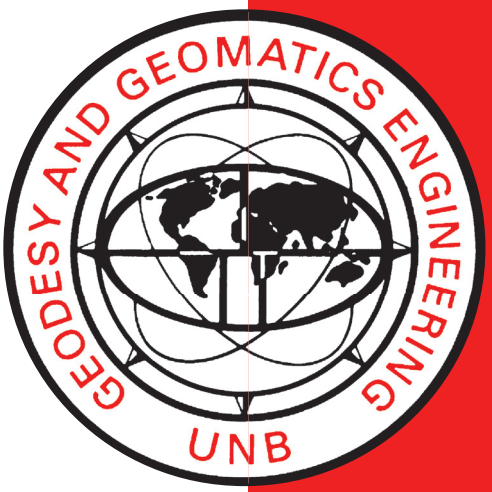
ON THE DETECTION OF OUTLIERS AND THE DETERMINATION OF RELIABILITY IN GEODETIC NETWORKS

M. KAVOURAS

November 1982



TECHNICAL REPORT
NO. 87

PREFACE

In order to make our extensive series of technical reports more readily available, we have scanned the old master copies and produced electronic versions in Portable Document Format. The quality of the images varies depending on the quality of the originals. The images have not been converted to searchable text.

ON THE DETECTION OF OUTLIERS AND THE DETERMINATION OF RELIABILITY IN GEODETIC NETWORKS

Marinos Kavouras

Department of Geodesy and Geomatics Engineering
University of New Brunswick
P.O. Box 4400
Fredericton, N.B.
Canada
E3B 5A3

November 1982
Latest Reprinting December 1994

PREFACE

This report is an unaltered version of the author's M.Sc.E. thesis of the same title, which was submitted to this Department in June, 1982.

This work contributes to research on deformation analysis, which is carried out by the Fredericton, group of the Fédération Internationale des Géomètres "ad hoc" committee, concerning the analysis of networks and the screening of incompatible observations.

The thesis advisor for this work was Dr. Adam Chrzanowski and financial support was partially from the Natural Sciences and Engineering Research Council of Canada and from the University of New Brunswick.

The assistance rendered by others is detailed in the acknowledgements.

ABSTRACT

The detection and removal of all gross errors or local systematic effects from geodetic data is of paramount importance when the quality of a Least Squares Estimation is concerned. In recent years, the above need, as well as the decrease of the computational cost compared to the cost of observational procedures, has led to the development of a number of sophisticated techniques for the systematic tackling of the problem, based on statistical tests. Several approaches are reviewed in this study, but the main weight is given to the most systematic and effective up to now - the post-adjustment techniques. The use of these techniques, and the analysis and comparison of their philosophies and sensitivities, are illustrated by simple numerical examples. A systematic strategy for error detection and elimination is proposed, with special emphasis on survey networks.

The finite sensitivity of the employed techniques may leave undetectable outliers in the model. Their magnitude as well as their undesirable effect on the final solution are assessed by the concepts of internal and external reliabilities.

Problems encountered with the detection of small gross errors and the resistance of networks to distortions caused by the presence of inconsistent observations, are also illustrated by analysing the strength of a real geodetic network. Certain limitations in the classical approaches have led to the study of alternatives which are more robust to outliers than is the Least Squares Method, an overview of which is also given.

Finally, recommendations and guidelines are given to practicing surveyors, concerning the application of the above ideas in the design and analysis of surveying projects.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOTATION

A	:	design matrix
A_k	:	design matrix of the desired parameters
A_t	:	design matrix of the nuisance parameters
C_k/Q_k	:	covariance/cofactor matrix of desired parameters
C_ℓ/Q_ℓ	:	covariance/cofactor matrix of observations
$C_{\hat{\ell}}/Q_{\hat{\ell}}$	:	covariance/cofactor matrix of adjusted observations
C_v/Q_v	:	covariance/cofactor matrix of the residuals
C_x/Q_x	:	covariance/cofactor matrix of estimated parameters

In every case $C_i = \sigma_o^2 \cdot Q_i$

$c_{(i)}$	:	critical value of the (i) distribution
d	:	partial quadratic form of the suspected residuals
$E[i]$	:	expected value of i
F_{r_1,r_2}	:	F-distribution
f_i	:	i^{th} misclosure
$f(i)$	:	probability density function of the (i) distribution
H_o	:	null hypothesis
H_a	:	alternative hypothesis
h_{ℓ_i}	:	reliability coefficients

$\underline{k}$	:	vector of unknown desired parameters
$\hat{\nabla k}$ (NABLA k)	:	influence of gross errors on the unknown desired parameters
$\underline{\ell}$	:	vector of observations
$\underline{\nabla\ell}$ (NABLA ℓ)	:	gross error part of $\underline{\varepsilon}$
$M = Q_v P$	:	first characteristic idempotent matrix
$N = Q_x^{-1}$	:	matrix of normal equations
$N(0,1)$	:	normal distribution
n	:	number of observations
$P = Q_\ell^{-1} = \sigma_o^2 \cdot C_\ell^{-1}$	:	weight matrix of the observations
p_i	:	diagonal element of the weight matrix (double subscript is used for the off-diagonal elements)
$Q_i = C_i/\sigma_o^2$	:	cofactor matrix of i
q_i	:	i^{th} diagonal element of cofactor matrix Q
q^2	:	total quadratic form of the residuals
$\dot{q}^2$	:	partial quadratic form of the consistent residuals
R	:	correlation matrix
r	:	total redundancy
r_i	:	redundancy numbers
$\bar{r} = r/n$	:	relative redundancy
$\underline{t}$	:	vector unknown nuisance parameters
t_r	:	t-distribution
$U = Q_\ell \hat{P}$	:	second characteristic idempotent matrix

$u = \Sigma u_i$	:	number of unknown parameters
$\bar{u} = u/n$	:	relative absorption
u_i	:	absorption numbers
$u_k = \Sigma u_{k_i}$	:	number of desired unknown parameters
u_{k_i}	:	absorption numbers for the desired parameters k
$u_t = \Sigma u_{t_i}$	:	number of nuisance unknown parameters
u_{t_i}	:	absorption numbers for the nuisance parameters t
$\underline{v}$	:	vector of residuals
v_i	:	i^{th} residual
$\underline{\nabla v}$ (NABLA v)	:	influence of $\underline{\nabla \ell}$ on the residuals
w_i	:	w-statistic for the "Data Snooping" test
$\underline{x}$	:	vector of unknown parameters
$\hat{\nabla x}$ (NABLA x)	:	influence of gross errors on the unknown parameters

---

α	:	significance level (generally)
α_o	:	significance level for one-D tests
$(1-\alpha)$	:	confidence level
β_o	:	power of the test *

---

* Although in most of the statistical textbooks the Greek letter "β" (beta) has been selected to denote the type II error, in the ß-method of testing (ch. 5), the reverse notation has been used, the one indicated above, i.e., β: power of the test; adopted by many others later on. Due to the use of nomograms (app. III), the same notation has also been used here but this remark should be kept in mind to avoid confusion.

$(1 - \beta_o)$	:	the probability of making the type II error
$\delta = \sqrt{\lambda}$	:	difference between two hypotheses
$\underline{\varepsilon}$	:	vector of true observational errors
$\underline{\varepsilon}_r$	:	random part of $\underline{\varepsilon}$
κ	:	expansion factor
λ, λ_o	:	non centrality parameter/boundary value
$\overline{\lambda}_{i,o}$	:	general variate for the external reliability
σ_o^2	:	a priori variance factor
$\hat{\sigma}_o^2$	:	a posteriori variance factor
σ_i^2	:	i^{th} diagonal element of a covariance matrix C
τ_i	:	τ-statistic for the "tau-test"
τ_r	:	τ-distribution (tau-distribution)
χ^2	:	χ^2-distribution

ABBREVIATIONS

Appendix	:	app.
Chapter	:	ch.
Equation	:	eq.
Figure	:	fig.
Section	:	sect.

ACKNOWLEDGEMENTS

This thesis had its inception in January 1981, obtained a more distinct form during the spring of the same year and it was completed one year later.

I wish to express my deep indebtedness to my supervisor Dr. Adam Chrzanowski for his suggesting this topic, for his continuous assistance and guidance, as well as for all the fruitful discussions we had and the genuine interest that he showed for the duration of my work.

My thanks are also given to Drs. W. Faig of the Department of Surveying Engineering, R. Mureika of the Department of Mathematics and Statistics, both at U.N.B., and W. Niemeier of the University of Hannover, for their constructive criticism in reviewing the original manuscript.

Discussions with several people have immensely assisted me during the course of this study. I thank Dr. W. Knight of the Department of Mathematics and Statistics and of the School of Computer Science at U.N.B., Dr. W. Förstner of the University of Stuttgart, Mr. J.J. Kok and Dr. J.E. Alberda of the Delft Geodetic Institute, Dr. P. Stefanovic of I.T.C., The Netherlands, and Dr. D.D. Balodimos of the National Technical University of Athens.

Special thanks are also given to my colleagues in the Department of Surveying Engineering, where there was always an atmosphere of enthusiatic learning and harmonious co-operation. In particular, I wish to extend my deep appreciation to Mr. Chen Yong Qi for his invaluable assistance; together with that of an excellent co-worker - James Secord.

For their precious company and friendship, when I was either rambling in the realm of abstractions or concentrating in the real world, and for the endless nocturnal discussions in an attempt to evaluate all of our steps, I am grateful to my colleagues Costas Armenakis and Stelios Mertikas; without omitting Demitris Paradisis of the National Technical University of Athens, for his distant but continuous encouragement. It is truly a privilege to have friends like these.

The endeavours of Ms. Katie Wilson in proofreading the manuscript and of Mrs. Ann Arnold in typing it, are also deeply appreciated.

Finally, I owe an awful lot to a little leprechaun, whose passing engraved beautifully and indelibly my life during the last two years, by revealing so many things which cannot be found in books.

1

INTRODUCTION

In the classical adjustment theory most of the research which has been done has concerned only the treatment of random and systematic errors in the observations. In most of the geodetic text books only one or two pages are devoted to the existence of gross errors and the possible solutions to the problem. Most often it is implied that all gross and systematic errors which have not been considered in the functional model, should be eliminated prior to the adjustment. This is required due to the high sensitivity of the Least Squares method to outliers which can dangerously lower the quality of the results. At this stage the quality of the results is vaguely associated with the concepts of precision and accuracy *(Mikhail, 1979)*.

The techniques for detection and elimination of gross errors consisted of all sorts of empirical checks, based on simple conditions which the observations should fulfill, governed by the experience and the

intuition of the investigator. These simple checks were neither very systematic nor very sensitive but their ability to locate large errors was and still is remarkable. A look at the residuals after the adjustment would not reveal much and if something seemed to be wrong, additional checks on the original observations were performed. This rather frugal analysis was also dictated by the absence of large computers or by the high computational cost.

However, in recent years the situation has changed profoundly. The above, requiring the investigator to have both experience and intuition as well as a good knowledge concerning the data acquisition procedures, is a "luxury" which unfortunately is not always available. On the other hand, the cost of employing experts to analyze the data is getting continuously higher than the computational cost of processing them. In addition to that, a non-systematic investigation founded only on intuition and simple checks on the observations, has its limitations and proves to be inadequate when the amount of data and consequently the number of combinations and checks on the observations one can perform, increases remarkably.

Hence, the apparent need for a very systematic procedure, which would not only detect very small gross or systematic errors (highly sensitive), but also provide us with a measure of the efficiency of the method, led to the development of a number of quite sophisticated techniques based on statistical tests. These techniques also made possible an estimation of how reliable the results of the adjustment are, that is, if small changes in the observations cause relatively small changes in the results of the adjustment *(Blais, 1976)*.

The initiator of this philosophy was *Baarda (1965, 1967, 1968, 1976, 1979)*, and through his theory the concepts of quality, accuracy, precision and reliability obtained a more distinct definition.

First of all, one should know approximately how the observations are affected by the different kinds of error. The question "how important is it to detect errors of small magnitude in the data?" has to be answered, considering both nature of the project and economy. After that, it is known what is to be expected from an error detection procedure, and criteria for the minimum expected precision of the observations and the maximum effect on the final results that can be tolerated, have to be established.

Since statistical tests are normally employed for error detection, it is essential to decide (depending again on the cost of a wrong decision), which of the two types of error is more serious to commit: To reject something that should not be rejected, or to accept something that should be rejected?

Related to the Least Squares method, techniques can be generally distinguished as pre- and post- adjustment techniques. Up to now, post-adjustment techniques have proven to be more systematic and powerful *(Stefanovic, 1978)*, as parametric adjustment is more often used than the conditional adjustment (ch. 4). This thesis is mainly concerned with the analysis of post-adjustment techniques for the detection of gross errors.

It is true that statistical literature abounds of different approaches for the detection of outliers in the data. Geodesists and photogrammetrists, being interested in this field, have started adopting and adapting these statistical methods to serve the analysis of geodetic-type problems. However, it is also true that statistical references

are written in a difficult language for surveyors.

On the other hand statistical methods are only tools, and logical and thoughtful interpretation of their results is always needed. If the proposed method is not sufficiently understood one has to choose between two evils: Either the method will not be applied at all (with all the consequences of this non-application) or the method will be applied blindly, leading to misinterpretations and wrong decisions.

In reality however, although there is an urgent need of a systematic analysis for the detection of gross errors in the data, there is not equal interest in a profound understanding of the entire theory behind the proposed technique, expecially by practicing surveyors. What they need are simple examples and a good understanding of the physical meaning of the testing procedure. Due to absence of such examples and useful guidelines for surveyors, all the methods given below have been presented in the simplest possible way regardless of the degree of their sophistication.

Moreover, looking at the available literature, one can notice a lack of consistent and commonly accepted notation and terminology, so that a great amount of time is spent in trying to find the relation among different symbols and terms. This is one of the main reasons that, although most of the presented methods have been under development for more than 17 years, they are still not very popular but mainly comprehensible to the large majority of the surveying profession. To sidestep this problem, the notation and terminology adopted in this thesis are those most often encountered in the geodetic literature, and a list of them is given.

It should be also mentioned that a meaningful comparison or link between the different approaches is not that easy. The main cause of the difficulty is not only the conceptual or structural difference of the statistical tests themselves, but also the differing philosophies behind them, concerning the choice of the probability levels, which can very easily lead to different decisions. Comparability was achieved not only by appropriately choosing the significance levels but also by always considering the philosophy of the methods whose sensitivities were being compared.

The localization of gross errors, especially if there is more than one, is still very problematic, since most of the existing methods were designed for detection purposes only. Thus it was one of the main tasks of this thesis to examine the localization problems and develop a systematic strategy for a successful treatment of gross errors. This strategy was designed to circumvent most of the problems caused by the weakness of the model and not by any inefficiency of the statistical methods.

Finally, due to certain limitations in the classical approaches, some new and more robust alternatives were examined, but the further development of such techniques that is still required was out of the scope of this study.

All the above needs, defined the purpose of this thesis, which is:

- To review and relate the different approaches for detection of gross or systematic errors, into common notation and "language" understandable by practicing surveyors, giving also simple practical examples.

- To compare the different philosophies which govern each technique and perform a critical evaluation of the methods.
- To develop a step-by-step procedure for detection, localization and elimination of gross errors in the observations, especially in small survey networks (low redundancy). This procedure has particular value in cases where outliers of small magnitude must be detected and the high accuracy of the survey is more important than the cost of processing the data.
- To explain the concepts of quality and reliability of the results of the adjustment, and illustrate their importance with an example in a real geodetic network.

First, a review of the nature of the errors in survey projects is given, and their role in the Least Squares method is studied, together with the statistical properties of the quantities involved *(second and third chapters)*.

In the analysis of different techniques, where the largest weight is given to the post adjustment ones *(fifth chapter)*, a detailed analysis with examples and a proposed strategy are presented.

The sensitivity of the above approaches as well as the influence of undetectable errors on the results of the adjustment are the major concern of the *sixth chapter*.

In the complete example of a real and quite precisely measured geodetic network that follows *(seventh chapter)*, problems encountered with the choice of the right hypothesis are illustrated, and the geometrical strength of the network is analyzed employing all the previous concepts.

In the *eighth chapter*, a few new methods for the treatment of gross errors in the data are introduced, which however require further development.

The modern definition of the quality of networks (as it was introduced by Baarda (1976)), and a few economical aspects are concern of the *ninth chapter*. Finally, some guidelines, recommendations and warnings are given for use by the land surveyor.

2

REVIEW AND ANALYSIS OF ERRORS IN SURVEYING PROJECTS

The basic assumption on which the theory of the Least Squares Estimation has been founded is that all the gross and systematic errors have been eliminated before the adjustment is performed, and only random errors affect the data.

The local and large disturbances are considered as gross errors, blunders or outliers, whereas smaller and global deviations are considered as systematic errors. Although these concepts look quite familiar, there are cases where clear distinction between local systematic errors and outliers of small magnitude cannot be made. Both kinds of errors have the same effect on the observations.

Generally, errors can be assigned to the functional or to the stochastical model according to their origin which, however, usually is unknown. It is true that the functional model can be extended in many cases, to include additional parameters in order to

take into account some of the systematic effects. This refinement has been a major research direction in recent years, mainly in the Photogrammetric field *(e.g., Moniwa, 1977; El-Hakim, 1979)*.

Nevertheless, systematic effects or gross errors depict the inability of the functional model to describe the real situation. Both biases are considered here as errors in the deterministic model which means that they are treated jointly. In point of fact, there are cases where these errors may be assigned to the stochastical instead of the functional model, and therefore the former has to be refined. This has been attempted by diminishing the weights of the erroneous observations instead of eliminating them (sect. 8.4), or by other methods.

The magnitude and the frequency of errors is also a problem to be considered when an error detection strategy is followed *(Förstner, 1981.a)*. It is not a problem to recognize outliers of large magnitude when they occur, since they depict local variations only. They may however become a problem if their magnitude is such as to destroy the linearization of the observation equations, so that the solution does not converge or the statistical tests cannot be used. These large gross errors have to be detected and eliminated through pre-adjustment (ch. 4), or other techniques.

The main goal here is the detection and elimination of errors of small magnitude. It should be realized that there is no clear boundary line between large random errors and very small outliers. First, an observation may be outlying compared to some a priori set criteria on its own precision. Secondly, an observation may be inconsistent in context of the rest of the observations (how well they

fit into the model). Most of the post-adjustment techniques base their philosophy on these two criteria; therefore, an observation may be considered as an outlier or not, depending on the sensitivity of the selected technique and the probability levels chosen. Due to the above criteria the sensitivity of the techniques is finite and thus no arbitrarily small gross errors can be detected.

There are not objective criteria about the smallest magnitude of gross errors, and they cannot be established easily. However, one could bound the smallest size of an outlier from the degree of its influence on the final results of the adjustment. The maximum effect of an observational error on the solution which can be tolerated should be established using a priori geodetic criteria only. The idea is directly connected to the concept of the reliability of the results (ch. 6).

The most systematic way to detect the presence of large errors is to perform a proper statistical test. It is also of major importance to have a measure of its sensitivity.

As it is known, a statistical test involves a null hypothesis H_o, which is a statement concerning the values which the population parameters are hypothesized to have. For every null hypothesis there is an infinite number of alternative hypotheses which assume different values for the population parameters. The null hypothesis is the reference level from which any deviation of the different alternative hypotheses has to be detected by statistical tests. In fact the null hypothesis often expresses what one hopes to reject, i.e. the opposite of what one would really wish to be the case.

The inquirer cannot confirm or refute the null hypothesis

since he does not have complete information about the entire population. The only thing he can do, is to report the probability of an observed outcome assuming that the null hypothesis is true. This can be done using test statistics, calculated by the experimental data under the null hypothesis.

The smaller the probability the lower the credibility of the null hypothesis. If this probability becomes sufficiently small, the null hypothesis may be rejected but not confuted. On the other hand, if the probability of the specific observed outcome is not enough to sufficiently reduce the credibility of the null hypothesis, it is said that "H_o is accepted". This does not imply that H_o is true but only that the available data cannot convince us that H_o is false. Therefore it would be more precise to say: "H_o is not rejected" (*Hogg and Craig, 1978*).

Some definitions are given now:

Type I Error: the rejection of the null hypothesis H_o, when H_o is actually true. The probability of committing this type of error is called "significance level" α. The probability of making the correct decision is the"confidence level"$(1-\alpha)$.

Type II Error: the acceptance of the null hypothesis when in fact is false (H_a is true). The probability of committing this type of error is $(1-\beta)$. The probability of making the correct decision is the "power of the test" β, i.e. the power of detecting a difference

δ between H_o and H_a when the test is carried out at a significance level α, with a given sample size.

If the detection of smaller differences δ between the two hypotheses is desirable, the price that one has to pay is the increase of the sample size. If, however, the observations have been already obtained, which is the usual case in Geodesy and Photogrammetry, the power of the test β indicates the smallest difference δ that can be detected if the test has been executed at a significance level α (ch. 6).

It is obvious (fig. 2.1), that fixing the smallest difference between the two hypotheses, type I and type II errors cannot both be decreased.

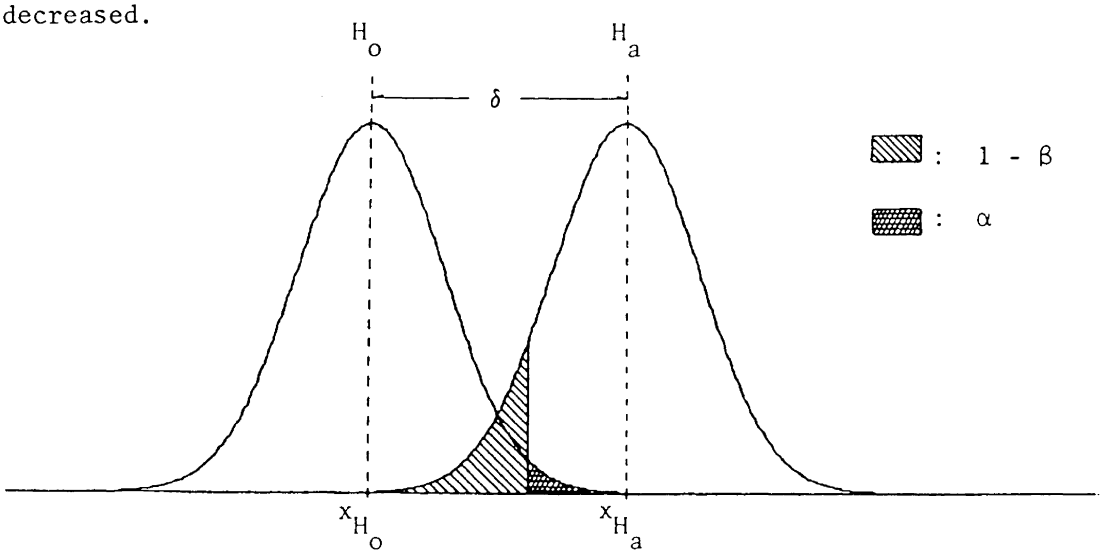


Figure 2.1: Variation of Type I and Type II Errors

A very small signigicance level expresses a reluctance to reject the null hypothesis unjustly *(Blais, 1976)*. On the other hand if the null hypothesis is false the type II error that is committed is quite large. In decreasing the probability of committing both types of error δ increases; hence the difference between H_o and H_a should

be sufficiently large to be detected.

It can be said that the selection of the significance level has been a matter of personal taste since it reflects only a degree of reluctance, descretion or conservatism. Objective criteria are generally unavailable. However, in our field the probability levels can be chosen, if tolerances are established a priori concerning the maximum effect of errors on the results if the observations have been executed with care *(Stefanovic, 1980)*.

Anyway, the selection of the probability levels "α" and "$1-\beta$" should be based on the cost of committing either type I or type II error. Thus depending on the purpose of testing (detection of gross errors, analysis of deformations, etc.) and on the cost of making a wrong decision of either type, a compromise between the confidence level and the power of the test can be reached.

In small geodetic networks where the redundancy is quite limited choosing a lower significance level is suggested, otherwise too many rejections will take place, with all the consequences of such a course. Moreover it is recommended that automatic rejection should never be executed without any justification.

On the other hand there are cases where the expense of recollecting data in a project proves to be not only uneconomical, but also impossible to repeat. In this situation, while the first tests are applied in the field, attention should be paid to the type II error and additional observations can be performed if needed. Thus sufficient redundancy ensures the detection of fairly small differences (due to gross or systematic errors) between H_o and H_a, without at the

same time decreasing the significance level.

The problem of choosing the right alternative hypothesis H_a at the time, is quite difficult to solve since it may appear that more than one H_a satisfy the tests when H_o is rejected. The design of H_a requires a good knowledge of the collection of the data (e.g. survey procedure), experience, and sometimes intuition which unfortunately are not always available. This is the reason why more systematic and automatic techniques are employed, which however introduce much simpler alternative hypotheses. This inevitably may bring a vagueness into the decision making.

The consideration of both types of error is mentioned further below when the different techniques are presented (ch. 5); in addition, relevant examples are given.

It should be mentioned here that gross errors in the data can be successfully treated only if the stochastic properties of the observations are sufficiently known, and if an adequate number of them is statistically consistant, also forming a well conditioned algebraic system. Moreover it is assumed that the number of gross errors does not exceed the redundancy of the system *(Stefanovic, 1978)*.

