Reliable Bounds for Convex Relaxation in Interval Global Optimization Codes

Frédéric Messine[1,a)] and Gilles Trombettoni[1,b)]

[1]*LAPLACE-ENSEEIHT-INPT, University of Toulouse, Toulouse, France.*
[2]*LIRMM, University of Montpellier, CNRS, France.*

[a)]Frederic.Messine@laplace.univ-tlse.fr
[b)]Gilles.Trombettoni@lirmm.fr

Abstract. In order to obtain reliable deterministic global optima, all the computed bounds have to be certified in a way that no numerical error due to floating-point operations can discard a feasible solution. Interval arithmetic Branch and Bound algorithms which are developed since the 1980th possess this property of reliability. However, some new accelerating techniques, such as convex relaxation, could improve the convergence of those reliable global optimization algorithms while keeping the property of reliability. In this work, we show that a floating-point solution obtained by solving a relaxed convex program can be corrected in order to certify that this new lower bound is lower than the real global optimum.

INTRODUCTION

The use of automatic convex relaxations inside global optimization software such as `BARON` or `COUENNE` is extremely powerful. Indeed, they make it possible to compute efficient lower bounds by solving a convex program with a local solver such as `ipopt`; note that if the program is convex with the objective function stricktly convex, a local solution is also a global one (unicity property). Unfortunately, the so-obtained local optima yield lower bounds that are not numerically reliable. In interval Branch&Bound algorithms such as `Ibex`, `IBBA` and `GlobSol`, it is necessary to be rigorous in a sense that all the numerical floating point computations have to be reliable.

This paper provides two theorems that make it possible to correct the bounds computed by a local solver applied to relaxed convex programs. Remark that in [1], Jansson provided the same theorem for the constrained convex case. However, our proof is different and comes from the KKT-conditions. Furthermore, in this paper, we detail the entire procedure to compute reliable lower bounds using interval arithmetic computations.

In the next two sections, we provide two theorems yielding reliable bounds for unconstrained and constrained convex programs. In the following section, we present a way to numerically compute those reliable bounds by using interval arithmetic, before concluding.

RELIABLE LOWER BOUNDS FOR BOX-CONSTRAINED CONVEX PROBLEM

Consider the following convex optimization problem with only bound constraints:

$$(\mathcal{UP})\begin{cases} \min\limits_{x\in\mathbb{R}^n} & f(x) \\ \text{s.t.} & \\ & \underline{x_i} \le x_i \le \overline{x_i},\ \forall i \in \{1,\cdots,n\}, \end{cases}$$

where f is a real twice differentiable convex function in X; where $X = \prod_{i=1}^{n}[\underline{x_i},\overline{x_i}]$. Let x^* denote the global solution of problem $(\mathcal{UP})$.

Theorem 1 *For all $\tilde{x} \in X$, one obtains:*

$$f(x^*) \geq f(\tilde{x}) + (z - \tilde{x})^T.\nabla f(\tilde{x}),$$

where

$$z_i := \begin{cases} \overline{x_i} \text{ if } \frac{\partial f}{\partial x_i}(\tilde{x}) \leq 0 \\ \underline{x_i} \text{ otherwise,} \end{cases}$$

Proof:
By using a Taylor expansion of f at the second order, we obtain:

$$\forall y \in X, \exists \xi \in X, \text{ such that}, f(y) = f(\tilde{x}) + (y - \tilde{x})^T.\nabla f(\tilde{x}) + \frac{1}{2}(y - \tilde{x})^T.H(\xi).(y - \tilde{x}),$$

where $H(\xi)$ is the Hessian matrix which is computed at point $\xi \in X$.

Because f is convex by assumption, $H(x)$ is positive definite $\forall x$ in X and then, we have $(y - \tilde{x})^T.H(\xi).(y - \tilde{x}) \geq 0$ for all $y \in X$ and $\tilde{x} \in X$.

Hence,

$$f(y) \geq f(\tilde{x}) + (y - \tilde{x})^T.\nabla f(\tilde{x}), \forall y \in X, \forall \tilde{x} \in X. \tag{1}$$

Note that $(y - \tilde{x})^T.\nabla f(\tilde{x}) = \sum_{i=1}^{n} (y_i - \tilde{x}_i).\dfrac{\partial f}{\partial x_i}(\tilde{x})$, and then by taking $z_i := \overline{x_i}$ if $\frac{\partial f}{\partial x_i}(\tilde{x}) \leq 0$ and $z_i := \underline{x_i}$ otherwise, we obtain the following inequalities.

