ESTIMATION IN NONLINEAR MODELS

Peter J.G. Teunissen
Geodetic Computing Centre
Faculty of Geodesy
Delft University of Technology
Thijsseweg 11
2629 JA Delft, The Netherlands

PREPRINT

II Hotine-Marussi Symposium
on
Mathematical Geodesy
Pisa, Italy, June 5-8, 1989

Abstract

This paper gives a survey of some results in nonlinear estimation theory which proved useful when dealing with both the numerical and statistical treatment of nonlinear geodetic adjustment problems. In the context of nonlinear optimization a number of well-known iterative algorithms belonging to the class of iterative descent methods are presented. The basic principles of these methods are discussed, necessary and sufficient conditions of convergence are given, and the rates of convergence of these methods are derived. In the context of nonlinear least-squares we present the Gauss-Newton method. In order to provide an intuitive understanding of the nature of nonlinearity, we emphasize the local or differential geometry of least-squares and in particular the role played by curvature.

Finally, in the context of densities of nonlinear estimators we discuss the impact of nonlinearity on the probabilistic properties of least-squares estimators. Special attention is given to their first moments. Also some useful and relatively easy computable measures for diagnosing the significance of nonlinearity are proposed.

Contents

Introduction

Nonlinear optimization, nonlinear least-squares and densities of nonlinear estimators are a trilogy of problems that are intimately related in the framework of nonlinear statistical inference. In this paper it is attempted to relate and unify some relevant aspects of these three important research areas. As such, the paper gives a survey of some (old and new) results in nonlinear estimation theory which proved useful when dealing with both the numerical and statistical treatment of nonlinear geodetic adjustment problems.

The numerical estimation of parameters is typically a problem of optimization. The estimation of parameters requires namely frequently the maximization or minimization of an objective function. Typical objective functions are riskfunctions, robust loss functions, posterior density functions, likelihood functions and (weighted or unweighted) sums of squares. Of these, the two most common methods of estimation are maximum likelihood and least-squares. In maximum likelihood, the estimates of the parameters are taken as those values that maximize the likelihood function given the data. Thus if $p_{\underline{y}}(y \mid x)$ is the density of the random data vector $\underline{y}$ (the underscore indicates randomness), the optimization problem of maximum likelihood reads

$$\max_x p_{\underline{y}}(y \mid x) \qquad (.1)$$

In general no *direct* methods exist for solving (1) when the parameter x enters to the third or higher power in $p_{\underline{y}}(y \mid x)$. For these cases one will therefore have to take recourse to computational techniques that are *iterative* in nature. That is, one starts with an initial guess x_0 of the solution $\hat{x}$ and then proceeds to generate according to some preassigned rule a sequence $x_1, x_2, x_3, \ldots$ that hopefully converges to $\hat{x}$. Various iterative techniques exist which can be used to solve a nonlinear optimization problem like (1). In *chapter 1* we present three of the best known iterative techniques. They are: the Steepest-Ascent (Descent) method, the Newton method and the Trust-Region method. The basic principles of these methods are discussed, necessary and sufficient conditions of convergence are given, and the rates of convergence of these methods are derived.

In most geodetic applications it is customary to assume that the random m-vector $\underline{y}$ has a multivariate normal (or Gaussian) distribution

$$p_{\underline{y}}(y \mid x) = (2\pi)^{-m/2} \mid Q_y \mid^{-1/2} exp[-\frac{1}{2} \parallel y - A(x) \parallel^2]$$

with $\parallel . \parallel^2 = (.)^* Q_y^{-1}(.)$ and $A(.) : R^n \rightarrow R^m, m \geq n$. In this case the maximization problem (1) can be turned into the minimization problem

$$\min_x \parallel y - A(x) \parallel^2 \qquad (.2)$$

This is the least-squares problem.

The minimization problem (2) can be solved directly if the map $A(.)$ is *linear*, i.e. if $A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A(x_1) + \alpha_2 A(x_2)\ \forall \alpha_1, \alpha_2 \in R, x_1, x_2 \in R^n$. The corresponding linear least-squares estimators are given by the well-known formulae

$$\begin{array}{ll} \underline{\hat{y}} = P_A \underline{y} & \underline{\hat{e}} = P_A^{\perp} \underline{y} \\ \underline{\hat{x}} = A^- P_A \underline{y} & \| \underline{\hat{e}} \|^2 = \| P_A^{\perp} \underline{y} \|^2 \end{array} \qquad (.3)$$

where: P_A is the orthogonal projector that projects onto the range of A and along its orthogonal complement; $P_A^{\perp} = I - P_A$; and A^- is an (arbitrary) inverse of A. The estimators $\underline{\hat{y}}$, $\underline{\hat{e}}$ and $\| \underline{\hat{e}} \|^2$ are unique, and the estimator $\underline{\hat{x}}$ is unique if and only if the map A has full rank n.

If the map $A(.)$ is *nonlinear* then generally no direct methods exist for solving (2). In this case one has to fall back on iterative techniques. One can in principle solve (2) with one of the iterative techniques that are presented in *chapter 1* of the paper. These methods however do not take advantage of the special structure of the objective function of (2). The Gauss-Newton method on the other hand does take advantage of the sums of squares structure of the objective function. The method is therefore especially suited for solving nonlinear least-squares problems. The Gauss-Newton method is treated in *chapter 2* of the paper. In this chapter we also introduce the differential geometric concept of normal curvature. And amongst other things, it is shown how the optimality conditions and the rate of convergence of the Gauss-Newton method can be expressed in terms of the normal curvatures of the manifold $A(x)$.

As to the distributional properties of the least-squares estimators, it is well-known that if the hypothesis $\underline{y} \sim N(A(x), Q_y)$ holds and map $A(.)$ is linear and of full rank, that

$$\begin{array}{ll} \underline{\hat{y}} \sim N(A(x), P_A Q_y) & \underline{\hat{e}} \sim N(0, P_A^{\perp} Q_y) \\ \underline{\hat{x}} \sim N(x, A^- P_A Q_y A^{-*}) & \| \underline{\hat{e}} \|^2 \sim \chi^2(m-n, 0) \end{array} \qquad (.4)$$

Unfortunately these simple results do not carry over to the nonlinear case. Essential properties which are used repeatedly in the development of the linear theory break down completely in the nonlinear case. Take for instance the mathematical expectation operator $E\{.\}$. If $\underline{\theta}$ is a random variable and $F(.)$ is a nonlinear map, then $E\{F(\underline{\theta})\} \neq F(E\{\underline{\theta}\})$, i.e. the mean of the image differs generally from the image of the mean. Hence, we can hardly expect our least-squares or maximum likelihood estimators to be unbiased in the nonlinear case.

For statistical inference purposes it is important to have means of computing the probability distribution of nonlinear estimators. Unfortunately no direct and straightforward methods are available. In practice one will therefore have to be satisfied with approximations. In *chapter 3* of the paper we discuss some ways of developing such approximations. Special attention is given to the first moments of nonlinear estimators and formulae for the biases of the least-squares estimators are derived. It will be shown what role is played by curvature and also some easily computable measures of nonlinearity are proposed.

Chapter 1

Nonlinear Optimization

1.1 Introduction

In this chapter we will consider the problem of finding (local or global) solutions to the problem:

$$\boxed{\min_x F(x) \quad , \quad x \in R^n \quad , \quad F : R^n \rightarrow R} \tag{1.1}$$

The methods that will be discussed in this chapter for solving the minimization problem (1.1) are all iterative descent algorithms. By ***iterative***, we mean, that the algorithm generates a sequence of points, each point being calculated on the basis of the points preceding it. An iterative algorithm is initiated by specifying a starting point, the initial guess. By ***descent***, we mean that as each new point is generated by the algorithm the corresponding value of $F(x)$ evaluated at the most recent point decreases in function value. Ideally, the sequence of points generated by the algorithm in this way converges in a finite or infinite number of steps to a solution of (1.1).

The methods discussed in this chapter all adhere to the following scheme:

$$\boxed{x_{k+1} = x_k + t_k d_k \quad , \quad k = 0, 1, 2, \ldots} \tag{1.2}$$

i Set $k = 0$. An initial guess is provided externally.

ii Direction generation: Determine a direction vector d_k in the direction of the proposed step.

iii Line search strategy: Determine a positive scalar t_k such that $F(x_{k+1}) \leq F(x_k)$.

iv Test wether the termination criterion is met. If so, accept x_{k+1} as the solution of (1.1). If not, increase k by one and return to step ii.

Generally one can say that the individual methods falling under (1.2) differ in their choice of the directionvector d_k and the scalar t_k. The iterative techniques fall roughly into two classes: direct search methods and gradient methods. Direct search methods are those which do not require the explicit evaluation of any partial derivatives of the function $F(x)$, but instead rely solely on values of the objective function $F(x)$, plus information gained from the earlier iterations. Gradient methods on the other hand are those which select the direction vector d_k using values of the partial derivatives of the objective function $F(x)$ with respect to the independent variables, as well as values of $F(x)$ itself, together with information gained from earlier iterations. The required derivatives, which for some methods are of order higher than the first, can be obtained either analytically or numerically using some finite difference scheme. This latter approach necessitates extra function evaluations close to the current point x_k, and effectively reduces a gradient method to one of direct search.

In this chapter we will restrict ourselves to gradient methods for which the required derivatives can be obtained analytically. The descent methods that will be discussed are: the Steepest Descent method, Newton's method and the Trust Region method. But before discussing these methods we first develop the conditions that must hold at a solution point of (1.1). These conditions are derived in the next section and they are simple extensions of the well-known derivative conditions for a function of a single variable that hold at a maximum or a minimum point.

1.2 Optimality conditions

In the investigation of the minimization problem (1.1) we distinguish two kinds of solution points: local minimizers and global minimizers.

Definition: The vector $\hat{x} \in R^n$ is said to be a *global minimum* of $F(x)$ if $F(\hat{x}) \leq F(x), \forall x \in R^n$. The global minimum is *unique* if $F(\hat{x}) < F(x), \forall x \in R^n$.

Definition: The vector $\hat{x} \in R^n$ is said to be a *local minimum* of $F(x)$ if $F(\hat{x}) \leq F(x)$ for all x near $\hat{x}$. By "x near $\hat{x}$" we mean that an ϵ-ball, $B(\hat{x},\epsilon)$, of the point $\hat{x}$ exists such that $x \in B(\hat{x},\epsilon)$. The ϵ-ball is defined as $B(\hat{x},\epsilon) = \{x \mid \| x - \hat{x} \| < \epsilon, x \in R^n\}$. The local minimum is said to be *isolated* if $F(\hat{x}) < F(x), \forall x \in B(\hat{x},\epsilon)$.

The problem of computing the minimum of $F(x)$ can be facilitated by deriving certain properties that must be satisfied by the minimizing vector $\hat{x}$. The following two theorems, theorem 1 and 2, state necessary and sufficient conditions for $\hat{x}$ to be a minimum of $F(x)$.

Theorem 1 (necessary conditions):
Assume that $F(x)$, $\partial_x F(x)$ and $\partial^2_{xx} F(x)$ are continuous $\forall x \in R^n$. If $\hat{x}$ is a (local or global) minimum of $F(x)$, then

$$\boxed{\begin{array}{ll} a) & \partial_x F(\hat{x}) = 0 \\ \\ b) & \partial^2_{xx} F(\hat{x}) \geq 0 \end{array}} \qquad (1.3)$$

proof:
First we shall proof (1.3a). Consider the vector $x = \hat{x} + td$ where t is a scalar and d an n-vector. Expansion of $F(x)$ in a Taylorseries at $\hat{x}$ gives

$$F(x) = F(\hat{x}) + t\partial_x F(\hat{x})^* d + O(t) \qquad (1.4)$$

The orderterm $O(t)$ indicates the remainder in the Taylorseries. It has the property

$$\lim_{t \to 0} \frac{O(t)}{t} = 0 \qquad (1.5)$$

Since $\hat{x}$ is a local minimum of $F(x)$ by assumption (a global minimum is ofcourse also a local minimum), it follows that for sufficiently small t,

$$F(\hat{x}) \leq F(x) = F(\hat{x} + td) \qquad (1.6)$$

This, together with (1.4) gives after dividing by t,

$$\begin{array}{l} \partial_x F(\hat{x})^* d + \frac{O(t)}{t} \geq 0 \;\; for\; t > 0 \\ \\ \partial_x F(\hat{x})^* d + \frac{O(t)}{t} \leq 0 \;\; for\; t < 0 \end{array} \qquad (1.7)$$

Taking the limit as $t \to 0$ and using (1.5) shows that (1.3a) must be true.

Next we shall proof (1.3b). Expanding $F(x)$ in a Taylorseries at $\hat{x}$, but now retaining the quadratic terms, gives with (1.3a),

$$F(x) = F(\hat{x}) + \frac{1}{2}t^2 d^* \partial^2_{xx} F(\hat{x}) d + O(t^2) \tag{1.8}$$

where $O(t^2)$ has the property,

$$\lim_{t \to 0} \frac{O(t^2)}{t^2} = 0 \tag{1.9}$$

Equations (1.6) and (1.8) imply that

$$\frac{1}{2} d^* \partial^2_{xx} F(\hat{x}) d + \frac{O(t^2)}{t^2} \geq 0$$

Taking the limit as $t \to 0$ and using (1.9) shows that

$$\frac{1}{2} d^* \partial^2_{xx} F(\hat{x}) d \geq 0$$

Since d is completely arbitrary, this means by definition that $\partial^2_{xx} F(\hat{x})$ is positive semi-definite. This proofs (1.3b). □

Theorem 1 gives necessary conditions for $\hat{x}$ to be a minimum of $F(x)$. The stated conditions are however not sufficient. This is easily illustrated by the following simple example. Consider the function $F_1(x) = x^4$. Clearly it has a (global) minimum at $\hat{x} = 0$, and $\partial_x F_1(0) = 0$ and $\partial^2_{xx} F_1(0) = 0$ hold. Consider now the function $F_2(x) = -x^4$. Then once again $\partial_x F_2(0) = 0$ and $\partial^2_{xx} F_2(0) = 0$ hold. But now, however, $\hat{x} = 0$ is a (global) maximum of $F_2(x)$.

The following theorem gives sufficient conditions for $\hat{x}$ to be a minimum of $F(x)$.

Theorem 2 (sufficient conditions):

Assume that $F(x)$, $\partial_x F(x)$ and $\partial^2_{xx} F(x)$ are continuous $\forall x \in R^n$. If

$$\boxed{\begin{array}{ll} a) & \partial_x F(\hat{x}) = 0 \\ b) & \partial^2_{xx} F(\hat{x}) > 0 \end{array}} \tag{1.10}$$

then $\hat{x}$ is a (local or global) minimum of $F(x)$.

proof:

Expansion of $F(x)$, with $x = \hat{x} + td$, at $\hat{x}$ gives with (1.10a),

$$F(x) = F(\hat{x}) + [\frac{1}{2} d^* \partial^2_{xx} F(\hat{x}) d + \frac{O(t^2)}{t^2}] t^2 \tag{1.11}$$

Since $\partial^2_{xx} F(\hat{x})$ is positive definite by assumption, the first term in the backet on the righthandside of (1.11) is strictly positive. Hence, in view of (1.9), we can conclude that the bracketed term is strictly positive for sufficiently small t. Thus from (1.11) follows that $F(x) > F(\hat{x})$ for all x near $\hat{x}$, that is, for t small. □

It should be noted that $\hat{x}$ can be a minimum of $F(x)$ and still violate the sufficiency conditions of theorem 2 (for example $F(x) = x^4$). It is also remarked, since finding the maximum of a function $F(x)$ is equivalent to finding the minimum of $-F(x)$, that the necessary and sufficient conditions for a maximum simply follow from changing the inequality signs "$\geq$" in (1.3) and "$>$" in (1.10) into "$\leq$" and "$<$" respectively. Those cases where $\partial_x F(\hat{x}) = 0$, and $\hat{x}$ is neither a minimum nor a maximum correspond to inflection and saddle points of $F(x)$. Thus we may conclude that if we find all the solutions of $\partial_x F(x) = 0$, the so-called *stationary- or critical points* of $F(x)$, then we shall have found all the minima (local and global), maxima (local and global), inflection and saddle points of $F(x)$.