A full algorithm for a systematic blunder detection would inevitably require the use of a large computer. The additional computations needed for this purpose prove to be less expensive than repeating the observations. The latter is not desirable for two more reasons. First, no-one can guarantee that other gross errors will not occur, and secondly, in many cases observations cannot be repeated due to their own nature and the peculiarity of the project. Re-observation is inevitable only in the case of erroneous but indispensable

observations.

The designed blunder detection technique should take care of the elimination of large gross errors at an early stage, so that problems with the linearization or the statistical assumptions will be avoided. Moreover it is desired to enable not only the detection but also the localization of gross errors, which has been up to now the Achilles' heel of all the post adjustment techniques.

Finally, it should be mentioned that a large number of errors which owe their existence in incorrect transferring, copying, or punching of the data, as well as in programming errors, etc., not being of a statistical nature, should be also detected by a blunder detection technique *(Pope, 1976)*.

Before these techniques will be presented it is necessary to analyze the influence of gross errors on the Least Squares Estimation.

3

LEAST SQUARES AND OUTLIERS

3.1 Least Squares and Statistics

The Least Squares method is only one of the estimation techniques. It has been widely used in most practical situations because of its simplicity and also because complete statistical information is generally unavailable. Whereas the other estimation techniques may require a more complete statistical model, the Least Squares Estimation requires only means, variances and covariances.

The assumption that all the systematic errors and biases have been removed and only random errors are left which have a Gaussian distribution, is not necessary for the L.S. Estimation itself but is is required by virtue of the statistical hypothesis testing.

The L.S. estimates for the unknown parameters x are the minimum variance estimates of x regardless of the distribution of the observational or other errors *(Blais, 1976)*. The fact that the statistical information about the estimated parameters is inadequate (their probability distribution is not known), does not allow any parametric testing on them.

The questionable assumption is that the observational errors are considered as unbiased. Any presence of bias is transferred to the estimation of the unknown parameters and cannot be generally detected through their examination.

The estimated variance factor is also unbiased and the distribution of the observational errors is assumed to be normal for the purpose of hypothesis testing only.

As far as the actual distribution of the geodetic observation of errors is concerned, the opinions conflict. It is true from experience that observations in geodesy are most often normally distributed or deviate only slightly *(Baarda, 1976)*. It is also true from experiments that the shape of the distribution differs somewhat from the normal one since it lacks the tail areas (they are cut off). The deviation between the actual and the theoretical shape is not very large *(Stefanovic, 1980)* but it should be tested whenever possible. This difference should be kept in mind because it shows that observational errors cannot exceed some certain values defined by the reality of the specific geodetic experiment. If an error larger than this boundary value takes place, it should be detected regardless of the chosen significance level.

The covariance matrix of the residuals C_v is singular

therefore it cannot be inverted (eq. I.10) to give solution for the true errors. The covariance matrix of the adjusted observations $C_{\hat{\ell}} = (AQ_xA^T)\cdot\sigma_o^2$, used only infrequently, indicates the difference between the covariance matrix of the observations C_ℓ and of the residuals C_v ,

$$C_v = C_\ell - AC_xA^T = C_\ell - C_{\hat{\ell}} \qquad \text{(eq.I.7)}$$

It is important to realize that although the C_ℓ matrix is diagonal, i.e., the observations are considered as uncorrelated, the residuals in general are always correlated. The latter comes from the model itself which produces the spurious correlation.

To have a diagonal covariance matrix for the unknown parameters C_x, it is required that the column vectors of A are orthogonal. This happens, for example, if each observation is used for the determination of only one unknown parameter, that is, each row of the design matrix A has only one non-zero element. To have a diagonal covariance matrix for the residuals C_v, it is additionally required to have a unique solution of the problem; namely the design matrix should be diagonal.

The higher the correlation among the residuals the more intricate the linear dependance of the residuals on the observations. Correlation equal to zero means that a variation of one observation would affect only one unknown parameter and the corresponding residual, whereas full C_v matrix simply implies that a change in one observation would influence more than one unknown parameter and residual.

The spurious correlation among the residuals, introduced by the model but usually neglected by the statistical tests, creates problems. When a single observational vector is to be tested, it can be easily

shown (*Morrison, 1976*) for P=I that the correlation among the residuals is quite small:

$$\rho_{ij} = -\frac{1}{n-1}$$

where n is the size of the vector.

However, when several observations are simultaneously examined using multivariate analysis, the correlation among the residuals in some cases may be quite significant. This statistical dependence causes problems since most of the usual techniques which examine standardized residuals (ch. 5), base their efficiency on the diagonal dominance of the elements of the Q_v matrix.

The problem is twofold. First, significance levels between uni- and multi-dimensional tests have to be established so that the decisions from both tests will be consistant. If, however, the quantities are considered as uncorrelated regardless of the actual situation, the problem can be somewhat circumvented through Bonferroni's inequality (sect. 5.3). Secondly, the existing dependence among the tested quantities (residuals), prevents us from pinpointing the erroneous observation since its gross error has been spread among all the residuals. The first problem is connected with the detection of outliers (at which probability level the testing is performed) and the second problem is connected with their localization. Both will be examined again below.

One could of course perform a diagonalization on the covariance matrix of the residuals and come out with a diagonal matrix,

$$C_v' = D^T C_v D$$

where D is an orthogonal matrix. In the case of an eigenvalue diagonalization, the D matrix consists of the columns of the normalized

eigenvectors of C_v. The residuals are also transformed using the same orthogonal matrix D,

$$\underline{v}' = D^T \underline{v}$$

However this solution would bring some other problems. Since the rank of C_v is n-u = r, a diagonalization would create n-r = u zero elements along the diagonal of the new transformed matrix C_v. The rest of the eigenvalues would be equal to 1 because C_v is an idempotent matrix. The new residuals v_i' which correspond to the zero diagonal elements of C_v' will also be zero.

The new quantities (residuals v_i'), are indeed uncorrelated but we are not interested in them since their relation with the initial observations is very intricate. The physical meaning of the new space is quite unclear as each new residual v_i' is a linear combination of all the previous ones. Under these circumstances, direct correspondence between actual observations and transformed residuals does not exist and the localization of errors through the examination of the standardized residuals becomes impossible. Therefore such a transformation of the residuals and their covariance matrix is not recommended.

3.2 Relation between Residuals and Gross Errors - Sparse Redundancy

The relation between residuals v and true errors ε proves to be (I.8), (I.10):

$$\underline{v} = Q_v P \underline{\varepsilon} = M \cdot \underline{\varepsilon} \qquad (3.1)$$

If we assume that each observational true error ε_i consists of one random part ε_{r_i} and one gross error part $\nabla\ell_i$ (using Baarda's notation),

$$\underline{\varepsilon} = \underline{\varepsilon}_r + \underline{\nabla\ell}$$

then, equation (3.1) can be written as:

$$\underline{v} = Q_v P \underline{\varepsilon}_r + Q_v P \underline{\nabla \ell} \qquad (3.2)$$
$$= \underline{v}_r + \underline{\nabla v}$$

where:

$\underline{v}_r = Q_v P \cdot \underline{\varepsilon}_r$ is the influence of random errors on the residuals, and

$\underline{\nabla v} = Q_v P \cdot \underline{\nabla \ell}$ is the influence of gross errors on the residuals.

Equation (3.2) shows that residuals are affected by random and gross errors. Since the Least Squares method does not distinguish between the two kinds of error, the minimization concerns the joint effect and not only the random part. Therefore, after the adjustment it is quite difficult to detect the gross errors, through the examination of the residuals.

From the properties of the idempotent matrix $M = Q_v P$ (app. I it is known that:

$$\text{trace } [Q_v P] = \text{rank } [Q_v P] = n-u = r \qquad (3.3)$$

where:

r : is the total redundancy.

If r_i are the diagonal elements of $M = Q_v P$, from equation (3.3) we obtain:

$$\sum_{i=1}^{n} r_i = r \qquad \text{and} \qquad r_i = (Q_v P)_{ii} = q_{v_i} \cdot p_i \qquad (3.4)$$

Moreover from equation (I.18),

$$0 \leq r_i \leq 1$$

Considering also equation (3.2): $\underline{\nabla v} = Q_v P \underline{\nabla \ell}$, the individual ∇v_i can

be defined as:

$$\nabla v_i = r_i \cdot \nabla \ell_i \tag{3.5}$$

The elements r_i are called by definition redundancy numbers (*Förstner, 1979*), and simply express the contribution of each single observation ℓ_i to the total redundancy r. The r_i's serve as a measure of the local controlability, that is going to be explained further below.

Consequently a relative redundancy can be defined as the average of the diagonal elements of M, that is,

$$\frac{\mathrm{tr}[Q_v P]}{n} = \frac{r}{n} \tag{3.6}$$

From the definition of M, (I.8),

$$M = Q_v P = I - Q_{\hat{\ell}} P = I - U \tag{3.7}$$

in which

$$U = Q_{\hat{\ell}} P = A Q_{\hat{x}} A^T P \tag{3.8}$$

since U is also idempotent,

$$\text{trace } [U] = \text{rank } [U] = \text{rank}\,[Q_{\hat{x}}] = u \tag{3.9}$$

By definition now, similar to equation (3.4),

$$u_i \overset{d}{=} (Q_{\hat{\ell}} P)_{ii} \tag{3.10}$$

and from equation (3.7)

$$u_i = 1 - r_i \qquad \text{and} \qquad 0 \leq u_i \leq 1 \tag{3.11}$$

Analogously to the previous definition the elements u_i can be called "absorption numbers". From equation (3.11) now,

$$\Sigma u_i = n - \Sigma r_i \qquad \text{or} \qquad u = n - r \tag{3.12}$$

which was expected.

If a gross error $\nabla \ell_i$ exists now in one of the observations ℓ_i, after multiplying by $\nabla \ell_i$, equation (3.11) becomes:

$$u_i \cdot \nabla\ell_i = \nabla\ell_i - r_i \cdot \nabla\ell_i \qquad (3.13)$$

Considering equation (3.5), equation (3.13) becomes

$$\nabla\ell_i = u_i\nabla\ell_i + \nabla v_i \qquad (3.14)$$

The last equation shows that if a gross error $\nabla\ell_i$ occurs in one observation only, ℓ_i, it will be reflected in the corresponding residual v_i, as much as $\nabla v_i = r_i\nabla\ell_i$ and the rest of it will be absorbed in the determination of the unknown parameters.

It is clear now that a gross error $\nabla\ell_i$, in an observation which has large redundancy number r_i, will affect more the corresponding v_i (as much as $r_i\nabla\ell_i$), thus it will be easily detected through the examination of the residuals. The larger the r_i the better the control on the i^{th} observation.

So $r_i = 1$ means first, that 100% of any gross error $\nabla\ell_i$ will be revealed in the residual v_i (full control), and secondly, it will not have any effect at all on the determination of the unknown parameters; e.g. a measured distance between two fixed points would have $r_i = 1$.

If now $r_i = 0$, it means that there is no control on the i^{th} observation, since $\nabla\ell_i$ does not affect at all the residuals, therefore it cannot be detected. Moreover this undetectable error $\nabla\ell_i$, affects 100% the final solution, in other words it is directly transferred into the estimated unknown parameters; e.g. in measuring a single angle and a distance to determine uniquely the position of a point, the local redundancy, r_i, for both observations, is equal to zero.

Hence, it is desirable to have a network with relatively large and if possible uniform r_i's. Unfortunately the relative

redundancy (global measure of controlability) remains constant in most of the geodetic nets, $\bar{r} = r/n \tilde{=} 0.5$ (*Pope, 1976*), and cannot easily be increased. Besides this, the local redundancy r_i varies significantly and does not remain close to $\bar{r}$ for all the ℓ_i's, which means that the controlability is not the same for all the observations. Thus, by revealing how high or low the controlability is in the different parts of the network, the redundancy numbers reflect its geometrical strength *(Förstner, 1981)*.

It is characteristic that the r_i's depend on the geometry of the network and have nothing to do with the actual measurements. Therefore they can be considered during the design of the network *(van Mierlo, 1981)*.

As for the unknown parameters u, it can be said that usually not all of them are of interest. If the model has the form:

$$-\underline{\ell} + \underline{v} = A\underline{x} = A_k \cdot \underline{k} + A_t \cdot \underline{t} \qquad (3.15)$$

where:

$A = [A_k \vdots A_t]$

k : the desired parameters such as point coordinates, etc., and

t : the nuisance parameters such as orientation, scale, etc.,

it can be derived (eq. II.5) that:

$$U = U_k + U_t = I - M \qquad (3.16)$$

Similarly the diagonal elements consist of two parts:

$$1 - r_i = u_i = u_{k_i} + u_{t_i} \qquad (3.17)$$

Multiplying by $\nabla\ell_i$ as it was done in (3.13),

$$\nabla\ell_i = u_{k_i} \nabla\ell_i + u_{t_i} \nabla\ell_i + \nabla v_i \qquad (3.18)$$

as long as only one gross error exists in the i^{th} observation,

where:

$u_{k_i} \nabla\ell_i$: is the part of the gross error $\nabla\ell_i$, which is absorbed in the determination of the desired parameters only.

Since r_i's and u_{k_i}'s describe the variation of the residuals ∇v_i, and of the estimates ∇x_i, as functions of the observational errors $\nabla\ell_i$, in other words the sensitivity of the L.S. estimators to gross errors, they are used for the definition of the reliability of a network (ch. 6).

4

PRE-ADJUSTMENT GROSS ERROR DETECTION TECHNIQUES

Before the large computers were introduced, most of the gross-error detection techniques were confined to simple checks during or just after the data acquisition. These techniques were not very systematic, relying for their efficiency on the experience and intuition of the investigator who most of the time was the project surveyor. These techniques prove to be very useful even today when large errors are to be detected and eliminated. On the other hand, the whole idea of eliminating most of the gross or systematic errors as early as possible, is very elegant. In this way remedies (in form of re-observations) may be still possible, as far as the indispensable observations are concerned; (large redundancy can afford the automatic rejection of erroneous data).

A single check of the closure of a triangle for example does not give, as is very often but erroneously believed, any measure of how accurately the observations were made. It gives however, information about the presence of blunders provided that good estimates of the accuracy of the observations have been chosen as criteria *(Uotila, 1973)*.

The most systematic way to detect the presence of large errors in the raw data prior to the adjustment, is to use all the condition equations and instead of testing residuals to test misclosures which are also functions of the observational errors. The test on a misclosure outlier is similar to the one-dimensional test on a residual outlier (ch. 5). Depending now on the a priori variance factor (known or unknown) the following tests can be performed (*Vanicek and Krakiwsky, 1982*):

1. If σ_o^2 is assumed to be known then the statistic, which is a standardized misclosure,

$$\overline{f}_i = \frac{f_i}{\sigma_{f_i}} \quad \text{has the N(0,1) distribution.}$$

Therefore if

$$|\overline{f}_i| = \left|\frac{f_i}{\sigma_{f_i}}\right| > \sqrt{F_{1-\alpha;\ 1,\infty}} \tag{4.1}$$

it is concluded that the i^{th} misclosure has been affected by the presence of gross errors.

2. If σ_o^2 is unknown, $\hat{\sigma}_o^2$ is employed and the statistic

$$\overline{f}_i = \frac{f_i}{\hat{\sigma}_{f_i}} \quad \text{has the } t_r \text{ distribution,}$$

where $\hat{\sigma}_o^2$ is to be determined from an independent source and r is the

redundancy used for its estimation. If $\hat{\sigma}_o^2$ is determined from the same sample then the τ_r-distribution (sect. 5.3), instead of t_r has to be used. Similarly to (4.1), if

$$|\bar{f}_i| = \left|\frac{f_i}{\hat{\sigma}_{f_i}}\right| > c_{t_{r;(1-\alpha/2)}} \qquad (4.2)$$

it seems that the i^{th} misclosure has been affected by the presence of large errors in the observations related to it. An investigation for the localization of the detected outliers starts after the testing procedure.

The main disadvantages of a preadjustment gross error technique are:

1. It is not always easy to set up the exact and correct number of the condition equations, as the figure of the network becomes more complicated; whereas parametric equations are very simple to construct *(Bomford, 1977, pp. 164-167)*. This is the reason why for the last years, after the high speed computers became more available, the parametric adjustment has been used almost exclusively.

2. Since each misclosure relates to more than one observation, a single condition equation with large misclosure does not give any information as to which observation was the erroneous one, if it is assumed that only one blunder exists. In other words, there is not any hint of blunder location. The fewer the observations involved in each of the condition equations, the easier the localization of possible gross errors. The efficiency of localization decreases as the number of outliers increases.

All the above plus the increase of the computational cost, have prevented a wide utilization of preadjustment gross error techniques. In spite of this, whenever the detection and elimination of large errors is of major importance, both simple quick checks and a systematic procedure using all the condition equations should be employed. The quick checks have a questionable sensitivity but their efficiency in localizing the errors is high most of the time, since they involve only few of the observations.

To sidestep the above problems, attempts have been made to develop a more systematic use of the conditional least squares model. This has also some other advantages concerning the cost of processing the data. For this purpose an automatic data screening process of two dimensional geodetic networks, using the conditional Least Squares model has been proposed by *Peter Steeves (in prep.)*. In this method a graph theoretical algorithm is used to generate a fundamental set of circuits for the network, which serves as a basis for constructing the condition equations.

In conclusion, it can be said that the preadjustment error detection techniques should be considered in addition to the post-adjustment ones (ch. 5), because of their advantages, if they do not require overly expensive computations.

5

POST-ADJUSTMENT GROSS ERROR DETECTION TECHNIQUES

Since the statistical model used in the Least Squares Estimation is quite simple, only information about the variances of the estimated quantities can be obtained. Their probability distribution is not known; therefore, a statistical test for the estimated parameters is not available (*Förstner, 1980*).

From equation (I.10) the original observational errors could be directly computed if matrix M could be inverted. Unfortunately, M is singular, hence the most rational way to detect the presence of gross errors is to test derived quantities which are functions of them. Thus, all the statistical tests after the adjustment focus on the analysis of the errors in the functional model, i.e. the residuals.

It is important to realize at this stage that tests on the

residuals after the adjustment belong to the multivariate tests and examine how the observations fit into the mathematical model. Since the residuals do not reflect only the quality of the observations, they are not always very good indicators of gross errors. This has already been made clear from the relationship between residuals and gross errors (sect. 3.2).

In the following sections the most popular post-adjustment blunder detection techniques are reviewed, their problems are analysed and finally a strategy for the detection, location and elimination of gross errors is proposed.

5.1 Global Test on the Variance Factor

The first test which is applied after the adjustment is the well known global test on the a posteriori variance factor $\hat{\sigma}_o^2$. This test obviously can be applied only when there is a priori knowledge about the precision of the observations, i.e. when the a priori variance factor σ_o^2 is assumed to be known. Otherwise the test has no meaning.

Under the null hypothesis H_o the statistic $\hat{\sigma}_o^2/\sigma_o^2$ follows the $F_{r,\infty}$ - distribution. It is to be remembered that $F_{r,\infty} = \chi_r^2/r$. The decision for the guise of this global test (one-tailed or two-tailed) depends on the purpose of the test which is defined by the null hypothesis H_o. If, for example, the null hypothesis:

$$H_o : \sigma^2 = \sigma_o^2 \text{ is tested, under } H_a : \sigma^2 \neq \sigma_o^2$$

where, σ^2 represents the variance factor and $\hat{\sigma}_o^2$ is its estimated value; then the two-tailed test is recommended, which gives the following $(1-\alpha)$ confidence interval for the variance factor σ_o^2:

$$\frac{r\hat{\sigma}_o^2}{\chi^2_{r,1-\alpha/2}} < \sigma_o^2 < \frac{r\hat{\sigma}_o^2}{\chi^2_{r,\alpha/2}} \qquad (5.1)$$

However, when the global test is used for the detection of outliers it is normally expected that $\hat{\sigma}_o^2 > \sigma_o^2$. Therefore, the null hypothesis to be tested is:

$$H_o : \sigma^2 = \sigma_o^2 \quad \text{under } H_a : \sigma^2 > \sigma_o^2 ,$$

and the one-tailed test is recommended:

$$\frac{\hat{\sigma}_o^2}{\sigma_o^2} < F_{r,\infty;1-\alpha} \tag{5.2}$$

Since $\hat{\sigma}_o^2/\sigma_o^2 = \underline{v}^T P \underline{v}/r\sigma_o^2$, equation (5.2) can be written as:

$$\underline{v}^T C_{\ell}^{-1} \underline{v} < rF_{r,\infty;1-\alpha} = \chi^2_{r;1-\alpha} \tag{5.3}$$

If there are any blunders in the data then the above quadratic form will increase and the test may fail or not, depending on the magnitude of the blunders and on how they are reflected in the residuals. If this test (5.3) fails, then H_o is rejected. Unfortunately, there may be more than one reason for rejection *(Uotila, 1976)* for example:

a) incorrect estimate of weights;

b) incorrect mathematical model;

c) blunders in the observations.

We may not know which one of the above reasons caused the failure of the test, and the test does not give any additional information. Whatever the reason is, it should be investigated and not ignored. Confining ourselves to the third possible cause for rejection, i.e. blunders in the observations, an alternative hypothesis H_a has now to be introduced. The simplest H_a considers a possible shift $\underline{\nabla\ell}$ of the probability distribution of the observations $\underline{\ell}$; therefore it can be defined as *(van Mierlo, 1977)*,

$$E[\underline{\ell}|H_a] = E[\underline{\ell}|H_o] + \underline{\nabla\ell} \tag{5.4}$$

Under the null hypothesis the expectation of the residuals is zero:

$$E[\underline{v}|H_o] = 0$$

but under the alternative hypothesis, $E[\underline{v}|H_a] = \underline{\nabla\ell}$. The relationship between $\underline{\nabla\ell}$ and $\underline{\nabla v}$ is already known (eq. 3.2):

$$\underline{\nabla v} = Q_v P \underline{\nabla\ell} \tag{5.5}$$

Also under the alternative hypothesis, the expectation of $\hat{\sigma}_o^2/\sigma_o^2$ is not equal to one anymore, but

$$E\left[\frac{\hat{\sigma}_o^2}{\sigma_o^2} \mid H_a\right] = E\left[\frac{\hat{\sigma}_o^2}{\sigma_o^2} \mid H_o\right] + \nabla\left(\frac{\hat{\sigma}_o^2}{\sigma_o^2}\right) \tag{5.6}$$

where

$$\nabla\left(\frac{\hat{\sigma}_o^2}{\sigma_o^2}\right) = \frac{\nabla\hat{\sigma}_o^2}{\sigma_o^2} = \frac{1}{\sigma_o^2}\,\frac{\underline{\nabla v}^T P \underline{\nabla v}}{r} = \frac{\lambda}{r} \tag{5.7}$$

Therefore, equation (5.6) can be expressed as

$$E[\hat{\sigma}_o^2/\sigma_o^2 \mid H_a] = 1 + \lambda/r \tag{5.8}$$

where,

$$\lambda = \frac{1}{\sigma_o^2}\,\underline{\nabla v}^T P \underline{\nabla v} = \frac{1}{\sigma_o^2}\,\underline{\nabla\ell} P Q_v P \underline{\nabla\ell} \tag{5.9}$$

Under the alternative hypothesis the statistic $\hat{\sigma}_o^2/\sigma_o^2$ has a non-central $F_{r,\infty,\lambda}$- distribution with the above non-centrality parameter λ. This depicts the shift of the probability distribution due to a gross error $\nabla\ell$ (fig. 5.1).

Although the residuals are known, the observational errors $\underline{\nabla\ell}$ and consequently λ are not. Thus no conclusions can be drawn as to which observations are the erroneous ones. Only a boundary value λ_o can be determined which gives the deviation from the null hypothesis which can be detected at certain probability levels α and β_o:

$$\lambda_o = \lambda(\alpha, \beta_o, r, \infty) \tag{5.10}$$

However, this value does not give any information about the individual

elements of the $\underline{\nabla \ell}$ vector since it is referred to the entire model. What is needed is a simpler and more specific H_a which will constrain the relationship between residuals and gross errors $\nabla \ell$, so that boundary values can be estimated for the vector $\underline{\nabla \ell}$ when a test is performed at a given probability level β_o. This H_a introduces the uni-dimensional tests on the residuals.

