Convergence Results for Pattern Search Algorithms are Tight

Charles Audet

CRPC-TR98779
November 1998

Center for Research on Parallel Computation
Rice University
6100 South Main Street
CRPC - MS 41
Houston, TX 77005

Submitted November 1998

Convergence Results for Pattern Search Algorithms are Tight

Charles Audet
charlesa@caam.rice.edu

Rice University
Department of Computational
and Applied Mathematics
6100 South Main Street - MS 134
Houston, Texas
77005-1892 USA

November 17, 1998

Abstract

Recently, general definitions of pattern search methods for both unconstrained and linearly constrained optimization were presented. It was shown under mild conditions, that there exists a subsequence of iterates converging to a stationary point. In the unconstrained case, stronger results are derived under additional assumptions. In this paper, we present three small dimensioned examples showing that these results cannot be strengthened without additional assumptions. First, we show that second order optimality conditions cannot be guaranteed. Second, we show that there can be an accumulation point of the sequence of iterates whose gradient norm is strictly positive. These two examples are also valid for the bound constrained case. Finally, we show that even under the stronger assumptions of the unconstrained case, there can be infinitely many accumulation points.

Key words

Pattern search algorithms, convergence analysis, unconstrained optimization, bound constrained optimization.

Acknowledgments: This work is supported by NSERC (Natural Sciences and Engineering Research Council of Canada) fellowship PDF-207432-1998, and by CRPC (Center for Research on Parallel Computing).

1 Introduction

We consider the unconstrained optimization problem of minimizing a continuously differentiable function $f : \mathrm{R}^n \to \mathrm{R}$, without any knowledge of its derivatives, and without any means of approximating them.

Torczon [4] observes that several existing search methods for this problem share a common structure, and they can be subsumed into a more general one. The *Pattern Search Method* defined there encompasses a wide class of algorithms, but it still guarantees strong convergence results under precise conditions. This work is generalized in Lewis and Torczon [2] to bound constrained minimization and in [3] to linearly constrained minimization.

Pattern search methods produce a sequence of iterates $x_0, x_1, \ldots$ in R^n as follows. For $k \geq 0$, iteration k is initiated with an iterate $x_k \in \mathrm{R}^n$, a pattern P_k consisting of a finite set of vectors in R^n, and a step size $\Delta_k > 0$. The objective of iteration k is to find a direction p_k^i in the pattern P_k such that $f(x_k + \Delta_k p_k^i) < f(x_k)$. The initial iterate x_0 and step size Δ_0 are given.

If such a point is found, then the iteration is declared successful, and the next iterate is $x_{k+1} = x_k + \Delta_k p_k^i$. The step size parameter Δ_{k+1} is set to $\lambda\Delta_k$, where $\lambda \geq 1$ is chosen among a fixed finite set of values. It is assumed that the level set $L(x_0) = \{x \in \mathrm{R}^n : f(x) \leq f(x_0)\}$ is compact, and therefore all iterates are bounded in norm since they all belong to the level set.

If no such point is found, then the iteration is declared unsuccessful, and the next iteration is initiated at the same point $x_{k+1} = x_k$. The step size parameter Δ_{k+1} is reduced to $\theta\Delta_k$, where $0 < \theta < 1$ is constant over all iterations.

Observe that not all points of the pattern P_k need to be considered. An iteration can end as soon as a solution that yields decrease in the objective function is found. However, prior to declaring an iteration unsuccessful, it is imperative that at least n linearly independent vectors of P_k as well as their opposite direction be considered. Therefore, if the gradient $\nabla f(x_k)$ is non-zero, and if the step size Δ_k is sufficiently small, then the iteration will be successful. Moreover, these linearly independent directions must be chosen among a fixed finite set over all iterations. Lewis and Torczon [2] lift this restriction using positive basis theory. The number of required directions drops from $2n$ to $n + 1$.

The contribution of this paper is to show through three small examples that Torczon's [4] convergence results cannot be strengthened without introducing additional assumptions. The main convergence result is that

$$\liminf_{k\to\infty} \|\nabla f(x_k)\| = 0. \tag{1}$$

The first example illustrates that no second order optimality conditions can be guaranteed. In the example, the sequence of iterates converges (in fact it remains) to a point whose

gradient norm is zero, but which is a maximizer of the function.

The second example shows that the sequence of iterates may have infinitely many accumulation points, and that the gradient norm at one of them may be non-zero. Therefore, the main result cannot be strengthened to

$$\lim_{k\to\infty} \|\nabla f(x_k)\| \;=\; 0. \tag{2}$$

Both first and second examples are also valid in the constrained cases studied in [1] and [3]. All iterates and trial points satisfy bound constraints, and the search directions p_k^1 and p_k^2 of the matrix P_k are orthogonal to the bound constraints. In fact, using the notation in [1], the submatrix BM_k of P_k is the identity matrix.

