

# **DIAGRAMATIC APPROACH TO SOLVE LEAST-SQUARES ADJUSTMENT AND COLLOCATION PROBLEMS**

**M. MOHAMMAD-KARIM**

**September 1981**



**TECHNICAL REPORT  
NO. 83**

## 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.

# **DIAGRAMATIC APPROACH TO SOLVE LEAST-SQUARES ADJUSTMENT AND COLLOCATION PROBLEMS**

M. Mohammad-Karim

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

September 1981

Latest Reprinting October 1995

## ABSTRACT

The standard method of obtaining the solution to the least-squares adjustment problems lacks objectivity and can not readily be interpreted geometrically. Thus far, one approach has been made to interpret the adjustment solution geometrically, using Hilbert space technique.

This thesis is still another effort in this respect; it illustrates the important advantage of considering the parallelism between the concept of a metric tensor and a covariance matrix. By splitting the linear mathematical models in the least-squares adjustment (parametric, conditional and combined models) into dual covariant and contravariant spaces, this method produces diagrams for the least-squares adjustment and collocation, from which the necessary equations may be obtained.

Theories and equations from functional analysis and tensor calculus, necessary for the analysis of this approach, are described. Also the standard method used in the least-squares adjustment and the least-squares collocation are reviewed. In addition, the geometrical interpretation of a singular weight matrix is presented.

TABLE OF CONTENTS

	Page
ABSTRACT .....	ii
ACKNOWLEDGEMENT .....	vii
CHAPTER 1. INTRODUCTION .....	1
1.1 Concept of the Least-Squares Adjustment .....	1
1.2 Outline of Treatment .....	6
CHAPTER 2. SELECTED DEFINITIONS AND CONCEPTS FROM FUNCTIONAL ANALYSIS .....	8
2.1 Metric Spaces .....	8
2.1.1 Definition of metric spaces .....	8
2.1.2 Complete metric spaces .....	9
2.1.3 Closure of a subset, dense set, separable set .....	9
2.1.4 Vector spaces, linear spaces .....	10
2.1.5 Linear dependence, independence .....	10
2.1.6 Vector spaces of finite and infinite dimensions .....	11
2.1.7 Normed spaces, Banach spaces .....	12
2.1.8 Completeness theorem .....	12
2.1.9 Closedness theorem .....	12
2.1.10 Inner product spaces, Hilbert spaces .....	13
2.2 Linear Operators .....	15
2.2.1 Definition of linear operators .....	15
2.2.2 Bounded linear operator .....	16
2.2.3 Finite dimension theorem .....	17
2.3 Linear Functional .....	17
2.3.1 Definition of linear functionals .....	17
2.3.2 Bounded linear functional .....	17
2.4 Dual Spaces .....	17
2.4.1 Definition of dual spaces .....	17
2.4.2 Dual bases .....	18
2.4.3 Annihilator .....	20
2.5 Projection Theorem .....	20
2.5.1 Orthogonality .....	20
2.5.2 Explanation of the projection theorem.....	20

TABLE OF CONTENTS (Cont'd)

	Page
2.5.3 Orthogonal complement .....	21
2.5.4 Dual spaces and orthogonality .....	21
2.5.5 Minimum norm problem and dual space .....	22
2.5.6 Dual spaces and least squares .....	24
 CHAPTER 3. FORMULATION AND SOLUTION OF THE LEAST- SQUARES PROBLEM - A REVIEW .....	 28
3.1 Formulation of Least-Squares Problem .....	28
3.2 Metricizing of the Spaces Involved in the Prob- lem .....	29
3.3 The Solution to the Least-Squares Problem .....	32
 CHAPTER 4. TENSOR STRUCTURE AND THE LEAST SQUARES .....	 38
4.1 Introduction .....	38
4.2 Covariant and Contravariant Components of a Vector .....	38
4.3 Relation Between Covariant and Contravariant Components of a Vector .....	42
4.4 Transformation of Vectors .....	44
 CHAPTER 5. DIAGRAMS FOR THE LEAST-SQUARES ADJUSTMENT PROBLEMS .....	 46
5.1 5.1 Diagram for a Simple Parametric Case .....	46
5.2 Diagram for a Conditional Adjustment Model .....	52
5.3 Diagram for a Combined Adjustment Model .....	55
5.4 Some Remarks to the above Diagrams .....	59
5.4.1 Singularity and the diagrams .....	60
 CHAPTER 6. DIAGRAM FOR LEAST-SQUARES COLLOCATION .....	 65
6.1 Two-Component Adjustment and Collocation .....	65
6.2 Two Component Adjustment Mathematical Model .....	68
6.3 Formulation of the Two Component Adjustment .....	70
6.4 Formulation of the Least-Squares Collocation ...	71
6.5 Diagram for the Least-Squares Collocation .....	76

TABLE OF CONTENTS (Cont'd)

	Page
6.6 Obtaining the Least-Squares Collocation Formulae from the Graph .....	78
6.7 Some Remarks About the Mathematical Structure of the Least-Squares Collocation .....	79
CHAPTER 7. CONCLUSION .....	84
REFERENCES .....	86

## LIST OF FIGURES

		Page
Figure 2.1	Illustration of Projection Theorem in a Three Dimensional Space	23
Figure 2.2	Illustration of a Linear Functional	26
Figure 3.1	Spaces Involved in a Combined Model	30
Figure 3.2	Iteration of a Linearized Implicit Model	36
Figure 4.1	Covariant and Contravariant Components of a Vector in a Two Dimensional Skew Cartesian Coordinate System	39
Figure 5.1	Diagram for the Parametric Least- Squares Adjustment Case	49
Figure 5.2	Diagram for a Conditional Least- Squares Adjustment Case	53
Figure 5.3	Diagram for a Combined Least- Squares Adjustment Case	57
Figure 6.1	Illustration of the Least-Squares Collocation in a Parametric Model	67
Figure 6.2	Spaces Involved in the Two Component Adjustment	69
Figure 6.3	Spaces Involved in the Least-Squares Collocation	73
Figure 6.4	Diagram for the Least-Squares Collocation	77
Figure 6.5.	Illustration of the Coordinate System in the Observation and Prediction Spaces	81



#### ACKNOWLEDGEMENT

I would like to express my gratitude to my supervisor, Dr. P. Vanicek, for suggesting this topic, and for his help, guidance and encouragement all through the research.

Appreciation is also extended to Dr. D.E. Wells and Mr. R.R. Steeves for their helpful suggestions.

I am indebted to Mr. C. Penton for his help in improving the English of the manuscript. Finally, I thank my fellow graduate students and the entire Department of Surveying Engineering in the University of New Brunswick.

## CHAPTER 1

### INTRODUCTION

#### 1.1 Concept of the Least-Squares Adjustment

The Least-Squares criterion is an imposed condition for obtaining a unique solution for an incompatible system of linear equations. The term adjustment, in a statistical sense, is a method of deriving estimates for random variables from their observed values. The application of the least-squares criterion in the adjustment problem is called the Least-Squares Adjustment method.

Since this thesis is closely related to the least-squares adjustment problem and will actually present a new approach for solving this problem, let us first have a closer look at the classical approach to the least squares method.

Since its first application to a problem in astronomy by Gauss in 1795, least-squares adjustment has been applied in a large number of fields in science and engineering, particularly in geodesy, photogrammetry and surveying. The practical importance of adjustment has been enhanced by the introduction of electronic computer and reformulation of the technique in matrix notation.

Generally, in the adjustment problems we are dealing with two types of variables; that is, (a vector of) observations, with a priori sample values, denoted by  $l$ , and (a vector of) parameters without a priori sample, denoted by  $x$ . These two types of data have random

nature. Each type of data is mathematically related by a functional or mathematical model.

Randomness of observations on one hand calls for redundant observations, and on the other hand causes the redundant observations to be incompatible with the mathematical model. In other words, a system of inconsistent equations comes into existence. To remove the incompatibility of the original set of observations with the model, this set will be replaced by another set of estimates  $\hat{l}$ , of  $l$  which satisfies the model. The difference between the two sets,  $l$  and  $\hat{l}$ , is denoted by  $r$ ; that is,

$$r = \hat{l} - l, \quad (1.1)$$

and is usually called the residual vector.

Due to redundant observations, there would be an infinite number of estimates of  $l$ . Among all the possibilities there exists one set of estimates that, in addition to being consistent with the model, satisfies another criterion, commonly referred to as the least-squares principle. This principle states that the sum of squares of the residuals is a minimum, that is,

$$\phi = \min_r (r^t W r), \quad (1.2)$$

where  $W$  is the weight matrix of the observations, and  $\phi$  is the variation function. Application of this criterion in the adjustment problems makes the vector  $\hat{l}$ , as close as possible to the original observations  $l$ , as well as allowing the relative uncertainty of observations to be taken into consideration through the elements of the weight matrix,  $W$ .

Where the observations are uncorrelated, the above general expression will reduce to

$$\phi = \min_v \sum_i w_i v_i^2, \quad (1.3)$$

where  $w_i$  is the  $i$ th diagonal element of  $W$  and  $v_i$  is the residual associated with the corresponding  $i$ th observation.

The simplest case occurs when the observations are taken to be of equal precision, in addition to being uncorrelated. Then  $W$  becomes the identity matrix and

$$\phi = \min_{r_i} \sum_{i=1}^n (r_i^2). \quad (1.4)$$

This latter case is the oldest and possibly the one which gave rise to the name "least-squares", since in this case we seek the "least" of the sum of the squares of residuals.

It should be noted that the application of the least squares principle does not require a priori knowledge of the statistical distribution associated with the observations. All that is necessary is to have the weight matrix,  $W$ , which is related to the covariance matrix of observations. This latter matrix will be denoted by  $C_1$ .

As mentioned earlier, in addition to the observations, the model may also include other variables (parameters) and numerical constants. The "solution" is the determination of the vector of parameters,  $x$ , and its covariance matrix, which will be denoted by  $C_x$  (if we are interested in the accuracy of  $x$ ). Therefore, the solution of the least-squares adjustment can be interpreted by the following transformation:

$$(1, C_1) \dots > (x, C_x) \quad (1.5)$$

It is important to note that the application of the least-squares principle in an adjustment problem is possible if the corresponding model is linear. Practically, however we are dealing with nonlinear models. This obstacle can be removed by replacing the model with its approximate linearized form. For this purpose Taylor series expansion is usually applied.

Linearized models can be classified as conditional, parametric and combined [see e.g. Vaníček and Krakiwsky, 1981]. The solution of a linearized mathematical model by least-squares method is a subject which has been thoroughly treated in the literature [see e.g. Mikhail, 1976]. This process is reviewed in Chapter 2 of this work.

The classical language used in the explanation of the least-squares adjustment, although scientifically adequate, lacks in insight from the geometric point of view, which seems to be a shortcoming. Least squares adjustment method, as a technique which has an extended application in the different fields of engineering, applied mathematics and statistics, should be interpreted geometrically to be more useful and understandable to the different users. The explanation of the least-squares method by geometric language has been quite helpful in obtaining the solution in very easy ways. Thus far, an approach has been made to interpret the least-squares adjustment solution geometrically, using Hilbert space technique.

Hilbert spaces, as the complete inner product spaces are more useful than the general Banach

spaces. The geometric meaning of a norm as a generalization of the length of a vector, along with the completeness and inner product properties of Hilbert spaces, make it possible for the least-squares problems in these spaces to be transformed to the minimum norm problems. These problems can be easily solved by the application of the projection theorem, which is a generalization of the following simple theorem in a three dimensional vector space. This theorem states that among all projections of a vector into a plane (in a three dimensional space), the orthogonal projection is the closest to the original vector.

Language and methods of the Banach spaces theories are quite popular in approximation and optimization problems [see e.g. Luenberger, 1969; Wells, 1974; Meissl, 1977; Adam, 1980].

The recent geometric interpretation of the least-squares method can be found in Vaníček [1979], where he considered the parallelism between the metric aspects of tensor operations and multidimensional probability measure, and used the duality relation, between the covariant and contravariant vector spaces, for the splitting of the observation and parameter vector spaces. Then he applied these relations in a linear parametric case of adjustment to construct a diagram, by which all the involved formulae can be easily obtained.

This method for solving the least-squares problems, in addition to geometrical interpretability, has a great advantage over the previous approach; that is, it is illustrative, which makes it

more interesting from both the scientific and technical points of view. We note that the projection theorem is imbedded in the theory of dual spaces.

The extreme ease in obtaining the solution and the elegance and power of the mathematical tool were the reasons for trying to apply it to more general cases in the least-squares adjustment and approximation. This thesis is an extension of that original work onto the conditional and combined cases of the least-squares adjustment, and on the collocation technique. Least-squares collocation is a method which combines the classical adjustment with prediction and has applications in different fields of geodesy and photogrammetry.

## 1.2 Outline of Treatment

This thesis is written in seven chapters. Following some introductory remarks in the first chapter, the second chapter is devoted to describing a metric space with the emphasis on the properties of complete metric spaces. The purpose of this section is to relate the Euclidean metric to the rectangular and skew Cartesian coordinate systems. In this part, vector spaces, normed and Banach spaces and inner product and Hilbert spaces are defined. Also, Hamel basis as well as total sets are mentioned. All the related theorems are stated without proof.

In the third chapter the formulation and solution of the least-squares equations are reviewed. In this part, the terminology is taken completely from Vaníček and Krakiwsky [1981].

The fourth chapter reviews the meaning of the covariant and contravariant components of a vector, and the metric tensor from the geometric point of view, with emphasis on the properties of a covariance matrix and its inverse, which can be used as metric for the corresponding spaces.

In the fifth chapter the properties and principles of the diagrammatic approach are described using the original example of Vanicek [1979] and then the diagrams for conditional and combined cases are developed. Also in this chapter, the case of a singular weight matrix is geometrically interpreted, and finally the use of the diagrams as a way for obtaining the reflexive least-squares  $g$ -inverse for a singular matrix is taken into consideration.

The sixth chapter deals with collocation, and its meaning is reviewed, and then the corresponding diagram is presented.

Conclusions are presented in the seventh chapter.



