

B. RHANIZAR (Rabat and Marrakech)

NONLINEAR MULTIPLE HYBRID PROCEDURES FOR SOLVING SOME CONSTRAINED NONLINEAR OPTIMIZATION PROBLEMS

Abstract. We introduce a new formulation of multiple hybrid procedures which consist in a combination of k arbitrary approximate solutions. The connection between this method and other vector sequence transformations is studied. This connection is also exploited for solving some constrained nonlinear optimization problems. A convergence acceleration result is established and numerical examples are given.

1. Introduction. It is well known that hybrid procedures can be used for solving a system of linear equations [1, 6]. Multiple hybrid procedures were defined by Brezinski and Redivo-Zaglia [6] and their acceleration properties have been studied. The nonlinear hybrid procedures studied in [5] are also used for accelerating fixed point iterations. In this paper, we will consider nonlinear multiple hybrid procedures for solving some constrained nonlinear optimization problems. In order to present conveniently our results, let us introduce the setting used throughout this paper. For the vector $x = (x_1, \dots, x_p) \in \mathbb{R}^p$, we use the Euclidean norm

$$\|x\| = \left(\sum_{1 \leq i \leq p} x_i^2 \right)^{1/2}.$$

We also denote by (\cdot, \cdot) the corresponding inner product. We use the matrix norm

$$\|A\| = \sup_{\|x\| \neq 0} \|Ax\|/\|x\|$$

for any $q \times p$ matrix A . The paper is organized as follows. Section 2 is devoted to a new formulation of multiple hybrid procedures and the connection

2000 *Mathematics Subject Classification*: Primary 90C25.

Key words and phrases: hybrid procedures, nonlinear optimization.

between this transformation and other vector sequence transformations. In Section 3, we study the convergence acceleration method for solving some constrained nonlinear optimization problems. Some numerical examples are given in Section 4.

2. Nonlinear multiple hybrid procedures. Consider the system of linear equations

$$(2.1) \quad Ax = b.$$

Hybrid procedures for solving (2.1) were defined by Abkowitz and Brezinski [1] and by Brezinski and Redivo-Zaglia [6]. Let x_1, \dots, x_k be k approximate solutions of (2.1). *Multiple hybrid procedures* construct a new approximate solution by setting

$$y_k = \sum_{i=1}^k \alpha_i x_i \quad \text{with} \quad \sum_{i=1}^k \alpha_i = 1$$

where $\alpha_1, \dots, \alpha_k$ are chosen to minimize $\|\varrho_k\|$, with

$$\varrho_k = \sum_{i=1}^k \alpha_i r_i \quad \text{and} \quad r_i = b - Ax_i.$$

Consider the nonlinear system

$$(2.2) \quad F(x) = 0$$

where $F : \mathbb{R}^p \rightarrow \mathbb{R}^p$. We suppose that there exists a solution x^* of (2.2). Let $x_n^{(1)}, \dots, x_n^{(k)}$ be k sequences of vectors of \mathbb{R}^p converging to x^* with $k \leq p$. We construct a new sequence by setting

$$(2.3) \quad t_n^{(k)} = \sum_{i=1}^k \alpha_n^{(i)} x_n^{(i)} \quad \text{with} \quad \sum_{i=1}^k \alpha_n^{(i)} = 1.$$

We make the following definition:

DEFINITION 2.1. A map $r : \mathbb{R}^p \rightarrow \mathbb{R}^p$ is said to be a *residual function* if

$$\forall x \in \mathbb{R}^p \exists \lambda \in \mathbb{R}^* \quad r(x) = \lambda F(x).$$

REMARK 2.1. If we assume that (2.2) has a unique solution x^* , then

$$r(x) = 0 \Leftrightarrow x = x^*,$$

and we get the definition given by Brezinski and Chehab [5].

Let $r^{(1)}, \dots, r^{(k)}$ be k residual functions, and set

$$r_n^{(i)} = r^{(i)}(x_n^{(i)}), \quad i = 1, \dots, k.$$

Setting

$$(2.4) \quad \beta_n^{(j)} = 1 - \sum_{i=1}^{j-1} \alpha_n^{(i)}, \quad j = 2, \dots, k,$$

we obtain from (2.3)

$$(2.5) \quad t_n^{(k)} = x_n^{(1)} + \sum_{j=2}^k \beta_n^{(j)} (x_n^{(j)} - x_n^{(j-1)}).$$

We define

$$(2.6) \quad \varrho_n^{(k)} = r_n^{(1)} + \sum_{j=2}^k \beta_n^{(j)} (r_n^{(j)} - r_n^{(j-1)}).$$

The coefficients $\beta_n^{(2)}, \dots, \beta_n^{(k)}$ are chosen to minimize $\|\varrho_n^{(k)}\|$. We introduce some notations.

