ANALYSIS OF CONTINUOUS METHODS FOR UNCONSTRAINED OPTIMIZATION AND THEIR DISCRETIZATIONS

von

Peter MITTER

Christoph W.UEBERHUBER

Forschungsbericht Nr.121

Juli 1977

PETER MITTER

INSTITUT FUER HOEHERE STUDIEN
STUMPERGASSE 56
A 1060 WIEN

CHRISTOPH W. UEBERHUBER

INSTITUT FUER NUMERISCHE MATHEMATIK
TECHNISCHE UNIVERSITAET WIEN
GUSSHAUSSTRASSE 27 - 29
A 1040 WIEN

Abstract

The well-known difficulties with the treatment of illconditioned unconstrained optimization problems can be explained by analogous difficulties with stiff differential equations. This observation provides the basis for an analysis of optimization problems and reveals new classes of optimization methods. This paper is primarily theoretic, a subsequent paper will be devoted to practical aspects of the proposed methods.

1. Introduction

The efficiency of a method to locate an unconstrained minimum of a real-valued function f(x) depends very much on individual properties of f as well as on the initial point x_o. Concerning well-behaved functions and starting points near the minimum one hardly will have difficulties. If in the same situation only a poor initial estimate of the location of the minimum is available, one is up against the same problem as that of solving a system of nonlinear equations. Like in the latter case, imbedding methods seem to be convenient (see e.g. Section 7.5. in Ortega-Rheinboldt [13]). This paper investigates the treatment of ill-conditioned optimization problems with good as well as bad initial estimates. Such problems correspond to functions with a minimum lying in a narrow valley with steep sides. Generally it is not very hard to get down to the bottom of the valley, but severe problems arise when following the bottom towards the minimum. Because of the shape of the function many common methods lead to oscillating iteratives. This situation can be described and investigated in terms of analogous phenomena which appear at the numerical integration of stiff ordinary differential equations by means of methods lacking certain stability properties. A similar situation concerning solution methods for systems of nonlinear equations was described by Boggs [2]. As explicit methods generally do not have the required stability properties, suitable implicit methods are considered.

Section 2 investigates connections between conventional minimization methods (or discrete methods, as we will call them) defined by a difference equation, and continuous minimization methods, defined by differential equations. In Section 3 continuous methods are analysed by means of Ljapunov's stability theory. In Section 4 the results of the preceding section are applied to discretizations and a class of suitable methods is presented. As a special but very efficient technique, the implicit Euler method is studied in greater detail. It is shown to be a highly stable method even in the case of poor initial estimates. The stabilty concepts employed for this purpose are defined and discussed in an appendix.

Throughout this paper, the function to be minimized is denoted by $f: \mathbb{R}^n \to \mathbb{R}$. It is assumed that $f \in C^2$. The gradient is denoted by $g(x)$, the Hessian matrix by $g'(x)$. The term *minimum* is used to designate the point $x^* \in \mathbb{R}^n$ where f (locally) takes on its least value. (x,y) denotes the usual inner product $\Sigma x_i y_i$ in $\mathbb{R}^n$, and $\|x\|$ stands for the norm $(x,x)^{1/2}$.

2. Lyapunov's Method for Discrete Optimization

The determination of a stationary point (especially a minimum) via a Quasi-Newton-method

$$x_{k+1} = x_k - hA(x_k)g(x_k) =: G(x_k)$$

is conceptually equivalent to the determination of a fixed point $x^* = G(x^*)$ of the function $G: \mathbb{R}^n \to \mathbb{R}^n$. Conditions for the convergence of Quasi-Newton-methods can therefore be obtained on the basis of well-known fixed point theorems for contracting mappings (cf. e.g. Krasnoselskii [9]). If for example the following relation holds

$$\|Gx - x^*\| < \|x - x^*\| \quad \forall x \in D - \{x^*\},$$

where D denotes an open neighbourhood of x^* such that $x_k \in D \quad \forall k \in \mathbb{N}$, then the convergence $x_k \to x^*$ is guaranteed. In this case the existence of the distance function $V(x) := \|x - x^*\|$ and the monotonicity $V(Gx) < V(x)$ imply the convergence $x_k \to x^*$. The following theorem, which is due to W. Hahn, yields the same result for a generalized distance function:

<u>Theorem 2.1</u>

Assume that $G: \mathbb{R}^n \to \mathbb{R}^n$ is continuous on an open neighbourhood D of a fixed point x^* and that $V: \mathbb{R}^n \to \mathbb{R}^1$ is continuous on D and

a) $V(x^*) = 0; \quad V(x) > 0 \quad \forall x \in D - \{x^*\}$
b) $V(Gx) < V(x)$ whenever $x, Gx \in D - \{x^*\}$.

Then there exists a $\delta > 0$ such that whenever $x_o \in \{x: \|x - x^*\| < \delta\}$ the sequence $\{x_k\}$ defined by $x_k = Gx_{k-1}$, $k \in \mathbb{N}$ converges to x^*.

<u>Proof</u>

Hahn [8] or Ortega [12].

A function V with such properties is called a Lyapunov-function. Ortega [12] p. 274 raised the question if the use of Lyapunov-functions leads to convergence results, which may not be concluded from the contraction principle. This question can be answered in the negative, because of the following theorem.