Under stronger assumptions, Torczon [4] shows that in the unconstrained case, equation (1) may be improved to (2). Our last example illustrates that even under these additional restrictions, there still might be an infinite number of accumulation points.

The reader is referred to Torczon [4] for a complete description of the method and its parameters. To ease the notation, we only introduce what is necessary to present the examples, the method is more general than what appears here.

2 Inexistence of second order optimality conditions

In this first example, we show that the convergence result cannot be extended to guarantee second order optimality conditions.

Consider the following pattern search strategy. At each iteration, all variables are increased and decreased one at a time by the same step size. If decrease in the objective function is observed, the next iteration is initiated at the improved point with the same step size (thus $\lambda = 1$), otherwise the next iteration is initiated at the same point with half the step size (thus $\theta = \frac{1}{2}$).

Consider the two dimensional example with the continuously differentiable objective function

$$f(x,y) \;=\; \begin{cases} -x^2y^2 & \text{if } x^2+y^2 \leq 1 \\ -x^2y^2 + (x^2+y^2-1)^2 & \text{if } x^2+y^2 > 1. \end{cases}$$

Let $(x_0, y_0) = (0,0)$ be the starting point, and let $\Delta_0 = 1$ be the initial step size parameter. The objective function is never evaluated by the algorithm outside of the ball centered at the origin or radius one. It is defined in a piecewise way only to ensure that the level set $L(0,0)$ is compact. For any k, the same pattern matrix is used

$$P_k = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \end{bmatrix}.$$

Throughout all iterations, the iterates remain at the origin since the objective values of the trial points $(\frac{1}{2^k},0),(0,\frac{1}{2^k})$, $(\frac{-1}{2^k},0)$ and $(0,\frac{-1}{2^k})$ are all equal to $f(0,0)=0$ for any $k\geq 0$. The algorithm never succeed in moving away from the origin, and therefore the sequence of iterates (x_k) remains and thus converges to $(0,0)$, a maximizer of the function.

In practice, such behavior is improbable since the search phase of the algorithm usually evaluates the function in several other directions (using the matrix L_k in Torczon [4]).

3 Accumulation point with non-zero gradient

In this section, we present a two-dimensional example on which a pattern search algorithm generates an infinite number of accumulation points. The gradient norm at one of these points is non-zero.

Consider the function g defined on the domain $\{(x,y): -3\leq y\leq 3\}$,

$$g(x,y)=\begin{cases} f_1(x,y)=-26x^3-32x^2y+7|y|^3 & \text{if } x\leq 0\\ f_2(x,y)=(7-8x^2)|y|^3 & \text{if } 0<x\leq\frac{1}{2}\\ f_3(x,y)=f_2(x,y)+8(x-\frac{1}{2})^2(y^3+y+x-1) & \text{if } \frac{1}{2}<x. \end{cases}$$

The function g is extended to f as follows

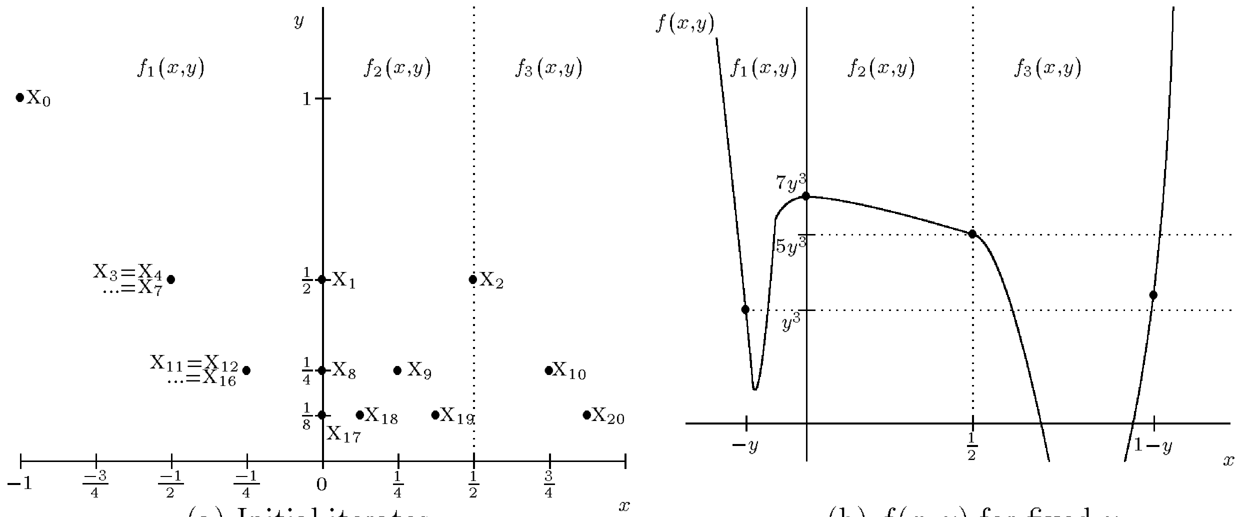
$$f(x,y)=\begin{cases} g(x,y)+(y+3)^6 & \text{if } y<3\\ g(x,y) & \text{if } -3\leq y\leq 3\\ g(x,y)+(y-3)^6 & \text{if } y>3. \end{cases}$$