We define the following matrices by giving their columns:

$$\Delta_{n,k}^1 = [x_n^{(2)} - x_n^{(1)}, \dots, x_n^{(k)} - x_n^{(k-1)}], \quad \Delta_{n,k}^2 = [\Delta^2 x_n^{(1)}, \dots, \Delta^2 x_n^{(k-1)}],$$

where $\Delta^2 x_n^{(i)} = \Delta(\Delta x_n^{(i)})$ and $\Delta x_n^{(i)} = x_n^{(i+1)} - x_n^{(i)}$; moreover,

$$R_{n,k} = [r_n^{(1)}, \dots, r_n^{(k-1)}], \quad \Delta R_{n,k} = [r_n^{(2)} - r_n^{(1)}, \dots, r_n^{(k)} - r_n^{(k-1)}],$$

$$(\Delta R_{n,k})^\dagger = ((\Delta R_{n,k})^T \Delta R_{n,k})^{-1} (\Delta R_{n,k})^T, \quad \text{the pseudo-inverse of } \Delta R_{n,k}.$$

We also set

$$\beta_{n,k} = (\beta_n^{(2)}, \dots, \beta_n^{(k)})^T \in \mathbb{R}^{k-1}.$$

The following theorem shows that $(t_n^{(k)})_n$ defined by (2.3) is well defined, and gives a new formula for it.

THEOREM 2.1. *If $\text{rank}(\Delta R_{n,k}) = k - 1$, then there exists a unique vector $\beta_{n,k}$ which solves*

$$(2.7) \quad (\Delta R_{n,k})^T \Delta R_{n,k} \beta_{n,k} = -(\Delta R_{n,k})^T r_n^{(1)}$$

and

$$(2.8) \quad t_n^{(k)} = x_n^{(1)} - \Delta_{n,k}^1 (\Delta R_{n,k})^\dagger r_n^{(1)}.$$

Proof. Using the previous notations and (2.6), we have

$$\varrho_n^{(k)} = r_n^{(1)} + \Delta R_{n,k} \beta_{n,k}.$$

Then

$$\begin{aligned} \|\varrho_n^{(k)}\|^2 &= \|r_n^{(1)}\|^2 + 2(r_n^{(1)}, \Delta R_{n,k} \beta_{n,k}) + \|\Delta R_{n,k} \beta_{n,k}\|^2 \\ &= \|r_n^{(1)}\|^2 + 2((\Delta R_{n,k})^T r_n^{(1)}, \beta_{n,k}) + ((\Delta R_{n,k})^T \Delta R_{n,k} \beta_{n,k}, \beta_{n,k}). \end{aligned}$$

Therefore, the vector $\beta_{n,k}$ minimizing $\|\varrho_n^{(k)}\|^2$ is such that

$$(\Delta R_{n,k})^T \Delta R_{n,k} \beta_{n,k} = -(\Delta R_{n,k})^T r_n^{(1)}.$$

By the previous notations and (2.5), we have

$$t_n^{(k)} = x_n^{(1)} + \Delta_{n,k}^1 \beta_{n,k}.$$

On the other hand, it is well known that the hypothesis $\text{rank}(\Delta R_{n,k}) = k - 1$ implies that $(\Delta R_{n,k})^T \Delta R_{n,k}$ is nonsingular. Then we finally have

$$\beta_{n,k} = -(\Delta R_{n,k})^\dagger r_n^{(1)} \quad \text{and} \quad t_n^{(k)} = x_n^{(1)} - \Delta_{n,k}^1 (\Delta R_{n,k})^\dagger r_n^{(1)}. \blacksquare$$

REMARK 2.2. (1) Consider the optimization problem

$$(2.9) \quad \text{find } x^* \in \mathbb{R}^p \quad \text{such that} \quad f(x^*) = \min_{x \in \mathbb{R}^p} f(x)$$

where f is a convex function from \mathbb{R}^p to \mathbb{R} . Problem (2.9) is equivalent to (2.2) with $F = \nabla f$, the gradient of f .

If we set $x_{n+i} = x_{n+i-1} - \lambda_{n,i} \nabla f(x_{n+i-1})$, $x_n^{(i)} = x_{n+i-1}$, where $x_0 \in \mathbb{R}^p$, $n = 0, 1, \dots$, $i = 1, \dots, k - 1$ and $\lambda_{n,i}$ is the solution of $\min_{\lambda \in \mathbb{R}^*} f(x_{n+i-1} - \lambda \nabla f(x_{n+i-1}))$, then:

- For $k = 3$ and $r_n^{(i)} = -\lambda_{n,i} \nabla f(x_{n+i-1})$, $i = 1, 2$, we obtain (see [11]) $t_n^{(k)} = \varepsilon_2^{(n)}$, the vector ε -algorithm [3, 4].
- For $k = p$ and $r_n^{(i)} = -\nabla f(x_{n+i-1})$, $i = 1, \dots, p$, we get the modified Henrici transformation used in [12] for solving (2.9):

$$t_n^{(p)} = x_n - \Delta X_n (\Delta F'(x_n))^{-1} \nabla f(x_n),$$

where $\Delta F'(x_n)$ is the $p \times p$ matrix whose columns are

$$\nabla f(x_{n+1}) - \nabla f(x_n), \dots, \nabla f(x_{n+p}) - \nabla f(x_{n+p-1}).$$

