THE SECOND ORDER STEEPEST DESCENT METHOD

Aharon Ben-Tal*

Technion, Israel Institute of Technology and
Department of Industrial and Operations
Engineering, The University of Michigan

Luc De Hulsters

Technion, Israel Institute of Technology

Technical Report #86-33

September 1986

* Research supported by NSF and the VPR Fund at the Technion.

ABSTRACT

The paper introduces a new algorithm for unconstrained minimization of an n-variable twice continuously differentiable function f. Unlike classical methods, which improve a current solution by moving along a straight line, the new method improves the solution by moving along a quadratic curve in R^n. The specific curve is determined by minimizing an appropriate approximate model of f. The algorithm thus obtained (SOSD) is a natural second order extension of the Steepest Descent method. It possesses a global convergence property combined with a quadratic rate of convergence, while using the same information and employing the same computational effort, as the Newton method. Versions of SOSD with inexact linesearch and with no linesearches at all are studied as well, retaining the above desirable convergence properties, which are also demonstrated computationally.

Introduction

In this paper we introduce a new computational method for solving the unconstrained minimization problem

$$\text{(A)} \qquad \min\{f(x) : x \in R^n\}$$

for functions f which are twice continuously differentiable ($f \in C^2$).

The classical methods for solving (A) are all based on the basic iteration step

$$x \leftarrow x + td \tag{1}$$

where $d \in R^n$ is a direction vector and $t > 0$ is the stepsize. The direction vector is obtained by minimizing an approximate model of $f(x+td)$. Thus if a first order Taylor approximation is used:

$$f(x+td) \simeq f(x) + td^T \nabla f(x)$$

one obtains the Steepest Descent (SD) direction:

$$d = -g / \|g\|, \ \big(g = \nabla f(x)\big).$$

If a second order Taylor expansion is used:

$$f(x+td) \simeq f(x) + td^T \nabla f(x) + 1/2t^2 d^T \nabla^2 f(x) d$$

then the Newton direction is obtained

$$d_N = -H^{-1} g \quad \big(H = \nabla^2 f(x)\big)$$

or (if a bound is imposed on the length of d) a direction as in a Trust Region method:

$$d = -(H + \lambda I)^{-1} g$$

emerges (see e.g. [5], [6]).

In our method the basic iteration step is

$$x \leftarrow x + td + 1/2t^2 z \tag{2}$$

with two "direction vectors" $d \in R^n$, $z \in R^n$ and a stepsize $t > 0$. Thus the improving step is along a quadratic curve in R^n, rather than along a straight line.

The idea of an improving step along a quadratic curve was used theoretically by the first author (see e.g. [2], [1], [3] and [4]) to obtain second order necessary optimality conditions for smooth and nonsmooth optimization problems. Here we study the computational implications of the basic iteration (2).

The direction vectors d and z are obtained by minimizing an approximate model of $f(x+td+1/2t^2z)$ which is in fact a second order Taylor expansion of the latter (as a function of t):

$$f(x+td+1/2t^2z) \simeq f(x)+td^T\nabla f(x)+1/2t^2[z^T\nabla f(x)+d^T\nabla^2 f(x)d] \tag{3}$$

As in the derivation of the steepest descent or the trust region directions, certain constraints are imposed on d, z. More specifically we consider the following model (recall $g=\nabla f(x)$, $H=\nabla^2 f(x)$)

(M) $\quad \min \Delta f(t,d,z):=tg^Td+1/2t^2[z^Tg+d^THd]$

subject to

$$\|z\| \le \alpha$$
$$d^Tg/\|g\| \le -\beta$$

The first constraint is just normalizing the length of z. The second constraint enforces the direction d to make a sharp angle with the steepest descent direction; this is a typical requirement in many convergent descent methods. The positive scalar β controls this angle and the positive scalar α controls the length of z. The objective function Δf in (M) is (omitting the constant f(x)) the righthand side of (3). Although we minimize Δf also with respect to t, this is only for deriving the optimal solutions $\bar{d}$ and $\bar{z}$ of (M). In the algorithm t will be chosen by minimizing the true objective $f(x+t\bar{d}+1/2t^2\bar{z})$ and not its approximation $\Delta f(t,\bar{d},\bar{z})$.

The directions $\bar{d}$ and $\bar{z}$ are obtained in Section 1, they are

$$\bar{z} = -\alpha g/\|g\|$$

$$\bar{d} = -\beta \frac{\|g\|}{g^T H^{-1} g}$$

Thus, $\bar{z}$ is the steepest descent direction and $\bar{d}$ is a "signed Newton direction" i.e. it is the Newton direction or its opposite according to the sign of $g^T H^{-1} g$ being positive or negative. At any rate (in sharp contrast to the newton direction d_N) $\bar{d}$ is a direction of descent, whenever $g \neq 0$, H^{-1} exists and $g^T H^{-1} g \neq 0$ (we call this the nonsingular case).

Based on the directions $\bar{d}$, $\bar{z}$, the method studied in this paper iterates from a current iteration point x_k to the next point via

$$\text{(SOSD)} \quad \begin{cases} x_{k+1} = x_k + t_k d_k + 1/2 t_k^2 z_k \\ t_k = \text{a suitable stepsize} \\ d_k = -\beta_k \dfrac{\|g_k\|}{g_k^T H_k^{-1} g_k} H_k^{-1} g_k; \quad z_k = -\alpha_k / \|g_k\| \\ \text{where } g_k = \nabla f(x_k), \;\; H_k = \nabla^2 f(x_k). \end{cases}$$

$\{\alpha_k\}$ and $\{\beta_k\}$ are sequences of positive scalars, bounded away from zero which may be fixed, predetermined or chosen iteratively. The name SOSD is an abbreviation for Second Order Steepest Descent. Indeed the method (SOSD) is a natural extension of the steepest descent method (SD). This is demonstrated in Table 1 below.

The quadratic rate of convergence of the SOSD method mentioned in Table 1 will be proved in Section 4 for a general C^2-function, following a proof for the special case where f is a convex quadratic function

$$f(x) = 1/2 x^T Q x - b^T x, \text{ (Q symmetric positive definite)}$$

These results are preceeded by Section 2, in which the global convergence of the SOSD

Table 1: Comparison of SD and SOSD methods

	SD	SOSD
data used	first order (gradients)	second order (gradients and) Hessians)
improving step	along 1st degree polynomial	along 2nd degree polynomial
role of convergence	linear	quadratic

method is demonstrated.

The above mentioned results are for exact line search:

$$t_k = \arg\min_{t>0} f(x_k + td_k + 1/2t^2 z_k)$$

Similar global convergence and rate of convergence results are obtained in Section 5 for versions of SOSD where the stepsize is chosen according to an Armijo-Goldstein type rule. In the quadratic case we avoid the need to compute the stepsize by fixing the stepsize t_k, and then choosing the a_k so as to make t_k an exact linesearch step. This version, which is also extended to the general C^2-case, is called "the a-method". It also possesses a quadratic rate of convergence, although not necessarily gobal convergence. Like the pure Newton method it does not require stepsize computations. For the single variable case it reduces exactly to the Newton's method, but for $n > 1$ it is quite different, and in our computational tests exhibited global convergence in all the examples for which Newton's method failed.

At this point we offer an intuitive explanation as to why one should expect from a method such as SOSD to be globally convergent and at a quadratic rate. Note that in the iterative step

$$x_{k+1} = x_k + t_k d_k + 1/2 t_k^2 z_k$$

when x_k is still far from optimality, and a large stepsize t_k is expected, then z_k (a SD direction) will be dominant (forcing global convergence). While close to optimality, where $H(x_k)$ is expected to be positive definite, and t_k close to zero, the vector d_k (a Newton direction) is the dominant one (forcing quadratic rate of convergence).

We would like to emphasize that the SOSD method uses the same data as the Newton method, and requires the same computational effort: evaluating the gradient and Hessian, and solving a linear system to obtain the direction d_k, plus a linesearch computation. But SOSD is globally and quadratically convergent. These properties are shared by Trust Region methods which we believe are more complicated and less natural; while they try to "correct" the Newton method, we try to extend (to second order elements) the Steepest Descent method. Thus we put forward the idea that the SOSD method, and not Newton's method, should serve as a fundamental algorithm from which first order modifications (such as quasi-Newton methods) should be obtained, and extensions to constrained problems be made.

The limited computational experience, reported in Section 6, indeed supports the fact that the SOSD method performs equally well to Newton's, when the latter converges, and succeeds to solve test problems from all starting points from which the latter failed to converge, and still exhibits quadratic rate of convergence.

