Yugoslav Journal of Operations Research
10 (2000), Number 2, 257-268

A MODIFIED STEP-LENGTH ALGORITHM IN NONLINEAR PROGRAMMING*

Nada I. \URANOVI] - MILI^I]
Department of Mathematics
Faculty of Technology and Metallurgy
University of Belgrade, Belgrade, Yugoslavia

Abstract: In this paper we consider a modification of the Armijo step-length algorithm based on so-called "forcing functions". It is proved that this modified algorithm is well-defined. Proof is given of the convergence of the obtained sequence of points to a first-order point of the problem of unconstrained optimization, as well as an estimate of the rate of convergence. Finally, numerical results obtained using of TURBO PASCAL programmes are given.

Keywords: Nonlinear programming, forcing function, step-length algorithm.

1. INTRODUCTION

In this paper we shall be concerned with finding solutions to the problem of unconstrained optimization:

$$\min\{j(x) \mid x \in D\} \tag{1}$$

where: $j : D \subset R^n \to R$ is a continuously differentiable function on the open set D.

We consider iterative algorithms for finding an optimal solution to problem (1) generating sequences of points x_k of the form:

$$x_{k+1} = x_k - a_k s_k, \quad k = 1,2,\ldots, \tag{2}$$

where we suppose that the direction vector s_k satisfies the condition

* This research was supported by the Science Fund of Serbia, grant number 04M03, through Institute of Mathematics, SANU

$$s_k \neq 0, \quad \left\langle \nabla j(x_k), s_k \right\rangle \geq 0, \tag{3}$$

and the step-length a_k is defined by a special step-length algorithm.

The original Armijo step-length algorithm [3] defines the step-length a_k for the sequence $\{x_k\}$ satisfying relations (2) and (3) in the following way:

$$a_k = 0 \quad \text{if} \quad \left\langle \nabla j(x_k), s_k \right\rangle = 0;$$

otherwise, $a_k > 0$ is a number satisfying

$$a_k = 2^{-i(k)},$$

where $i(k)$ is the smallest integer from $i = 0,1,\ldots,$ such that

$$x_k - 2^{-i(k)} s_k \in D$$

and

$$j(x_k) - j(x_k - 2^{-i(k)} s_k) \geq g \cdot 2^{-i(k)} \left\langle \nabla j(x_k), s_k \right\rangle,$$

where $0 < g < 1$ is a preassigned constant.

Now we shall give some definitions and lemmas, which will be necessary in the following text.

Definition 1. (See [4]). A mapping $s : [0,\infty) \to [0,\infty)$ is a forcing function if for any sequence $t_k \subset [0,\infty)$

$$\lim_{k\to\infty} s(t_k) = 0 \quad \text{implies} \quad \lim_{k\to\infty} t_k = 0$$

and $s(t) > 0$ for $t > 0$.

(The concept of the forcing function was first introduced by Elkin in [2].)

Definition 2. (See [4]). Let $\{x_k\} \subset R^n$ be any sequence that converges to $\overline{x}$. Then the root-convergence factors of R-factors, for short, of the sequence are defined, as follows:

$$R_p\{x_k\} = \begin{cases} \lim\limits_{k\to\infty} \sup \| x_k - \overline{x} \|^{\frac{1}{k}} & \text{if} \quad p = 1; \\ \lim\limits_{k\to\infty} \sup \| x_k - \overline{x} \|^{\frac{1}{p^k}} & \text{if} \quad p > 1. \end{cases}$$

If $0 < R_1\{x_k\} < 1$, the sequence $\{x_k\}$ is said to converge to $\overline{x}$ at least R-lineary.

Lemma 1. Let $j : D \subset R^n \to R$ be continuously differentiable on an open set $D_0 \subset D$ and suppose that $\{x_k\} \subset D_0$ converges to $x^* \in D_0$. Assume that $\nabla j(x^*) = 0$, that j has a second derivative at x^* and the Hessian matrix $H(x^*)$ is invertible and that there is some $h > 0$ and k_0 for which

$$j(x_k) - j(x_{k+1}) \geq h \| \nabla j(x_k) \|^2, \quad \forall k \geq k_0 .$$

Then $R_1\{x_k\} < 1$.

Proof: See [4].

Lemma 2. (See [5]). Let $j : D \subset R^n \to R$ be twice continuously differentiable and let there exist some $m > 0$, $0 < m < \infty$ such that

$$m \| y \|^2 \leq \langle y, H(x)y \rangle \text{ for all } x \in D, y \in R^n .$$

Then the function j is strongly convex and the set $\{x \mid j(x) \leq j(x_0)\}$ is bounded for any $x_0 \in D$.