(2) Consider the vector sequence defined by $x_{n+i} = x_{n+i-1} - \lambda F(x_{n+i-1})$, with $x_0 \in \mathbb{R}^p$, $\lambda > 0$, $i = 1, \dots, p$ and $n = 0, 1, \dots$. If we set

$$x_n^{(i)} = x_{n+i-1}, \quad r_n^{(i)} = -\lambda F(x_{n+i-1}),$$

then we have $r_n^{(i)} = \Delta x_n^{(i)}$, $r_n^{(i+1)} - r_n^{(i)} = \Delta^2 x_n^{(i)}$ and $\Delta R_{n,k} = \Delta_{n,k}^2$. We obtain the *RRE* (reduced rank extrapolation; see [13])

$$(2.10) \quad t_n^{(k)} = x_n - \Delta_{n,k}^1 (\Delta_{n,k}^2)^\dagger \Delta x_n.$$

3. Nonlinearly constrained optimization problems. Let us now apply the vector sequence transformation (2.8) to the solution of the constrained optimization problem

$$(3.1) \quad \begin{cases} \min f(x) \\ \text{subject to } g(x) \leq 0 \end{cases}$$

where the functions $f : \mathbb{R}^p \rightarrow \mathbb{R}$ and $g : \mathbb{R}^p \rightarrow \mathbb{R}^q$ (with $g(x) = (g_i(x))_{1 \leq i \leq q}$ and $q \leq p$) are convex and twice continuously differentiable. The Lagrangian is defined by

$$L(x, \lambda) = f(x) + \sum_{i=1}^q \lambda^i g_i(x)$$

where $x \in \mathbb{R}^p$ and $\lambda = (\lambda^1, \dots, \lambda^q) \in \mathbb{R}^q$.

It is well known that if a vector x^* is a solution of (3.1), then there exists a multiplier $\lambda^* \in \mathbb{R}_+^q$ such that

$$(3.2) \quad L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*) \quad \forall x \in \mathbb{R}^p \quad \forall \lambda \in \mathbb{R}_+^q.$$

We denote by J_g the Jacobian matrix of g , and we set $T(x_n) = t_n^{(p)}$.

Consider problem (3.2) and let $x_0 \in \mathbb{R}^p$ and $\lambda_0 \in \mathbb{R}_+^q$ be given vectors. Assuming that $(x_n, \lambda_n) = (x_n^1, \dots, x_n^p, \lambda_n^1, \dots, \lambda_n^q) \in \mathbb{R}^p \times \mathbb{R}_+^q$ is given, we consider the problem

$$L(x^*, \lambda^*) \leq L(x, \lambda^*) \quad \forall x \in \mathbb{R}^p.$$

One can obtain x_{n+1} such that

$$L(x_{n+1}, \lambda_n) \leq L(x_n, \lambda_n).$$

Then we apply the gradient method [8] to $\min_{x \in \mathbb{R}^p} L(x, \lambda_n)$, where x_n is considered as an initial point, and we define the vector sequence $(y_n^{(m)})_m$ by

$$y_n^{(1)} = x_n$$

for $m = 1, \dots$, do

$$r_n^{(m)} = \alpha(\nabla f(y_n^{(m)}) + (J_g(y_n^{(m)}))^T \lambda_n)$$

$$y_n^{(m+1)} = y_n^{(m)} - r_n^{(m)}$$

end do

where $\alpha > 0$. Then the sequence $(T(y_n^{(m)}))_m$ defined by (2.8) for $(y_n^{(m)})_m$ converges to a limit denoted by x_{n+1} .

In order to determine λ_{n+1} , we consider the problem

$$L(x^*, \lambda) \leq L(x^*, \lambda^*) \quad \forall \lambda \in \mathbb{R}_+^q.$$

One can obtain λ_{n+1} such that

$$L(x_{n+1}, \lambda_n) \leq L(x_{n+1}, \lambda_{n+1}).$$

Then we can take λ_{n+1} as the first iteration of the gradient method applied to $\max_{\lambda \in \mathbb{R}_+^q} L(x_{n+1}, \lambda)$, where λ_n is considered as an initial point. Then we have

$$\lambda_{n+1} = P(\lambda_n + \varrho g(x_{n+1}))$$

where $\varrho > 0$ and P is the projection onto \mathbb{R}_+^q defined by

$$P(x) = (\max(x^i, 0))_{1 \leq i \leq q} \quad \forall x \in \mathbb{R}^q.$$

The corresponding algorithm is as follows:

ALGORITHM 1.