## CHAPTER 2

### SELECTED DEFINITIONS AND CONCEPTS FROM FUNCTIONAL ANALYSIS

#### 2.1 Metric Spaces

##### 2.1.1 Definition of metric spaces

A metric space is a pair  $(X, d)$  where  $X$  is a set and  $d$  is a metric (or distance function) on  $X$ . The set,  $X$ , is usually called the underlying set of  $(X, d)$  and its elements are called points. The metric can be selected in any way, but it must be selected in such a way that for all  $x, y, z \in X$  the following relations are satisfied:

$$(a) \ d \text{ is real-valued, finite and nonnegative} \quad (2.1)$$

$$(b) \ d(x, y) = 0 \text{ if and only if } x = y \quad (2.2)$$

$$(c) \ d(x, y) = d(y, x) \quad (\text{symmetry}) \quad (2.3)$$

$$(d) \ d(x, y) \leq d(x, z) + d(z, y) \quad (\text{triangle inequality}) \quad (2.4)$$

The above axioms are the general conditions for a functional  $d$  to be a metric.

In the selection of a practically useful metric, the particular properties of the space under consideration, and the relations between its points must be taken into account. For example, in a vector space which has an algebraic construction; that is, vector addition and multiplication of vectors by scalars, the selected metric, to be applicable and useful, must be related somehow

to this algebraic structure. In these spaces, such a metric can usually be defined via an auxiliary concept, the norm, which uses the algebraic operations in the vector spaces.

### 2.1.2 Complete metric spaces

A sequence  $(x_n)$ , in a metric space  $(X, d)$ , is said to be Cauchy if, for every  $\epsilon > 0$ , there is an  $N = N(\epsilon)$ , such that:

$$d(x_m, x_n) < \epsilon \text{ for every } m, n > N. \quad (2.5)$$

Completeness is an important property, which a metric space may have. A complete metric space is the one on which every Cauchy sequence is convergent. In other words, the limit of a Cauchy sequence on a complete metric space is a point of that space.

In the following, some related theorems and definitions from functional analysis and particularly from finite dimensional vector spaces will be presented.

### 2.1.3 Closure of a subset, dense set, separable set

Let  $M$  be a subset of a metric space  $X$ . Then a point of  $X$  (which may or may not be a point of  $M$ ) is called an accumulative point of  $M$  (or limit point of  $M$ ) if every neighborhood of  $x$  contains at least one point  $y \in M$  distinct from  $x$ . The set consisting of the points of  $M$  and the accumulation points of  $M$  is called the closure of  $M$  and is denoted by  $\bar{M}$ . It is the smallest closed set containing  $M$ .

A subset  $M$  of a metric space  $X$  is said to be dense in  $X$  if

$$\bar{M} = X.$$

X is said to be separable if it has a countable subset which is dense in X.

#### 2.1.4 Vector spaces, Linear spaces

A vector space over a field K is a nonempty set of elements,  $x, y, \dots$  (called vectors), together with two algebraic operations. These operations are vector addition and multiplication of vectors by scalars, that is, by elements of K. Vector addition is commutative and associative. Multiplication of vectors by scalars is distributive, relative to the addition.

The set of all linear combinations of vectors in a non-empty subset, S, of a vector space, V, denoted by L(S), is a subspace of V containing S. Furthermore, if W is another subspace of V containing S, then L(S) is a subspace of W. In other words, L(S) is the smallest subspace of V containing S. Hence, it is called the subspace spanned by S.

#### 2.1.5 Linear dependence, independence

The vectors,  $x_1, x_2, \dots, x_n \in X$ , where X is a vector space over a field K, are said to be linearly dependent over field K, or simply dependent, if there exists scalars,  $a_1, a_2, \dots, a_m \in K$  not all of them zero, such that:

$$a_1 x_1 + a_2 x_2 + \dots + a_m x_m = 0. \quad (2.6)$$

Otherwise, the vectors are said to be independent.

2.1.6 Vector spaces of finite and infinite dimensions

A vector space,  $S$ , is said to be finite dimensional if there is a positive integer,  $n$ , such that  $X$  contains a linearly independent set of  $n$  vectors, whereas any set of  $n+1$  or more vectors of  $X$  is linearly dependent. Then,  $n$  is called the dimension of  $X$ .

If  $X$  is not of finite dimension, it is said to be infinite dimensional. If  $X$  is of finite dimension,  $n$ , a set of  $n$  linearly independent vectors,  $e$ , is called a base for  $X$ . If  $e = [e_1, e_2, \dots, e_n]$  is a basis for  $X$ , every  $x \in X$  has a unique representation as a linear combination of the base vectors:

$$x = a_1 e_1 + a_2 e_2 + \dots + a_n e_n . \quad (2.7)$$

For instance, an  $n$ -dimensional Euclidean space  $E_n$  has a base

$$\begin{aligned} e_1 &= (1, 0, 0, \dots, 0) \\ e_2 &= (0, 1, 0, \dots, 0) \\ &\vdots \\ e_n &= (0, 0, 0, \dots, 1) \end{aligned} \quad (2.8)$$

which is called a canonical basis for  $E_n$ .

More generally, if  $X$  is any vector space, not necessarily finite dimensional, and  $B$  is a linearly independent subset of  $X$  which spans  $X$ , then  $B$  is called a base (or Hamel base) for  $X$ . Hence if  $B$  is a base for  $X$ , then every nonzero  $x \in X$  has a unique representation as a linear combination of (finitely many) elements of  $B$  with nonzero scalars as coefficient. Every vector space  $X \neq \{0\}$  has a base.

All the bases for a given (finite or infinite dimensional) vector space  $X$  have the same cardinality. This number is  $n$ , the dimension of  $X$ .

### 2.1.7 Normed spaces, Banach spaces

A normed space,  $X$  is a vector space with a norm defined on it. A Banach space is a complete normed space (complete in the metric induced by the norm). Here a norm on a vector space is a real valued function on  $X$  whose value for  $x \in X$  is denoted by  $\|x\|$ , and which has the following properties:

$$(a) \quad \|x\| \geq 0 \quad (2.9)$$

$$(b) \quad \|x\| = 0 \iff x = 0 \quad (2.10)$$

$$(c) \quad \|\alpha x\| = |\alpha| \|x\| \quad (2.11)$$

$$(d) \quad \|x+y\| \leq \|x\| + \|y\| \quad (2.12)$$

The metric,  $d$ , induced on  $X$  by the norm is given by

$$d(x, y) = \|x-y\|. \quad (2.13)$$

### 2.1.8 Completeness theorem

Every finite dimensional subspace of a normed space is complete. In particular, every finite dimensional normed space is complete.

### 2.1.9 Closedness theorem

Every finite dimensional subspace of a normed space is closed in that space.

2.1.10 Inner product spaces, Hilbert spaces

An inner product space is a vector space with an inner product defined in it. A Hilbert space is a complete inner product space, complete in a metric induced by the inner product.

An inner product,  $\langle x, y \rangle$ , on a vector space,  $X$ , defines a norm as:

$$||x|| = \langle x, x \rangle^{1/2}. \quad (2.14)$$

The metric induced by the inner product is of the following form:

$$d(x, y) = ||x-y|| = \langle x-y, x-y \rangle^{1/2}. \quad (2.15)$$

If with such a metric, the corresponding inner product space is complete, then it is a Hilbert space. From equation (1.15), it is clear that:

$$d(x, 0) = \langle x, x \rangle^{1/2} = ||x||. \quad (2.16)$$

Also, it is evident that the square of the distance between two points (vectors),  $x$  and  $y$ , is equal to the scalar product  $\langle x-y, x-y \rangle$  with itself. For example, consider the Banach space,  $\ell^p$ , where  $p \geq 1$ . By definition each element in the space,  $\ell^p$ , is a sequence of elements  $x = (\xi_i) = (\xi_1, \xi_2, \dots)$ , such that:  $|\xi_1|^p + |\xi_2|^p + \dots$  converges; thus

$$\sum_{i=1}^{\infty} |\xi_i|^p < \infty \quad (p \geq 1, \text{ fixed}), \quad (2.17)$$

and the metric is defined by

$$d(x, y) = \left( \sum_{i=1}^{\infty} |\xi_i - \eta_i|^p \right)^{1/p}. \quad (2.18)$$

Or, in real space

$$d(x, y) = \left( \sum_{i=1}^{\infty} |\xi_i - \eta_i|^p \right)^{1/p}. \quad (2.19)$$

In the case of  $p = 2$ , we have the Hilbert sequence space,  $l^2$ . This space is an Hilbert space with inner product defined by

$$\langle x, y \rangle = \sum_{i=1}^{\infty} \xi_i \eta_i \quad (\text{in real space}) \quad (2.20)$$

For  $(i = 1, 2, 3)$  this space is Euclidean  $E^3$ .

We know from elementary analytic geometry that if a rectangular Cartesian coordinate system is used, then an inner product can be defined in the three dimensional Euclidean space  $E^3$  by

$$\langle x, y \rangle = \sum_{i=1}^3 \xi_i \eta_i \quad (2.21)$$

and the corresponding metric is defined by

$$d(x, y) = \langle x - y, x - y \rangle^{1/2} = \left( \sum_{i=1}^3 (\xi_i - \eta_i)^2 \right)^{1/2} \quad (2.22)$$

or

$$d^2(x, y) = \sum_{i=1}^3 (\xi_i - \eta_i)^2, \quad (2.23)$$

where  $\xi_i$  and  $\eta_i$  are the components of the vector  $x$  and  $y$  respectively.

The metric is called The Euclidean metric.

In the non-geometric sense, however, instead of the orthogonal coordinate axes we will have an orthogonal base, which is a total set (or fundamental set) in a finite dimensional vector space.<sup>1</sup> If the elements of this set have norm 1, this base is said to be orthonormal. As a consequence, every Hilbert space has a total orthonormal set.

## 2.2 Linear Operators

### 2.2.1 Definition of linear operators

A linear operator,  $T$  is an operator such that:

- (a) the domain,  $D(T)$ , of  $T$  is a vector space and the range,  $R(T)$ , lies in a vector space over the same field;
- (b) for all  $x, y \in D(T)$  and for all scalars:

$$T(x+y) = Tx + Ty \quad (2.24)$$

$$T(\alpha x) = \alpha Tx, \quad (2.25)$$

( $Tx$  and  $Ty$  stand for  $T(x)$  and  $T(y)$ ).

By definition, the null space (or kernel) of  $T$ , denoted by  $N(T)$ , is the set of  $x \in D(T)$  such that  $Tx = 0$ . It can be shown that  $N(T)$  is a vector space. Clearly, a mapping  $T$  from a space  $X$  onto another space  $Y$  is restricted to  $D(T)$  and  $R(T)$ ; in other words,

$$T: D(T) \text{ -----} \rightarrow R(T) . \quad (2.26)$$

If  $D(T)$  is the whole space,  $X$ , it can be written as

---

<sup>1</sup> A total set in a normed space  $X$  is a subset  $M \subset X$  whose span is dense in  $X$ . Accordingly, an orthonormal set in an inner product space  $X$  which is total in  $X$  is called a total orthonormal set.



$$T: X \longrightarrow Y \tag{2.27}$$

As an example, consider the equation,  $y = A x$ , where  $A = (\alpha_{ij})$  is a real matrix with  $r$  rows and  $n$  columns which defines an operator from an  $n$ -dimensional Euclidean space onto an  $r$ -dimensional Euclidean space;

$$T: E^n \longrightarrow E^r, \text{ by means of,} \tag{2.28}$$

$$y = A x$$

$T$  is linear since matrix multiplication is a linear operation.

### 2.2.2 Bounded linear operator

Let  $X$  and  $Y$  be normed spaces and  $T: D(T) \longrightarrow Y$  a linear operator where  $D(T) \subset X$ . The operator,  $T$ , is said to be bounded if there is real number,  $C$ , such that for all  $x \in D(T)$

$$\|Tx\| \leq C\|x\| \tag{2.29}$$

The norm of operator  $T$  can be obtained by

$$\|T\| = \sup_{\substack{x \in D(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|} \tag{2.30}$$

or,

$$\|T\| = \sup_{\substack{x \in D(T) \\ \|x\|=1}} \|Tx\| \tag{2.31}$$

The operator,  $T$ , defined by a real matrix in the above example is a bounded linear operator.

### 1.2.3 Finite dimension theorem

If a normed space is finite dimensional, then every linear operation on it is bounded.

## 2.3 Linear Functional

### 2.3.1 Definition of linear functionals

A linear functional,  $f$ , is a linear operator with domain in a vector space,  $X$ , and range in scalar field  $K$  of  $X$ ; thus

$$f: D(f) \rightarrow K. \quad (2.32)$$

### 2.3.2 Bounded linear functional

A bounded linear functional,  $f$ , is a bounded linear operator with range in the scalar field of the normed space,  $X$ , in which the domain,  $D(f)$  lies.

## 2.4 Dual Spaces

### 2.4.1 Definition of dual spaces

The set of all linear functionals on a vector space,  $X$ , can itself be made into a vector space. This space is denoted by  $X^*$ , and is called algebraic dual space of the space  $X$  simply because there is no topological structure on it and on the primal vector space. For example, there is no norm nor metric, which would allow the definition of analytical or topological concepts, such as convergence and open and closed sets, to be possible.

In the case of a normed space, the set of all bounded linear functionals on it constitutes a second normed space, which is called the dual space of the primal space, and is denoted by  $X'$ .

The interrelation between a normed space and its dual space has an interesting consequence in the optimization problems. For example, if the problem in the primal space is a maximization problem, the problem in its dual space will be a minimization problem.

#### 2.4.2 Dual bases

In this section the concept of bases (or coordinate systems) will be used to show the relation between a vector space and its dual space. In this respect the following theorems are mentioned.

(a) Theorem 1

If  $X$  is an  $n$ -dimensional vector space,  $[x_1, x_2, \dots, x_n]$  is a basis in  $X$ , and if  $[\alpha_1, \alpha_2, \dots, \alpha_n]$  is a set of  $n$  scalar, then there is one, and only one, linear functional,  $f$ , on  $X$  such that  $f(x_i) = \alpha_i$ , for  $i = 1, 2, \dots; n$ .

(b) Theorem 2

If  $X$  is an  $n$ -dimensional vector space and if  $x = [x_1, x_2, \dots, x_n]$  is a basis in  $X$ , then there is a uniquely determined basis  $x^* = [f_1, f_2, \dots, f_n]$  in  $X^*$  with the property that  $f_j(x_i) = \delta_{ij}$ ; consequently, the dual space of an  $n$ -dimensional space is  $n$ -dimensional. The basis  $X^*$  is called the dual basis of  $x$ .

