Estimating convexifiers in continuous optimization

SANJO ZLOBEC*

Abstract. *Every function of several variables with the continuous second derivative can be convexified (i.e., made convex) by adding to it a quadratic "convexifier". In this paper we give simple estimates on the bounds of convexifiers. Using the idea of convexification, many problems in applied mathematics can be reduced to convex mathematical programming models. This is illustrated here for nonlinear programs and systems of nonlinear equations.*

Key words: *twice continuously differentiable function, convexifier, nonlinear programming, convex model, system of nonlinear equations*

AMS subject classifications: 90C30, 90C31

Received February 14, 2003 Accepted March 15, 2003

1. Introduction

Every function $f : \mathbb{R}^n \to \mathbb{R}$ with the continuous second derivative (i. e., a C^2 function) can be "convexified" on a bounded convex set C. This is achieved by adding a simple quadratic convexifier $c(\boldsymbol{x}) = \gamma \boldsymbol{x}^T \boldsymbol{x}$ to f, where γ is an appropriate sufficiently large non-negative number. This idea has a wide range of applications ranging from Tikhonov's regularization of ill-posed problems, e.g. [4], to numerical methods of nonlinear programming. In particular, using a convexification, Liu and Floudas [3] have observed that every nonlinear program with C^2 functions is equivalent to a "partly convex" program (using the terminology from, e.g. [5]). Furthermore, it has been recently shown in [6], see also [5], that an arbitrary program with C^2 functions is equivalent to a "convex programming model". Moreover, such model is "stable" (i.e., the feasible set mapping is lower semi-continuous) at the optimal solution. This means that the study of arbitrary nonlinear programs with C^2 functions can be essentially reduced to the study of only linear programs, convex programs, and the relationship between the two. In this paper we show that it is relatively easy to obtain practical estimates for quadratic convexifiers.

*Department of Mathematics and Statistics, McGill University, Burnside Hall, 805 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6, e-mail: `zlobec@math.mcgill.ca`

2. Estimates for quadratic convexifiers

For a function $f : \mathbb{R}^n \to \mathbb{R}$ we denote the first derivative at $\boldsymbol{x}$ by $\nabla f(\boldsymbol{x})$ and the second derivative by $\nabla^2 f(\boldsymbol{x})$. Assuming that these objects exist, one can represent them by the gradient and the Hessian of f at $\boldsymbol{x}$, respectively. In our notation, the gradient is a row n-tuple, while the Hessian is an $n \times n$ symmetric matrix. Denote the eigenvalues of $\nabla^2 f(\boldsymbol{x})$ by $\lambda_i(\boldsymbol{x})$, i.e.,

$$\lambda_i(\boldsymbol{x}) = \lambda_i(\nabla^2 f(\boldsymbol{x})), \qquad i = 1, \dots, n$$

and their infimum on a given bounded convex set C by

$$\lambda^* = \inf_{\boldsymbol{x} \in C} \min_{i=1,\dots,n} \lambda_i(\boldsymbol{x}).$$

The following result, used in [3], is included here for the sake of completeness:

Lemma 1 [Convexification Lemma]. *If $f : \mathbb{R}^n \to \mathbb{R}$ is a twice continuously differentiable function on a bounded convex set C, then for every $\gamma \le \lambda^*$*

$$\varphi(\boldsymbol{x}) = f(\boldsymbol{x}) - \frac{1}{2}\gamma \boldsymbol{x}^T \boldsymbol{x}$$

is a convex function (convexification of f) on C.

Proof. The eigenvalues of the second derivative $\nabla^2 \varphi(\boldsymbol{x}) = \nabla^2 f(\boldsymbol{x}) - \gamma \boldsymbol{I}$ satisfy

$$\lambda_i(\nabla^2 f(\boldsymbol{x})) - \gamma \ge \lambda_i(\nabla^2 f(\boldsymbol{x})) - \lambda^* \ge 0, \quad i = 1, \dots, n$$

for $\gamma \le \lambda^*$. This means that $\nabla^2 \varphi(\boldsymbol{x})$ is positive semi-definite and hence $\varphi(\boldsymbol{x})$ is convex. □

The essence of convexification is illustrated by the following example.