- (a) Choose $x_0 \in \mathbb{R}^p$, $T(y_0^{(-1)}) = x_0$, $\lambda_0 \in \mathbb{R}_+^q$, $\alpha > 0$, $\varepsilon > 0$ and $\varrho > 0$
- (b) for $n = 0, 1, \dots$, do
- (b1) Set $y_n^{(1)} = x_n$
- for $i = 1, \dots, p$, do
- $r_n^{(i)} = \alpha(\nabla f(y_n^{(i)}) + (J_g(y_n^{(i)}))^T \lambda_n)$
- $y_n^{(i+1)} = y_n^{(i)} - r_n^{(i)}$
- end do
- (b2) for $i = 1, \dots$, do
- $r_n^{(p+i)} = \alpha(\nabla f(y_n^{(p+i)}) + (J_g(y_n^{(p+i)}))^T \lambda_n)$
- $y_n^{(p+i+1)} = y_n^{(p+i)} - r_n^{(p+i)}$
- solve the linear system $(\Delta R_{n,p})^T \Delta R_{n,p} \beta_{n,p} = (\Delta R_{n,k})^T r_n^{(i)}$
- compute $T(y_n^{(i)}) = y_n^{(i)} - \Delta_{n,p}^1 \beta_{n,p}$
- if $\|T(y_n^{(i)}) - T(y_n^{(i-1)})\| \leq \varepsilon$ then
- set $x_{n+1} = T(y_n^{(i)})$
- stop
- end if
- end do
- (b3) Computation of λ_{n+1}
- for $i = 1, \dots, q$, do
- $\lambda_{n+1}^i = \max(\lambda_n^i + \varrho g_i(x_{n+1}), 0)$.
- end do
- if $\|x_{n+1} - x_n\| \leq \varepsilon$ then
- set $x^* = x_{n+1}$
- stop
- end if
- end do.

As mentioned above, the following theorem shows that $(T(y_n^{(m)}))_m$ converges to x_{n+1} faster than $(y_n^{(m)})_m$. Furthermore, we shall see that the vector sequence $(x_n)_n$ obtained by this procedure converges to the solution of (3.1).

We suppose that $g(x) = Cx - d$, and $\{x \in \mathbb{R}^p : g(x) < 0\} \neq \emptyset$, where $C \in M_{q,p}(\mathbb{R})$ and $d \in \mathbb{R}^q$. This case occurs frequently in practice.

Let us now study the convergence and acceleration of the sequence transformation $(T(y_n^{(m)}))_m$. We denote by $\nabla^2 f$ the Hessian of f . The following theorem shows the acceleration of convergence obtained by $(T(y_n^{(m)}))_m$.

THEOREM 3.1. *Let $f : \mathbb{R}^p \rightarrow \mathbb{R}$ be of class C^2 and convex. Suppose that there exist constants $m, \varrho, M, \varepsilon > 0$ such that*

$$(3.3) \quad m\|x - y\|^2 \leq \left(\nabla f(x) - \nabla f(y) - \frac{\rho}{2} C^T C(x - y), x - y \right) \quad \forall x, y \in \mathbb{R}^p,$$

$$(3.4) \quad \forall n \exists N \exists \gamma_n > 0 \forall m \geq N \quad \left| \det \left(\frac{r_n^m}{\|r_n^m\|}, \dots, \frac{r_n^{m+p-1}}{\|r_n^{m+p-1}\|} \right) \right| \geq \gamma_n,$$

$$(3.5) \quad \|\nabla^2 f(x) - \nabla^2 f(y)\| \leq M\|x - y\| \quad \forall (x, y) \in \mathbb{R}^p \times \mathbb{R}^p.$$

Then:

- 1) The sequence $(T(y_n^{(m)}))_m$ is defined for $m \geq N$ and $n \geq 0$.
- 2) For all $n \geq 0$ there exists a unique $x_{n+1} \in \mathbb{R}^p$ such that

$$\lim_{m \rightarrow \infty} \frac{\|T(y_n^{(m)}) - x_{n+1}\|}{\|y_n^{(m)} - x_{n+1}\|} = 0.$$

- 3) The sequence $(x_n)_{n \in \mathbb{N}}$ converges to the solution x^* of problem (3.1).

Proof. 1) Assume that the vector λ_n is given and consider the function $\Phi_n : \mathbb{R}^p \rightarrow \mathbb{R}$ defined by

$$\Phi_n(x) = f(x) + (\lambda_n, g(x)) \quad \forall x \in \mathbb{R}^p.$$

We have

$$(3.6) \quad \nabla \Phi_n(x) = \nabla f(x) + C^T \lambda_n,$$

$$(3.7) \quad \nabla^2 \Phi_n(x) = \nabla^2 f(x).$$

By the hypothesis (3.3), we have

$$m\|x - y\|^2 \leq (\nabla f(x) - \nabla f(y), x - y) - \frac{\rho}{2} \|C(x - y)\|^2 \quad \forall (x, y) \in \mathbb{R}^p \times \mathbb{R}^p,$$

and then, by (3.6), we deduce

$$m\|x - y\|^2 \leq (\nabla \Phi_n(x) - \nabla \Phi_n(y), x - y) \quad \forall (x, y) \in \mathbb{R}^p \times \mathbb{R}^p.$$