(c) Theorem 3

If  $x_1$  and  $x_2$  are two different vectors of the  $n$ -dimensional vector space  $X$ , then there exists a linear functional,  $f$ , on  $X$  such that  $f(x_1) \neq f(x_2)$  or, equivalently, to any non-zero vector,  $x$ , in  $X$  there corresponds an  $f$  in  $X^*$  such that  $f(x) \neq 0$ . Of course, all of these theorems are related to algebraic dual spaces. The set of all linear functionals on  $X^*$  (dual to vector space  $X$ ) is called second dual space of  $X$  and is denoted by  $X^{**}$ , thus  $X^{**} = (X^*)^*$ . To each  $x \in X$  there corresponds a  $g \in X^{**}$ . This defines a mapping

$$C: X \text{ ----> } X^{**} , \quad (2.33)$$

where  $C$  is called the canonical mapping of  $X$  into  $X^{**}$ . If  $C$  is surjective (hence bijective)<sup>2</sup>, so that  $R(C) = X^{**}$ , then  $X$  is said to be algebraically reflexive.

It can be shown that a finite dimensional vector space,  $X$ , is algebraically reflexive. The following theorem, in relation to dual spaces, is of interest.

"The dual space,  $X'$ , of a normed space,  $X$ , is a Banach space (whether or not  $X$  is a Banach space)".

---

<sup>2</sup> A mapping  $T$  is injective or one-to-one if every  $x_1, x_2 \in D(T)$ ,  
 $x_1 \neq x_2$  implies  $Tx_1 \neq Tx_2$

$T: D(T) \text{ ----> } Y$  is surjective, or a mapping of  $D(T)$  is onto  $Y$   
 if  $R(T) = Y$   
 $T$  is bijective if it is both injective and surjective.

### 2.4.3 Annihilator

If  $M$  is a non-empty set of a normed space,  $X$ , the annihilator  $M^\alpha$  of  $M$  is defined to be the set of all bounded linear functionals on  $X$  which are zero everywhere on  $M$ . Thus  $M^\alpha$  is a subset of dual space  $X'$  of  $X$ . It can be shown that  $M^\alpha$  is a vector subspace of  $X'$  and is closed. If  $M$  is an  $m$ -dimensional subspace of  $n$ -dimensional normed space  $X$ ,  $M^\alpha$  is an  $(n-m)$  dimensional subspace of  $X'$ .

## 2.5 Projection Theorem

### 2.5.1 Orthogonality

The property of orthogonality which exists in Hilbert spaces has made these spaces simpler and richer than the general Banach spaces. This property is used in the definition of the projection theorem, which is the most important in this context from the functional analysis view.

### 2.5.2 Explanation of the projection theorem

If  $Z$  is a closed subspace of a Hilbert space  $H$ , then corresponding to any vector  $x \in H$  there is a unique vector  $z \in Z$  such that  $\|x - z_0\| \leq \|x - z\|$  for all  $z \in Z$ . Furthermore, a necessary and sufficient condition that  $z_0 \in Z$  be the unique minimizing vector is that  $x - z_0$  be orthogonal to  $Z$ .  $z_0$  is the best approximation or solution for  $x$  in  $Z$ .

Note that the closeness of the subspace  $Z$  guarantees the existence of  $z_0$ ; therefore, finite dimensionality of  $Z$  is a

sufficient condition for  $z_0$  to exist (see also figure (2.1)).

### 2.5.3 Orthogonal complement

Given a subset of a inner product space, the set of all vectors orthogonal to  $S$  is called the orthogonal complement of  $S$  and is denoted  $S^\perp$ . Clearly if  $S$  consists of only the null vector its orthogonal complement is the whole space.

An important point: An orthogonal complement can be taken as an annihilator (see 1.5.4).

For any set  $S$ ,  $S^\perp$  is a closed subspace; also  $S \subset S^{\perp\perp}$ , where  $S^{\perp\perp}$  is the orthogonal complement of  $S^\perp$ .

If  $S$  is a closed subspace of a Hilbert space  $H$ , then  $S = S^{\perp\perp}$ .

In a normed space, the notion of orthogonality can be introduced through the dual space.

### 2.5.4 Dual spaces and orthogonality

The vector  $x \in X$  and  $x' \in X'$  are said to be orthogonal if  $\langle x, x' \rangle = 0$ . In a Hilbert space the linear functionals are generated by the elements of the space itself, which implies that the dual of a Hilbert space  $X$  is  $X$  itself. The above definition of orthogonality can be regarded as a generalization of the corresponding Hilbert space definition.

By this definition it is clear that the orthogonal complement of a subset of a normed space is its annihilator.

### 2.5.5 Minimum norm problem and dual space

Let  $X$  and  $Y$  be two vector spaces and let  $y = A x$  be a linear transformation between the elements  $y \in Y$  and  $x \in X$ , where  $A$  is a matrix. From the transformation, we can say that the columns of matrix  $A$  span a closed subspace in  $Y$ . A vector  $y \in Y$  can be projected in infinitely many ways into this closed subspace. These projections are different vectors with different norms. The minimum norm problem is actually that of finding the orthogonally projected vector among these projections. Figure 2.1 is the geometrical visualization of the problem in a three dimensional space, where the above mentioned closed subspace is indicated by  $Z$  and  $z \in Z$  [see Adam, 1980]. The minimum norm problem, geometrically, is a generalization of this elementary geometrical problem.

In Figure 2.1 we can see that the closest vector to  $y$  in  $Z$  is the one for which the distance of point  $y$  from  $Z$  is of minimum length. Therefore finding the minimum norm solution (best approximation) of the equation  $y = Ax$  will result in finding the minimum norm vector between the set of vectors whose norms are equal  $\|y - Ax\|$  when  $Ax$  varies over the closed subspace spanned by the columns of the matrix  $A$ . The above mentioned set of vectors are called error vectors. In a general normed space many error vectors may have the property of minimum norm. In a Hilbert space, however, the uniqueness of the solution is available and the solution satisfies an orthogonality condition. The square of the distance of a point  $y$ , in this space, from a closed subspace can be obtained by scalar product of the orthogonal error vector with itself.

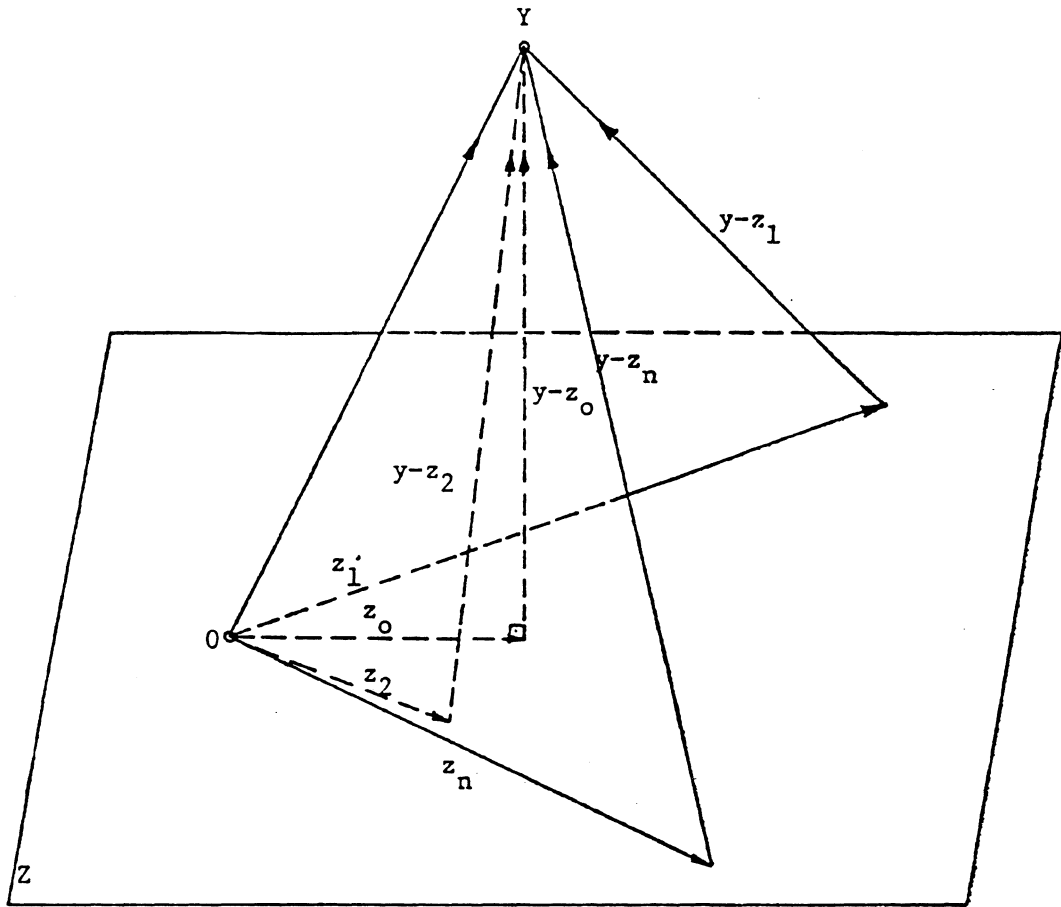


Figure 2.1. Illustration of Projection Theorem in a Three Dimensional Space



In the next section of this chapter we will see that finding the minimum distance mentioned above is a problem which can be solved in the dual space.

Also, it is clear from Figure 2.1 that finding the best approximation for the vector  $y$  in the subspace  $Z$  and finding the minimum distance from a point  $y \in Y$  to  $Z$ , are two equivalent problems in the sense that one of these problems is solved, the other is automatically solved.

#### 2.5.6 Dual spaces and least squares

Before proceeding with this topic, the following lemma is worth noting:

LEMMA: Existence of a functional

"Let  $Z$  be a proper subspace of an  $n$ -dimensional vector space  $Y$ , and let  $y \in Y-Z$ . There exists a linear functional,  $f$ , such that  $f_0(y) = 1$  and  $f(z) = 0$  for all  $z \in Z$ ."

In a normed space, this lemma may be restated as:

"Let  $Z$  be a proper closed subspace of a normed space  $Y$ .  $L$  be arbitrary and let

$$\delta = \inf_{z \in Z} \|z - y\|, \quad (2.34)$$

be the distance from  $y$  to  $Z$ . Then there exists an  $f_0 \in Y$  such that

$$\|f_0\| = 1, \quad f_0(z) = 0 \quad \text{for all } z \in Z, \quad f_0(y) = \delta. \quad (2.35)$$

To visualize this lemma in three dimensional space, let us once again take Figure (2.1) into consideration. This time, however,

a vector  $f_0$  with unit length is added to the figure, drawn from the origin (point 0), and orthogonal to the Z plane, (see figure (2.2)). The set of error vectors is given by

$$\{z-y\}$$

and the length of an error vector by

$$\|z-y\|$$

The shortest length, which is the distance between the point y and plane Z, is

$$\delta = \inf_{z \in Z} \|z-y\| \tag{2.36}$$

This distance belongs to the error vector which is perpendicular to the Z plane, and is thus parallel to the vector  $f_0$ . Inner products of  $f_0$  and all  $z \in Z$  are equal to zero. The problem of the least squares in a system of inconsistent equations,  $y = Ax$ , where  $x$  and  $y$  are elements of two normed spaces, X and Y, is finding an  $\hat{x} \in X$  which satisfies the relation

$$\|Ax-y\| = \inf_x \|Ax-y\| \tag{2.37}$$

As mentioned in the previous section, for some vector  $y \in Y$  the form  $\{Ax-y\}$ , where  $x$  is a variable vector, would show a set of error vectors. The norms of these error vectors are comparable to the distances of the point  $y \in Y$  from the subspace Z in Figure 2.1. As shown, the shortest distance is the one for which the corresponding vector  $z-y$  is orthogonal to the subspace Z.

According to the above lemma and the definition of the dual basis (2.4.2, theorem 2), the set of orthonormal vectors which



constitute the dual basis, are indeed the linear functionals (as elements of dual space of  $Y$ ) for which  $f_{\alpha}(y) = \delta$  (see formula (2.35)). The understanding of the use of dual spaces in least-squares problem needs some geometrical visualization, since one is dealing with two spaces, in the sense that in this approach the orthogonal complement of a subspace of a space, is a subspace of the dual space.

The fact that the solving of a least squares problem in a vector space can be converted to the problem of finding the shortest distance in its dual space has a surprising consequence, which plays the main role in this thesis.

## CHAPTER 3

### FORMULATION AND SOLUTION OF THE LEAST-SQUARES PROBLEM - A REVIEW

In the adjustment two types of data are considered: observations and parameters. Both types of data are related through a mathematical model. Mathematical models, from the point of view of their solutions, can be categorized in three classes; unique, undetermined and overdetermined, depending whether the number of parameters is equal, greater than, or less than the number of observations respectively [see Vaníček & Krakiwsky, 1981]. In the following, the formulation and solution of overdetermined models is reviewed.

#### 3.1 Formulation of Least-Squares Problem

The general form of the mathematical model is:

$$f(x, l) = 0 \quad (3.1)$$

where  $x$  and  $l$  stand for parameters and observations respectively. The randomness of data causes the problem to have no solution. This simply means that the mathematical model (3.1) has inconsistent equations, and reformulation of the model is necessary to remove these inconsistencies. Accordingly, equation (3.1) becomes

$$f(x, \hat{l}) \implies f(x, l+r) = 0 \quad (3.2)$$

where the corrected (estimated) observations,  $\hat{l}$ , are the estimates of the expected values  $\hat{l}$ .

To simplify the solution, the above model is normally approximated with the linear part of a Taylor series expansion, whether the model is linear or not. The result of this approximation is given by:

$$\begin{matrix} A & \delta' & + & B & r & + & w & = & 0 \\ \text{mxu} & \text{uxl} & & \text{mxn} & \text{nxl} & & \text{mxl} & & \end{matrix} \quad (3.3)$$

where

$$A = \frac{\partial f}{\partial x} \Big|_{x^{(0)}, l^{(0)}}, \quad B = \frac{\partial f}{\partial l} \Big|_{x^{(0)}, l^{(0)}}, \quad w = f(x^{(0)}, l^{(0)}), \quad (3.4)$$

$$\delta = x - x^{(0)}, \quad r = l - l^{(0)}$$

and  $x^{(0)}, l^{(0)}$  are points of expansion and are two vectors of approximate values of the unknown parameters and observed values, respectively.

The dimensions are given by  $m$ , the number of equations,  $u$ , the number of unknown parameters, and,  $n$ , the number of observations. Equation (3.3) is the differential form of the original non-linear mathematical model. Figure 3.1 depicts the linear relation of the quantities in the neighbourhoods of  $x^{(0)}, l^{(0)}$  and  $w$ .