If $F(x)$ is quadratic, then $\partial_x F(x)$ is linear and the stationary points simply follow from solving a system of linear equations. In het non-quadratic case, however, $\partial_x F(x)$ is nonlinear and a system of nonlinear equations, $\partial_x F(x) = 0$, has to be solved. Due to the nonlinearity of the system $\partial_x F(x) = 0$, it is very seldom that one can find analytical expressions for its solutions (there are exceptions!). In practice one will therefore have to recourse to methods which are iterative in nature.

Once the stationary points of $F(x)$ are found, one can check whether $\partial^2_{xx} F(x) > 0$ holds in order to show that the corresponding stationary point must be a local minimum. If $\partial^2_{xx} F(x) \geq 0$, one first computes the vectors d that satisfy $\partial^2_{xx} F(\hat{x}) d = 0$ and then checks whether $F(x) \leq F(\hat{x})$ for all those x that lie on the ray that emanates from $\hat{x}$ with direction vector d. After the local minima are found, the global minima follow from comparison of the function values at the local minima.

Since we are interested in locating minima of the function $F(x)$ it seems intuitively appealing to restrict our attention to those iterative methods that impose the descent condition. In the next section we will give a representation of the class of vectors that lie in a descent direction of the objective function $F(x)$.

1.3 Descent direction generation

The direction vector d_k of

$$x_{k+1} = x_k + t_k d_k \tag{1.12}$$

is said to be in a *descent direction* if a possitive scalar t_k exists such that

$$F(x_k + t_k d_k) < F(x_k) \tag{1.13}$$

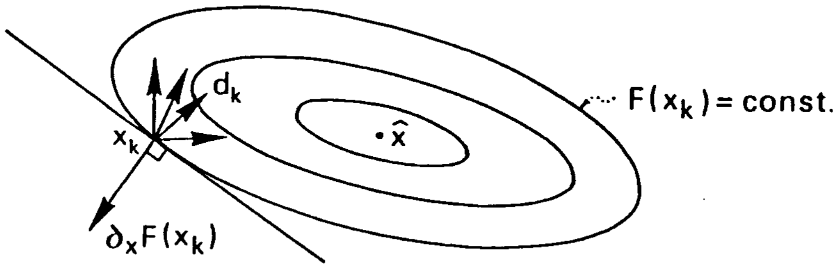


Figure 1.1: Contours of $F(x)$ and descent directions at x_k.

If we apply Taylor's expansion to $F(x_k + t_k d_k)$ at x_k we get

$$F(x_k + t_k d_k) = F(x_k) + t_k \partial_x F(x_k)^* d_k + O(t_k)$$

This shows that if

$$\boxed{\partial_x F(x_k)^* d_k < 0} \tag{1.14}$$

then it is possible to choose a positive scalar t_k so that (1.13) holds. Direction vectors d_k that satisfy inequality (1.14) are thus vectors that lie in the direction of descent. The various descent directions at x_k of the function $F(x)$ are shown in figure 1.1.

It follows form inequality (1.14) that the descent direction vectors can be represented as

$$\boxed{d_k = -Q(x_k)\partial_x F(x_k)} \tag{1.15}$$

where $Q(x_k)$ is an arbitrary but positive-definite matrix that depends on x_k. In the following sections we will see that each particular iterative descent method can be characterized by the choice made for matrix $Q(x_k)$. But before we discuss the different iterative descent methods we will first establish sufficiency conditions that guarantee convergence of the descent methods to a stationary point of $F(x)$. This is done in the next section.

1.4 A convergence theorem

It follows from (1.12) and (1.15) that the descent methods take the form

$$\boxed{x_{k+1} = x_k - t_k Q(x_k)\partial_x F(x_k)} \tag{1.16}$$

The two variables in (1.16) are the positive scalar t_k and the positive-definite matrix $Q(x_k)$. Different choices for t_k and $Q(x_k)$ correspond with different descent algorithms.

If we define a vectorfunction $\Phi : R^n \to R^n$ as

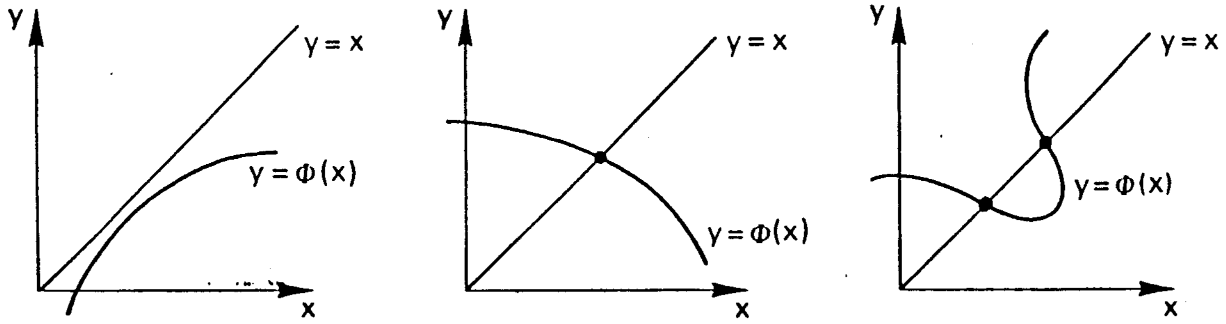
$$\Phi(x) = x - t(x)Q(x)\partial_x F(x) \tag{1.17}$$

equation (1.16) can be written in the compact form

$$\boxed{x_{k+1} = \Phi(x_k)} \tag{1.18}$$

Note that since $t(x)$ is positive and $Q(x)$ is positive-definite, the solutions of $x = \Phi(x)$, the socalled *fixed points* of $\Phi(x)$, are identical to the solutions of $\partial_x F(x) = 0$, i.e. the stationary points of $F(x)$. This implies that if the sequence generated by (1.18) converges to a fixed point of $\Phi(x)$, the sequence generated by the descent method (1.16) will converge to a stationary point of $F(x)$.

The iterationscheme (1.18) is known as the *fixed point iteration method*. It is sometimes also called the method of successive approximation and also Picard's method. The geometry of the fixed point iteration method is best illustrated for the univariate or scalar case. In figure 1.2 the three possible cases of no solution, a unique solution and a non-unique solution are shown.



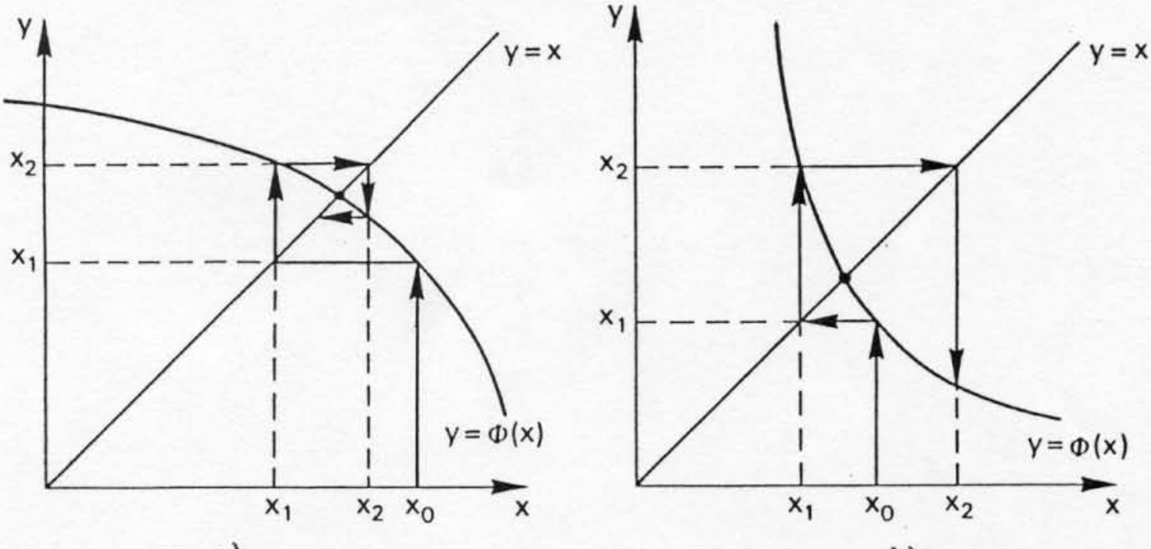
a)
no solution

b)
a unique solution

c)
a non-unique solution

Figure 1.2: Existence and uniqueness of $x = \Phi(x)$.

Figure 1.3 shows how the sequence $x_{k+1} = \Phi(x_k)$, $k = 0, 1, 2, \ldots$, is constructed geometrically. Both a convergent and a divergent case are shown.



a)
convergence

b)
divergence

Figure 1.3: The sequence $x_{k+1} = \Phi(x_k)$, $k = 0, 1, 2, \ldots$

The following theorem gives sufficient conditions for the fixed point method to convergence to the unique solution of $x = \Phi(x)$.

Theorem 3 (Fixed point iteration):
Let Ω be a set of $R^n, \Omega \subset R^n$.
Assume that

i $\Phi(x) \in \Omega$, $\forall x \in \Omega$

ii $\Phi(x)$ is continuous $\forall x \in \Omega$

iii $\| \Phi(x_2) - \Phi(x_1) \| \leq c \| x_2 - x_1 \|$, $0 \leq c < 1$, $\forall x_1, x_2 \in \Omega$

Then:

1. A solution $\hat{x}$ of $x = \Phi(x)$ exists in Ω
2. The solution $\hat{x}$ is unique
3. The fixed point algorithm converges to $\hat{x}$, that is $\lim_{k\to\infty} x_k = \hat{x}$

proof:
Assumption i guarantees with (1.18) that if $x_0 \in \Omega$, then $x_1 = \Phi(x_0) \in \Omega$ and so on. Thus every member of the sequence $x_0, x_1, x_2, \ldots$ remains in the set Ω.

We will now show that the sequence converges to a limit. With assumption iii and (1.18) follows that $\| x_{k+1} - x_k \| = \| \Phi(x_k) - \Phi(x_{k-1}) \| \leq c \| x_k - x_{k-1} \|$. Therefore,

$$\| x_{k+1} - x_k \| \leq c^k \| x_1 - x_0 \|$$

From this follows that

$$\begin{aligned}\| x_{k+p} - x_k \| &= \| x_{k+p} - x_{k+p-1} + x_{k+p-1} - x_{k+p-2} + \cdots + x_{k+1} - x_k \| \\ &\leq \| x_{k+p} - x_{k+p-1} \| + \| x_{k+p-1} - x_{k+p-2} \| + \cdots + \| x_{k+1} - x_k \| \\ &\leq [c^{k+p-1} + c^{k+p-2} + \cdots + c^k] \| x_1 - x_0 \| \\ &\leq [c^k \textstyle\sum_{i=0}^{p-1} c^i] \| x_1 - x_0 \| \\ &\leq [c^k \textstyle\sum_{i=0}^{\infty} c^i] \| x_1 - x_0 \| \\ &\leq \tfrac{c^k}{1-c} \| x_1 - x_0 \|\end{aligned}$$

since $(1-c)^{-1} = \sum_{i=o}^{\infty} c^i$ if $0 \leq c < 1$.

Therefore $\lim_{k\to\infty} \| x_{k+p} - x_k \| = 0$ for any p. This implies that for every $\epsilon > 0$ there exists a positive integer N such that $\| x_m - x_k \| < \epsilon$ for all $k, m \geq N$. A sequence with such a property is called a *Cauchy sequence*. Since a Cauchy sequence in R^n is also a convergent sequence it follows that $\lim_{k\to\infty} x_k = \hat{x} \in \Omega$.

To show that the limitpoint $\hat{x}$ is indeed a solution of $x = \Phi(x)$, note that $\hat{x} = \lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty} \Phi(x_k)$. By the continuity assumption of $\Phi(x)$ we have $\lim_{k\to\infty} \Phi(x_k) = \Phi(\hat{x})$ and thus $\hat{x} = \Phi(\hat{x})$. What remains to be shown is that $\hat{x}$ is unique. Let $\hat{x}_1$ and $\hat{x}_2$ be two fixed points of $\Phi(x)$ in Ω. Then $\| \hat{x}_1 - \hat{x}_2 \| = \| \Phi(\hat{x}_1) - \Phi(\hat{x}_2) \| \leq c \| \hat{x}_1 - \hat{x}_2 \|$ and thus $\hat{x}_1 = \hat{x}_2$. This concludes the proof of theorem 3. □

Although theorem 3 gives sufficiency conditions for the guaranteed convergence of the sequence $x_{k+1} = \Phi(x_k), k = 0, 1, 2, 3, \ldots$, to a unique fixed point, its usefulness in practical applications is unfortunately rather limited. This is due to the difficulty one has in practical applications with verifying the sufficiency conditions. Especially the verification of the inequality condition iii for all pairs of vectors in Ω is most difficult. This task becomes somewhat simpler if we may assume that $\Phi(x)$ has continuous partial derivatives and that Ω is convex. With the *mean value theorem* follows than that

$$\| \Phi(x_2) - \Phi(x_1) \| = \| \partial_x \Phi(\bar{x})(x_2 - x_1) \| \leq \| \partial_x \Phi(\bar{x}) \| \| x_2 - x_1 \|$$

with $\bar{x} = x_1 + t(x_2 - x_1), 0 \leq t \leq 1$. This result implies that we may check condition iii of the theorem by verifying whether

$$c = \max_{x \in \Omega} \| \partial_x \Phi(x) \| < 1 \tag{1.19}$$

With this result we are now also able to formulate more tractable convergency conditions for the class of descent methods (1.16). By taking the partial derivatives of (1.17) we get

$$\partial_x \Phi(x) = I - \sum_{\alpha=1}^{n} \partial_x q_\alpha(x) \partial_\alpha F(x) - t(x) Q(x) \partial^2_{xx} F(x) \tag{1.20}$$

where $q_\alpha(x), \alpha = 1, 2, \ldots, n$, are the columnvectors of the positive-definite matrix $t(x)Q(x)$. Hence,

$$\| \partial_x \Phi(x) \| \leq \| I - t(x) Q(x) \partial^2_{xx} F(x) \| + \sum_{\alpha=1}^{n} \| \partial_x q_\alpha(x) \| | \partial_\alpha F(x) | \tag{1.21}$$

This shows that convergence of the descent methods is guaranteed if

$$\boxed{\| I - t(x)Q(x)\partial^2_{xx}F(x) \| < 1} \tag{1.22}$$

and if the second term on the right hand side of (1.21) can be made sufficiently small. Since $\partial_x F(\hat{x}) = 0$ and $\partial_x F(x)$ is continuous, then by the very definition of continuity for each $\epsilon > 0$ there exists a $\delta > 0$ such that if $\| x - \hat{x} \| < \delta$, then $\| \partial_x F(x) - \partial_x F(\hat{x}) \| = \| \partial_x F(x) \| < \epsilon$. This implies that the second term on the right hand side of (1.21) can be made sufficiently small for a sufficiently small neighborhood of $\hat{x}$. Thus convergence of the descent methods is guaranteed if (1.22) holds and if the initial guess is *sufficiently close* to the solution $\hat{x}$. The practical problem with the above proof of guaranteed convergence is ofcourse still that one never knows beforehand whether the initial guess is indeed sufficiently close to $\hat{x}$. Nevertheless the above derivation shows clearly what the cause for a possible lack of convergence can be. And it also shows, see (1.22), how convergence can be enforced by a suitable choice for the scalar $t(x)$.