Example 1. *Consider the scalar function $f(t) = -t^4$ over $-1 \le t \le 1$. This function is strictly concave. Since*

$$\lambda_i(\nabla^2 f(t)) = -12t^2, \quad i = 1, \dots n,$$

we have $\lambda^* = -12$, a convexifier is $c(t) = 6t^2$ and a convexification $\varphi(t) = -t^4 + 6t^2$. The graphs of $f(t)$ and its convexification $\varphi(t)$ are depicted in Figure 1. *Note that the convexification is not convex outside the prescribed interval.*

Similarly, one can introduce

$$\Lambda^* = \sup_{\boldsymbol{x} \in C} \max_{i=1,\dots,n} \lambda_i(\boldsymbol{x})$$

and perform a "concavization" of a function. Indeed, if $f : \mathbb{R}^n \to \mathbb{R}$ is a twice continuously differentiable function on a bounded convex set C, then $f(\boldsymbol{x}) - \frac{1}{2}\beta \boldsymbol{x}^T \boldsymbol{x}$ is a concave function over $\boldsymbol{x} \in C$ for every $\beta \ge \Lambda^*$. One can think of λ^* and Λ^* as extreme global eigenvalues of the second derivative of f over C.

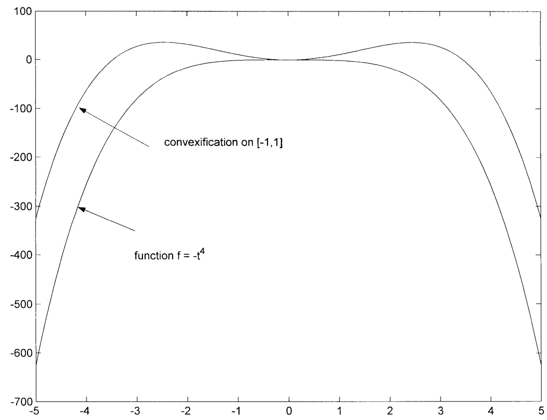

Figure 1. *Function f and its convexification φ*

From now on we will focus on the bounds λ^* and Λ^* and give their estimates. At this point let us make a quick reference to the Rayleigh quotient from linear algebra. This quotient gives a lower and an upper bound for the smallest and largest eigenvalues of a given symmetric matrix. These estimates are quickly obtained often after only several randomly chosen approximations and they are "pretty accurate". Our problem is more complicated, because we do not estimate the eigenvalues of a constant matrix but rather of the second derivative matrix that generally depends on the points of C. In order to avoid calculation of the second derivative of f at various points of C, we will work with its convexification $\varphi(\boldsymbol{x})$. In this way we can use only the first derivative or avoid derivatives all together.

The function that will replace the need for calculating the second derivative is

$$\psi(\boldsymbol{x},\boldsymbol{y}) = \frac{(\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}))(\boldsymbol{x}-\boldsymbol{y})}{\|\boldsymbol{x}-\boldsymbol{y}\|^2}.$$

Note that $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is defined for every $\boldsymbol{x}$ and $\boldsymbol{y}$, $\boldsymbol{x} \neq \boldsymbol{y}$. We will use the notation $(\boldsymbol{u},\boldsymbol{v}) = \boldsymbol{u}^T\boldsymbol{v}$ for the Euclidean inner product and $\|\boldsymbol{u}\| = (\boldsymbol{u}^T\boldsymbol{u})^{1/2}$ for the Euclidean norm. The function ψ is symmetric, i.e., $\psi(\boldsymbol{x},\boldsymbol{y}) = \psi(\boldsymbol{y},\boldsymbol{x})$, so the points in the domain of ψ do not have to be ordered.