### 3.2 Metricizing of the Spaces Involved in the Problem

We observe from Figure 3.1, that in equation (3.3) we are dealing with three spaces. These spaces are called observation, model and parameter spaces. In the case that we do not have any a priori information about the parameters, the elements of the

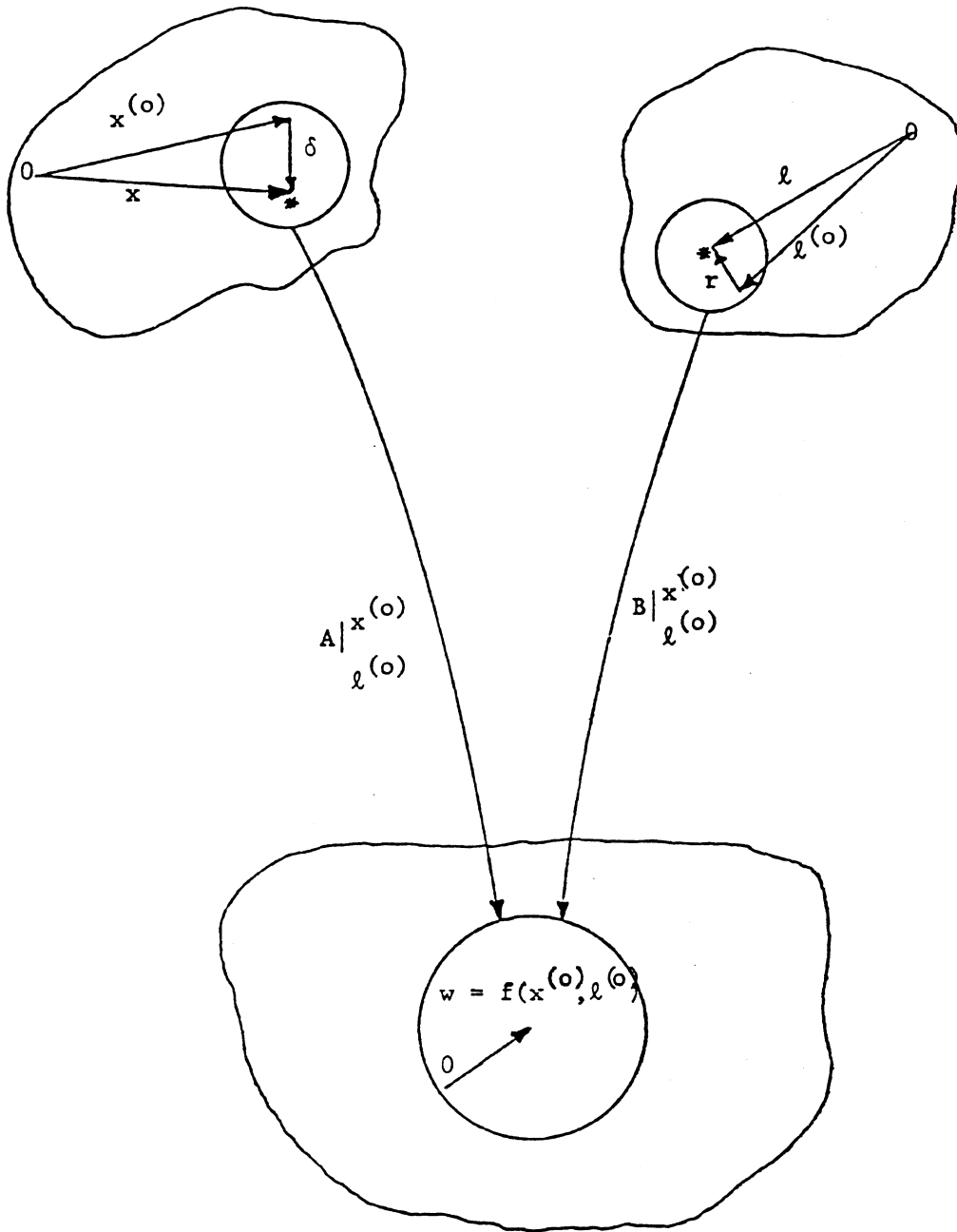


Figure 3.1 Spaces Involved in a Combined Model

parameter space are dependent absolutely on the elements of the observation space. The model space is a tool, or medium, which links the parameters to the observations.

As we know, adjustment problems are closely related to probabilities. However, least-squares, in its general definition, does not have any relation with probability. In the following discussion, the relation between least-squares adjustment and probability measure is considered.

Recall that a metric in a space is a distance function, that is, a tool for measuring the distance between two points in that space. Also we know that the elements of the covariance matrix and their inverses give us some probabilistic information about the observations. Therefore using the inverse of the covariance matrix of the observations as a metric in the observation space implies that we are measuring the distance between the points of observation space by a probabilistic tool. For example, if  $\vec{r}$  is a vector in the observation space and if this space has been metricized by  $C_r^{-1}$  (inverse of the covariance matrix of observation) then we will have

$$||r||^2 = \vec{r}^T C_r^{-1} \vec{r} .$$

In the case where we do not have any a priori information about the parameters, the metric of this space is induced by the metric of the observation space, via the model space. The covariance of the model space can be determined before the least-squares problem is solved. This can be done by the application of the covariance law to the equation of the misclosure vector;

$$w = f(x^{(0)}, l^{(0)}), \quad (3.5)$$

which yields



$$C_w = M^{-1} = \frac{\partial f}{\partial l} C_l \frac{\partial f^T}{\partial l} = BC_l B^T \quad (3.6)$$

where  $C_l = C_r$  is the covariance matrix of the observations. Also the term  $A^T M A$  will be denoted by  $N$ .

As we will see later in this chapter, the covariance matrix of the estimated values of the parameters is of the form:

$$C_x = N^{-1} = [A^T (BC_l B^T)^{-1} A]^{-1} = (A^T M A)^{-1} . \quad (3.7)$$

### 3.3 The Solution to the Least-Squares Problem

The purpose of the solution in the adjustment problem is to find the most probable values  $(\hat{\delta}, \hat{r})$  for  $\delta$  and  $r$  by minimizing the length of the vector  $r$ . To do this, either the length of  $r$  can be minimized in  $L$  (observation space) or the length of its projection,

$$\bar{r} = Br, \quad (3.8)$$

can be minimized in  $F$  (model space). The first case leads to

$$\min_{\delta, r} (r^T C_r^{-1} r) \quad (3.9)$$

while the second case corresponds to

$$\min_{\delta, \bar{r}} (\bar{r}^T \bar{C}^{-1} \bar{r}) \quad (3.10)$$

Using the second approach, the linearized model becomes

$$\bar{r} = -(A\delta + w) \quad (3.11)$$

Note that the formulation in  $F$  allows for the rewriting of the minimum condition directly in terms of  $\delta$ . Solution for  $\bar{r}$  in equation (3.10), from equation (3.11) yields:

$$\min_{\delta, \bar{r}} (\bar{r}^T \bar{C}^{-1} \bar{r}) = \min_{\delta, \bar{r}} ((A\delta + w)^T C_r^{-1} (A\delta + w)). \quad (3.12)$$

Formulation in L does not allow this direct substitution and thus has to be treated differently.

Because matrix B is usually a rectangular matrix, by minimizing the residual vector in the model space we can not obtain  $\hat{r}$ . Therefore the solution must be done in the observation space.

The solution is obtained by minimizing the following function:

$$\phi = r^T C_r^{-1} r + 2k^T (A\delta + Br + w) \quad (3.13)$$

instead of  $r^T C_r^{-1} r$ . This function is called the variation function, where k is an undefined vector from the model space called the vector of Lagrange correlates.

After taking the partial derivatives of the above equations with respect to r,  $\delta$  and k and equating them to zero we obtain:

$$C_r^{-1} \hat{r} + B^T k = 0 \quad (3.14)$$

$$A^T k = 0 \quad (3.15)$$

$$A\hat{\delta} + Br + w = 0, \quad (3.16)$$

After solving the above system of equations we obtain:

$$\delta = -[A^T (BC_r B^T)^{-1} A]^{-1} A^T (BC_r B^T)^{-1} w \quad (3.17)$$

or,

$$\hat{\delta} = -N^{-1} A^T M w, \quad (3.18)$$

$$\text{or, } N\hat{\delta} + u = 0, \text{ where } u = A^T M w \quad (3.19)$$

(M, N were previously defined in equations (3.6) and (3.7)).

Equations (3.19) are called the least-squares normal equations.

Also

$$\hat{k} = (BC_r B^T)^{-1} (A\hat{\delta} + w) \quad (3.20)$$

or

$$\hat{k} = M(A\hat{\delta} + w) \quad (3.21)$$

and

$$\hat{r} = -C_r B^T \hat{K} = -C_r B^T (BC_r B^T)^{-1} (A\hat{\delta} + w), \quad (3.22)$$

or,

$$\hat{r} = -C_r B^T (M A N^{-1} A^T M - M) w = C_r B^T L w, \quad (3.23)$$

where

$$L = M A N^{-1} A^T M - M \quad (3.24)$$

and finally,

$$\hat{x} = x^{(0)} + \hat{\delta} \quad (3.25)$$

$$\hat{1} = 1 + \hat{r}. \quad (3.26)$$

Equations (3.25) and (3.26) are the least squares estimates of the unknown parameters and the observations, respectively. From equation (3.25), after application of the covariance law, we obtain

$$C_x = (-N^{-1} A^T M) M^{-1} (N^{-1} A^T M)^T, \quad (3.27)$$

as mentioned earlier in this chapter  $A^T M A = N$ , therefore

$$C_x = N^{-1} = [A^T (BC_r B^T)^{-1} A]^{-1} = C_\delta. \quad (3.28)$$

Note that  $x^{(0)}$  is a constant vector and thus its covariance matrix is the null matrix.

The statistical evaluation of the solution  $(\delta, r)$  requires that, the covariance matrices of these results be determined. However these do not have anything to do with the subject of this thesis, so the determination will not be reviewed.

Two particular cases, that is, when the first design matrix  $A$  equals zero and when the second design matrix  $B$  equals to  $-I$  ( $-$ identity matrix) bear particular names: conditional and parametric

cases, respectively. For the solution in the parametric case, there is no need to use the Lagrange correlates, the solution can be obtained directly by minimizing  $r^T C_r^{-1} r$ . We shall say more about these cases in Chapter 5 when their diagrams are presented.

Before ending this chapter, it is necessary to mention how to handle the effect of non-linearity arising from replacing a non-linear model by its linear approximation. It has been shown by Pope [1974] that the non-linear solution can be arrived at by a series of repeated linear solutions. The linearized model, at the  $n$ th iteration, is given by

$$A^{(n)} (x^{(n+1)} - x^{(n)}) + B^{(n)} (l^{(n+1)} - l^{(n)}) + f(x^{(n)}, l^{(n)}) = 0, \quad (3.29)$$

where  $(x^{(n)}, l^{(n)})$  is the latest point of expansion, and  $A^{(n)}$  and  $B^{(n)}$  are evaluated at this point of expansion (see Figure 3.2).

The iterations should be carried out until two successive increments  $(\delta^{(n)})$  and  $(\delta^{(n+1)})$  are zero, to some practical limit.

In the implicit non-linear model both parameters and observations are subjected to updating because the model is non-linear in both quantities. Models explicit in  $l$  need be iterated only in  $l$ , in the same way as the parametric models are iterated in  $x$ .

It is clear that the method of least-squares does not have anything to do with the iteration procedure and is applied only within each iteration and on the linearized set of equations.

It should be noted, however, that the residual vector  $r$  consists of statistically dependent and statistically independent components (see Chapter 6). In the above review, only the statistically

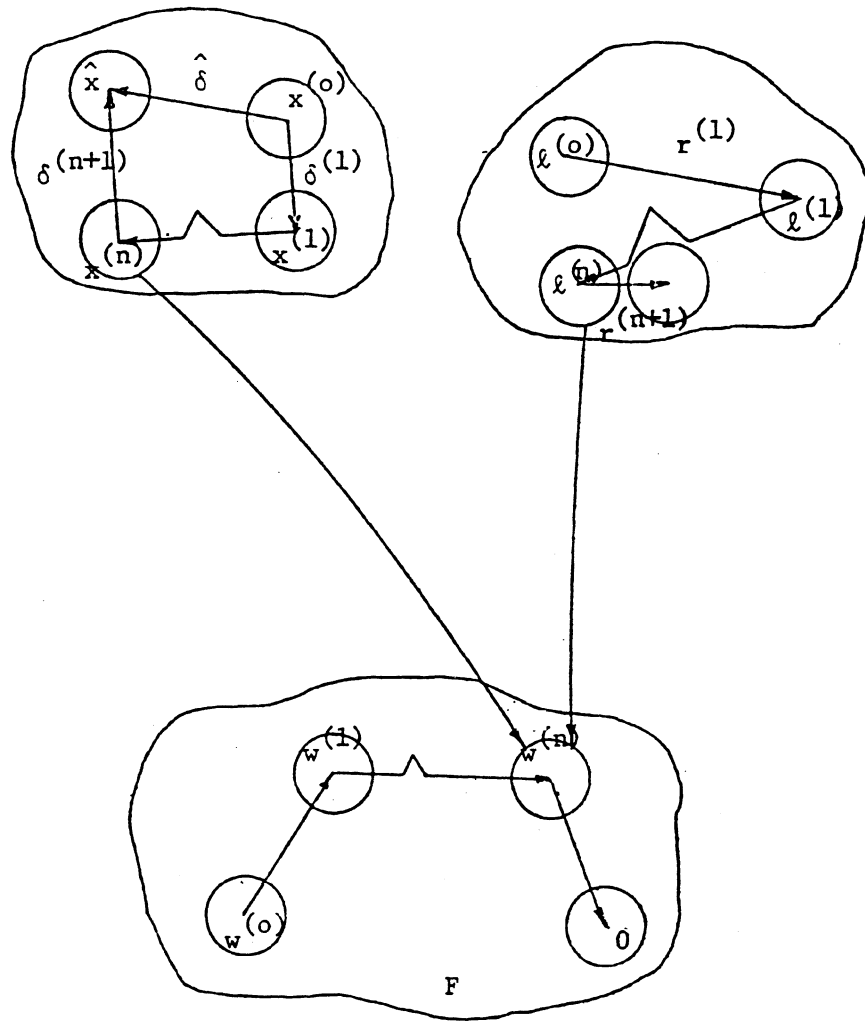


Figure 3.2. Iteration of a Linearized Implicit Model

independent component of the residual vector was considered. A review of the two components adjustment is given in Chapter 6, where the concept of collocation is taken into consideration.

## CHAPTER 4

### TENSOR STRUCTURE AND THE LEAST SQUARES

#### 4.1 Introduction

In chapter 2, the dual spaces and dual bases as well as their relation with the least-squares problems were described, from the functional analysis and particularly from the finite dimensional vector spaces points of view.