Theorem 2.2

Let G be a continuous operator, which maps a complete metric space R into itself, the space having finite diameter with respect to a metric $\rho_o(x,y)$. Assume that G has a unique fixed point in R, and that the successive approximations $x_k = Gx_{k-1}$ converge uniformly (in initial approximations $x_o \in R$) to this fixed point. Then one can define an equivalent metric $\rho(x,y)$ in R for which G is a contraction operator, i.e. $\rho(Gx,Gy) < q\rho(x,y)$, $0 < q < 1$.

Proof

Krasnoselskii [9], Theorem 3.5.

The investigation of the convergence of iterative methods with the help of Lyapunov's stability theory for difference equations therefore does not lead to an extension of the results which are obtainable on the basis of the contraction principle. Concerning convergence results both approaches are equivalent. This raises the question if one may take advantage of the results of Lyapunov's stability theory for differential equations, which are much stronger than those for difference equations. This is the case with *continuous* optimization methods which may considered as the result of a limiting process when the stepsize of a discrete method tends to zero: e.g. in the "classical" method of steepest descent $x_{k+1} = x_k - h_k g(x_k)$ the direction $d_k := -h_k g(x_k)$ is orthogonal to the level-surface $\{x : f(x) \equiv f(x_k)\}$. For $h_k \to 0 \quad \forall k \in \mathbb{N}$ the points x_k tend to the trajectory of steepest descent, which is everywhere orthogonal to the level-surfaces. This trajectory is the solution of the initial value problem:

$$\dot{x} = -g(x), \qquad x(0) = x_o.$$

In the continuous method of steepest descent stationary points of the above differential equation are calculated (without discretization e.g. on an analog computer). Correspondingly a continuous Quasi-Newton-method is defined by the differential equation $\dot{x} = -A(x)g(x)$.

There are two ways to use Lyapunov-functions and Lyapunov's stability theory in further analysis:

a) by direct application of Lyapunov's stability theory for difference equations, and

b) by the analysis of the corresponding continuous methods and by applying the consequent results to the discretizations of the continuous methods.

The first possibility has been discussed in detail by Ortega [12]. In this paper the second alternative is studied.

3. Lyapunov's method for continuous optimization

In this section we investigate the behaviour of solution-trajectories $x(t;x_o)$ of continuous Quasi-Newton-methods

$$\dot{x} = -A(x)g(x), \qquad x(0) = x_o \tag{3.1}$$

by using the very general theorems of Lyapunov's stability theory. Throughout this section we will assume that:

(i) x_o lies in a bounded level set

$L(z) := \{x : f(x) \leq z\}$

(ii) $A(x)$ is symmetric, positive definite $\forall x \in L(z)$.

(iii) A and g are sufficiently smooth.

Under these assumptions the existence of a unique solution $x(t;x_o)$ which satisfies $x(0;x_o) = x_o$ and is defined for all $t \in \mathbb{R}_+$ is guaranteed.

Since it will be clear from the context what the underlying differential equation is, we do not refer to it explicitly in the discussion below. As a notational simplification, but without loss of generality, we will consider only a stationary point coinciding with the origin, i.e. $x^* = 0$.

Theorem 3.1

If a continuous Quasi-Newton-method converges towards a point x_L, then x_L is a stationary point (i.e. $g(x_L) = 0$).

Proof

Theorem II 2.8 of Bhatia, Szegö [1] implies that x_L satisfies $x(t;x_L) = x_L$ $\forall t \in \mathbb{R}_+$ if there exists a trajectory that converges to x_L. Because of the assumed positive definiteness of $A(x)$ this means that $g(x_L) = 0$.

A slightly more general concept results from the following definition:

Definition

For any starting point x_o, the set

$LIM(x_o) := \{x_L : \exists\{t_k\} \subset \mathbb{R}, t_k \to \infty : x(t_k;x_o) \to x_L\}$

is called its (positive or omega) *limit set*.
The set $E(M) := \{x : LIM(x) \neq \emptyset \wedge LIM(x) \subset M\}$ is called the *region of attraction* of the compact set M. E(M) of this definition is equivalent to the "classical" region of attraction as the following theorem shows:

Theorem 3.2

z belongs to E(M), iff distance$(x(t;z),M) \to 0$ for $t \to \infty$.

Proof:

Bhatia, Szegö [1], Proposition V 1.2.

If a given method can be linearized

$$\dot{x} = -A(x)g(x) = Bx + o(\|x\|) \text{ for } \|x\| \to 0$$

then the eigenvalues of B can be used to characterize the region of attraction according to the following theorem:

Theorem 3.3

For $\dot{x} = -A(x)g(x) = Bx + r(x)$ with $r(x) = o(\|x\|)$ for $\|x\| \to 0$, where $\lambda_1 \leq \ldots \leq \lambda_m < 0 < \lambda_{m+1} \leq \ldots \leq \lambda_n$ are the (ordered) eigenvalues of B, the region of attraction $E(\{0\})$ is an m-parametric manifold in the $\mathbb{R}^n$.

Proof:

$r(x) = o(\|x\|)$ for $\|x\| \to 0$ implies $\partial r(x)/\partial x_i \to 0$ for $\|x\| \to 0$. Therefore all assumptions of Theorem IV 1.41 of Nemytskii, Stepanov [11] are satisfied.

One of the consequences of this theorem is for instance that the region of attraction of a saddle point is non-empty (but $m < n$).