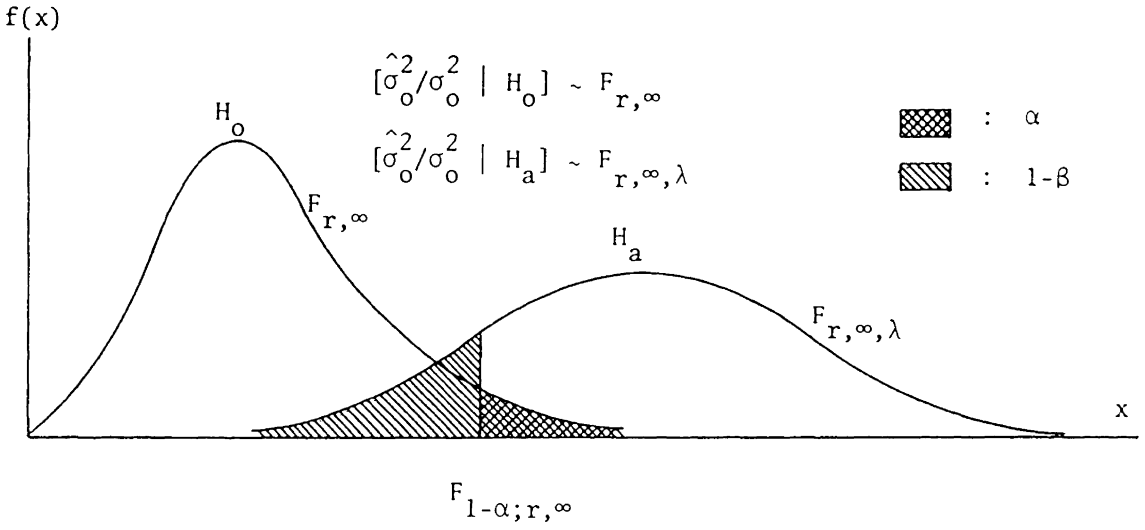


Figure 5.1: P.D.F. of $[\hat{\sigma}_o^2/\sigma_o^2]$ under H_o and H_a

5.2 Data Snooping - Boundary Values

Under the new H_a it is assumed that only certain observations have been affected by an error $\nabla \ell$.

$$H_{a_c} : \underline{\nabla \ell} = \underline{c} \nabla \ell \qquad (5.11)$$

where, the units in the vector $\underline{c}$ = (0 ... 1 0 1 0 ... 0) indicate the erroneous observations, and $\nabla \ell$ is a scalar denoting the gross error. Considering the boundary value of λ_o (eq. 5.10), equation (5.9) can

be written as

$$\lambda_o = (1/\sigma_o^2)\underline{\nabla_o \ell}^T PQ_v P \underline{\nabla_o \ell} \tag{5.12}$$

which under equation (5.11) becomes

$$\lambda_o = \frac{\nabla_o \ell^2}{\sigma_o^2} \underline{c}^T PQ_v P \underline{c} \tag{5.13}$$

from which the boundary value $\nabla_o \ell$ of a gross error in the observations specified by the vector $\underline{c}$, can be obtained:

$$|\nabla_o \ell| = \sigma_o \sqrt{\frac{\lambda_o}{\underline{c}^T PQ_v P \underline{c}}} \tag{5.14}$$

It would be easy now to derive a statistic which tests the alternative hypothesis H_{a_c}, but a further simplified alternative hypothesis is used in practice, for it is not easy to know which observations are the erroneous ones. The new H_a assumes that only one observation at the time is erroneous. Therefore, the vector $\underline{c}_i$ will have the i^{th} element equal to one and the rest of them equal to zero. Baarda (1968), calls all these H_a, the "conventional alternative hypotheses":

$$H_{a_c} : \underline{\nabla_o \ell_i} = \underline{c_i} \nabla_o \ell_i \tag{5.15}$$

Equation (5.14) can be written now as

$$|\nabla_o \ell_i| = \sigma_o \sqrt{\frac{\lambda_o}{(PQ_v P)_{ii}}} \tag{5.16}$$

This expression is very important and it is used for the definition of the internal reliability (sect. 5.1).

Now the one-dimensional test statistic can be derived to test the alternative hypotheses, *(Alberda, 1980)*,

$$w_i = \frac{\underline{c}_i^T P \underline{v}}{\sigma_o \sqrt{\underline{c}_i PQ_v P \underline{c}_i}} = \frac{(P\underline{v})_{ii}}{\sigma_o \sqrt{(PQ_v P)_{ii}}} \tag{5.17}$$

which in the case of a diagonal P matrix is simplified to:

$$w_i = \frac{\bar{v}_i}{\sigma_{v_i}} = \frac{v_i}{\sigma_o \sqrt{q_{v_i}}} \qquad (5.18)$$

In other words w_i is a standardized residual which has the normal distribution N(0,1) (fig. 5.2).

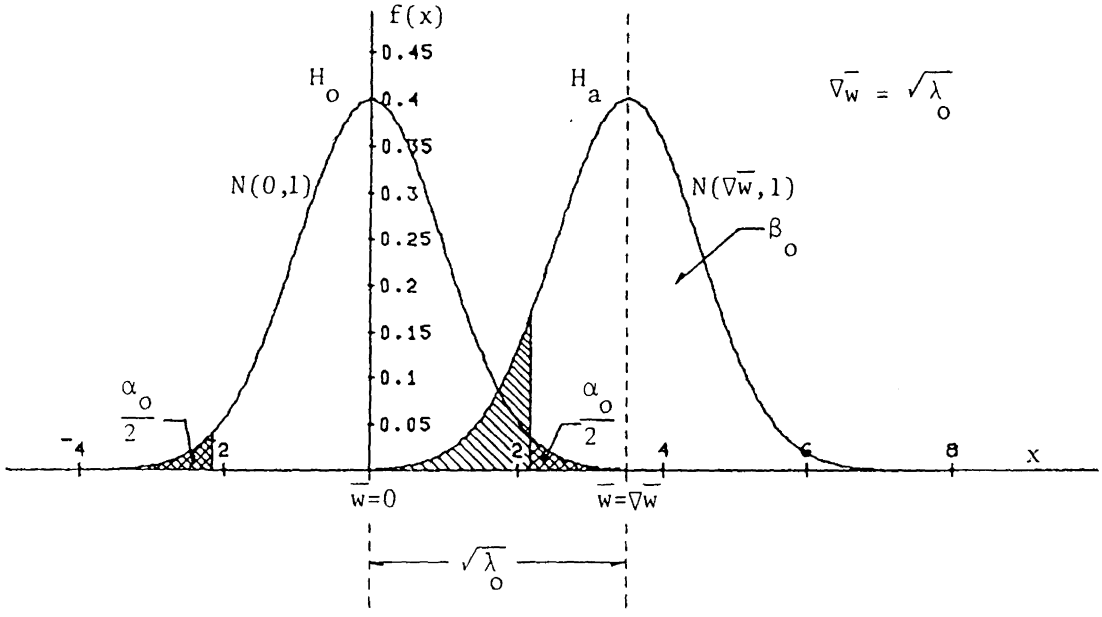


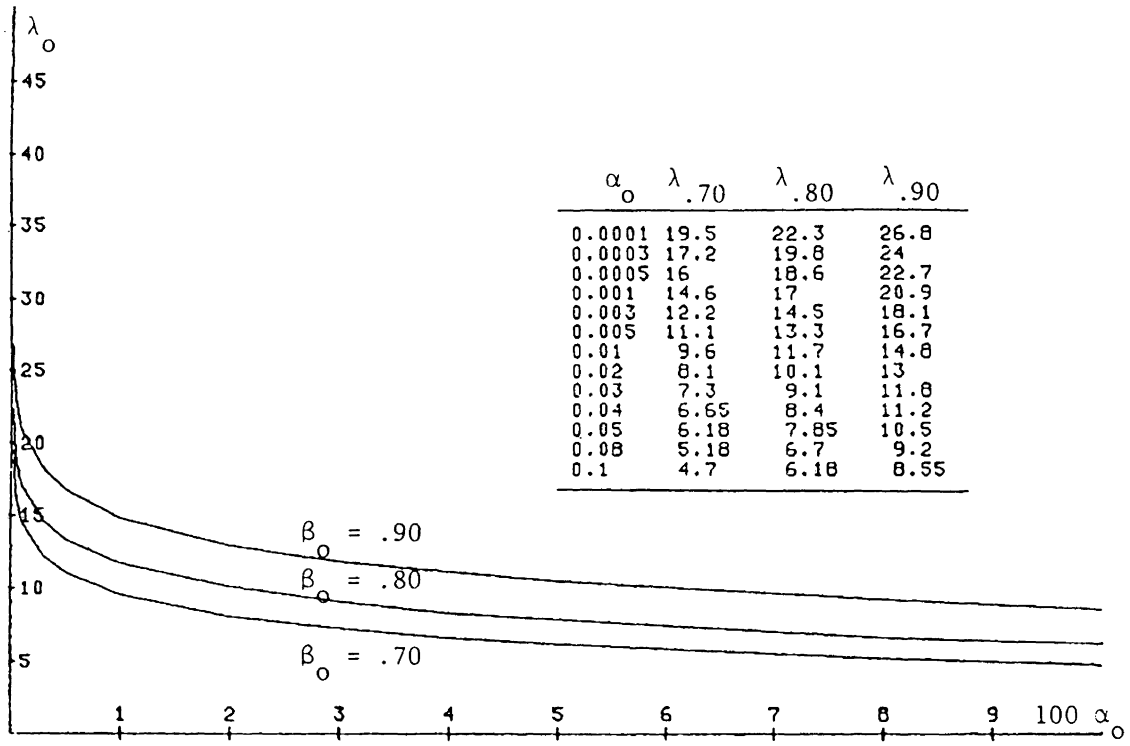
Figure 5.2: w-test (Data Snooping)

From this figure it can be easily seen that fixing the significance level α_o and the power of the test β_o, the non-centrality parameter can be estimated from:

$$\sqrt{\lambda_o} = \sqrt{F_{1-\alpha_o;1,\infty}} + \sqrt{F_{\beta_o;1,\infty}} \qquad (5.19)$$

The variation of λ due to this dependence is also shown in figure (5.3).

The term "Data Snooping" refers to the above one-dimensional test, examining only one standardized residual (eq. 5.18), at the time.



α_o	$\lambda_{.70}$	$\lambda_{.80}$	$\lambda_{.90}$
0.0001	19.5	22.3	26.8
0.0003	17.2	19.8	24
0.0005	16	18.6	22.7
0.001	14.6	17	20.9
0.003	12.2	14.5	18.1
0.005	11.1	13.3	16.7
0.01	9.6	11.7	14.8
0.02	8.1	10.1	13
0.03	7.3	9.1	11.8
0.04	6.65	8.4	11.2
0.05	6.18	7.85	10.5
0.08	5.18	6.7	9.2
0.1	4.7	6.18	8.55

Figure 5.3: Variation of the non-centrality parameter λ_o.

If,

$$|w_i| = \left| \frac{v_i}{\sigma_{v_i}} \right| > \sqrt{F_{1-\alpha_o;1,\infty}} \qquad (5.20)$$

the null hypothesis (H_o : there is not any gross error in the i^{th} observation), is not accepted and the i^{th} residual is flagged for rejection. For $\alpha_o = .001$ which is usually suggested *(Baarda, 1968)*, the critical value is $\sqrt{F_{1-\alpha_o;1,\infty}} = 3.29$. Therefore the null hypothesis is rejected if

$$|v_i| > 3.29\ \sigma_{v_i}$$

The procedure can be subsequently repeated in the case of there being more than one gross error. Problems encountered with their localization

are examined below (sect. 5.8).

It should be mentioned that tests on standardized residuals are more robust than tests on the residuals themselves.

It is also notable that since $0 \leq q_{v_i} \leq 1$ (Q_v is an idempotent matrix), in the case of P = I, it can be seen easily that

$$w_i \geq v_i$$

The expansion factor $\kappa = 1/\sqrt{q_{v_i}}$ varies from 1 to $+\infty$. The larger the κ factor the larger the statistic w_i, and consequently the easier the failure of the one-dimensional test (5.20). The variation of the κ factor can be also seen in figure (5.4).

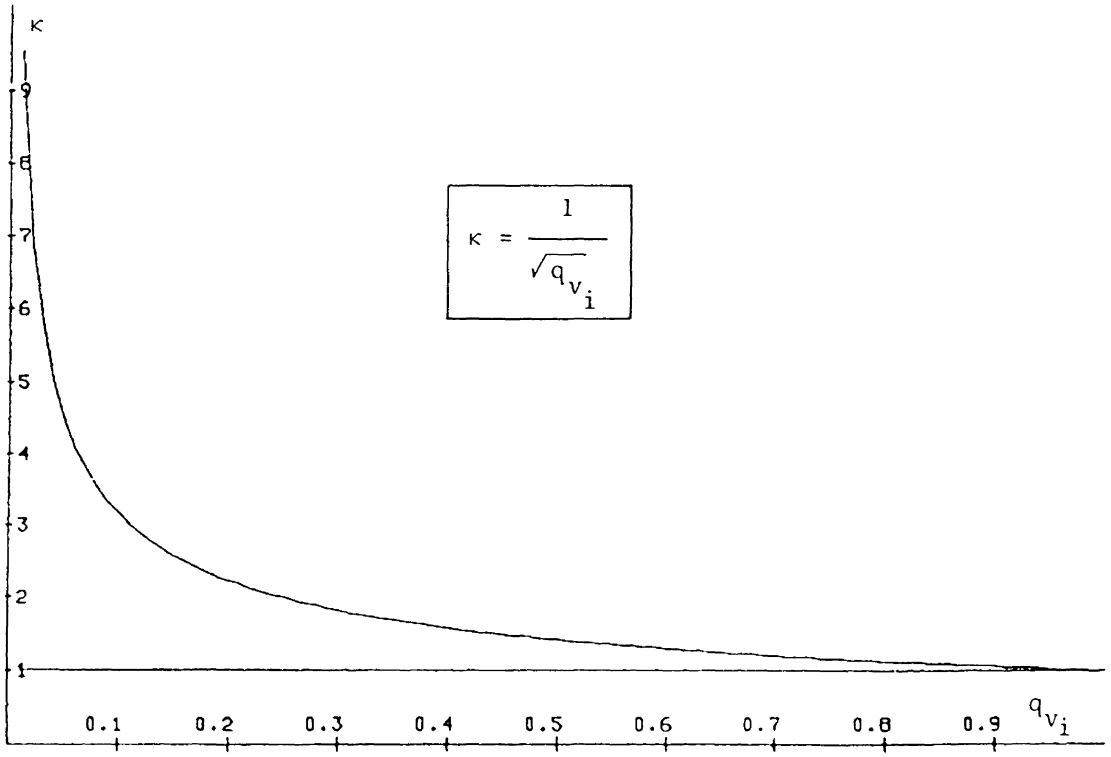


Figure 5.4: Variation of the Expansion Factor κ.

Sometimes in the unavailability of σ_{v_i}'s, which require considerable computational effort (sect. 5.6), partially standardized residuals are used (v_i/σ_{ℓ_i}) *(Pope, 1976)*, where the standard deviation of the observations is employed.

It should be kept in mind however, that since $\sigma_{v_i} < \sigma_{\ell_i}$ the above approximation would decrease the value of the statistic (5.18), that is:

$$\frac{v_i}{\sigma_{\ell_i}} < \frac{v_i}{\sigma_{v_i}}$$

and thus less failures of the test (5.20) would occur.

Baarda (1968) proposed the use of the global test (5.2) for the detection of gross errors and the "Data Snooping" test (5.20) for their localization. The decisions from both tests should be consistent, i.e., the same boundary values should be found whether the global or the single test is performed.

The normal procedure would be to find the probability β, (power of the test), with which a certain alternative hypothesis can be asserted by the test. But for the reason given above, Baarda's philosophy was to find the minimum deviation of the alternative hypothesis H_a from H_o, which can just be detected with a given probability β_o. So his B-method of testing has been founded on the following: "The power of the tests remains constant but the level of significance is variable". For the global test it is "α" but for the uni-dimensional test (Data-Snooping), it is "α_o". If $\beta = \beta_o$ (constant), these levels of significance are interconnected through the non-centrality parameter λ_o.

The normal procedure is *(Baarda, 1968, 1976)*:

choose α_o and β_o (e.g. $\alpha_o = 0.001$ and $\beta_o = 0.80$)
compute λ_o from $\lambda_o = \lambda(\alpha_o,\beta_o,1,\infty)$
compute α from $\lambda_o = \lambda(\alpha,\beta_o,r,\infty)$

The values which are usually recommended are:

$$\alpha_o = 0.001$$

$$\beta_o = 0.80$$

The selection of β_o is not as critical as the selection of α_o for the outcome of the tests. The dependence between α and α_o can be found from nomograms *(Baarda, 1968, 1976)*. An example is given in Appendix III. Here only some values are given (table 5.1) for 10 degrees of freedom ($r = 10$):

$r = 10$					
α_o	β_o	λ_o	α	$\sqrt{F_{1-\alpha_o;1,\infty}}$	$F_{1-\alpha;r,\infty}$
.05	.80	7.85	.30	1.96	1.17
.05	.90	10.50	.35	1.96	1.12
.001	.80	17.00	.04	3.29	1.90
.001	.90	20.90	.05	3.29	1.83

Table 5.1: Example values for the B-Method of testing.

5.3 Tau-Test

In the null hypothesis of the previous tests the variance factor σ_o^2 is assumed to be known, which means that all the variances are properly scaled. If, however, σ_o^2 is not adequately known or one does not want

to rely on a priori estimates, then the a posteriori estimate $\hat{\sigma}_o^2$ is always available. In this case the global test on the variance is not performed and "Data Snooping" has to be modified.

The new test statistic, proposed by *Pope (1976)*, is similar to (5.18):

$$\tau_i = \frac{v_i}{\hat{\sigma}_{v_i}} = \frac{v_i}{\hat{\sigma}_o \sqrt{q_{v_i}}} = \frac{w_i}{\hat{\sigma}_o} \tag{5.21}$$

and follows the τ-distribution (tau-distribution), since the residuals are used for the estimation of the $\hat{\sigma}_{v_i}$'s through the $\hat{\sigma}_o^2$.
The multi-dimensional test is defined as a test on the max τ:

$$\begin{aligned}\alpha &= P\{\max \tau > c\} = P\{\text{one or more of } \tau_i \geq c\} \\ &= 1 - P\{\text{all } \tau_i\text{'s} \leq c\} \\ &= 1 - P\{(\tau_1 \leq c) \text{ and } (\tau_2 \leq c) \text{ and } \ldots \text{ and } (\tau_n \leq c)\}\end{aligned}$$

and if the above n events are considered as uncorrelated,

$$\begin{aligned}\alpha &= 1 - \prod_i^n P\{\tau_i \leq c\} = 1 - [P\{\tau_i \leq c\}]^n \\ &= 1 - (1 - \alpha_o)^n\end{aligned}$$

which means that the multi-dimensional test at a significance level α, is a function of n independent one-dimensional tests executed at a significance level α_o where,

$$\alpha = 1 - (1 - \alpha_o)^n \qquad \text{or} \qquad \alpha_o \tilde{=} \alpha/n$$

Similarly to (5.20), if

$$|\tau_i| = \left|\frac{v_i}{\hat{\sigma}_{v_i}}\right| > c_\tau \tag{5.22}$$

the i^{th} residual is flagged for rejection. The test is applied successively to all standardized residuals.

In fact the above assumption concerning the independence among the tested quantities, is not so indispensable.

In multivariate analysis, whenever there is a finite and a priori known number of comparisons to be made, Bonferroni's method is most often used, based on an inequality that is usually given his name. When the problem of simultaneous inferences about each individual of a group of hypotheses is faced, then the tests are constructed in such a way that the confidence level for the entire group is at least $(1 - \alpha)$, *(Morrison, 1976)*.

Since the calculation of the exact joint probability is difficult most of the times, the lower bound for statistically dependant quantities is very often used which is described by the above mentioned inequality:

$$p\{\bigcap_{i=1}^{n}(|\tau_i| \leq c_{\tau_{(\alpha/n)}})\} \geq (1 - \frac{\alpha}{n})^n \simeq 1 - \alpha \qquad (5.23)$$

This says that if all the individual tests of quantities, whether dependent or not, are executed at $(\alpha_o = \alpha/n)$ significance level, the confidence level for the entire group (simultaneous probability), will be at least $(1 - \alpha)$.

Let us assume now that each of the hypotheses is tested at $(\alpha_o = \alpha/n)$ significance level. Bonferroni's inequality says that:

1. If the quantities are uncorrelated the significance level for the entire group is α.
2. If the quantities are correlated the significance level for the entire group is less than α.

In other words, ($\alpha_o \geq \alpha/n$), so in selecting ($\alpha_o = \alpha/n$), Bonferroni's inequality is more conservative, that is, it chooses the smallest α_o and thus it is more reluctant to reject the tested hypothesis. Consequently and inevitably it protects us from commiting the type I error but not the type II.

The "tau"-distribution is rarely mentioned in the statistical textbooks and it is not universally known under this name. It can be derived from the t_r-distribution (*Pope, 1976*):

$$\tau_r = \frac{\sqrt{r} \cdot t_{r-1}}{\sqrt{r - 1 + t^2_{r-1}}} \qquad (5.24)$$

Since tau-distribution is not very common, not very easy to calculate, the tables or subroutines are not always available. Very often it is wrongly assumed that ($\tau_i = v_i/\hat{\sigma}_{v_i}$) has the t_r-distribution of which critical values are always available.

It should be noticed here that the two distributions give closer critical values, c_τ and c_t, as r increases. For $r \to \infty$ they approach the normal distribution (fig. 5.5).

However for small r, which is the usual case in small geodetic networks, one should be cautious, since the difference ($c_t - c_\tau$) may be significant as it is shown in figure (5.6) and in table (5.2), for $\alpha_o = 0.05$. The tau-distribution is closer to the normal N(0,1) than to the t-distribution, as far as the critical values are concerned.

The difference between c_t and c_τ, also increases as α_o decreases. Use of c_t instead of c_τ, causes generally fewer rejections since $c_t > c_\tau$. Thus, attention should be paid, especially in cases where α_o and r are

Table 5.2: Critical values for the t_r, τ_r and N(0,1) distributions.

$\alpha_o = 0.05$

r	$t_{1-\alpha_o;r}$	$\tau_{1-\alpha_o;r}$	$N_{1-\alpha_o}$	t-τ
2	4.303	1.41	1.96	2.893
3	3.182	1.645	1.96	1.537
4	2.776	1.757	1.96	1.019
5	2.571	1.814	1.96	0.757
6	2.447	1.848	1.96	0.599
7	2.365	1.87	1.96	0.496
8	2.306	1.885	1.96	0.421
9	2.262	1.896	1.96	0.366
10	2.228	1.904	1.96	0.324
11	2.201	1.91	1.96	0.291
12	2.179	1.915	1.96	0.264
13	2.16	1.92	1.96	0.24
14	2.145	1.923	1.96	0.222
15	2.131	1.926	1.96	0.205
16	2.12	1.929	1.96	0.191
17	2.11	1.931	1.96	0.179
18	2.101	1.933	1.96	0.168
19	2.093	1.934	1.96	0.159
20	2.086	1.936	1.96	0.15
21	2.08	1.937	1.96	0.143
22	2.074	1.938	1.96	0.136
23	2.069	1.939	1.96	0.13
24	2.064	1.94	1.96	0.124
25	2.06	1.941	1.96	0.119
26	2.056	1.942	1.96	0.114
27	2.052	1.943	1.96	0.109
30	2.042	1.945	1.96	0.097
35	2.03	1.947	1.96	0.083
40	2.021	1.949	1.96	0.072
50	2.009	1.951	1.96	0.058
60	2	1.953	1.96	0.047
80	1.99	1.955	1.96	0.035
100	1.984	1.956	1.96	0.028

relatively small. Generally for ($\alpha_o = 0.05$), t-distribution can substitute tau-distribution for $r > 30$ ($c_t - c_\tau \leq 0.1$), and normal-distribution substitutes "tau" for $r > 10$ ($c_N - c_\tau \leq 0.06$).

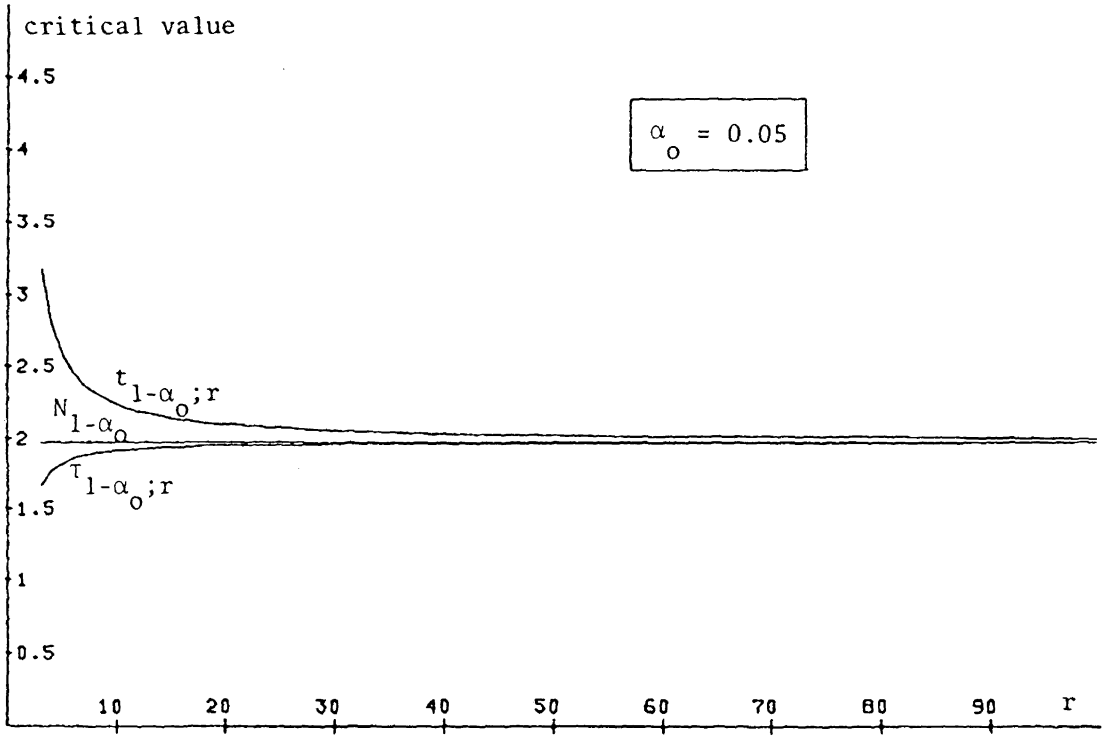


Figure 5.5: Critical values for one-dimensional tests

From the characteristic formula (5.24), which gives the critical values of tau-distribution for r degrees of freedom as a function of t, it can be easily derived that

$$\lim_{t\to\infty} \tau_r = \lim_{t\to\infty} \frac{\sqrt{r} \cdot t_{r-1}}{\sqrt{r - 1 + t^2_{r-1}}} = \sqrt{r} \lim_{t\to\infty} \frac{t_{r-1}}{\sqrt{r - 1 + t^2_{r-1}}} = \sqrt{r} \qquad (5.25)$$

which simply says that the probability distribution function of tau is bounded by $-\sqrt{r}$ and $\sqrt{r}$.