It can be shown that both functions are continuously differentiable everywhere on R^2. Moreover, the function g is extended to f on R^2 to ensure that its level sets are compact. All iterates and trial points generated by the algorithm belong to the domain of g. For any $y\in[-3,3]$, the values and gradients at the transition points of f are

$$\begin{aligned} f(0,y) &= f_1(0,y)=7|y|^3, & f(\tfrac{1}{2},y) &= f_2(\tfrac{1}{2},y)=5|y|^3\\ \nabla f(0,y) &= \begin{bmatrix} 0\\ 21y|y| \end{bmatrix}, & \nabla f(\tfrac{1}{2},y) &= \nabla f_2(\tfrac{1}{2},y). \end{aligned}$$

In order to lighten the notation, we denote the k^{th} iterate (x_k,y_k) by x_k. The same pattern matrix P_k as in the first example is used but with the additional column $[4\ -2]^t$. The search strategy is guided by the following rules:

- If $x_k<0$ and $y_k=4\Delta_k$, then the search starts with $\mathrm{x}_k+\Delta_k(4,-2)$.
- If $x_k\in[0,\frac{1}{2}[$, then the search starts with $\mathrm{x}_k+\Delta_k(1,0)$.
- If $x_k\geq\frac{1}{2}$, then the search phase starts with $\mathrm{x}_k+\Delta_k(-1,0)$.
- Otherwise, the search phase explores directly and sequentially the first four columns.



(a) Initial iterates (b) $f(x,y)$ for fixed y

Figure 1: Initial iterates and objective function

In all cases, an iteration ends as soon as a solution that yields a lower objective function value than the current iterate's is found.

The step size control parameters are $\theta = \frac{1}{2}$ and $\lambda = 2$. Therefore, at unsuccessful iterations $\Delta_{k+1} = \frac{\Delta_k}{2}$, and at successful ones $\Delta_{k+1} = 2\Delta_k$.

The initial data is $(x_0, y_0) = (-1, 1)$, $f(x_0) = 1$ and $\Delta_0 = \frac{1}{2^2}$. Figure 1(a) displays part of the domain of function f and plots the first few iterates. Figure 1(b) shows the shape of the function f when x varies from $-y$ to $1-y$, and y is fixed to a value between 0 and $\frac{1}{2}$. Table 1 details the search strategy of the first eight iterations. The successful trial points appear in boldface letters.

The key structure of this example is that for any fixed value of y, the function value monotonically decreases from $7y^3$ to $5y^3$ when x varies from 0 to $\frac{1}{2}$, as illustrated in Figure 1(b). The iterates go from $(0, y)$ eventually to $(\frac{1}{2} - y, y)$ and the objective function value decreases from $7y^3$ to a value greater than $5y^3$. From this last point a mesh size parameter of $\frac{1}{2}$ yields decrease at trial point $(1-y, y)$ at which the function value is $(1+8y)y^3 \in]y^3, 5y^3]$. From there, a step length of 1 leads to $(-y, y)$ with function value y^3.

We define the ℓ^{th} cycle (for an $\ell \geq 0$) as the iterations beginning at $k = \ell^2 + 6\ell$ and ending and including $k = (\ell+1)^2 + 6(\ell+1) - 1$. As shown below, a cycle starts at the iterate $\mathrm{x}_k = (\frac{-1}{2^\ell}, \frac{1}{2^\ell}) = (x_k, y_k)$ with $\Delta_k = \frac{1}{2^{\ell+2}}$ and function value $\frac{1}{8^\ell}$. The cycle goes through $\ell + 3$ successful iterations,

$$f\left(0, \frac{y_k}{2}\right) = \frac{7}{8^{\ell+1}}, \ldots, f\left(\frac{1}{2} - \frac{y_k}{2}, \frac{y_k}{2}\right) > \frac{5}{8^{\ell+1}}, f\left(1 - \frac{y_k}{2}, \frac{y_k}{2}\right) > \frac{1}{8^{\ell+1}}, f\left(\frac{-y_k}{2}, \frac{y_k}{2}\right) = \frac{1}{8^{\ell+1}}$$