Thus, if $\frac{\partial f}{\partial x_i}(\tilde{x}) \leq 0$, one has for all $(y_i, \tilde{x}_i) \in X_i^2$:

$$\begin{aligned} & z_i := \overline{x_i} \geq y_i, \\ & z_i - \tilde{x}_i \geq y_i - \tilde{x}_i, \\ & (z_i - \tilde{x}_i)\frac{\partial f}{\partial x_i}(\tilde{x}) \leq (y_i - \tilde{x}_i)\frac{\partial f}{\partial x_i}(\tilde{x}). \end{aligned}$$

Now, if $\frac{\partial f}{\partial x_i}(\tilde{x}) \geq 0$, one has for all $(y_i, \tilde{x}_i) \in X_i^2$:

$$\begin{aligned} & z_i := \underline{x_i} \leq y_i, \\ & z_i - \tilde{x}_i \leq y_i - \tilde{x}_i, \\ & (z_i - \tilde{x}_i)\frac{\partial f}{\partial x_i}(\tilde{x}) \leq (y_i - \tilde{x}_i)\frac{\partial f}{\partial x_i}(\tilde{x}). \end{aligned}$$

Therefore in both cases, we obtain:

$$(y_i - \tilde{x}_i)\frac{\partial f}{\partial x_i}(\tilde{x}) \geq (z_i - \tilde{x}_i)\frac{\partial f}{\partial x_i}(\tilde{x}), \forall (y_i, \tilde{x}_i) \in X_i^2, \forall i \in \{1, \cdots, n\}. \tag{2}$$

Hence, by using (1) and 2, we obtain

$$f(y) \geq f(\tilde{x}) + (z - \tilde{x})^T.\nabla f(\tilde{x}), \forall y \in X, \forall \tilde{x} \in X, \tag{3}$$

and then the result follows.

□

Remark 1 *Note that if $\tilde{x}$ is the floating point solution provided by a local search algorithm, if all the numerical computations behave well we obtain $x^* \simeq \tilde{x}$ and then the correcting negative term $(z - \tilde{x})^T.\nabla f(\tilde{x})$ is close to zero. This is the term which makes it possible to provide a reliable lower bound on $f(x^*)$ from a point $\tilde{x} \in X$.*

Remark 2 *Note that if the problem is not box-constrained, the reliable lower bound is $-\infty$.*

RELIABLE LOWER BOUNDS FOR CONSTRAINED CONVEX PROBLEM

Consider now the following constrained convex optimization problem:

$$(\mathcal{CP}) \begin{cases} \min\limits_{x\in \boldsymbol{R}^n} & f(x) \\ \text{s.t.} & \\ & g(x) \leq 0 \\ & Ax = b \\ & \underline{x_i} \leq x_i \leq \overline{x_i}, \forall i \in \{1, \cdots, n\} \end{cases}$$

where f and g are twice differentiable convex functions defined over $\boldsymbol{R}^n$ but g has its value in $\boldsymbol{R}^p$. A is a real matrix of size $q \times n$ and b a real vector of q components. In fact, one defines an optimization problem ($\mathcal{CP}$) of a convex function with p convex inequality constraints and q linear equality ones. Let us denote by x^* the global solution of problem ($\mathcal{UP}$) and $\mathcal{L}(x, \mu, \lambda) := f(x) + \mu^T.g(x) + \lambda^T.(Ax - b)$ the Lagrangian function with $\lambda \in \boldsymbol{R}^n$ and $\mu \in (\boldsymbol{R}^n)^+$ its corresponding multipliers.

Theorem 2 *For all $(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) \in (X, (\boldsymbol{R}^n)^+, \boldsymbol{R}^n)$, one obtains:*

$$f(x^*) \geq \mathcal{L}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) + (z - \tilde{x})^T.\nabla_x\mathcal{L}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}).$$

where $z_i := \overline{x_i}$ if $\frac{\partial \mathcal{L}}{\partial x_i}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) \leq 0$ and $z_i := \underline{x_i}$ otherwise. By developing, one has:

$$f(x^*) \geq f(\tilde{x}) + \tilde{\mu}^T.g(\tilde{x}) - \tilde{\lambda}^T.b + (z - \tilde{x})^T.\left(\nabla_x f(\tilde{x}) + \tilde{\mu}^T.\nabla_x g(\tilde{x})\right) + z^T.A^T\tilde{\lambda}.$$