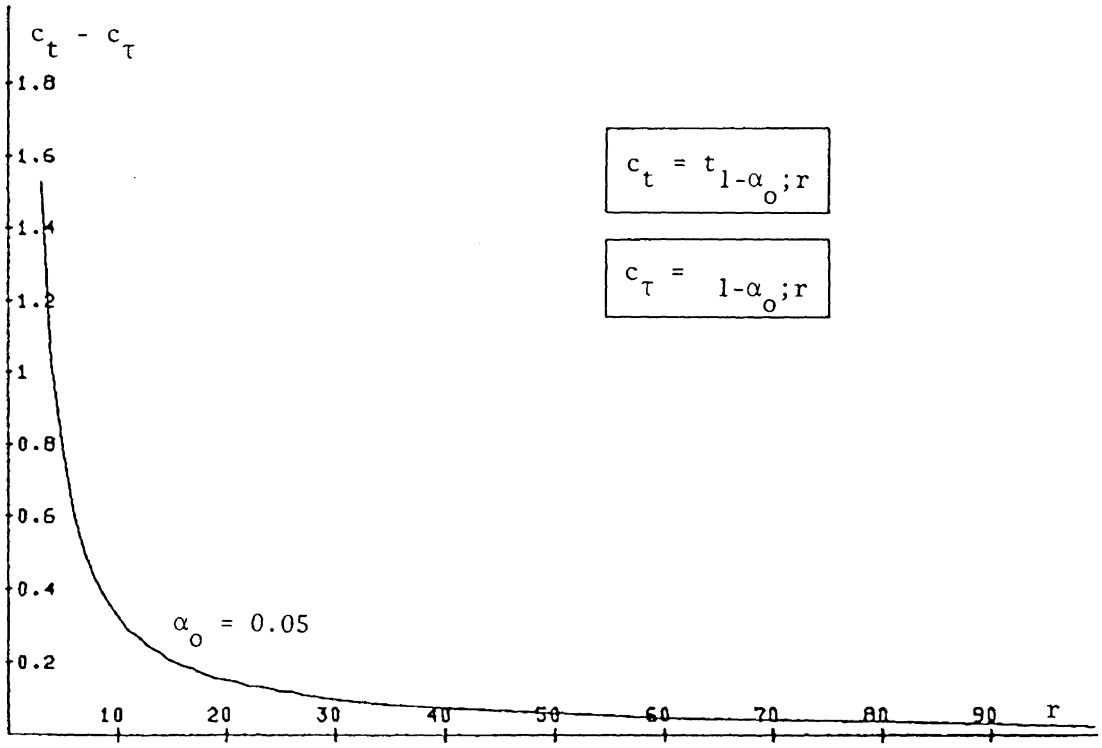


Figure 5.6: Difference between "t" and "tau" critical values

To better illustrate this, the probability density function of the tau-distribution can be derived. From equation (5.24), t_{r-1} can be expressed as a function of τ_r (for the sake of simplicity subscripts are omitted):

$$t = \frac{\tau \sqrt{r - 1}}{\sqrt{r - \tau^2}} \tag{5.26}$$

which again implies that $|\tau| \leq \sqrt{r}$.

Differentiation of (5.26) with respect to τ yields:

$$dt = \frac{r \sqrt{r - 1}}{(r - \tau^2)^{3/2}} \cdot d\tau \tag{5.27}$$

Moreover, it is known that;

$$\int_{-\infty}^{+\infty} f(t)dt = \int_{-\sqrt{r}}^{\sqrt{r}} f(\tau)d\tau = 1 \qquad (5.28)$$

where (*Hogg and Craig, 1978*),

$$f(t)dt = \frac{1}{\sqrt{(r-1)\pi}} \frac{\Gamma(r/2)}{[(r-1)/2]} \left(1 + \frac{t^2}{r-1}\right)^{-\frac{r}{2}} \cdot dt \qquad (5.29)$$

for (r - 1) degrees of freedom. It is also reminded that $\Gamma(\alpha) = (\alpha-1)!$ is the "Gamma" function.

Substitution of the expressions for t (eq. 5.26), and dt (eq. 5.27), in equation (5.29), after some manipulations yields:

$$f(t)dt = \frac{1}{\sqrt{\pi}} \frac{\Gamma(r/2)}{\Gamma[(r-1)/2]} \frac{(r - \tau^2)^{\frac{r-3}{2}}}{r^{\frac{r-2}{2}}} \, d\tau \qquad (5.30)$$

From equations (5.28) and (5.29), the density function for the tau-distribution can be derived,

$$f(\tau_r) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(r/2)}{\Gamma[(r-1)/2]} \frac{(r - \tau^2)^{\frac{r-3}{2}}}{r^{\frac{r-2}{2}}} \qquad (5.31)$$

for r degrees of freedom, (for a similar analysis see *Heck, 1981*).

In figure (5.7) the probability distribution function of tau for r = 6 degrees of freedom has been plotted, with upper and lower bounds of ($\sqrt{r}$ = 2.5) and ($-\sqrt{r}$ = -2.5) respectively. For comparison, the t-distribution for r = 6 also, and the normal N(0,1) have been plotted.

This bounded distribution is explained by the fact that the residuals v_i have been used for the estimation of $\hat{\sigma}_{v_i}$. Hence, the value of the statistic ($v_i/\hat{\sigma}_{v_i}$), (eq. 5.21) cannot exceed a certain value ($\sqrt{r}$), no matter how large the v_i might be.

The convergence of the tau-distribution to the normal N(0,1),

as r increases, can be viewed in figure (5.8) where the tau distribution for 4, 6, 9 and 16 degrees of freedom has been plotted, to be compared to the normal distribution.

Finally, the cummulative functions for the tau and also for the normal N(0,1) distributions are shown in figure (5.9).

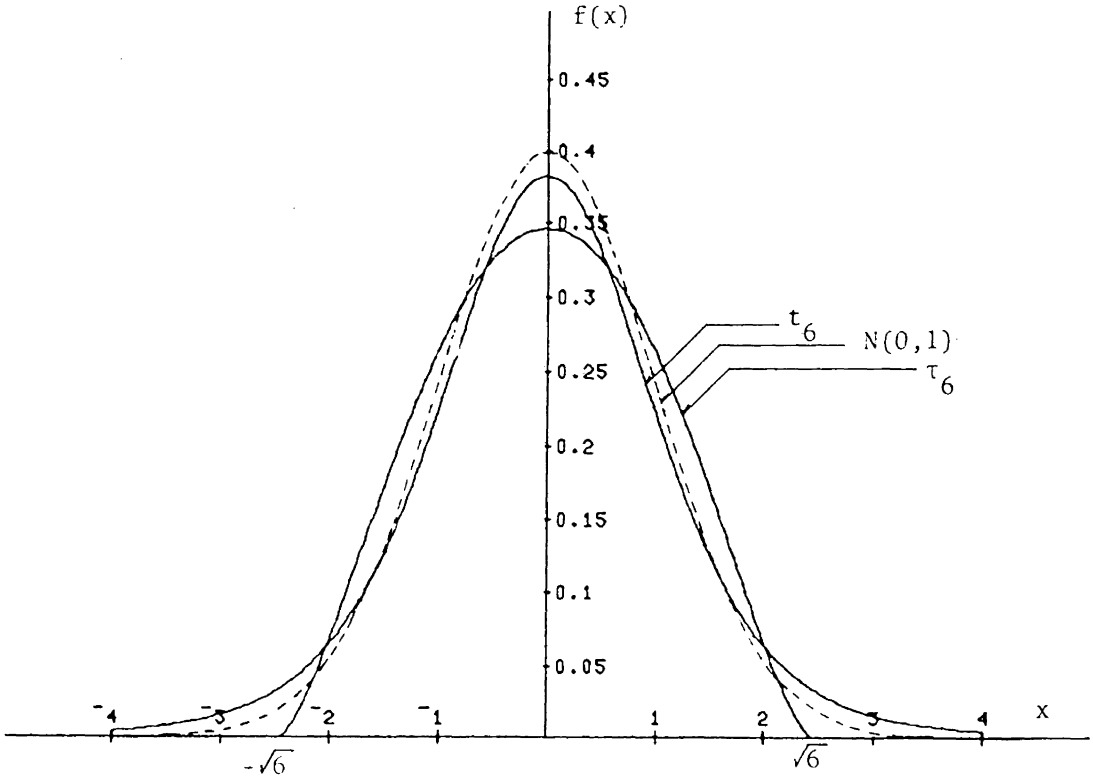


Figure 5.7: P.D.F. of N(0,1), "t" and "tau" for r = 6

5.4 Tests on Quadratic Forms of the Residuals

This method was introduced by *Stefanovic (1978)*, and it is used when groups of residuals are to be tested. A simple function of the residuals is the total quadratic form:

$$q^2 = \underline{v}^T P \underline{v} \qquad (5.32)$$

This statistic has been already used for the global test on the variance (eq. 5.3). If now the vector of observations $\underline{\ell}$ is partitioned to:

$$\underline{\ell} = \begin{bmatrix} \underline{\ell}_1 \\ \hline \underline{\ell}_2 \end{bmatrix}$$

where:

$\underline{\ell}_2$ is the "test group" which under the H_a contains the suspected outlying observations,

and $\underline{\ell}_1$ is the group of the observations free from gross errors.

The residual vector $\underline{v}$ and its cofactor matrix Q_v are similarly partitioned to:

$$\underline{v} = \begin{bmatrix} \underline{v}_1 \\ \hline \underline{v}_2 \end{bmatrix} \qquad \text{and} \qquad Q_v = \left[\begin{array}{c|c} Q_{v_1} & Q_{v_{12}} \\ \hline Q_{v_{21}} & Q_{v_2} \end{array}\right]$$

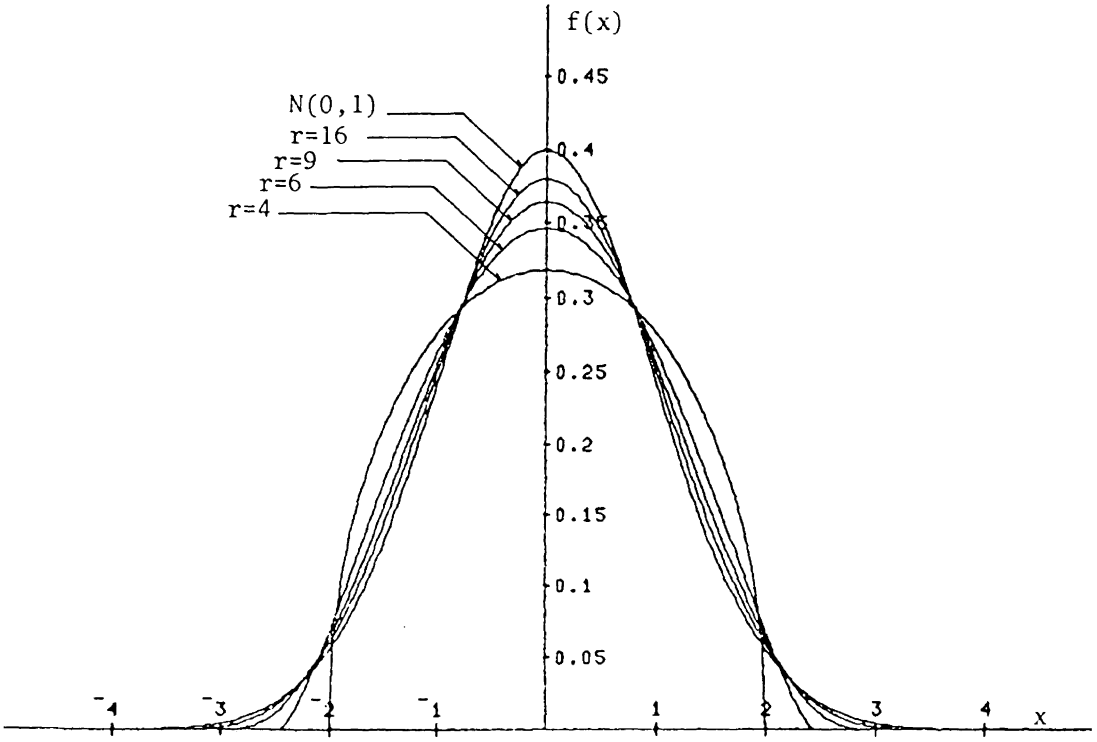


Figure 5.8: P.D.F. of N(0,1) and "tau" for r = 4, 6, 9, 16

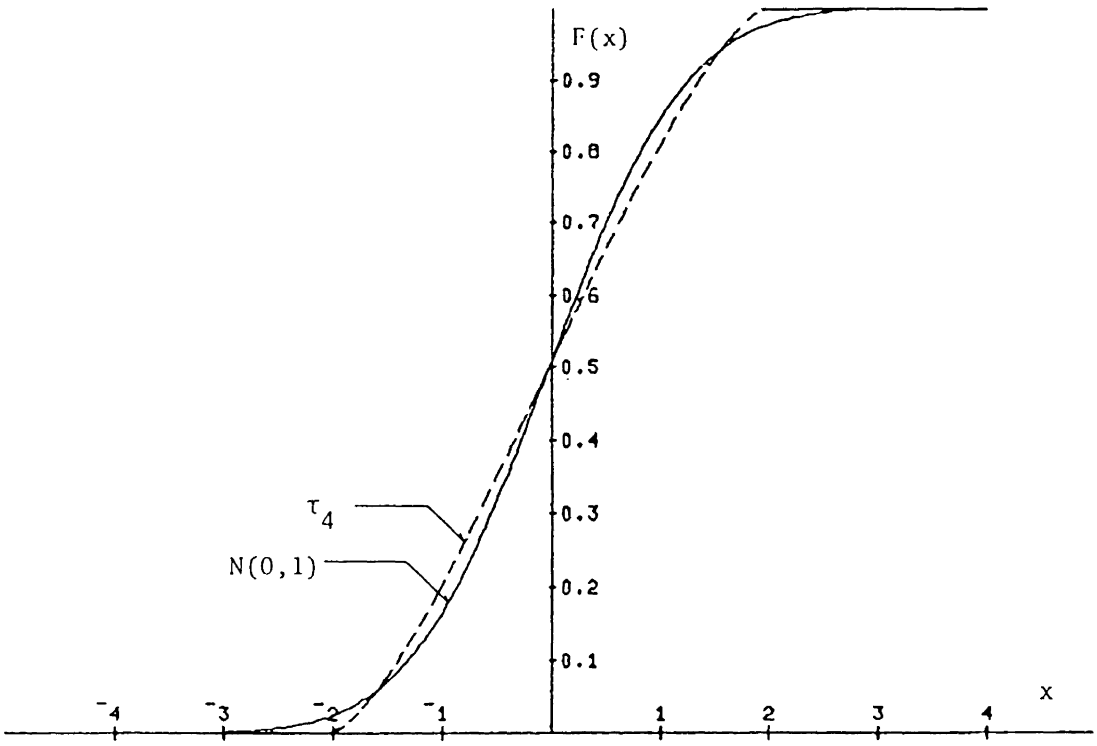


Figure 5.9: C.D.F. of N(0,1) and "tau" for r = 4

Now two new test values can be introduced. The contribution of the test group to the total quadratic form q^2 (eq. 5.32), is given by the following partial quadratic form:

$$d = \underline{v}_2^T Q_{v_2}^{-1} \underline{v}_2 \tag{5.33}$$

which is estimable only if Q_{v_2} is non-singular. If only the $\underline{\ell}_1$ group is used in a new adjustment, the quadratic form of the new residuals $\dot{\underline{v}}_1$, free from the influence of the test group $\underline{\ell}_2$ is given by:

$$\dot{q}^2 = \dot{\underline{v}}_1^T P_1 \dot{\underline{v}}_1 \tag{5.34}$$

It can also be easily derived *(Stefanovic, 1978)* that,

$$\dot{q}^2 = q^2 - d \tag{5.35}$$

which saves much computation compared to the equation (5.34).

From equation (5.35), it can be shown that under the null hypothesis (H_o: no errors in the observations),

$$E[\dot{q}^2] = E[q^2] - E[d] \tag{5.36}$$

which results into

$$[r - p] = [r] - [p]$$

where r : the degrees of freedom

and p : the size of the test group $\underline{\ell}_2$.

The test on the total quadratic form q^2:

$$q^2 < \sigma_o^2 \cdot \chi^2_{\alpha;r} \tag{5.37}$$

is identical to the global test (eq. 5.3). If the test fails one of the reasons may be: "gross errors in the data" (sect. 5.1).

The next test on the partial quadratic form is:

$$d < \sigma_o^2 \cdot \chi^2_{\alpha;p} \tag{5.38}$$

where p is the size of the group $\underline{\ell}_2$.

This test will reveal if the hypothesis, that the test group contains most of the gross errors, is valid or not. If the test group consists of one observation only, the above test on d, (5.38), becomes identical to the Data Snooping test, provided that the significance level has been also properly chosen.

The test on the third quadratic form $\dot{q}^2$, (eq. 5.34),

$$\dot{q}^2 < \sigma_o^2 \cdot \chi^2_{\alpha;(r-p)} \tag{5.39}$$

will tell us if there are any gross errors left in group $\underline{\ell}_1$ and not

included in group $\underline{\ell}_2$.

From the inequalities (5.37), (5.38) and (5.39), it is obvious that a proper detection of a group of outlying observations requires a good knowledge of the variance factor σ_o^2 and the selection of a significance level.

If the hypothesis: "only the observations belonging to the group $\underline{\ell}_2$ are affected by gross errors" is true, then:

-- the first test (5.37) should fail,

-- the second test (5.38) should also fail, but

-- the third test (5.39) should pass.

This does not necessarily mean that in the test group $\underline{\ell}_2$ only erroneous observations have been included. Here is the difficult point in choosing each time the right alternative hypothesis among more than one which would probably satisfy the tests. The problem is not easy to solve (sect. 5.8; ch. 7), and a lot of research has been done along that direction *(Stefanovic, 1978, 1980; Förstner, 1980, 1981)*.

5.5 Example

The following example illustrates the application of different criteria for the detection of gross errors. Let us consider the case of direct observations with the following 10 measurements:

$$\underline{\ell}^T = [14, 19, 20, 20, 20.5, 20, 19.5, 19, 17.5, 21]$$

with $\sigma_{\ell_i} = 1.27$

$P = 0.621 \cdot I$

$\sigma_o^2 = 1$

$n = 10,\ r = n-1 = 9$

The Least Squares solution gives:

$$\hat{x} = 19.05$$

$$\underline{v}^T = [5.05, 0.05, -0.95, -0.95, -1.45, -0.95, -0.45, 0.05, 1.55, -1.95]$$

$$\hat{\sigma}_o^2 = 2.53$$

$$\sigma_x^2 = \frac{1}{n\, p_i} = 0.161$$

$$\sigma_{v_i}^2 = \frac{n-1}{n}\, \sigma_{\ell_i}^2 = 1.45$$

The null hypothesis: H_o: there are not any gross errors in $\underline{\ell}$, can be tested by using the global test on the variance factor

a. Global Test (sect. 5.1)

From test (5.2):

$$\frac{\hat{\sigma}_o^2}{\sigma_o^2} < F_{1-\alpha;r,\infty}$$

for $\alpha = 0.05$ and $r = 9$, the critical value can be determined:

$$F_{0.95;9,\infty} = 1.9$$

Therefore, one tests if

$$\frac{2.53}{1} < 1.9$$

which fails (H_o is rejected). The rejection of H_o is an indication of large errors in the observations.

b. Data Snooping (sect. 5.2)

The null hypothesis is more specific here:

H_o : there is not any gross error in the first observation.

The critical value can be determined from nomograms (app. III):

for $\alpha = 0.05$, $\beta_o = 0.80$ and $r = 9$, $\alpha_o = 0.002$, therefore

$$\sqrt{F_{1-\alpha_o;1,\infty}} = 3.1.$$

From (5.20), one tests if

$$\left| \frac{v_i}{\sigma_{v_i}} \right| < \sqrt{F_{1-\alpha_o;1,\infty}}$$

$$\frac{5.05}{\sqrt{1.45}} = 4.19 < 3.1$$

which fails for the first standardized residual.

c. Tau-Test (sect. 5.3)

With $\alpha_o = 0.002$ and $n = 10$, the significance level for the max τ is $\alpha = n\alpha_o = (10)(0.002) = 0.02$, (sect. 5.3). The critical value is given from tables or from a subroutine *(Pope, 1976)*, which in this case is:

$$c_\tau = \tau_{1-\alpha;n;r} = 2.25 \text{ (in context)}.$$

From (5.22), test if

$$\left| \frac{v_i}{\hat{\sigma}_{v_i}} \right| < c_\tau$$

that is, if

$$\left| \frac{5.05}{\sqrt{(2.53)(1.45)}} \right| = 2.63 < 2.25$$

The test fails again for the first residual.

d. Tests on the Quadratic Forms (sect. 5.4)

Here the tests (5.37), (5.38) and (5.39) are used. The critical values have been determined using: $\alpha = 0.05$, $r = 9$ and $p = 1$, i.e., the test group consists of one observation only.

d.1) $q^2 = \underline{v}^T P \underline{v} < \sigma_o^2 \chi^2_{1-\alpha;r}$, that is, test if $22.71 < 16.92$: it fails.

d.2) $d = \underline{v}_2^T Q_{v_2}^{-1} \underline{v}_2 < \sigma_o^2 \chi^2_{1-\alpha;p}$, that is, test if

$$\frac{5.05^2}{1.45} = 17.59 < 3.84$$: it fails.

d.3) $\dot{q}^2 = \dot{\underline{v}}_1^T P_1 \dot{\underline{v}}_1 < \sigma_o^2 \chi^2_{1-\alpha;(r-p)}$, or

$\dot{q}^2 = (q^2 - d) < \sigma_o^2 \chi^2_{1-\alpha;(r-p)}$, that is

test if $5.22 < 15.5$: it passes.

According to the comments of section (5.4), the null hypothesis is not accepted and the first observation is flagged for rejection.

5.6 Computation of the Covariance Matrix of the Residuals

Up to now nothing has been mentioned about the computation of the covariance matrix of the residuals

$$C_v = \sigma_o^2 Q_v = \sigma_o^2 (P^{-1} - AQ_xA^T) \qquad (5.40)$$

It is indeed true that a rigorous complete calculation of the above matrix involves a large computational effort; i.e., assuming that Q_x is known the computation of the second part of (5.40):

$$Q_{\hat{\ell}} = AQ_xA^T$$

requires $(n \times u^2) + (n^2 \times u)$ multiplications. Instead of this rigorous calculation, an average value of the elements of Q_v is often used. This value can be derived using the properties of the idempotent matrix $Q_v P$ (sect. 3.2; app. I), in the case of a diagonal weight matrix P:

From equation (5.40),

$$\frac{C_v P}{\sigma_o^2} = Q_v P = I - A Q_x A^T P$$

but from equation (3.3) it is known already that

$$\text{trace } [\frac{C_v P}{\sigma_o^2}] = n - u = r$$

therefore,

$$\sum_{i=1}^{n} (\frac{\sigma_{v_i}^2}{\sigma_{\ell_i}^2}) = r$$

Taking the average of the $(\sigma_{v_i}^2 / \sigma_{\ell_i}^2)$ elements now

$$\overline{\left(\frac{\sigma_{v_i}^2}{\sigma_{\ell_i}^2}\right)} = \frac{r}{n}$$

from which the well known *(Pope, 1976)* and simple formula for an approximate value of σ_{v_i} can be derived:

$$\sigma_{v_i} \simeq \sqrt{\frac{r}{n}} \quad \sigma_{\ell_i} \qquad (5.41)$$

However, it is also true that nowadays the computational cost is not the major problem in a project and additionally, more accurate values

for σ_{v_i}'s are needed when more sophisticated techniques are used for testing the residuals. When a network is quite small (e.g., if it consists of 5-6 points), there is not even a question of computing the Q_v rigorously, if computer facilities are available, but in larger adjustment problems where the number of observations is quite high, attractive techniques have been developed for minimizing the number of computations *(Kok, 1982)* by computing the Q_v matrix rigorously but not completely, for example:

a) computing only the diagonal of Q_v from
$$\text{diag}\,[Q_v] = \text{diag}\,[Q_\ell^{-1}] - \text{diag}\,[AN^{-1}A]$$
and moreover using Choleski factors $N = U^T U$ to avoid the inversion of the whole N matrix;

b) computing the "sparse inverse" of N and use this for the computation of the diag $[Q_v]$.