§1. DERIVATION OF THE DIRECTION VECTORS $\bar{d}$ and $\bar{z}$

We solve the model problem (M) first under the nonsingularity assumption

$$g \neq 0, \quad H \text{ nonsingular}, \quad g^T H^{-1} g \neq 0 \tag{4}$$

It is easy to see that for fixed but arbitrary $d \in R^n$ and $t > 0$, the optimal z is

$$\bar{z} = -\frac{\alpha}{\|g\|} g$$

and we are left with the problem

$$\min_{0<t\leq T} \min_{d} \; t\,\Delta f(t,d,\bar{z}) = tg^T d + 1/2t^2[d^T Hd - \alpha\|g\|]$$

(P) s.t.

$$g^T d + \beta\|g\| \leq 0 \tag{5}$$

The upper bound T imposed here, will be dropped later. Let λ be the multiplier corresponding to the constraint (5). The Kuhn-Tucker conditions for the optimality of d, t are

(i) $g + tHd + \lambda g = 0$

(ii) $g^T d + t(d^T Hd - \alpha\|g\|) \leq 0$

(iii) $g^T d \leq -\beta\|g\|$

(iv) $0 < t \leq T$

(v) $t < T \Rightarrow$ equality holds in (ii)

(vi) $\lambda[g^T d + \beta\|g\|] = 0$

(vii) $\lambda \geq 0$

We consider first the case

$$g^T H^{-1} g > 0 \tag{6}$$

and define

$$h =: \frac{g^T H^{-1} g}{\|g\|}$$

If

$$h/\beta < T \tag{7}$$

then, the solution of (i)−(vii) is

$$\bar{d} = -\frac{\beta}{h} H^{-1} g$$

$$\bar{t} = \begin{cases} \dfrac{\beta h}{\beta^2 - \alpha h} & \text{if} \quad 0 < \dfrac{\beta h}{\beta^2 - \alpha h} < T \\ T & \text{otherwise} \end{cases}$$

$$\bar{\lambda} = \frac{\beta \bar{t}}{h} - 1$$

If

$$h/\beta \geq T \tag{8}$$

then the solution is

$$\bar{d} = -\frac{1}{T} H^{-1} g$$
$$\bar{t} = T$$
$$\bar{\lambda} = 0$$

We now drop the bound on the stepsize by letting $T \to \infty$. Then, only case (7) is relevant, giving the direction

$$\bar{d} = -\beta \frac{\|g\|}{g^T H^{-1} g} H^{-1} g \tag{9}$$

Next consider the case

$$g^T H^{-1} g < 0 \tag{10}$$

Multiplying (i) by the row vector $g^T H^{-1}$ we get $g^T H^{-1} g + t g^T d + \lambda g^T H^{-1} g = 0$ implying $t g^T d = -(1+\lambda) g^T H^{-1} g > 0$ by (10) contradicting $t > 0$ and (iii). Thus no solution exists to the Kuhn-Tucker conditions under (10), indicating an unbounded solution of (P). Indeed set

$$d_M = -\frac{M}{h}H^{-1}g,$$

the value of the objective junction in (P) is

$$\Delta f(t,d_M,\bar{z}) = -t\|g\|M + 1/2t^2\frac{\|g\|^2}{g^TH^{-1}g}M^2 - 1/2\,t^2\alpha\|g\|.$$

For all $t > 0$, the coefficients of both M and M^2 are negative, and therefore Δf can be made to approach $-\infty$ by letting M approach $+\infty$. Therefore $\{ d_M, M \to \infty \}$ is an infimizing sequence. For any $M > 0$, d_M is the same direction as $\bar{d}$ in (9). If a common bound on the length of both $\bar{d}$ and d_M is imposed then actually $d_M = \bar{d}$.

We conclude that if $g^TH^{-1}g$ is either positive or negative, the directions obtained from the model problem (M) are

$$\bar{z} = -\alpha\frac{g}{\|g\|} \qquad \bar{d} = -\beta\frac{\|g\|}{g^TH^{-1}g}H^{-1}g$$

An important property of the directions $\bar{d}$, $\bar{z}$ is that

$$\bar{z}^T\bar{d} = \alpha\beta > 0 \tag{11}$$

hence the angle between them is greater than 90°. This ensures that a movement from x along the curve $x(t) = x + t\bar{d} + 1/2t^2\bar{z}$, will not bring us back to x.

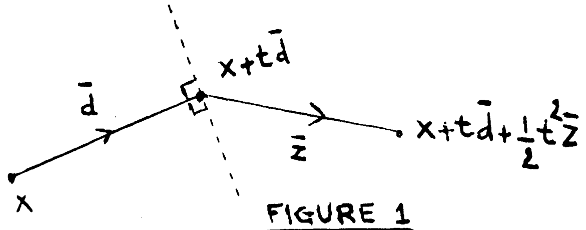

FIGURE 1

Moreover (11) implies that $\|x(t)-x\|^2$ is monotonely increasing for $t > 0$, an essential property for many convergent algorithms, see also the convergence proof in the next section.

The length of $\bar{d}$ is bounded below by β, indeed

$$\|\bar{d}\|^2 = \beta^2 \frac{\|g\|^2\|H^{-1}g\|^2}{\|g^TH^{-1}g\|^2} \geq \beta^2$$

by the Cauchy-Schwartz inequality.

If H is positive definite, with eigenvalues $0<\lambda_1\leq\lambda_2\leq\ldots\leq\lambda_n$, an upper bound can be obtained as follows: let $H^{1/2}$ denote the root matrix of H and $H^{-1/2}$ its inverse, set $x = H^{-1/2}g$, then

$$\|\bar{d}\|^2 = \beta^2 \frac{(x^THx)(x^TH^{-1}x)}{(x^Tx)^2} \leq \beta^2 \frac{(\lambda_1+\lambda_n)^2}{4\lambda_1\lambda_n}$$

by the Kantorovich inequality (see e.g. [7]). So, for H positive definite

$$\beta \leq \|\bar{d}\| \leq \beta \frac{(\lambda_1+\lambda_n)/2}{\sqrt{\lambda_1\lambda_n}} \tag{12}$$

and we see that β essentially controls the length of $\bar{d}$.

We close this section with a few remarks concerning the singular cases $g=0$ or $g^TH^{-1}g=0$. If $g=0$ (but H is not positive semi definite, for otherwise we are at a point satisfying second order necessary conditions) then the model problem (M) reduces to minimizing d^THd. Adding a bound on $\|d\|$ the solution of

$$\min\{d^THd:\ \|d\| \leq D\}$$

is a direction of negative curvature. In fact the optimal d is

$$\hat{d} = \frac{D}{\|v_1\|} v_1$$

where v_1 is the eigenvector corresponding to the minimal (negative) eigenvalue of H. So, for $g=0$, the basic iteration (2) becomes

$$x \leftarrow x + t\hat{d}.$$

Consider now the case $g \neq 0$ but $g^T H^{-1} g = 0$: the Kuhn-Tucker conditions for problem (P) are inconsistent for any $\beta > 0$. Thus for this case an appropriate model is

$$\min_d \ \{\Delta f(t, d, \bar{z}):\ g^T d \leq 0\}$$

the optimal solution of which is

$$d = -\lambda H^{-1} g, \quad \lambda \geq 0 \ \text{arbitrary}.$$

We might as well take $\lambda = 0$, in which case $d=0$ and the iteration step (2) becomes just the SD step. Indeed if SD steps are taken in the algorithm SOSD whenever $g^T H^{-1} g = 0$ (or H is singular) then this will not affect either the global convergence, or the rate of convergence for problems with local minima satisfying second order sufficient conditions. Henceforth in the paper we consider only the nonsingular cases.

§2 GLOBAL CONVERGENCE OF THE SOSD METHOD (EXACT LINESEARCH)

In this section we study the convergence properties of the SOSD method with exact linesearch, i.e.

$$t_k = \arg\min_{t>0} f(x_k + td_k + 1/2t^2 z_k)$$

as applied to the minization of a C^2-function $f : R^n \to R$. This version will be refered to as the SOSD-E method.

We will be using the framework of the Global Convergence Theorem as developed by Zangwill [10] and described in [7, Section 6.5]. The SOSD-E method generates a sequence $\{x_k\}_0^\infty$ by the rule:

$$x_{k+1} \in A(x_k) \qquad (x_0 \text{ given initially})$$

where A is the composite point to set mapping from R^n into 2^{R^n}:

$$A = SG.$$

G is the function maping R^n into $R^n \times R^n \times R^n$:

$$G(x) = (x, \bar{z}(x), \bar{d}(x)), \text{ where}$$

$$\bar{z}(x) =: -\alpha \frac{g(x)}{\|g(x)\|}; \quad \bar{d}(x) =: -\beta \frac{\|g(x)\|}{g(x)^T H(x)^{-1} g(x)} H(x)^{-1} g(x),$$

and S is the point to set mapping $S : R^n \times R^n \times R^n \to 2^{R^n}$:

$$S(x,z,d) = \{y = x + \bar{t}d + 1/2\bar{t}^2 z : \quad \bar{t} = \arg\min_{t\geq 0} f(x + td + 1/2t^2 z)\}$$

Let Γ denote the solution set

$$\Gamma = \{x \in R^n : \; g(x) = 0\}.$$

For the function G to be well defined (ouside Γ) we assume henceforth the nonsingularity condition :

$$H^{-1}(x) \text{ exists and } g(x) \neq 0 \Rightarrow g(x)^T H(x)^{-1} g(x) \neq 0, \quad (2.1)$$

The implication in (2.1) holds if e.g. $|y^T H^{-1} y| > M\|y\|^p$ for some $M > 0$ and $p > 0$. The main result of this section follows.