Proof:
By using a Taylor expansion of the Lagrangian function at the second order, for all $(y, \mu, \lambda) \in (X, (\boldsymbol{R}^n)^+, \boldsymbol{R}^n)$ and for all $\forall(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) \in (X, (\boldsymbol{R}^n)^+, \boldsymbol{R}^n)$, there exists $(\xi, \hat{\mu}, \hat{\lambda}) \in (X, (\boldsymbol{R}^n)^+, \boldsymbol{R}^n)$, such that one has:

$$\mathcal{L}(y, \mu, \lambda) = \mathcal{L}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) + (y - \tilde{x})^T.\nabla_x\mathcal{L}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) + \frac{1}{2}(y - \tilde{x})^T.H_x(\xi, \hat{\mu}, \hat{\lambda}).(y - \tilde{x}),$$

where $\nabla_x\mathcal{L}$ is the Gradient of $\mathcal{L}$ w.r.t. the variables x and $H_x(\xi, \hat{\mu}, \hat{\lambda})$ is the Hessian matrix of $\mathcal{L}$ w.r.t. x at the point $\xi, \hat{\mu}, \hat{\lambda}$.

Because f and g are convex and $Ax - b = 0$ are linear equality constraints, $H_x(x, \mu, \lambda)$ is positive definite for all x in X, for all $\mu \in (\boldsymbol{R}^n)^+$ and for all $\lambda \in \boldsymbol{R}^n$ yielding $(y - \tilde{x})^T.H_x(\xi, \hat{\mu}, \hat{\lambda}).(y - \tilde{x}) \geq 0$, for all $y \in X$ and all $\tilde{x} \in X$. Hence, for all $(y, \mu, \lambda) \in (X, (\boldsymbol{R}^n)^+, \boldsymbol{R}^n)$ and all $(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) \in (X, (\boldsymbol{R}^n)^+, \boldsymbol{R}^n)$, one has:

$$\mathcal{L}(y, \mu, \lambda) \geq \mathcal{L}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) + (y - \tilde{x})^T.\nabla_x\mathcal{L}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}).$$

Because $(y-\tilde{x})^T.\nabla_x\mathcal{L}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) = \sum_{i=1}^n (y_i - \tilde{x}_i).\frac{\partial \mathcal{L}}{\partial x_i}(\tilde{x}, \tilde{\mu}, \tilde{\lambda})$, the following lower bound can be derived for all $(y, \mu, \lambda) \in (X, (\boldsymbol{R}^n)^+, \boldsymbol{R}^n)$ and all $(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) \in (X, (\boldsymbol{R}^n)^+, \boldsymbol{R}^n)$:

$$\mathcal{L}(y, \mu, \lambda) \geq \mathcal{L}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) + (z - \tilde{x})^T.\nabla_x\mathcal{L}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}). \tag{4}$$

where $z_i := \overline{x_i}$ if $\frac{\partial \mathcal{L}}{\partial x_i}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) \leq 0$ and $z_i := \underline{x_i}$ otherwise.

The first part of the proof, derives directly from Equation (4) because if x^* is a solution of problem ($\mathcal{CP}$) then the constraints are satisfied and therefore, $f(x^*) = \mathcal{L}(x^*, \mu, \lambda)$, for all $\mu \in (\boldsymbol{R}^n)^+$ and all $\lambda \in \boldsymbol{R}^n$.

From equation (4), one obtains:

$$f(x^*) \geq \mathcal{L}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}) + (z - \tilde{x})^T.\nabla_x\mathcal{L}(\tilde{x}, \tilde{\mu}, \tilde{\lambda}),$$

by developing, one has

$$\begin{aligned} f(x^*) &\geq f(\tilde{x}) + \tilde{\mu}^T.g(\tilde{x}) + \tilde{\lambda}^T.(A\tilde{x} - b) + (z - \tilde{x})^T.\nabla_x\left(f(\tilde{x}) + \tilde{\mu}^T.g(\tilde{x}) + \tilde{\lambda}^T.(A\tilde{x} - b)\right), \\ &\geq f(\tilde{x}) + \tilde{\mu}^T.g(\tilde{x}) + \tilde{\lambda}^T.(A\tilde{x} - b) + (z - \tilde{x})^T.\left(\nabla_x f(\tilde{x}) + \tilde{\mu}^T.\nabla_x g(\tilde{x}) + A^T\tilde{\lambda}\right), \\ &\geq f(\tilde{x}) + \tilde{\mu}^T.g(\tilde{x}) - \tilde{\lambda}^T.b + (z - \tilde{x})^T.\left(\nabla_x f(\tilde{x}) + \tilde{\mu}^T.\nabla_x g(\tilde{x})\right) + z^T.A^T\tilde{\lambda}. \end{aligned}$$