Also *Grün (1979, 1980)*, has proposed a rigorous calculation of only a few diagonal elements. Since from all the standardized residuals only a small portion exceeds the critical value there is no point in rigorously calculating all the σ_{v_i}'s. Grün suggested computing approximate values $\sigma^o_{v_i}$ for all the residuals first, with a tolerance of $\pm$ f% against the exact values. Then all the observations which give:

a) $|v_i/\sigma^o_{v_i}| > c(1 + \frac{f}{100})$ are rejected

b) $|v_i/\sigma^o_{v_i}| < c(1 - \frac{f}{100})$ are accepted

and for the rest of the observations which are close to the critical region:

$$(1 - \frac{f}{100}) < \frac{v_i}{\sigma_{v_i}} < c\,(1 + \frac{f}{100})$$

exact values σ_{v_i} should be computed. The smaller the f the less the number of computations but the larger the risk of making a wrong decision.

The algorithm is very easy to compute as long as the unknown parameters have been estimated. The exact elements of the Q_v matrix,

$$q^2_{v_{ki}} = \sigma^2_{v_{ki}} / \sigma^2_o$$

are given by:

$$q_{v_{ki}} = \frac{t - \alpha_k x_i}{p_i} \qquad (5.42)^*$$

with

$$t = \begin{cases} 1 & \text{if } k = i \\ 0 & \text{if } k \neq i \end{cases}$$

where α_k is the k^{th} row of the A matrix.

5.7 Sensitivity of the Global Test on the Variance

It was mentioned before that one of the reasons which cause failure of the global test on the variance (5.3), may be the presence of gross errors in the data.

However, the expectation of $\dot{q}^2$, (eq. 5.36), which is $(r - p)\sigma_c^2$ may differ significantly from its actual value, still satisfying the test (5.39), but affecting the maximum value of d, (eq. 5.33), which can stay undetectable by the test (5.37): $\dot{q}^2 < \sigma_o^2 \chi^2_{\alpha;r}$.

---

*For the derivation of (5.42) see *Grün (1979, 1980)*.

This means that failure of the global test does not necessarily depend only on the presence and number of outliers, but also on how good the model is (proper choice of weights, etc.).

However, even when the test (5.3) detects the gross errors, it is entirely unable to locate them. For this purpose the "Data Snooping" technique is used, and to assure that the same decision would come out from both tests, their power is kept constant. Keeping the power of the test constant however, does not provide both tests with the same sensitivity. Sensitivity depends mainly on the nature of the test. Therefore, "Data Snooping" and "tau-test" are the most sensitive ones since they compare each residual against its own standard deviation.

It is quite possible that a significant error, detectable by Data Snooping, may stay hidden in the total sum of the squares of the residuals when the global test is performed, if the other errors turn out to be very small. This probability increases as the redundancy increases, that is, as the tolerance for the quadratic form ($q^2 = \underline{v}^T P \underline{v}$) becomes larger. Therefore it is always recommended (sect. 5.9), to perform the one-dimensional test even if the test on the q^2 has not failed. Only if the tolerance (critical value), is considerably larger than the test statistic, and moreover the redundancy is relatively low, can it be claimed that the system is blunderless.

The insensitivity of the global test has been conceded especially by many Photogrammetists *(Ackermann, Förstner, Grün, Mikhail, Stefanovic)* who have to deal with a great amount of data.

However, the global test retains its importance, since it requires negligibly small computational effort and moreover it gives

a general information about other sources of model errors. If the global test fails, there is definitely something wrong with our basic assumptions (null hypothesis) and the error has be be discovered.

For a better understanding of the insensitivity of the global test, an example is given below.

Example:

The global test (eq. 5.3) can be expressed as:

$$\underline{v}^T C_{\ell}^{-1} \underline{v} < rF_{1-\alpha;r,\infty} = \chi^2_{1-\alpha;r} \qquad (5.43)$$

Assuming that $\underline{\ell}$ is a vector of direction observations

$$\underline{\ell}^T = [18''.5, 19''.5, 20'', 20'', 20''.5, 21''.5]$$

where there is only one unknown parameter (case of direct observations). Obviously:

$$n = 6 , \qquad u = 1 , \qquad r = n - u = 5$$

Taking for instance, $\sigma_o^2 = 1$, as well as P = I, (observations uncorrelated and equally weighted), the L.S. solution gives:

$$\hat{x} = 20''$$

$$C_x = 1/6 = 0.167$$

$$\underline{v}^T = [1.5,\ 0.5,\ 0,\ 0,\ 0,\ -0.5,\ -1.5]$$

$$\hat{\sigma}_o^2 = 1$$

$$\underline{v}^T C^{-1} \underline{v} = r\,\hat{\sigma}_o^2 = 5$$

The critical value for the global test (5.43) is:

$$rF_{1-\alpha;r,\infty} = \chi^2_{1-\alpha;r} = \chi^2_{0.95;5} = 11.05 \qquad (5.44)$$

Applying the global test (5.43), it is tested if:

$$5 < 11.05$$

which is true and the test passes.

But let us see here what happens if the redundancy increases, namely if there are more observations in the vector $\underline{\ell}$.

Obviously, when increasing r, the value of $F_{1-\alpha;r,\infty}$ decreases but not at the same rate (fig. 5.10). More specifically $\chi^2_{1-\alpha;r} = r\, F_{1-\alpha;r,\infty}$ increases as r increases (fig. 5.11).

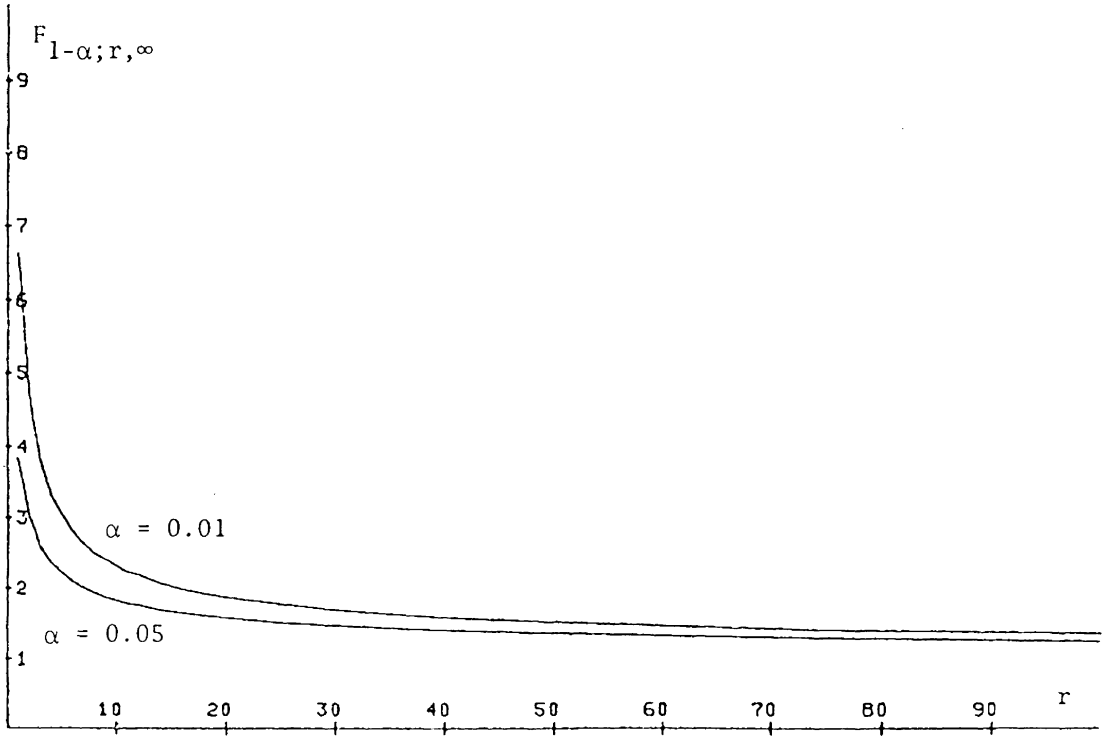


Figure 5.10: Critical values of the $F_{r,\infty}$-distribution

How does this affect the sensitivity of the test? Let us first see the example:

In the previous vector $\underline{\ell}$, five new observations ℓ_i, ($i = 6, \ldots, 11$), are added such as:

$$\ell_6 = \ldots = \ell_{11} = \hat{x}_{1-6} = 20''$$

In this way:

r	$= n - 1 = 10$	increases
$\hat{x}_{1-11}$	$= 20''$	remains the same
$\hat{\sigma}_o^2$	$= 0.5$	decreases
$\underline{v}^T C_\ell^{-1} \underline{v}$	$= (0.5)(10) = 5$	remains the same!

and the critical value

$$\chi^2_{0.95;10} = 18.3 \quad \text{increases.}$$

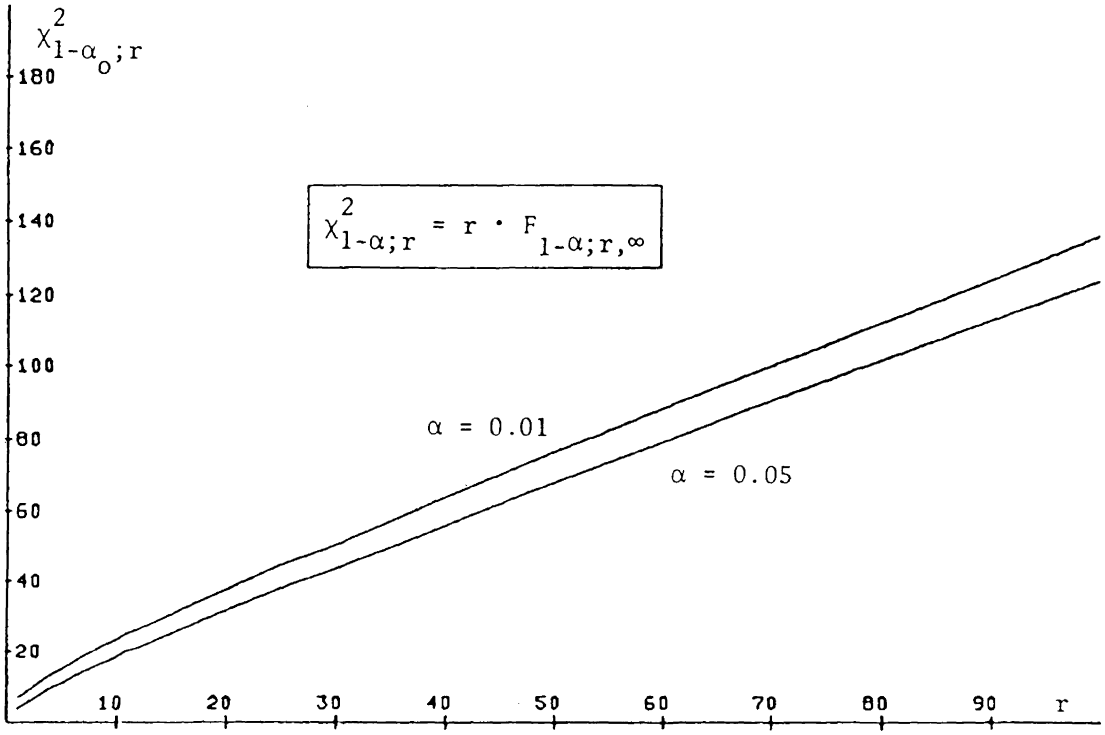


Figure 5.11: Critical values of the χ^2_r-distribution

From a first sight it would be expected that a decreasing $\hat{\sigma}_o^2$ would produce shorter confidence intervals, namely an extreme value ℓ_i should have greater probability to be considered as an outlier. However applying the global test:

$$\underline{v}^T C_{\ell}^{-1} \underline{v} < \chi^2_{1-\alpha;r} \qquad (5.45)$$

$$5 < 18.3$$

it is shown that the tolerance for the quadratic form $\underline{v}^T C_{\ell}^{-1} \underline{v}$ depends on the redundancy r. Comparing (5.44) to (5.45) it becomes obvious that "the larger the redundancy (r) the larger the tolerance". Hence, the global test becomes less sensitive as r increases and a large individual gross error may easily stay hidden in the total sum of $\underline{v}^T C_{\ell}^{-1} \underline{v}$.

NOTE: As far as the two-tailed test (eq. 5.2) is concerned:

$$\frac{r\,\hat{\sigma}_o^2}{\chi^2_{r;1-\alpha/2}} < \sigma_o^2 < \frac{r\,\hat{\sigma}_o^2}{\chi^2_{r;\alpha/2}}$$

For $\alpha = 0.05$ and $r = 10$,

$$\frac{(10)(0.5)}{20.5} < 1 < \frac{(10)(0.5)}{3.25}$$

$$0.24 < 1 < 1.53$$

which is true and the test passes; (this two-tailed test would fail only for $\hat{\sigma}_o^2 < 0.325$ or $\hat{\sigma}_o^2 > 2.05$). Even with $\alpha = 0.10$, which shortens the interval from both sides, the above test passes.

To better illustrate the above idea, let us consider the existence of an outlier in the place of the first observation ℓ_1, for example: $\ell_1 = 16''.5$.

1. With r = 5

$$\hat{x} = 19.67$$

$$\hat{\sigma}^2 = 2.87$$

$$\underline{v}^T P \underline{v} = (2.87)(5) = 14.33$$

$$v_1 = 3.17$$

$$\sigma_{v_i} = \sqrt{\frac{5}{6}} = 0.91$$

According to Baarda (1968), for $\alpha = 0.05$, $\beta_o = 0.80$ and $r = 5$, the significance level for the one-D test (Data Snooping), is: $\alpha_o = 0.008$, $1-\alpha_o = 0.992$ (app. III).

Therefore,

$\chi^2_{0.95;5} = 11.05$: critical value for the global test.

$\sqrt{F_{0.992;1,\infty}} = 2.65$: critical value for the one-D test.

Global test; (test 5.3)

Test if $\underline{v}^T C_{\ell} \underline{v} < \chi^2_{1-\alpha;r}$

$14.33 < 11.05$

which fails; namely there is an indication of large gross errors.

One-D test (Data Snooping; test 5.20)

Test if $\frac{v_i}{\sigma_{v_i}} < \sqrt{F_{1-\alpha_o;1,\infty}}$

$\frac{3.17}{0.91} < 2.65$

which fails; namely the first observation ℓ_1 is considered as an outlier. The rest of the standardized residuals pass the test.

2. With r = 10

(by introducing five new observations equal to: 19.67),

$$\hat{x} = 19.67$$

$$\hat{\sigma}_o^2 = 1.43$$

$$\underline{v}^T P \underline{v} = 14.33$$

$$v_1 = 3.17$$

$$\sigma_{v_i} = \sqrt{\frac{10}{11}} = 0.95$$

Moreover for $\alpha = 0.05$ and $r = 10$, $\alpha_o = 0.0017$, (app. III).

$$\chi^2_{0.95;10} = 18.3$$

$$\sqrt{F_{0.998;1,\infty}} = 3.14$$

Global test; (test 5.3)

Test if $14.33 < 18.3$

which is true. Hence, $\ell_1 = 16''.5$ is not considered as an outlier.

One-D test; (Data Snooping; 5.20)

Test if $\frac{v_i}{\sigma_{v_i}} = \frac{3.17}{0.95} = 3.33 < 3.14$

which fails. Therefore $\ell_1 = 16''.5$ is still considered as an outlier.

Thus, it can be concluded that increasing the redundancy the global test on the variance factor becomes less sharp and sensitive than before.

5.8 Discussion

In studying all aforementioned techniques, it can be said that:

The analysis of the residuals is always required as the final safety measure before the results of the adjustment are given to the user.

The main problems which all these techniques face, are the number of gross errors, their magnitude and their localization, all of which are interconnected.

When a priori knowledge of the accuracy of the observations is available, i.e., σ_o^2 is considered as known, the most sensitive test if the Data Snooping test. Otherwise a posteriori estimates may be used and the tau-test can be employed, which is equally sensitive. The one-dimensional Data Snooping and tau tests are quite similar giving the same test-statistics w and τ in the case of $\hat{\sigma}_o^2 = \sigma_o^2$, so their level of significance α_o can be matched. However, the multi-dimensional tests are different so no comparison or link between their levels of significance α can be made.

In the B-method of testing the multi-dimensional test, which according to *Baarda (1968)* procedes the Data Snooping, is the known global test on the total quadratic form of the residuals (sect. 5.1). This is referred as to the total computing model which is affected by the geometric model, the stochastic model (Q_{ℓ}), and by errors in the data (blunders, outliers, systematic errors).

In the τ-statistic proposed by *Pope (1976)*, the multi-dimensional test is a test on max τ which is a function of the

number of the one-dimensional τ-tests, or of the number of the observations (sect. 5.3). If the knowledge of the a priori variance factor needs to be tested, a separate χ^2-test on the variance factor is usually performed at an independently chosen significance level, e.g., $\alpha = 0.05$ *(Kok, 1982)*.

When the Data Snooping technique is applied, a test similar to the max τ multi-dimensional test would be a max w-test which, however, is not performed at this time. After that, uni- and multi-dimensional tests in w and τ statistics would be parallel and comparable.

Tau-test does not account for the type II error, therefore no reliability statements concerning the sensitivity of the detection (similar to eq. 5.16), can take place. The use of the a posteriori $\hat{\sigma}_o^2$ has a serious disadvantage *(Stefanovic, 1980)*: the estimated $\hat{\sigma}_o^2$ is affected by the presence of gross errors, so that in the case of a large error among the data $\hat{\sigma}_o^2$ increases, the statistic τ_i decreases and the test becomes too insensitive.

The major problem of all the post-adjustment techniques is the localization of the errors. Both Data Snooping and tau-test were designated for detection purposes only. Förstner has shown *(Stefanovic, 1980)* that when only one outlier exists in the data and the diagonal elements of the covariance matrix of the residuals are dominant, the maximum standardized residual likely indicates the error. But if there is more than one outlier and the residuals are significantly correlated, it is very difficult and sometimes impossible to locate precisely the erroneous observations. The use of the quadratic forms (sect. 5.4), solves the problem somewhat by giving the flexibility

of examining groups of residuals simultaneously, but again the effectiveness of the method is reduced if the residuals are highly correlated.

Only a systematic iterative testing procedure would give the best results without the guarantee of complete success in every possible case. A look at the correlation matrix of the residuals R, and at the redundancy numbers (sect. 3.2), would give a good picture of the effectiveness of the testing procedure that would follow. It is inevitable to have problems with any of the techniques mentioned since all of them work with residuals which are not the true observational errors but functions of them. The efficiency of any of these techniques depends on the behaviour of the model under the presence of gross errors.

Neglecting the very special cases where problems are unavoidable, a strategy is proposed below for the treatment of gross errors in the data.

5.9 Proposed Strategy

In the following proposed strategy a few assumptions have been made:

The large gross errors have been detected and eliminated at an early stage with other means, e.g., pre-adjustment techniques, etc. Here only the treatment of errors of small magnitude is considered.

Experience shows that in a complete survey project the number of gross errors is quite limited *(Stefanovic, 1978)*. It is assumed that not only does it not exceed the redundancy of the system, but also that the maximum allowable number of gross errors can be pre-established.

In usual survey networks the redundancy is also quite bounded, so there is a reluctance to reject observations which not only have a high cost but also are difficult to repeat. In this case it is recommended that observations should never be rejected blindly without first checking to find a good reason for the failure of the test.

The domination of the Q_v matrix along the diagonal can be tested by means of the correlation matrix R:

$$R = D^{-1/2} Q_v D^{-1/2} \tag{5.46}$$

where $D = \text{diag}(Q_v)$.

If there is strong correlation among the residuals, automatic rejection cannot be applied. For this reason programs should only flag observations and never reject them automatically, although the statistical design of the test makes it possible.

It has been also assumed that there is an a priori knowledge of the accuracy of the observations. If not, the global test cannot be used and the Data Snooping is substituted by the tau-test.

The procedure which can be viewed in figure (5.12), is now described below:

To check the consistancy of the field observations it is logical to run first an adjustment with minimum fixed information. Therefore, only minimum constraints should be used in the adjustment, which would give $q^2 = \underline{v}^T P \underline{v}$ equal to q^2 obtained when no constraints are used *(Uotila, 1976)*.

The first global test on $q^2 = \underline{v}^T P \underline{v}$, depending on the redundancy of the system, gives information mainly about the proper

selection of the weights and the presence of at least large errors in the data. If the test fails, there is definitely something wrong with the basic assumptions (H_o). If the alternative hypothesis assumes the presence of large errors, these can be confined and examined by testing the quadratic forms d (eq. 5.33), and $\dot{q}^2$ (eq. 5.34), without performing a new adjustment, as it is shown in the flow-chart below. If $\dot{q}^2 < \sigma_o^2 \chi^2_{\alpha;(n-p)}$ all of at least the large errors have been confined in d and can therefore be withdrawn.

For detecting outliers of small magnitude, albeit the global test passes, it is necessary to resort to the Data Snooping again. Each time only the observation ℓ_i which corresponds to the largest standardized residual $w_i = v_i/\sigma_{v_i}$ that exceeds the critical value $\sqrt{F_{1-\alpha_o;1,\infty}}$, may be excluded from the following adjustment, provided that it had been checked very carefully before it was rejected. The scheme is very effective as long as the residuals are not highly correlated. Otherwise the case of reintroducing a previously withdrawn observation and rejecting a new one should be examined.

The decision depends on the size of the redundancy r and on the influence of an error on the final solution. In small survey networks where r is not very large, we try not to reject observations whenever possible, not only because of their cost, but also because with every new rejection, r decreases and the method becomes less effective. Anyway, reintroducing previously withdrawn observations, inevitably brings a complexity into the procedure.

If the above procedure is followed and the redundancy has not been dangerously decreased, after the test on $\dot{q}^2$ passes, the global test on the following adjustments will always pass.

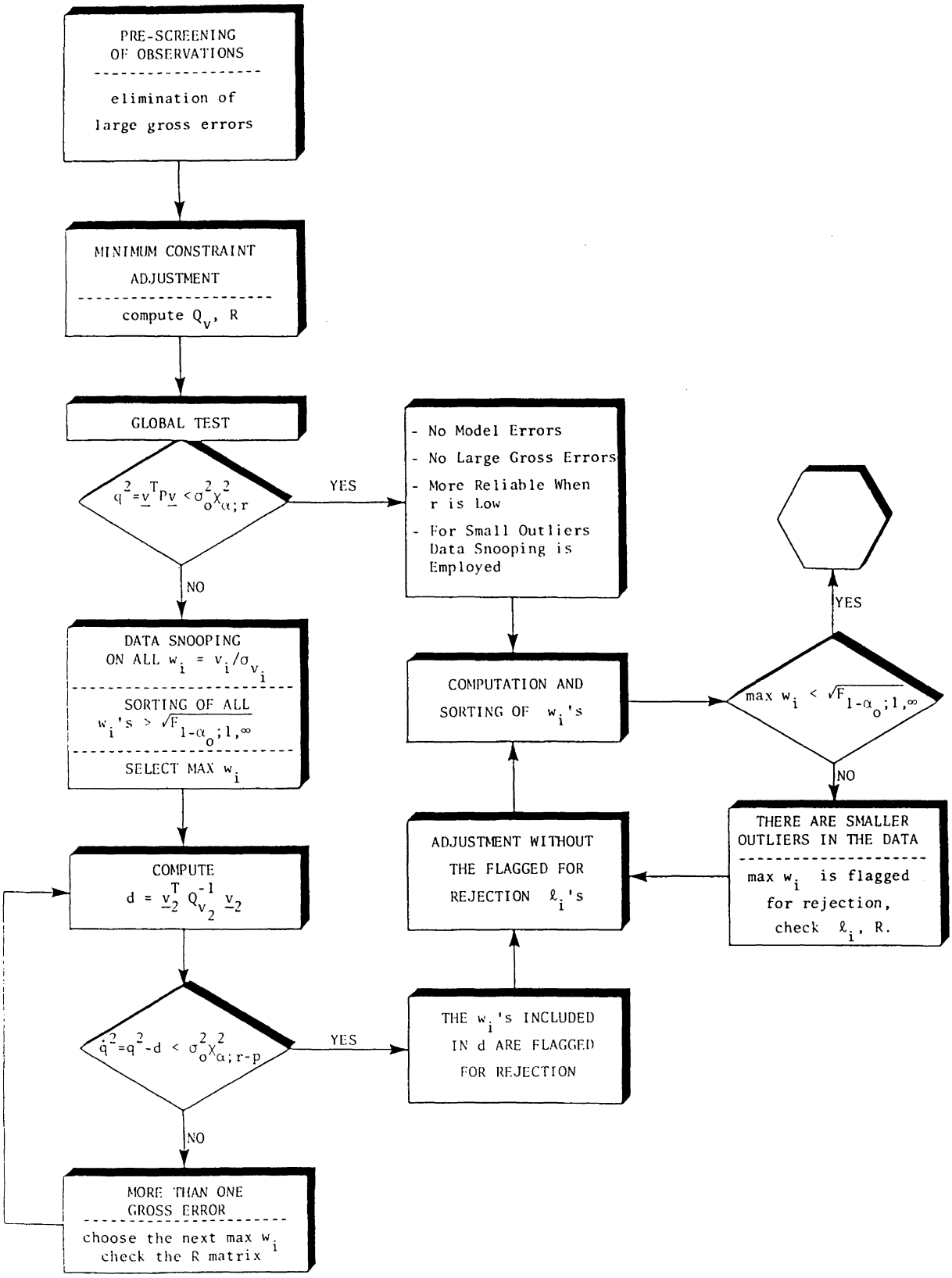


Figure 5.12: Proposed Strategy

The second time the Data Snooping is applied, all the standardized residuals and not only the ones that failed at the first time should be tested. This is proposed first because the model changes, and secondly because the q_{v_i} elements decrease as the redundancy decreases.

The procedure ends when the global test and the one-dimensional test for all the w_i's pass. After that the system is considered as being blunderless.