whose last point has step size parameter equal to 2. Then, the cycle goes through $\ell + 4$

k	x_k	Δ_k	Search			
0	$(-1,1)$	$\frac{1}{4}$	$\mathbf{f_1(0,\frac{1}{2})=\frac{7}{8}}$			
1	$(0,\frac{1}{2})$	$\frac{1}{2}$	$\mathbf{f_2(\frac{1}{2},\frac{1}{2})=\frac{5}{8}}$			
2	$(\frac{1}{2},\frac{1}{2})$	1	$\mathbf{f_1(\frac{-1}{2},\frac{1}{2})=\frac{1}{8}}$			
3	$(\frac{-1}{2},\frac{1}{2})$	2	$f_3(\frac{3}{2},\frac{1}{2})=\frac{61}{8}$	$f_1(\frac{-5}{2},\frac{1}{2})=\frac{2457}{8}$	$f_1(\frac{-1}{2},\frac{5}{2})=\frac{741}{8}$	$f_1(\frac{-1}{2},\frac{-3}{2})=\frac{311}{8}$
4	x_3	1	$f_2(\frac{1}{2},\frac{1}{2})=\frac{5}{8}$	$f_1(\frac{-3}{2},\frac{1}{2})=\frac{421}{8}$	$f_1(\frac{-1}{2},\frac{3}{2})=\frac{119}{8}$	$f_1(\frac{-1}{2},\frac{-1}{2})=\frac{65}{8}$
5	x_3	$\frac{1}{2}$	$f_1(0,\frac{1}{2})=\frac{7}{8}$	$f_1(-1,\frac{1}{2})=\frac{87}{8}$	$f_1(\frac{-1}{2},1)=\frac{18}{8}$	$f_1(\frac{-1}{2},0)=\frac{26}{8}$
6	x_3	$\frac{1}{4}$	$f_1(\frac{-1}{4},\frac{1}{2})=\frac{2.25}{8}$	$f_1(\frac{-3}{4},\frac{1}{2})=\frac{22.75}{8}$	$f_1(\frac{-1}{2},\frac{3}{4})=\frac{1.625}{8}$	$f_1(\frac{-1}{2},\frac{1}{4})=\frac{10.875}{8}$
7	x_3	$\frac{1}{8}$	$\mathbf{f_1(0,\frac{1}{4})=\frac{7}{8^2}}$			

Table 1: Initial eight iterations of pattern search method

unsuccessful iterations (the function is evaluated at all four trial points corresponding to the first four columns of P_k) leading to the start of cycle $\ell+1$.

For example, cycle 0 starts with $(x_0, y_0) = (-1, 1)$ and $\Delta_0 = \frac{1}{4}$. Three successful iterations go through the points $(0, \frac{1}{2}), (\frac{1}{2}, \frac{1}{2})$ and finally $(\frac{-1}{2}, \frac{1}{2})$, where $\Delta_3 = 2$. Then five unsuccessful iterations reduce the mesh size parameter to $\Delta_7 = \frac{1}{8}$. This initiates cycle 1.

We show that the sequence of iterates possesses an infinite number of accumulation points, namely all points of the set $\{(\frac{1}{2^\ell}, 0) : \ell = 0, 1, \ldots\}$. Moreover, the gradient norm of all these points is zero, except for the point $(1, 0)$.

Proposition 3.1 *For any integer $\ell \geq 0$, the iterates and step size parameters of cycle ℓ are*

$$(x_k, y_k), \Delta_k = \begin{cases} \left(\frac{-1}{2^\ell}, \frac{1}{2^\ell}\right), & \frac{1}{2^{\ell+2}} \quad \textit{if } k = \ell^2 + 6\ell \\ \left(\frac{2^{j-1}-1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}\right), & \frac{1}{2^{\ell-j+2}} \quad \textit{if } k = \ell^2 + 6\ell + j, \quad j \in \{1, 2, \ldots, 2+\ell\} \\ \left(\frac{-1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}\right), & \frac{1}{2^{j-2}} \quad \textit{if } k = \ell^2 + 7\ell + 2 + j, \quad j \in \{1, 2, \ldots, 4+\ell\}. \end{cases}$$

Proof: The proof is done by induction on ℓ. We already verified in Table 1 that the result holds for the cycle $\ell = 0$, that is, from iterations 0 to 6.