□

Remark 3 *Although the connection is not direct, Theorem 2 corresponds to Lemma 1 in [1], where another proof of this result can be found using linear programming.*

RELIABLE INTERVAL COMPUTATIONS FOR THE LOWER BOUNDS

The two previous theoretical results provide bounds for a convex optimization problem. However, on a computer all the computations are based on floating-point operations and and these operations also introduce numerical errors on the correction terms. Thus, all those computations have to be done using interval arithmetic. By denoting $[y]$ the smallest floating-point interval containing the real value y, by extension $[v]$ the thinest floating-point interval vector of a real vector v, $[A]$ the thinest floating-point interval matrix containing the real matrix A, and $\underline{I}$ the lower bound of the interval I, respectively $\overline{I}$ the upper bound of the interval I. Assume, that $\tilde{x}_i, \underline{x_i}, \overline{x_i}$ are all floating point numbers, we have the following reliable floating-point bounds:

Corollary 1

- *Box-constrained case:*

$$f(x^*) \geq \underline{[f([\tilde{x}])] + (\mathbf{z} - [\tilde{x}])^T.[\nabla f([\tilde{x}])]} \left(= \underline{[f([\tilde{x}])] + \sum_{i=1}^{n} (\mathbf{z_i} - [\tilde{x}_i]) \times \left[\frac{\partial f([\tilde{x}])}{\partial x_i}\right]} \right)$$

where $\mathbf{z_i} := [\overline{x_i}, \overline{x_i}]$ *if* $\overline{\left[\frac{\partial f([\tilde{x}])}{\partial x_i}\right]} \leq 0$, $\mathbf{z_i} := [\underline{x_i}, \underline{x_i}]$ *if* $\underline{\left[\frac{\partial f([\tilde{x}])}{\partial x_i}\right]} \geq 0$ *and* $\mathbf{z_i} := [\underline{x_i}, \overline{x_i}]$ *otherwise.*

- *Constrained case:*

$$f(x^*) \geq \underline{[f([\tilde{x}])] + [\tilde{\mu}]^T.[g(\tilde{x})] - [\tilde{\lambda}]^T.[b] + (\mathbf{z} - [\tilde{x}])^T.\left([\nabla_x f([\tilde{x}])] + [\tilde{\mu}]^T.[\nabla_x g([\tilde{x}])]\right) + \mathbf{z}^T[\tilde{\lambda}]^T.[A]}$$

where $\mathbf{z_i} := [\overline{x_i}, \overline{x_i}]$ *if* $\overline{\left[\frac{\partial \mathcal{L}}{\partial x_i}([\tilde{x}], [\tilde{\mu}], [\tilde{\lambda}])\right]} \leq 0$, $\mathbf{z_i} := [\underline{x_i}, \underline{x_i}]$ *if* $\underline{\left[\frac{\partial \mathcal{L}}{\partial x_i}([\tilde{x}], [\tilde{\mu}], [\tilde{\lambda}])\right]} \geq 0$ *and* $\mathbf{z_i} := [\underline{x_i}, \overline{x_i}]$ *otherwise.*

Proof:
The proof comes directly from the theorems by using correct rounding in the uses of the interval computations. A particular attention must be paid on the gradient computations which have to be evaluated using interval arithmetic and thus, this provides an interval floating-point vector which includes the real values of the partial derivatives. Hence, the definition of $\mathbf{z_i}$ *has now three possibilities if the partial derivative is always negative, positive or if the sign is not assigned in this step using interval computations.*

□

CONCLUSION

We have presented in this paper correction theorems that make it possible to compute reliable lower bounds for unconstrained and constrained convex optimization programs. We will validate that the computation of the correction term is not expensive. Furthermore, if the point returned by local search will be close to the real global optimum, we expect that the correction term will be very small and adds only a few iterations in a Branch&Bound algorithm. One of the main purpose of our presentation is to provide first numerical results using this reliable convex relaxation method.

REFERENCES

[1] C. Jansson, *A Rigourous Lower Bound for the Optimal Value of Convex Otimization Problems*, Journal of Global Optimization, Vol. 28, N.1, pp.121–137, 2004.