qquad \alpha(r) := \min_{x \in I\!R^n} \left\{ f_0(x) + r \sum_{i \in I} \exp[f_i(x)/r] \right\}.$$

When the optimal set $S(P)$ is nonempty and compact, it is well-known (see for instance [3]) that: (i) for every $r > 0$, the optimal set $S(P_r)$ is nonempty and compact; (ii) given $r_0 > 0$ there exists a bounded set U such that $S(P_r) \subset U$ for every $r \in]0, r_0]$; (iii) each cluster point of $\{x(r) : r \to 0^+\}$ with $x(r) \in S(P_r)$ belongs to $S(P)$ and $\alpha(r) \to \alpha$ as $r \to 0^+$. Certainly, (i)-(iii) give us the convergence of $x(r)$ when (P) has a unique solution. In the case of multiple optimal solutions, we have:

Theorem 3.1 *If $S(P)$ is nonempty and compact, and for all $i = 0, ..., m$, f_i is quasi-analytic, then for each $r > 0$ there exists a unique solution $x(r)$ of (P_r), and furthermore*

$$\lim_{r \to 0^+} x(r) = x^*,$$

where x^* is the absolute minimizer of

$$\min_{x \in S(P)} \max_{i \in I} \{f_i(x)\}. \qquad (1)$$

Proof. To establish the uniqueness of $x(r)$, fix $r > 0$ and let $x_1, x_2 \in S(P_r)$. For every $t \in]0, 1[$ set $x_t := (1-t)x_1 + tx_2$. If there exists $i_0 \in I$ such that $f_{i_0}(x_1) \neq f_{i_0}(x_2)$ then

$$\exp[f_{i_0}(x_t)/r] < (1-t)\exp[f_{i_0}(x_1)/r] + t \exp[f_{i_0}(x_2)/r],$$

and for $f(x,r) := f_0(x) + r\sum_{i\in I}\exp[f_i(x)/r]$ we obtain

$$f(x_t, r) < (1-t)f(x_1, r) + tf(x_2, r) = \alpha(r),$$

which is impossible. Therefore, for each $i \in I$ the function f_i is constant on $S(P_r)$. Hence, f_0 is also constant on $S(P_r)$ and the boundedness of $S(P_r)$ implies $x_2 = x_1$ as claimed.

The task is now to prove the convergence of $x(r)$. It suffices to show that $\{x(r) : r \to 0^+\}$ has as unique cluster point the absolute minimizer of (1). Let $r_k \to 0^+$ and $\widehat{x} \in S(P)$ be such that $x(r_k) \to \widehat{x}$ as $k \to +\infty$. Let $\overline{x} \in S(P)$ be arbitrary and set $x_k := x(r_k) + \overline{x} - \widehat{x}$. Thus $x_k \to \overline{x}$ as $k \to +\infty$. The optimality of $x(r_k)$ for (P_{r_k}) gives

$$f_0(x(r_k)) + r_k\sum_{i\in I}\exp[f_i(x(r_k))/r_k] \le f_0(x_k) + r_k\sum_{i\in I}\exp[f_i(x_k)/r_k].$$

Since $S(P)$ is convex, we have that f_0 is constant on the segment $[\overline{x}, \widehat{x}]$. It is simple to see that whenever a quasi-analytic convex function f is constant on a segment $[x, y]$ then $f(z + x - y) = f(z)$ for every $z \in I\!R^n$ (see [2]). Thus, the convexity and quasi-analyticity of f_0 yield

$$f_0(x(r_k)) = f_0(x(r_k) + \overline{x} - \widehat{x}) = f_0(x_k).$$

Therefore, it follows that

$$\sum_{i\in I}\exp[f_i(x(r_k))/r_k] \le \sum_{i\in I}\exp[f_i(x_k)/r_k].$$

Since for all $y \in I\!R^m$ we have that

$$\lim_{r\to 0^+, z\to y} r\ln\left(\sum_{i\in I}\exp[z_i/r]\right) = \max_{i\in I}\{y_i\},$$