Remark 1. *If one fixed an $\boldsymbol{x}$ and chose $\boldsymbol{y} = \boldsymbol{x} + \delta\boldsymbol{d}$, for some $\delta > 0$, and $\boldsymbol{d}$ such that $\|\boldsymbol{d}\| = 1$, then one could show, using the directional derivative, that $\lim_{\boldsymbol{y}\to\boldsymbol{x}} \psi(\boldsymbol{x},\boldsymbol{y}) = (\boldsymbol{d}, \nabla^2 f(\boldsymbol{x})\boldsymbol{d})$. In particular, if $\boldsymbol{d} = (1,0,\dots,0)^T$, then $\lim_{\boldsymbol{y}\to\boldsymbol{x}} \psi(\boldsymbol{x},\boldsymbol{y}) = \frac{\partial^2 f(\boldsymbol{x})}{\partial^2 \boldsymbol{x}}$, etc. Hence one could define the function $\psi(\boldsymbol{x},\boldsymbol{y})$ also for $\boldsymbol{y} = \boldsymbol{x}$. However, we will not need this limiting case.*

Values of the function $\psi(\boldsymbol{x},\boldsymbol{y})$ can be calculated for any two points $\boldsymbol{x}$ and $\boldsymbol{y}$ on a grid spread over the convex set C. Hence the global bounds can be estimated after calculating only the first derivatives of f:

Theorem 1. *Consider an arbitrary twice continuously differentiable function* $f : \mathbb{R}^n \to \mathbb{R}$ *and a bounded convex set* C*. Then*

$$\lambda^* \le \psi(\boldsymbol{x}, \boldsymbol{y}) \le \Lambda^*$$

for every $\boldsymbol{x} \in C$ *and* $\boldsymbol{y} \in C$, $\boldsymbol{y} \neq \boldsymbol{x}$.

Proof. We know that the function $\varphi(\boldsymbol{x}) = f(\boldsymbol{x}) - \frac{1}{2}\lambda^* \boldsymbol{x}^T \boldsymbol{x}$ is a convexification of f on C. This means that $\varphi(\boldsymbol{x})$ is convex. Therefore

$$(\nabla\varphi(\boldsymbol{x}) - \nabla\varphi(\boldsymbol{y}))(\boldsymbol{x} - \boldsymbol{y}) \ge 0$$

for every $\boldsymbol{x} \in C$ and $\boldsymbol{y} \in C$, $\boldsymbol{y} \neq \boldsymbol{x}$. This is

$$(\nabla^T f(\boldsymbol{x}) - \lambda^* \boldsymbol{x} - \nabla f^T(\boldsymbol{y}) + \lambda^* \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}) = (\nabla^T f(\boldsymbol{x}) - \nabla^T f(\boldsymbol{y}) - \lambda^*(\boldsymbol{x} - \boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y}) \ge 0$$

proving the left-hand side inequality. Similarly, one uses the concavification $f(\boldsymbol{x}) - \frac{1}{2}\Lambda^* \boldsymbol{x}^T \boldsymbol{x}$ to prove the right-hand side inequality. □

The bounds can be estimated without the derivative:

Theorem 2. *Consider an arbitrary twice continuously differentiable function* $f : \mathbb{R}^n \to \mathbb{R}$ *and a bounded convex set* C*. Then*

$$\lambda^* \le \{2/[\lambda(1-\lambda)\|\boldsymbol{x} - \boldsymbol{y}\|^2]\}\{\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) - f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y})\} \le \Lambda^*$$

for every $\boldsymbol{x} \in C$ *and* $\boldsymbol{y} \in C$, $\boldsymbol{y} \neq \boldsymbol{x}$, *and for every* $0 < \lambda < 1$.

Proof. Since $\varphi(\boldsymbol{x}) = f(\boldsymbol{x}) - \frac{1}{2}\lambda^* \boldsymbol{x}^T \boldsymbol{x}$ is convex, we know that

$$\varphi(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \le \lambda\varphi(\boldsymbol{x}) + (1-\lambda)\varphi(\boldsymbol{y}), \text{for every } 0 \le \lambda \le 1.$$

This is, after substitution,

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) - \frac{1}{2}\lambda^* \|\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\|^2 \le \lambda\Big\{ f(\boldsymbol{x}) - \frac{1}{2}\lambda^* \|\boldsymbol{x}\|^2 \Big\} \\ + (1-\lambda)\Big\{ f(\boldsymbol{y}) - \frac{1}{2}\lambda^*) \|\boldsymbol{y}\|^2 \Big\}. \end{aligned}$$