In the following, the fact that covariant and contravariant component spaces of a vector space are dual to each other is proven.

#### 4.2 Covariant and Contravariant Components of a Vector

In order to prove the duality relation between covariant and contravariant vector spaces, we consider the following figure from Hotine, [1969] (see Figure 4.1) and review the concept of covariant and contravariant components of a vector  $OP$  in a skew two dimensional space.

In a rectangular Cartesian coordinate system, the square of the length of a vector can be obtained by the summation of its squared components along the coordinate axes. The length of a vector does not depend on the origin of the coordinate axes, whereas the coordinates of a point are dependent on that point.

In a skew Cartesian system, if the point,  $O$ , of vector  $OP$  is taken as the origin of the coordinate system, the coordinates of

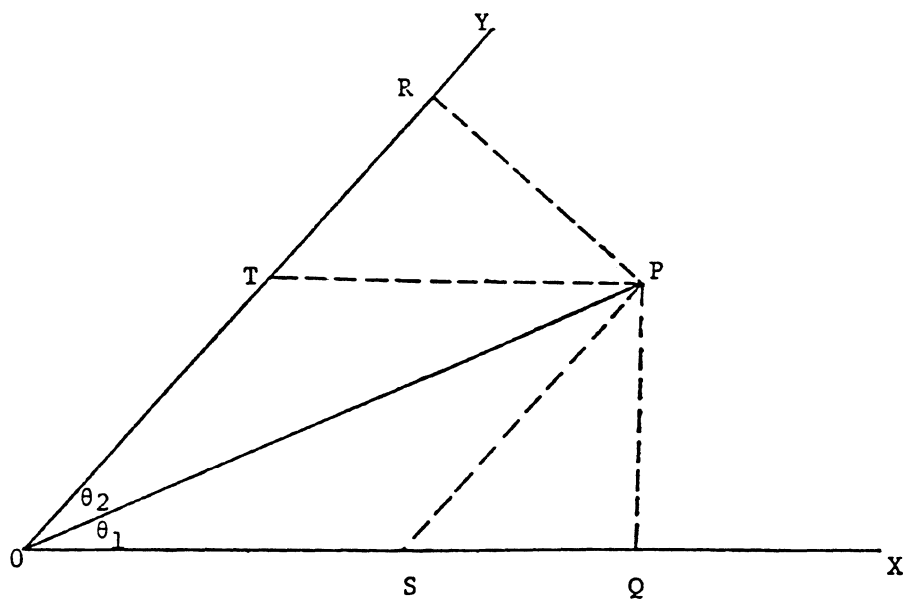


Figure 4.1 Covariant and Contravariant Components of a Vector in a Two Dimensional Skew Cartesian Coordinate System



point P will be actual lengths along the coordinate axes. In such a system it is still possible to specify the vector OP by its orthogonal projections, OQ, OR onto the coordinate axes (see figure 4.1).

The lengths OQ and OR are called covariant components of the vector OP and are obtained by:

$$l_1 = OQ = OP \cos \theta_1 \quad (4.1)$$

$$l_2 = OR = OP \cos \theta_2$$

We can see that the length of OP can not be obtained by the sum:

$$l_1^2 + l_2^2.$$

Alternatively, OP could be specified by taking the skew coordinates OS, OT as components, which are called the contravariant components of OP, and are obtained from

$$l^1 = OS = OP \sin \theta_2 / \sin (\theta_1 + \theta_2) \quad (4.2)$$

$$l^2 = OT = OP \sin \theta_1 / \sin (\theta_1 + \theta_2).$$

The square length of OP cannot be obtained from  $(l^1)^2 + (l^2)^2$  either, but with both (4.1) and (4.2) we get

$$l^1 l_1 + l^2 l_2 = OP^2 \quad (4.3)$$

According to the definition of linear functional (1.3), it is clear from formula (4.3) that if we consider the covariant and contravariant components of a vector as elements of two different spaces, then one is the set of all linear functionals on the other. It is also clear from the definition that skew coordinates play the role of covariant components.

The above mentioned Cartesian systems (rectangular and skew), are particular cases of the general form in which the coordinate axes are curvilinear and their directions are not parallel to the directions of the corresponding coordinate lines at other points.

In a rectangular coordinate system it is possible to specify a vector in a curvilinear coordinate system. In addition, the space itself may be curved, like the surface of a sphere in two dimensions. In this case the space can only be described in curvilinear coordinates.

A curvilinear coordinate will no longer necessarily be an actual length measured along a coordinate line in the case of Cartesian coordinates, although length and coordinates must obviously be related in some way, since a displacement over a given length in a certain direction must involve a unique change in coordinates. The relation, which may vary from point to point, is expressed by the metric or line element of the space.

To show the relation between the displacement over a given length in a certain direction and the corresponding change in the coordinates the following formula is used:

$$ds^2 = g_{rs} dx^r dx^s. \quad (4.4)$$

Because  $dx^r dx^s$  is two times contravariant and  $ds^2$  is an invariant, then according to the quotient law (see e.g. McConnell, [1957]),  $g$  must be a two times covariant tensor.

The tensor  $g_{rs}$  is called the metric tensor and has all the properties of a metric; that is, it is positive-definite and symmetric. Only in this way can formula (4.4) represent the square of a real element of length.

In a three dimensional vector space  $g_{rs}$  can be represented by a  $3 \times 3$  matrix, in which six of the nine elements may have different values. In this matrix the diagonal elements give the squares of the scale factors along the three coordinate axes, and off-diagonal elements are functions of the angles between the coordinate axes. More precisely, the diagonal elements are the scalar products of vectors along the coordinate axes with themselves where the vectors represent a displacement of one unit in terms of the coordinate system and off-diagonal elements are scalar products of these vectors with each other. Clearly in the Euclidean system the off-diagonal elements of the matrix  $g$  are zero, and the diagonal elements are equal to one. That is:

$$g_{rs} = \delta_{rs} \text{ (Kronecker delta)}$$

#### 4.3 Relation Between Covariant and Contravariant Components of a Vector

In Figure 4.1 we can see that the covariant components of a vector in Cartesian coordinates are some lengths along the coordinate axes. For a unit vector  $l$  we have, from (4.4):

$$g_{rs} l^r l^s = 1, \quad (4.5)$$

also by (4.3) for such a vector in the same system of coordinates we have,

$$l^S l_S = 1 \quad (4.6)$$

If these equations are to hold for all directions at a point, that is, for arbitrary values of the contravariant components  $l$ , we must have

$$l_S = g_{rs} l^r \quad (4.7)$$

This formula is in agreement with our earlier discussion about the dual spaces and the relation between their elements (see 1.4). As we know, if in a vector space, we define a set of vectors as the basis or coordinate axes, we can get the set of all linear functionals on that space. This set of functionals constitutes another vector space, which is called the dual space of the primal vector space. This formula is another illustration of the duality relation between covariant and contravariant spaces.

Obviously corresponding to a basis (or a coordinate system) in a covariant vector space, there is a basis in the contravariant space, such that these bases are reciprocal (see e.g. Spiegel, [1959]). In tensor analysis, this basis would be introduced and characterized by a tensor which is called the associated metric tensor. From this it is possible to obtain the contravariant components of a vector from its covariant components by the formula:

$$l^S = g^{rs} l_r \quad (4.8)$$

where  $g^{rs}$  is the associated metric tensor (a two times contravariant tensor).

It can be shown that the once contracted product of the metric and the associated metric tensors equals to the Kronecker delta (see e.g. Vaníček, [1972]). This is another illustration of the duality of the two space.

#### 4.4 Transformation of Vectors

In (4.2) it was mentioned that in a three dimensional rectangular coordinate system, for instance, it is also possible to show coordinates of a point with respect to an orthogonal curvilinear coordinate system.

Let us assume a three dimensional space with a rectangular Cartesian coordinate system  $X$  defined in it. We say there is a curvilinear system of coordinates  $U$  defined in the space if and only if there are three functions

$$u^i = u^i(x^1, x^2, x^3) \quad i = 1, 2, 3 \quad (4.9)$$

of the rectangular Cartesian coordinates defined and if these functions can be inverted to give

$$x^i = x^i(u^1, u^2, u^3), \quad i = 1, 2, 3. \quad (4.10)$$

The necessary and sufficient condition for the two sets of functions to be reversible is that the functions in (4.9) and (4.10) be single-valued and have continuous derivatives; in other words, both determinants of transformation

$$\det \left( \frac{\partial u^j}{\partial x^i} \right), \quad \det \left( \frac{\partial x^j}{\partial u^i} \right)$$

must be different from zero almost everywhere.

If the above conditions are satisfied, it is possible to obtain the covariant components of a vector with respect to the curvilinear system if its covariant components relative to the X system are known, using the following formula,

$$\bar{a}_i = \frac{\partial u^j}{\partial x^i} a_j \quad (4.11)$$

( $\bar{a}_i$  and  $a_j$  are the covariant components of a vector with respect to curvilinear and Cartesian coordinate systems respectively).

It can be shown that the same rule can be used in obtaining the covariant (contravariant) components of a vector relative to a coordinate system, if the same components of that vector relative to another coordinate system are known, when none of these systems are Cartesian. The following formulae show the rule of transformation in such cases:

$$\begin{aligned} \bar{a}_j &= \frac{\partial u^i}{\partial \bar{u}^j} a_i, & \bar{a}^j &= \frac{\partial \bar{u}^j}{\partial u^i} a^i \\ a_j &= \frac{\partial \bar{u}^i}{\partial u^j} \bar{a}_i, & a^j &= \frac{\partial u^j}{\partial \bar{u}^i} \bar{a}^i \end{aligned} \quad (4.12)$$

## CHAPTER 5

### DIAGRAMS FOR THE LEAST-SQUARES ADJUSTMENT PROBLEMS

#### Introduction

The concepts covered in the previous chapters serve as a pre-requisite to the principal topic of this thesis, which we will be dealing with from now on. These concepts will be used in constructing diagrams for the parametric, conditional and combined cases in the least-squares adjustment problems.

#### 5.1 Diagram for a Simple Parametric Case

Let  $L$  and  $X$  be real vector spaces of dimensions  $n$  and  $u < n$  respectively. Also, suppose  $A$  is a linear mapping of  $X$  onto  $L$ . As we know the linear operators on the finite dimensional vector spaces can be represented in terms of matrices. The linear mapping  $A$  is therefore a matrix which we suppose has the rank  $u$  ( $A$  has been assumed to have full rank).

By taking a set of vectors as basis in  $L$ , we can introduce a metric in this space which is a positive definite and symmetric tensor. Thus  $L$  space will be a metric space. We refer to this metric as  $P_L$ . Similarly, let us assume that the vector space  $X$  is metricized by a metric tensor  $P_X$ . We assume these vector spaces

are spaces of the covariant vectors, which according to the above remarks and assumptions are related by following formula

$$l = Ax, \quad l \in L, \quad x \in X \quad (5.1)$$

By introducing these metric tensors in X and L spaces and according to the previous description of dual spaces, dual bases, covariant and contravariant spaces and their relations, we can define the dual bases corresponding to X and L. These spaces according to the above assumptions, will be contravariant vector spaces X and L, with the elements which would be obtained from the following relations

$$(l \in L, l' = P_1 l) \longrightarrow l' \in L', \quad (x \in X, x' = P_x x) \longrightarrow x' \in X'. \quad (5.2)$$

As we saw in (4.1) both  $P_1$  and  $P_x$  are two times contravariant tensors, and, from (2.4.2, theorem 2), we know that the dual spaces are of the same dimensions as the primal spaces. Formula (5.1) is nothing else but a case of formula (4.12); that is

$$\bar{a}_j = \frac{\partial u^i}{\partial \bar{u}^j} a_i$$
$$a_j = \frac{\partial \bar{u}^{-i}}{\partial u^j} \bar{a}_i$$

which shows the rules of transformation between the two covariant vector spaces. In these transformations  $(\frac{\partial u^i}{\partial \bar{u}^j})$  and  $(\frac{\partial \bar{u}^{-i}}{\partial u^j})$  are Jacobian matrices. Also in (4.12), we had the case of transformation between two contravariant vector spaces as



$$a^{-j} = \frac{\partial u^{-j}}{\partial u^i} a^i$$

$$a^j = \frac{\partial u^j}{\partial u^i} a^i$$

By comparing these two sets of formulae we conclude: If the elements of two covariant vector spaces are transformed by, say,  $\frac{\partial u^i}{\partial u^{-j}}$ , then the elements of corresponding contravariant vector spaces will be transformed by  $\frac{\partial u^{-j}}{\partial u^i}$ , where  $\frac{\partial u^i}{\partial u^{-j}}$  shows the elements of a matrix and  $\frac{\partial u^{-j}}{\partial u^i}$  is its transpose, which will simply be shown by A and A<sup>t</sup>. Therefore, we have

$$x' = A^t l' . \quad (5.3)$$

Accordingly, there is a set of vectors in each of the contravariant spaces X' and L' which constitute the bases in them. The metric tensors in X' and L' are two times covariant.

Algebraically, we can say, because P<sub>l</sub> and P<sub>x</sub> are positive-definite and symmetric matrices, they have regular inverses P<sub>l</sub><sup>-1</sup>, P<sub>x</sub><sup>-1</sup> which are the metric tensors in the contravariant spaces (in our case).

With such background material we can now construct a diagram (see Figure 5.1), from which the various relations between the above mentioned elements and appurtenances of those covariant and contravariant spaces can be easily obtained as shown below.