Definition

A compact set M is said to be

(i) *invariant* iff $x_o \in M \Rightarrow x(t;x_o) \in M, \quad \forall t \geq 0$

(ii) *attractive* iff E(M) is a neighbourhood of M

(iii) *stable* iff every neighbourhood of M contains an invariant neighbourhood of M

(iv) *asymptotically stable* iff M is attractive and stable.

If the set $M = \{x^*\}$ is asymptotically stable with respect to $\dot{x} = -A(x)g(x)$ it shows all desirable properties (stability and convergence $x(t;x_o) \to x^*$ for $x_o \in E(M)$).

The asymptotic stability of a compact set M implies some remarkable facts:

<u>Theorem 3.4</u>

An asymptotically stable set M has only a finite number of components, each of which is asymptotically stable itself.

<u>Proof</u>:

Bhatia, Szegö [1] Theorem V 1.21 and V 1.22.

<u>Theorem 3.5</u>

If x^* is an asymptotically stable stationary point the set $E(\{x^*\})$ is homeomorphic to the whole space $\mathbb{R}^n$.

<u>Proof</u>:

Bhatia, Szegö [1] Theorem V 3.4.

This theorem implies for instance that $E(\{x^*\})$ is connected. A similar result holds for a discrete optimization method only under very severe restrictions.

The following result due to A. Lyapunov is the central theorem of this section:

<u>Theorem 3.6</u>

Let $V: \mathbb{R}^n \to \mathbb{R}$, $V \in C^1$ be defined in an open neighbourhood G of a compact, connected, invariant set M. Assume that

1) $V(x) = 0, \quad \forall x \in M, \quad V(x) > 0 \quad \forall x \in G - M$

2) $\dot{V} = (\text{grad } V(x), f(x)) < 0 \quad \forall x \in G - M$
   where $f(x) := -A(x)g(x)$.

Then M is asymptotically stable.

<u>Proof</u>:

see e.g., Bhatia, Szegö [1] Corollary VIII 3.10.

We have now derived conditions, which ensure asymptotic stability (especially convergence of the continuous Quasi-Newton-methods), that are based on the monotonic decline of a "generalized distance function" V

(Lyapunov-function) along solution trajectories. We will now point out some possibilities to define Lyapunov-functions for continuous Quasi-Newton-methods:

a) $V(x) = \|x - x^*\|^2$
$V = (x - x^*)^T(x - x^*)$ obviously satisfies condition 1) of Theorem 3.6 and from grad $V = 2(x - x^*)$ it follows that
$$\dot{V} = (\text{grad}\, V(x), - A(x)g(x)) = -2(x - x^*)^T A(x)g(x)$$
Condition 2) of Theorem 3.6 is therefore satisfied if we can show that the following relation holds
$$((x - x^*), A(x)g(x)) > 0.$$
Asymptotic stability (i.e. convergence) of the continuous Quasi-Newton-method is therefore guaranteed if the direction
$$d(x) = A(x)g(x)$$
of the "solution-trajectory" always has a component pointing towards x^*.

b) $V(x) = F(x) - F(x^*)$
Here we use the residual as a measure of the distance between x and x^*. Condition 1) of Theorem 3.6 is trivially satisfied in a neighbourhood of x^*, if x^* is a local minimum. Because of
$$\dot{V} = (\text{grad}\, V, - Ag) = (g, - Ag) = g^T Ag$$
the second condition of Theorem 3.6 is satisfied, as we assumed $A(x)$ to be positive definite on the region G.

c) $V(x) = \|g(x)\|^2$
$V = g(x)^T g(x)$ satisfies condition 1) of Theorem 3.6, and grad $V = 2(g')^T g$ implies
$$\dot{V} = (\text{grad}\, V, - Ag) = (2(g')^T g, - Ag) = - 2g^T(g'A)g.$$
Condition 2) of Theorem 3.6 is therefore satisfied if $g'A(x)$ is positive definite, e.g. if $g'(x)$ and $A(x)$ are both positive definite. As an example of a convergence-result, which may be deduced from the above stability considerations, we state the following theorem:

<u>Theorem 3.7</u>

If the function $F: \mathbb{R}^n \to \mathbb{R}$ satisfies the assumptions:

1) $F \in C^1$ on the whole space $\mathbb{R}^n$

2) F is bounded below

3) Every level-set of F is bounded

4) F has exactly one stationary point x^*

then the convergence of the continuous method
$$\dot{x} = - A(x)g(x), \quad x(0) = x_0$$
is guaranteed from any starting point $x_0 \in \mathbb{R}^n$, if $A(x)$ is positive defi-

nite on the whole space $\mathbb{R}^n$.

Proof:

see Ueberhuber [15].

Note, that F is *not* assumed to be convex. For the continous method of steepest descent the matrix A(x) reduces to the unit-matrix I, which ensures convergence under very general conditions.

A more detailed discussion of convergence results obtained via Lyapunov-functions for other continuous methods (e.g. for the continuous Newton-method or for the Levenberg-Marquardt-stabilized Newton-method) may be found in Ueberhuber [15].

So far we have shown that under rather weak assumptions convergence of a continuous method towards a stationary point may be expected. We will now give some results concerning the manner in which $x(t;x_o)$ tends to x^*.