Hence

$$\begin{aligned} & f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) - \lambda f(\boldsymbol{x}) - (1-\lambda) f(\boldsymbol{y}) \\ & \le \frac{1}{2}\lambda^* \Big\{ \|\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\|^2 - \lambda\|\mathbf{x}\|^{\mathbf{2}} - (\mathbf{1} - \lambda)\|\boldsymbol{y}\|^{\mathbf{2}} \Big\} \\ & = \frac{1}{2}\lambda^* \lambda(\lambda - 1)\|\boldsymbol{x} - \boldsymbol{y}\|^2 \end{aligned}$$

after squaring and rearranging. Now division by $\lambda(\lambda-1)\|\boldsymbol{x}-\boldsymbol{y}\|^2 < 0$ yields the left-hand side inequality. Similarly, one uses a concavification to obtain the right-hand side bound. □

Corollary 1. *Consider an arbitrary twice continuously differentiable function* $f : \mathbb{R}^n \to \mathbb{R}$ *and a bounded convex set* C*. Then, for every* $\boldsymbol{x} \in C$ *and* $\boldsymbol{y} \in C$, $\boldsymbol{y} \neq \boldsymbol{x}$, *we have*

$$\lambda^* \le [8/\|\boldsymbol{x} - \boldsymbol{y}\|^2]\{[f(\boldsymbol{x}) + f(\boldsymbol{y})]/2 - f((\boldsymbol{x} + \boldsymbol{y})/2)\} \le \Lambda^*.$$

Proof. Specify $\lambda = 1/2$ in *Theorem 2.* □

Remark 2. *The above results give only rough estimates of the bounds of convexifiers. Typically these bounds are not reached on C. In practice one should go below the estimates for λ^* and above the estimates for Λ^*. Generally, the smaller the set C and the denser the mesh, the better the bounds.*

3. Applications

The motivation for our study of convexifiers is to gain more information about optimal solutions of programs with arbitrary C^2 functions after reducing these programs to convex programs. Consider the "classical" convex programming problem

$$\text{(CP)} \quad \begin{array}{l} \text{Min } f(\boldsymbol{x}) \\ \\ \boldsymbol{x} \in C \end{array}$$

Here we assume that $f : \mathbb{R}^n \to \mathbb{R}$ is a differentiable *convex* function and C is a compact convex set in $\mathbb{R}^n$. Every local optimum of (CP) is also its global optimum, so we can talk only about "optimal solutions". Now we consider (CP) but instead of assuming that the objective f is convex and differentiable, we assume that it is just a C^2 function (possibly non-convex). In this case we cannot characterize optimality of a feasible point $\boldsymbol{x}^*$ in a constructive way. (There do exist various necessary and various sufficient conditions of optimality, but not characterizations; see, e.g. [5].) After a convexification of f one can obtain interesting results. First we make two observations:

Observation 1. Consider an arbitrary twice continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ and a convex set $C \in \mathbb{R}^n$. Then *for every $\boldsymbol{\theta} \in C$*, the function $c(\cdot, \boldsymbol{\theta}) : \mathbb{R}^n \to \mathbb{R}$ defined by

$$c(\boldsymbol{x}, \boldsymbol{\theta}) = \frac{1}{2}\alpha(\boldsymbol{\theta}^T\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{x})$$

is a convexifier of f on C for any $\alpha \leq \lambda^*$. This is easy to see using the fact that, for any $\boldsymbol{\theta}$ and for any $\alpha \leq \lambda^*$, the function $\varphi(\boldsymbol{x}, \boldsymbol{\theta}) = f(\boldsymbol{x}) + \frac{1}{2}\alpha(\boldsymbol{\theta}^T\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{x})$ is convex on C. The linear term is of no importance for convexification.

Observation 2. (Due to Liu and Floudas [3]) Consider the two programs

(P) Min $f(\boldsymbol{x})$, subject to $\boldsymbol{x} \in C$

and

$(C, \boldsymbol{\theta})$ Min $\varphi(\boldsymbol{x}, \boldsymbol{\theta}) = f(\boldsymbol{x}) + \frac{1}{2}\alpha(\boldsymbol{\theta}^T\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{x})$, subject to $\boldsymbol{x} - \boldsymbol{\theta} = \mathbf{0}$, $\boldsymbol{x} \in C$

for some $\alpha \leq \lambda^*$, where $f(\boldsymbol{x})$ is twice continuously differentiable and C is a compact convex set. Note that (P) is an optimization problem in $\boldsymbol{x}$, while $(C, \boldsymbol{\theta})$ is an optimization problem simultaneously in $\boldsymbol{x}$ and $\boldsymbol{\theta}$. (Hence $(C, \boldsymbol{\theta})$ has twice as many variables as (P)). Suppose that $f(\boldsymbol{x})$ has a unique global minimum in C. Then $\boldsymbol{x}^* \in C$ is the global minimum of (P) if, and only if, $\boldsymbol{x}^*$ and $\boldsymbol{\theta}^*$ is an optimal solution of $(C, \boldsymbol{\theta})$, and $\boldsymbol{\theta}^* = \boldsymbol{x}^*$.