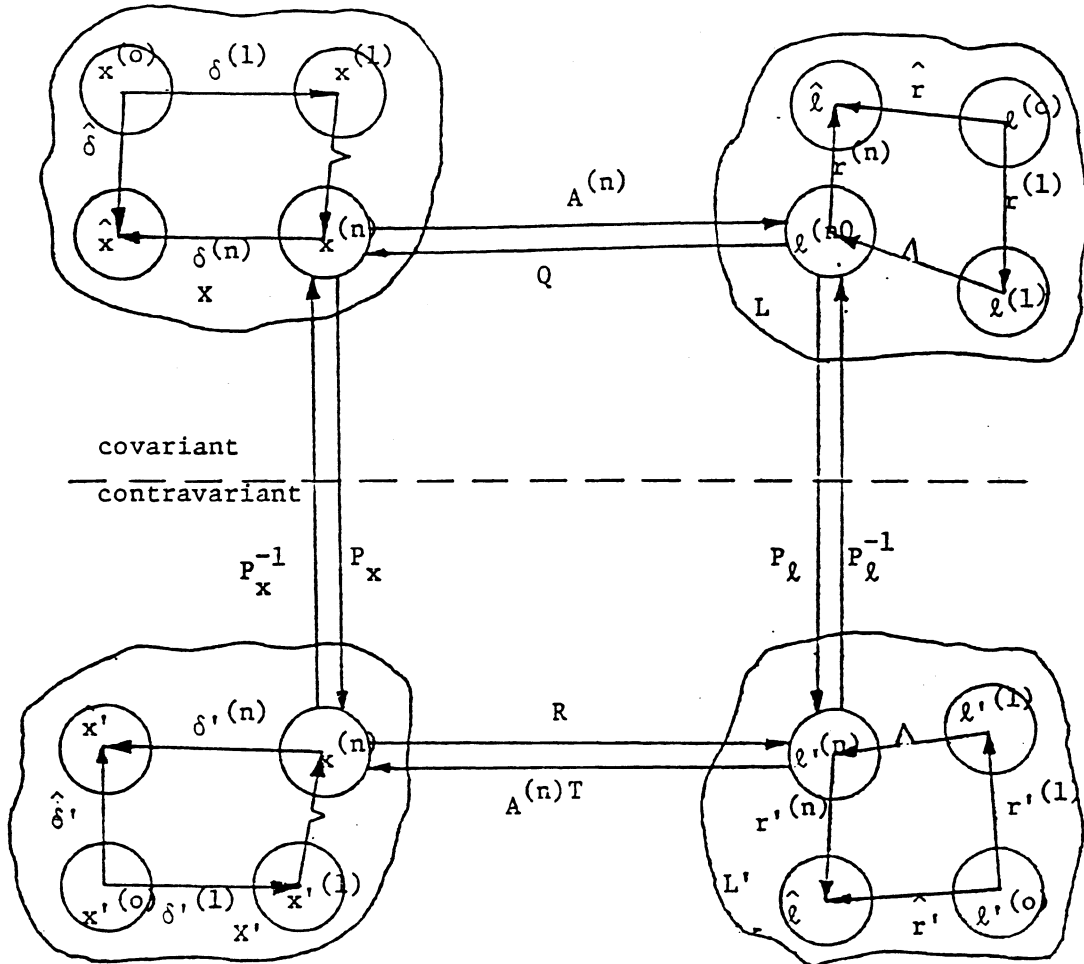


Figure 5.1 Diagram for the Parametric Least-Squares Adjustment Case

Assuming that only  $A$  and  $P_1$  are known:

(i) Starting from  $X$  and proceeding towards  $X'$  in both directions, we get

$$P_x = A^T P_1 A \quad (5.4)$$

This formula may be called the rule for metric induction.

(ii) To get  $Q$  it is enough to start from  $L$  and proceed towards  $X$  in both directions

$$Q = P_x^{-1} A^T P_1 = (A^T P_1 A)^{-1} A^T P_1 \quad (5.5)$$

Note that by earlier assumptions, since  $A$  is of full rank, (Rank ( $A$ )  $u$ ),  $P_x$  has a unique inverse and  $Q$  therefore is unique.

(iii) The inverse transformation can be derived equally easily

$$R = P_1 A P_x^{-1} \quad (5.6)$$

(iv) It is easy to show that  $P_1^{-1}$  and  $P_x^{-1}$  are related by the following formula,

$$P_1^{-1} = A P_x^{-1} A^T \quad (5.7)$$

If the covariant metric tensor,  $P_1^{-1}$  is selected to equal the covariance matrix ( $C_1$ ) and  $P_x^{-1}$  to  $C_x$ , we have from (5.7)

$$C_1 = A C_x A^T \quad (5.8)$$

the known covariance law. Similarly equation (5.5) yields:

$$x = (A^T C_1^{-1} A)^{-1} A^T C_1^{-1} l \quad (5.9)$$

the solution of least-squares normal equations. In addition, it can be easily verified that,

$$R = Q^T \quad (5.10)$$

Note that in this diagram the iterative process, which was mentioned at the end of Chapter 3 is shown. The least-squares technique is applied within each iteration.

It should be noted that, in order to obtain the matrices (i.e. arrows in the diagram) using the diagram, one has to start from the tail point of the arrow representing that matrix. Then going around the diagram and following the directions which lead him to the pointed end of arrow. For example, in order to obtain the unknown matrix  $P_x$  we have to start from matrix A and end with  $A^T$  i.e.

$$P_x = \dots A$$

$$P_x = \dots P_l A$$

$$P_x = A^T P_l A$$

The elements of the spaces which are unknown to us can be obtained as follows:

Starting from the element of a known space and going around the diagram using the matrices (i.e. arrows) which transform this element to the intermediate spaces. For example in order to obtain the unknown parameter  $x \in X$  we start from the known element  $l \in L$ , i.e.

$$x = \dots l$$

$$x = \dots P_l l$$

$$x = \dots A^T P_l l$$

$$x = P_x^{-1} A^T P_l l$$

## 5.2 Diagram for a Conditional Adjustment Model

As mentioned in Chapter 3, if in the linearized model the first design matrix  $A$  is equal to 0, we call it a conditional model; that is,

$$Br + w = 0 \quad (5.11)$$

In fact, the above formula is obtained after the linearization of a non-linear equation, which explains the relationship of  $n$  observations, that is,

$$F(l) = 0 . \quad (5.12)$$

This model can be linearized by Taylor series expansion and yields

$$F(l) = F(l^{(0)}) + \left. \frac{\partial F}{\partial l} \right|_{l=l^{(0)}} (l-l^{(0)}) + \dots, \quad (5.13)$$

[see e.g. Vaníček, 1971]. The equations (5.13) can thus be taken as the differential forms of equation (5.12).

All the relations reviewed in Chapter 3 about the combined adjustments, are valid for a conditional adjustment case, when the first design matrix is taken as a zero matrix, and we can obtain them by the following diagram (Figure 5.2).

In this case we assume the observation space to be metrized by  $P_1$  which is a positive definite and symmetric matrix; also

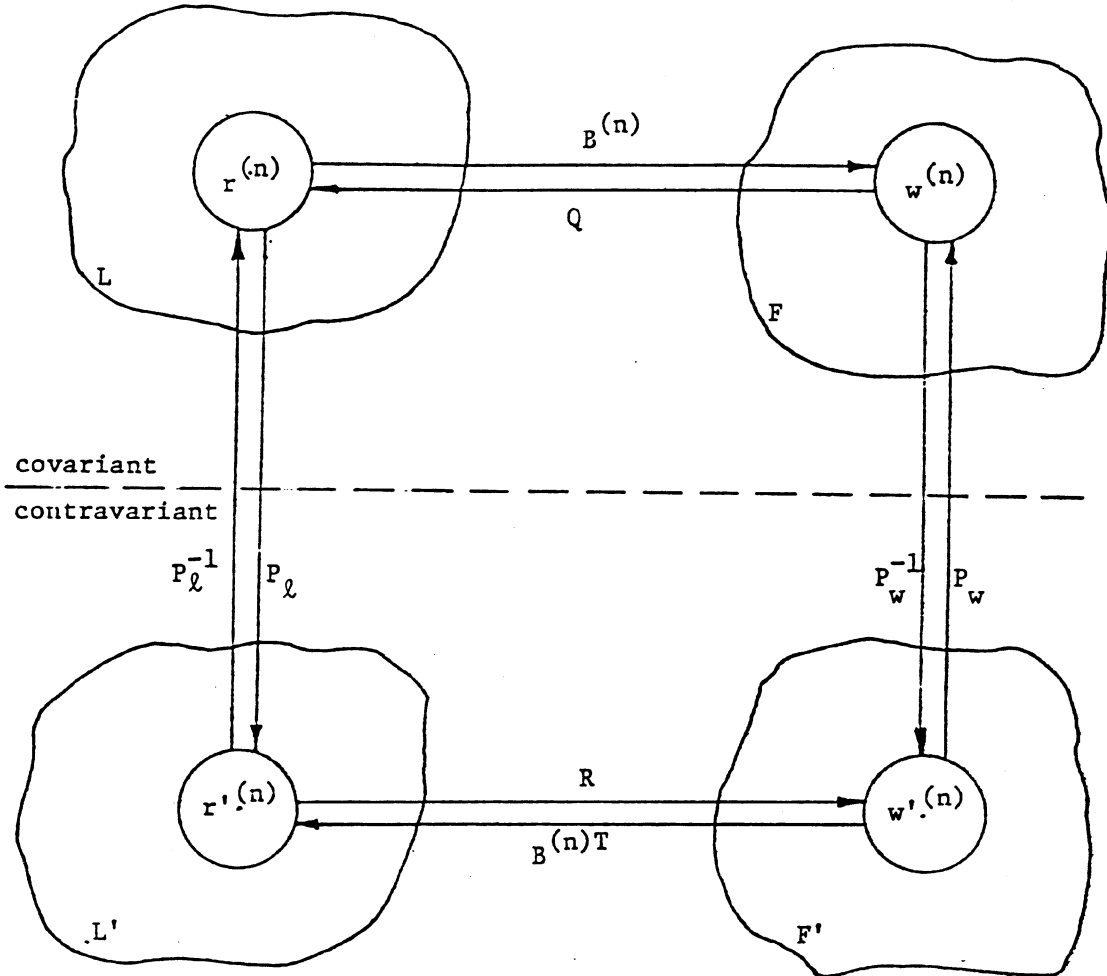


Figure 5.2. Diagram for a Conditional Least-Squares Adjustment Case

B has been assumed to have full rank.

The tensor  $P_w$  is another positive definite and symmetric matrix which has been taken as the metric of the model space.

Let L and F be covariant vector spaces and B be a linear mapping of L onto F. Therefore, by the same reasoning as in the previous section, their contravariant counterparts L' and F' are connected by  $B^T$  and are of the same dimensions as L and X respectively.

Let B and  $P_1$  be known, then the elements of diagram (5.2) can be obtained as follows:

(i) Starting from F' and proceeding towards F, we get

$$P_w^{-1} = B P_1^{-1} B^T \quad (5.14)$$

Remember that since B is of full rank,  $P_w$  can be uniquely obtained from

$$P_w = (B P_1^{-1} B^T)^{-1} \quad (5.15)$$

(ii) Starting from F and proceeding towards L, we get

$$Q = P_1^{-1} B^T P_w = P_1^{-1} B^T (B P_1^{-1} B^T)^{-1} \quad (5.16)$$

By taking the covariant metric tensor  $P_1^{-1}$  equal to the covariance matrix of the observation  $C_1$ , we get

$$Q = C_1 B^T (B C_1 B^T)^{-1} \quad (5.17)$$

(iii) The formula (5.17) can be applied in obtaining the solution in the conditional adjustment,

$$l = Qw, \quad l \in L, \quad w \in F \quad (5.18)$$

or,

$$l = C_1 B^T (B C_1 B^T)^{-1} w. \quad (5.19)$$

Also in formula (5.14), if we take the covariant metric tensor,  $P_w^{-1}$ , equal to the covariance matrix of the misclosures,  $C_w$ , then,

$$C_w = B C_1 B^T. \quad (5.20)$$

The formulae (5.19) and (5.20), obtained easily by using the diagram, are known to us from the least-squares adjustment [see Vaníček, 1971].

### 5.3 Diagram for a Combined Adjustment Model

As mentioned before, parametric and conditional cases in the classification of the least-squares adjustment, are two particular cases of the combined case which was reviewed in Chapter 3. There we saw that the implicit equation  $f(x, l) = 0$ , after linearization, will have the form

$$A\delta + Br + w = 0 \text{ or } A\delta + Br = -w,$$

This shows the relation between the three differential neighbourhoods of three corresponding vectors in parameter, observation and model vector spaces.

From Chapter 3 we recall that the solution vector  $\delta$  can be obtained even if the solution is done in the model space. To obtain  $r$ , however, the solution must be done in the observation space. Here again we assume that linear mappings  $A$  and  $B$  are full rank matrices,



also  $P_1$  is taken as a positive definite symmetric tensor in  $L$ , where  $L$ ,  $X$  and  $F$  are covariant vector spaces.

The same reasoning as before can be applied to construct a diagram for this case (see Figure 5.3)). The least-squares method is applied within each stage. The matrices  $A$  and  $B$  represent the values of first and second design matrices in the  $n$ th iteration, corresponding to points  $x$ ,  $l$  and  $w$  in the above mentioned spaces.