1.5 The Steepest Descent Method

The steepest descent method is one of the oldest iterative descent methods for solving a minimization problem. The method goes back to Cauchy (1847). The steepest descent method is characterized by the following simple choice for the positive-definite matrix $Q(x_k)$ of (1.16):

$$Q(x_k) = I \tag{1.23}$$

The steepest descent method takes therefore the form

$$\boxed{x_{k+1} = x_k - t_k \partial_x F(x)} \tag{1.24}$$

The choice (1.24) is motivated by the fact that the vector $d_k = -\partial_x F(x_k)$ minimizes

$$\frac{\partial_x F(x_k)^* d_k}{(d_k^* d_k)^{\frac{1}{2}}}$$

Thus within a linear approximation, the direction vector $d_k = -\partial_x F(x_k)$ points in the direction of the steepest descent of the function $F(x)$ at x_k.

One of the advantages of the steepest descent methods is its great simplicity. No partial derivatives of $F(x)$ of the order higher than the first are needed and no matrices need to be inverted. A drawback of the method is however that its performance is dependent on the more or less arbitrary choice of the variables x used to define the minimization problem. This can be seen as follows.

Suppose that R is an invertible $n \times n$ matrix. We can represent points in R^n either by the standard vector x or by $\bar{x}$ where $R\bar{x} = x$. The problem of finding x to minimize $F(x)$ is equivalent to that of finding $\bar{x}$ to minimize $G(\bar{x}) = F(R\bar{x})$. Thus using steepest descent, the direction vector in case of minimizing $G(\bar{x})$ will be $\bar{d}_k = -R^* \partial_x F(R\bar{x}_k)$ which in the original variables is $d_k = -RR^* \partial_x F(x_k)$. Thus, we see that if $RR^* \neq I$ the change of variables changes the direction of a search. Hence, a new choice of variables may substantially alter the performance characteristics of the steepest descent method.

Another drawback of the steepest descent method is that it has the tendency to *zig-zag*, when it is combined with an *exact line search strategy* and the contours of the objective function are elongated. An exact line search strategy is a strategy in which the positive scalar t_k is chosen so as to minimize $F(x_k + t_k d_k)$. If t_k is a minimizer of $F(x_k + t_k d_k)$ then

$$0 = \frac{dF}{dt}(t_k) = \partial_x F(x_k + t_k d_k)^* d_k = \partial_x F(x_{k+1})^* d_k$$

This shows that if an exact line search is used, the successive directions of search, d_k and d_{k+1}, are orthogonal to each other. Hence the steepest descent method will obviously zig-zag when the contours of $F(x)$ are very elongated. (see figure 1.4) The zig-zagging is absent ofcourse when the contours of $F(x)$ are circular. In fact, the steepest descent method with an exact line search will locate the minimum of $F(x)$ in one step if the contours of $F(x)$ are circles (or hyperspheres).

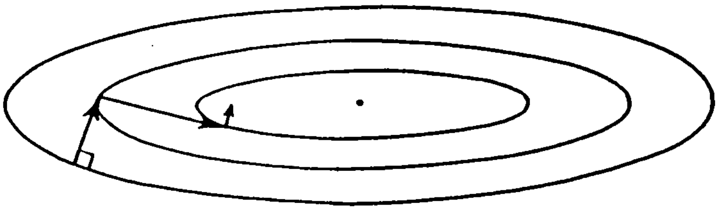


Figure 1.4: Zig-zagging of the steepest descent method.

An important performance measure of an iteration method is its *rate of convergence*. The rate of convergence of an iterative technique is related to the way the errormagnitude at the $(k+1)$th step, $\| x_{k+1} - \hat{x} \|$, is related to the errormagnitude in the previous step. The rate of convergence shows therefore whether convergence of an iteration method is rapid enough to make the whole scheme practical.

In order to derive the rate of convergence for the steepest descent method we expand (1.24) into a Taylorseries at the solution $\hat{x}$. This gives

$$x_{k+1} - \hat{x} = [I - t_k \partial^2_{xx} F(\hat{x})](x_k - \hat{x}) + O(\| x_k - \hat{x} \|) \tag{1.25}$$

If $\hat{x}$ is a local minimizer of $F(x)$ then the matrix $\partial^2_{xx} F(\hat{x})$ is positive semi-definite and its eigenvalues may be ordered so that

$$0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$$

By taking the norm of (1.25) we therefore get

$$\| x_{k+1} - \hat{x} \| \leq max\{ | 1 - t_k \lambda_1 | , | 1 - t_k \lambda_n | \} \| x_k - \hat{x} \| + O(\| x_k - \hat{x} \|) \tag{1.26}$$

This shows that the steepest descent method has a *linear rate of convergence* for points sufficiently close to the solution. Thus for points sufficiently close to the solution the errormagnitude gets reduced by a factor $max\{| 1 - t_k \lambda_1 | , | 1 - t_k \lambda_n |\}$ at each iterationstep. The closer this factor is to 1 the slower the rate of convergence; the closer the factor is to 0 the faster the rate of convergence.

If the positive scalar t_k is taken to be equal to one in each iterationstep (this is the simplest line search strategy), the rate of convergence of the steepest descent method becomes approximately

$$\boxed{\| x_{k+1} - \hat{x} \| \leq max\{ | 1 - \lambda_1 | , | 1 - \lambda_n | \} \| x_k - \hat{x} \|} \tag{1.27}$$

This shows that the errormagnitude gets reduced if the extreme eigenvalues of $\partial^2_{xx}F(\hat{x})$ satisfy

$$0 < \lambda_1, \lambda_n < 2 \tag{1.28}$$

Hence, see also (1.22), local convergence cannot be guaranteed if one or more eigenvalues of the positive semi-definite matrix $\partial^2_{xx}F(\hat{x})$ lie outside the open interval (0,2).

The rate of convergence of (1.27) can be improved and local convergence can be guaranteed, however, if the positive scalar t_k is chosen so as to minimize $max\{\mid 1 - t_k\lambda_1 \mid, \mid 1 - t_k\lambda_n \mid\}$. It follows from figure 1.5 that the corresponding optimal choice for t_k is

$$t_k = 2/(\lambda_1 + \lambda_n) \tag{1.29}$$

With this choice for t_k it follows from (1.26) that instead of (1.27) we have

$$\boxed{\| x_{k+1} - \hat{x} \| \leq \left[\frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1}\right] \| x_k - \hat{x} \|} \tag{1.30}$$

Note that if the matrix $\partial^2_{xx}F(\hat{x})$ is positive definite then the factor $(\lambda_n - \lambda_1)/(\lambda_n + \lambda_1)$ is always less than one and local convergence is guaranteed. This factor is close to one if the conditionnumber, λ_n/λ_1, of the matrix $\partial^2_{xx}F(\hat{x})$ is large, i.e. if the contours of $F(x)$ are very elongated near the solution $\hat{x}$.

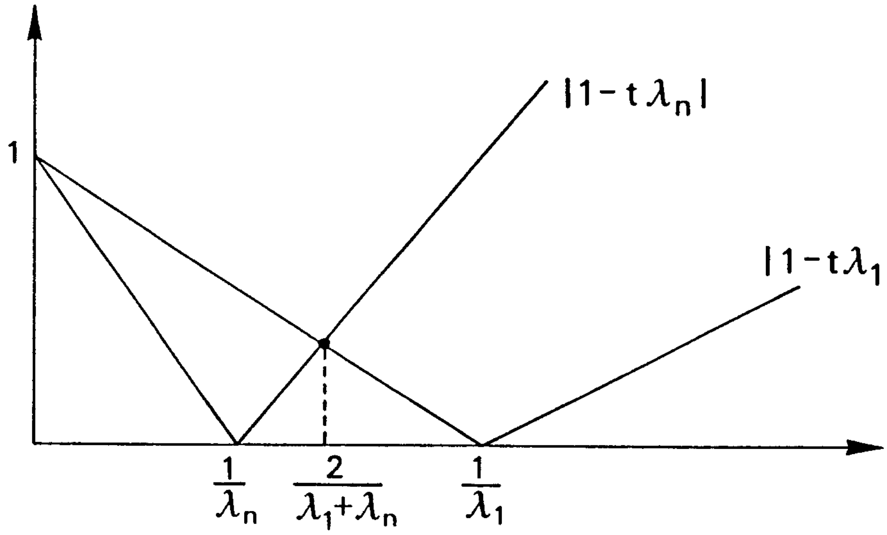


Figure 1.5: Optimal choice for t is $t = 2/(\lambda_1 + \lambda_n)$.

1.6 Newton's Method

Newton's method is characterized by the following choice for the positive-definite matrix $t_kQ(x_k)$ of (1.16):

$$t_kQ(x_k) = [\partial^2_{xx}F(x)]^{-1} \tag{1.31}$$

Newton's method takes therefore the form

$$\boxed{x_{k+1} = x_k - [\partial^2_{xx}F(x_k)]^{-1}\partial_x F(x_k)} \tag{1.32}$$

We will give two motivations for the choice (1.31). The first one goes back to the basic idea on the basis of which Newton's method was originally introduced. Newton's method was originally conceived as an iterative technique for solving a system of nonlinear equations. The basic idea of the method is best explained for a function $G(x)$ with one variable x. Let the nonlinear equation which needs to be solved be

$$G(x) = 0 \tag{1.33}$$

The graph of this function is plotted in figure 1.6.

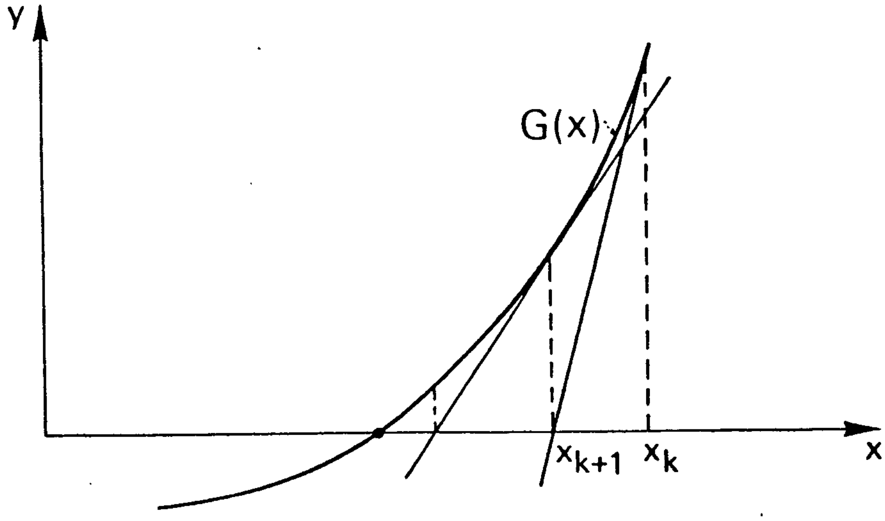


Figure 1.6: Newton's method: $\frac{dG}{dx}(x_k) = G(x_k)/(x_k - x_{k+1})$.

At a given point x_k the graph of the function $G(x)$ is approximated by its tangent, and an approximate solution to equation (1.33) is taken to be the point x_{k+1} where the tangent crosses the x-axis. The process is then repeated from this new point. This procedure defines a sequence of points according to the recurrence relation

$$x_{k+1} = x_k - [d_x G(x_k)]^{-1} G(x_k) \tag{1.34}$$

If we replace $G(x)$ in (1.33) by $d_x F(x)$, the recurrence relation for determining a stationary point of $F(x)$ becomes

$$x_{k+1} = x_k - [d^2_{xx} F(x)]^{-1} d_x F(x) \tag{1.35}$$

In order to generalize this result to the multivariate case, note that in the above procedure the original nonlinear equation, $G(x) = 0$ or $d_x F(x) = 0$, is linearized about the point x_k and then solved for x_{k+1}. If we apply this procedure to the system of nonlinear equations $\partial_x F(x) = 0$, linearization gives

$$0 = \partial_x F(x_k) + \partial^2_{xx} F(x_k)(x_{k+1} - x_k) \tag{1.36}$$

from which x_{k+1} follows as (1.32).

Since Newton's method is based on a linearization of $\partial_x F(x)$, one can interpret the method as one that computes the minimum of a *quadratic approximation* of $F(x)$ at each iteration step. This shows the distinct difference with the steepest descent method. The steepest descent method is namely based on a *linear approximation* of $F(x)$ at each iteration step. As a consequence, if the function $F(x)$ is quadratic, Newton's method will locate the minimum in one iterationstep, whereas the steepest descent method needs in general an infinite number of iteration steps.

The second motivation for the choice (1.31) is based on inequality (1.22). Note that with (1.31) inequality (1.22) is trivially fulfilled, which implies that Newton's method has a guaranteed convergence for points sufficiently close to the solution. Thus, contrary to the steepest descent method no line search is needed to enforce local convergence.

In order to derive the rate of convergence for Newton's method we expand (1.32) into a Taylorseries at the solution $\hat{x}$. This gives

$$x_{k+1} - \hat{x} = -\frac{1}{2}(x_k - \hat{x})^*[\partial_x[\partial^2_{xx}F(\hat{x})]^{-1}\partial^2_{xx}F(\hat{x})](x_k - \hat{x}) + O(\| x_k - \hat{x} \|^2)$$

or with $[\partial^2_{xx}F(\hat{x})]^{-1}\partial^3_{xxx}F(\hat{x}) = -\partial_x[\partial^2_{xx}F(\hat{x})]^{-1}\partial^2_{xx}F(\hat{x})$,

$$\boxed{x_{k+1} - \hat{x} = \frac{1}{2}(x_k - \hat{x})^* \left[[\partial^2_{xx}F(\hat{x})]^{-1}\partial^3_{xxx}F(\hat{x})\right] (x_k - \hat{x}) + O(\| x_k - \hat{x} \|^2)} \tag{1.37}$$

This shows that Newton's method has a *quadratic* rate of convergence.

Although the information requirements associated with the evaluation, storage and inversion of the matrix $\partial^2_{xx}F(x)$ as required by Newton's method are rather heavy, the method has proved, due to its guaranteed local convergence and quadratic rate of convergence, to be extremely effective in dealing with general minimization problems. Difficulties with Newton's method occur however when the matrix $\partial^2_{xx}F(x)$ is non-invertible or when it fails to be positive definite. These difficulties can be overcome by using a so-called trust region method. This method, which can be considered as a regularized version of Newton's method, will be discussed in the next section.

1.7 The Trust Region Method

The trust region method was introduced by Levenberg (1944), reinvented by Marquardt (1963) and further developed by Goldfeld, Quandt and Trotter (1966). The method is characterized by the following choice for the positive definite matrix $t_kQ(x_k)$ of (1.16):

$$t_kQ(x_k) = [\partial^2_{xx}F(x_k) + \alpha_k R]^{-1} \tag{1.38}$$

where α_k is a non-negative scalar and R is a positive definite matrix. The trust region method takes therefore the form

$$\boxed{x_{k+1}(\alpha_k) = x_k - [\partial^2_{xx}F(x_k) + \alpha_k R]^{-1}\partial_x F(x_k)} \tag{1.39}$$

This formula already shows some of the basic ideas underlying the trust region method. Since matrix R is positive definite by assumption, a sufficiently large α_k ensures the positiveness of (1.38). Thus by adjusting α_k, a possible lack of positive definiteness of $\partial^2_{xx}F(x_k)$ can be circumvented and a descent direction can be generated. Furthermore note that for $R = I$, the trust region method can be interpreted as a compromise between Newton's method and the method of steepest descent. For $\alpha_k = 0$, we get

$$x^N_{k+1} = x_k - [\partial^2_{xx}F(x_k)]^{-1}\partial_x F(x_k) \tag{1.40}$$

which is Newton's method, and for large α_k we have approximately

$$x^{sd}_{k+1} = x_k - \alpha_k^{-1}\partial_x F(x_k) \tag{1.41}$$

which is the steepest descent method with α_k^{-1} playing the role of the line search scalar t_k. Thus the direction of search of the trust region method interpolates between the Newton direction and the steepest descent direction (see figure 1.7). Since Newton's method is based on a quadratic approximation of $F(x)$ and the method of steepest descent is based on a linear approximation, it seems that with the trust region method one can, by adjusting α_k, control the approximation used for $F(x)$.