Theorem 3.8

For the continuous method

$$\dot{x} = -A(x)g(x) = Bx + r(x) \text{ with } r(x) = o(\|x\|) \text{ for } \|x\| \to 0$$

with the (ordered) eigenvalues of B

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

the subset E_a of $E(\{0\})$ whose points z are characterized by

$$\|x(t;z)\| \leq c\exp((a+e(t))t) \quad \text{with } e(t) \to 0 \quad \text{as} \quad t \to \infty \quad \text{and } c \in \mathbb{R}$$

where m is defined by $\lambda_1 \leq \ldots \leq \lambda_m \leq a < \lambda_{m+1} \leq \ldots \leq \lambda_n < 0$, is an m-parametric manifold of $\mathbb{R}^n$.

Proof:

This theorem is an immediate consequence of Theorem IV 1.45 in Nemytskii, Stepanov [11].

E_a has, according to this theorem, maximum dimension only for $\lambda_n \leq a < 0$. For a starting point x_o lying in a set E_a with $a < \lambda_n$, only the exact solution $x(t;x_o)$ shows the postulated behavior; numerical solutions $x(t;x_o)$, however, soon leave the m-dimensional manifold (when $m < n$) because of roundoff- and discretization-errors. The rate of convergence for $\tilde{x}(t;x_o) \to x^*$ is therefore determined mainly by the size of the eigenvalue λ_n with smallest modulus.

An even more precise specification of the asymptotic behaviour of $x(t;x_o)$ is given in the following theorem:

Theorem 3.9

For the continuous method

$\dot{x} = -A(x)g(x) = Bx + r(x)$ with $r(x) = o(\|x\|)$ for $\|x\| \to 0$

with the (ordered) eigenvalues of B

$\lambda_1 \leq \ldots \leq \lambda_n < 0$ it holds that

1) $\exists i$: $t^{-1} \ln (\|x(t;x_o)\|) \to \lambda_i$ as $t \to \infty$

2) $\|P_{\pm} x(t;x_o)\| = o(\|Px(t;x_o)\|)$ as $t \to \infty$

where P_- denotes the projection onto the eigenspace with respect to $\lambda_1, \ldots, \lambda_{i-1}$ and P, P_+ are the respective projections onto the eigenspaces belonging to λ_i and $\lambda_{i+1}, \ldots, \lambda_n$.

Proof:

Coppel [4]. Theorem IV.5.

According to this theorem all trajectories $x(t;x_o)$ are tangential to some eigenspace of B. Combined with the result of Theorem 3.8, it follows that every numerical solution $\tilde{x}(t;x_o)$ approaches x^* in a subspace of $\mathbb{R}^n$ which is the eigenspace of B belonging to λ_n.

4. Discretization

The application of a numerical method to integrate (3.1) obviously yields an optimization method. Consider an ill-conditioned problem, i.e. a function $f(x)$ with a minimum x^* lying in a narrow valley with steep sides. Without loss of generality assume $x^* = 0$. If $f(x)$ is approximated by a quadratic $c + \frac{1}{2}(x, Qx)$, ill-conditioned problems correspond to matrices having widely separated eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0$, the condition measure cond $Q := \lambda_1/\lambda_n$ (see Rheinboldt [14]) being very large. In this case the differential equation $\dot{x} = -g(x) = -Qx$ corresponding to the quadratic approximation is said to be stiff (Dahlquist [5]), and the condition measure cond Q is called stiffness. Stiff problems characteristically have solutions with time constants that differ greatly in magnitude. To understand the difficulties arising in such a situation consider the simple model problem

$$\dot{y} = -\lambda(y - F(t)) + F'(t), \qquad \lambda \gg 0,$$

where $F(t)$ is a well behaved, smooth function (Gear [7]), with solution $y(t) = F(t) + c\, e^{-\lambda t}$ which rapidly tends to $F(t)$ (large time constant $\lambda \gg 0$, see Fig. 1).