Suppose that  $l$  and  $P_1$  are known. Then other elements of the diagram can be obtained by the following procedure:

(i) Starting from  $F'$  and proceeding towards  $F$  via  $L$  and  $L'$ , we get

$$P_w^{-1} = B P_1^{-1} B^T \quad (5.21)$$

Because matrix  $B$  has been assumed to have full rank,  $P_w$  is uniquely determined by

$$P_w = (B P_1^{-1} B^T)^{-1} \quad (5.22)$$

(ii) From  $X$  to  $X'$  via  $F$  and  $F'$ ,  $P_x$  is determined as

$$P_x = A^T P_w A = A^T (B P_1^{-1} B^T)^{-1} A. \quad (5.23)$$

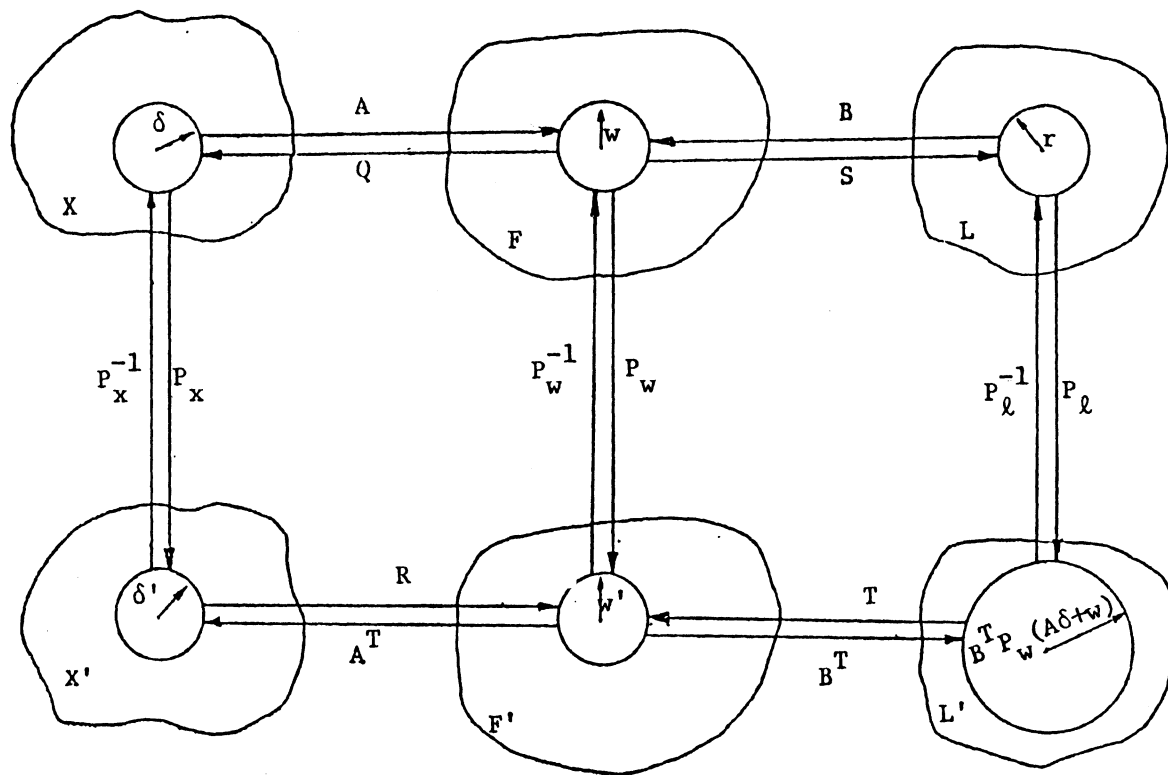


Figure 5.3. Diagram for a Combined Least-Squares Adjustment Case

(iii) Starting from F and proceeding towards X, we get

$$Q = P_x^{-1} A^T P_w = [A^T (BP_1^{-1} B^T)^{-1} A]^{-1} A^T (BP_\ell^{-1} B^T)^{-1} \quad (5.24)$$

which after substitution in the equation

$$\begin{aligned} \hat{\delta} &= Qw, \quad \hat{\delta} \in X, \quad w \in F, \\ \text{we obtain} \quad \hat{\delta} &= [A^T (BP_\ell^{-1} B^T)^{-1} A]^{-1} A^T (BP_\ell^{-1} B^T)^{-1} w \end{aligned} \quad (5.25)$$

If the covariant metric tensor  $P_1^{-1}$  is taken equal to the covariance matrix of the observations  $C_1$  or  $C_r$ , then formula (5.26) reduces to formula (3.17) in Chapter 3.

(iv) As stated, to get the residuals, solution must be done in the observation space. To clarify this fact, a differential distance in  $L'$  and its value has been shown and written in the diagram.

From the diagram, we get

$$S = P_1^{-1} B^T P_w = P_1^{-1} B^T (BP_1^{-1} B^T)^{-1} \quad (5.27)$$

and

$$r = S(w + \hat{\delta}), \quad \text{where, } r \in L, \quad w \in F \quad (5.28)$$

After substitution of S from (5.27) into (5.28) we get

$$\hat{r} = -P_1^{-1} B^T (BP_1^{-1} B^T)^{-1} (w + \hat{\delta}), \quad (5.29)$$

which after substitution  $\bar{P}_1^{-1} = C_1$  will reduce to formula (3.22)

5:4 Some Remarks to the above Diagrams

In the previous sections of this chapter and in the construction of the diagrams some concepts from the tensor operations were used. However, it is worth noting that in [Tienstra, 1956], the elements of the metric tensor have been used in the standardization of some normally distributed observation series. This is nothing else but the metricization of the observation space by the inverse of the covariance matrix of the observation which was discussed above. Consequently, some tensor algebra has been used there for the least-squares adjustment. For example a linear parametric case,

$$AY = x + w , \quad (5.31)$$

has been expressed as

$$a_{\lambda}^{i} Y^{\lambda} = f^i + \eta^i \quad (5.32)$$

and its normal equation,

$$A^T P A Y = A^T P w, \quad \text{by} \quad (5.33)$$

$$g_{ik} a_{\lambda}^{i} a_{\Gamma}^{k} Y^{\lambda} = g_{ik} a_{\Gamma}^{k} f^i \quad (5.34)$$

which shows,

$$g_{ik} = P = C_r^{-1} ,$$

where  $g_{ik}$  is the metric tensor [see Tienstra, 1956, pp. 152].

#### 5.4.1 Singularity and the diagrams

The linear mappings  $A$  and  $B$ , which were considered in the previous sections of this chapter, were assumed to be of full rank. In the case of a rank deficient matrix  $A$  or  $B$ , a unique inverse can no longer be obtained for  $P_x$  or  $C_w$  in equations (5.4), (5.20), and (5.23). In such a case the generalized inverse can be applied [Rao and Mitra, 1971]. This was suggested in [Vanicek, 1979]. Here the case of a singular weight matrix is taken into consideration, to see what this means geometrically.

From linear algebra, we know that a linear mapping

$$F: V \longrightarrow U,$$

is said to be singular if the image of some nonzero vector under  $F$  is  $0$ , i.e., if there exists  $v \neq 0$  in  $V$  with  $F(v) = 0$ . Thus  $F$  is nonsingular if and only if  $0 \in v$  (not other vectors) maps into  $0 \in U$ , or equivalently, if its null space consists only of the zero vector:  $N(F) = \{0\}$ . Now, if the linear mapping  $F$  is one-to-one, then only  $0 \in V$  can map into  $0 \in U$  and so  $F$  is nonsingular. The converse is also true. It can be easily verified that by a nonsingular mapping, independent sets would transform to the independent sets [see e.g. Lipschutz, 1974].

In relation to the subject of this section, the following theorem also is useful.

Theorem: Let  $V$  be a finite dimensional vector space and let

$F: V \longrightarrow U$  be a linear mapping. Then,

$$\dim (V) = \dim (N(F)) + \dim (R(F))$$

(dim stands for dimension).

This relation according to the following equations

$$\text{rank } (F) = \dim (R(F)), \text{ Nullity } (F) = \dim (N(F)),$$

can be written as

$$\dim (V) = \text{rank } (F) + \text{Nullity } (F).$$

In the case of a nonsingular  $F$ , the above reduces to,

$$\dim (V) = \text{rank } (F).$$

In the case of a singular weight matrix for the observations in constructing the diagrams we will in fact be dealing with a set of dependent vectors which cannot be taken as a basis in space  $L$ . If  $l_1, l_2, \dots, l_n$  is such a set, then, there is at least one nonzero element among a set of scalars  $\{a_1, a_2, \dots, a_n\}$  such that

$$a_1 l_1 + a_2 l_2 + \dots + a_n l_n = 0.$$

According to the above remarks, when the weight matrix  $P_1$  in  $L'$  is singular, we can conclude that the dimension of  $L'$ , dual space of  $L$ , as the range of  $P_1$ , is of lower dimension than  $L$ . This really does not make sense, because as mentioned in (2.4.2, theorem 2), dual space of a finite dimensional vector space, is of the same dimension as the primal space. Geometrically in this case we can say that the observation vectors are of dimension  $n$ , but the space  $L$  spanned by them is of lower dimension  $m$ . The singular weight matrix is a map from a space of dimension  $n$  onto a space  $m$ , so cannot give a map from  $L$  ( $\dim m$ ) to  $L'$ . Thus

it cannot be used to define a metric on  $L$ . This case corresponds to a coordinate system of higher dimension being used in space of lower dimension.

Finally, it should be noted that Figures (5.1) and (5.2) can be considered as a diagrammatic method for finding the reflective least-squares  $g$ -inverse for a singular matrix. Before proving the above, the following definitions are necessary.

Definition 1. Let  $A$  be a real  $m \times n$  matrix. A matrix  $P_A$  is called a projection operator onto the space generated by the columns of  $A$ , with respect to a non negative definite matrix  $M$ , if

a)  $P_A^T M P_A = M P_A$  (idempotent)

b)  $M P_A A = M A$

c)  $\text{rank}(P_A) = \text{rank}(A)$

If  $M = I$  (identity matrix), then conditions (a) and (b) reduce to

a)  $P_A^2 = P_A$

b)  $P_A A = A$

Definition 2. A  $g$ -inverse of a matrix  $A$  of order  $n \times m$  is a matrix  $\bar{A}$  of order  $m \times n$ , such that  $A \bar{A}$  is idempotent and  $\text{rank}(A \bar{A}) = \text{rank}(A)$ . or, alternatively,  $\bar{A} A$  is idempotent and  $\text{rank}(\bar{A} A) = \text{rank}(A)$ .

Theorem 1. As mentioned in Chapter 1, for an inconsistent equation  $Ax = y$ ,  $\hat{x}$  is a least-squares solution if

$$\|A\hat{x} - y\| = \inf_x \|Ax - y\|$$

Now, if  $Q$  is a matrix (not necessary a g-inverse) such that  $Qy$  is a least-squares solution of  $Ax = y$ , then it is necessary and sufficient that

$$AQ = P_A$$

A g-inverse which provides a least-squares solution of  $Ax = y$  is referred to as a least-squares g-inverse.

Definition 3. If the inverse of the inverse of a matrix is equal to the original matrix, then it will be called a reflexive inverse. A necessary and sufficient condition for a g-inverse  $G$  of  $A$  to be reflexive is that:

$$\text{rank}(G) = \text{rank}(A) .$$

A reflexive g-inverse providing a least-squares solution is denoted by  $\bar{A}_{1r}$

Theorem 2. Let  $\|y\| = (y^T M y)^{1/2}$ , where  $M$  is positive-definite; then

$$\bar{A}_{1r} = (A^T M A)^{-1} A^T M .$$

For the above definitions and theorems see [Rao & Mitra, 1971].

With respect to the above definitions and considering Figure (5.1), it can be seen that  $Q$  is a reflexive least-squares g-inverse of the singular matrix  $A$ , verified as follows:

(a) Matrix  $QA$  is idempotent since,

$$\begin{aligned} (QA)^2 x &= ((A^T P_1 A)^{-1} A^T P_1 A) (A^T P_1 A)^{-1} A^T P_1 A x = \\ &= (A^T P_1 A)^{-1} A^T P_1 A x = QA = x . \end{aligned}$$



(b)  $\text{rank } (QA) = \text{rank } (Q)$ .

(c)  $\text{rank } (Q) = \text{rank } (A)$ .

(d) Matrix  $QA$  is a projection operator which projects the elements  $l \in L$  onto the subspace spanned by the Columns of  $A$ .

This follows since the columns of  $A$  generate a subspace in  $L$  space, and matrix  $Q$  projects the elements  $l \in L$  onto this subspace. Indeed, the solution space ( $X$  space) is such a subspace, and the solution vector,  $\hat{x}$ , is the orthogonal projection of the observation vector onto it. In addition to the above reasoning, it can be directly concluded, by taking  $M = P$  in theorem 2, that  $Q$  is a reflexive least-squares  $g$ -inverse for matrix  $A$ .

## CHAPTER 6

### DIAGRAM FOR LEAST-SQUARES COLLOCATION

#### 6.1 Two-Component Adjustment and Collocation

In Chapter 3 the formulation of least-squares adjustment was reviewed. The corresponding diagrams for three cases of adjustment, i.e., parametric, conditional and combined adjustments were presented in Chapter 5. In Chapter 3 we assumed that the observation space is formed only of statistically independent observation vectors. Practically, however, we are dealing with observations which are composed of two components, i.e., statistically dependent and statistically independent components. According to the relation  $C_1 = C_r$ , we can say in the case of a two component observation, the residual  $r$  is also composed of two component, statistically dependent ( $s$ ) and statistically independent ( $v$ ), or

$$r = v + s. \quad (6.1)$$

The independent component  $v$ , often originates within the measuring apparatus and we can link it to a measuring error, or resolution capability, and thus internal to the instrument.

The statistically dependent residual  $s$ , however, is thought of as being external to the instrument and related to the behaviour of the observable in a particular milieu.

The presence of  $s$  in a mathematical model, for example in following parametric model

$$l = Ax + v + s \quad (6.2)$$

(which in ordinary adjustment is presented as,

$$l = Ax + v \quad [\text{see e.g. Vaníček, 1971}] \quad (6.3)$$

can be interpreted as the shortcoming in the mathematical model; that is, the "inability" of the model to describe completely the actual relationship among the measurements  $l$  and unknown parameters  $x$ .

In equation (6.2)  $v$  is called noise and  $s$  can be called either signal or noise, depending to whether it is wanted or not. If  $s$  is treated as noise, then, along with the statistically independent residual  $v$ , it can be separated from the signal  $x$  by ordinary adjustment procedures. On the other hand, if  $s$  is treated as a signal, it must be separated from the noise  $v$  and from signal  $x$ , and the procedure is called two-component adjustment.

The case of the least-squares collocation can easily be visualized as follows:

In the parametric model (6.2), we have to determine the curve, shown as solid line (see Figure 6.1), by means of discrete observations (small crosses), which are furthermore affected by observation errors  $v$ . The curve to be reconstructed consists of a systematic part,  $Ax$ , and a random part  $s$ , both of which are of interest.