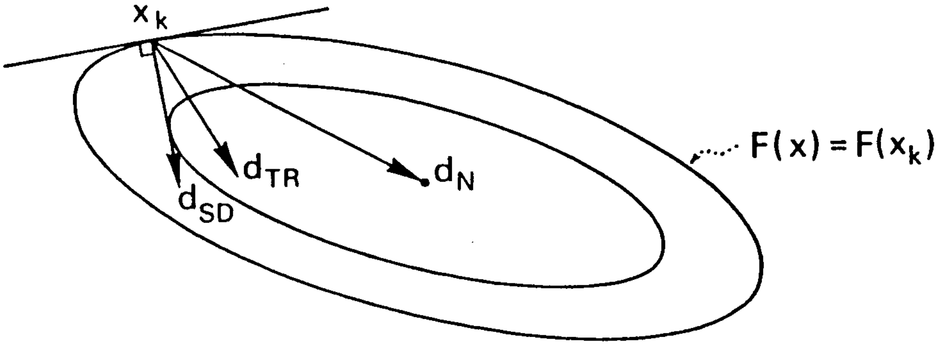


Figure 1.7: d_N = Newton direction, d_{SD} = Steepest descent direction, d_{TR} = Trust region direction.

In order to get a better understanding of this phenomenon, let us study the different approximations involved. We start from the following Taylor series expansion of $F(x)$:

$$F(x) = a(x) + O(\| x - x_k \|^2) \tag{1.42}$$

with

$$a(x) = F(x_k) + \partial_x F(x_k)^*(x - x_k) + \frac{1}{2}(x - x_k)^* \partial^2_{xx} F(x_k)(x - x_k) \tag{1.43}$$

As we know Newton's method is based on the approximation $a(x)$ of $F(x)$. The Newton solution x^N_{k+1} follows then from minimizing (provided this is possible) the quadratic function $a(x)$. Thus $a(x^N_{k+1}) \leq a(x) \forall x \in R^n$ and $a(x^N_{k+1}) < a(x_k) = F(x_k)$ if $x^N_{k+1} \neq x_k$. From this and $F(x^N_{k+1}) = a(x^N_{k+1}) + O(\| x^N_{k+1} - x_k \|^2)$ follows that the objective function gets reduced, i.e. $F(x^N_{k+1}) < F(x_k)$, if $O(\| x^N_{k+1} - x_k \|^2)$ is small enough. Thus descent of the objective function occurs when $a(x^N_{k+1})$ can still be considered an *adequate approximation* of $F(x^N_{k+1})$.

Problems may occur however when this approximation is not adequate, that is, when $O(\| x^N_{k+1} - x_k \|^2)$ is too large. This may happen if matrix $[\partial^2_{xx} F(x_k)]^{-1}$ of (1.40) is "large", i.e. when the matrix $\partial^2_{xx} F(x_k)$ is poorly conditioned and thus the contours of $a(x)$ are very elongated. If the approximation of $F(x)$ by $a(x)$ is inadequate one can improve the approximation by restricting the region for x. In the steepest descent method this is achieved by replacing $x - x_k$ in (1.42) and (1.43) by $t_k d_k$, by taking d_k in the direction of steepest descent, $-\partial_x F$, and then by adjusting t_k so that $\| x^{sd}_{k+1} - x_k \| = \| t_k d_k \|$ is sufficiently small and $F(x^{sd}_{k+1}) = F(x_k + t_k d_k) < F(x_k)$ holds. This idea of a line search along the direction vector d_k to restrict the region of x by validity of the Taylor approximation can in principle also be applied to Newton's method. Instead of (1.40) one gets then

$$x^N_{k+1} = x_k - t_k [\partial^2_{xx} F(x_k)]^{-1} \partial_x F(x_k) \tag{1.44}$$

By taking t_k sufficiently small one can then again ensure that $F(x^N_{k+1}) < F(x_k)$, provided that $\partial^2_{xx} F(x_k)$ is positive-definite. The problem with this modification of Newton's method is however that it cannot deal with those cases where $\partial^2_{xx}F(x_k)$ lacks positive definiteness or is singular. The basis idea of the trust region method is now to replace the one dimensional steepest descent-like restriction along a fixed direction d_k by an n-dimensional restricted *region* for x. That is, in the trust region method again the quadratic approximation (1.43) is used, but now with the additional restriction that x should lie within an ellipsoidal region for which $a(x)$ is *trusted* to be an adequate approximation of $F(x)$ (see figure 1.8).

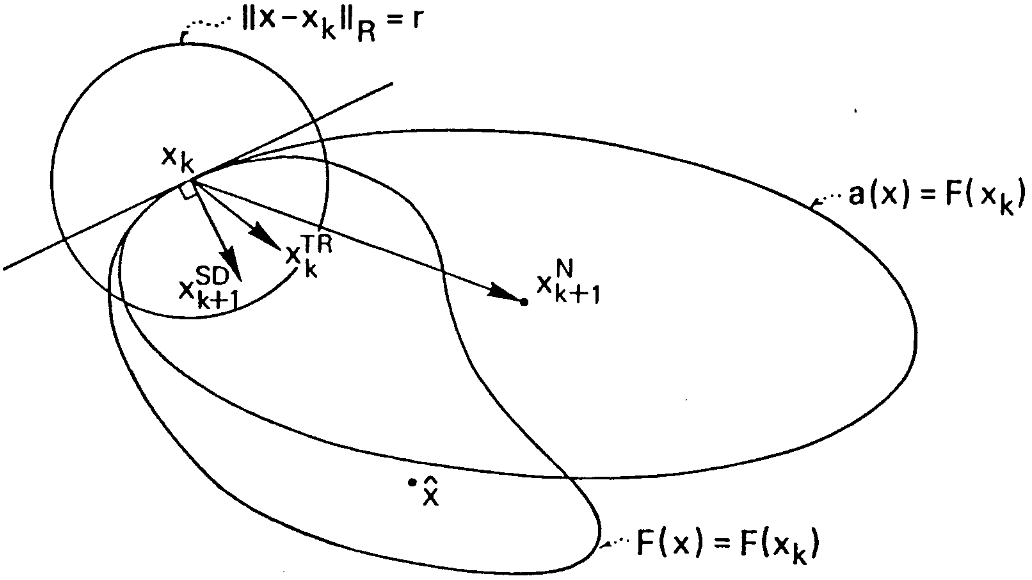


Figure 1.8: The trust region: $\| x - x_k \|_R = [(x - x_k)^* R(x - x_k)]^{\frac{1}{2}} \leq r$.

If we define the quadratic form

$$A(x) = a(x) + \frac{1}{2}\alpha_k(x - x_k)^* R(x - x_k)$$

it follows that if $\partial^2_{xx}F(x_k) + \alpha_k R$ is positive definite then

$$A(x_{k+1}(\alpha_k)) \leq A(x) \ \ \forall x \in R^n$$

This implies that

$$a(x_{k+1}(\alpha_k)) \leq a(x) \ \ \forall x \in \{x \,|\, \| x - x_k \|_R \leq \| x_{k+1}(\alpha_k) - x_k \|_R\}$$

Thus $x_{k+1}(\alpha_k))$ of (1.39) minimizes $a(x)$ over the ellipsoid $\| x - x_k \|_R \leq \| x_{k+1}(\alpha_k) - x_k \|_R$.

The "radius" $r(\alpha_k) = \| x_{k+1}(\alpha_k) - x_k \|_R$ is a decreasing function of α_k. In order to show this, we consider the problem

$$\partial^2_{xx}F(x_k)e_i = \lambda_i R e_i \ , \ i = 1, \ldots, n \tag{1.45}$$

The eigenvectors e_i , $i = 1, \ldots, n$, form a basis of R^n and they can be chosen so that

$$e_i^* R e_j = \delta_{ij} \ , \ i, j = 1, \ldots, n \tag{1.46}$$

Suppose that

$$\partial_x F(x_k) = \sum_{i=1}^{n} c_i R e_i \neq 0$$

Then with (1.45),

$$x_{k+1}(\alpha_k) - x_k = -[\partial^2_{xx}F(x_k) + \alpha_k R]^{-1}\partial_x F(x_k) = -\sum_{i=1}^{n} \frac{c_i}{\lambda_i + \alpha_k} e_i$$

and thus with (1.46)

$$r(\alpha_k) = \| x_{k+1}(\alpha_k) - x_k \|_R = \left[\sum_{i=1}^{n} \left(\frac{c_i}{\lambda_i + \alpha_k}\right)^2\right]^{1/2} \tag{1.47}$$

This shows the monotone decreasing property of $r(\alpha_k)$ if $\lambda_i + \alpha_k > 0$, i.e. if $\partial^2_{xx}F(x_k) + \alpha_k R$ is positive definite. The trust region is therefore an expanding ellipsoid if α_k gets smaller, and a contracting ellipsoid if α_k gets larger.

The trust region method operates now as follows. At the kth-iteration the point $x_{k+1}(\alpha_k)$ of (1.39) is computed for a certain $\alpha_k > 0$. Then the actual reduction $F(x_{k+1}(\alpha_k)) - F(x_k)$ is compared with the predicted reduction $a(x_{k+1}(\alpha_k)) - a(x_k) = a(x_{k+1}(\alpha_k)) - F(x_k)$. If the prediction is poor, the parameter α_k is increased in order to contract the trust region, and the computations are repeated; otherwise $x_{k+1}(\alpha_k)$ is accepted as the new iteration point.

Chapter 2

Nonlinear least squares

2.1 Introduction

This chapter is devoted to nonlinear least-squares problems, i.e. minimization problems in which the objective function is a weighted sum of squared terms

$$F(x) = \frac{1}{2} \parallel y - A(x) \parallel^2 \tag{2.1}$$

where $\parallel . \parallel^2 = (.)^* Q_y^{-1}(.)$; Q_y is positive definite; y is an m-dimensional data vector, and $A(.)$ is a nonlinear vectorfunction or map from R^n into R^m. The factor $\frac{1}{2}$ in (2.1) is merely introduced for convenience.

For varying values of x, $A(x)$ traces locally an n-dimensional surface or manifold embedded in R^m. If the metric of R^m is described by the positive definite matrix Q_y^{-1}, the scalar $\parallel y - A(x) \parallel$ equals the distance from point y to the point $A(x)$ on the manifold. Hence, the problem of minimizing $F(x)$ corresponds to the problem of finding that point on the manifold, say $\hat{y} = A(\hat{x})$, which has least distance to y. This geometry of the nonlinear least-squares problem is sketched in figure 2.1.

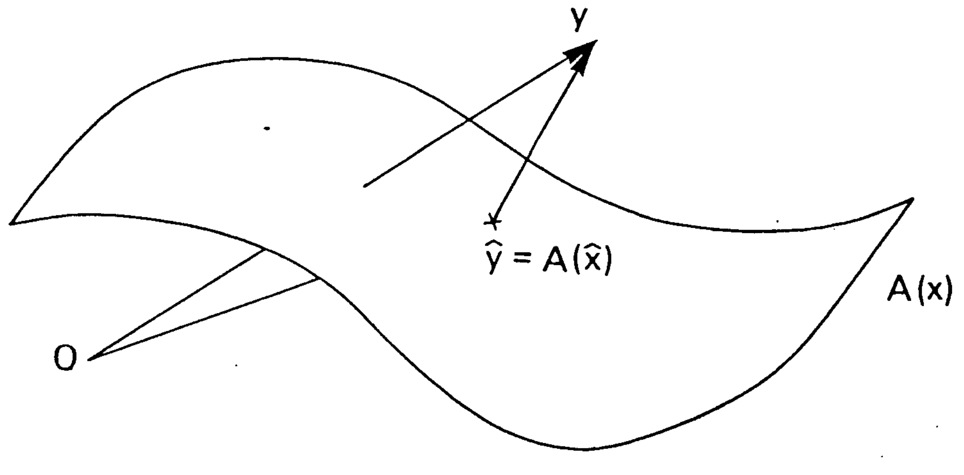


Figure 2.1: Geometry of nonlinear least-squares.

The minimizer of (2.1) can in principle be located by one of the iterative descent methods of the previous chapter. Since

$$\partial_x F(x) = -\partial_x A(x)^* Q_y^{-1}(y - A(x)) \tag{2.2}$$

the steepest-descent method takes the form

$$\boxed{x_{k+1} = x_k + t_k \partial_x A(x_k)^* Q_y^{-1}(y - A(x))} \tag{2.3}$$

And since

$$\partial^2_{xx} F(x) = \partial_x A(x)^* Q_y^{-1} \partial_x A(x) - (y - A(x))^* Q_y^{-1} \partial^2_{xx} A(x)$$

Newton's method takes the form

$$\boxed{x_{k+1} = x_k + [\partial_x A(x_k)^* Q_y^{-1} \partial_x A(x_k) - (y - A(x_k))^* Q_y^{-1} \partial^2_{xx} A(x_k)]^{-1} \partial_x A(x_k)^* Q_y^{-1}(y - A(x_k))} \tag{2.4}$$

Although the steepest-descent method and Newton's method are certainly iterative methods that can locate the minimizer of (2.1), they do not take advantage of the special structure of the objective function (2.1). A method which does take advantage of the "sum of squares" structure of the objective function is the Gauss-Newton method. This method is therefore especially suited for solving nonlinear least-squares problems. In this chapter the Gauss-Newton method will be presented and its characteristics explored.

2.2 The Gauss-Newton Method

The Gauss-Newton method belongs to the same class of iterative descent methods as the steepest-descent method, Newton's method and the trust-region method. The method is characterized by the following choice for the positive-definite matrix $Q(x_k)$ of (1.16)

$$Q(x_k) = [\partial_x A(x_k)^* Q_y^{-1} \partial_x A(x_k)]^{-1} \tag{2.5}$$

With (2.2) follows therefore that the Gauss-Newton method takes the form

$$\boxed{x_{k+1} = x_k + t_k [\partial_x A(x_k)^* Q_y^{-1} \partial_x A(x_k)]^{-1} \partial_x A(x_k)^* Q_y^{-1}(y - A(x_k))} \tag{2.6}$$

Note that the Newton direction reduces to the Gauss-Newton direction if $(y - A(x))^* Q_y^{-1} \partial^2_{xx} A(x)$ is neglected in (2.4). But the particular choice (2.5) is perhaps best motivated if we draw a parallel with *linear* least-squares problems. A least-squares problem is said to be linear if the map $A(x)$ is linear. If $A(x)$ is linear, the minimizer of $F(x)$ follows from solving a system of linear equations. Thus the nonlinearity of $\partial_x F(x)$ is due to the nonlinearity of $A(x)$. The idea is therefore to approximate $F(x)$ by a function which is obtained by replacing $A(x)$ in $\| y - A(x) \|$ by its linearized version $A(x_k) + \partial_x A(x_k)(x - x_k)$. Hence, instead of using a Taylor-expansion of $F(x) = \| y - A(x) \|$ with first or second order terms as is done in case of the steepest descent method or Newton's method, one approximates $F(x)$ through a linearization *within* the norm. The resulting approximation

$$\| y - A(x_k) - \partial_x A(x_k) d_k \|$$

is then minimized as function of d_k. This gives the solution $d_k = -Q(x_k)\partial_x F(x_k)$. Thus the Gauss-Newton direction d_k can be seen as the solution of a linear(ized) least-squares problem. The geometry of the Gauss-Newton method is therefore also one of *orthogonal projection*. That is, the vector $\partial_x A(x_k) d_k$, which lies in the tangentspace of the manifold $A(x)$ at $A(x_k)$, is the orthogonal projection of the residual vector $y - A(x_k)$ onto this tangentspace. (see figure 2.2) This geometric interpretation of the Gauss-Newton method already makes intuitively clear that the geometry of the manifold $A(x)$ must play an important role in the local behaviour of the method. The role of the geometry of the manifold will be made precise in the sections following. In the next sections we start with the geometry of the optimality conditions.