This relation implies (see [7]) the sufficient condition for the existence and uniqueness of the solution of the problem $\min_{x \in \mathbb{R}^p} \Phi_n(x)$, denoted by x_{n+1} , and for all $x, y \in \mathbb{R}^p$, we have

$$m\|y\|^2 \leq (\nabla^2 \Phi_n(x)y, y) \quad \forall (x, y) \in \mathbb{R}^p \times \mathbb{R}^p.$$

Therefore, using (3.4), (3.5) and (3.7), we deduce that $\nabla^2 \Phi_n(x_{n+1})$ is regular,

$$\|\nabla^2 \Phi_n(x) - \nabla^2 \Phi_n(y)\| \leq M\|x - y\| \quad \forall (x, y) \in \mathbb{R}^p \times \mathbb{R}^p,$$

and

$$\forall n \geq 0 \exists N > 0 \exists \gamma_n \forall m \geq N$$

$$\left| \det \left(\frac{\nabla \Phi_n(y_n^{(m)})}{\|\nabla \Phi_n(y_n^{(m)})\|}, \dots, \frac{\nabla \Phi_n(y_n^{(m+p-1)})}{\|\nabla \Phi_n(y_n^{(m+p-1)})\|} \right) \right| \geq \gamma_n.$$

Then, for all $n \geq 0$, the sequence $T(y_n^{(m)})$ is defined for $m \geq N$ and (see [13])

$$\lim_{m \rightarrow \infty} \frac{\|T(y_n^{(m)}) - x_{n+1}\|}{\|y_n^{(m)} - x_{n+1}\|} = 0.$$

Now we will prove the convergence of the sequence $(x_n)_n$ to the solution of problem (3.1).

The sufficient condition for the existence of the solution of (3.1) is given by (3.3) and the hypothesis $\{x \in \mathbb{R}^p : g(x) < 0\} \neq \emptyset$ (see [7]). It is well known that if x^* is a solution to (3.1) with corresponding multipliers $\lambda^* \in \mathbb{R}_+^p$, then x^* also satisfies

$$(3.8) \quad L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*) \quad \forall (x, \lambda) \in \mathbb{R}^p \times \mathbb{R}_+^q.$$

Using (3.3), we have

$$m\|x_{n+1} - x^*\|^2 \leq \left(\nabla f(x_{n+1}) - \nabla f(x^*) - \frac{\varrho}{2} C^T C(x_{n+1} - x^*), x_{n+1} - x^* \right).$$

By (3.8), we have $L(x^*, \lambda^*) = \min_{x \in \mathbb{R}^p} L(x, \lambda^*)$, thus

$$\nabla f(x^*) = -C^T \lambda^*.$$

On the other hand, $\Phi_n(x_{n+1}) = \min_{x \in \mathbb{R}^p} \Phi_n(x)$ implies

$$\nabla f(x_{n+1}) = -C^T \lambda_n.$$

Thus, we deduce

$$\begin{aligned} m\|x_{n+1} - x^*\|^2 &\leq \left(-C^T(\lambda_n - \lambda^*) - \frac{\varrho}{2} C^T C(x_{n+1} - x^*), x_{n+1} - x^* \right) \\ &\leq -\frac{1}{2\varrho} (2\varrho C^T(\lambda_n - \lambda^*) + \varrho^2 C^T C(x_{n+1} - x^*), x_{n+1} - x^*) \\ &\leq -\frac{1}{2\varrho} (\varrho^2 (C^T C(x_{n+1} - x^*), x_{n+1} - x^*) + 2\varrho (C^T(\lambda_n - \lambda^*), x_{n+1} - x^*)). \end{aligned}$$

Therefore

$$\begin{aligned} m\|x_{n+1} - x^*\|^2 &\leq -\frac{1}{2\varrho} (\|\lambda_n - \lambda^* + \varrho C(x_{n+1} - x^*)\|^2 - \|\lambda_n - \lambda^*\|^2) \\ &\leq -\frac{1}{2\varrho} (\|\lambda_n + \varrho(Cx_{n+1} - d) - \lambda^* - \varrho(Cx^* - d)\|^2 - \|\lambda_n - \lambda^*\|^2). \end{aligned}$$

Using (3.8), we have $L(x^*, \lambda^*) = \max_{\lambda \in \mathbb{R}_+^q} L(x^*, \lambda)$, thus

$$\lambda^* = P(\lambda^* + \varrho(Cx_{n+1} - d)),$$

and since $\lambda_{n+1} = P(\lambda_n + \varrho(Cx_{n+1} - d))$, we finally obtain

$$(3.9) \quad m \|x_{n+1} - x^*\|^2 \leq -\frac{1}{2\rho} (\|\lambda_{n+1} - \lambda^*\|^2 - \|\lambda_n - \lambda^*\|^2),$$

which proves that the sequence $(\|\lambda_n - \lambda^*\|)_n$ is decreasing, hence convergent, which gives $\lim_{k \rightarrow \infty} x_{n+1} = x^*$. ■

REMARKS 3.1. (1) The assumption (3.4) is the uniform invertibility hypothesis [13]. In [10, Prop. 4], a characterization of this hypothesis is given using the matrix $\nabla^2 f(x_{n+1})$.