Proof: See [5].

2. A MODIFICATION OF THE ARMIJO STEP-LENGTH ALGORITHM

We consider a sequence of points $\{x_k\}$ with properties (2) and (3) where the step a_k is defined in the following way

$$a_k = 0 \quad \text{if} \quad \langle \nabla j(x_k), s_k \rangle = 0 ;$$

otherwise, $a_k > 0$ is a number satisfying

$$a_k = q^{-i(k)}, \quad q > 1 ,$$

where $i(k)$ is the smallest integer from $i = 0,1,\ldots,$ such that

$$x_k - q^{-i(k)} s_k \in D \tag{4A}$$

and

$$j(x_k) - j(x_k - a_k s_k) \geq a_k s_k(\| \nabla j(x_k) \|) , \tag{4B}$$

where $s_k : [0,\infty) \to [0,\infty)$ is a sequence of forcing functions such that $s_k(t) \leq dt$ for every $t \geq 0$, and some $0 < d < 1$, $k = 0,1,2,\ldots,$.

Since D is open, $x_k - q^{-i} s_k \in D$ for sufficiently large i. The existence of a finite $i(k)$ such that $a_k = q^{-i(k)}$ satisfies (4) is proved in the following lemma.

Lemma 3. Let $j : D \subset R^n \to R$ be continuously differentiable on an open set $D_0 \subset D$ and suppose that $x_k, a_k = q^{-i}$ and s_k satisfy $\langle \nabla j(x_k), s_k \rangle > 0$, $[x_k, x_k - q^{-i} s_k] \subset D$, $\nabla j(x_k) \neq 0$ and

$$j(x_k) - j(x_k - q^{-i} s_k) < q^{-i} s_k(\| \nabla j(x_k) \|),$$

for some $i(=0,1,\ldots)$ and $q > 1$, where $s_k : [0,\infty) \to [0,\infty)$ is a sequence of forcing functions such that $s_k(t) \le s t$ for every $t \ge 0$, and some $0 < s < 1$, $k = 0,1,2,\ldots$. Assume that

$$\langle \nabla j(x_k), s_k \rangle \ge n_k(\| \nabla j(x_k) \|), \quad \nabla j(x_k) \neq 0,$$

where $n_k : [0,\infty) \to [0,\infty)$ is a sequence of forcing functions such that $n_k(t) \ge b\, t$ for some $b > 0$, $k = 0,1,2,\ldots$.

Then there exists a finite $i^* > i$ such that (4) holds.

Proof: Define the function $F : [0,1] \to R$ in the following way:

$$F(l) = \begin{cases} \dfrac{j(x_k) - j(x_k - l \cdot q^{-i} s_k)}{l q^{-i} \langle \nabla j(x_k), s_k \rangle}, & l \in (0,1]; \\ 1, & l = 0. \end{cases}$$

By L'Hospital's rule, we have $F(l) \to 1$ as $l \to 0$, so that F is continuous on $[0,1]$. Consequently, since $F(0) = 1$

$$F(1) < \frac{s_k(\| \nabla j(x_k) \|)}{\langle \nabla j(x_k), s_k \rangle} < \frac{d \| \nabla j(x_k) \|}{n_k(\| \nabla j(x_k) \|)} < \frac{d \| \nabla j(x_k) \|}{b \| \nabla j(x_k) \|} = \frac{d}{b}.$$

Therefore, F takes on all values between $\dfrac{d_k(\| \nabla j(x_k) \|)}{\langle \nabla j(x_k), s_k \rangle}$ and $\dfrac{d}{b}$. Hence, there exists a finite $\bar{i}$, $\bar{l} = q^{-\bar{i}}$, $0 < \bar{l} < 1$, such that

$$\frac{s_k(\| \nabla j(x_k) \|)}{\langle \nabla j(x_k), s_k \rangle} \le F(\bar{l}) < \frac{d}{b},$$

i.e. (4) will be satisfied for $a_k^* = q^{-\bar{i}} q^{-i}$, i.e. for $i^* = i + \bar{i}$.