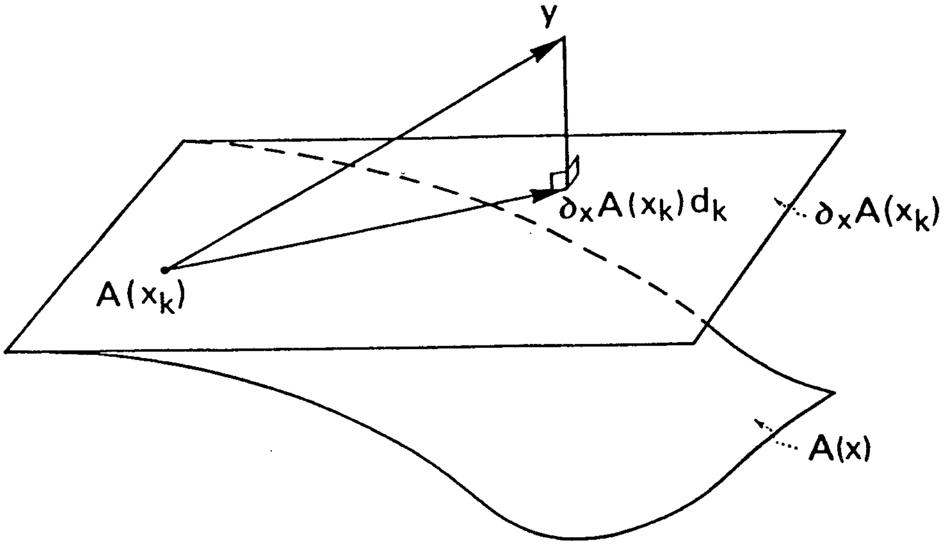


Figure 2.2: Orthogonal projection onto tangentspace of $A(x)$ at $A(x_k)$.

2.3 Geometry of Optimality Conditions

According to theorem 2 of the previous chapter a point $\hat{x}$ is a (local or global) minimum of the objective function $F(x)$ if

$$\begin{array}{ll} a) & \partial_x F(\hat{x}) = 0 \\ b) & \partial^2_{xx} F(\hat{x}) > 0 \end{array}$$

When applied to the objective function $F(x) = \frac{1}{2} \parallel y - A(x) \parallel^2 = \frac{1}{2} \parallel e(x) \parallel^2$ these necessary and sufficient conditions become

$$\begin{array}{ll} a) & \partial_x F(\hat{x}) = -\partial_x A(\hat{x})^* Q_y^{-1} e(\hat{x}) = 0 \\ b) & \partial^2_{xx} F(\hat{x}) = Q(\hat{x})^{-1} - e(\hat{x})^* Q_y^{-1} \partial^2_{xx} A(\hat{x}) > 0 \end{array} \tag{2.7}$$

Both these conditions can be given an interesting geometric interpretation (Teunissen 1984, 1985). The geometric interpretation of (2.7a) is rather simple. Equation (2.7a) states namely that the residual vector $e(x)$ should be *orthogonal* to the tangentspace of manifold $A(x)$ at the solution $\hat{x}$. The interpretation of (2.7b) is somewhat more complicated. In order to interpret (2.7b) geometrically we first introduce the concept of *normal curvature*.

In Gaussian surface theory the normal curvature is defined as the ratio of the second fundamental form and the first fundamental form. The second fundamental form for the map $A(x)$ is given by $v^*(n^* Q_y^{-1} \partial^2_{xx} A(x))v$, where v is a tangentvector and n is a unit normal vector, i.e. $n^* Q_y^{-1} \partial_x A(x) = 0$ and $n^* Q_y^{-1} n = 1$. The first fundamental form for the map $A(x)$ is given by $v^* Q(x)^{-1} v$. Hence, the normal curvature is defined as

$$\boxed{k_n(v) = \frac{v^*(n^* Q_y^{-1} \partial^2_{xx} A(x))v}{v^* Q(x)^{-1} v}} \tag{2.8}$$

The extreme values of this ratio are the *principal normal curvatures*. They follow as the eigenvalues of the generalized eigenvalue problem

$$| n^* Q_y^{-1} \partial^2_{xx} A(x) - \lambda Q(x)^{-1} | = 0$$

Since in the classical Gaussian surface theory $A(x)$ is a map from R^2 into R^3, the dimension of the rangespace of $\partial_x A(x)$, $R(\partial_x A(x))$, is two and the dimension of its orthogonal complement, $R(\partial_x A(x))^\perp$, is one. Thus in the classical case one has just one second fundamental form and two principal normal curvatures. In our case however, $A(x)$ is a map from R^n into R^m. Therefore $dim\, R(\partial_x A(x)) = n$ and $dim R(\partial_x A(x))^\perp = m - n$. This implies that in our case the number of principal normal curvatures equals $n.(m-n)$. We will denote the n-number of principal normal curvatures for the normal direction n by

$$k_n^1 \leq k_n^2 \leq \cdots \leq k_n^n \tag{2.9}$$

It should be noted that the normal curvature is *invariant* under a change of variables in $A(x)$. This can be seen as follows. Let $x(\bar{x})$ be a one-to-one map from R^n to R^n. Then

$$\begin{aligned} \partial^2_{\bar{x}\bar{x}} A(\bar{x}) &= \partial_{\bar{x}} x^* \partial^2_{xx} A(x(\bar{x})) \partial_{\bar{x}} x + \partial_x A(x(\bar{x})) \partial^2_{\bar{x}\bar{x}} x \\ Q(\bar{x})^{-1} &= \partial_{\bar{x}} x^* Q(x(\bar{x}))^{-1} \partial_{\bar{x}} x \\ v &= \partial_{\bar{x}} x \bar{v} \end{aligned}$$

Substitution into

$$k_n(\bar{v}) = \frac{\bar{v}^* (n^* Q_y^{-1} \partial^2_{\bar{x}\bar{x}} A(\bar{x})) \bar{v}}{\bar{v}^* Q(\bar{x})^{-1} \bar{v}}$$

shows then, since $n^* Q_y^{-1} \partial_x A(x(\bar{x})) = 0$, that $k_n(\bar{v}) = k_n(v)$. This invariance of the normal curvature under a change of variables implies that the curvature $k_n(v)$ is an *intrinsic* property of the manifold $A(x)$ embedded in R^m.

In order to relate the normal curvature to condition (2.7b), note that (2.7b) is equivalent to

$$\frac{v^* (e(\hat{x})^* Q_y^{-1} \partial^2_{xx} A(\hat{x})) v}{v^* Q(\hat{x})^{-1} v} < 1 \quad \forall\, v \in R^n \tag{2.10}$$

Hence, if we introduce the unit normal vector

$$\hat{n} = \frac{e(\hat{x})}{\| e(\hat{x}) \|}$$

we may write (2.10) with the help of (2.8) also as

$$k_{\hat{n}}(v) \| e(\hat{x}) \| < 1 \quad \forall\, v \in R^n \tag{2.11}$$

This important result shows that condition (2.7b) is governed by two distinct quantities, namely the curvature of the manifold and the amount of inconsistency of the observation vector.

We can now rephrase the necessary and sufficient conditions of (2.7) in geometric terms as

$$\boxed{\begin{array}{ll} a) & e(\hat{x}) \perp R(\partial_x A(\hat{x})) \\ b) & k_{\hat{n}}^n \| e(\hat{x}) \| < 1 \end{array}} \tag{2.12}$$

Note that both these conditions are invariant under a change of variables.

As an exemplification of (2.12), assume that $A(x)$ is a circle embedded in R^2 with curvature k. Figure 2.3 shows for this case four possible situations that may occur. In all four cases the point $\hat{x}$ is a stationary point and satisfies condition (2.12a). In figure 2.3a, $A(\hat{x})$ is a minimizer because of negative curvature. In figure 3b, $A(\hat{x})$ is a minimizer since the curvature, although positive, is still small enough relative to $\| \hat{e} \|$. In figure 2.3c, $A(\hat{x})$ is a nonunique minimizer. And in figure 2.3d, $A(\hat{x})$ is a maximizer instead of a minimizer.

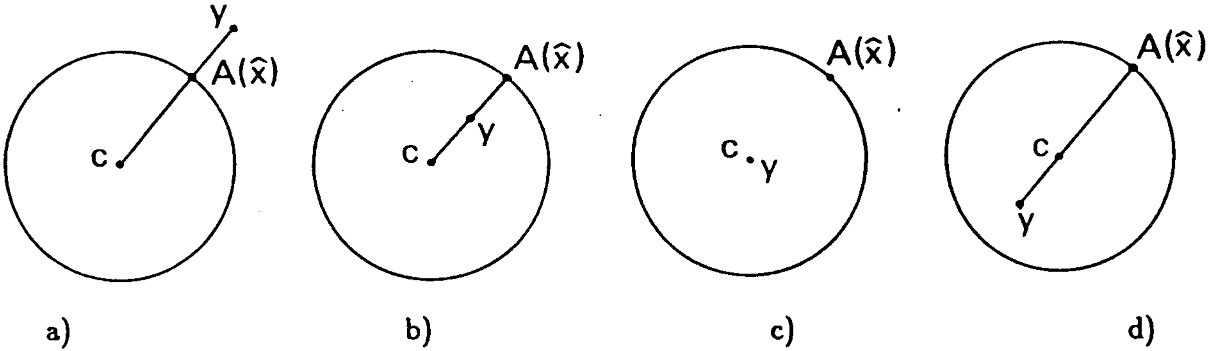


Figure 2.3: a) k negative, b) k positive but $k \| \hat{e} \| < 1$, c) k positive and $k \| \hat{e} \| = 1$, d) k positive and $k \| \hat{e} \| > 1$.

2.4 Local Convergence of the Gauss-Newton Method

In order to derive the rate of convergence of the Gauss-Newton method we expand (2.6) into a Taylorseries at the solution $\hat{x}$. With $\partial_x A(\hat{x})^* Q_y^{-1} e(\hat{x}) = 0$, this gives

$$x_{k+1} - \hat{x} = [(1 - t_k)I + t_k Q(\hat{x})[e(\hat{x})^* Q_y^{-1} \partial^2_{xx} A(\hat{x})]](x_k - \hat{x}) + O(\| x_k - \hat{x} \|) \tag{2.13}$$

This shows that the Gauss-Newton method has a *linear rate of convergence* for points sufficiently close to the solution $\hat{x}$. If we take the eigenvectors v_i , $i = 1, \ldots, n$, of the generalized eigenvalue problem

$$\hat{n}^* Q_y^{-1} \partial^2_{xx} A(\hat{x}) v_i = k^i_{\hat{n}} Q(\hat{x})^{-1} v_i$$

as base vectors of the tangentspace of the manifold $A(x)$ at $\hat{x}$, and reparametrize $x_{k+1} - \hat{x}$ and $x_k - \hat{x}$ as

$$x_{k+1} - \hat{x} = \sum_{i=1}^{n} u^i_{k+1} v_i \quad \textit{and} \quad x_k - \hat{x} = \sum_{i=1}^{n} u^i_k v_i$$

we can write (2.13) in terms of the principal normal curvatures as

$$u^i_{k+1} = [(1 - t_k) + t_k k^i_{\hat{n}} \| e(\hat{x}) \|] u^i_k + O(\| x_k - \hat{x} \|) \tag{2.14}$$

This expression shows that the local convergence of the Gauss-Newton method is *invariant* against a change of variables. Hence, the rate of convergence of the Gauss-Newton method cannot be speed up or slowed down by a particular choice of parametrization. If the positive scalar t_k in (2.14) is taken to be equal to one, the rate of convergence becomes approximately

$$\boxed{\| x_{k+1} - \hat{x} \| \leq [max\{| k^1_{\hat{n}} |, | k^n_{\hat{n}} |\} \| e(\hat{x}) \|] \| x_k - \hat{x} \|} \tag{2.15}$$

The parameter norm in this expression is with respect to the *induced* metric $Q(\hat{x})^{-1}$. Expression (2.15) shows that the errormagnitude gets reduced if

$$max\{\mid k^1_{\hat{x}} \mid, \mid k^n_{\hat{x}} \mid\} \parallel e(\hat{x}) \parallel < 1 \tag{2.16}$$

i.e. if the observation point y lies within a *hypersphere* with centre $A(\hat{x})$ and a radius equal to the inverse of the in absolute value largest curvature. If this is the case then by virtue of (1.22) local convergence of the Gauss-Newton method is *guaranteed*. Note however that local convergence is not necessarily ensured by the fact that $A(\hat{x})$ is a local minimum of $\parallel y - A(x) \parallel$. This follows if we compare inequality (2.16) with (2.12b). In figure 2.3a for instance, $A(\hat{x})$ is a minimizer, but $\parallel e(\hat{x}) \parallel$ may still be too large for convergence to occur.

As an exemplification of (2.15), assume that $A(x)$ represents a unit circle, that $Q_y = I$ and $y = (1.5, 0.0)^*$. The least squares solution is then given by $\hat{x} = 0$ and the local convergence factor by

$$k_{\hat{x}} \parallel e(\hat{x}) \parallel = -0.5 \tag{2.17}$$

The results of the Gauss-Newton iteration are given in table 2.1. They clearly show that the errormagnitude gets reduced by the factor (2.17) in each iteration step. Also note the oscillatory character of the iteration. Oscillation or *overshoot* generally occurs if the curvatures are negative (confer (2.14) for $t_k = 1$). *Undershoot* on the other hand occurs if the curvatures are positive.

iterationstep k	$A^1(x) = cosx$	$A^2(x) = sinx$	x_k
1	0.96235	-0.27180	-0.27526
2	0.99124	0.13205	0.13244
3	0.99785	-0.06560	-0.06564
4	0.99946	0.03274	0.03275
5	0.99987	-0.01637	-0.01637
6	0.99997	0.00818	0.00818

Table 2.1: Gauss-Newton iteration for orthogonal projection onto a unitcircle.

If inequality (2.16) is not satisfied, one can enforce local convergence by a suitable choice for the positive scalar t_k of (2.14). It follows (compare with our discussion of the steepest descent method) that the optimal choice for t_k is

$$t_k = 1/[1 - \parallel e(\hat{x}) \parallel \frac{1}{2}(k^1_{\hat{x}} + k^n_{\hat{x}})] \tag{2.18}$$

With this coice for t_k it follows from (2.14) that instead of (2.15) we have

$$\boxed{\parallel x_{k+1} - \hat{x} \parallel \leq \left[\frac{\parallel e(\hat{x}) \parallel (k^n_{\hat{x}} - k^1_{\hat{x}})}{2 - \parallel e(\hat{x}) \parallel (k^n_{\hat{x}} + k^1_{\hat{x}})}\right] \parallel x_k - \hat{x} \parallel} \tag{2.19}$$

In this case local convergence is guaranteed if (2.12b) holds. Equation (2.18) shows that the simplest choice $t_k = 1$ is close to optimal if either $\parallel e(\hat{x}) \parallel$ is small enough or the average of the extreme curvatures is small enough. Thus for points sufficiently close to the solution, the simplest line search strategy can be considered adequate if the manifold is moderately curved at $\hat{x}$ and/or the observation point is close enough to the manifold.

