Journal of Global Optimization **25:** 169–174, 2003.
© 2003 *Kluwer Academic Publishers. Printed in the Netherlands.*

A Note on the Griewank Test Function

M. LOCATELLI
Dipartimento di Informatica, Università di Torino, Corso Svizzera 185, 10149 Torino, Italy (e-mail: locatelli@di.unito.it)

Abstract. In this paper we analyze a widely employed test function for global optimization, the Griewank function. While this function has an exponentially increasing number of local minima as its dimension increases, it turns out that a simple Multistart algorithm is able to detect its global minimum more and more easily as the dimension increases. A justification of this counterintuitive behavior is given. Some modifications of the Griewank function are also proposed in order to make it challenging also for large dimensions.

Key words: Griewank function, Multistart

1. Introduction

The Griewank function, first introduced in (Griewank, 1981), has been employed as a test function for global optimization algorithms in many papers. It is defined as follows

$$\begin{aligned} &\text{Griewank}_n(x) = \sum_{i=1}^{n} \frac{x_i^2}{4000} - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 \\ &- 600 \leqslant x_i \leqslant 600 \quad i = 1, \ldots, n. \end{aligned} \tag{1}$$

The global minimum value is 0 and the global minimum is located in the origin, but the function also has a very large number of local minima, exponentially increasing with n. A typical choice for the n value in the literature is $n = 10$. Figure 1 shows the graph of the function in the two-dimensional case over the box $[-20, 20] \times [-20, 20]$. The fast increasing number of local minima suggests that the global minimum becomes extremely difficult to detect as n increases. Is this really the case? In what follows we present the results obtained by a simple Multistart algorithm. We recall that, at each iteration, the Multistart algorithm samples a uniform random point over the feasible region and start a local search from it. In Table 1 we report the expected number of local searches to first detect the global minimum for $n = 2, 4, 6, 8, 10, 20$. The local search procedure is the limited memory BFGS (see Nocedal, 1980). Surprisingly, it appears to be very difficult to detect the global minimum for small n values but it becomes extremely easy for large n values (even for the value $n = 10$, usually employed in the tests). In the following section we will give an explanation of this counterintuitive behavior and

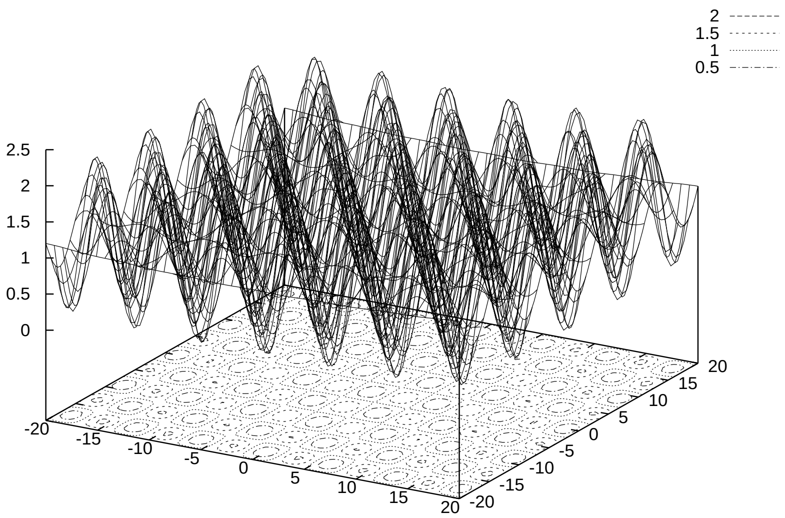

Figure 1. The graph of the two-dimensional Griewank function over the box $[-20, 20] \times [-20, 20]$.

Table 1. Expected number of local searches to first hit the global minimum

n	2	4	6	8	10	20
Expected number	> 10000	> 10000	767	31	5	1

some modifications of the function will be proposed in order to make it challenging also for large n values.

2. Why the Griewank function becomes easier as the dimension increases

We first notice that the Griewank function (1) can be naturally split as follows

$$\text{Griewank}_n(x) = f_n(x) + h_n(x) + 1,$$

where

$$f_n(x) = \sum_{i=1}^{n} \frac{x_i^2}{4000}$$

is a quadratic convex function, whose local (and global) minimum is located in the origin (the same position of the global minimum of the Griewank function), while

$$h_n(x) = -\prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right)$$

is an oscillatory nonconvex function. Basically, the oscillations introduced by function h_n are superimposed over the function f_n and give rise to the large number of local minima.