5.10 Undetectable Errors

No matter how sensitive a technique is, it cannot detect gross errors of arbitrarily small magnitude. Therefore, even "Data Snooping" has a finite capability of detecting small gross errors, and after the detection and elimination there are still blunders remaining in the data. This sensitivity and its influence on the estimated parameters will be treated extensively in the next chapter introducing the concept of the "reliability" of an estimation.

6

RELIABILITY OF NETWORKS

The detection and elimination of gross errors in a model is very important, but having a measure of the sensitivity of the specific technique is also indispensable.

This sensitivity is expressed by the "internal reliability" which refers to the maximum undetectable error resulting from using this technique.

On the other hand these possible errors remaining in the data may have a significant effect on the final estimation. The concept of the "external reliability" relates to this maximum influence of hidden errors on the final solution.

Both internal and external reliability are described by boundary values which will be derived in the following sections.

6.1 Internal Reliability

As was mentioned before, the internal reliability is assessed by the lower values $\nabla_o \ell_i$, of gross errors $\nabla \ell_i$, which can just be detected by the test (Data-Snooping), with a given probability β_o, if the test has confidence level $1 - \alpha_o$.

The lower bound $\nabla_o \ell_i$ is related to the lower bound of the non-centrality parameter λ_o, (eq. 5.9) which depends only on the chosen α_o and β_o. The expressions for the evaluation of these lower bounds have already been derived and are given by the equations (5.14) and (5.16). With a diagonal weight matrix P, equation (5.16) is simplified to:

$$|\nabla_o \ell_i| = \sigma_o \sqrt{\frac{\lambda_o}{p_i (Q_v P)_{ii}}} = \sigma_o \sqrt{\frac{\lambda_o}{p_i \, r_i}} \qquad (6.1)$$

where $r_i = (Q_v P)_{ii}$ is the redundancy number of the i^{th} observation (eq. 3.4). From equation (6.1) it follows that:

$$|\nabla_o \ell_i| = \sqrt{\frac{\lambda o}{r_i}} \; \sigma_{\ell_i} \qquad (6.2)$$

From this expression it appears that a just detectable gross error in an observation ℓ_i, relies on its precision σ_{ℓ_i}.

The coefficients $h_{\ell_i} = \sqrt{\lambda_o / r_i}$ show the sensitivity of the test. It is desirable to have small values of h_{ℓ_i} in a network. The smaller the h_{ℓ_i}, the larger the r_i. Large r_i values imply that a gross error ∇_{ℓ_i}, in an observation ℓ_i, will be more clearly reflected in the corresponding residual v_i, and consequently easily revealed through the testing procedure.

Hence, large redundancy numbers produce lower bounds $\nabla_o \ell_i$ for the internal reliability. In other words the tests become more sensitive and the number of undetectable gross errors is reduced to the minimum.

If the redundancy is uniformly distributed in the network, then all r_i's and therefore all $\nabla_o \ell_i$'s, are practically the same, as long as P = I. However in general, this does not happen since large differences between r_i's may be noticed in the different parts of networks, especially in the case of different kinds of observables. For this reason instead of calculating all the $\nabla_o \ell_i$ values for the different r_i's, a global reliability measure is more often desirable, using the average redundancy of the network:

$$\overline{r_i} = \frac{tr[Q_v P]}{n} = \frac{r}{n}$$

Thus equation (6.2) becomes:

$$|\overline{\nabla_o \ell}| = \sqrt{\frac{\lambda_o \, n}{r}} \; \sigma_{\ell_i} \qquad (6.3)$$

Experience has shown that for ordinary networks $r/n \tilde{=} 0.50$ *(Pope, 1976)*. Using equation (6.3) for $\alpha_o = 0.001$ and $\beta_o = 0.80$ (therefore $\lambda_o = 17$),

$$|\overline{\nabla_o \ell_i}| = \sqrt{\frac{17}{0.50}} \; \sigma_{\ell_i} \tilde{=} 5.8 \; \sigma_{\ell_i}$$

Thus in most of the networks a gross error $\nabla \ell_i$, (in one observation ℓ_i with $r_i \tilde{=} 0.50$), smaller than $5.8 \, \sigma_{\ell_i}$ will not be detected examining the residuals using the "Data Snooping" method (at the above probability levels α_o, β_o).

6.2 External Reliability

External reliability relates to the maximum effect of possible undiscovered observational gross errors $\nabla\ell_i$, on the results of the adjustment (e.g., coordinates). This influence $\nabla_o\hat{x}_i$ is given by:

$$\underline{\nabla_{o,i}\hat{x}} = (A^TPA)^{-1} A^TP \underline{|\nabla_o\ell_i|} \tag{6.4}$$

If there are any nuisance parameters u_t in the model, while we are interested only in the desired parameters, u_k, the influence is given now by equation (II.2)

$$\underline{\nabla_{o,i}\hat{k}} = Q_k\bar{A}_k^TP\underline{|\nabla_o\ell_i|} \tag{6.5}$$

where $\bar{A}_k = (I - A_t(A_t^TPA_t)^{-1}A_t^TP)A_k$, (eq. II.3). Assuming again only one gross error in the i^{th} observation, the $\underline{c}_i$ vector can be employed again and equation (6.5) becomes,

$$\underline{\nabla_{o,i}\hat{k}} = Q_k\bar{A}^TP\underline{c}_i|\nabla_o\ell| \tag{6.6}$$

where $|\nabla_o\ell|$ is a scalar, coming from equation (6.3)

The above expressions are very revealing but they are not general estimates of the usual size of external reliability $\nabla\hat{x}$. They require a large number of computations since for each estimated parameter there are n different components $\nabla\hat{x}$, when testing under conventional hypotheses, one for each hypothesis (eq. 5.15). Usually these $\nabla\hat{x}$ vary significantly in size. Moreover they are dependent on the coordinate definition. Hence, Baarda and De Haus *(Baarda, 1976, 1979)* proposed the new standardized general variate:

$$\bar{\lambda}_{i,o} = \frac{1}{\sigma_o^2} (\nabla_{i,o}\hat{x})^T Q_x^{-1} (\nabla_{i,o}\hat{x}) \qquad (6.7)^*$$

which is invariant with respect to the coordinate definition and can be considered as a measure of the reliability of the results. It is desirable to have $\bar{\lambda}_{i,o}$ approximately constant for all i's so that the ability of detecting gross errors is the same in every part of the network. One is interested mainly in the maximum value of this variate which is related to the minimum deviation from the null hypothesis that can be detected with a certain probability β_o. So the following variate *(Baarda, 1976)*:

$$\bar{\lambda}_o = \max (\bar{\lambda}_{i,o}) \qquad (6.8)$$

can be considered as a measure of the external reliability. Substituting σ_o^2 from equation (5.13) and using also the vector $\underline{c}$ (eq. 6.6), equation (6.7) becomes:

$$\bar{\lambda}_{o,i} = \frac{\nabla_o \ell \underline{c}^T P A Q_x^T Q_x^{-1} Q_x A^T P \underline{c} \nabla_o \ell}{\nabla_o \ell^2 \underline{c}^T P Q_v P \underline{c}} \lambda_o$$

$$= \frac{\underline{c}^T P A Q_x A^T P \underline{c}}{\underline{c}^T P Q_v P \underline{c}} \lambda_o \qquad (6.9)$$

Assuming that P is diagonal and that only one gross error affects the observations, equation (6.9) yields:

$$\bar{\lambda}_{o,i} = \frac{p_i (A Q_x A^T P)_{ii}}{p_i (Q_v P)_{ii}} \lambda_o = \lambda_o \frac{u_i}{r_i} \qquad (6.10)$$

---

* $\nabla_o \hat{k}$ can be used in the place of $\nabla_o \hat{x}$ as well.

If there are nuisance parameters in the model, they are not considered in this general variate. So equation (6.5) is used in (6.6) and Q_x becomes Q_k, which is referred only to the desired parameters $\hat{k}$ (e.g., coordinates). Under the same assumptions, we end up with:

$$\bar{\lambda}_{o,i} = \lambda_o \frac{u_{k_i}}{r_i} \tag{6.11}$$

instead of equation (6.9). It is to be remembered (eq. 3.17), that:

$$u_{k_i} = 1 - r_i - u_{t_i}$$

where u_{t_i} is referred to the nuisance parameters and can be estimated from equation (II.8).

However, this value (eq. 6.11), is rather abstract to the practical geodesist who would prefer to have something more palpable than the above variate. What would be more useful is the maximum effect of this variate on the unknown parameters. Equation (6.6), using the $\underline{c}_i$ vector gives (*Forstner, 1981.b*)

$$\underline{\nabla_{o,i}\hat{k}} = Q_k \underline{\bar{a}}_i \; p_i |\nabla_o \ell_i| \tag{6.12}$$

where

$$\underline{c}_i^T = [0 \ldots 0 \;\; 1 \;\; 0 \ldots 0]$$

$p_i |\nabla_o \ell|$: is a scalar

$$\underline{\bar{a}}_i = [\bar{a}_{i1} \;\; \ldots \;\; \bar{a}_{iu}]$$

If one is interested only in the effect of the gross error on one specific unknown parameter k_p, equation (6.12) is projected onto direction k_p using a vector (u x 1), $e_p^T = [0 \ldots 0 \;\; 1 \;\; 0 \ldots 0]$ where

the "1" corresponds to the p^{th} element. Therefore eq. (6.12) yields:

$$\nabla_{o,i}\hat{k}_p = \underline{e}_p^T Q_k \underline{\bar{a}}_i \; p_i \, |\nabla_o \ell_i| \tag{6.13}$$

where the form $\underline{e}_p^T Q_k \underline{\bar{a}}_i^T$ can be written as:

$$\underline{e}_p^T Q_k \underline{\bar{a}}_i^T = \underline{e}_p^T S^T S \underline{\bar{a}}_i^T = \underline{\dot{e}}_p^T \underline{\dot{\bar{a}}}_i^T \tag{6.14}$$

where:

$$Q_k = S^T S$$

$$\underline{\dot{e}}_p = S\underline{e}_p$$

$$\underline{\dot{\bar{a}}}_i^T = S\underline{\bar{a}}_i^T$$

The scalar product of the two vectors $\underline{\dot{\bar{a}}}$ and $\underline{\dot{e}}_p$ is given by *(see Förstner, 1981.b)*,

$$\underline{\dot{e}}_p^T \underline{\dot{\bar{a}}}_i^T = ||\underline{\dot{e}}_p|| \; ||\underline{\dot{\bar{a}}}_i|| \cos(\underline{\dot{e}}_p, \underline{\dot{\bar{a}}}_i) \cos(\underline{\dot{e}}_p, \underline{\dot{\bar{a}}}_i) \leq ||\underline{\dot{e}}_p|| \; ||\underline{\dot{\bar{a}}}|| \tag{6.15}$$

which obviously has its largest value when $\underline{\dot{\bar{a}}}_i$ and $\underline{\dot{e}}_p$ are parallel. Considering the lengths of the above vectors:

$$||\underline{\dot{e}}_p|| = \sqrt{\underline{\dot{e}}_p^T \underline{\dot{e}}_p} = \sqrt{\underline{e}_p^T S^T S \underline{e}_p} = \sqrt{\underline{e}_p^T Q_k \underline{e}_p} = \sigma_{k_p}/\sigma_o$$

$$||\underline{\dot{\bar{a}}}|| = \sqrt{\underline{\dot{\bar{a}}}\,\underline{\dot{\bar{a}}}^T} = \sqrt{\underline{\bar{a}}_i S^T S \underline{\bar{a}}_i^T} = \sqrt{\underline{\bar{a}}_i Q_k \underline{\bar{a}}_i^T} = \sqrt{u_{k_i}/p_i}$$

equation (6.15) becomes:

$$\underline{\dot{e}}_p^T \underline{\dot{\bar{a}}}_i^T \leq \frac{\sigma_{k_p}}{\sigma_o} \sqrt{\frac{u_{k_i}}{p_i}} \tag{6.16}$$

and equation (6.13) can be written as:

$$\nabla_{o,i}\hat{k}_p \leq \frac{\sigma_{k_p}}{\sigma_o}\sqrt{\frac{u_{k_i}}{p_i}}\; p_i |\nabla_o \ell_i| \tag{6.17}$$

Considering also the expression for the internal reliability (6.2), equation (6.17) results in:

$$\nabla_{o,i}\hat{k}_p \leq \frac{\sigma_{k_p}}{\sigma_o}\sqrt{\frac{u_{k_i}}{p_i}}\; p_i \;\; \sigma_o\sqrt{\frac{\lambda_o}{r_i p_i}}$$

$$\leq \sigma_{k_p}\sqrt{\frac{\lambda_o u_{k_i}}{r_i}} \tag{6.18}$$

where the quantity $\sqrt{\lambda_o\, u_{k_i}/r_i}$ has been already defined (6.11) as the general variate $\sqrt{\bar{\lambda}_{o,i}}$ for the external reliability. Therefore equation (6.18) can be written as:

$$\nabla_{o,i}\hat{k}_p \leq \sigma_{k_p}\sqrt{\bar{\lambda}_{o,i}} \tag{6.19}$$

where

$$\bar{\lambda}_{o,i} = (\lambda_o u_{k_i}/r_i).$$

An average value of (u_{k_i}/r_i) leads to the global figure for the external reliability, similar to the expression (6.3),

$$\overline{\nabla_{o,i}\hat{k}_p} \leq \sigma_{k_p}\sqrt{\frac{\lambda_o u_k}{r}} \tag{6.20}$$

where

u_k : number of desired unknown parameters, and

r : total redundancy.

Namely, the effect of a non-detectable gross error $\nabla_o \ell_i$ on the results

of the adjustment can be as much as $\sqrt{\lambda_o \, u_k/r}$ times the computed precision σ_k of the desired parameters $\hat{k}$, but of course it can be much less *(Mikhail, 1979)*.

Using the standardized measure for external reliability (eq. 6.7), one can make statements about the size of error influence on the estimated parameters. Since the scalars $\sqrt{\bar{\lambda}_{o,i}}$ (eq. 6.11), depend only on the shape of the network and not on the actual measurements, they can be utilized during the design of the network as a bound of the external reliability *(van Mierlo, 1981)*. A criterion based on practical experience is often used *(Kok, 1982)*, which gives: $\sqrt{\bar{\lambda}_{o,i}} \leq 10$ for each observation ℓ_i. If this has been successful, the maximum interval for the coordinates after the adjustment, with a certainty of $\beta_o = 80\%$, will be:

$$\hat{x}_p - \sqrt{\bar{\lambda}_{o,i}}\, \sigma_{x_p} \leq x_p \leq \hat{x}_p + \sqrt{\bar{\lambda}_{o,i}}\, \sigma_{x_p} \qquad (6.21)$$

Therefore it is desirable to have small values of u/r which would produce lower bounds for $\nabla_o \hat{x}$, thus better reliability for the network.

7

EXAMPLE ON A REAL GEODETIC NETWORK

This example is referred to a network of 6 points (fig. 7.1), established to detect deformations of the Lohmühle dam in Luxemburg. The observations considered below are part of several epochs of observations which have been used for studies by the Working Groups of the FIG "ad hoc" committee on deformation analysis, (*Chrzanowski, 1981*).

As it is shown in the following analysis, the network was very accurately surveyed; 27 directions were observed and points 5 and 6 were fixed as minimum constraints in the adjustment. The chosen network was very consistent to ensure that any a posteriori inconsistency would be only due to the simulated inserted outliers.

7.1 Gross Error Detection - Alternative Hypotheses

Assuming that there was a gross error in centering the instrument on point #3, of 1 mm (unacceptable for forced centering), observations were generated (table 7.1), and substituted for the four directions which had been observed in reality from this point.

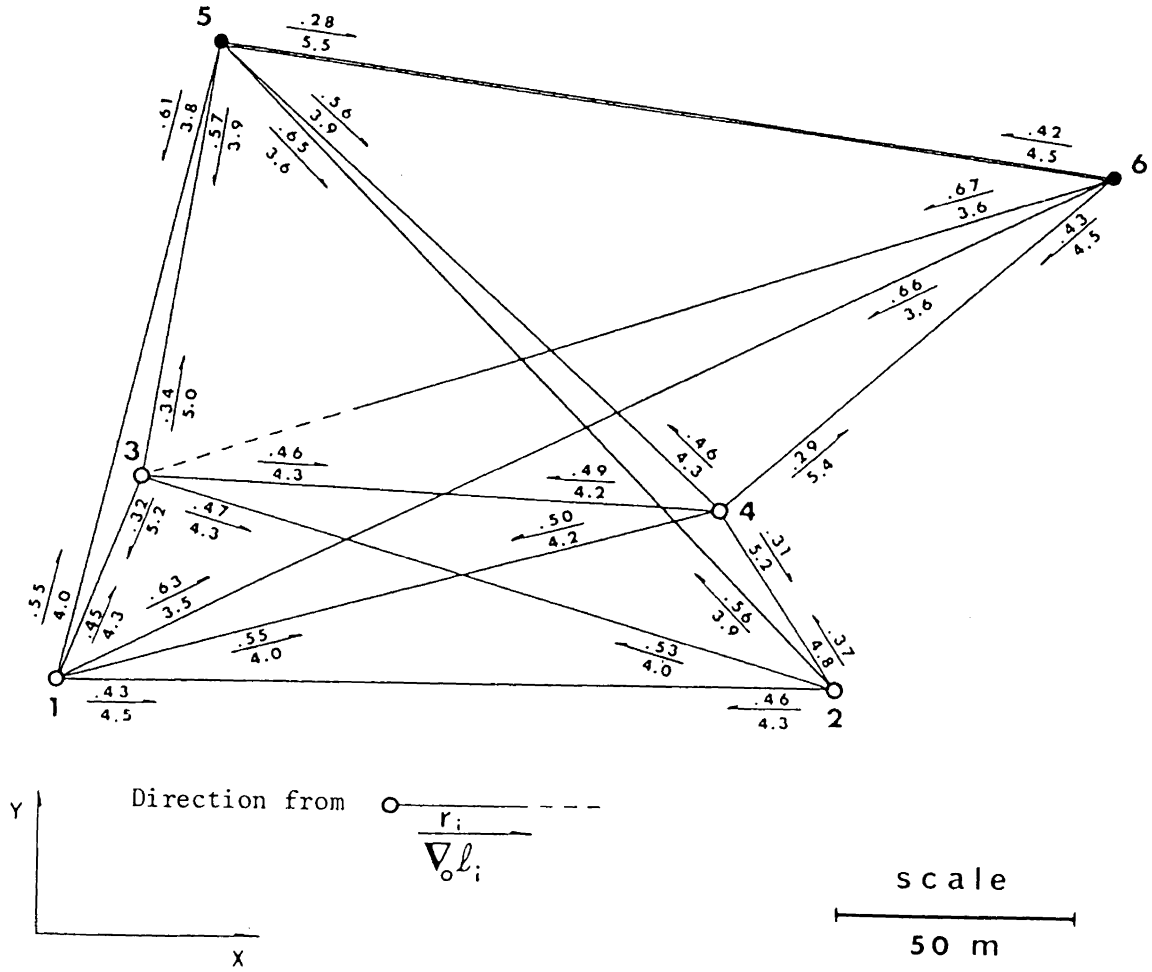


Figure 7.1: Geometrical strength of the test network

After the adjustment the standardized residuals were sorted according to their magnitude (table 7.1). The redundancy was r = 13. Therefore the critical value for the one-dimensional test (Data Snooping), for a significance level α_o = 0.001 (α = 0.06) gives:

$$\sqrt{F_{1,\infty;\alpha_o}} = 3.29 \ .$$

For this case the largest residuals are indicated also below:

Direction From	To	Residual v_i	Standard Deviation From C_v	Standard Deviation From Approximation	Standardized Residual
3	1	1.91	$\sqrt{.162}$	$\sigma_{v_i} = \sqrt{\frac{r}{n}}\,\sigma_\ell$	4.75
1	3	-2.06	$\sqrt{.229}$	$= \sqrt{\frac{13}{27}}\,(.71) = .49$	4.30
3	5	-1.27	$\sqrt{.173}$		3.05
4	5	1.18	$\sqrt{.231}$		2.46
⋮	⋮	⋮	⋮	(not used here)	⋮

Compared to the critical value, the first two residuals are flagged for rejection.

It may happen that the largest standardized residual does not correspond to the observation which in fact was erroneous, (an example is given later). In this example the largest standardized residual corresponds to the direction 3-1 which was affected by 5".38 (due to the eccentricity). The fact that the direction 1-3 has also a large residual is due to the model and this can be explained by the large correlation between the v_{1-3} and v_{3-1} residuals:

$$\rho = \frac{\sigma_{v_{1-3}\,v_{3-1}}}{\sigma_{v_{1-3}}\,\sigma_{v_{3-1}}}$$

$$= \frac{-0.162}{\sqrt{(0.162)(0.229)}} = -0.84$$

Now if we had the slightest idea about the reason for the rejection of the H_0 : (no existence of any systematic errors or blunders), e.g.,

suspecting errors due to eccentricity, this alternative hypothesis H_a could be tested by examining the quadratic forms of groups of residuals (sect. 5.4). For example, testing the H_a: "all the observations from station #3 were affected by a systematic error". Therefore the tests (5.35) (5.36) and (5.37) can be performed, where the size of the test group $\underline{\ell}_2$ is p = 4 (all the observations from station #3). Moreover, the selected significance level is a = 0.06 and the redundancy r = 27 - 14 = 13. Now it is tested if:

a. $$q^2 = \underline{v}^T P \underline{v} < \sigma_o^2 \cdot \chi^2_{\alpha;r}$$

that is $36.8 < 22$ $\therefore$ it fails.

b. $$d = \underline{v}_2^T Q_{v_2}^{-1} \underline{v}_2 < \sigma_o^2 \cdot \chi^2_{\alpha;p}$$

that is $25.3 < 9.4$ $\therefore$ it fails.

c. $$\dot{q}^2 = \dot{\underline{v}}_1^T p_1 \dot{\underline{v}}_1$$

$$= q^2 - d^2 < \sigma_o^2 \cdot \chi^2_{\alpha;(r-p)}$$

that is $11.5 < 18.31$ $\therefore$ it passes.

It should be noticed that in the last test the redundancy for the critical value was not equal to (r-p) = 13 - 4 = 9 but (r-p+1) = 10, since taking out all the directions from station #3 the nuisance orientation unknown parameter is also eliminated. Therefore $\chi^2_{.06;10} = 18.31$.

According to the results of the tests the alternative hypothesis H_a is accepted. Summarizing, it can be said that the method gave satisfactory results and the outlying observations were removed so that the last test passed.

However, there is not only one alternative hypothesis H_a, to be tested when H_o is rejected. The problem that arises is now:

Which H_a to accept, considering that any of H_a's which happen to reject the observations, which were mostly affected by gross errors, will probably pass the test? In other words, how to ensure that in accepting the specific H_a, all the erroneous observations but only these were rejected.

In the previous example, someone who applied the testing procedure with that H_a, would never be fully convinced that all the four rejected observations were in fact outliers and no other gross errors, among the rest of the observations, have remained undetectable but still affect the final solution. For example, if the H_a: "the directions 3→1 and 1→3 were affected by refraction (thus the large residuals)", was tested, this test would have passed, too, because the gross errors in the rest of the observations were absorbed during the adjustment and did not produce large enough v_i's. In this case, acceptance of this specific H_a would mean that:

a. We committed error in rejecting the direction 1-3 although it was not an outlier.

b. We committed error in not rejecting the directions 3-5, 3-4 and 3-2 although they were erroneous.

If now "Data-Snooping" is employed, examining one H_a at the time for each single observation, the first observation which will be rejected is the direction 3-1. After performing the adjustment for a second time, the new standardized residuals are under the critical value and nothing can be rejected (see table 7.1). But although everything with the observations seems to be acceptable, the adjusted coordinates differ from the initial ones up to 0.5 mm.