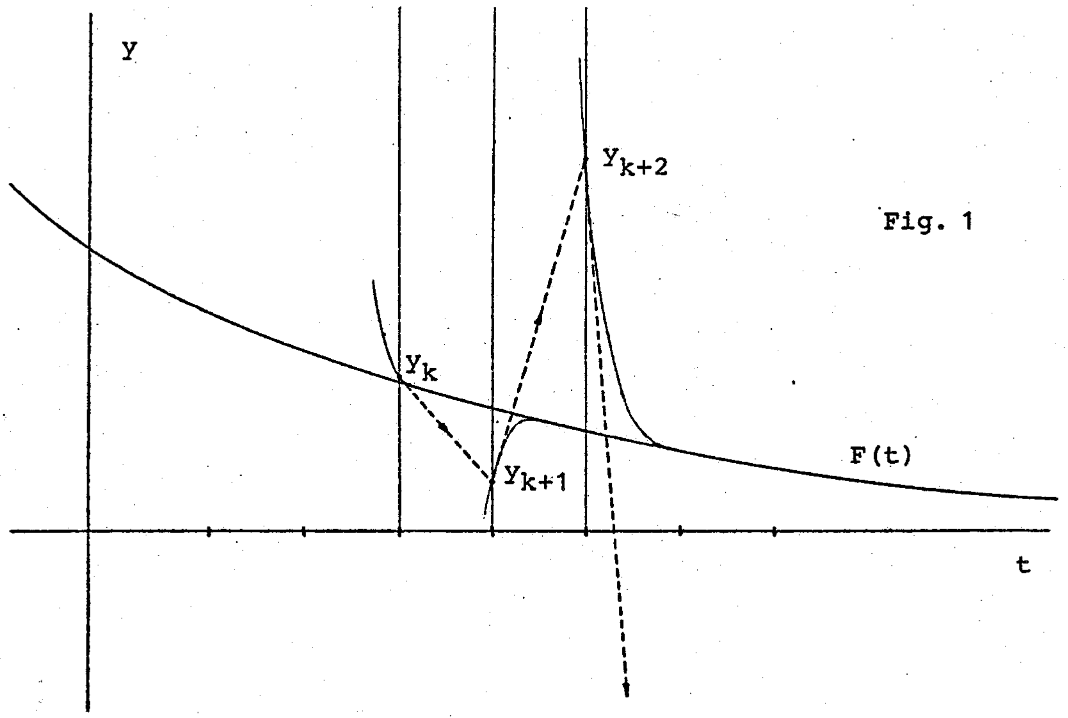


Fig. 1

Given the initial value $y(0) = F(0)$ the solution is the stationary trajectory $F(t)$ (small time constant). Truncation- and roundoff-errors in the numerical computation will induce a perturbation term $ce^{-\lambda t}$. This perturbation satisfies the differential equation

$$\dot{e} = \frac{d}{dt}(y - F) = -\lambda(y - F) = -\lambda e$$

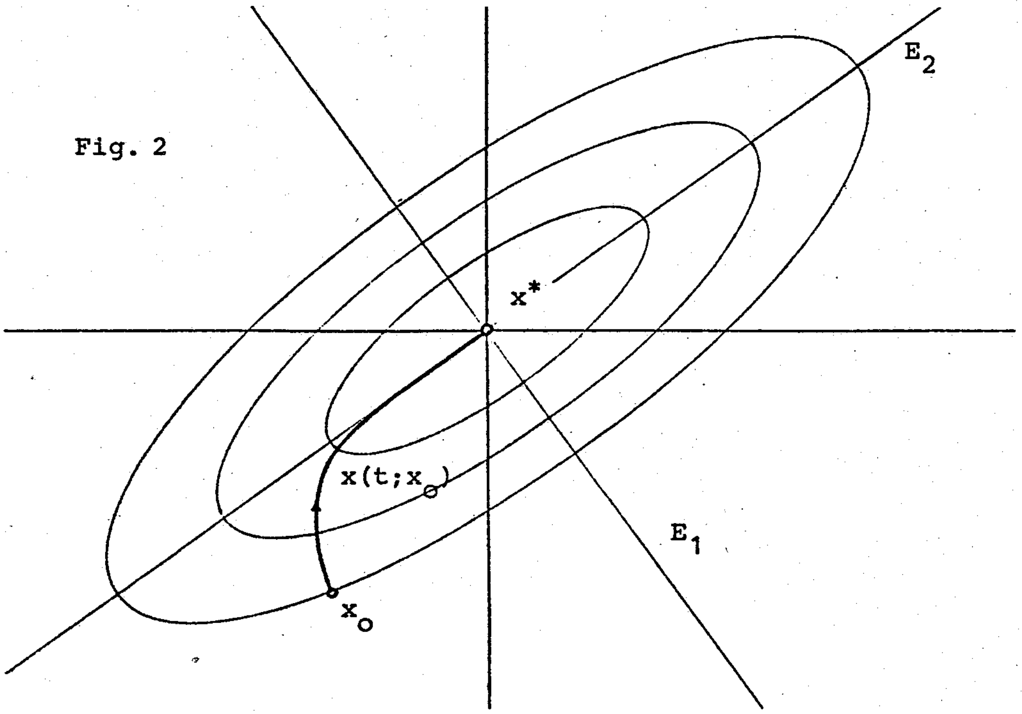
If Euler's method $y_{k+1} = y_k + h\dot{y}_k$ is applied, the perturbation term at successive steps is

$$e_{k+1} = e_k + h\dot{e}_k = (1 - h\lambda)e_k.$$

If $h\lambda > 2$, the perturbation term is unstable (Fig. 1). Despite the fact that the solution $F(t)$ we are interested in is slowly varying (the integration of $F'(t)$ would allow comparatively large step sizes), we must use very small step sizes h because λ is very large.

This is exactly the situation which is encountered when following the bottom of a steep-sided valley (Fig. 2). Here $F(t)$ corresponds to the solution lying in the eigenspace E_n belonging to the smallest eigenvalue λ_n.

Fig. 2



The integration of $\dot{x} = -Qx$ by means of the Euler method is equivalent to the application of the steepest descent method to the original optimization problem. As λ_1 is comparatively large, one can only use small step sizes. Moreover, the solution F(t) very slowly approaches the origin (because λ_2 is comparatively small), so there is a need to use large step sizes. This conflict (small step sizes to guarantee stability together with large integration intervals) is characteristic of many common explicit integration methods when applied to stiff differential equations. To overcome this dilemma one can consider the following alternatives

(i) choose an appropriate matrix-valued function A(x) and thus reduce the stiffness of $\dot{x} = -A(x)g(x) = Bx + r(x)$, cf. Section 4.1. Optimal stiffness implies cond B = 1.

(ii) choose $A(x) \equiv I$ and an integration method which is suitable for stiff differential equations, cf. Section 4.2.