Now we consider the behavior of many local search methods when applied to f_n. The reason why we omit the function h_n will become clear later. Many local search methods are able to return very quickly the local and global minimum of function f_n. Indeed, in view of Theorem 3.4.2 in (Fletcher, 1987) it holds that, if exact line searches are employed, each method in the Broyden family (which includes as special cases the DFP and BFGS methods) starting with the identity matrix I as an approximation for the Hessian, is equivalent to the Fletcher–Reeves conjugate gradient method when applied to a quadratic function and thus reaches a stationary point of the function after m iterations, where m is the number of linearly independent vectors in the sequence

$$g_1,\ Hg_1,\ H^2g_1,\ldots$$

(g_1 denotes the gradient of the function at the starting point x_1 and H is the Hessian of the quadratic function). Since the Hessian of the function f_n is $I/2000$ we can conclude that $m = 1$, i.e., each of the above mentioned methods returns the local and global minimum of f_n after a single iteration. We underline that the results in Table 1 have been obtained with the limited memory BFGS procedure which is not included in the above-mentioned methods. However, the limited memory BFGS procedure and the BFGS procedure are equivalent during the first iterations. But now let us reintroduce the function h_n previously omitted. We need to modify all the computations above by adding to all the gradient values the gradient values of the function h_n, whose i-th component is the following

$$[\nabla h_n(x)]_i = \frac{1}{\sqrt{i}}\sin\left(\frac{x_i}{\sqrt{i}}\right)\prod_{j=1,\,j\neq i}^{n}\cos\left(\frac{x_j}{\sqrt{j}}\right).$$

We notice that all the components of ∇h_n are obtained as the product of n values belonging to the interval $[-1, 1]$. As n increases, it becomes more and more likely that the product of such values is very small so that the gradient of h_n can be neglected with respect to the gradient of f_n. Equivalently, the part of the feasible region where the gradient of the Griewank function is significatively different from the gradient of f_n becomes smaller and smaller as n increases. This explains the results in Table 1: as n increases the limited memory BFGS procedure applied to the Griewank function behaves more and more similarly to the same procedure applied to function f_n and the global minimum (which is the same for the Griewank function and function f_n) is more and more easily detected.

It could be argued that the Griewank function becomes easier as n increases only if we employ a local search procedure such as the limited memory BFGS procedure or any of the other local search procedures previously mentioned (Broyden family, conjugate gradient). It is undoubtely true that, as shown above, these

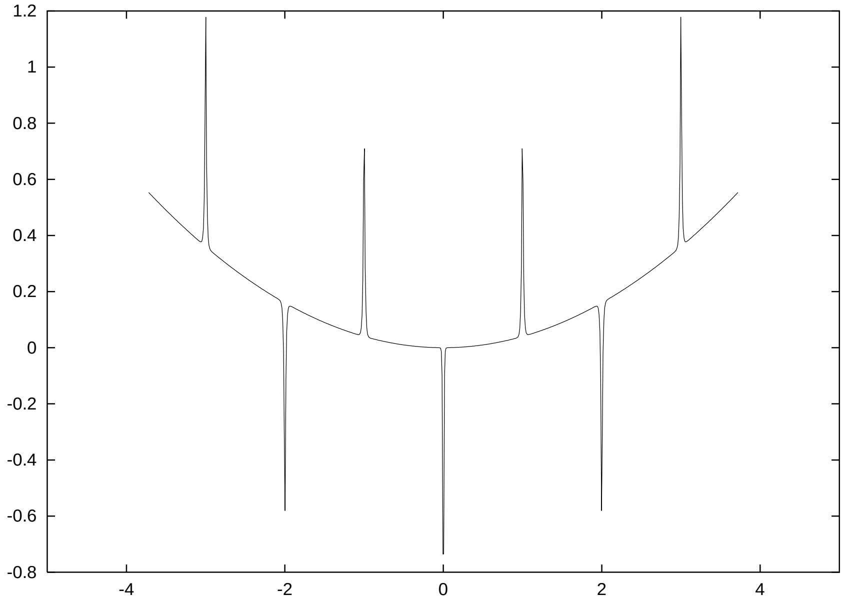

Figure 2. A one-dimensional illustration of the landscape of the Griewank function for large n values.

local search procedures are particularly well suited for this problem, but the same reasoning which explains their success suggests that this function is not a very challenging one also for other algorithms, at least for large values of n. Indeed, since the region where the function values and the gradient values of the Griewank function are significatively different from those of function f_n becomes narrower and narrower as n increases, algorithms can hardly distinguish, on the basis of the observed values, between the two functions. Therefore, any algorithm which is able to solve in a reasonable amount of time convex problems and, in particular, the one represented by function f_n, is very likely able to approach the global minimum of the Griewank function in a comparable time. Figure 2 represents a one-dimensional function which is not the Griewank one for $n = 1$ but has properties similar to those of the Griewank function for large values of n: there are very narrow regions where the function drops down or rise up but in the largest part of the feasible region the function is basically not distinguishable from the underlying convex function f_n.