Theorem 1 [Global Convergence of the SOSD-E Method]

Let $f : R^n \to R$ be a C^2-function, and $x_0 \in R^n$ a given starting point for the sequence $\{x_k\}_0^\infty$ generated by the SOSD-E algorithm. Assume that the nonsingularity condition (2.1) holds and that the level set

$$L_0 = \{x : f(x) \leq f(x_0)\}$$

is compact. Then, any limit point x^* of a convergent subsequence of $\{x_k\}_0^\infty$ is a solution, i.e. $g(x^*) = 0$.

Proof We refer to the Global Convergence Theorem (GCT) [7, p. 124]; accordingly, three conditions must hold

(i) $x_k \in$ some compact set S, $\forall\, k = 0,1,\ldots$

(ii) $\exists$ a descent function $Z : R^n \to R$ (see [7, p. 122])

(iii) The mapping A is closed outside Γ

Since both d(x) and z(x) are descent directions it follows that $f(x_{k+1}) \leq f(x_k)$ with strict inequality for $x_k \notin \Gamma$. Thus (i) holds with $S = L_0$ and (ii) holds with $Z(x) = f(x)$. To prove the validity of (iii), i.e. the closedness of A = SG, it suffices to show that G is continuous and S is closed on the range R(G) of G. The continuity of G follows from $f \in C^2$ and (2.1). To prove the closedness of S or R(G) note first that for any triple $(x,z,d) \in R(G)$: $z^T d > 0$ (see Section 2). Fix a triple $(x,d,z) \in R(G)$ and let $(x_k,z_k,d_k) \in R(G)$, with $x_k \to x$, $z_k \to z$, $d_k \to d$, and let $y_k \in S(x_k,z_k,d_k)$, $y_k \to y$; we need to verify $y \in S(x,z,d)$.

Now,

$$y_k = x_k + t_k d_k + 1/2 t_k^2 z_k; \quad t_k = \arg\min_{t \geq 0} f(x_k + t d_k + 1/2 t^2 z_k) \tag{2.3}$$

hence

$$\|y_k - x_k\|^2 = \|t_k d_k + 1/2 t_k^2 z_k\|^2 \quad \text{or}$$

$$t_k^2 \|d_k\|^2 + t_k^3 z_k^T d_k + 1/4 t_k^4 \|z_k\|^2 - \|y_k - x_k\|^2 = 0 \tag{2.4}$$

The function in the lefthand side of (2.4) is a polynomial in t_k which (for large k) has positive coefficients (here we use $z_k^T d_k > 0$), thus it is strictly monotonely increasing for $t_k \geq 0$. Since it has a nonpositive value at $t = 0$, there is a unique nonnegative $t_k = \Phi(x_k, y_k, z_k, d_k)$ solving (2.4) and Φ is continuous. So, when $k \to \infty$,

$$t_k \to \Phi(x,y,z,d) =: \bar{t} \geq 0.$$

Now,

$$y = \lim y_k = x + \bar{t} d + 1/2 \bar{t}^2 z.$$

For all $t \geq 0$: $f(y_k) \leq f(x_k + t d_k + 1/2 t^2 z_k)$, $\forall$ k by (2.3).

Letting $k \to \infty$

$$f(y) \leq f(x + td + 1/2 t^2 z) \qquad \forall\, t \geq 0$$

hence

$$f(y) \leq \min_{t \geq 0} f(x + td + 1/2 t^2 z) \leq f(x + \bar{t} d + 1/2 \bar{t}^2 z) = f(y)$$

showing that $y \in S(x,z,d)$.

∎

Remark 1 As explained at the end of Section 1, we can take a steepest descent step at any point x violating the nonsingularity condition. The modified algorithm will still converge globally by the Spacer Step Theorem (see Section 7.9 in [7]).

§ 3. THE SOSD-E ALGORITHM FOR A QUADRATIC FUNCTION

In this section we shall prove the quadratic rate of convergence of the SOSD-E algorithm for quadratic functions. In addition, we will propose a version of the method not requiring a linesearch which will be called "the α-method".

Consider the quadratic problem

$$\min_{x \in R^n} 1/2 x^T Q x - b^T x$$

where Q is a positive definite (P.D.) symmetric n x n matrix, and $b \in R^n$. We denote the solution of the problem by x^*, i.e. $Qx^* = b$; it is clear then that the above problem is equivalent to:

$$\text{(QP)} \qquad \min_{x \in R^n} f(x) = 1/2(x-x^*)^T Q(x-x^*) \quad .$$

and it is this problem which we shall consider in the remainder of this section.

Let the eigenvalues $\{\lambda_i\}$ of Q be ordered as:

$$0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n < \infty,$$

and define $r = \lambda_1/\lambda_n$ (the condition number of Q). Defining $y = x - x^*$, we have for the gradient of f : $g = Qy$ and for the Hessian $H = Q$.

We recall that for any $v \in R^n$ and $p \in R$:

$$0 < l \leq \frac{v^T Q^P v}{\|v\|^2} \leq u < \infty \qquad (3.1)$$

for some real positive l and u, a fact we shall use frequently.

Finally we define:

$$\underline{\alpha} = \liminf_{k\to\infty} \alpha_k$$

$$\bar{\alpha} = \limsup_{k\to\infty} \alpha_k$$

and analogously for β_k.

We begin by deriving expressions for the exact stepsize t_k (see eqs. (3.2) or (3.3) below) and for the quantity:

$$\|x_{k+1} - x^*\|^2 / \|x_k - x^*\|^4.$$

A necessary condition for t to be the optimal stepsize is

$$0 = \frac{\partial}{\partial t} f(x_k + td_k + 1/2t^2 t_k)|_{t=t_k}$$

$$= (d_k + t_k z_k)^T \nabla f(x_k + t_k d_k + 1/2t_k^2 z_k)$$

for our quadratic function, this gives:

$$(d_k + t_k z_k)^T Q(x_k + t_k d_k + 1/2t_k^2 z_k - x^*) = 0$$

or

$$\frac{z_k^T Q z_k}{2} t_k^3 + 3/2 d_k^T Q z_k t_k^2 + (z_k^T Q(x_k - x^*) + d_k^T Q d_k) t_k + d_k^T Q(x_k - x^*) = 0$$

After substituting the expressions for d_k, z_k and using $Qy = g$, we obtain (omitting the index k):

$$\alpha^2 \frac{g^T Q g}{2\|g\|^2} t^3 + 3/2\alpha\beta \frac{\|g\|^2}{g^T Q^{-1} g} t^2 + (\beta^2 \frac{\|g\|^2}{g^T Q^{-1} g} - \alpha\|g\|)\, t - \beta\|g\| = 0 \qquad (3.2)$$

Defining: $w_k = \dfrac{\|g_k\|^2}{g_k^T Q^{-1} g_k}$ and $u_k = \dfrac{g_k^T Q g_k}{2\|g_k\|^2}$, note that by (3.1) the w_k's and u_k's are finite, positive and bounded away from zero. Eq. (3.2) can now be written as:

$$1/2\alpha_k^2 u_k t_k^3 + 3/2\alpha_k \beta_k w_k t_k^2 + (\beta_k^2 w_k - \alpha_k\|g_k\|)\, t_k - \beta_k\|g_k\| = 0 \qquad (3.3)$$

For the rate of convergence we need the following expression:

$$\frac{\|x_{k+1}-x^*\|^2}{\|x_k-x^*\|^4} = \frac{\|x_k+t_kd_k+1/2t_k^2z_k-x^*\|^2}{\|y_k\|^4}$$

After substituting the values for d_k, t_k and omitting the index k, this becomes:

$$\frac{\alpha^2t^4}{4\|y\|^4} + \frac{\alpha\beta t^3}{\|y\|^4} + \left(\frac{\beta^2\|g\|^2}{(g^TQ^{-1}g)^2\|y\|^2} - \frac{\alpha g^TQ^{-1}g}{\|g\|\cdot\|y\|^4}\right)t^2 - \frac{2\beta\|g\|}{\|y\|^2g^TQ^{-1}g}t + \frac{1}{\|y\|^2} \qquad (3.4)$$

This can be further written as:

$$\frac{\alpha^2t^4}{4\|y\|^4} + \frac{\alpha(g^TQ^{-1}g)t^2}{\|g\|\cdot\|y\|^3}\left(\frac{1}{\|y\|}\left(\frac{\beta\|g\|t}{g^TQ^{-1}g} - 1\right)\right) + \left(\frac{1}{\|y\|}\left(\frac{\beta\|g\|t}{g^TQ^{-1}g} - 1\right)\right)^2 \qquad (3.5)$$

we are now ready to state the main theorem of this section:

<u>Theorem 2.</u> [Quadratic Convergence of SOSD-E for Quadratic Functions]

Let $x_k \to x^*$, then

$$\limsup_{k\to\infty} \frac{\|x_{k+1}-x^*\|}{\|x_k-x^*\|^2} \le 1/4\,\frac{\bar{\alpha}}{\underline{\beta}^2}$$

Moreover, if $\frac{2\sqrt{r}}{1+r} \ge \frac{\sqrt{2}}{2}$ we have the sharper bound:

$$\limsup_{k\to\infty} \frac{\|x_{k+1}-x^*\|}{\|x_k-x^*\|^2} \le 1/2\,\frac{\bar{\alpha}}{\underline{\beta}^2}\left(\frac{2\sqrt{r}}{1+2}\right)\left(\frac{1-r}{1+r}\right)$$

The proof of the Theorem is preceded by Lemma 1 and a few calculations.