4.1. Choice of the differential equation

Theorem 4.1

Newton's method is essentially the only optimal continuous Quasi-Newton-method, i.e. if

$$\dot{x} = -A(x)g(x) = Bx + r(x), \quad r(x) = O(\|x\|^2),$$

B symmetric, cond B = 1, then there exists a scalar $\mu > 0$, such, that

$$A(x)g(x) - \mu g'^{-1}(x)g(x) = O(\|x\|^2)$$

Proof:

Because of $g'^{-1}(x)g(x) = x + r_1(x)$, $r_1(x) = O(\|x\|^2)$, Newton's method is optimal (cond I = 1). Conversely, a symmetric matrix B with cond B = 1 must be of the form μI.

The result of the choice $A(x) = \mu g'^{-1}(x)$ is a differential equation without disadvantages caused by stiffness.

Theorem 4.2

$A(x)g(x)$ is assumed to be Lipschitz-continuous. Then there exists $h_o > 0$ such that the exponential stability of the solution $x(t) = 0$ of $\dot{x} = -A(x)g(x)$ implies the exponential stability of the solution $x_k = 0$ of the discretization $x_{k+1} = x_k - hA(x_k)g(x_k)$, $h \leq h_o$.

Proof:

See Falb, Groome [6]

Therefore, provided that sufficiently small stepsizes and sufficiently accurate initial values are used, discrete Newton's method converges.

4.2. Choice of the discretization

Integration methods applicable to stiff differential equations must yield stable solution sequences even if they are used with large step sizes. A common class possessing appropriate stability properties are A-stable resp. $A(\alpha)$-stable methods (cf. appendix). As the eigenvalues of the matrix B in (3.2) are real, we can restrict attention to A(0)-stability. Since we expect the function value to decrease with increasing step size (at least in the quadratic case), we demand strong A(0)-stability.

Results of Dahlquist [5] and Widlund [16] show that *explicit* Runge-Kutta- or linear multistep methods cannot be A- or A(0)-stable, thus not strongly A(0)-stable, too. From this point of view it is clear that the steepest descent method which results from the application of the explixit Euler-method cannot be an efficient technique for ill-conditioned problems. The following theorem shows the utility of strongly A(0)-stable methods.

Theorem 4.3

Let a well-defined strongly A(0)-stable Runge-Kutta method be applied with arbitrary, but constant step size $h > 0$ to

$$\dot{x} = -A(x)g(x) = Bx + r(x), \quad r(x) = O(\|x\|^2),$$

B having only real, negative eigenvalues. Then for the resulting discretization x_n it holds that

$$\lim_{n\to\infty} x_n = 0$$

provided that x_o was chosen sufficiently close to the stationary point $x^* = 0$.

Proof:

See Chipman [3] Theorem 4.2.

It should be noted that we are not interested in high order (i.e. high accuracy) methods. Accuracy is just needed in so far as the solution sesequence must remain in the domain of attraction of x^*. But high accuracy of strongly A(O)-stable methods generally implies expensive calculations, and moreover it may impede fast convergence to x^*, because the numerical solution is kept near the slowly convergent solution $x(t)$ of the continuous problem. The simplest method due to these restrictions, the implicit Euler method, is therefore studied in greater detail. Applied to differential equation (3.1) with $A(x) \equiv I$, it yields a sequence $\{x_k\}$ defined implicitly by

$$x_{k+1} = x_k - h \cdot g(x_{k+1}) \tag{4.1}$$

Theorem 4.4

$f: \mathbb{R}^n \to \mathbb{R}$ is assumed to have a bounded, nonempty level set $L(z)$ and to be $\in C^2$ in a neighbourhood U of $L(z)$. The matrix $g'(x)$ is assumed to be positive semidefinite in $L(z) \cdot x_o \in L(z)$ arbitrary, $g(x_o) \neq 0$ +).

Then it holds that

(i) the equation $y = x_o - h \cdot g(y)$ has a solution y_h for all $h \geq 0$

(ii) $0 \leq h_1 < h_2 \Rightarrow f(y_{h_2}) < f(y_{h_1})$

(iii) $0 \leq h_1 < h_2 \Rightarrow \|g(y_{h_2})\| \leq \|g(y_{h_1})\|$

(iv) $\lim_{h\to\infty} \|g(y_h)\| = 0$

Proof:

See Mitter [10]

+) If $g(x_o) = 0$ we already are on a stationary point.

Remark

Theorem 4.4 holds under weaker conditions, too. It suffices that f is twice continuously differentiable in a neighbourhood of $\{y_h | h \geq 0\}$. Moreover, $g'(y_h)$ need not be positive semidefinite, it must merely lack negative eigenvalues of too large a modulus (Mitter [10]).

The problem of finding y_h now requires the solution of the system of (generally) nonlinear equations

$$y_h = x_o - hg(y_h). \tag{4.2}$$

To estimate the minimum x* with a given tolerance, one single step with a sufficiently large step size h would suffice. Unfortunately, the condition number of equation (4.2) when solved numerically is increasing with the step size. Thus we would replace an ill-conditioned problem (the original optimization problem) by another one. By the choice of appropriate step sizes, however, we can replace the ill-conditioned optimization problem by a sequence of well-conditioned systems of nonlinear equations $x_{k+1} = x_k - h_{k+1}\, g(x_{k+1})$. From Theorem 4.4 it follows that the sequences $f(x_k)$ and $\|g(x_k)\|$ are monotonely decreasing, thus $x_k \in L(z)$ for all k. As L(z) was assumed to be bounded, the sequence x_k must have at least one cluster point.