Theorem 1. Let $j : D \subset R^n \to R$ be a continuously differentiable function on the open set D. Let the sequence $\{x_k\}$ be defined by relations (2), (3), and (4), where $s_k : [0,\infty) \to [0,\infty)$ is a sequence of forcing functions such that $d_1 t \le s_k(t) \le d_2 t$ for every $t \ge 0$ and some $0 < d_1 < d_2 < 1, k = 0,1,2,\ldots,$. Let $\overline{x} \in D$ be a point of accumulation of $\{x_k\}$ and K_1 as set of indices such that $x_k \to \overline{x}$ for $k \in K_1$. Assume that

$$\langle \nabla j(x_k), s_k \rangle \ge n_k(\| \nabla j(x_k) \|) \text{ for } k \in K_1 \text{ and } \nabla j(x_k) \ne 0 .$$

where $n_k : [0,\infty) \to [0,\infty)$ is a sequence of forcing functions such that $n_k(t) \ge b\, t$, for some $b > 0$, $d_2 < b < 1$, $k \in K_1$. Assume that the sequence $\{s_k\}$ is uniformly bounded ($k \in K_1$) and

$$\| s_k \| \ge m_k(\| \nabla j(x_k) \|) \text{ for all } k \in K_1 ,$$

where $m_k : [0,\infty) \to [0,\infty)$ is a sequence of forcing functions such that $m_k(t) \ge m t$ for some $m > 0$, $k \in K_1$. Then $\nabla j(\overline{x}) = 0$.

Proof: There are two cases to consider.

a) The set of indices $\{i(k)\}$ for $k \in K_1$ is uniformly bounded above by number I.

Since, by (4) the sequences $\{j(x_k)\}$ is monotone decreasing, it follows that

$$j(x_0) - j(\overline{x}) \ge \sum_{k \in K_1} [j(x_k) - j(x_{k+1})] \ge \sum_{k \in K_1} q^{-i(k)} s_k(\| \nabla j(x_k) \|) \ge$$

$$\ge q^{-I} \sum_{k \in K_1} d_1 \| \nabla j(x_k) \| = d_1 q^{-I} \sum_{k \in K_1} \| \nabla j(x_k) \|.$$

Since $j(\overline{x})$ is finite and since $\| \nabla j(x_k) \| \ge 0$, it follows that $\nabla j(x_k) \to 0$ for $k \in K_1$. Hence, by continuity of ∇j, we have $\nabla j(\overline{x}) = 0$.

b) There is a subset of indices $K_2 \subset K_1$ such that $\lim_{k \in K_2} i(k) = \infty$. Because of the definition of $i(k)$, then either

$$x_k + q^{-i(k)+1} s_k \notin D$$

or

$$j(x_k) - j(x_k - q^{-i(k)+1} s_k) < q^{-i(k)+1} s_k(\| \nabla j(x_k) \|) . \tag{5}$$

If the cause of termination of iteration k were that $x_k + q^{-i(k)+1} s_k \notin D$ infinitely often, then, since $i(k) \to \infty$ for $k \in K_2$ and because $\{s_k\}$ are uniformly bounded, it follows that $\bar{x}$ is on the boundary of D. Since D is an open set, $\bar{x} \notin D$, a contradiction to the theorem assumption.

Thus, without generality (5) can be considered to hold for all $k \in K_2$.

Because j is by assumption continuously differentiable and $\{s_k\}$ uniformly bounded, it follows that (5) can be written as

$$j(x_k) - j(x_k - q^{-i(k)+1} s_k) = q^{-i(k)+1} \langle \nabla j(x_k), s_k \rangle + o(q^{-i(k)+1} \| s_k \|) <$$
$$< q^{-i(k)+1} s_k (\| \nabla j(x_k) \|) \le q^{-i(k)+1} d_2 \| \nabla j(x_k) \| .$$

Hence,

$$q^{-i(k)+1} [\langle \nabla j(x_k), s_k \rangle - d_2 \| \nabla j(x_k) \|] < o(q^{-i(k)+1} \| s_k \|) . \tag{6}$$

Because of the assumptions of the theorem it follows that

$$\langle \nabla j(x_k), s_k \rangle - d_2 \| \nabla j(x_k) \| \ge n_k (\| \nabla j(x_k) \|) - d_2 \| \nabla j(x_k) \| \ge$$
$$\ge b \| \nabla j(x_k) \| - d_2 \| \nabla j(x_k) \| = (b - d_2) \| \nabla j(x_k) \| . \tag{7}$$

From (6) and (7) it follows that:

$$q^{-i(k)+1} (b - d_2) \| \nabla j(x_k) \| \le q^{-i(k)+1} [\langle \nabla j(x_k), s_k \rangle - d_2 \| \nabla j(x_k) \|] <$$
$$< o(q^{-i(k)+1} \| s_k \|).$$