<u>Lemma 1</u>

<u>If</u> the sequence $\{x_k\}$ converges to x^*

<u>Then</u> there exists $k_0 > 0$ such that for $k > k_0$:

$$t_k \leq \frac{\beta_k \|g_k\|}{\beta_k^2 (\|g_k\|^2 / g_k^T Q^{-1} g_k) - \|g_k\|}$$

where t_k is the unique positive root of eq. (3.2).

<u>Proof</u> We have : $\{x_i\} \rightarrow x^*$ and therefore : $\{g_i\} \rightarrow 0$. This means that for k large enough, eq. (3.2) will be of the form:

$$P(t) = At^3 + Bt^2 + Ct - D$$

with A,B,C,D > 0, corresponding to the coefficients in eq. (3.2). Now P(t) is a continuous increasing function, mapping $[0, \infty)$ into $[-D, \infty)$ hence there exists a unique positive root t_k.

We shall denote by L(t) the linear part of P(t) i.e. L(t) = Ct-D. Since for $t \geq 0$ $P(t) \geq L(t)$ and $P(0) = L(0) = -D$, we find that the root t_L of L(t) = 0 is larger than the root t_k of P(t) = 0, see Figure 2.

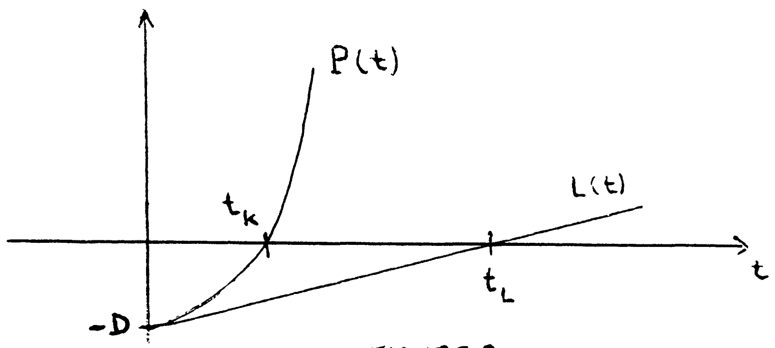

FIGURE 2

we obtain: $t_k \leq t_L = D/C = \dfrac{\beta \|g\|}{\beta^2 (\|g\|^2 / g^T Q^{-1} g) - \|g\|}$

■

Note that since $\dfrac{\|g_k\|^2}{g_k^T Q^{-1} g_k}$ is bounded, then

$$t_k = O(\|g_k\|)$$

the symbol $O(a)$ meaning $\lim_{a \to 0} \frac{O(a)}{a} \leq A < \infty$. We further denote by $t_k \sim \gamma_k$ the fact that t_k <u>behaves asymptotically as γ_k</u> i.e.

$$\lim_{k \to \infty} \left(\frac{t_k}{\gamma_k}\right) = 1$$

Rearranging eq. (3.2) we have:

$$t_k = \frac{\beta_k \|g_k\|}{\beta_k^2 w_k - a_k \|g_k\| + 3/2 a_k \beta_k w_k t_k + a_k^2 u_k t_k^t} \tag{3.6}$$

Now, w_k and u_k are bounded and assuming that $x_k \to x^*$, $\lim_{k \to \infty} \|g_k\| = \lim_{k \to \infty} t_k = 0$ and we conclude that

$$t_k \sim \frac{\beta_k \|g_k\|}{\beta_k^2 w_k} = \frac{\beta_k \|g_k\|}{\beta_k^2 (\|g_k\|^2 / g_k^T Q^{-1} g_k)} = \frac{g_k^T Q^{-1} g_k}{\beta_k \|g_k\|} \tag{3.7}$$

We shall use (3.7) in examining the following expression for large k, under the assumption that $x_k \to x^*$:

$$\frac{1}{\|y_k\|}\left(\frac{\beta_k \|g_k\| t_k}{g_k^T Q^{-1} g_k} - 1\right) \tag{3.8}$$

Using eq. (3.2), (3.8) can be written as:

$$-\frac{a_k^2}{\beta_k} \frac{g_k^T Q^{-1} g_k}{\|y_k\| \cdot \|g_k\|^3} t_k^3 - 3/2 a_k \frac{\|g_k\|}{(g_k^T Q^{-1} g_k)\|y_k\|} t_k^2 + \frac{a_k}{\beta_k} \frac{t_k}{\|y_k\|} \tag{3.9}$$

The first term in (3.9) will converge to zero for $k \to \infty$, as $t_k \sim O(\|g_k\|)$ and $\|g_k\| = O(\|y_k\|)$.

With the use of (3.7), the last two terms in (3.9) behave asymptotically as :

$$-3/2\, \alpha_k \frac{\|g_k\|}{(g_k^TQ^{-1}g_k)\|y_k\|} \left(\frac{g_k^TQ^{-1}g_k}{\beta_k\|g_k\|}\right)^2 + \frac{\alpha_k}{\beta_k} \frac{g_k^TQ^{-1}g_k}{\beta_k\|g_k\|\cdot\|y_k\|}$$

$$= -1/2 \frac{\alpha_k}{\beta_k^2} \left(\frac{g_k^TQ^{-1}g_k}{\|g_k\|\,\|y_k\|}\right) \tag{3.10}$$

The proof of Theorem 2 follows.

<u>Proof of Theorem 2</u>

Using in (3.5) the asymptotic behavior of t_k and $\frac{1}{\|y_k\|}\left(\frac{\beta_k\|g_k\|t_k}{g_k^TQ^{-1}g_k} - 1\right)$ as given by (3.7) and (3.10) respectively, we obtain:

$$\limsup_{k\to\infty} \frac{\|x_{k+1}-x^*\|^2}{\|x_k-x^*\|^4} = \limsup_{k\to\infty} \Bigg\{ \frac{\alpha_k^2}{4\|y_k\|^4} \cdot \left(\frac{g_k^TQ^{-1}g_k}{\beta_k\|g_k\|}\right)^4$$

$$+ \frac{\alpha_k g_k^TQ^{-1}g_k}{\|g_k\|\cdot\|y_k\|^3} \left(\frac{g_k^TQ^{-1}g_k}{\beta_k\|g_k\|}\right)^2 \left(\frac{(-1/2)\alpha_k}{\beta_k^2}\left(\frac{g_k^TQ^{-1}g_k}{\|g_k\|\cdot\|y_k\|}\right)\right)$$

$$+ \left(\frac{(-1/2)\alpha_k}{\beta_k^2}\left(\frac{g_k^TQ^{-1}g_k}{\|g_k\|\cdot\|y_k\|}\right)\right)^2 \Bigg\}$$

$$= \limsup_{k\to\infty} \left\{1/4 \frac{\alpha_k^2}{\beta_k^4} \left[\left(\frac{g_k^TQ^{-1}g_k}{\|g_k\|\cdot\|y_k\|}\right)^2 - \left(\frac{g_k^TQ^{-1}g_k}{\|g_k\|\cdot\|y_k\|}\right)^4 \right]\right\}$$

We define $s_k = \dfrac{g_k^TQ^{-1}g_k}{\beta_k\|g_k\|\cdot\|y_k\|}$.

Note that : $\dfrac{2\sqrt{\lambda_1\lambda_n}}{\lambda_1+\lambda_n} \le s_k = \dfrac{y_k^TQy_k}{\|Qy_k\|\cdot\|y_k\|} \le 1$

Since $\max\{v(s) =: s^2 - s^4\} = 1/4$ (see Figure 3), then

$$\limsup_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} \le \frac{\bar{a}}{4\underline{\beta}^2}$$

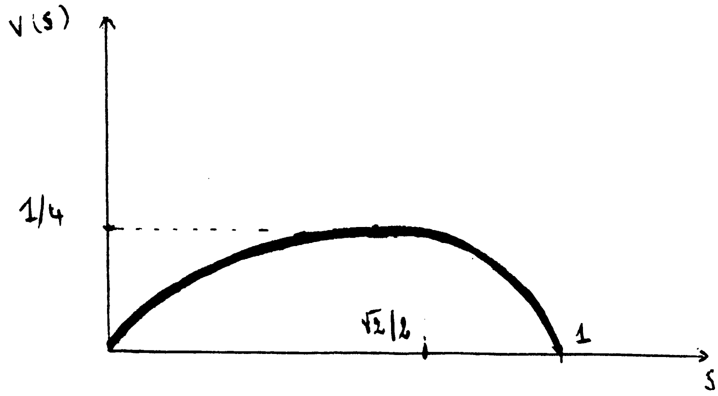

FIGURE 3

We see therefore, that if $\frac{2\sqrt{\lambda_1\lambda_n}}{\lambda_1+\lambda_n} \ge \frac{\sqrt{2}}{2}$ then (see Figure 3) $v\left(\frac{2\sqrt{\lambda_1\lambda_n}}{\lambda_1+\lambda_n}\right) \le 1/4$ giving the sharper upper bound in Theorem 2.