Theorem 4.5

Assume the conditions of Theorem 4.4 to be valid and let the sequence x_k, $k \geq 1$ be defined by (4.1), using a convergent sequence of step sizes h_k with limit $h_o > 0$. Then

(i) all cluster points of the sequence x_k are stationary points

(ii) at all cluster points the function f takes on the same values, therefore $f(\tilde{x}) = \inf_{k \geq 0} f(x_k)$ for all cluster points $\tilde{x}$.

Proof:

(ii) follows from the monotonicity of $f(x_k)$, (i) is then proved by contradiction. If there is a cluster point $\tilde{x}$ such that $g(\tilde{x}) \neq 0$, then the application of Theorem 4.4 with $\tilde{x}$ instead of x_o and the limit $h_o > 0$ instead of h yields a solution y_{h_o} with $f(y_{h_o}) < f(\tilde{x})$. The application of the implicit function theorem shows that for some k with x_k and h_k sufficiently close to $\tilde{x}$ and h_o, resp., x_{k+1} is close enough to y_{h_o} such that $f(x_{k+1}) < f(x)$ which contradicts (i).

4.3. A Numerical example

Consider Rosenbrock's function

$$f(x_1,x_2) = 100\ (x_1^2 - x_2)^2 + (1 - x_1)^2$$

with starting point $x_o = (-1.2,\ 1.0)$ and the global minimum $x^* = (1,1)$. The corresponding initial value problem is given by

$$\begin{aligned} \dot{x}_1 &= -400\ (x_1^2 - x_2)\ x_1 + 2(1 - x_1) \\ \dot{x}_2 &= 200\ (x_1^2 - x_2) \\ x_1(0) &= -1.2 \\ x_2(0) &= 1.0 \end{aligned}$$

The stiffness of this differential equation (at the point x^*) is

$$\frac{\lambda_1(x^*)}{\lambda_2(x^*)} \approx 2508.$$

The method of steepest descent (equivalent to the explicit Euler method with a special stepsize control) requires stepsizes $h \approx 0.001$ to ensure convergence to x^* and is therefore extremely inefficient. The implicit Euler scheme on the other hand needs only a few steps to reach a close neighbourhood of x. Note, however, that additional gradient evaluations are needed to solve the implicit equation (4.1) at every step. The overall number of 36 gradient evaluations (including 24 gradient evaluations needed to approximate the Hessian numerically) which were necessary to obtain the approximate result (0.99998, 0.99996) in 5 implicit steps appears to be an encouraging practical result.

APPENDIX:

A-, A(α)-, A(0)-Stability

Consider a single autonomous differential equation (or a system of such equations) $\dot{y} = F(y)$ which is to be solved by a Runge-Kutta method

$$\text{(A.1)} \qquad \begin{aligned} K_1 &= F(y_n + h \sum_{i=1}^{m} \beta_{1i} K_i) \\ &\ \ \vdots \\ K_m &= F(y_n + \sum_{i=1}^{m} \beta_{mi} K_i) \\ \hline y_{n+1} &= y_n + h \sum_{i=1}^{m} c_i K_i \end{aligned}$$

For $0 < \alpha < \frac{\pi}{2}$ let W_α be the complex cone

$$\{z \in C \mid z \neq 0,\ -\alpha < \pi - \arg z < \alpha\}$$

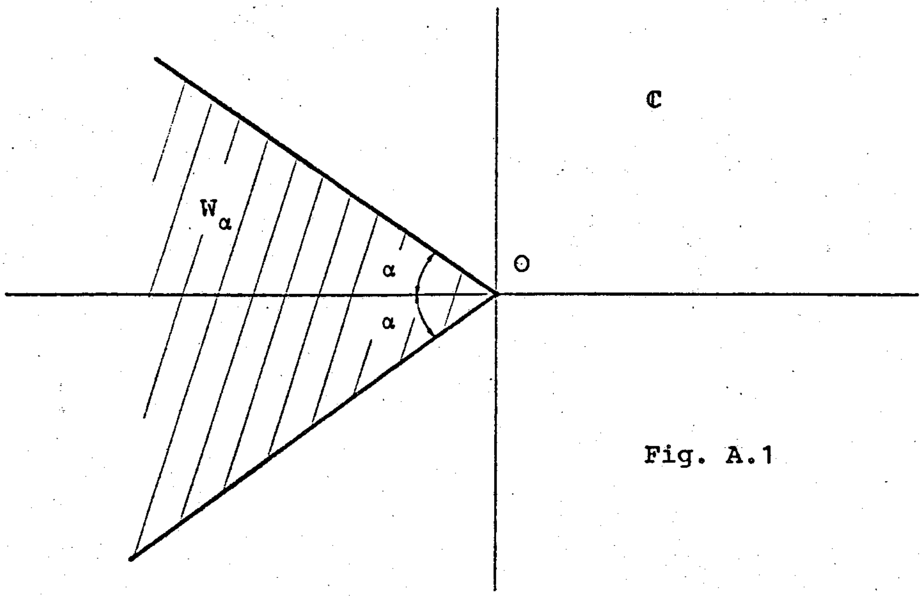


Fig. A.1

Definition

A Runge-Kutta method is *A(α)-stable*, $0 < \alpha \leq \frac{\pi}{2}$, if, when applied with arbitrary, but fixed $h > 0$ to the single equation $\dot{y} = q \cdot y$, $q \in W_\alpha$, for all initial values y_0, the corresponding solution sequence y_n converges to 0 with $n \to \infty$.