So far it was assumed that we were dealing with a curved manifold with inconsisitent data. But what happens with the local convergence behaviour of the Gauss-Newton method if either the manifold is *flat* (zero-curvature) or the data is *consistent* (zero-residual vector)? In order to answer this question we first note that for a flat manifold, the orthogonal projector

$$P_{\partial_x A} = \partial_x A(x)[\partial_x A(x)^* Q_y^{-1} \partial_x A(x)]^{-1} \partial_x A(x)^* Q_y^{-1}$$

is constant and independent of x, and

$$A(\hat{x}) = \bar{y} + P_{\partial_x A}(y - \bar{y}) \quad \forall\, \bar{y} \in A(x)$$

With this follows that

$$\begin{aligned} Q(x_k)\partial_x A(x_k)^* Q_y^{-1}(y - A(x_k)) &= Q(x_k)\partial_x A(x_k)^* Q_y^{-1} P_{\partial_x A}(y - A(x_k)) \\ &= Q(x_k)\partial_x A(x_k)^* Q_y^{-1}(A(\hat{x}) - A(x_k)) \end{aligned}$$

This result shows that for both the cases of a flat manifold and consistent data, the observation vector y in (2.6) may be replaced by $A(\hat{x})$. From a Taylor series expansion at $\hat{x}$ of (2.6) with y replaced by $A(\hat{x})$ follows then with $t_k = 1$ and the identity

$$\partial_x Q(x) Q(x)^{-1} + Q(x)\partial^2_{xx} A(x) Q_y^{-1} \partial_x A(x) + Q(x)\partial_x A(x)^* Q_y^{-1} \partial^2_{xx} A(x) = 0$$

that

$$\boxed{x_{k+1} - \hat{x} = \frac{1}{2} Q(\hat{x})\partial_x A(\hat{x})^* Q_y^{-1}[(x_k - \hat{x})^* \partial^2_{xx} A(\hat{x})(x_k - \hat{x})] + O(\| x_k - \hat{x} \|^2)} \tag{2.20}$$

This shows that the rate of convergence of the Gauss-Newton method is *quadratic* in case of flat manifolds and/or consistent data. If we take the norm of (2.20) we get

$$\| x_{k+1} - \hat{x} \| = \frac{1}{2} \| P_{\partial_x A}(x_k - \hat{x})^* \partial^2_{xx} A(\hat{x})(x_k - \hat{x}) \|$$

This shows that the convergence factor is determined by the *tangential* components of $\partial^2_{xx} A(\hat{x})$. Compare this with for instance (2.13), where the convergence factor depends on the *normal* component of $\partial^2_{xx} A(\hat{x})$.

As an exemplification of (2.20), assume that $A(x)$ represents a straight line, that $A^1(x) = exp(10x)$, $A^2(x) = exp(10x)$, $Q_y = I$ and $y = (0, 2e)^*$. The least-squares solution is then given by $\hat{x} = 0.1$ and the convergence factor by

$$\frac{1}{2} Q(\hat{x})\partial_x A(\hat{x})^* Q_y^{-1} \partial^2_{xx} A(\hat{x}) = 5$$

The results of the Gauss-Newton iteration are given in table 2.2. They clearly show the quadratic rate of convergence. Also note that e.g. $(x_4 - \hat{x}) = 5(x_3 - \hat{x})^2$.

iterationstep k	$A^1(x) = exp(10x)$	$A^2(x) = exp(10x)$	x_k
1	5.57494	5.57494	0.17183
2	3.33967	3.33967	0.12059
3	2.77267	2.77267	0.10198
4	2.71881	2.71881	0.10002
5	2.71828	2.71828	0.10000

Table 2.2: Gauss-Newton iteration for orthogonal projection onto a straight line with a nonlinear parametrization.

2.5 A Convergence Criterion

Every iteration method needs one or more termination criteria in order to be able to test whether the iteration should be continued or not. Apart from the computertime to termination and/or the number of iterations, the most important termination criterion is the one which measures the success in obtaining an optimal solution. Since iterative descent methods try to locate a stationary point of the objective function $F(x)$, convergence can be declared if the gradient of $F(x)$, evaluated at the current iteration point x_k, falls below a preset tolerance level. A test for convergence is therefore

$$\| \partial_x F(x_k) \| < \epsilon$$

For our least-squares problem this becomes

$$\| \partial_x A(x_k)^* Q_y^{-1} e(x_k) \| < \epsilon \tag{2.21}$$

In order to make the norm of the gradient *invariant* to a change of variables and thus insensitive to scale changes we chooce to take the norm in (2.21) with respect to the induced metric $Q(x_k)^{-1}$. This gives for the Gauss-Newton method the convergency test

$$\boxed{\| x_{k+1} - x_k \| < \epsilon} \tag{2.22}$$

Note that since $\partial_x A(x_k)(x_{k+1} - x_k) = P_{\partial_x A(x_k)}$ the convergency test can also be written as

$$\boxed{\| P_{\partial_x A(x_k)} e(x_k) \| < \epsilon} \tag{2.23}$$

where the norm is with respect to the metric Q_y^{-1}.

In order to apply the convergency test we need to chooce a value for the tolerance level ϵ. On what should we base our choice for ϵ? It seems natural to base the choice for ϵ on the quality of the observation vector y. Since x_k is the *exact* least-squares solution of the perturbed minimization problem

$$\min_x \| [y - P_{\partial_x A(x_k)} e(x_k)] - A(x) \|$$

(see figure 2.4), it follows that the tolerance level ϵ of (2.23) should be chosen such that a perturbation of y with $P_{\partial_x A(x_k)} e(x_k)$ is considered insignificant.

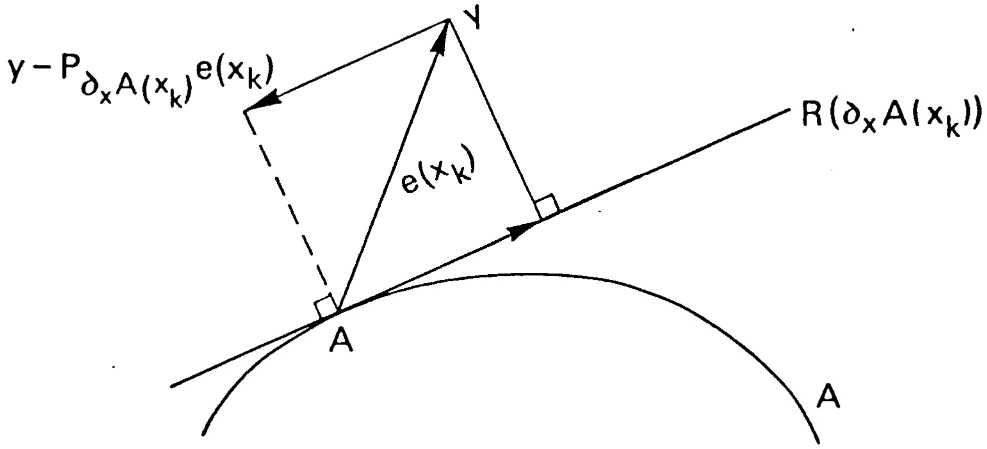


Figure 2.4: Perturbaton of y with $P_{\partial_x A(x)} e(x_k)$.

Once convergence is declared it is of interest to know the errormagnitude of the computed quantities x_k, $A(x_k)$ and $\| e(x_k) \|$.

In order to determine the relation between the above convergency indicator and the errormagnitude of x_k, we expand $P_{\partial_x A(x_k)} e(x_k)$ in a Taylor series at $\hat{x}$. This gives

$$P_{\partial_x A(x_k)} e(x_k) = \partial_x A(\hat{x})[Q(\hat{x})\partial^2_{xx} A(\hat{x}) Q_y^{-1} e(\hat{x}) - I](x_k - \hat{x}) + O(\| x_k - \hat{x} \|)$$

From this follows, if $\hat{x}$ is a local minimum of $\| y - A(x) \|$ and thus $k^n_{\hat{A}} \, \| e(\hat{x}) \| < 1$ (see (2.12b)),that

$$\left.\begin{array}{c} [1 - k^n_{\hat{A}} \| \hat{e} \|] \, \| x_k - \hat{x} \| \\ +O(\| x_k - \hat{x} \|) \end{array}\right\} \leq \| P_{\partial_x A(x_k)} e(x_k) \| \leq \left\{\begin{array}{c} [1 - k^1_{\hat{A}} \| \hat{e} \|] \, \| x_k - \hat{x} \| \\ +O(\| x_k - \hat{x} \|) \end{array}\right.$$

Hence, we have the approximate interval

$$\boxed{\frac{\| x_{k+1} - x_k \|}{1 - k^1_{\hat{A}} \| \hat{e} \|} \leq \| x_k - \hat{x} \| \leq \frac{\| x_{k+1} - x_k \|}{1 - k^n_{\hat{A}} \| \hat{e} \|}} \quad (2.24)$$

In an analogous way we can derive the approximate intervals

$$\boxed{\frac{\| x_{k+1} - x_k \|}{1 - k^1_{\hat{A}} \| \hat{e} \|} \leq \| A(x_k) - A(\hat{x}) \| \leq \frac{\| x_{k+1} - x_k \|}{1 - k^n_{\hat{A}} \| \hat{e} \|}} \quad (2.25)$$

and

$$\boxed{\frac{\| x_{k+1} - x_k \|^2}{1 - k^1_{\hat{A}} \| \hat{e} \|} \leq \| e(x_k) \|^2 - \| e(\hat{x}) \|^2 \leq \frac{\| x_{k+1} - x_k \|^2}{1 - k^n_{\hat{A}} \| \hat{e} \|}} \quad (2.26)$$

And similarly we find for the errormagnitude of the individual parameters, the upperbound

$$\boxed{| x^\alpha_k - \hat{x}^\alpha | \leq \sigma_{\hat{x}^\alpha} \frac{\| x_{k+1} - x_k \|}{1 - k^n_{\hat{A}} \| \hat{e} \|}} \quad (2.27)$$

with $\sigma_{\hat{x}^\alpha}$ the square root of the αth-diagonal element of $Q(\hat{x})$.

All the above inequalities show how the errormagnitudes of the computed quantities are related to the convergency indicator $\| x_{k+1} - x_k \|$. In particular note that $\| x_k - \hat{x} \|$ can be large if $k^n_{\hat{A}} \, \| \hat{e} \|$ is close to one. This happens if y is close to the centre of curvature of $k^n_{\hat{A}}$, in which case the objective function $\| e(x) \|$ is flat near $\hat{x}$.

Chapter 3

On Measures of Nonlinearity and Biases in Nonlinear Least-squares estimators

3.1 Introduction

Almost all functional relations in our geodetic models are nonlinear. Hence, one might question whether the use of the ideas, concepts and results from the theory of linear estimation is justifiable in all cases. Of course, one may argue that probably most nonlinear models are only moderately nonlinear and thus permit the use of a linear(ized) model. This is true. Nevertheless, we need to have ways of assessing the amount of nonlinearity in nonlinear models and methods to prove whether a linear(ized) model is a sufficient approximation. We therefore need to know the impact of nonlinearity on the distributional properties of nonlinear estimators. We will briefly discuss in this section some general methods of deriving the distributions of nonlinear estimators. Most of these exact methods are however not applicable in practice.

Let $\underline{x}$ be a random variable and $F(.)$ a nonlinear function from R to R. Our objective is to find the distribution of $F(\underline{x})$. One way to estimate the distribution of $F(\underline{x})$ would be to use computer simulation (Monte Carlo methods). Cne replicates the series of experiments as many times as one pleases, each time with a new sample x drawn from the parent distribution $p_{\underline{x}}(x)$ and so obtains the relevant distributional properties of $F(\underline{x})$ by averaging over all replications. Although this approach could give us valuable insight into the effect of nonlinearity for a particular problem, the method does not give us a general formula on the basis of which a qualitative analysis can be carried out.

An alternative way to estimate the properties of nonlinear estimators is to rely on the results from asymptotic theory. The central idea of asymptotic theory is that when the number m of observations is large and errors of estimation corresponding small, simplifications become available that are not available in general. The rigorous mathematical development involves limiting distributional results holding as $m \rightarrow \infty$ and is closely related to the classical limit theorems of probability theory. In recent years many researchers have concentrated on developing an asymptotic theory for nonlinear least-squares estimation. In (Jennrich, 1969) a first complete account was given of the asymptotic properties of nonlinear least-squares estimators. And in (Schmidt, 1982) it was shown how the asymptotic theory can be utilized to formulate asymptotic exact test statistics. See also the book (Bierens, 1984). Unfortunately, since the theory is based on the assumption that $m \rightarrow \infty$, the results abtained up to now cannot satisfy all the requirements of applications in practice.

A third approach to estimate the distribution of $F(\underline{x})$ is based on the fundamental relations that define distribution functions. If the density $p_{\underline{x}}(x)$ of $\underline{x}$ is given, then theoretically at least, we can find the distribution of $F(\underline{x})$. This follows since the cumulative distribution $P_{\underline{y}}(y)$ of $\underline{y} = F(\underline{x})$ satisfies

$$P_{\underline{y}}(y) = Prob(\underline{y} \leq y) = Prob(F(\underline{x}) \leq y) = \int_{\{x|F(x)\leq y\}} p_{\underline{x}}(x)dx \tag{3.1}$$

for fixed y. Since (3.1) describes the probability of an event in terms of x, such a probability can theoretically be determined by integrating the density of $\underline{x}$ over the region corresponding to the event. The practical problem with this method is however that in general one cannot easily evaluate the desired probability for each y.

Instead of using the cumulative distribution one can also use the density function. Under some restrictions on $F(.)$ equation (3.1) can namely be worked out to give the density $p_{\underline{y}}(y)$ of $\underline{y} = F(\underline{x})$ in terms of the density $p_{\underline{x}}(x)$ of $\underline{x}$:

$$p_{\underline{y}}(y) = \frac{p_{\underline{x}}(F^{-1}(y))}{\mid d_x F(F^{-1}(y)) \mid} \tag{3.2}$$

The difficulty with this method is however that one needs the inverse of the nonlinear function $F(.)$.

Finally one can try to derive some of the moments of the distribution of $F(\underline{x})$, for instance its mean and variance:

$$E\{F(\underline{x})\} = \int_{-\infty}^{\infty} F(x)p_{\underline{x}}(x)dx \quad ; \quad Var\{F(\underline{x})\} = \int_{-\infty}^{\infty} [F(x) - E\{F(\underline{x})\}]^2 p_{\underline{x}}(x)dx \tag{3.3}$$

The complexity of these computations depends very much on the nature of the functions $F(.)$ and $p_{\underline{x}}(.)$. But in general they can become quite complicated, especially in the multivariate case.

If in a particular problem it is impossible to apply the above given exact methods, the next one thing one can try to do is to make use of approximations. This can be done by using a suitable Taylor expansion. In this way and based on (3.2), (Pazman 1987) obtained an approximation to the density of nonlinear least-squares estimators. In a similar way and based on (3.3), (Teunissen 1984,1985,1988) obtained approximate expressions for the mean and covariance matrix of nonlinear least-squares estimators. In this chapter we will review in some detail some of the results of (Teunissen 1984,1985). We will restrict ourselves however to the first moments of the nonlinear least-squares estimators. We also propose some relatively easy computable measures of nonlinearity. But before we discuss the nonlinear least-squares problem, we will first derive an approximation to the mean of an *arbitrary* nonlinear estimator. This is done in the next session.