Suppose that cycle ℓ is initiated at iterate $\mathrm{x}_k = \left(\frac{-1}{2^\ell}, \frac{1}{2^\ell}\right)$ where $k = \ell^2 + 6\ell$, and that $\Delta_k = \frac{1}{2^{\ell+2}}$. The objective function value is $f(\mathrm{x}_k) = f_1\left(\frac{-1}{2^\ell}, \frac{1}{2^\ell}\right) = \frac{1}{8^\ell}$. Table 2 details all iterations of cycle ℓ. The successful trial points appear in boldface letters.

k	x_k	Δ_k	Search
$\ell^2+6\ell$	$\left(\frac{-1}{2^\ell}, \frac{1}{2^\ell}\right)$	$\frac{1}{2^{\ell+2}}$	$\mathbf{f(x_k+(4,-2)\Delta_k) = f_1\left(0, \frac{1}{2^{\ell+1}}\right) = \frac{7}{8^{\ell+1}}}$
$\ell^2+6\ell+j$ $1 \le j < 1+\ell$	$\left(\frac{2^{j-1}-1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}\right)$	$\frac{1}{2^{\ell-j+2}}$	$\mathbf{f(x_k+(1,0)\Delta_k) = f_2\left(\frac{2^j-1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}\right) = \frac{7-8\left(\frac{2^j-1}{2^{\ell+1}}\right)^2}{8^{\ell+1}}}$
$\ell^2+7\ell+1$	$\left(\frac{2^\ell-1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}\right)$	$\frac{1}{2}$	$\mathbf{f(x_k+(1,0)\Delta_k) = f_3\left(1-\frac{1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}\right) = \frac{1+\frac{1}{2^{\ell-2}}}{8^{\ell+1}}}$
$\ell^2+7\ell+2$	$\left(1-\frac{1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}\right)$	1	$\mathbf{f(x_k+(-1,0)\Delta_k) = f\left(\frac{-1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}\right) = \frac{1}{8^{\ell+1}}}$
$\ell^2+7\ell+2+j$ $1 \le j < 4+\ell$	$\left(\frac{-1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}\right)$	$\frac{1}{2^{j-2}}$	$f(\mathrm{x}_k+(1,0)\Delta_k) = f\left(\frac{-1}{2^{\ell+1}}+\frac{1}{2^{j-2}}, \frac{1}{2^{\ell+1}}\right)$ $f(\mathrm{x}_k-(1,0)\Delta_k) = f_1\left(\frac{-1}{2^{\ell+1}}-\frac{1}{2^{j-2}}, \frac{1}{2^{\ell+1}}\right)$ $f(\mathrm{x}_k+(0,1)\Delta_k) = f_1\left(\frac{-1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}+\frac{1}{2^{j-2}}\right)$ $f(\mathrm{x}_k-(0,1))\Delta_k) = f_1\left(\frac{-1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}-\frac{1}{2^{j-2}}\right)$

Table 2: Objective function values of cycle ℓ

Let $a = \frac{1}{2^{\ell+1}}$. In Table 2, all trial points of iterations $\ell^2+7\ell+2+j$ for $1 \le j < 4+\ell$ are of the form $f(x,a)$ or $f(-a,y)$, where

$$x \in \{-a-2, -a-1, -a-\frac{1}{2}, \dots, -a-a, -a-\frac{a}{2}\} \cup \{\frac{a}{2}-a, a-a, 2a-a, \dots, 1-a, 2-a\}$$

and

$$y \in \{a-2, a-1, a-\frac{1}{2}, \dots, a-a, a-\frac{a}{2}\} \cup \{a+\frac{a}{2}, a+a, a+2a, \dots a+1, a+2\}.$$

The current iterate of the unsuccessful iterations is $(-a,a)$. Table 3 shows that the objective function value of the search points are greater than $f_1(-a,a) = a^3$.

Table 3 shows that the last iteration appearing in Table 2 is unsuccessful. The following iteration number is $k = (\ell^2+7\ell)+2+(4+\ell)+1 = (\ell+1)^2+6(\ell+1)$. Therefore, cycle $\ell+1$ is initiated at $\mathrm{x}_k = \left(\frac{-1}{2^{\ell+1}}, \frac{1}{2^{\ell+1}}\right)$ with step length parameter $\Delta_k = \frac{-1}{2^{\ell+3}}$. This completes the proof.