■

We now present a version of the algorithm without linesearch which will be called the α-method. Here, we choose a suitable stepsize $\bar{t}_k$ and determine α_k so as to make $\bar{t}_k$ an exact stepsize. For convenience we use the parameter $\rho_k = \beta_k / \alpha_k$. Subsitituting $\beta_k = \rho_k\alpha_k$ in (3.3) and dividing by α_k yields:

$$\alpha_k = \frac{\|g_k\|\bar{t}_k + \|g_k\|\cdot\rho_k}{u_k\bar{t}_k^3 + 3/2\rho_k w_k \bar{t}_k^2 + \rho_k^2 w_k \bar{t}_k} \tag{3.11}$$

For α_k to be finite and bounded away from zero as $x_k \to x^*$, it is necessary that

$$\liminf_{k\to\infty} \frac{\|g_k\|}{\bar{t}_k} > 0 \tag{3.12}$$

$$\limsup_{k\to\infty} \frac{\|g_k\|}{\bar{t}_k} < \infty \tag{3.13}$$

Possible choices for $\bar{t}_k$ satisfying (3.12), (3.13) include:

$$\bar{t}_k = \|g_k\| \quad \text{or} \quad \bar{t}_k = \frac{g_k^T Q^{-1} g_k}{\beta_k \|g_k\|},$$

the latter explained by (3.7).

We can immediately state the following theorem:

Theorem 3: [Convergence Properties of the a-method] The a-method, where a_k is determined by (3.11) and with $\bar{t}_k$ satisfying (3.12) and (3.13), converges to the solution of (QP) with quadratic rate of convergence, as given by Theorem 2.

Proof The proof is immediate from Theorem 1 and Theorem 2 since the a-method is a special case of the SOSD algorithm.

§4. THE SOSD-E ALGORITHM FOR A GENERAL C^2-FUNCTION

In this section we generalize the quadratic convergence result from the previous section to general C^2-functions.

Throughout this section we consider the problem:

$$\text{(GP)} \qquad \min_{x \in R^n} f(x).$$

Let x^* be a strict local minimum of f, i.e. the second order sufficient optimality conditions $g^* =: \nabla f(x^*) = 0$, $H^* =: \nabla^2 f(x^*)$ positive definite, hold at x^*.

The generalization of Theorem 2 is obtained by showing that, for a C^2-function, crucial expressions (such as the inequality in Lemma 1, or eq. (3.4)), which were obtained for a quadratic function, differ only by a term of magnitude $O(\|x_k - x^*\|)$. First we derive some auxiliary results.

Lemma 2:

If

(1) $f : D \subset R^n \to R$ is C^2

(2) There exists an open convex set D_0 in D, containing x^*, such that:

$$\exists\, R > 0 \ni \forall\, x,y \in D_0 : \|H(x) - H(y)\| \le R\|x - y\| \tag{4.1}$$
i.e. H is locally Lipschitz at x^*.

(3) $\forall\, x \in D_0$: H(x) and H^* are bounded and positive definite

Then

$$g = H^*(x - x^*) + O(\|x - x^*\|^2) \tag{4.2}$$

$$\|H^{-1} - H^{*-1}\| = O(\|x - x^*\|) \tag{4.3}$$

$$|g^T Hg - g^T H^* g| = O(\|x - x^*\|^3) \tag{4.4}$$

$$|g^TH^{-1}g - g^TH^{*-1}g| = O(\|x - x^*\|^3) \tag{4.5}$$

$$|g^TH^{-1}H^*g - \|g\|^2| = O(\|x - x^*\|^3) \tag{4.6}$$

$$|g^TH^{-1}H^*g - g^TH^{*-1}g| = O(\|x - x^*\|^3) \tag{4.7}$$

$$|g^TH^*y - \|g\|^2| = O(\|x - x^*\|^3) \quad , \quad y = x - x^* \ . \tag{4.8}$$

Proof: We start with the Taylor expansion for $g =: \nabla f(x) : R^n \to R^n$ around x^* :

$$g = \nabla f(x^*) + H(\bar{x})(x - x^*) \qquad (\bar{x} = \lambda x + (1 - \lambda)x^* \text{ for some } 0 \le \lambda \le 1.)$$

Using $\nabla f(x^*) = 0$, this can be rewritten as

$$g = H^*(x - x^*) + (H(\bar{x}) - H^*)(x - x^*).$$

Since $\|\bar{x} - x^*\| \le \|x - x^*\|$, this yields

$$\|g - H^*(x - x^*)\| \le \|H(\bar{x}) - H^*\|.\|x - x^*\| \le R\|\bar{x} - x^*\| \ \|x - x^*\|, \text{ using (4.1)},$$

$$\le R\|x - x^*\|^2$$

In addition we have, for any vector $v \in R^n$:

$$|v^Tg - v^TH^*(x - x^*)| \le \|v\|R\,\|x - x^*\|^2$$

and therefore

$$v^Tg = v^TH^*(x - x^*) + O(\|x - x^*\|^2) \tag{4.9}$$

To prove (4.3) let us look at:

$$\begin{aligned}\|H^{-1} - H^{*-1}\| &= \|-H^{-1}(H - H^*)H^{-1}\| \\ &\le \|H^{-1}\|\,\|H^{-1}\|\,\|H - H^*\| \\ &\le \|H^{-1}\|\,\|H^{-1}\|R\|x - x^*\| \qquad \text{by (4.1)}\end{aligned}$$

To prove (4.5) note that

$$|g^THg - g^TH^*g| = |g^T(H - H^*)g| \le \|g\|^2R\,\|x - x^*\|$$

But from : (4.2): $\|g\| \le \|H^*\|.\|x - x^*\| + O(\|x - x^*\|)$

and therefore:

$$|g^T H g - g^T H^* g| = O(\|x - x^*\|^3)$$

The proof of expressions (4.6)–(4.8) is similar and hence omitted. ■

Lemma 3: Under the assumptions of Lemma 2, any sequence $x_k \to x^*$ generated by a convergent descent method for (GP) has the following property:

$$\exists\ K_0 \ni \forall\ k > K_0: \quad \|x_{k+1} - x^*\| \le \frac{\sqrt{\lambda_n^*}}{\sqrt{\lambda_1^*}} \cdot \|x_k - x^*\|$$

where λ_1^* and λ_n^* are the minimum and maximum eigenvalues of H^* respectively.

Proof: For any descent method we have:

$$f(x_{k+1}) \le f(x_k)\ ;$$

using the Taylor expansion of f, the inequality can be rewritten as

$$f(x^*) + g^{*T}(x_{k+1} - x^*) + 1/2(x_{k+1} - x^*)H(y_1)(x_{k+1} - x^*)$$
$$\le f(x^*) + g^{*T}(x_k - x^*) + 1/2(x_k - x^*)H(y_2)(x_k - x^*)$$

where y_1 is in the interval (x_{k+1}, x^*)

y_2 is in the interval (x_k, x^*).

Recalling that $g^* = 0$, this further reduces to

$$(x_{k+1} - x^*)^T H(y_1)(x_{k+1} - x^*) \le (x_k - x^*)H(y_2)(x_k - x^*)$$

Now, since $x_k \to x^*$, we have:

$$y_1 \to x^* \text{ and } H(y_1) \to H^*$$

$$y_2 \to x^* \text{ and } H(y_2) \to H^*$$

The proof is completed by recalling that for any vector $v \in R^n$:

$$\lambda_1\|v\|^2 \le v^TH^*v \le \lambda_n\|v\|^2$$

in particular for $v = x_{k+1} - x^*$ and $v = x_k - x^*$.