3.2 The Bias of Nonlinear Estimators

Let $\underline{x}$ be a random n-vector and $F(.)$ be a nonlinear map from R^n into R^m. We define the random m-vector $\underline{y}$ as

$$\underline{y} = F(\underline{x}) \tag{3.4}$$

Our objective is to find an approximate expression for the first moment of the random m-vector $\underline{y}$. We will assume that the random n-vector $\underline{x}$ is an estimator of x. An estimator $\underline{x}$ of x is said to be *unbiased* if $E\{\underline{x}\} = x$. Otherwise the estimator is said to be *biased*. We will denote the bias in $\underline{x}$ by

$$b_x = E\{\underline{x}\} - x \tag{3.5}$$

Furthermore we denote the covariance matrix of $\underline{x}$ by $\sigma^2 Q_x$, where σ^2 is the variancefactor of unit weight. Thus

$$\sigma^2 Q_x = E\{(\underline{x} - E\{\underline{x}\})(\underline{x} - E\{\underline{x}\})^*\} \tag{3.6}$$

The bias in the estimator $\underline{y}$ of $y = F(x)$ is denoted by

$$b_y = E\{\underline{y}\} - y = E\{\underline{y}\} - F(x) \tag{3.7}$$

If we assume that map $F(.)$ is sufficiently smooth we can derive an approximation to the bias in $\underline{y}$ by expanding (3.4) into a Taylorseries at x. This gives

$$\underline{y} = F(\underline{x}) = F(x) + \partial_x F(x)(\underline{x} - x) + \frac{1}{2}(\underline{x} - x)^* \partial^2_{xx} F(x)(\underline{x} - x) + \cdots \tag{3.8}$$

If we take the expectation and use (3.7) we get

$$b_y = E\{\underline{y}\} - F(x) = \partial_x F(x) E\{\underline{x} - x\} + \frac{1}{2} E\{(\underline{x} - x)^* \partial^2_{xx} F(x)(\underline{x} - x)\} + \cdots \tag{3.9}$$

The second term on the righthand side of (3.9) is an m-vector of expected values of quadratic forms. The expected value of the quadratic form $\underline{x}^* \partial^2_{xx} F^i(x) \underline{x}$ is given by

$$E\{\underline{x}^* \partial^2_{xx} F^i(x) \underline{x}\} = \sigma^2 trace[\partial^2_{xx} F^i(x) Q_x] + E\{\underline{x}\}^* \partial^2_{xx} F^i(x) E\{\underline{x}\} \tag{3.10}$$

Hence, with the help of (3.5) and (3.10) it follows from (3.9) that the bias in $\underline{y}$ is given by

$$\boxed{b_y = \partial_x F(x) b_x + \frac{1}{2}\sigma^2 trace[\partial^2_{xx} F(x) Q_x] + \frac{1}{2} b_x^* \partial^2_{xx} F(x) b_x + \cdots} \tag{3.11}$$

This important formula shows how the bias in $\underline{y}$ depends on

i the *bias* b_x in $\underline{x}$

ii the *precision* $\sigma^2 Q_x$ of $\underline{x}$

iii the *nonlinearity* of the map $F(.)$

It is remarked, that the covariance matrix of $\underline{y} = F(\underline{x})$ can be derived in a way analogous to our derivation of the bias (Teunissen 1988).

In order to see formula (3.11) at work for the case that $b_x = 0$ we consider the following example. Consider the function $\underline{y} = \underline{l} cos \underline{\alpha}$, where $\underline{l}$ stands for distance and $\underline{\alpha}$ for azimuth. We assume that $\underline{l}$ and $\underline{\alpha}$ are uncorrelated random variables with mean l and α, and variance σ_l^2 and σ_α^2 respectively. A Taylor expansion of $\underline{y} = \underline{l} cos \underline{\alpha}$ about l and α gives

$$\underline{y} = l cos\alpha + (cos\alpha \quad l sin\alpha) \begin{bmatrix} \underline{l} - l \\ \underline{\alpha} - \alpha \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \underline{l} - l \\ \underline{\alpha} - \alpha \end{bmatrix}^* \begin{bmatrix} 0 & -sin\alpha \\ -sin\alpha & -l cos\alpha \end{bmatrix} \begin{bmatrix} \underline{l} - l \\ \underline{\alpha} - \alpha \end{bmatrix} + \cdots$$

From this follows that

$$E\{\underline{y} - l cos\alpha\} \doteq \frac{1}{2} trace\{ \begin{bmatrix} 0 & -sin\alpha \\ -sin\alpha & -l cos\alpha \end{bmatrix} \begin{bmatrix} \sigma_l^2 & 0 \\ 0 & \sigma_\alpha^2 \end{bmatrix} \} = -\frac{1}{2}\sigma_\alpha^2 l cos\alpha$$

Thus the bias in $\underline{y}$ due to nonlinearity increases with increasing distance and decreases with increasing precision.

In section 3.5 formula (3.11) will be applied to derive the bias in the least-squares residual vector.

3.3 Three simple Measures of Nonlinearity of the Observation Equations

Before we continue our discussion of biases in nonlinear estimators and in particular the biases in least-squares estimators, let us first consider the cause of these biases. The biases in nonlinear least-squares estimators are caused by the nonlinearity of the observation equations. In order to diagnose the significance of nonlinearity we propose in this section three simple measures of nonlinearity of the observation equations.

Consider the model

$$E\{\underline{y}\} = A(x) \quad , \quad E\{(\underline{y} - E\{\underline{y}\})(\underline{y} - E\{\underline{y}\})^*\} = \sigma^2 Q_y \tag{3.12}$$

where $A(.)$ is a nonlinear map from R^n into R^m. Then in theory the well-known results from linear inference are not applicable anymore. Nevertheless, in most practical applications one still to a large extent relies on the results from the theory of linear inference. That is, one usually assumes that it is permitted to replace the nonlinear model (3.12) by its linearized version

$$E\{\Delta \underline{y}\} = \partial_x A(x_0)\Delta x \quad , \quad E\{(\Delta \underline{y} - E\{\Delta \underline{y}\})(\Delta \underline{y} - E\{\Delta \underline{y}\})^*\} = \sigma^2 Q_y \tag{3.13}$$

Statistical inference is then usually based on this linear model. In order to find out whether it is justified to replace the nonlinear model (3.12) by the linear model (3.13), we need to have means for diagnosing the approximations involved. Since the second order remainder

$$R_2(x) = \frac{1}{2}\Delta x^* \partial^2_{xx} A(x)\Delta x \tag{3.14}$$

is neglected in (3.13), a bound on this remainder may be used as a measure of nonlinearity.

The following bound is based on an evaluation of the individual elements of the Hessian matrices of the observation equations

$$\boxed{\mid R_2^i(x) \mid \leq \frac{c^i}{2} n \parallel \Delta x \parallel^2 \quad , \quad if \quad \mid \partial^2_{\alpha\beta} A^i(x) \mid \leq c^i \quad , \quad \begin{matrix} i = 1, \ldots, m \\ \alpha, \beta = 1, \ldots, n \end{matrix}} \tag{3.15}$$

The evaluation of the scalars c^i , $i = 1, \ldots, m$, is in general relatively easy. A disadvantage of the above upperbound is that it can become somewhat pessimistic for large n, i.e. in case of many parameters.

An alternative bound for the remainder $R_2(x)$ may be based on the extreme eigenvalues of the Hessian matrix $\partial^2_{xx} A^i(x)$:

$$\boxed{\frac{1}{2}\lambda^i_{min} \parallel \Delta x \parallel^2 \leq R_2^i(x) \leq \frac{1}{2}\lambda^i_{max} \parallel \Delta x \parallel^2 \quad , \quad i = 1, \ldots, m} \tag{3.16}$$

The computation of the extreme eigenvalues is not too difficult if the matrix $\partial^2_{xx} A^i(x)$ is sparce, i.e. if only a few parameters are involved in the observation equations. This is usually the case in geodetic applications. Take for instance the distance-observation equation $l_{ij} = (x^2_{ij} + y^2_{ij})^{1/2}$ where l_{ij} is the distance between two points i and j, and x_{ij} and y_{ij} are the corresponding cartesian coordinate differences. The Hessian matrix reads then

$$\partial^2_{\alpha\beta} l_{ij} = \frac{1}{l^3_{ij}} \begin{bmatrix} y^2_{ij} & -x_{ij}y_{ij} & -y^2_{ij} & x_{ij}y_{ij} \\ -x_{ij}y_{ij} & x^2_{ij} & x_{ij}y_{ij} & -x^2_{ij} \\ -y^2_{ij} & x_{ij}y_{ij} & y^2_{ij} & -x_{ij}y_{ij} \\ x_{ij}y_{ij} & -x^2_{ij} & -x_{ij}y_{ij} & x^2_{ij} \end{bmatrix}$$

The extreme eigenvalues of this matrix are easy to compute. They are

$$\lambda_{min} = 0 \quad and \quad \lambda_{max} = 1/l_{ij}$$

Thus, for the second order remainder of the distance equation the following bounds hold:

$$0 \leq R_2 \leq (\Delta x_{ij}^2 + \Delta y_{ij}^2)/2l_{ij} \tag{3.17}$$

Some numerical values of this interval are given in table 3.1.

$\Delta x_{ij}, \Delta y_{ij}$	l_{ij}	R_2
100 m	1 km	$\leq$ 20. m
50 m	1 km	$\leq$ 2.5 m
10 m	1 km	$\leq$ 0.1 m
5 m	1 km	$\leq$ .03 m

Table 3.1: Bounds on the second order remainder R_2 of $l_{ij} = (x_{ij}^2 + y_{ij}^2)^{1/2}$.

Finally a third way to measure the nonlinearity of the observation equations is to take a (weighted or unweighted) average of the remainder (3.14). *Assume* therefore that $\Delta\underline{x}$ is a random n-vector with zero-mean and covariance matrix$\sigma^2 Q_x$. The mean of $\underline{R}_2(x) = \frac{1}{2}\Delta\underline{x}^*\partial_{xx}^2 A(x)\Delta\underline{x}$ follows then as

$$\boxed{E\{\underline{R}_2^i(x)\} = \frac{1}{2}\sigma^2 trace[\partial_{xx}^2 A^i(x)Q_x], \quad i = 1, \ldots, m} \tag{3.18}$$

Note (compare also with (3.11)) that $E\{\underline{R}_2^i(x)\}$ describes the bias in $\underline{y}$ if $\underline{y}$ were computed as $\underline{y} = A(\underline{x})$. The measure (3.18) is very easy to compute if we take the identity matrix for Q_x. In the following sections we will see that (3.18) can also be used as an upperbound on the biases of the least-squares estimators.

3.4 The Bias of Least-Squares Parameters

In this section we will derive an approximation to the bias

$$b_{\hat{x}} = E\{\hat{\underline{x}}\} - x \tag{3.19}$$

of the least-squares estimator $\hat{\underline{x}}$ of the unknown parameter vector x of the nonlinear model

$$E\{\underline{y}\} = A(x) \quad , \quad E\{(\underline{y} - A(x))(\underline{y} - A(x))^*\} = \sigma^2 Q_y \tag{3.20}$$

The method of derivation is taken from (Teunissen 1985). We will assume that the least-squares estimator $\hat{\underline{x}}$ can be written as a smooth enough map of the random m-vector $\underline{y}$. Thus $\hat{\underline{x}} = \hat{x}(\underline{y})$. We also assume that $x = \hat{x}(E\{\underline{y}\})$. If we Taylorize the map $\hat{x}(.)$ at the mean $E\{\underline{y}\}$ of $\underline{y}$, we get an expansion in $\underline{e} = \underline{y} - E\{\underline{y}\}$:

$$\hat{\underline{x}} = x + \partial_y \hat{x}\underline{e} + \frac{1}{2}\underline{e}^*\partial_{yy}^2\hat{x}\underline{e} + \cdots \tag{3.21}$$

Since $E\{\underline{e}^*\partial_{yy}^2\hat{x}\underline{e}\} = \sigma^2 trace[\partial_{yy}^2\hat{x}Q_y]$, it follows upon taking the expectation of (3.21) that

$$b_{\hat{x}} \doteq \frac{1}{2}\sigma^2 trace[\partial_{yy}^2\hat{x}Q_y] \tag{3.22}$$

The problem is now to find the second order partial derivatives of the map $\hat{x}(.)$. These derivatives are found in the following way. We start from the orthogonality condition $e(\hat{\underline{x}}) \perp R(\partial_x A(\hat{x}))$, or

$$0 = \partial_x A(\hat{\underline{x}})^* Q_y^{-1} e(\hat{\underline{x}}) \tag{3.23}$$

A Taylor expansion of the righthand side of (3.23) at x gives the following expansion in $\Delta\underline{x} = \hat{\underline{x}} - x$:

$$\begin{aligned} 0 \;=\; & \partial_\alpha A(x)^* Q_y^{-1}\underline{e} + [\partial^2_{\alpha\beta} A(x) Q_y^{-1}\underline{e} - \partial_\alpha A(x)^* Q_y^{-1} \partial_\beta A(x)] \Delta\underline{x}^\beta \\ & + \tfrac{1}{2}[\partial^3_{\alpha\beta\gamma} A(x) Q_y^{-1}\underline{e} - 2\partial^2_{\alpha\beta} A(x) Q_y^{-1}\partial_\gamma A(x) - \partial_\alpha A(x)^* Q_y^{-1}\partial^2_{\beta\gamma} A(x)]\Delta\underline{x}^\beta \Delta\underline{x}^\gamma + \cdots \end{aligned} \tag{3.24}$$

where use is made of Einstein's summation convention. We now substitute our first expansion in $\underline{e}$, (3.21), into the above expansion in $\Delta\underline{x} = \hat{\underline{x}} - x$. The result is a new expansion in $\underline{e}$:

$$\begin{aligned} 0 \;=\; & [\partial_\alpha A(x)^* Q_y^{-1} - \partial_\alpha A(x)^* Q_y^{-1}\partial_\beta A(x)\partial_y \hat{x}^\beta]\underline{e} + \\ + \;& \underline{e}^* [Q_y^{-1}\partial^2_{\alpha\beta} A(x)\partial_y \hat{x}^\beta - \tfrac{1}{2}\partial_\alpha A(x)^* Q_y^{-1}\partial^2_\beta A(x)\partial^2_{yy}\hat{x}^\beta - \partial_y \hat{x}^\beta \partial^2_{\alpha\beta} A(x) Q_y^{-1}\partial_\gamma A(x)\partial_y \hat{x}^\gamma + \\ - \;& \tfrac{1}{2}\partial_y \hat{x}^\beta \partial_\alpha A(x)^* Q_y^{-1}\partial^2_{\beta\gamma} A(x)\partial_y \hat{x}^\gamma]\underline{e} + \cdots \end{aligned} \tag{3.25}$$

This expansion is identical to zero for all $\underline{e}$. Hence we may collect terms of the same order and set them to zero. For the first order term this gives:

$$\partial_y \hat{x} = [\partial_x A(x)^* Q_y^{-1}\partial_x A(x)]^{-1}\partial_x A(x)^* Q_y^{-1} = Q(x)\partial_x A(x)^* Q_y^{-1} \tag{3.26}$$

Note that when the map $A(.)$ is linear, substitution of (3.26) into (3.21) gives indeed the *linear* least-squares estimator of x.