Dividing by $q^{-i(k)+1} \| s_k \|$ yields

$$\frac{o(q^{-i(k)+1} \| s_k \|)}{q^{-i(k)+1} \| s_k \|} > \frac{(b - d_2) \| \nabla j(x_k) \|}{\| s_k \|} \tag{8}$$

Since, by assumption, the sequence $\{s_k\}$ is uniformly bounded $(k \in K_1)$ it follows that there exists some $M > 0$ such that $\| s_k \| \le M, k \in K_1$. Hence, from (8) it follows that

$$\frac{o(q^{i(k)+1} \| s_k \|)}{q^{i(k)+1} \| s_k \|} > \frac{(b - d_2) \| \nabla j(x_k) \|}{\| s_k \|} \ge \frac{b - d_2}{M} \| \nabla j(x_k) \| .$$

Because of the uniform boundness of $\{s_k\}$, taking the limit as $k \to \infty$ for $k \in K_2$ yields, by continuity of ∇j :

$$\frac{b-d_2}{M}\|\nabla j(x_k)\| \leq 0$$

Since $b > d_2$, $M > 0$, it follows that $\nabla j(\overline{x}) = 0$.

Theorem 2. Let the assumptions of Theorem 1 be satisfied. Let additionally, the function $j : D \subset R^n \to R$ be twice continuously differentiable such that there exists some $l > 0$ satisfying

$$l\|y\|^2 \leq \langle y, H(x)y\rangle \text{ for all } x \in D, y \in R^n \tag{9}$$

then the sequence $\{x_k\}$ generated by the modified Armijo algorithm converges to $\overline{x}$, where $\overline{x}$ is the unique optimal solution of problem (1), at least R-linearly.

Proof: From condition (9) it follows that the function j is, by Lemma 2, strongly convex and that the level set $C = \{x \in D \mid j(x) \leq j(x_0)\}$ for some $x_0 \in D$ is convex and compact. Furthermore, since j is twice continuously differentiable from (9) it also follows that there exists some $L > 0$, $L \geq l$ such that

$$l\|y\|^2 \leq \langle y, H(x)y\rangle \leq L\|y\|^2 \text{ for all } x \in C, y \in R^n. \tag{10}$$

From relation (4) we have that $j(x_k) > j(x_{k+1})$; hence $x_{k+1} \in C$ if $x_k \in C$.

By Theorem 1 we have that $\|\nabla j(\overline{x})\| = 0$.

By strong convexity of j it follows that $\overline{x}$ is the unique optimal solution of problem (1).

Denote by F the following function:

$$F(x, s, a) = j(x) - j(x - as).$$

By Taylor's theorem we have

$$F(x_k, s_k, a) = a\langle \nabla j(x_k), s_k\rangle - a^2\int_0^1 (1-t)\langle s_k, H(x_k - tas_k)s_k\rangle dt.$$

From this equation, in view of (10), it follows that

$$F(x_k, s_k, a) \geq a\langle \nabla j(x_k), s_k\rangle - \frac{1}{2}a^2 L\|s_k\|^2. \tag{11}$$

Using the function F because of the definition of $i(k)$ we have:

$$F(x_k, s_k, q^{-i(k)+1}) < q^{-i(k)+1} s_k(\|\nabla j(x_k)\|), \quad k = 0,1,2,\ldots,.$$

i.e. by introducing the function $\overline{F}$:

$$\overline{F}(x_k, s_k, q^{-i(k)+1}) = F(x_k, s_k, q^{-i(k)+1}) - q^{-i(k)+1} s_k(\| \nabla j(x_k) \|) < 0 .$$

Now from (11) it follows that

$$\overline{F}(x_k, s_k, q^{-i(k)+1}) \geq q^{-i(k)+1} \langle \nabla j(x_k), s_k \rangle - \frac{1}{2} q^{-2i(k)+2} L \| s_k \|^2 -$$
$$- q^{-i(k)+1} s_k(\| \nabla j(x_k) \|), \quad k = 0,1,2,\ldots$$
$$\geq q^{-i(k)+1} n_k(\| \nabla j(x_k) \|) - \frac{1}{2} q^{-2i(k)+2} L \| s_k \|^2 - q^{-i(k)+1} d_2(\| \nabla j(x_k) \|), \quad k \in K_1$$
$$\geq q^{-i(k)+1} b \| \nabla j(x_k) \| - q^{-i(k)+1} d_2(\| \nabla j(x_k) \|) - \frac{1}{2} q^{-2i(k)+2} L \| s_k \|^2 < 0.$$