Let us now look at the necessary condition for the stepsize t_k to be exact for (GP):

$$(d_k + tz_k)^T \nabla f(x_k + td_k + t^2 z_k)\,|_{t=t_k} = 0$$

This yields, using the boundedness of d_k and z_k together with (4.9):

$$(d_k + t_kz_k)^T H^*(x_k + t_kd_k + 1/2t_k^2z_k - x^*) +$$
$$t_k(O(\|x_{k+1} - x^*\|)^2) + O(\|x_{k+1} - x^*\|^2) = 0$$

Using Lemma 3 we obtain

$$(d_k + t_kz_k)^T H^*(x_k - x^* + t_kd_k + 1/2t_k^2z_k) + t_k(O(\|x_k - x^*\|^2)) + O(\|x_k - x^*\|^2) = 0$$

$$1/2\, z_k^TH^*z_kt_k^3 + 3/2d_k^TH^*z_kt_k^2 + (z_kH^*(x_k - x^*) + d_k^TH^*d_k$$
$$+ O(\|x - x^*\|^2))t_k + d_k^TH^*(x_k - x^*) + O(\|x_k - x^*\|^2) = 0 \tag{4.10}$$

After substituting d_k and z_k, we obtain, omitting the index k, and denoting $y = x - x^*$

$$\alpha^2\frac{g^TH^*g}{2\|g\|^2}t^3 + 3/2\alpha\beta\frac{g^TH^{-1}H^*g}{g^TH^{-1}g}t^2 + \Big(-\alpha\frac{g^TH^*y}{\|g\|} +$$
$$\beta^2\frac{g^TH^{-1}H^*H^{-1}g}{(g^TH^{-1}g)^2}\|g\|^2 + O(\|y\|)\Big)t - \beta\frac{g^TH^{-1}H^*y}{g^TH^{-1}g} + O(\|y\|) = 0 \tag{4.11}$$

With the help of Lemma 2 this can be further transformed to:

$$\alpha^2\frac{g^TH^*g}{2\|g\|^2}t^3 + 3/2\alpha\beta\frac{\|g\|^2 + O(\|y\|^3)}{g^TH^{*-1}g + O(\|y\|^3)}t^2 + \Big(\frac{-\alpha\|g\|^2 + O(\|y\|^3)}{\|g\|}$$

$$+ \beta^2 \frac{(g^T H^{*-1} g + O\|y\|^3)\|g\|^2}{(g^T H^{-1} g)(g^T H^{*-1} g + O(\|y\|^3))} + O(\|y\|)) t - \beta \frac{g^T H^{-1} g + O(\|y\|^3)}{g^T H^{-1} g} \|g\| + O(\|y\|^2) = 0$$

and finally:

$$\alpha^2 \frac{g^T H^* g}{2\|g\|^2} t^3 + 3/2\alpha\beta \frac{\|g\|^2(1+O(\|y\|))}{g^T H^{*-1} g(1+O(\|y\|))} t^2 + \Big(\beta^2 \frac{\|g\|^2(1+O(\|y\|))}{g^T H^{*-1} g(1+O(\|g\|))}$$

$$- \alpha\|g\|(1+O(\|y\|)) + O(\|y\|)\Big) t - \beta\|g\|(1+O(\|y\|)) + O(\|y\|^2) = 0 \tag{4.12}$$

Introducing the index k, we obtain for t_k analogously to (3.6):

$$t_k = \frac{\beta_k\|g_k\|(1+(O(\|y_k\|))}{\alpha_k^2 \frac{g_k^T H^* g_k}{2\|g_k\|^2} t_k^2 + 3/2\alpha_k\beta_k \frac{\|g_k\|^2(1+O(\|y_k\|))}{g_k^T H^{*-1} g_k(1+O(\|y_k\|))} t_k + \beta_k^2 \frac{\|g_k\|^2 (1+O(\|y_k\|))}{(g_k^T H^{*-1} g_k)(1+O(\|y_k\|))} + O(\|y_k\|)} \tag{4.13}$$

This shows that asymptotically ($k \to \infty$) we may write:

$$t_k = \tau_k(1+O(\|y_k\|)) \tag{4.14}$$

where τ_k is the solution of eq. (3.2) with $Q = H^*$ and $g = \nabla f(x)$.

This also means that, similar to Section 3:

$$t_k = O(\|g_k\|) = O(\|y_k\|).$$

We now develop the expression for the rate of convergence which will be the equivalent of (3.4) for a general C^2-function:

$$\frac{\|x_{k+1} - x^*\|^2}{\|x_k - x^*\|^4} = 1/4 \frac{z_k^T z_k}{\|y_k\|^4} t_k^4 + \frac{d_k^T z_k}{\|y_k\|^4} t_k^3 + \frac{z_k^T y_k + d_k^T d_k}{\|y_k\|^4} t_k^2 + \frac{2d_k^T y_k}{\|y_k\|^4} t_k + \frac{1}{\|y_k\|^2}$$

Substituting d_k and z_k and using (4.4)−(4.8), this gives:

$$1/4\,\frac{\alpha_k^2}{\|y_k\|^4}t_k^4 + \frac{\alpha_k\beta_k}{\|y_k\|^4}t_k^3 + \Big(-\alpha_k\frac{(g_k^TH^{*-1}g_k)(1+O(\|y_k\|))}{\|g_k\|\cdot\|y_k\|^4}$$

$$+\ \beta_k^2\frac{\|g_k\|^2\|y_k\|^2(1+O(\|y_k\|))}{(g_k^TH^{*-1}g_k)^2(1+O(\|y_k\|))^2\|y_k\|^4}\Big)t_k^2 - \frac{2\beta_k\|g_k\|\|y_k\|^2(1+O(\|y_k\|))}{\|y_k\|^4(g_k^TH^{*-1}g_k)(1+O(\|y_k\|))}t_k + \frac{1}{\|y_k\|^2}$$

Bearing in mind that

$$1 = 1+O(\|y\|)$$

$$\frac{1}{1+O\|y\|} = 1+O(\|y\|)$$

and omitting the index k, this can finally be written as:

$$\Big(\frac{\alpha^2t^4}{4\|y\|^4} + \frac{\alpha\beta t^3}{\|y\|^4} + \Big(-\alpha\frac{g^TH^{*-1}g}{\|g\|\|y\|^4} + \beta^2\frac{\|g\|^2}{(g^TH^{*-1}g)\|y\|^2}\Big)t^2$$

$$-\ 2\beta\frac{\|g\|}{(g^TH^{*-1}g)\|y\|^2}t + \frac{1}{\|y\|^2}\Big)(1 + O\|y\|) \tag{4.15}$$

We now formulate the main theorem of this section:

Theorem 4: [Quadratic Convergence Rate of SOSD-E for C^2-functions]

Let f satisfy assumptions (i)−(iii) of Lemma 2. Assume that the sequence $\{x_k\}_0^\infty$ generated by the SOSD algorithm converges to a strict local minimum x^* of (GP). Then:

$$\limsup_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} \le 1/4 \, \frac{\bar{a}}{\underline{\beta}^2}$$

Moreover, if $\dfrac{2\sqrt{r^*}}{1+r^*} \ge \dfrac{\sqrt{2}}{r}$ we have the sharper bound:

$$\limsup_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} \le 1/2 \, \frac{\bar{a}}{\underline{\beta}^2} \left(\frac{2\sqrt{r^*}}{1+r^*}\right)\left(\frac{1-r^*}{1+r^*}\right)$$

Here $r^* = \lambda_1^* / \lambda_n^*$ is the condition number of H^*.

Proof: For k large enough, x_k will be in a neighborhood D_0 of x^* (see Lemma 2). Substituting : $t_k = \tau_k(1+O(\|y_k\|))$ from (4.14) in (4.15) yields:

$$\limsup_{k\to\infty} \frac{\|x_{k+1} - x^*\|^2}{\|x_k - x^*\|^4} =$$

$$= \limsup_{k\to\infty} \Big\{ \frac{a_k^2 \tau_k^4}{4\|y_k\|^4} + \frac{a_k \beta_k \tau_k^3}{\|y_k\|^4} + \Big(\beta_k^2 \frac{\|g_k\|^2}{(g_k^T H^{*-1} g_k)\|y_k\|^2} - a_k \frac{g_k^T H^{*-1} g_k}{\|g_k\|\,\|y_k\|^4}\Big)\tau_k^2$$

$$- 2\beta_k \frac{\|g_k\|}{(g_k^T H^{*-1} g_k)\|y_k\|^2}\tau_k + \frac{1}{\|y_k\|^2} \Big\} (1+O(\|y_k\|))$$

Since $y_k \to 0$, $1+O(\|y_k\|) \to 1$. Furthermore the expression in the curly brackets is exactly as in the quadratic case with $Q = H^*$ and the result of Theorem 2 applies.

∎

Remark: An alternative way of proving Theorem 3 is to use (3.4) in [11] and to show that $\dfrac{\|x_{k+1} - x_k\|}{\|-H_k^{-1} g_k\|} \to 0$. This can be proven, however, we chose not to prove the theorem this way since it does not provide us with the explicit bound in Theorem 4, in particular the effect of the condition number of H^* on the rate of convergence.