we deduce by letting $k \to +\infty$ in the last inequality that

$$\max_{i\in I}\{f_i(\widehat{x})\} \le \max_{i\in I}\{f_i(\overline{x})\}.$$

Since $\overline{x} \in S(P)$ is arbitrary, we have that $\widehat{x}$ solves

$$(\widehat{P}) \qquad \mu^* = \min_{x\in S(P)}\max_{i\in I}\{f_i(x)\}.$$

Denote $I_0 := \{i \in I \mid \forall x \in S(\widehat{P}),\ f_i(x) = \mu^*\}$ and assume that $I_0 \neq I$. Let now $\overline{x} \in S(\widehat{P})$. We have that for all $i \in I_0$ and $t \in [0,1]$ $f_i(\widehat{x} + t(\overline{x} - \widehat{x})) = \mu^*$; hence

$$f_i(x(r_k)) = f_i(x(r_k) + \overline{x} - \widehat{x}).$$

Thus

$$\sum_{i\in I\setminus I_0} \exp[f_i(x(r_k))/r_k] \leq \sum_{i\in I\setminus I_0} \exp[f_i(x_k)/r_k].$$

Letting $k \to +\infty$ in the last inequality, we obtain

$$\max_{i\in I\setminus I_0} \{f_i(\widehat{x})\} \leq \max_{i\in I\setminus I_0} \{f_i(\overline{x})\}.$$

We conclude that $\widehat{x}$ solves

$$\mu_1^* = \min_{x\in S(P)} \max_{i\in I\setminus I_0} \{f_i(x)\}.$$

Repeated application of these arguments enables us to prove that $\widehat{x}$ solves the recursive hierarchy of minimax problems that define the absolute minimizer x^* of $(\widehat{P})$, which completes the proof. □

REMARK. In practice, (P_r) is solved only approximately. An interesting open question is the extension of this result to the case of inexact solutions.

Aknowledgements. I wish to thank the anonymous referees for helpful comments concerning the presentation of this paper.

References

[1] F. Alvarez: Métodos continuos en optimización paramétrica, Engineering memoir, Universidad de Chile (1996).

[2] H. Attouch, R. Cominetti: L^p approximation of variational problems in L^1 and L^∞, Nonlinear Anal. 36, no. 3, Ser. A : 373-399, 1999.

[3] A. Auslender, R. Cominetti and M. Haddou: Asymptotic analysis for penalty methods in convex and linear programming, Math. Oper. Res. 22, No.1 : 43-62, 1997.

[4] D.P. Bertsekas: Approximation procedures based on the method of multipliers, J. Optim. Theory Appl., 23 : 487-510, 1977.

[5] D.P. Bertsekas: Constrained Optimization and Lagrange Multiplier Methods, Academic Press, New York, 1981.

[6] R. Cominetti, J. San Martín: Asymptotic analysis of the exponential penalty trajectory in linear programming, Math. Programming 67 : 169-187, 1994.

[7] C.G. Gonzaga: Path-following methods for linear programming, SIAM Review 34 : 167-224, 1992.

[8] N.I.M. Gould: On the convergence of a sequential penalty function method for constrained minimization, SIAM J. Numer. Anal. 26 : 107-128, 1989.

[9] M. Grigoriadis, L. Khachiyan: An exponential-function reduction method for block-angular convex programs. Networks 26, no. 2 : 59-68, 1995.

[10] B.W. Kort, D.P. Bertsekas: A new penalty function method for constrained minimization, in Proceedings of the IEEE conference on decision and control (New Orleans, 1972) : 162-166, 1972.

[11] N. Megiddo: Pathways to the optimal set in linear programming, in Progress in Math. Prog.: interior-point and related methods, Springer-Verlag : 131-158, 1989.

[12] T.S. Motzkin: New technique for linear inequalities and optimization, in Project SCOOP Symposium on Linear Inequalities and Programming, Planning Research Division, U.S. Air Force, Washington D.C., 1952.

[13] K. Mouallif, P. Tossings: Une méthode de pénalisation exponentielle associée à une régularisation proximale, Bull. Soc. Roy. Sci. Liege 56, no. 2 : 181-192, 1987.

[14] R.A. Polyak: Smooth optimization methods for minimax problems, SIAM J. Control and Optimization ,Vol. 26 No. 6 : 1274-1286, 1988.

[15] J.J. Strodiot, V.H. Nguyen: An exponential penalty method for nondifferentiable minimax problems with general constraints, JOTA 27 : 205-219, 1979.

[16] P. Tseng, D.P. Bertsekas: On the convergence of the exponential multiplier method for convex programming, Math. Programming 60 : 1-19, 1993.