From the last inequality it follows that

$$q^{-i(k)+1}[(b - d_2) \| \nabla j(x_k) \| - \frac{1}{2} q^{-i(k)+1} L \| s_k \|^2] < 0, \quad k \in K_1 .$$

Hence,

$$q^{-i(k)} > \frac{2(b - d_2) \| \nabla j(x_k) \|}{q \cdot L \cdot \| s_k \|^2} \geq \frac{2(b - d_2) \| \nabla j(x_k) \|}{L \cdot q \cdot M^2}$$

because $\| s_k \| \leq M$, $k \in K_1$, since $\{s_k\}$ is, by assumption, uniformly bounded.

Finally, from (4b) it follows that:

$$j(x_k) - j(x_k - q^{-i(k)} s_k) \geq q^{-i(k)} s_k(\| \nabla j(x_k) \|) \geq$$
$$\geq \frac{2(b - d_2) d_1}{L \cdot q \cdot M^2} \| \nabla j(x_k) \|^2 = h \| \nabla j(x_k) \|^2, \quad k \in K_1,$$

where $h = \dfrac{2(b - d_2) d_1}{L \cdot q \cdot M^2} > 0$.

From the last inequality, by Lemma 2, it follows that the sequence $\{x_k\}$ converges to $\overline{x}$ at least R-linearly (where $\overline{x}$ is, as we have already proved, the unique optimal solution of problem (1)).

3. COMPUTATIONAL RESULTS

We shall give results obtained by computer programmes written in TURBO PASCAL for minimizing the functions:

$$j(x,a) = 10(x_2 - x_1^2)^2 + a \cdot (1 - x_1)^2 , \qquad (I)$$

$$j(x,a) = (x_1 + 2ax_2)^2 + a \cdot (x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 2a \cdot (x_1 - x_4)^4 , \qquad (II)$$

$$j(x,a) = (x_1 + 10x_2)^2 + (x_3 - x_4)^2 + a \cdot (x_2 - 2x_3)^4 + a \cdot (x_1 - x_4)^4 , \qquad (III)$$

for $a = 1,2,\dots,10$.

The direction vector $s_k = \nabla j(x_k)$, the initial point for (I) is $x^0 = (-1,2;1)$ and for (II) and (III) it is $x^0 = (-3,-1,0,1)$.

In the following tables we present some application results of the original and modified Armijo algorithm. Here, i means the number of iterations, f the number of computing functions (and gradients), * means the demanded accuracy ($e = 0.00001$) is not achieved even in 300 iterations.

When minimizing the functions (I) we have the following results. Table 1 gives the results of the original algorithm. The modified Armijo algorithm gives very good results. They are presented in Table 2 for different forcing functions.

Table 1: Original algorithm for functions (I)

$a \backslash g$		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$a = 1$	i	173	184	177	179	183	160	145	147	141
	f	1071	1145	1101	1120	1163	1027	933	991	1012
$a = 2$	i	117	115	127	126	110	111	110	111	121
	f	744	738	811	807	713	726	724	767	860
$a = 3$	i	107	114	65	110	92	94	94	82	100
	f	722	735	421	713	600	612	618	592	724
$a = 4$	i	67	67	86	28	83	72	83	73	99
	f	443	439	563	176	540	473	551	522	713
$a = 5$	i	71	91	93	59	72	67	57	62	83
	f	481	598	611	517	471	440	393	445	612
$a = 6$	i	73	66	84	58	41	61	60	58	83
	f	486	440	555	382	267	401	404	413	606
$a = 7$	i	68	57	73	55	53	49	55	56	74
	f	458	387	485	360	346	327	376	389	543
$a = 8$	i	23	17	53	50	40	45	39	52	76
	f	159	117	349	328	265	301	256	361	560
$a = 9$	i	42	43	16	44	42	46	42	52	76
	f	290	297	108	289	277	305	287	360	573
$a = 10$	i	30	30	30	44	40	40	38	48	72
	f	206	207	205	291	275	268	260	332	536

Table 2: Modified algorithm for functions (I)

$s_{(t)} \backslash a$		1	2	3	4	5	6	7	8	9	10
$\frac{t}{t+2}$	i	9	10	5	11	16	13	7	4	14	8
	f	179	84	63	100	226	208	206	108	118	88
$\frac{t}{2t+3}$	i	9	22	9	15	23	29	14	6	16	8
	f	178	154	176	213	180	216	212	111	131	172

When minimizing the functions (II) we have the following. (For $a = 5$ we have Fletcher-Powell's function, often used as a test function). Table 3 gives the results of the original algorithm. The modified algorithm gives excellent results as can be seen from Table 4.