y fixed	x fixed
for $x \le \frac{-3a}{2}$: $f_1(x,a) = -26x^3 - 32x^2a + 7a^3 \ge \frac{26\times 3}{2}x^2a - 32x^2a + 7a^3 = 7(x^2a + a^3)$	for $y \le 0$: $f_1(-a,y) = 26a^3 - 32a^2y - 7y^3 \ge 26a^3$
$f_1\left(\frac{-a}{2}, a\right) = \frac{26}{8}a^3 - \frac{32}{2}a^3 + 7a^3 = \frac{9}{4}a^3$	$f_1\left(-a, \frac{a}{2}\right) = 26a^3 - 16a^3 + \frac{7}{8}a^3 = \frac{87}{8}a^3$
for $x \in [0, \frac{1}{2}]$: $f_2(x,a) \in [5a^3, 7a^3]$ – see Figure 1(b) –	$f_1\left(-a, \frac{3a}{2}\right) = 26a^3 - 48a^3 + \frac{7\times 27}{8}a^3 = \frac{13}{8}a^3$
$f_3(1-a,a) = (1+8a)a^3 \in]a^3, 5a^3]$	$f_1(-a, 2a) = 26a^3 - 64a^3 + 56a^3 = 18a^3$
$f_3(2-a,a) = 2(a^2(2a-1)^2 + 3(1-a)(3-a)) + a^3 > a^3$	for $y \ge 3a$: $f_1(-a,y) = 26a^3 - 32a^2y + 7y^3 \ge 26a^3 - 32a^2y + 63a^2y = 26a^3 + 31a^2y$

Table 3: Objective function values of unsuccessful search points: $f(\mathrm{x}_k + \Delta_k p_k^i) > a^3 = f(\mathrm{x}_k)$

■

The previous proposition details all iterates generated by the algorithm. We now discuss some properties of the accumulation points of the sequence of iterates. Consider the sequence of iterations $k = \ell^2 + 9\ell + 3$ for $\ell \ge 0$: all corresponding iterates $\mathrm{x}_k = \left(1, \frac{1}{2^{\ell+2}}\right)$ are successful and converge to the point $(1,0)$ at which the gradient is $\nabla f(1,0) = \nabla f_3(1,0) = (1,0)$. All other accumulation points are of the form $\left(\frac{1}{2^\ell}, 0\right)$ for $\ell > 0$, and they have zero gradient. Moreover, only the sequence of iterates that converges to $(0,0)$ contains unsuccessful iterations.

In summary, there is a subsequence of successful iterates that converges to a solution whose gradient norm is non-zero. Furthermore, any subsequence of unsuccessful iterates (whose step size parameters go to zero) converges to a zero gradient accumulation point.

4 Infinite number of accumulation points when under strong assumptions

In this last example, we show that even under the stronger assumptions presented in Torczon [4], there still can be infinitely many accumulation points. These additional requirements are

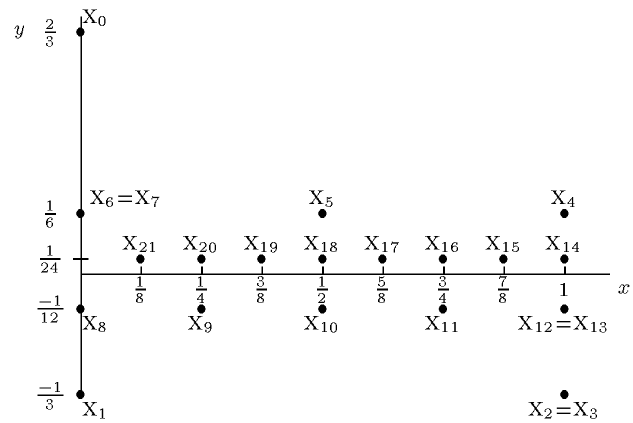


Figure 2: Initial iterates of coordinate search

-i- The columns of the patterns P_k are bounded in norm.
-ii- $\lim_{k\to\infty} \Delta_k = 0$.
-iii- Prior to declaring an iteration successful, at least n linearly independent vectors of P_k as well as their opposite directions must be considered, and the objective function value of the next iterate must be less than or equal to all objective function values of the required points.

Consider the continuously differentiable function in R^2:

$$f(x,y) \;=\; \begin{cases} (x^2+1)y^2 & \text{if } y \geq 0 \\ ((1-x)^2+1)y^2 & \text{if } y < 0. \end{cases}$$

For any $y > 0$ the function monotonically decreases when x varies from 1 to 0, and for any $y < 0$ the function monotonically decreases when x varies from 0 to 1.

The same pattern matrix as in the first example is used. The search strategy consists in evaluating the objective function value at all four trial points, and to set the next iterate to be the one having the least value. Conditions -i- and -iii- are satisfied. The step control parameters are $\lambda = 1$ and $\theta = \frac{1}{2}$, thus condition -ii- is satisfied.