§5. THE SOSD METHOD WITH INEXACT LINESEARCH

In this section we study the SOSD method, with the stepsize computed via an Armijo-Goldstein type rule; this version is denoted by SOSD-I. We demonstrate that SOSD-I maintains the same convergence properties as the SOSD method with exact linesearch. We define (for g, d and z such that $g^T d \neq 0$)

$$\gamma(t) = \frac{f(x + td + 1/2t^2 z) - f(x)}{tg^T d}$$

At the k-th iteration, the SOSD-I method is then given by:

compute:
$$d_k = - \frac{\beta_k \|g_k\|}{g_k^T H_k^{-1} g_k} H_k^{-1} g_k$$

$$z_k = - \frac{\alpha_k}{\|g_k\|} g_k$$

compute t_k : such that:

$$0 < \sigma \leq \gamma(t_k) \leq 1 - \sigma \leq 1 \qquad (5.1)$$

starting from the initial guess:

$$t_k^0 = \left| \frac{g_k^T H_k^{-1} g_k}{\beta_k \|g_k\|} \right|$$

update x_k:
$$x_{k+1} = x_k + t_k d_k + 1/2 t_k^2 z_k.$$

Here, σ is a number satisfying $0 < \sigma < 1/2$ (usually : $\sigma \simeq 10^{-4}$). We now formulate and prove two theorems concerning SOSD-I:

Theorem 5 [Convergence of SOSD-I]

If f is a C^2-function and bounded below, then

(i) for all k such that $g_k \neq 0$, there exists t_k satisfying (5.1).

(ii) for the sequence $\{x_k\}_0^\infty$, generated by the SOSD-I algorithm:

$$\lim_{k \to \infty} x_k = \hat{x} \text{ such that } \nabla f(\hat{x}) = 0$$

Proof: (i) (We omit the index k) Suppose $g^T d < 0$ and consider the following Taylor expansion: $(0 \leq \theta \leq 1)$

$$f(x+td+1/2t^2z) = f(x) + g^T(td+1/2t^2z) + 1/2t^2(d+1/2tz)^T H\,(x+\theta(td+1/2t^2z))\,(d+1/2tz)$$

This gives :

$$\gamma(t) = \quad 1 + t\left(\left(\frac{z^Tg}{d^Tg}\right) + \frac{(d+1/2tz)^T H(x+\theta(td+1/2t^2z))(d+1/2tz)}{2g^Td}\right) \tag{5.2}$$

Since d, z and H are bounded, it follows that $\gamma(t)$ is continuous and

$$\lim_{t \to 0} \gamma(t) = 1. \tag{5.3}$$

Now, since $g^T d < 0$, and f is bounded below we can always find a $t > 0$ large enough such that:

$$f(x + td+ 1/2t^2z) \geq f(x) + \sigma t g^T d.$$

We can therefore always find a t such that $\gamma(t) < \sigma$. This, together with the continuity of γ and (5.3) proves (i).

(ii) To prove (2), we distinguish between two cases:

Case (a) Suppose that the sequence $\{t_k\}$ of stepsizes satisfies $\{t_k\} \geq \bar{t}$ for some $\bar{t} > 0$. In this case, the left hand side inequality in (5.1) gives: $$\frac{f(x_{k+1} - f(x_k)}{t_k g_k^T d_k} \geq \sigma$$

or: $$- g_k^T d_k \leq \frac{1}{t_k \sigma} (f(x_k) - f(x_{k+1})).$$

Now, the sequence $\{f(x_k)\}$ is descending and bounded below and therefore:

$(f(x_k) - f(x_{k+1}))$ converges to zero and therefore also $g_k^T d_k$ $(= -\beta_k \|g_k\|)$ which proves (ii) for Case (a).

Case (b) $\bar{t} = 0$, in which case, there is a subsequence $\{t_{v_k}\}$ of $\{t_k\}$ converging to zero. The right hand inequality in (5.1) then yields (see (5.2)):

$$1 + t_{v_k} \left(\frac{g_{v_k}^T z_{v_k}}{g_{v_k}^T d_{v_k}} + \frac{(d_{v_k} + 1/2 t_{v_k} z_{v_k})^T H(x_k + \theta(t_{v_k} d_{v_k} + 1/2 t_{v_k}^2 z_{v_k}))(d_{v_k} + 1/2 t_{v_k} z_{v_k})}{2 g_{v_k}^T d_{v_k}} \right)$$

or: $$0 \leq -g_{v_k}^T d_{v_k} \leq \frac{t_{v_k}}{\sigma} \left(g_{v_k}^T z_{v_k} + 1/2 \, (d_{v_k} + \frac{t_{v_k}}{2} z_{v_k})^T H(x_{v_k} + \theta(t_{v_k} d_{v_k} + t_{v_k}^2 z_{v_k})) (d_{v_k} + \frac{t_{v_k}}{2} z_{v_k}) \right)$$

Now, as $t_{v_k} \to 0$, $g_{v_k}^T d_{v_k} \to 0$, hence $g_{v_k} \to 0$.

∎

Theorem 6: [Rate of Convergence of SOSD-I]

If the sequence $\{x_k\}$, generated by the SOSD-I method, converges to a strict local minimum x^* and if the function f satisfies the conditions of Lemma 2, then $\{x_k\}$ converges to x^* quadratically.

Proof: If we can prove that, for k large enough, the initial guess t_k^0 in the SOSD-I algorithm will satisfy (5.1), then the proof follows; this is so because in §4 we showed that the exact stepsize t_k behaves asymptotically as t_k^0 (i.e. $\lim_{k \to \infty} t_k^0/t_k = 1$) and the proof of the quadratic rate of convergence was based on this fact.

For k large enough, we also have that H_k will be positive definite and we can therefore omit the absolute value sign in the expression for t_k^0.

We now prove that t_k^0 satisfies (5.1) for large enough k. Indeed, substituting for z_k, d_k and t_k^0 in (5.2) and letting $k \to \infty$, in which case $t_k^0 \to 0$ and $g_k \to 0$, yields

$$\lim_{k \to \infty} \gamma(t_k^0) = 1/2$$

■

The "α-methods" of § 3, which is an exact linesearch version of SOSD for quadratic function, can be naturally extended to general C^2-functions: the results is a method free of linesearch:

<u>α-method</u>

<u>choose:</u> sequences ρ_k and t_k $(k=0,1,2,...)$

such that :

$$\begin{cases} \rho_k > 0, \ \lim\limits_{k\to\infty} \rho_k > 0 \\ t_k > 0, \ \lim\limits_{k\to\infty} t_k \to 0 \\ \lim\limits_{k\to\infty} \dfrac{\|g_k\|}{t_k} < \infty \quad (\text{e.g. } t_k = \|g_k\|) \end{cases}$$

and an initial point $x_0 \epsilon R^n$.

<u>compute:</u>

$$u_k = \frac{g_k^T H_k g_k}{2\|g_k\|^2}$$

$$w_k = \frac{\|g_k\|^2}{g_k^T H_k^{-1} g_k}$$

$$\alpha_k = \frac{\|g_k\|(t_k + \rho_k)}{u_k t_k^3 + 3/2 \rho_k w_k t_k^2 + \rho_k^2 w_k t_k}$$

$$d_k = -\rho_k \alpha_k \frac{\|g_k\|}{g_k^T H_k^{-1} g_k} H_k^{-1} g_k$$

$$z_k = -\alpha_k g_k / \|g_k\|$$

<u>update:</u>

$$x_{k+1} = x_k + t_k d_k + 1/2 t_k^2 z_k$$

Recall that α_k is chosen so as to make t_k the minimum of the quadratic approximation to f at x_k. Thus it is not hard to verify that (4.14) holds. Also, by the choice of t_k another crucial relation: $t_k = O(\|g_k\|) = O(\|y_k\|)$ holds, thus it possible to show that, if $x_k \to x^*$ then the convergence is with quadratic rate. However, no global convergence is guaranteed for the α-method. It is interesting to note that for the single variable case (n=1) the α-method is exactly the (pure) Newton method

$$x_{k+1} = x_k - f'(x_k)/f''(x_k)$$

and this is true independent of the choice of the sequences ρ_k and t_k !

The sequence ρ_k can be chosen fixed: $\rho_k \equiv \rho$. The parameter ρ then controls the "balance" between a SD step and a pure Newton step. To verify this note that if $\rho = 0$ then $d_k = 0$ and hence we have a SD step. If $\rho \to \infty$ then

$$\lim_{\rho\to\infty} \alpha_k = 0, \quad \lim_{\rho\to\infty} \alpha_k\rho = \frac{\|g_k\|}{w_k t_k} \qquad \text{so } z_k \to 0, \text{ while}$$

$t_k d_k \to -H_k^{-1} g_k$ and we have a pure Newton step (regardless of the choice of t_k.) In our numerical experiments it was found that a large ρ ($\simeq 10^6$) gives good convergence results; even with such large ρ, far from an optimal solution (where $\|g_k\|$ is large) the effect of z_k is not negligible (since z_k is multiplied by t_k^2) and it may be sufficient to secure global convergence.

§ 6. COMPUTATIONAL RESULTS

In this section we report the computational results of the performance of the various SOSD algorithms as applied to the minimization of four classical test functions (see [8]). The results are then compared to the performance of three versions of Newton's method.