(2) It suffices to take $\alpha \in]0, 1/M[$ in order to have $I - \alpha \nabla^2 f(x_{n+1})$ regular.

Now we propose another version of Algorithm 1 where the vector sequence $(y_n^{(m)})_m$ is computed by the gradient method with an optimal step.

ALGORITHM 2. We replace (b1) and (b2) by (b'1) and (b'2), where (b'1) and (b'2) are defined by:

- (b'1) Set $y_n^{(1)} = x_n$
 for $i = 1, \dots, p$, do
 compute $s_n^{(i)} = (\nabla f(y_n^{(i)}) + (J_g(y_n^{(i)}))^T \lambda_n)$
 compute $\alpha_n^{(i)}$ by $\min_{\alpha \in \mathbb{R}^+} L(y_n^i + \alpha s_n^{(i)}, \lambda_n)$
 set $r_n^i = \alpha_n^{(i)} s_n^{(i)}$, $y_n^{(i+1)} = y_n^{(i)} - r_n^{(i)}$
 end do
- (b'2) for $i = 1, \dots$, do
 compute $s_n^{(p+i)} = (\nabla f(y_n^{(p+i)}) + (J_g(y_n^{(p+i)}))^T \lambda_n)$
 compute $\alpha_n^{(p+i)}$ by $\min_{\alpha \in \mathbb{R}^+} L(y_n^{(p+i)} + \alpha s_n^{(p+i)}, \lambda_n)$
 set $r_n^{(p+i)} = \alpha_n^{(p+i)} s_n^{(p+i)}$, $y_n^{(p+i+1)} = y_n^{(p+i)} - r_n^{(p+i)}$
 solve the linear system $(\Delta R_{n,p})^T \Delta R_{n,p} \beta_{n,p} = (\Delta R_{n,k})^T r_n^{(i)}$
 compute $t_n^{(p)} = T(y_n^{(i)}) = y_n^{(i)} - \Delta_{n,p}^1 \beta_{n,p}$
 if $\|T(y_n^{(i)}) - T(y_{n-1}^{(i)})\| \leq \varepsilon$ then
 set $x_{n+1} = T(y_n^{(i)})$
 stop
 end if
 end do.

The following theorem shows that $T(y_n^{(m)})_m$ given by Algorithm 2 converges to x_{n+1} faster than $(y_n^{(m)})_m$. We denote by $G(u_1, \dots, u_p) = \det((u_i, u_j)_{1 \leq i, j \leq p})$ the Gram determinant corresponding to the p -tuple (u_1, \dots, u_p) . We obtain the following result.

THEOREM 3.2. *Let $f : \mathbb{R}^p \rightarrow \mathbb{R}$ be of class C^2 and suppose that there exist constants $m, \rho, \varepsilon > 0$ such that*

$$(3.10) \quad m\|x - y\|^2 \leq \left(\nabla f(x) - \nabla f(y) - \frac{\rho}{2} C^T C(x - y), x - y \right), \quad \forall x, y \in \mathbb{R}^p,$$

$$(3.11) \quad \forall n \exists N \exists \gamma_n > 0 \forall m \geq N \quad G\left(\frac{r_n^{(m)}}{\|r_n^{(m)}\|}, \dots, \frac{r_n^{(m+p-1)}}{\|r_n^{(m+p-1)}\|} \right) \geq \gamma_n.$$

Then:

- 1) The sequence $(T(y_n^{(m)}))_m$ is defined for $m \geq N$ and $n \geq 0$.
- 2) For all $n \geq 0$ there exists a unique $x_{n+1} \in \mathbb{R}^p$ such that

$$\lim_{m \rightarrow \infty} \frac{\|T(y_n^{(m)}) - x_{n+1}\|}{\|y_n^{(m)} - x_{n+1}\|} = 0.$$

- 3) The sequence $(x_n)_{n \in \mathbb{R}}$ converges to the solution x^* of problem (3.1).

Proof. As in Theorem 3.1, the problem $\min_{x \in \mathbb{R}^p} \Phi_n(x)$ has a unique solution denoted by x_{n+1} . Using (3.10), (3.11), (3.6), and (3.7), we have

$$\begin{aligned} m\|y\|^2 &\leq (\nabla^2 \Phi_n(x)y, y) \quad \forall (x, y) \in \mathbb{R}^p \times \mathbb{R}^p, \\ m\|x - y\|^2 &\leq (\nabla \Phi_n(x) - \nabla \Phi_n(y), x - y) \quad \forall (x, y) \in \mathbb{R}^p \times \mathbb{R}^p, \\ \forall n \geq 0 \exists N > 0 \exists \gamma_n > 0 \forall m \geq N &G\left(\frac{\nabla \Phi_n(y_n^m)}{\|\nabla \Phi_n(y_n^m)\|}, \dots, \frac{\nabla \Phi_n(y_n^{m+p-1})}{\|\nabla \Phi_n(y_n^{m+p-1})\|} \right) \geq \gamma_n. \end{aligned}$$