Using the above observations we conclude that the program (P) of minimizing an arbitrary twice continuously differentiable function on a convex set can be solved using $(C, \boldsymbol{\theta})$. A stable formulation of $(C, \boldsymbol{\theta})$ has been recently given in [6] as

$$(C,\boldsymbol{\theta};\varepsilon) \quad \begin{array}{l} \text{Min } \varphi(\boldsymbol{x},\boldsymbol{\theta}) = f(\boldsymbol{x}) + \frac{1}{2}\alpha(\boldsymbol{\theta}^T\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{x}), \text{ subject to} \\ \|\boldsymbol{x}-\boldsymbol{\theta}\| \le \varepsilon, \quad \boldsymbol{x} \in C \\ \text{for some } \epsilon \ge 0. \end{array}$$

Note that $(C, \boldsymbol{\theta}; \varepsilon)$ is a convex programming model for every fixed ε and it is a convex program for every fixed ε and $\boldsymbol{\theta}$. Under the assumption of a unique global optimum of (P), for a sequence $\varepsilon \to 0$, the sequence of optimal solutions $(\boldsymbol{x}^\circ(\varepsilon), \boldsymbol{\theta}^\circ(\varepsilon))$ converges to some $((\boldsymbol{x}^\circ(0), \boldsymbol{\theta}^\circ(0))$, where the limit points coincide, i.e., $\boldsymbol{x}^\circ(0) = \boldsymbol{\theta}^\circ(0)$. Hence $\boldsymbol{x}^* = \boldsymbol{x}^\circ(0)$ is an optimal solution of (P). Since the set C is assumed to be nonempty and compact, the optimal solutions $(\boldsymbol{x}^\circ(\varepsilon), \boldsymbol{\theta}^\circ(\varepsilon))$ exist. The results on convergence and stability have been reported in [5, 6] and we will not repeat them here. The main point is that, using convexifications of C^2 functions, one can reduce general nonlinear programs to convex models. In particular, one can convexify programs in which one is minimizing a concave (!) function on an arbitrary set of constraints.

Example 2. *Let us consider the problem of minimizing a concave scalar function, say,* $f(t) = -t^2 + 2t$, *on the set* $0 \le t \le 1$. *This problem can be formulated, using a convexification of* $f(t)$ *and after choosing* $\lambda^* = -2$, *as*

$$\begin{aligned} \text{Min } & -t^2 + 2t + \frac{1}{2}(-2)(\theta t - t^2) = t(2-\theta), \\ & -\varepsilon \le t - \theta \le \varepsilon, \quad 0 \le t \le 1. \end{aligned}$$

As, say, $\varepsilon_k = \frac{1}{k} \to 0$, *the optimal solutions* $\theta_k = \frac{1}{k} \to 0$ *and* $t_k = \frac{1}{k} \to 0$. *Hence* $t^* = 0$ *is an optimal solution of the original problem.*

One can apply convexification to all problems that can be reduced to continuous optimization problems. For example, the problem of finding a solution (or a best approximate solution) of a nonlinear system of equations can be replaced by an optimization problem and then reduced to a convex model.