The algorithms and test functions involved are:

Algorithms

N: The pure Newton method: $x_{k+1} = x_k - H_k^{-1} g_k$

N-E: The damped Newton method with exact linesearch

N-I: The damped Newton method with inexact linesearch according to the Goldstein rule as described in ([9], Section 8.3.e.), with $\sigma = 10^{-4}$

SOSD-E: The SOSD method with exact linesearch (Section 4)

SOSD-I: The SOSD method with inexact linesearch as described in Section 5 with $\sigma = 10^{-4}$

The α-method: See last part of Sec. 5, with $t_k = \|g_k\|$, $\rho_k = \rho$.

Test Functions

Rosenbrock's function (2 variables):

$$R(x) = 100(x_1^2 - x_2)^2 + (1 - x_1)^2$$

Wood's function (4 variables):

$$w(x_1,x_2,x_3,x_4) = 100(x_2 - x_1^2)^2 + (1-x_1)^2 + 90(x^2 - x_3^2)^2 + (1-x_3)^2 + 10.1((x_2-1)^2 + (x_4-1)^2) + 19.8(x_2-1)(x_4-1)$$

Extended Wood's function (20 variables):

$$W(x) = \sum_{i=1}^{5} w(x_{4i-3}, x_{4i-2}, x_{4i-1}, x_{4i})$$

Dixon's function (10 variables):

$$D_n(x) = (1-x_1)^2 + (1-x_n)^2 + \sum_{i=1}^{n-1} (x_i^2 - x_{i+1})^2 \qquad (n=10)$$

The results are summarized in the following tables: each entry designates the number of iterations required to satisfy the following stopping criterion:

$$\|x_k - x^*\| \leq 10^{-10}$$

NC(j) means that no convergence was achieved after j iterations, whereas NC means that convergence was not obtained at all.

FUNCTION: ROSENBROCK

Starting Point	Second Order Steepest Descent			Newton		
	SOSD-E	SOSD-I	α-method	N	N-E	N-I
(20,200)	$\alpha=1$ $\beta=1$	$\alpha=1$ $\beta=1$	$\rho=10^6$			
	31	67	12	5	45	78
(−1.2,1)	$\alpha=1$ $\beta=1$	$\alpha=1$ $\beta=1$	$\rho=10^6$			
	12	21	7	6	13	21
(10,10)	$\alpha=2$ $\beta=4$	$\alpha=1$ $\beta=1$	$\rho=1/2\ 10^6$			
	13	37	8	5	27	46
(−25,50)	$\alpha=1.7$ $\beta=2.89$	$\alpha=1$ $\beta=1$	$\rho=1/2\ 10^6$			
	46	56	18	5	52	85
(−25,−50)	$\alpha=1.5$ $\beta=2.25$	$\alpha=s$ $\beta=1$	$\rho=1/2\ 10^6$			
	32	74	12	5	52	89

FUNCTION: WOOD

Starting Point	Second Order Steepest Descent			Newton		
	SOSD-E	SOSD-I	α-method	N	N-E	N-I
(−3,−1,−3,−1)	$\alpha=4$ $\beta=16$	$\alpha=1$ $\beta=1$	$\rho=1/2\ 10^6$			
	25	32	30	NC	NC	NC
(0,2,0,2)	$\alpha=5$ $\beta=25$	$\alpha=1$ $\beta=1$	$\rho=1/2\ 10^6$			
	11	19	19	NC	NC	NC
(0.1,1.0,0.1,10)	$\alpha=10$ $\beta=100$	$\alpha=1$ $\beta=1$	$\rho=1/2\ 10^6$			
	9	10	18	NC	NC	NC
(200,−300,450,250)	$\alpha=9$ $\beta=81$	$\alpha=9$ $\beta=81$	$\rho=1/2\ 10^6$			
	23	45	27	32	25	33
(−200,−300, −450,−250)	$\alpha=9$ $\beta=81$	$\alpha=9$ $\beta=81$	$\rho=1/2\ 10^6$			
	17	46	32	38	31	43

FUNCTION: EXTENDED WOOD

Starting Point	Second Order Steepest Descent			Newton		
	SOSD-E	SOSD-I	α-method	N	N-E	N-I
p_1	$\alpha=5$ $\beta=25$	$\alpha=5$ $\beta=25$	$\rho=10^6$			
	26	39	28	NC	NC	NC
p_2	$\alpha=5$ $\beta=50$	$\alpha=5$ $\beta=50$	$\rho=8\ 10^6$			
	40	60	53	NC	NC	NC
p_3	$\alpha=10$ $\beta=100$	$\alpha=5$ $\beta=25$	$\rho=5\ 10^6$			
	37	37	39	49	NC	44
p_4	$\alpha=10$ $\beta=100$	$\alpha=10$ $\beta=100$	$\rho=5\ 10^6$			
	17	16	16	17	16	17
p_5	$\alpha=10$ $\beta=100$	$\alpha=10$ $\beta=100$	$\rho=1/2\ 10^6$			
	25	36	20	24	NC	26

$p_1=(-3,-1,-3,-1,\dots)$
$p_2=(-1,-2,-3,-4,\dots)$
$p_3=(20,19,\dots,11,-11,-12,\dots,-20)$
$p_4=(10,-20,30,-40,50,10,10,\dots,10,-50,40,\dots,-10)$

FUNCTION: DIXON

Starting Point	Second Order Steepest Descent			Newton		
	SOSD-E	SOSD-I	α-method	N	N-E	N-I
p_1	$\alpha=10$ $\beta=100$	$\alpha=10$ $\beta=100$	$\rho=5\ 10^6$			
	21	24	47	218	NC	NC
p_2	$\alpha=10$ $\beta=100$	$\alpha=10$ $\beta=100$	$\rho=5\ 10^6$			
	21	25	31	610	NC	NC
p_3	$\alpha=10$ $\beta=100$	$\alpha=10$ $\beta=100$	$\rho=1/2\ 10^6$			
	28	34	46	418	NC	NC
p_4	$\alpha=10$ $\beta=100$	$\alpha=10$ $\beta=100$	$\rho=1/2\ 10^6$			
	22	27	33	NC(1000)	NC	NC
p_5	$\alpha=10$ $\beta=100$	$\alpha=10$ $\beta=100$	$\rho=1/2\ 10^6$			
	27	33	47	685	NC	NC

$p_1=(-3,-1,-3,\ldots)$

$p_2=(-1,-2,-3,-4,\ldots)$

$p_3=(-100,-100,1,1,-100,-100,1,1,\ldots,-100,-100)$

$p_4=(0,-10,0,-10,0,\ldots)$

$p_5=(100,200,300,400,-500,600,700,800,900,1000)$

Remarks

(i) The parameters α and $\rho = \frac{\beta}{\alpha}$ in SOSD-E and SOSD-I were chosen between 1 and 10. No attempt was made to choose the best values for the parameters, so that the results reflect the typical behavior of the methods.

(ii) In SOSD-I, less than two function evaluations were needed on the average to perform the inexact linesearch.

Some general conclusions can be deduced form Tables 2–5. First we point out that β should be chosen greater than one. This can be motivated by the coefficient $\frac{\alpha}{\beta^2}$ in the rate of convergence result given in Theorem 4. Our experience so far suggests using $\frac{\beta}{\alpha} \simeq 10$(†).

As previously mentioned, for the α-method, it was found that a large value of ρ, typically $\rho \simeq 10^6$, gave good results. Here we remind that the α-method does not require a linesearch and is therefore much more efficient in terms of function evaluations than the other methods.

---

† This was further corroborated by additional numerical experiments performed by J. Zowe and his students at the University of Bayreuth. (private communication to the authors)

References

(1) Ben-Israel, A., Ben-Tal, A. and Zlobec, S. Optimality in Nonlinear Programming: A Feasible Directions Approach, Wiley Interscience, New York 1981.

(2) Ben-Tal, A. "Second Order and Related Extremality Conditions in Nonlinear Programming" J.O.T.A., 31 1980.

(3) Ben-Tal, A. and Zowe, J. "Directional Derivatives in Nonsmooth Optimization" J.O.T.A., 47 (1985)483-490.

(4) Ben-Tal, A. and Zowe, J. "A Unified Theory of First and Second Order Conditions for Extremum Problems in Topologycal Vector Spaces" Math. Prog. Study 19 (1982)39-76.

(5) Dennis, J.E. and Schnabel, R.B. Numerical Methods for Unconstrained Optimaization and Nonlinear Equations Prentice Hall 1983

(6) Fletcher, R. Practical Methods of Optimization, Vol.1, Wiley, Chichester, 1980.

(7) Luenberger, D.G. Introduction to Linear and Nonlinear Programming, Addison-Wesley, Reading Mass., 1973.

(8) Moré, J.J., Grabow, B.S. and Hillstrom, K.E. "Testing Unconstrained Optimization Software" ACM Trans. on Math. Software, 7 (1981)17-41.

(9) Ortega, J.M. and Rheinholdt, W.C. Iterative Solutions of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

(10) Zangwill, W.I. Nonlinear Programming: A Unified Approach, Prentice-Hall, Englewood Cliffs, N.J., 1967.

(11) Dennis, J.E. and More, J.J., "Quasi Newton Methods, Motivation and Theory" SIAM Review 19 (1977), 46-89.