If the eccentricity now increases (dx_3 = +1.5 mm) observations

SIMULATED ECCENTRICITY ON x-direction OF ST. 3		INITIAL ($dx_3 = 0$)		dx_3 = +1 mm		dx_3 = +1.5 mm		dx_3 = +2.0 mm	
Generated directions from station 3 due to the eccentricity	3-5	0° 0' 0".00		0° 0' 0".00		0° 0' 0".00		0° 0' 0".00	
	3-4	83 26 10.42		83 26 12.97		83 26 14.07		83 26 15.22	
	3-2	97 22 4.91		97 22 6.87		97 22 7.97		97 22 10.00	
	3-1	192 35 16.52		192 35 21.90		192 35 25.14		192 35 28.38	
Initial coordinates and displacements due to the new observations	STATION	x(m)	y(m)	$\Delta x \cdot 10^4$	$\Delta y \cdot 10^4$	$\Delta x \cdot 10^4$	$\Delta y \cdot 10^4$	$\Delta x \cdot 10^4$	$\Delta y \cdot 10^4$
	1	10.0071	10.0005	+1	+6	+2	+8	+2	+10
	2	170.7584	10.0097	-3	+4	-3	+5	-3	+4
	3	26.8448	52.4324	+5	+2	+8	+2	+11	+3
	4	146.5610	47.2207	-1	+1	-1	+1	-1	+2
Sorted and rejected stand. residuals by the Data-Snooping test	3-1	none of the standardized residuals is rejected		$(1.91/\sqrt{.162})$: rej.		$(-2.71/\sqrt{\ })$: rej.		$(-3.4/\sqrt{\ })$: rej.	
	1-3			$(-2.06/\sqrt{.229})$: rej.		$(2.94/\sqrt{\ })$: rej.		$(3.76/\sqrt{\ })$: rej.	
	3-5			----------		$(1.87/\sqrt{\ })$: rej.		$(2.59/\sqrt{\ })$: rej.	
	5-3			$(-1.72/\sqrt{.173})$: acc.		----------		$(-1.89/\sqrt{\ })$: rej.	
	1-5			.		. : acc.		$(1.67/\sqrt{\ })$: rej.	
				.		.			
New adjustment after removing only the first direction 3-1 (if rejected)	STATION	x(m)	y(m)	$\Delta x \cdot 10^4$	$\Delta y \cdot 10^4$	$\Delta x \cdot 10^4$	$\Delta y \cdot 10^4$	$\Delta x \cdot 10^4$	$\Delta y \cdot 10^4$
	1	10.071	10.0005	+2	+2	+3	+3	+4	+4
	2	170.7584	10.0097	-4	+2	-5	+2	-5	0
	3	26.8448	52.4324	+3	+5	+4	+7	+6	+9
	4	146.5610	47.2207	-2	0	-3	0	-3	0
Data Snooping on the new standardized residuals w_i's		none of the w_i's is rejected		none of the w_i's is rejected		none of the w_i's is rejected		for direction 3-5 $w_i = \frac{1.46}{\sqrt{.173}}$: is rejected	

Table 7.1: Testing of the simulated eccentricities.

are generated again (see table 7.1). After the rejection of the direction 3-1, the remaining gross errors in the data affect the final solution up to 0.7 mm. This value shows the effect of an undetectable gross error in the final solution (external reliability) of which a more extreme case will be given below.

7.2 Geometrical Strength of the Network

The analysis of the present network has up to now been confined to the study of the efficiency of different approaches to the identification of gross errors in the data (based on ch. 5). However, although it gives an idea of the sensitivity of the methods, a more systematic analysis of the strength of the network is required. By "strength" of the network is implied its ability to resist distortions caused by the presence of inconsistant observations. The strength of the network can be revealed by looking at the redundancy and absorption numbers r_i and u_i respectively (sect. 3.2).

In the present network there are 6 orientation parameters, one for each station, therefore u_i consists of a u_{k_i} and a u_{t_i} part. The computed values for the redundancy and the absorption numbers for each observation are given in table (7.2). If it is assumed that only one observation ℓ_i has been affected by a gross error $\nabla\ell_i$, then (see eq. 3.5):

$$\nabla v_i = r_i \, \nabla\ell_i$$

and moreover (eq. 3.18):

$$\nabla\ell_i = u_{k_i} \, \nabla\ell_i + u_{t_i} \, \nabla\ell_i + r_i \, \nabla\ell_i$$

where of course $u_{k_i} + u_{t_i} + r_i = 1$

and

$\nabla v_i = r_i \nabla \ell_i$: is the part of $\nabla \ell_i$ shown in the corresponding residual v_i.

$u_{k_i} \nabla \ell_i$: is the part of $\nabla \ell_i$ absorbed during the determination of the unknown desired parameters k.

$u_{t_i} \nabla \ell_i$: is the part of $\nabla \ell_i$ absorbed during the determination of the unknown nuisance parameters t.

There is one r_i, u_i, u_{k_i} and u_{t_i} for each of the 27 directions. These values were computed from

$$r_i = (Q_v P)_{ii} \qquad \text{(eq. 3.4)}$$

$$u_i = (A(A^T P A)^{-1} A^T P)_{ii} \qquad \text{(eq. 3.10)}$$

$$u_{t_i} = (A_t (A_t^T P A_2)^{-1} A_t^T P)_{ii} \qquad \text{(eq.IL7)}$$

and $u_{k_i} = u_i - u_{t_i}$

It is obvious that if r_i is large the error $\nabla \ell_i$ is revealed better in v_i. If r_i is small (hence u_i is large), the $\nabla \ell_i$ is well absorbed in the determination of the unknown parameters and will not be easily detected.

In the example there are different values r_i for each observation ℓ_i, from $r_{19} = 0.28$ up to $r_{27} = 0.67$. Therefore, a gross error in direction #27 has a greater probability of detection than in direction #19, simply because in the first case $r_{27} = 0.67$ of a gross error $\nabla \ell_i$ is reflected in v_{27}, where in the second case only $r_{19} = 0.28$ of a gross error $\nabla \ell_i$ is revealed in v_{19}.

For the observations from point #3, $\bar{r}_{10-13} = 0.40$ is obtained which warns us not to expect a very easy detection of probable gross

#	Direction From-To	r_i	u_i	u_{k_i}	u_{t_i}	$\lvert \nabla_o \ell_i \rvert$ in ["]
1	1-5	0.546	0.454	0.254	0.2	3.96
2	1-3	0.454	0.546	0.346	0.2	4.34
3	1-6	0.63	0.37	0.17	0.2	3.69
4	1-4	0.548	0.452	0.252	0.2	3.95
5	1-2	0.43	0.47	0.37	0.2	4.46
6	2-5	0.555	0.445	0.195	0.25	3.93
7	2-4	0.373	0.627	0.377	0.25	4.8
8	2-1	0.458	0.542	0.292	0.25	4.33
9	2-3	0.525	0.475	0.225	0.25	4.04
10	3-5	0.343	0.657	0.407	0.25	4.99
11	3-4	0.459	0.541	0.291	0.25	4.32
12	3-2	0.471	0.529	0.279	0.25	4.26
13	3-1	0.431	0.679	0.429	0.25	5.17
14	4-5	0.459	0.541	0.341	0.2	4.32
15	5-6	0.29	0.71	0.51	0.2	5.43
16	4-2	0.314	0.686	0.486	0.2	5.22
17	4-1	0.489	0.511	0.311	0.2	4.19
18	4-3	0.494	0.506	0.306	0.2	4.16
19 *	5-6	0.284	0.716	0.516	0.2	5.49
20	5-4	0.564	0.436	0.236	0.2	3.9
21	5-2	0.646	0.354	0.154	0.2	3.64
22	5-3	0.565	0.435	0.235	0.2	3.9
23	5-1	0.606	0.394	0.194	0.2	3.76
24	6-5	0.418	0.582	0.332	0.25	4.53
25	6-4	0.429	0.571	0.321	0.25	4.47
26	6-1	0.658	0.342	0.0918	0.25	3.61
27 *	6-3	0.669	0.331	0.0815	0.25	3.58
		$\Sigma r_i = 13$	$\Sigma u_i = 14$	$\Sigma u_{k_i} = 8$	$\Sigma u_{k_i} = 6$	

Table 7.2: Controlability and internal reliability of the network.

errors in the observations from point #3. This is only due to the model, and has nothing to do with the actual measurements. That explains why examination of the individual standardized residuals corresponding to the inserted outliers was not able to detect them properly.

Besides the r_i's, a global reliability measure can be obtained, that is the average diagonal term of $Q_v P$ and it is called relative redundancy (eq. 3.6),

$$\frac{tr[Q_v P]}{n} = \frac{r}{n} = \frac{13}{27} = 0.48$$

From table (7.2) it can easily be seen that since there are 14 unknown parameters in the network and 8 of them are the desired ones,

$$\Sigma r_i = r = 13$$

$$\Sigma u_i = u = 14$$

$$\Sigma u_{k_i} = u_k = 8$$

$$\Sigma u_{t_i} = u_t = 6$$

From all these, only r and u_k are used for the evaluation of the reliability of the network below.

7.3 Reliability of the Network

Internal reliability (sect. 6.1), refers to the size of gross errors which can just be detected. To determine the probability with which this happens, an alternative hypothesis H_a is introduced, which simply assumes that the observation(s) ℓ_i have been affected by a certain gross error $\nabla_o \ell_i$ (sect. 5.1).

The value of $\nabla_o \ell_i$ depends on both α_o and β_o. By choosing here:

$$\alpha_o = 0.001$$

$$\beta_o = 0.80$$

and considering also that r = 13, the non-centrality parameter λ_o can be determined through nomograms (app.III):

$$\lambda_o = 17 \qquad \text{and also} \qquad \alpha = 0.06$$

With known λ_o and r_i (redundancy numbers), the $\nabla_o \ell_i$ values can be computed using equation (6.2):

$$|\nabla_o \ell_i| = \sqrt{\frac{\lambda_o}{r_i}} \cdot \sigma_{\ell_i}$$

These values are given in table (7.2), where the maximum undetectable gross error can be up to 5".5 in the direction 5-6, and the smallest down to 3".6 in the direction 6-3. The redundancy number and the maximum undetectable error for each observed direction are also shown in figure (7.1).

The global figure for the internal reliability is also often used (eq. 6.3):

$$|\overline{\nabla_o \ell_i}| = \sqrt{\frac{\lambda_o \, n}{r}} \; \sigma_{\ell_i} = \sqrt{\frac{(17)(27)}{13}} \; \sigma_{\ell_i} = 5.9 \; \sigma_{\ell_i}$$

which shows that the internal reliability is 5.9 times larger than σ_{ℓ_i}. In the present network σ_{ℓ_i} = 0".71 was assumed, therefore:

$$|\overline{\nabla_o \ell_i}| = (5.9)(0".71) = 4".19$$

In other words, the smallest detectable error by the "Data Snooping" test with a probability of 80%, is $\simeq$ 4".2. That is, with probability

β_o = .80 this model can detect outliers only if they are larger than 4".2. Keeping α_o fixed, the larger the β_o the larger the $\nabla_o \ell_i$ must be to be detected. Increasing α_o (β_o fixed), $\nabla_o \ell_i$ decreases. Therefore, the size of the just detectable error is small, but at the same time more good observations are rejected when H_o is true.

The above-mentioned undetectable error $|\nabla_o \ell_i|$ has an effect on the estimated parameters from the adjustment, x. This influence, $\nabla_o x_i$, is the external reliability, and is given by (eq. 6.4):

$$\underline{\nabla_o \hat{x}} = (A^T P A)^{-1} A^T P |\underline{\nabla_o \ell_i}|$$

In our example an error of 4".46 on the direction 3-5 affects all the directions from station #3. So the $\nabla_o \hat{x}_i$'s were calculated and their effect on x_1 and y_3 was $\simeq$ 1 mm.

However, the effect of the 1.5 mm eccentricity on the observations from point #3, was slightly larger. Hence the rejection of the direction 3-1 was inevitable and the remaining undetectable gross errors in the directions 3-5, 3-4 and 3-2 had as a consequence an effect on the estimated unknown parameters of up to 0.7 mm.

Similarly to the internal reliability (eq. 6.3) a global figure for the external reliability can be obtained (eq. 6.20); worrying only about the desired parameters, namely using only u_k,

$$\overline{\nabla_o \hat{k}_2} \leq \sqrt{\frac{\lambda_o u_k}{r}} \cdot \sigma_k$$

$$\leq \sqrt{\frac{(17)(8)}{13}} \cdot \sigma_k = 3.2\, \sigma_k$$

This says that the effect of a non-detected gross error $\nabla\ell_i$ can be as much as 3.2 times the calculated precision σ_k of the desired parameters, but of course can be much less. Hence, the smaller the u_k/r the smaller the $\nabla_o\hat{k}_i$, and the reliability of the network is better.

If a gross error $|\nabla\ell_i| < |\nabla_o\ell_i|$ exists in the i^{th} observation, it sometimes affects the solution considerably but does not significantly increase the residuals since the most part of it has been absorbed by the model. As σ_{v_i} remains the same the standardized residual v_i/σ_{v_i} is not detectable.

Consider the case of the 19^{th} observation (table 7.2), which has the minimum redundancy

19	direction 5-6	$r_{19} = 0.284$	$u_{19} = 0.716$	$\nabla_o\ell_{19} = 5''.5$

We assume that during the measurements, a gross error of 4".5 affected the direction 5-6, due to $\simeq$ +4 mm error in centering the target on point #6. The value $r_{19} = 0.284$ warns a priori that any value of $\nabla\ell_i < \nabla_o\ell_i \simeq 5''.5$ cannot be detected through the examination of the residuals. Consequently the same warning is valid for the error of 4".5. This is indeed true. Performing the adjustment the largest standardized residuals are:

direction	standardized residual	decision
6-5	$\lvert 1.47\rvert/\sqrt{.211} = 3.20$	none of them
5-6	$\lvert -1.16\rvert/\sqrt{.143} = 3.07$	exceeds the
4-5	$\lvert -1.29\rvert/\sqrt{.231} = 2.68$	critical value:
		$\sqrt{F_{0.999;1,\infty}} = 3.29$

In this case, it is also shown that the largest standard residual does not necessarily correspond to the erroneous observation. If we compare the adjusted coordinates with the "correct" ones we can see the differences $\hat{dx}_i$, $\hat{dy}_i$:

$$\hat{dx}_1 = + 1.6 \text{ mm} \qquad \hat{dy}_1 = - 1.13 \text{ mm}$$

$$\hat{dx}_2 = + 2.9 \text{ mm} \qquad \hat{dy}_2 = + 0.7 \text{ mm}$$

$$\hat{dx}_3 = + 1.2 \text{ mm} \qquad \hat{dy}_3 = - 0.9 \text{ mm}$$

$$\hat{dx}_4 = + 2.2 \text{ mm} \qquad \hat{dy}_4 = + 0.6 \text{ mm}$$

which are significantly large compared to the σ_{k_i} from C_k. These values, $\hat{dx}_i$, $\hat{dy}_i$, express clearly the external reliability of the network since they owe their existence to the undetectable error of 4"5 in the direction 5-6.

If one looks at the global expression of the external reliability for the $\hat{x}_2$ coordinate (eq. 6.20) (using the $\sigma^2_{x_2} = .731 \text{ mm}^2$ from the adjustment),

$$\nabla\hat{x}_2 \leq 3.2\, \sigma_{x_2} = 3.2\, \sqrt{.731} = 2.7 \text{ mm}$$

which is approximately equal to $\hat{dx}_2$.

The probable large differences between ∇k_i's and σ_{k_i}'s prove that C_k cannot describe only by itself the accuracy of the estimated parameters.

Concluding, it can be said that the relationships (6.3), (6.20) are very important for the evaluation of the reliability of networks. In the present network, even after the gross-detection procedure, outliers

as large as 5.9 σ_{ℓ_i} may remain in the data. These can affect the final results $(\hat{k})$ by as much as 3.2 σ_k. Therefore, the reliability of the network cannot be described by the covariance matrix C_k of the unknown parameters which is a measure of the precision and relates mainly to random errors. The latter happens because C_k is based only on the geometry of the network (design matrix A), and on the assumed precision of the observations C_ℓ, presuming no bias and absence of any systematic errors. This model will never show $\nabla\ell_i < \nabla_o\ell_i$ errors hidden in the data, because of its finite capability of controlling the observations.

8

NEW METHODS

The pre- and post- adjustment techniques as they are presented here, are only some of the several possibilities for detecting and eliminating gross errors in the data. They have advantages and disadvantages, but they have been used almost exclusively, whenever such techniques were needed, mainly because they can be easily attached to a Least Squares software package and also because their efficiency has been repeatedly tested.

Research, however, should not be confined only to a further refinement of these techniques. One should try different philosophies and have several alternatives. Some of these studies are presented here.

8.1 A Premium-Protection Method

The concept of premium-protection has often appeared in the statistical references, but it was introduced in the geodetic field only recently *(Clerici and Harris, 1980)*.

-- Premium is assessed by the percentage increase in the value of the trace of the covariance matrix of the estimated parameters due to a rejection that should not have taken place.

$$\text{Premium} = (\text{Tr}[C_{x_A}] - \text{Tr}[C_x])/\text{Tr}[C_x] \qquad (8.1)$$

where:

C_x : covariance matrix without rejections

C_{x_A} : covariance matrix after a rejection (in this case wrong rejection).

-- Protection is assessed by the percentage decrease in the value of the trace of the covariance matrix, due to a correct rejection of an erroneous observation.

$$\text{Protection} = (\text{Tr}[\bar{C}_x] - \text{Tr}[C_{x_A}])/\text{Tr}[\bar{C}_x] \qquad (8.2)$$

where:

$\bar{C}_x$ = covariance matrix being affected by a gross error in an observation. When there is not any gross error: $\bar{C}_x = C_x$.

The method is similar to the B-method of testing (sect. 5.2), and the correspondence between premium-protection and the type I and type II errors is apparent. Likewise, for a chosen premium and protection, critical values c for the standardized residuals can be determined and moreover values similar to the values for the internal reliability

(sect. 6.1), can be also estimated. Examples using both methods have shown a very good agreement *(Clerici and Harris, 1980)*.

The main advantage of the method is that instead of using the abstract concepts of the significance level and the power of the test, more palpable geodetic criteria are used by expressing the desired gain (protection), and the tolerable loss (premium), in the trace of the covariance matrix of the unknown parameters.

8.2 Use of Strain Analysis

Strain has been used to analyze crustal movements for many years. In that respect, strain is caused by a real movement. This real movement is detected examining differences between geodetic observations of two epochs.

However, differences between geodetic observations are caused not only by real displacements but also by large errors or more generally by inconsistent observations. Discrepancies and inconsistencies among various kinds of observations and also constraints, act like forces in a network, extending or contracting its different parts, therefore the corresponding strain can be computed.

In the research being done *(Vanicek et al., 1981)* the effect of different inconsistencies and constraints has been studied and the results are represented by strain ellipses at each point. The size of these ellipses reveals the magnitude of the existing strain in the network. This use of strain proves to be useful in the analysis of the geometrical strength of networks and research is still being continued along this direction.

8.3 Adjustment by Minimizing the Sum of the Absolute Residuals

As it is known, the method of Least Squares minimizes the 2-norm, which is the sum of the squares of the residuals.

$$||v||_2 = \sum_{i=1}^{N} v_i^2 = \text{Min}$$

It has been perhaps the most extensively used approximation technique because of its mathematical and computational simplicity.

However, when the errors in the system are not normally distributed, the parameters estimated by the minimization are not the most likely. The Least Squares method is very sensitive to the presence of large errors, distorting the solution significantly. Moreover it is quite efficient in spreading the effect of one gross error over all adjusted observations and residuals making the detection and the localization of the error quite difficult, through examination of the residuals.

The use of standardized residuals as tested quantities solves somewhat the problem, since it is more robust than examining the residuals themselves. But generally the Least Squares method is non robust. To overcome this problem other attempts have been made and other methods have been proposed concerning more robust estimations.

One alternative is the minimization of the 1-norm, which is the sum of the absolute values of the residuals.

$$||v||_1 = \sum_{i=1}^{N} |v_i| = \text{Min}$$

This method is relatively insensitive to the presence of gross errors among the measurements since it can tolerate larger residuals. In this

way the errors are confined to a few residuals only (not spread as before), and they can be easily detected and eliminated. Moreover they do not affect the solution so much; e.g., in a simple case of direct observations the adjusted value leads to the median *(Fuchs, 1981)* which definitely is much more robust than the mean.

Problems encountered with numerical computations have been solved and the technique is claimed to be very efficient *(Meissl, 1980; Fuchs, 1981)*.

In another study based on a similar philosophy, the advantages of minimizing the length of the vector of the residuals, in constructing more reliable statistical tests, are stressed and examples are also given *(Caspary, W. and Chen, Y.Q., 1981)*.

8.4 Iteratively Reweighted Least Squares

If the weights are also considered, the Least Squares algorithm minimizes the sum of the weighted squares of the residuals:

$$I = \sum_{i=1}^{n} p_i \cdot v_i^2 = \text{Min} \qquad (8.3)$$

where:

p_i : the selected weight factors.

Under the philosophy, that the larger the residual the more probable the corresponding observation to be erroneous, one, instead of eliminating the observation itself, can diminish its weight *(Schlossmacher, 1973)*.

If in an iterative procedure, the reciprocal of the residuals of the k^{th} iteration, are used as weights in the (k + 1) st iteration:

$$p(k+1)_i = \frac{1}{|v(k)_i|} \tag{8.4}$$

equation (8.3) becomes:

$$I(k+1) = \sum_{i=1}^{n} \frac{1}{|v(k)_i|} v^2(k+1)_i \tag{8.5}$$

If $|v(k)_i - v(k+1)_i| \simeq 0$ for i = 1, 2, ... , n equation (8.5) results into:

$$I(k+1) \tilde{=} \sum_{i=1}^{n} |v(k+1)_i| \tag{8.6}$$

which approximates the 1-norm, i.e., the sum of the absolute values of the residuals (mentioned in 8.3). The procedure starts with $p_i = 1$. If any of the residuals have values 0 the weights (eq. 8.4) become very large causing numerical instabilities which moderate the convergence rate. Therefore, since a very small residual does not contribute significantly to the total sum (eq. 8.3), it is acceptable to "eliminate" this observation by setting $p_i = 0$. If in the following iterations the residuals become significantly large again the weights can change accordingly *(Schlossmacher, 1973)*.

In this way the outlying observations have very small weight and thus, they do not affect the solution considerably. The main advantage of the method is that it can be used with any standard Least Squares computer program. Moreover in a non-linear case, both iterations can take place simultaneously.

8.5 Other Robust Estimators

Statistical literature abounds in estimators which are relatively insensitive to limited deviations of the distribution function

of the measurements. These deviations imply presence of both gross and systematic errors.

However, in the Geodetic and Photogrammetric field only isolated studies for the adaptation of such techniques have been made, as far as the knowledge of the author is concerned. One of these studies concerns the "Danish Method" *(Krarup, 1980)*, where an iteratively reweighted Least Squares adjustment is proposed. The weights are based on the weight function:

$$p = \begin{cases} 1 & \text{for } |v| \leq 2 \cdot \sigma \\ \text{proportional to } \exp(-C.v^2) & \text{for } |v| > 2 \cdot \sigma \end{cases}$$

where:

C : a constant

σ : standard deviation of the observations.

More specific values are given for the application of this weight function in photogrammetric problems. After the procedure converges, the erroneous observations have weights of zero and their residuals indicate the magnitude of the errors.

9

CONCLUSIONS AND RECOMMENDATIONS

Post-adjustment gross error detection techniques, as they were presented above (ch. 5), have proven to be very sensitive and efficient in detecting errors of small magnitude. Their application in the analysis of networks and also in many photogrammetric problems, can considerably improve the quality of the results, but the main advantage appears when their reliability is considered.

It has been a matter of principle in every geodetic problem that not only the solution but also a measure of its quality are provided. Since absolute measures of the deviation from the physical value cannot be obtained, one can be content when the maximum extension from the desired value can be bounded at a certain probability level.

The quality of the results depends first on the quality of the observations, and secondly on how these observations are utilized by the

model. By the latter, two things are implied. One is how the precision of the observations affects the results through the geometry of the model (error propogation), and the other is how sensitive the model itself is to small changes in the observations. While the first matter refers either to the a priori assumed precision of the data, or to their fit into the model (precision of the results), the second matter refers to the robustness of the model, namely, how the model reacts to small inconsistencies of the observations (reliability of the results).

If a priori estimates for the precision of the observations are used, then precision relates entirely to random errors. Reliability, however, as it was analyzed here (ch. 6), accounts for the true errors (sect. 2.2), that is for both random and gross or systematic errors.

This more distinct definition of the concept of quality of an estimation, was obtained after Baarda's theory (1976, 1979) was introduced (fig. 9.1).

-- Accuracy is a measure of quality.
-- Quality consists of precision and reliability.
-- Measure of precision is the covariance matrix of the unknown parameters.
-- Measures of reliability are the maximum inconsistency in the observations that can remain undetectable (internal reliability), and its effect on the desired unknown parameters (external reliability).

In an actual situation where the data acquisition has taken place, usually a gross error detection and elimination technique is employed, and then reliability statements can be made. This, however, is not always the case. Since reliability analysis does not require actual data, it can precede the gross error detection or even the data acquisition; i.e., it can be utilized during the design of the network.

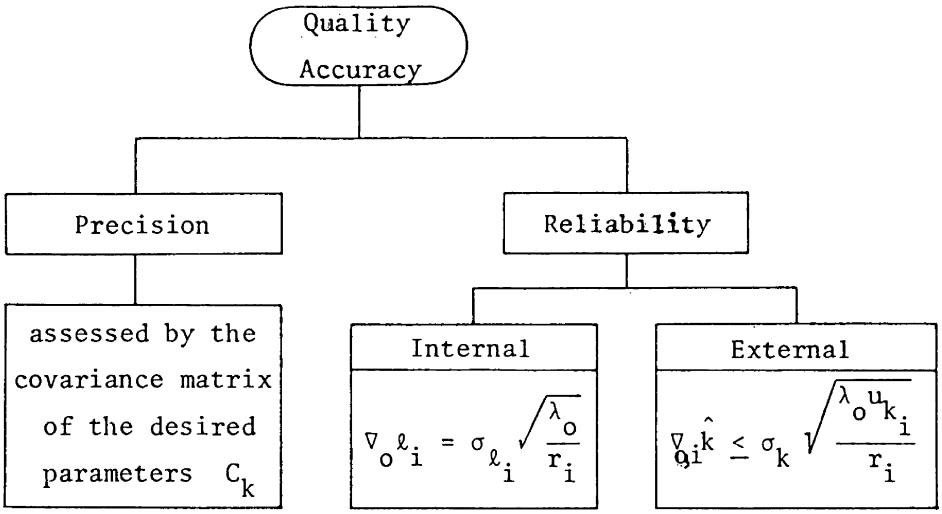


Figure 9.1: Quality of networks *(Förstner, 1981.b)*

The surveyor who wishes to perform a detailed and meaningful analysis of a network should know first what kinds of gross errors might have affected the results. Also without a good knowledge of the expected accuracy of the observations, no serious investigation can be initiated. Knowing how observations are approximately affected by gross errors, it helps to select the right alternative hypotheses and therefore to make the right decision.