Then for all $n \geq 0$, the sequence $T(y_n^m)$ is defined for $m \geq N$ and we have (see [12])

$$\lim_{m \rightarrow \infty} \frac{\|T(y_n^m) - x_{n+1}\|}{\|y_n^m - x_{n+1}\|} = 0.$$

The end of the proof is the same as in Theorem 3.1. ■

4. Numerical experiments. In this section, we report on some numerical experiments. We compare the nonlinear multiple hybrid procedures NLM given by Algorithms 1 and 2 with the gradient method GM and the gradient method with optimal step GMO [2, 7]. This comparison will be summarized in tables which give the number of iterations, iter., and the associated residual norms for each method. The stopping criterion is given by $\text{res.} = \|x_k - x^*\|$, where x^* is the solution of problem (3.1). To solve the linear system in Algorithms 1 and 2, we use Gaussian elimination.

EXAMPLE 1. The first example has been used in [15]. We consider the following problem:

$$\begin{cases} \text{Minimize } \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 - 2x_1x_2 - x_1 - 2x_2 \\ \text{subject to } \begin{cases} x_1 + x_2 \leq 1, \\ x_1 \geq 0 \text{ and } x_2 \geq 0. \end{cases} \end{cases}$$

For this function, we can see that $x^* = (1/3, 2/3)$. For different initial points x_0 we obtain the following tables:

Table 1. $x_0 = (0.5, 1)$

iter	GM	NLMH
0	1.414213562373089E-002	1.414213562373089E-002
1	7.071067811865560E-003	1.138360651264128E-016
4	8.838834764831656E-004	
10	1.381067932011570E-005	
20	1.348699145788843E-008	
25	4.214691149978832E-010	
30	1.430201717150080E-011	

Table 2. $x_0 = (0.6, 0.9)$

iter	GM	NLMH
0	1.414213562373018E-003	1.414213563273018E-003
1	5.656854249492542E-004	1.159106867033638E-015
5	1.448154687869300E-005	
10	1.482910400262855E-007	
15	1.518500241054302E-009	
20	1.554941226978357E-011	
25	1.602622606765867E-013	
30	4.399650097552470E-014	

EXAMPLE 2. This example is taken from [15]. We consider the following problem:

$$\begin{cases} \text{Minimize } x_1^2 + x_2^2 + x_3^2 + x_4^2 - 2x_1 - x_2 - 3x_4 \\ \text{subject to } \begin{cases} 2x_1 + x_2 + x_3 + 4x_4 \leq 7, \\ x_1 + x_2 + 2x_3 + x_4 \leq 6, \\ x_i \geq 0 \text{ pour } i = 1, 2, 3, 4. \end{cases} \end{cases}$$

For $x_0 = (2, 2, 1, 0)$ we obtain the results shown in Fig. 1.

EXAMPLE 3. This example was taken from [8]. We consider the problem

$$\begin{cases} \text{Minimize } \sum_{1 \leq i \leq n} f_i^2(x) \\ \text{subject to } Cx \leq d \end{cases}$$

where

- n is any positive multiple of 2, $x = (x_i)_{1 \leq i \leq n}$,
- for $i = 1, \dots, n/2$, $f_{2i-1}(x) = 10(x_{2i} - x_{2i-1}^2)$, $f_{2i}(x) = 1 - x_{2i-1}$,

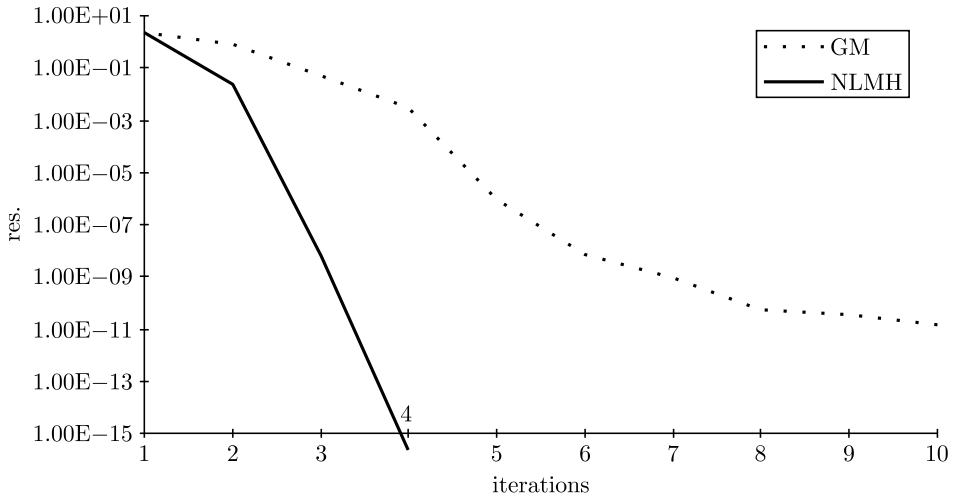


Fig. 1

- C is the $n \times n$ matrix given by

$$C = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix},$$

- $d = (n, n - 1, n - 2, \dots, 2, 1)$.