From the second order term of (3.25) follows that

$$\begin{aligned} \tfrac{1}{2}\partial^2_{yy}\hat{x}^\gamma \;=\; & Q_y^{-1} Q(x)^{\gamma\alpha}\partial^2_{\alpha\beta} A(x) Q(x)^{\beta\delta}\partial_\delta A(x)^* Q_y^{-1} + \\ & - Q(x)^{\gamma\alpha}\partial^2_{\alpha\beta} A(x) Q_y^{-1} P_{\partial_x A(x)} Q(x)^{\beta\delta}\partial_\delta A(x)^* Q_y^{-1} + \\ & - \tfrac{1}{2} Q(x)^{\gamma\rho}\partial_\rho A(x)^* Q_y^{-1}(Q_y^{-1}\partial_\delta A(x) Q(x)^{\delta\alpha}\partial^2_{\alpha\beta} A(x) Q(x)^{\beta\epsilon}\partial_\epsilon A(x)^* Q_y^{-1}) \end{aligned} \tag{3.27}$$

Although this expansion looks rather complicated, fortunately it simplifies considerably when substituted into (3.22). When we substitute (3.27) into (3.22) the first two terms of (3.27) cancel and we finally get the expression sought

$$\boxed{\begin{aligned} b_{\hat{x}} &\doteq Q(x)\partial_x A(x)^* Q_y^{-1} b_y\,, \quad \textit{with} \\ b_y &= -\tfrac{1}{2}\sigma^2 \mathit{trace}[\partial^2_{xx} A(x) Q(x)] \end{aligned}} \tag{3.28}$$

This important and rather simple expression for the bias in the least-squares estimator $\hat{\underline{x}}$ has some interesting properties. First note that the m-vector b_y of (3.28) closely resembles the third measure of nonlinearity as proposed in the previous section, see equation (3.18). Secondly, note that the bias $b_{\hat{x}}$ in the least-squares estimator can be computed from the m-vector b_y, just like in the *linear* least-squares case the estimator $\hat{\underline{x}}$ is computed from $\underline{y}$. Hence, with an available standard least-squares software package the evaluation of the bias $b_{\hat{x}}$ becomes rather straightforward. Finally we remark that $b_{\hat{x}}$ can be given a simple geometric interpretation. As shown in (Teunissen 1984, 1985), $b_{\hat{x}}$ equals the weighted trace of the *Christoffel symbols of the second kind* and is therefore a measure of the "turning and twisting" of the coordinate lines in the manifold $A(x)$.

3.5 The Bias of the Least-Squares Residual Vector

The bias in the least-squares residual vector $\hat{\underline{e}}$ is defined as

$$b_{\hat{e}} = E\{\hat{\underline{e}}\} = E\{\underline{y} - A(\hat{\underline{x}})\} \tag{3.29}$$

Note that $b_{\hat{e}} = -b_{\hat{y}}$. If we let map $A(.)$ play the role of map $F(.)$ of section 2, it follows with (3.11) that

$$b_{\hat{y}} = -b_{\hat{e}} = \partial_x A(x) b_{\hat{x}} + \frac{1}{2}\sigma^2 trace[\partial^2_{xx} A(x) Q_{\hat{x}}] + \frac{1}{2} b^*_{\hat{x}} \partial^2_{xx} A(x) b_{\hat{x}} + \cdots \tag{3.30}$$

Since $b_{\hat{x}}$ is of the order σ^2 (see (3.28)), the first two terms on the righthand side of (3.30) are of the order σ^2 and the third term is of the order σ^4. Substitution of (3.28) into (3.30) gives therefore up to the order σ^2,

$$b_{\hat{y}} = -b_{\hat{e}} \doteq \partial_x A(x) Q(x) \partial_x A(x)^* Q_y^{-1} [-\frac{1}{2}\sigma^2 trace[\partial^2_{xx} A(x) Q(x)]] + \frac{1}{2}\sigma^2 trace[\partial^2_{xx} A(x) Q_{\hat{x}}]$$

And since $Q_{\hat{x}} \doteq Q(x)$ within the same approximation, the following expression for the bias in $\hat{\underline{e}}$ is obtained

$$\boxed{\begin{array}{rcl} b_{\hat{e}} & \doteq & P^{\perp}_{\partial_x A(x)} b_y \;, \quad with \\ b_y & = & -\frac{1}{2}\sigma^2 trace[\partial^2_{xx} A(x) Q(x)] \end{array}} \tag{3.31}$$

Again note that it is rather straightforward to evaluate the bias $b_{\hat{e}}$ with a standard least-squares software package.

The bias vector $b_{\hat{e}}$ can be given an interesting geometric interpretation in terms of the normal curvatures of manifold $A(x)$. This can be seen as follows. Recall that the principal normal curvatures $k^i_{n_j}, i = 1, \ldots, n$, are the eigenvalues of the generalized eigenvalue problem

$$| \, n_j^* Q_y^{-1} \partial^2_{xx} A(x) - \lambda Q(x)^{-1} \, | = 0$$

From this follows that

$$trace[n_j^* Q_y^{-1} \partial^2_{xx} A(x) Q(x)] = \sum_{i=1}^{n} k^i_{n_j} \tag{3.32}$$

If we let $n_j, j = 1, \ldots, m-n$, be an orthonormal basis of the orthogonal complement of the rangespace of $\partial_x A(x)$, we can write the projector $P^{\perp}_{\partial_x A(x)}$ as

$$P^{\perp}_{\partial_x A(x)} = \sum_{j=1}^{m-n} n_j n_j^* Q_y^{-1} \tag{3.33}$$

From (3.32) and (3.33) follows then that

$$P^{\perp}_{\partial_x A(x)} trace[\partial^2_{xx} A(x) Q(x)] = \sum_{j=1}^{m-n} n_j \sum_{i=1}^{n} k^i_{n_j}$$

and thus with (3.31) that

$$\boxed{b_{\hat{e}} \doteq -\frac{1}{2}\sigma^2 \sum_{j=1}^{m-n} n_j \sum_{i=1}^{n} k^i_{n_j}} \tag{3.34}$$

This result shows how the local geometry of the manifold $A(x)$ determines the bias in the least-squares residual vector $\hat{\underline{e}}$. Equation (3.34) also shows that the bias $b_{\hat{e}}$ is *invariant* under a change of variables.

3.6 On Scalar Measures of Biases

Apart from our results (3.28) and (3.31) for the biases in the parameter vector and residual vector, it is also useful to have global scalar measures of biases available which summarize the bias present in the nonlinear model. In order to descern the significance of the biases it was proposed in (Teunissen and Knickmeyer 1988) to weight the biases in the parameters and residuals with the inverses of $\sigma^2 Q(x)$ and $\sigma^2 Q_y$ respectively. The proposed global bias-measures read therefore

$$\begin{array}{rlcl} a) & \| b_{\hat{x}} \|^2_{Q(x)} & = & \sigma^{-2} b_{\hat{x}}^* Q(x)^{-1} b_{\hat{x}} \\ b) & \| b_{\hat{e}} \|^2_{Q_y} & = & \sigma^{-2} b_{\hat{e}}^* Q_y^{-1} b_{\hat{e}} \end{array} \qquad (3.35)$$

Substitution of (3.28) and (3.31) into (3.35) gives

$$\boxed{\begin{array}{rlcl} a) & \| b_{\hat{x}} \|^2_{Q(x)} & = & \| P_{\partial_x A(x)} b_y \|^2_{Q_y} \\ & & & \\ b) & \| b_{\hat{e}} \|^2_{Q_y} & = & \| P^{\perp}_{\partial_x A(x)} b_y \|^2_{Q_y} \end{array}} \qquad (3.36)$$

This result shows that the bias in the parameters is determined by the *tangential* component of the Hessian of $A(x)$, whereas the bias in the residuals is determined by the *normal* component of the Hessian of $A(x)$. Compare this with our discussion in chapter two, section four on the local rate of convergence of the Gauss-Newton method.

Since $P^{\perp}_{\partial x A(x)} = I - P_{\partial_x A(x)}$ is an orthogonal projector, it follows from (3.36) and the *Pythagorean theorem* that

$$\boxed{\| b_y \|^2_{Q_y} = \| b_{\hat{x}} \|^2_{Q(x)} + \| b_{\hat{e}} \|^2_{Q_y}} \qquad (3.37)$$

This result shows that the scalar bias-measures of (3.36) are bounded from above by $\| b_y \|^2_{Q_y}$. Thus

$$\boxed{\| b_{\hat{x}} \|^2_{Q(x)} \leq \| b_y \|^2_{Q_y} \quad and \quad \| b_{\hat{e}} \|^2_{Q_y} \leq \| b_y \|^2_{Q_y}} \qquad (3.38)$$

Hence the relatively easy computable scalar $\| b_y \|^2_{Q_y}$ can be used as a first indicator for deciding whether the bias due to nonlinearity in $\hat{x}$ and $\hat{e}$ is significant or not. In a somewhat similar way we find with the help of the Cauchy-Schwarz inequality for the individual bias components the upperbounds

$$\boxed{| b^{\alpha}_{\hat{x}} | \leq \sigma_{\hat{x}^{\alpha}} \cdot \| b_{\hat{x}} \|_{Q(x)} \leq \sigma_{\hat{x}^{\alpha}} \| b_y \|_{Q_y} \, , \quad \alpha = 1, \ldots, n} \qquad (3.39)$$

and

$$\boxed{| b^{i}_{\hat{e}} | \leq \sigma_{\hat{e}^{i}} \| b_{\hat{e}} \|_{Q_y} \leq \sigma_{\hat{e}^{i}} \| b_y \|_{Q_y} \, , \quad i = 1, \ldots, m} \qquad (3.40)$$

Acknowledgements

This work was partially funded by the Royal Netherlands Academy of Sciences (KNAW) and the Netherlands Geodetic Commission (NGC). The expert typesetting of this work by Ms. ir. H.M.E. Verhoef using the program LaTeX is greatly appreciated.

Bibliography

Amari, S.I., O.E. Barndorff-Nielsen, R.E. Kass, S.L. Lauritzen and C.R. Rao (1987): *Differential geometry in statistical inference*, Institute of Mathematical Statistics, Vol. 10.

Bähr, H.G. (1985): *Second order effects in the Gauss-Helmert model*, 7th International Symposium on Geodetic Computations, Cracow, Poland.

Bähr, H.G. (1988): *A quadratic approach to the non-linear treatment of non-redundant observations*, Manuscripta Geodaetica, Vol. 13, No. 3, pp. 191-197.

Bierens, H.J. (1984): *Robust Methods and Asymptotic Theory in Nonlinear Econometrics*, Lecture notes in Econometrics and Mathematical Systems, Springer-Verlag, Vol. 192.

Blaha, G. (1987): *Non-linear parametric least-squares adjustment*, Nova University Oceanographic Center, Scientific Report No. 1.

Bopp, H. and H. Krauss (1978): *Strenge oder herkömmliche bedingte Ausgleichung mit Unbekannten bei nichtlinearen Bedingungsgleichungen?*, Allgemeine Vermessungs-Nachrichten, Vol. 85, pp. 27-31.

Borre, K. and S.L. Lauritzen (1980): *Some Geometric Aspects of Adjustment*, In: Festschrift to Torben Krarup, Ed. E. Kejlso et al., Geod. Inst., No. 58, pp. 70-90.

Cauchy, A. (1847): *Méthode générale pour la résolution des systèmes d'équations simultanées*, C.R. Acad. Sci. Paris, Vol. 25, pp. 536-538.

Goldfeld, S.M., R.E. Quandt and H.F. Trotter (1966): *Maximization by quadratic hill climbing*, Econometrica, Vol. 34, pp. 541-551.

Grafarend, E.W. and B. Shaffrin (1989): *The Geometry of Nonlinear Adjustment - the Planar Trisection Problem*, In: Festschrift to Torben Krarup, Ed. E. Kejlso et al., Geod. Inst., No. 58, pp. 149-172.

Jennrich, R.I. (1969): *Asymptotic Properties of Nonlinear Least Squares Estimators*, The Annals of Mathematical Statistics, Vol. 40, pp. 633-643.

Jeudy, L.M.A. (1988): *Generalyzed variance-covariance propagation law formulae and application to explicit least-squares adjustments*, Bulletin Geodesique, Vol. 62, No. 2, pp. 113-124.

Krarup, T. (1982): *Non-linear adjustment and curvature*, In: Forty Years of Thought, Delft, pp. 145-159.

Kubik, K.K. (1967): *Iterative Methoden zur Lösung des nichtlinearen Ausgleichungsproblemes*, Zeitschrift für Vermessungswesen, Vol. 91, No. 6, pp. 214-225.

Kubik, K.K. (1968): *On the efficiency of least-squares estimators in non-linear models*, Statistica Neerlandica, Vol. 22, No. 1, pp. 33-36.

Levenberg, K. (1944): *A method for the solution of certain nonlinear problems in least-squares*, Quart. Appl. Math. Vol. 2, pp. 164-168.

Marquardt, D.W. (1963): *An algorithm for least-squares estimation of nonlinear parameters*, J. SIAM II, pp. 431-441.

Ortega, J.M. and W.C. Rheinboldt (1970): *Iterative solution of nonlinear equations in several variables*, Academic Press, New York.

Pazman, A. (1987): *On Formulas for the Distribution of Nonlinear L.S. Estimates*, Statistics, Vol. 18, No. 1, pp. 3-15.

Pope, A. (1972): *Some pitfalls to be avoided in the iterative adjustment of non-linear problems*, Proceedings of the 38th Annual Meeting, American Society of Photogrammetry.

Pope, A. (1982): *Two approaches to non-linear least-squares adjustments*, The Canadian Surveyor, Vol. 28, No. 5, pp. 663-669.

Saito, T. (1973): *The non-linear least-squares of condition equations*, Bulletin Geodesique, Vol. 110, pp. 367-395.

Schaffrin, B. (1985): *A note on linear prediction within a Gauss-Markoff model linearized with respect to a random approximation*, Proc. First Tampere Sem Linear Models, Univ. Tampere, pp. 285-300.

Schek, H.J. and Ph. Maier (1976): *Nichtlineare Normalgleichungen zur Bestimmung der Unbekannten und deren Kovarianzmatrix*, Zeitschrift für Vermessungswesen, Vol. 101, No. 4, pp. 140-159.

Schmidt, W.H. (1982): *Testing Hypothesis in Nonlinear Regressions*, Math. Operations forsch. Statist., Secr. Statistics, Vol. 13, No. 1, pp. 3-19.

Teunissen, P.J.G. (1984): *A note on the use of Gauss' formula in nonlinear geodetic adjustments*, Statistics and Descisions, No. 2, pp. 455-466.

Teunissen, P.J.G. (1985): *The geometry of geodetic inverse linear mapping and nonlinear adjustment*, Netherlands Geodetic Commission, Publications on Geodesy, New Series, Vol. 8, No. 1, Delft.

Teunissen, P.J.G. (1985): *Nonlinear adjustment: An introductory discussion and some new results*, In: Proceedings of SSG 4.56, Workshop Meeting, Ghania, Greece, pp. 10-12.

Teunissen, P.J.G. (1988): *First and Second Order Moments of Nonlinear Least-Squares Estimators*, will be published in Bulletin Geodesique.

Teunissen, P.J.G. and E.H. Knickmeyer (1988): *Nonlinearity and Least Squares*, CISM Journal ACSGC, Vol. 42, No. 4, pp. 321-330.

Teunissen, P.J.G. (1989): *A Note on the Bias in the Symmetric Helmert Transformation*, In: Festschrift to Torben Krarup, Ed. E. Kejlso et al., Geod. Inst., No. 58, pp. 335-342.

Teunissen, P.J.G. (1989): *Nonlinear Inversion of Geodetic and Geophysical Data: Diagnosing Nonlinearity*, Invited paper, The Ron Mater Symposium on Four-Dimensional Geodesy, Sydney, Australia.