Again, the k^{th} iterate (x_k, y_k) is denoted by x_k. The starting point is $\mathrm{x}_0 = (0, \frac{2}{3})$, the step size parameter is $\Delta_0 = 1$ and $f(\mathrm{x}_0) = \frac{4}{9}$. Figure 2 displays the first few iterates generated by the algorithm. Table 4 details the search strategy of the first eight iterations. The successful trial points appear in boldface letters.

We define an even cycle to be the successful iterations starting at $(0,y)$ where $y>0$, then going to $(0,-y)$ and ending at the unsuccessful one at $(1,\frac{-y}{2})$. The following odd cycle is composed of the successful iterations starting at $(1,\frac{-y}{2})$, then going to $(1,\frac{y}{4})$ and ending at the unsuccessful one at $(0,\frac{y}{4})$.

k	x_k	Δ_k	$f(x_k+\Delta_k,y_k)$	$f(x_k,y_k+\Delta_k)$	$f(x_k-\Delta_k,y_k)$	$f(x_k,y_k-\Delta_k)$
0	$(0,\frac{2}{3})$	1	$f(1,\frac{2}{3})=\frac{8}{3^2}$	$f(0,\frac{5}{3})=\frac{25}{3^2}$	$f(-1,\frac{2}{3})=\frac{8}{3^2}$	$\mathbf{f(0,\frac{-1}{3})=\frac{2}{3^2}}$
1	$(0,\frac{-1}{3})$	1	$\mathbf{f(1,\frac{-1}{3})=\frac{1}{3^2}}$	$f(0,\frac{2}{3})=\frac{4}{3^2}$	$f(-1,\frac{-1}{3})=\frac{5}{3^2}$	$f(0,\frac{-4}{3})=\frac{32}{3^2}$
2	$(1,\frac{-1}{3})$	1	$f(2,\frac{-1}{3})=\frac{2}{3^2}$	$f(1,\frac{2}{3})=\frac{8}{3^2}$	$f(0,\frac{-1}{3})=\frac{2}{3^2}$	$f(1,\frac{-4}{3})=\frac{16}{3^2}$
3	x_2	$\frac{1}{2}$	$f(\frac{3}{2},\frac{-1}{3})=\frac{2}{3^2}$	$\mathbf{f(1,\frac{1}{6})=\frac{2}{6^2}}$	$f(\frac{1}{2},\frac{-1}{3})=\frac{2}{3^2}$	$f(1,\frac{-4}{6})=\frac{16}{3^2}$
4	$(1,\frac{1}{6})$	$\frac{1}{2}$	$f(\frac{3}{2},\frac{1}{6})=\frac{3.25}{6^2}$	$f(1,\frac{1}{6})=\frac{32}{6^2}$	$\mathbf{f(\frac{1}{2},\frac{1}{6})=\frac{1.25}{6^2}}$	$f(1,\frac{-1}{3})=\frac{4}{6^2}$
5	$(\frac{1}{2},\frac{1}{6})$	$\frac{1}{2}$	$f(1,\frac{1}{6})=\frac{2}{6^2}$	$f(\frac{1}{2},\frac{2}{3})=\frac{20}{6^2}$	$\mathbf{f(0,\frac{1}{6})=\frac{1}{6^2}}$	$f(\frac{1}{2},\frac{-1}{3})=\frac{5}{6^2}$
6	$(0,\frac{1}{6})$	$\frac{1}{2}$	$f(\frac{1}{2},\frac{1}{6})=\frac{1.25}{6^2}$	$f(0,\frac{2}{3})=\frac{16}{6^2}$	$f(\frac{-1}{2},\frac{1}{6})=\frac{1.25}{6^2}$	$f(0,\frac{-1}{3})=\frac{8}{6^2}$
7	x_6	$\frac{1}{4}$	$f(\frac{1}{4},\frac{1}{6})=\frac{4.25}{12^2}$	$f(0,\frac{5}{12})=\frac{25}{12^2}$	$f(\frac{-1}{4},\frac{1}{6})=\frac{4.25}{12^2}$	$\mathbf{f(0,\frac{-1}{12})=\frac{2}{12^2}}$

Table 4: Initial eight iterations of coordinate search

We now explicitly write the values of all iterates.

Proposition 4.1 *For any integer $\ell \geq 0$, the iterates and of cycle ℓ are*

$$\begin{array}{llll} x_{2^\ell+2\ell-1} = \left(0,\frac{2\Delta_k}{3}\right), & \text{when } k=2^\ell+2\ell-1, & & \\ x_{2^\ell+2\ell+j} = \left(j\Delta_k,\frac{-\Delta_k}{3}\right), & j\in\{0,1,\ldots,2^\ell\} & \text{and } \ell \text{ is even} \end{array}$$

$$\begin{array}{llll} x_{2^\ell+2\ell-1} = \left(1,\frac{-2\Delta_k}{3}\right), & \text{when } k=2^\ell+2\ell-1, & & \\ x_{2^\ell+2\ell+j} = \left(1-j\Delta_k,\frac{-\Delta_k}{3}\right), & j\in\{0,1,\ldots,2^\ell\} & \text{and } \ell \text{ is odd} \end{array}$$

and the step size parameters are all equal to $\Delta_k=\frac{1}{2^\ell}$.