Table 3: Original algorithm for functions (II)

$a \backslash g$		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$a = 1$	i	45	38	48	39	50	53	53	56	83
	f	163	134	175	145	205	222	250	302	560
$a = 2$	i	47	38	48	43	51	48	56	67	102
	f	226	184	234	212	269	272	327	432	776
$a = 3$	i	101	105	105	102	94	83	89	94	123
	f	598	621	622	610	591	549	607	693	1003
$a = 4$	i	53	71	58	46	36	49	59	76	150
	f	343	461	380	298	264	354	431	590	1249
$a = 5$	i	269	*	*	*	253	214	217	176	195
	f	2054	*	*	*	1939	1645	1720	1524	1835
$a = 6$	i	162	194	201	178	171	132	137	144	182
	f	1250	1487	1539	1369	1371	1105	1158	1317	1752
$a = 7$	i	*	*	*	*	*	*	*	273	271
	f	*	*	*	*	*	*	*	2605	2717
$a = 8$	i	60	89	67	69	61	53	71	125	186
	f	515	760	544	636	499	449	625	1141	1834
$a = 9$	i	*	*	*	*	*	*	*	*	*
	f	*	*	*	*	*	*	*	*	*
$a = 10$	i	*	*	*	*	*	*	*	*	*
	f	*	*	*	*	*	*	*	*	*

Table 4: Modified algorithm for functions (II)

$s_{(t)}$ \ a		1	2	3	4	5	6	7	8	9	10
$\frac{t}{t+2}$	i	8	14	26	21	67	53	125	46	150	218
	f	70	107	202	174	551	445	1100	428	1409	2124
$\frac{0.5t}{1+t^2}$	i	8	15	26	23	66	58	123	46	148	214
	f	74	109	202	191	543	488	1092	416	1390	2084
$\ln(1+t)$	i	6	10	18	15	51	37	80	38	96	135
	f	63	96	153	137	354	325	723	367	919	1329
$0.9\sin t$	i	6	10	16	18	38	44	75	48	90	126
	f	68	95	144	160	330	386	684	421	865	1245

When minimizing the functions (III) we have the following. (For $a = 1,2,3,4,5$ and $g = 0.1; 0.2; ...; 0.7$ and for $a = 6,7,8,9,10$ and $g = 0.1; 0.2; ...; 0.5$ the demanded accuracy ($e = 10^{-5}$) is not achieved by the original algorithm even in 3000 iterations. For the remaining values of a and g the number is around 300 and f is around 2400. Meanwhile, the modified algorithm gives excellent results as can be seen from Table 5.

Table 5: Modified algorithm for functions (III)

$s_{(t)}$ \ a		1	2	3	4	5	6	7	8	9	10
$0.9\sin t$	i	34	34	24	26	47	32	33	34	37	40
	f	319	315	240	259	427	310	315	324	350	367
$\ln(1+t)$	i	37	37	30	33	49	48	39	41	40	43
	f	342	340	289	313	438	427	359	375	368	392

4. CONCLUSION

When compared to the original algorithm, this modification of the Armijo algorithm is superior because it demands a smaller number of iterations, and a smaller number of function (and gradient) evaluations than the original algorithm (as can seen from the previous tables).

Furthermore, because of the general assumptions about the objective function the modified algorithm can be used to solve a wide class of unconstrained optimization problems.

Also, there is a wide choice of forcing functions $s(t)$ with the property $s(t) \le dt$, $0 < d < 1$.

Finally, this modified algorithm can be used to solve constrained optimization problems (see [1]), when constraints are adequately considered.

REFERENCES

[1] Djuranovi}-Mili~i}, N.I., "An algorithm in constrained optimization" in: M. Thoma, A. Wyner (eds.), Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin, 1986, 203-208.

[2] Elkin, R., "Convergence theorems for Gauss-Siedel and other minimization algorithms", Doctoral Thesis, University of Maryland, College Park, 1968.

[3] McCormick, G.P., Nonlinear Programmings, Theory, Algorithm and Applications, Wiley, New York, 1983.

[4] Ortega, J., and Rheinboldt, W., Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[5] Polak, E., Computational Methods in Optimization, A Unified Approach, Academic Press, New York, 1971.