Example 3. *Consider the problem of solving the system of equations from [5, p.19]:*

$$\begin{aligned} 4x_1^3 + 3x_1^2 + x_2x_3 &= 1, \\ 4x_2^3 - 2x_2 + x_1x_3 &= -1, \\ 2x_3 + x_1x_3 &= 1. \end{aligned}$$

This problem can be formulated as the unconstrained program

$$\mathit{Min}\ (4x_1^3 + 3x_1^2 + x_2x_3 - 1)^2 + (4x_2^3 - 2x_2 + x_1x_3 + 1)^2 + (2x_3 + x_1x_3 - 1)^2$$

and then as the convex programming model

$$\begin{aligned} \text{Min } & (4x_1^3 + 3x_1^2 + x_2x_3 - 1)^2 + (4x_2^3 - 2x_2 + x_1x_3 + 1)^2 + (2x_3 + x_1x_3 - 1)^2 \\ & + \frac{1}{2}\alpha(\theta_1x_1 + \theta_2x_2 + \theta_3x_3 - x_1^2 - x_2^2 - x_3^2) \end{aligned}$$

subject to $-\epsilon \leq x_i - \theta_i \leq \epsilon$, $i = 1, 2, 3$ *for* $\alpha < 0$ *sufficiently small. Using input optimization for convex models, one finds that, as* $\epsilon \to 0$, *the optimal solutions* $(\boldsymbol{x}(\epsilon), \boldsymbol{\theta}(\epsilon)) \to (\boldsymbol{x}^*, \boldsymbol{\theta}^*)$, *where* $\boldsymbol{x}^* = \boldsymbol{\theta}^* = (0.571, -0.940, 0.768)^T$. *Here* $\boldsymbol{x}(\epsilon)$ *denotes an optimal solution of the convex program for fixed* ϵ *and* $\boldsymbol{\theta}(\epsilon)$. *A solution of the system of equations is thus obtained as the limit of a sequence of convex programs.*

Convex models can be solved by the GOP methods [1], or by input optimization methods studied in [5]. The success of convexification depends on the methods used for solving convex programming models.

4. Counter Example

Convexification may or may not work for functions that are not C^2. Here are examples.

Example 4. *Consider the function* $f(t) = t^4(2 + \sin(1/t))$. *It appears in, e.g., [2] and [5], where it is used in a different context. At the first sight, by looking at its graph, the function appears to be convex. However, a closer inspection of the function reveals that this is not true in any neighborhood of* $t = 0$, *although* $t = 0$ *is its unique global minimum; see* Figure 2.

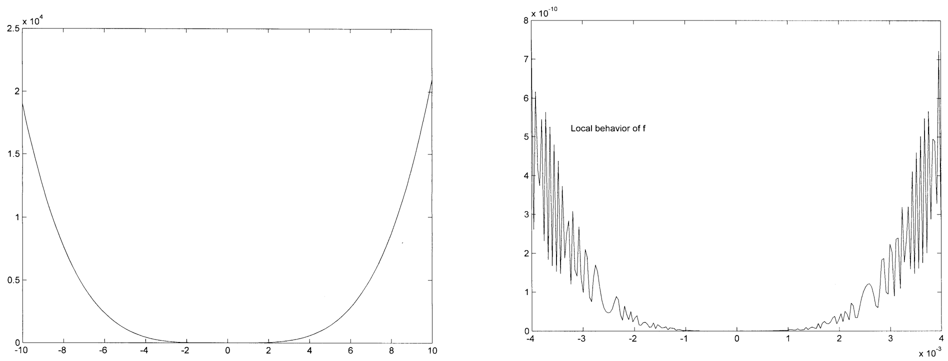

Figure 2. *Global and local behavior of* f

The second derivative of f is $f''(t) = 12t^2(2 + \sin(1/t)) - 6t\cos(1/t) - \sin(1/t)$. The function is not defined at $t = 0$ and hence it does not belong to the class C^2. (One can find a sequence $t \to 0$ that converges to any prescribed value from the interval $[-1, +1]$). Still, the values of the function, close to $t = 0$, range between -1 and $+1$. If we chose, e.g., $\lambda^* = -2$, we would get the convexification $\varphi(t) =$