After the surveyor detects most of the gross errors he should keep in mind that small outliers may still remain in the data and affect the solution. One cannot detect these errors using standard techniques (e.g., Data Snooping), but measures of the sensitivity of the detection are possible. For this analysis the surveyor has to consider the cost of making a wrong decision, of either type of error. He should not rely only on the covariance matrix of the unknown parameters but also on the

above measures of reliability. These measures are of major importance especially when the redundancy is quite low (e.g., small survey nets). In the latter case reliability measures give a warning which reveals how defenceless the model is in front of lurking outliers.

The surveyor should also consider this analysis during the design of the network, so that the redundancy numbers (sect. 3.2) are not only large enough but also homogeneous throughout the entire net, increasing its strength. When high accuracy is required $r_i < 0.35$ are not recommended for design purposes. For the external reliability values of $\sqrt{\bar{\lambda}_o}$ (sect.6.2) smaller than 10 are considered as satisfactory *(Kok, 1982)*.

Despite the efficiency of all the presented strategies (ch. 5), research has not yet reached the ultimate goal for reasons mentioned above (sect. 5.8). There are still problems encountered with the localization of errors and further development is needed. The present trend is to connect the statistical nature of the significance level with more geodetic quantities, so that subjectiveness and personal criteria will be quite limited *(Stefanovic, 1980; Clerici et al. 1980)*.

More information is also required about the number, magnitude, frequency, clustering and spread of errors *(Förstner, 1981.a)*.

Further development is also required for the tau-test *(Pope, 1976)*, where up to now only the type I error is considered.

There is also an objection concerning the usefulness of the concept of the reliability from the practical point of view *(Stefanovic, 1980)*, to be clarified.

On the other hand, research should not be restricted in the conventional techniques, but other possibilities should be examined,

some of which are definitely more robust than the Least Squares Method (ch. 8). This is worth trying since a very detailed analysis using standard techniques may also lead to uneconomical computations.

The problem of economy itself differs from application to application, depending on the importance and specifications of the project. This has to be considered before the procedure to be followed has been designed. If, for example, a photogrammetric problem is concerned, recollection of some data (e.g., remeasurement of a model or of a photo), is quite justifiable and not very costly. Moreover no detailed and complicated strategies are required for the localization of gross errors as long as the model or the picture, where the error occured, have been found. However, in a surveying project, remeasuring may be either impossible or more costly than performing a scrutiny on the data, if of course a systematic procedure is available. This analysis is indispensable since every single observation is valuable and contributes significantly to the strength of the network. The gain is much higher than the computational cost, which definitely in small surveying nets is not the dominant factor, compared to the other expenses, and is going to be further reduced.

To increase the effectiveness of the procedure, very often a decomposition of the large adjustment takes place, or a combination of pre- and post- adjustment techniques can be implemented *(Kok, 1982)*. Regardless of which techniques are employed to detect, localize and eliminate gross or systematic errors, the final tests of highest sensitivity on the entire model will reveal any inconsistency that may

be still present. If all the tests are passed, the system is considered to be free from significant gross errors and the sensitivity of the detection (reliability) is also estimated. After that, the results can be given to the user.

Briefly, in a serious survey project one should follow the following:

-- Design of Network

Purpose and specifications - economical aspects.

Configuration.

Preanalysis - precision and reliability criteria.

Design of the observations.

-- Observations

Simple checks during the measurements to ensure at least against large gross errors.

Special care to indispensable observations.

-- Analysis of Network

Pre-adjustment techniques.

Other robust estimators (if available).

Post-adjustment techniques.

Detection and elimination of gross errors.

Final adjustment.

Quality (precision and reliability) of the network.

-- Results and Their Quality to the User

In closing, a warning is necessary concerning the application of the proposed statistical techniques.

It should be realized that to obtain the right answers one has to ask the right questions. Statistical tests are useful as long as one knows what is needed and how to use them properly. The investigator should have a good understanding of the survey procedure and the philosophy of the error detection technique. The outcome of the tests should be interpreted with care, and decisions have to be governed by the logic and the nature of the particular problem. Blind acceptance of the test-results is never recommended. Statistical tests are not "panacea" but only a very useful tool; therefore they do not have the ability to protect us against nescience or wrong reasoning. A blunder in the data is always less destructive than a "blunder" in our way of thinking.

REFERENCES

ACKERMANN, F. (1981) Reliability and Gross Error Detection in Photogrammetric Blocks, presented paper at IUGG, IAG Symposium on Geodetic Networks and Computations, Munich.

ALBERDA, J.E. (1980) A Review of Analysis Techniques for Engineering Survey Control Schemes, presented paper at the Industrial and Engineering Survey Conference, London.

AMER, F.A.A.F. (1981) Theoretical reliability of elementary photogrammetric procedures, ITC Journal 1981-3, pp. 278-307.

BAARDA, W. (1965) Precision, Accuracy and Reliability of Observations, Report IAG Symposium, Lisbon 1960, Reprint, Delft.

BAARDA, W. (1967) Statistical Concepts in Geodesy, New Series, Vol. 2, No. 4, Delft.

BAARDA, W. (1968) A Testing Procedure for Use in Geodetic Networks, New Series, Vol. 2, No. 5, Delft.

BAARDA, W. (1976) Reliability and Precision of Networks, presented paper, VII International Course for Engineering Surveys of High Precision, Darmstadt, pp. 17-27.

BAARDA, W. (1979) Measures for the Accuracy of Geodetic Networks, Optimization of Design and Computation of Control Networks, (Lectures from the Symposium held in Sopron in 1977), Budapest, pp. 419-436.

BLAIS, J.A.R. (1976) Error Analysis of Least Squares Adjustments, collected papers, Geodetic Survey of Canada, Surveys and Mapping Branch, Ottawa.

BOMFORD, G. (1977) Geodesy, 3rd edition, Oxford University Press, Oxford, pp. 164-167.

CASPARY, W.; CHEN, Y.Q. (1981) A Special Similarity Transformation to Investigate the Congruency of Two Sets of Points, (to be published), Dept. of Surveying Eng., Univ. of New Brunswick, Fredericton, N.B., Canada.

CHRZANOWSKI, A. (1981) (with contributions by the members of the FIG "AD HOC" committee) A Comparison of Different Approaches into the Analysis of Deformation Measurements, FIG XVI International Congress, Montreux, Switzerland.

CLERICI, E.; HARRIS, M.W. (1980) A Premium-Protection Method Applied to Detection and Rejection of Erroneous Observations, manuscripta geodetica, vol. 5, pp. 283-298.

EL-HAKIM, S.F. (1979) Potentials and Limitations of Photogrammetry for Precision Surveying, Surveying Eng. Dept., Univ. of New Brunswick Tech. Report No. 63, Fredericton, N.B., Canada.

FÖRSTNER, W. (1979) On the Internal and External Reliability of Photogrammetric Coordinates, presented paper at ASP-ASCM Convention, Washington, D.C.

FÖRSTNER, W. (1980) Evaluation of Block Adjustment Results, presented paper, International Archives of Photogrammetry, Vol. 23, Part B3, Commission III, pp. 236-245.

FÖRSTNER, W. (1981.a) Review of Problems, ISPRS, W.G. III/1, circular letter No. 1, Appendix 1.

FÖRSTNER, W. (1981.b) Theorie der äusseren Zuverlässigkeit, Institute für Photogrammetrie Stuttgart, Heft 7, pp. 59-70.

FÖRSTNER, W. (1981.c) On the Reliability of 3rd and 4th Order Densifications, presented paper at IUGG, IAG Symposium on Geodetic Networks and Computations, Munich.

FUCHS, H. (1981) Adjustment by Minimizing the Sum of Absolute Residuals, presented paper at IUGG, IAG Symposium on Geodetic Networks and Computations, Munich.

GRÜN, A. (1979) Gross Error Detection in Bundle Adjustment, presented paper at Aerial Triangulation Symposium, Dept. of Surveying, Univ. of Queensland, Australia.

GRÜN, A. (1980) Internal Reliability Models for Aerial Bundle Systems, presented paper, International Archives of Photogrammetry, Vol. 23, Part B3, Commission III, pp. 272-283.

HECK, B. (1981) Der Einfluss einzelner Beobachtungen auf das Ergebnis einer Ausgleichung und die Suche nach Ausreissern in den Beobachtungen, Allgemeine Vermessungs-Nachrichten, Heft 1, Januar 1981.

HOGG, R.V.; CRAIG, A. (1978) Introduction to Mathematical Statistics, 4th Edition, Macmillan Publishing Co., Inc., New York, pp. 235-241.

KOK, J.J. (1982) "Personal Communication".

KRARUP, T. (1980) Götterdämmerung over Least Squares Adjustment, presented paper, International Archives of Photogrammetry, Vol. 23, Part B3, Commission III, pp. 369-378.

MEISSL, P. (1980) Adjustment of levelling networks by minimizing the sum of absolute rediduals, Proceedings of NAD Symposium, Ottawa, pp. 393-415.

MIERLO, J. van (1977) Systematic Investigations on the Stability of Control Points, presented paper at XV International Congress of Surveyors, Stockholm, Sweden.

MIERLO, J. van (1981) Second Order Design: Precision and Reliability Aspects, Allgemeine Vermessungs-Nachrichten, 88, pp. 95-101.

MIKHAIL, E.M. (1979) Review and some Thoughts on the Assessment of Aerial Triangulation Results, presented paper at Aerial Triangulation Symposium, Dept. of Surveying, Univ. of Queensland, Australia.

MONIWA, H. (1977) Analytical Photogrammetric System with Self-Calibration and Its Applications, Ph.D. Thesis, Dept. of Surveying Engineering, University of New Brunswick, Fredericton, N.B. Canada.

MORRISON, D.F. (1976) Multivariate Statistical Methods, 2nd Edition, McGraw-Hill Series, pp. 32-33, 62, 67.

POPE, A.J. (1976) The Statistics of Residuals and the Detection of Outliers, U.S. Dept. of Commerce, NOAA Technical Report NOS. 65 NGS 1, Rockville, Md.

SCHLOSSMACHER, E.J. (1973) An Iterative Technique for Absolute Deviations Curve Fitting, Journal of the Amererican Statistical Association, Vol. 68, pp. 857-859.

STEEVES, P.A. (in prep.) Numerical Processing of Horizontal Control Data - Economization by Automation, Ph.D. Thesis, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B.

STEFANOVIC, P. (1978) Blunders and Least Squares, ITC Journal 1978 - 1, pp. 122-157.

STEFANOVIC, P. (1980) Pitfalls in Blunder Detection Techniques, presented paper, International Archives of Photogrammetry, Vol. 23, Part B3, Commission III, pp. 687-700.

UOTILA, A.U. (1973) Useful Statistics for Land Surveyors, Surveying and Mapping, Vol. 33, No. 1, March.

UOTILA, A.U. (1976) Statistical Tests as Guidelines in Analysis of Adjustment of Control Nets, presented paper, Fédération Internationale des Géomètres, 14th Congress, Washington, D.C.

VANICEK, P.; KRAKIWSKY, E.J. (1982) Geodesy: The Concepts, 13th Chapter, North Holland, Amsterdam.

VANICEK, P.; THAPA, K.; SCHNEIDER, D. (1981) The Use of Strain to Identify Incompatible Observations and Constraints in Horizontal Geodetic Networks, manuscripta geodaetica, Vol. 6, No. 3, pp. 257-281.

WELLS, D.E.; KRAKIWSKY, E.J. (1971) The Method of Least Squares, Lecture Notes No. 18, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B., Canada.

APPENDIX I: LEAST SQUARES AND TRUE ERRORS

Considering the case of the second standard problem $A\underline{x} + \underline{\ell} = \underline{0}$, (parametric L.S.A.), the observation equations are of the form,

$$A\hat{\underline{x}} + \underline{\ell} = \underline{v} \qquad \text{with} \qquad P = \sigma_o^2 \cdot C_\ell^{-1}$$

The solution follows from,

$$\hat{\underline{x}} = -(A^T P A)^{-1} A^T P \underline{\ell} \tag{I.1}$$

and if C_x is the covariance matrix of the unknown parameters,

$$C_x = \sigma^2 (A^T P A)^{-1} = (A^T C_\ell^{-1} A)^{-1} \tag{I.2}$$

The residual vector is given by

$$\begin{aligned} \underline{v} &= A\hat{\underline{x}} + \underline{\ell} \\ &= (I - A(A^T P A)^{-1} A^T P)\ \underline{\ell} = M\ \underline{\ell} \end{aligned} \tag{I.3}$$

where

$$M = I - A(A^T P A)^{-1} A^T P \tag{I.4}$$

and if C_v is the covariance matrix of the residuals

$$C_v = \sigma_o^2\ (P^{-1} - A\ (A^T P A)^{-1} A^T) = \sigma_o^2\ M\ P^{-1} \tag{I.5}$$

Not very often used, the covariance matrix of the adjusted parameters can be derived by,

$$C_{\hat{\ell}} = \sigma_o^2\ (A\ (A^T P A)^{-1} A^T) \tag{I.6}$$

From (I.5) and (I.6) it comes that

$$C_v = C_\ell - C_{\hat{\ell}} \tag{I.7}$$

and from (I.4)

$$M = \frac{1}{\sigma_o^2} C_v P = Q_v P \qquad (I.8)$$

Expressing now the true observational errors as $\underline{\varepsilon} = \underline{\ell} - E(\underline{\ell})$, and considering that $E(\underline{\ell}) + A\underline{x}_{TRUE} = \underline{0}$, it follows that:

$$\underline{\varepsilon} = \underline{\ell} - A\underline{x}_{TRUE} \qquad (I.9)$$

It is easy now to prove that

$$\underline{v} = M\,\underline{\varepsilon} \qquad (I.10)$$

Indeed,

$$\underline{v} = \underline{\ell} + A\hat{\underline{x}}$$

$$= \underline{\ell} + A\hat{\underline{x}} - A\underline{x}_{TRUE} + A\underline{x}_{TRUE}$$

$$= \underline{v} - A\underline{x}_{TRUE} + A(A^T PA)^{-1}A^T PA\underline{x}_{TRUE}$$

$$= \underline{v} - MA\underline{x}_{TRUE}$$

$$= M\underline{\ell} - MA\underline{x}_{TRUE}$$

$$= M(\underline{\ell} - A\underline{x}_{TRUE}) \text{ which due to (I.9) results in (I.10)}$$

$$\underline{v} = M\,\underline{\varepsilon}$$

At this stage a transformation can be performed $\overline{A} = SA$, $\overline{\underline{\ell}} = S\underline{\ell}$,

$$\overline{\underline{v}} = S\underline{v},\ \overline{\underline{\varepsilon}} = S\underline{\varepsilon} \qquad (I.11)$$

where

$$P = S^T S$$

since P is always positive definite.

This transformation has as purpose to simplify the expressions and helps in the case where a geometrical interpretation of L.S. adjustment and statistical testing is desired *(Pope, 1976; van Mierlo, 1981)*. Through

this transformation the correlated stochastic elements ε are transformed to uncorrelated $\bar{\varepsilon}$. Thus their covariance matrix becomes,

$$C_{\bar{\varepsilon}} = SC_{\varepsilon}S^T \qquad \text{and since} \qquad C_{\varepsilon} = C_{\ell} = \sigma_o^2 P$$

$$C_{\bar{\varepsilon}} = \sigma_o^2 SP^{-1}S^T = \sigma_o^2 SS^{-1}(S^T)^{-1}S^T = \sigma_o^2 I \qquad \text{(I.12)}$$

Also from (I.11)

$$\underline{\bar{v}} = S\underline{v} = S[I - A(A^TPA)^{-1}A^TP]\ \underline{\varepsilon}$$

$$= S\underline{\varepsilon} - SA(A^TS^TSA)^{-1}A^TS^TS\underline{\varepsilon}$$

$$\underline{\bar{v}} = (I - \bar{A}(\bar{A}^T\bar{A})^{-1}\bar{A}^T)\ \underline{\bar{\varepsilon}}$$

$$= \bar{M}\ \underline{\bar{\varepsilon}} \qquad \text{(I.13)}$$

where

$$\bar{M} = I - \bar{A}(\bar{A}^T\bar{A})^{-1}\bar{A}^T \qquad \text{(I.14)}$$

Similarly

$$C_{\bar{v}} = \sigma_o^2 \bar{M} \qquad \text{(I.15)}$$

on the contrary to (I.12)

$$C_{\bar{\varepsilon}} = \sigma_o^2 I \qquad \text{(I.16)}$$

Some properties of the idempotent matrix $M = Q_v P$

The matrix $M = I - A(A^TPA)^{-1}A^TP = Q_vP$ (eq. I.4) is an idempotent matrix since $MM = M$. It has the following important characteristics *(Pope, 1976; Stefanovic, 1978)* which also hold for the Q_v matrix if $P = I$.

1. The M matrix is square and symmetric.
2. rank [M] = trace [M] (I.17)

Since rank [M] = n - u = r , i.e., trace [M] = (total redundancy r)

3. Since $M \neq I$ or rank [M] < order [M] = n, M is singular.

4. If r_i are the diagonal elements of M, $0 \leq r_i \leq 1$ (I.18)

5. (average diagonal term): $\bar{r} = \frac{\text{trace [M]}}{n} = \frac{r}{n}$ (I.19)

6. $r_i = r_{ii} = \sum_{i=1}^{n} r_{ij}$ (I.20)

7. $0 \leq r_{ij} < r_i\, r_j$ (I.21)

8. The second part of M, i.e., $U = A(A^T PA)^{-1}A^T P$ is also an idempotent matrix.

9. In the case of P = I, $r_i = q_{v_i}$ and it has been shown *(Amer, 1981)* that $0 \leq q_{v_{ij}} \leq 0.5$ (I.22)

10. The vector of eigenvalues of M consists of u zeros and r ones *(Morrison, 1976)*.

APPENDIX II: ELIMINATION OF NUISANCE PARAMETERS

If there are nuisance parameters in the model denoted by t, whereas the desired parameters are denoted by k, *(Förstner, 1980)*, the Least Squares model is of the form:

$$-\underline{\ell} + \underline{v} = A\,\underline{\hat{x}} = A_k\,\underline{\hat{k}} + A_t\,\underline{\hat{t}} \qquad \text{(II.1)}$$

where

$$A = [A_k \,\vdots\, A_t] \qquad \text{and} \qquad \underline{\hat{x}}^T = [\underline{\hat{k}}^T \,\vdots\, \underline{\hat{t}}^T]$$

The solution for the desired parameters is given by *(Wells and Krakiwsky, 1971)*,

$$\hat{k} = -Q_k\,\bar{A}_k^T\,P\,\underline{\ell} \qquad \text{(II.2)}$$

where

$$\bar{A}_k = (I - A_t(A_t^T P\, A_t)^{-1} A_t\, P).\, A_k \qquad \text{(II.3)}$$

and

$$Q_k = [A_k^T P\, A_k - (A_k^T P\, A_t)(A_t^T P\, A_t)^{-1}(A_t^T P\, A_k)]^{-1}$$

From the definition of the U matrix (eq. 3.8),

$$U \overset{d}{=} Q_{\hat{\ell}}\, P = A\, Q_x\, A^T P = I - M \qquad \text{(II.4)}$$

If there are nuisance parameters, after the partitioning of A (eq. II.1) the U matrix becomes *(Mikhail, 1979)*:

$$U = [A_k \,\vdots\, A_t] \cdot \left[\begin{array}{c:c} Q_k & Q_{kt} \\ \hdashline Q_{kt} & Q_t \end{array}\right] \cdot \left[\begin{array}{c} A_k \\ \hdashline A_t \end{array}\right] \cdot P$$

which after a few manipulations becomes *(see Förstner, 1981.b)*,

$$U = \bar{A}_k \, (\bar{A}_k^T \, P \, \bar{A}_k)^{-1} \, \bar{A}_k^T \, P + A_t \, (A_t^T \, P \, A_t)^{-1} \, A_t^T \, P \qquad \text{(II.5)}$$

$$= U_k + U_t$$

Therefore if u_{k_i} and u_{t_i} are the diagonal elements of the U_k and U_t matrices correspondingly,

$$u_{k_i} = (\bar{A}_k \, (\bar{A}_k \, P \, \bar{A}_k)^{-1} \, \bar{A}_k^T \, P)_{ii} \qquad \text{(II.6)}$$

and

$$u_{t_i} = (A_t \, (A_t^T \, P \, A_t)^{-1} \, A_t^T \, P)_{ii} \qquad \text{(II.7)}$$

Moreover from equation (II.5),

$$u_i = u_{k_i} + u_{t_i} \qquad \text{(II.8)}$$

where u_i's are the diagonal elements of the U matrix.

APPENDIX III: NOMOGRAMS FOR THE B-METHOD OF TESTING

The required values for the B-method of testing are given through a set of nomograms *(Baarda, 1968; pp. 21-23)*. An extract of the most frequently used case of $\beta_0 = 0.80$, is given below so that the reader can follow the examples.

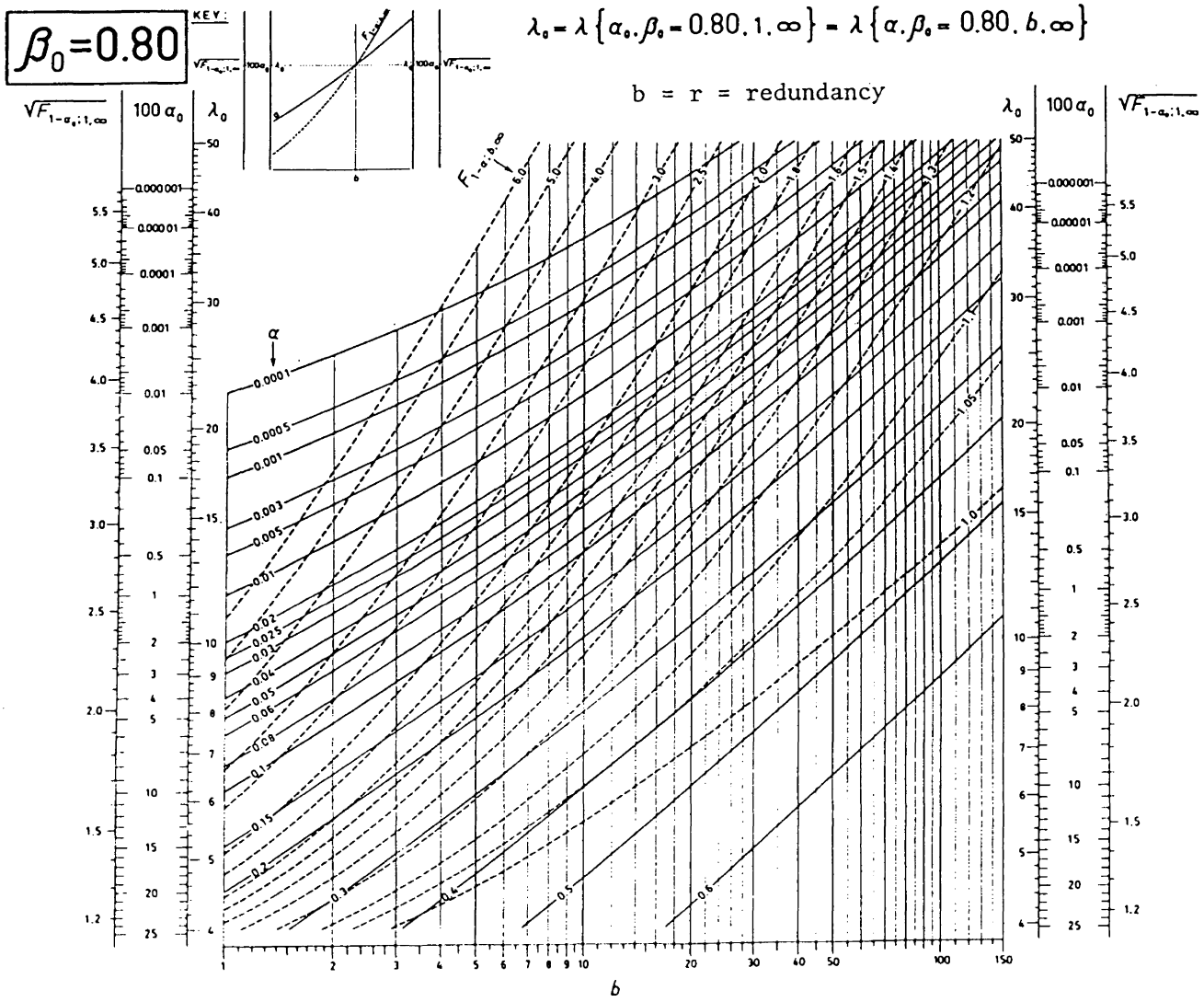


Figure III.1: Nomograms for the B-Method of Testing

Following the normal procedure (sect. 5.2), by choosing

$$\alpha_o = 0.001$$

$$\beta_o = 0.80$$

for e.g., 10 degrees of freedom (r = 10), the rest of the values can be easily determined,

$$\lambda_o = 17$$

$$\alpha = 0.04$$

$$\sqrt{F_{1-\alpha_o;1,\infty}} = 3.29$$

$$\sqrt{F_{1-\alpha;r,\infty}} = 1.90$$

A few more example values are given in section 5.2.