Instead of $A(\frac{\pi}{2})$-stable one usually says *A-stable*. The method is called A(0)-stable, if the statement is fulfilled for $q \in R$, $q < 0$.

Although the definition is concerned with scalar equations $\dot{y} = q \cdot y$ only, these concepts are valid for systems of linear equations, too. For example, take $\dot{y} = Ay$ with a diagonalizable matrix A. Then after transformation one has a set of decoupled scalar equations $\dot{x}_i = \lambda_i x_i$, λ_i being the eigenvalues of A. If all eigenvalues lie in W_α, we again have the situation concerning A(α)-stability.

In the case of a quadratic optimization problem $f(x) = c + \frac{1}{2}(x,Ax)$ with a symmetric, positive definite matrix A we obtain a differential equation $\dot{x} = -Ax$. Thus A(0)-stable methods seem to be convenient.

Definition

A Runge-Kutta method is called *strongly* A(α)-stable (*strongly* A-stable, *strongly* A(0)-stable, resp.), iff

(i) it is A(α)-stable (A-stable, A(0)-stable, resp.)

(ii) when applied to $\dot{y} = qy$, $q \in W_\alpha$, for all initial values y_0 it holds

$y_1 \to 0$ with $h \to \infty$ and q fixed or, equivalently
$y_1 \to 0$ with $|g| \to \infty$ and h fixed

The second condition excludes methods with undesirable properties, e.g. the trapezoidal rule which produces oscillating and very slowly decreasing approximations:

$$y_1 = \frac{1 + hq/2}{1 - hq/2} \cdot y_o$$

when applied to $\dot{y} = qy$ with $(hq) \gg 0$. In the case of a quadratic optimization problem, condition (ii) guarantees that y_1 tend towards the minimum with $h \to \infty$.

Definition

An $A(\alpha)$-stable Runge-Kutta method is called *well-defined*, if the matrix $B = (\beta_{ij})$ of the method (see (A.1)) has no eigenvalue in W_α.

If a well-defined $A(\alpha)$-stable Runge-Kutta method is applied to $\dot{y} = Ay$ whereby all eigenvalues of the matrix A lie in W_α, then the vectors K_i are uniquely given by equation (A.1).

The most common well-defined strongly A(O)-stable Runge-Kutta method is the implicit Euler method. For other methods of that type see e.g. Chipman [3].

REFERENCES

[1] N.P. Bhatia, G.P. Szegö: Stability Theory of Dynamical Systems; Springer Verlag, Berlin-Heidelberg-New York, 1970.

[2] P.T. Boggs: The Solution of Nonlinear Systems of Equations by A-stable Integration Techniques; SIAM Journal on Numerical Analysis 8(1971) 767-785.

[3] F.H. Chipman: Numerical Solution of Initial Value Problems using A-stable Runge-Kutta Processes; University of Waterloo, Dept. of Applied Analysis and Computer Science, Research Report CSRR 2042, 1971.

[4] W.A. Coppel: Stability and Asymptotic Behaviour of Differential Equations; Heath, Boston, 1965.

[5] G. Dahlquist: A Special Stability Problem for Linear Multistep Methods; BIT 3 (1963), 27-43.

[6] P.L. Falb, G.M. Groome: Stability of Difference Approximations to Differential Equations; Journal on Differential Equations 13 (1973), 48-56.

[7] C.W. Gear: Numerical Initial Value Problems in Ordinary Differential Equations; Prentice-Hall, Englewood Cliffs, 1971.

[8] W. Hahn: Über die Anwendung der Methode von Lyapunov auf Differentialgleichungen; Math. Anm. 136 (1958), 430-441.

[9] M.A. Krasnoselskii et. al.: Approximate Solution of Operator Equations; Wolter-Noordhoff, Groningen, 1972.

[10] P. Mitter: Minimization Methods based on A(O)-Stable Runge-Kutta Processes; in
W. Gröbner et. al.
The Method of Lie Series for Differential Equations and its Extensions; Techn. Report of the European Research Office, London, 1973.

[11] V.V. Nemytskii, V.V. Stepanov: Qualitative Theory of Differential Equations; Princeton University Press, 1960.

[12] J.M. Ortega: Stability of Difference Equations and Convergence of Iterative Processes; SIAM Journal on Numerical Analysis 10 (1973), 268-282.

[13] J.M. Ortega, W.C. Rheinboldt: Iterative Solution of Nonlinear Equations in Several Variables; Academic Press, New York, 1970.

[14] W.C. Rheinboldt: On Measures of Ill-Conditioning for Nonlinear Equations; Techn. Report TR-330, University of Maryland, 1974.

[15] C.W. Ueberhuber: Optimierung mittels kontinuierlicher Quasi-Newton-Verfahren und deren Diskretisierungen; Institut für Numerische Mathematik, TU Wien, Bericht Nr. 19/76, Vienna, 1976.