2.1. MODIFICATIONS OF THE GRIEWANK FUNCTION

In this subsection we discuss how the Griewank function could be modified in order to make it a very challenging test function also in large dimensions. In order to achieve this aim we need to increase the incidence of the nonconvex and oscillatory part. For instance, we can modify the terms in the product in such a way that they are never below the value 1. We can redefine the Griewank function as follows

$$\text{Griewank}^1_{\text{modified}}(x) = f_n(x) - \prod_{i=1}^{n}\left[2 + \cos\left(\frac{x_i}{\sqrt{i}}\right)\right] + 3^n.$$

or, if we want to avoid the numerical problems due to the often very large values of the product by taking the logarithm of the product, as follows

$$\text{Griewank}^2_{\text{modified}}(x) = f_n(x) - \sum_{i=1}^{n}\log\left[2 + \cos\left(\frac{x_i}{\sqrt{i}}\right)\right] + n\log(3). \qquad (2)$$

Note that the global minimum of both modifications is still located in the origin and its value is still equal to 0. As expected, the Multistart algorithm is now unable to detect, within 10 000 local searches, the global minimum of both the modifications above for $n = 10$ and for $n = 20$. However, we underline a negative aspect of modification (2). The function proposed in (2) is a separable one, i.e., it can be written as the sum of n one-dimensional functions. Some results in the literature on separable test functions are excellent but misleading. Indeed, the proposed approaches modify only one or few variables at each iteration. In the most extreme case where only a single variable is modified at each iteration, this means that the approach is performing n optimization of one-dimensional functions, so that the complexity of the n-dimensional problem is comparable to n times the complexity of a one-dimensional problem. In order to recover nonseparability we can modify the Griewank function as follows

$$\text{Griewank}^3_{\text{modified}}(x) = f_n(x) - \sum_{i=1}^{n}\log\left[2 + \cos\left((Hx)_i\right)\right] + n\log(3), \qquad (3)$$

where H is a nonsingular matrix. Note that modification (3) include as a special case (2), but if H is nondiagonal, then the resulting modification of the Griewank function is nonseparable. Figure 2.1 illustrates the graph of a two-dimensional instance over the box $[-10, 10] \times [-10, 10]$, where

$$H = \begin{bmatrix} 1 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \qquad (4)$$

together with the underlying graph of function f_n.

3. Conclusion

In this paper a widely employed test function for global optimization algorithms, the Griewank function, has been analyzed. It has been observed that, while the number of local minima increases exponentially with the dimension n, the function becomes very easy to optimize, for large n values, by a simple Multistart algorithm.

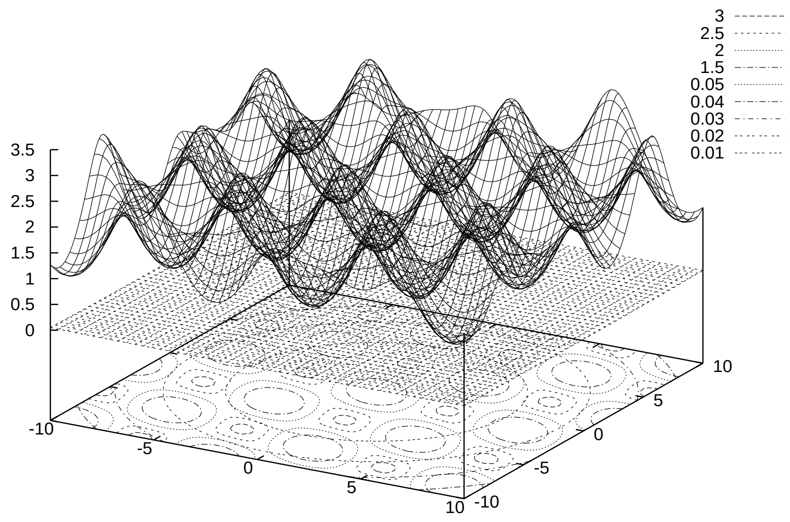

Figure 3. A two-dimensional instance of modification (2) over the box $[-10, 10] \times [-10, 10]$ with h_n given in (3) and H given in (4), and the underlying function f_n.

An explanation of this fact has been given. The same explanation suggests that not only the Multistart algorithm considered in this paper, but also other optimization algorithms can relatively easily approach the global minimum of this function for large n values. Finally, some modifications of the function have been proposed in order to make it a challenging one also in high dimensions.

Acknowledgements

The author is very grateful to an anonymous referee for some helpful suggestions through which the paper could be improved.

References

Fletcher R. (1987), *Practical Methods of Optimization*, 2nd edition, John Wiley, Chichester.

Griewank A.O. (1981), Generalized descent for global optimization, *Journal of Optimization Theory and Applications* 34, 11-39.

Nocedal J. (1980), Updating quasi-Newton matrices with limited storage, *Math.Comput.* 35, 773-782.