$t^4(2 + \sin(1/t)) + t^2$. The graph of $\varphi(t)$ is depicted in *Figure 3.*

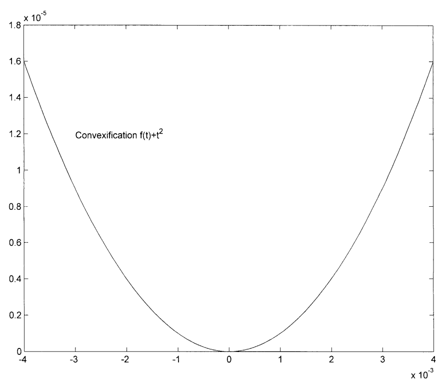

Figure 3. *Convexification of f*

Example 5. *Now consider* $g(t) = t^2 \sin(1/t)$ *for* $t \neq 0$ *and* $g(0) = 0$. *It is well known that this function is differentiable but not continuously differentiable; again see, e.g., [1] and [4]. Hence* $g \notin C^2$. *One finds that its first derivative is* $g^{'}(t) = 2t \sin(1/t) - \cos(1/t)$ *if* $t \neq 0$ *and* $g^{'}(0) = 0$. *Global and local behavior of g is depicted in* Figure 4.

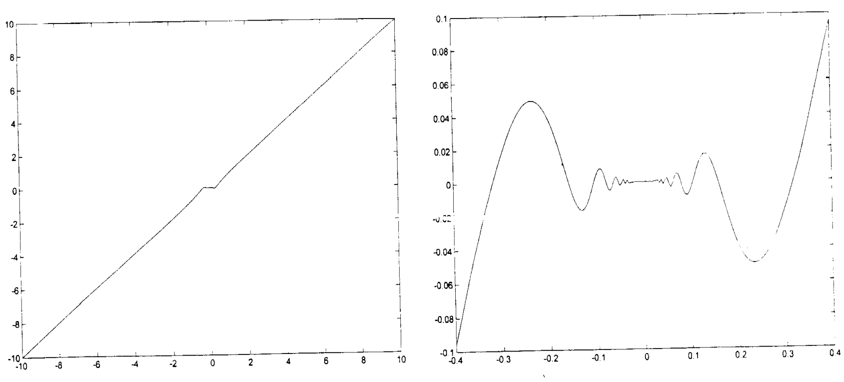

Figure 4. *Global and local behavior of g*

If one attempted to convexify g with any quadratic $c(t) = \gamma t^2$, for $\gamma > 0$ sufficiently large, the convexification would not work. In the absence of the second derivative, one can check this claim by looking at the first derivative. Indeed, $g^{'}(t) + 2\gamma t$ is not monotonic around the origin, which is enough to deduce that $\varphi(t) = g(t) + \gamma t^2$ is not convex. The graph of $\varphi^{'}(t)$ is depicted in *Figure 5* for $\gamma = 500$.

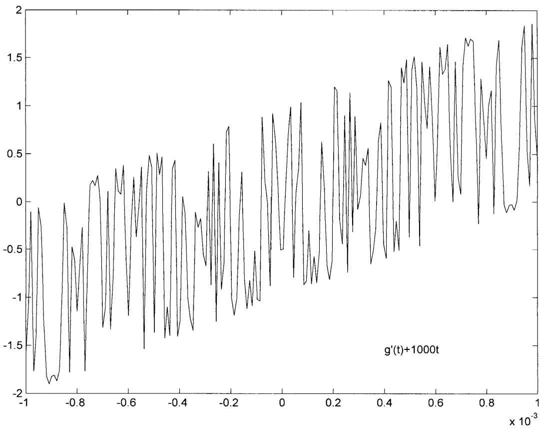

Figure 5. *The derivative of* φ

Acknowledgment. The author is indebted to Mr Mathieu Leduc for testing the results of this paper.

References

[1] C. A. Floudas, *Deterministic Global Optimization*, Kluwer, Dordrecht, 2000.

[2] B. R. Gelbaum, J. M. H. Olmsted, *Counterexamples in Analysis*, Holden-Day Inc., San Francisco, 1964.

[3] W. B. Liu, C. A. Floudas, *A remark on the GOP algorithm for global optimization*, Journal of Global Optimization **3**(1993), 519–521.

[4] A. N. Tikhonov, V. Y. Arsenin, *Solutions of Ill-Posed Problems*, Winston & Sons, Washington, D. C., 1977.

[5] S. Zlobec, *Stable Parametric Programming*, Kluwer, Dordrecht, 2001.

[6] S. Zlobec, *Saddle-point optimality: A look beyond convexity*, Lecture dedicated to Professor Lj. Martić, given at the conference KOI'02, Trogir, Croatia, October 2002, to be published in Journalof Global Optimization.