Proof: The proof is done by induction on ℓ. We already verified in Table 4 that the result holds for the even cycle $\ell=0$, (iterations 0 to 2) and for the odd cycle $\ell=1$ (iterations 3 to 6).

Suppose that the even cycle ℓ is initiated with $\Delta_k=\frac{1}{2^\ell}$ and $\mathrm{x}_k=\left(0,\frac{2\Delta_k}{3}\right)$ for $k=2^\ell+2\ell-1$. The current objective function value is $f(\mathrm{x}_k)=\frac{4\Delta_k^2}{9}$. Table 5 details the objective function values of all trial points associated to the first four columns of P_k for each iterations of the even cycle ℓ. The successful trial points appear in bold face letters. All iterations but the last are successful, therefore the step size parameter remains constant.

k	x_k	$f(x_k+\Delta_k,y_k)$ $f(x_k-\Delta_k,y_k)$	$f(x_k,y_k+\Delta_k)$ $f(x_k,y_k-\Delta_k)$
$2^\ell+2\ell-1$	$(0,\frac{2\Delta_k}{3})$	$(\Delta_k^2+1)\frac{4\Delta_k^2}{9}$ $(\Delta_k^2+1)\frac{4\Delta_k^2}{9}$	$\frac{25\Delta_k^2}{9}$ $\mathbf{\frac{2\Delta_k^2}{9}}$
$2^\ell+2\ell$	$(0,\frac{-\Delta_k}{3})$	$\mathbf{((1-\Delta_k)^2+1)\frac{\Delta_k^2}{9}}$ $((1+\Delta_k)^2+1)\frac{\Delta_k^2}{9}$	$\frac{4\Delta_k^2}{9}$ $\frac{32\Delta_k^2}{9}$
$2^\ell+2\ell+j$ $1\le j<2^\ell$	$(j\Delta_k,\frac{-\Delta_k}{3})$	$\mathbf{\left((1-(j+1)\Delta_k)^2+1\right)\frac{\Delta_k^2}{9}}$ $\left((1-(j-1)\Delta_k)^2+1\right)\frac{\Delta_k^2}{9}$	$(j^2\Delta_k^2+1)\frac{4\Delta_k^2}{9}$ $\left((1-j\Delta_k)^2+1\right)\frac{16\Delta_k^2}{9}$
$2^\ell+2\ell+2^\ell$	$(1,\frac{-\Delta_k}{3})$	$(\Delta_k^2+1)\frac{\Delta_k^2}{9}$ $(\Delta_k^2+1)\frac{\Delta_k^2}{9}$	$\frac{8\Delta_k^2}{9}$ $\frac{16\Delta_k^2}{9}$

Table 5: Objective function values of the trial points of the even cycle ℓ.

The odd cycle $\ell'=\ell+1$ starts at iteration $k'=2^{(\ell+1)}+2(\ell+1)-1=(2^\ell+2\ell+2^\ell)+1$. The last iteration appearing in Table 5 is unsuccessful, and therefore $\Delta_{k'}=\frac{\Delta_k}{2}=\frac{1}{2^{\ell'}}$ and $x_{k'}=\left(1,\frac{-2\Delta_{k'}}{3}\right)$.

We have shown that the result is true for even values of ℓ. The proof for odd cycles is similar, and is omitted. ■

For this example, there are infinitely many accumulation points of the sequence of iterates. Every point of the set $\left\{\left(\frac{j}{2^\ell},0\right):j,l\in \mathrm{N}; j\le 2^\ell\right\}$ is an accumulation point.

References

[1] Lewis R.M. and Torczon V.(1996), "Pattern search algorithms for bound constrained minimization," *ICASE NASA Langley Research Center* TR 96-20. To appear in SIAM Journal on Optimization.

[2] Lewis R.M. and Torczon V.(1996), "Rank ordering and positive basis in pattern search algorithms," *ICASE NASA Langley Research Center* TR 96-71.

[3] Lewis R.M. and Torczon V.(1998), "Pattern search methods for linearly constrained minimization," *ICASE NASA Langley Research Center* TR 98-3.

[4] Torczon V.(1997), "On the Convergence of Pattern Search Algorithms," *SIAM Journal on Optimization* Vol.7 No.1, 1–25.