For this problem, we used two values of n :

- $n = 10$, $x_0 = (1, 1, 2, 3, -1, 1, 0, 2, -3, 1)$ (Table 3),
- $n = 20$, $x_0 = (1, 1, 2, 3, -1, 1, 0, 2, -3, 1, 1, 1, 2, 3, -1, 1, 0, 2, -3, 1)$ (Table 4).

Table 3

iter	GM	NLMH
0	5.196152423125231E-000	5.196152423125231E-000
1	1.848869271451412E-000	4.551452613214589E-001
10	4.874501256871418E-001	8.691254803654878E-002
20	4.701452178321422E-001	2.574512014583601E-002
30	4.110214567852103E-001	1.461420568972214E-003
40	3.661204568912511E-001	4.401257896542451E-004
45	3.601025321642539E-001	3.261250458792152E-010

Table 3 (cont.)

iter	GM	NLMH
50	3.591204520121544E-001	1.182507890125461E-016
90	1.123654825613125E-001	
100	1.084536218962116E-001	
190	6.801245612561215E-002	
200	7.151452105682512E-003	
210	8.831258976521252E-004	

Table 4

iter	GM	NLMH
0	7.348469228123105E-000	7.348469228123105E-000
1	1.831287965412451E-000	3.920124563214528E-001
10	2.172014508963212E-001	3.041789156325469E-002
20	2.084567825946129E-001	2.421576824563154E-002
30	2.045862587165234E-001	2.075862143654152E-002
40	1.764589253156015E-001	4.452136589456201E-003
50	1.701452879653223E-001	1.391254368245631E-004
60	1.665239015846152E-001	1.282458964321145E-004
70	1.614528635248963E-001	1.204568124351238E-004
80	1.541502589673541E-001	1.091254362154786E-004
83	1.531452689354611E-001	9.554218961254518E-005
85	1.530125647856918E-001	7.661254897314051E-007
86	1.521456924056314E-001	1.102578965314256E-007
87	1.520456243654213E-001	6.891452865432897E-011
88	1.520412365478127E-001	4.401789546243154E-013
89	1.520411356245634E-001	2.184579652301457E-015
100	1.445628397025648E-002	
150	7.431546289378564E-002	
200	7.082154839147325E-003	

References

- [1] A. Abkowitz and C. Brezinski, *Acceleration properties of the hybrid procedure for solving linear systems*, Appl. Math. (Warsaw) 23 (1996), 417–432.
- [2] A. Auslender, *Optimisation: Méthodes Numériques*, Masson, Paris, 1976.

- [3] C. Brezinski, *Accélération de la Convergence en Analyse Numérique*, Springer, Berlin, 1977.
- [4] —, *Algorithmes d'Accélération de la Convergence. Etude Numérique*, Éditions Technip, Paris, 1978.
- [5] C. Brezinski and J.-P. Chehab, *Nonlinear hybrid procedures and fixed points iterations*, Numer. Funct. Anal. Optim. 19 (1998), 465–487.
- [6] C. Brezinski and M. Redivo-Zaglia, *Hybrid procedures for solving linear systems*, Numer. Math. 67 (1994), 1–19.
- [7] P. G. Ciarlet, *Introduction à l'Analyse Numérique Matricielle et à l'Optimisation*, Masson, Paris, 1985.
- [8] J. E. Dennis, Jr. and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983.
- [9] A. Kolmogorov and S. Fomine, *Éléments de la Théorie des Fonctions et de l'Analyse Fonctionnelle*, Éditions Mir, Moscou, 1979.
- [10] H. Le Ferrand, *Recherche d'extrema par les méthodes d'extrapolation*, C. R. Acad. Sci. Paris Sér. I 318 (1994), 1043–1049.
- [11] B. Rhanizar, *Hybrid procedures for solving some unconstrained nonlinear optimization problems*, Appl. Numer. Math. 30 (1999) 459–474.
- [12] —, *On Henrici's transformation in optimization*, Appl. Math. (Warsaw) 27 (2000), 127–141.
- [13] H. Sadok, *Accélération de la Convergence de Suites Vectorielles et Méthodes de Point Fixe*, thèse, Univ. des Sciences et Techniques de Lille, 1988.
- [14] J. Vignes, *Algorithmes Numériques, Analyse et Mise en Œuvre*, t.2, Editions Technip, Paris, 1980.
- [15] H. Moulin et F. Fogelman-Souli, *Convexité dans les Mathématiques de la Décision*, Hermann, Paris, 1979.

Département de Mathématiques
 École Normale Supérieure Takaddoum
 B.P. 5118
 Rabat, Maroc
 E-mail: rhanizar@caramail.com

Département de Mathématiques
 Faculté des Sciences Semlalia
 Université Cadi Ayyad
 Marrakech, Maroc

Received on 10.12.2001

(1608)