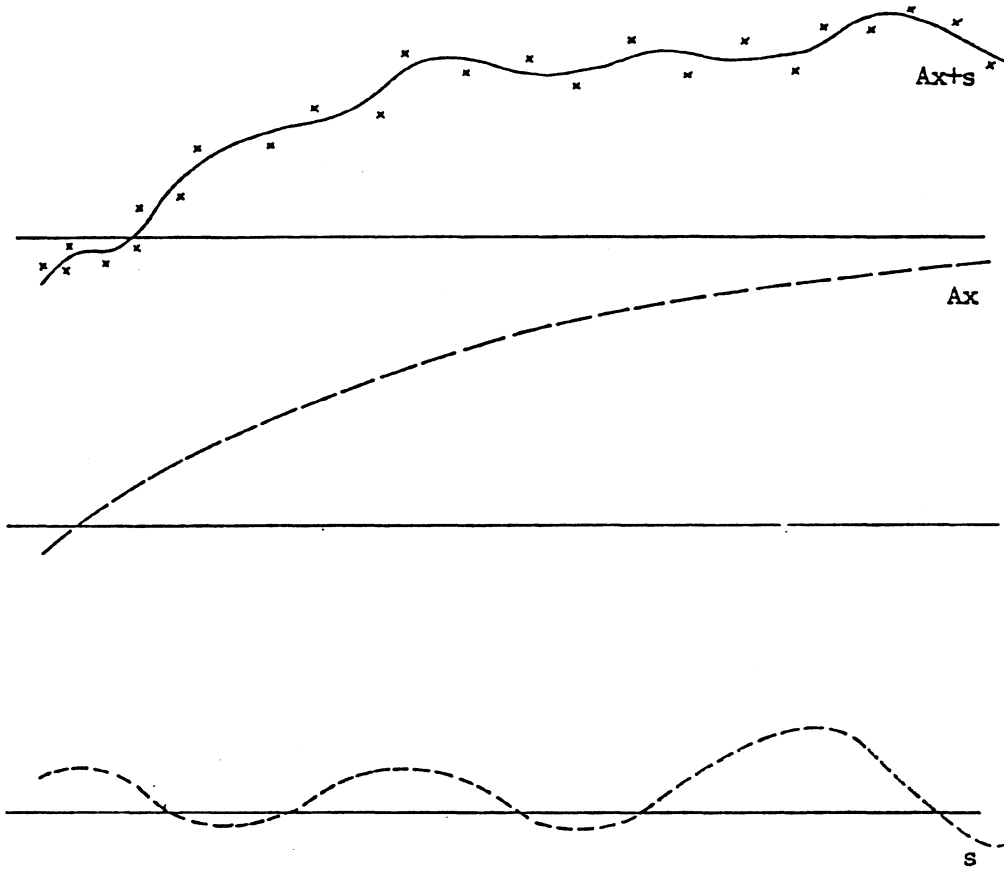


Figure 6.1. Illustration of the Least-Squares Collocation in a Parametric Model

If we consider the determination of the parameter  $x$  as adjustment, the removal of the noise  $v$  as filtering, and the computation of  $s$  at points other than the observation points as prediction, least-squares collocation is a combination of all these processes [see Moritz, 1972; Balmino, 1978; Krakiwsky, 1975].

Because mathematical model and the formulation of collocation is similar to that of the two component adjustment, these cases will be reviewed. For this purpose, the two following sections have been abstracted from [Vaníček and Krakiwsky, 1981].

## 6.2 Two Component Adjustment Mathematical Model

In the two component adjustment the implicit mathematical model is of the form

$$f(x, l) = f(x, l + s + v) = 0 \quad (6.4)$$

with two covariance matrices  $C_s$  and  $C_v$  for the residuals. The cross-covariance matrix  $C_{sv}$  will be assumed to be equal to zero. The differential form of the model can be written as

$$A\delta + Bs + Bv + w = 0 \quad (6.5)$$

In the more general case, the component  $s$  is not from the observation space  $L$  but from another "space of statistically dependent observations"  $S$ . Therefore, for the differential environment of the observation point  $s$ , the following transformation holds (see figure 6.2).

$$T \in \{S \rightarrow L\} \quad (6.6)$$

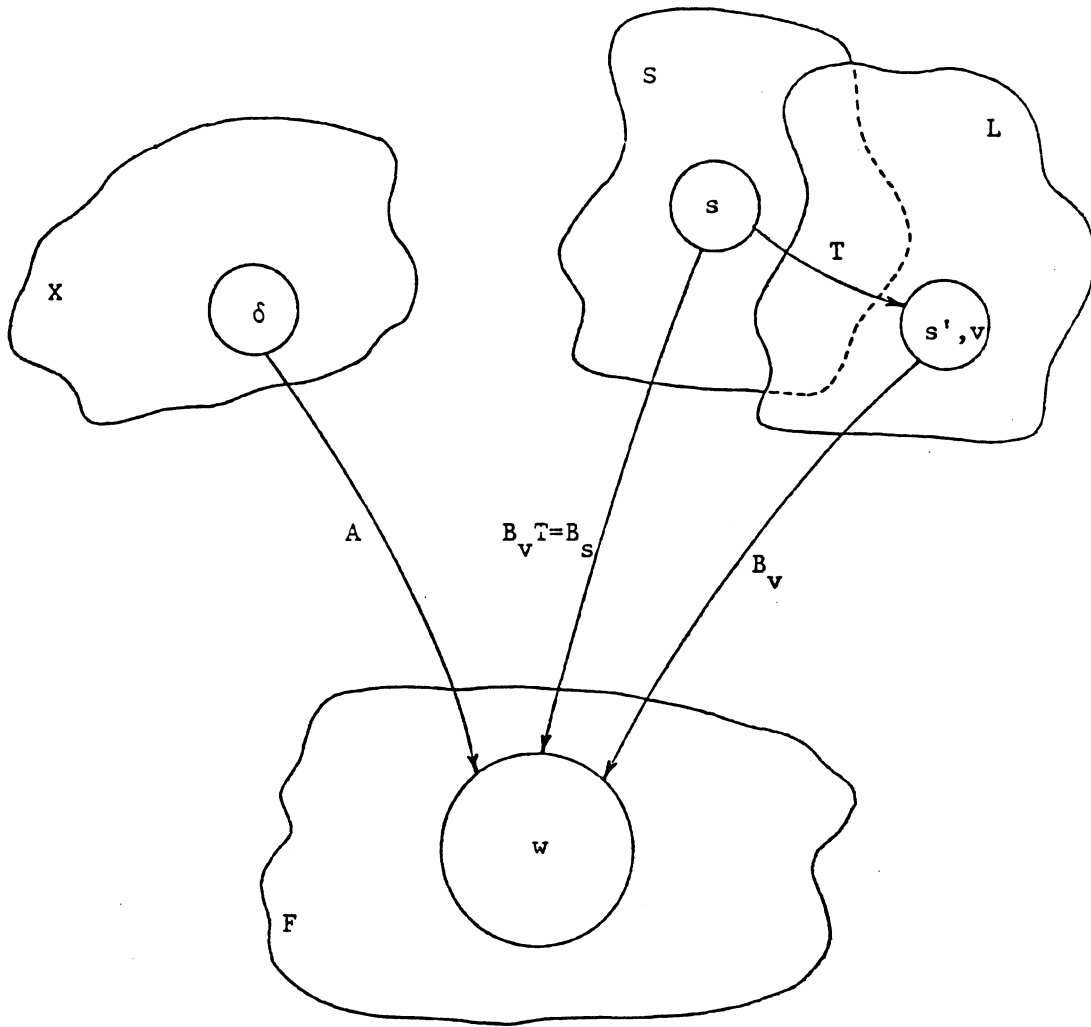


Figure 6.2. Space Involved in the Two Component Adjustment

Then we have

$$s' = Ts \quad (6.7)$$

where S may or may not have the same dimension as L. By substituting (6.7) into (6.5) and denoting B by  $B_v$  and BT by  $B_s$ , one gets a more general two component linear model

$$A\delta + B_v v + B_s s + w = 0 \quad (6.8)$$

### 6.3 Formulation of the Two Component Adjustment

The variation function from which the normal equation system follows, reads

$$\phi = s^T C_s^{-1} s + v^T C_v^{-1} v + 2k^T (A\delta + B_s s + B_v v + w) \quad (6.9)$$

As we can see, both vectors of residuals play a role in the quadratic form, along with their respective covariance matrices.

All the formulae reviewed in Chapter 3 for ordinary adjustment can be applied here, by replacing  $C_1 = C_r$  by  $C'_r$  and B by  $B'$  - where

$$C'_r = \left[ \begin{array}{c|c} C_s & 0 \\ \hline 0 & C_v \end{array} \right] \quad (6.10)$$

and

$$B' = [B_s \mid B_v]. \quad (6.11)$$

Realizing that  $\hat{r}' = \begin{bmatrix} \hat{s} \\ \hat{v} \end{bmatrix}$  one obtains, from equation (3.22)

and

$$\hat{r}' = \begin{bmatrix} \hat{s} \\ \hat{v} \end{bmatrix} = - \left[ \begin{array}{c|c} C_s & 0 \\ \hline 0 & C_v \end{array} \right] \begin{bmatrix} B_s^T \\ B_v \end{bmatrix} \hat{k} \quad (6.12)$$

and

$$s = -C_s B_s^t L w, \quad (6.13)$$

$$\hat{v} = -C_v B_v^t L w \quad (6.14)$$

where,

$$L = (M A N^{-1} A^t M - M) \text{ (see (3.24)).}$$

Equations (6.10) and (6.11) directly affect M (see Chapter 2), changing its inverse to

$$M^{-1} = [B_s \mid B_v] \begin{bmatrix} C_s & 0 \\ \hline 0 & C_v \end{bmatrix} \begin{bmatrix} B_s^T \\ \hline B_v^T \end{bmatrix} = (B_s C_s B_s^T + B_v C_v B_v^T). \quad (6.15)$$

#### 6.4 Formulation of the Least-Squares Collocation

As was mentioned earlier in this chapter, and also according to the above formulation, the statistically dependent component of the residual, as a noise, can be separated from the statistically independent component of the residual and from the signal  $x$  (or  $\delta$ ) by a two component adjustment technique. What distinguishes the least-squares collocation from two component adjustment is that it views  $s$  as predictable, and thus one has two kinds of signal,  $x$  and  $s$ .

The first signal  $x$  does not naturally lend itself to prediction because  $x$  is defined on a parameter space  $X$ , usually reflecting some specific observables, and so it does not make any physical sense to predict  $x$  elsewhere. On the other hand,  $s$  is



often thought of as being only a sample of an effect that can be modelled in a wider space P than just the observation space S.

Thus, a prediction denoted by  $s_p$ , of the statistically dependent signal  $s$  is sought in a prediction space P such that  $S \subset P$  (see Figure 6.3). Mathematically, the prediction feature can be built into the two component adjustment by expanding the second design matrix to B as follows

$$B'' = [0 \mid B_s \mid B_v], \quad (6.16)$$

and taking

$$\hat{r}'' = [s_p^T \mid s^T \mid v^T]^T \quad (6.17)$$

The prediction is really made possible by the stipulation that the stochastic characteristics of the new quantity  $s_p$  are the same as those of  $s$ . This is expressed through an expanded covariance matrix C

$$C'_s = \begin{bmatrix} C_{s_p} & C_{s_p s} \\ C_{s s_p} & C_s \end{bmatrix} \quad (6.18)$$

which describes the covariance between  $s$  at the observation points and  $s_p$  at the prediction points. Here,  $C_{s s_p}^T$  clearly equals  $C_{s_p s}$  and preserves the required symmetry of  $C'_s$ . It is the covariance among all the signal components which mediates the prediction, by relating  $s$  to  $s_p$ .

All the equations for the two component adjustment remain valid when the expanded versions of B (equation (6.16)) and  $\hat{r}$  (equation (6.17)) are used.

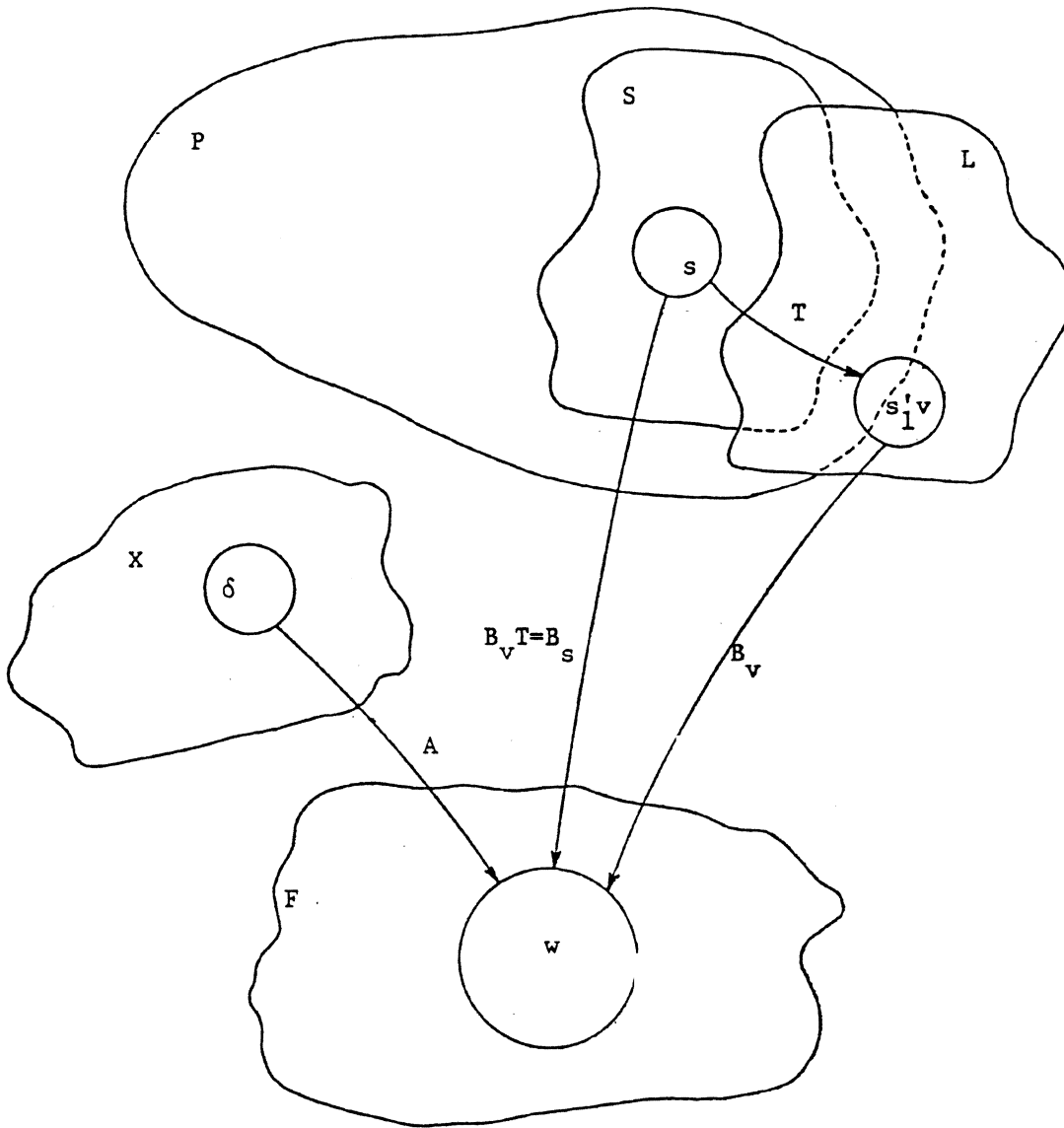


Figure 6.3. Spaces Involved in the Collocation

Changes only occur to equation (6.13) which, upon substitution of  $C'_s$  for  $C_s$  yields the predicted signal.

To derive the formula for the predicted signal, we substitute:

$$C''_r = \begin{bmatrix} C_{s_p} & C_{s_p s} & 0 \\ C_{ss_p} & C_s & 0 \\ 0 & 0 & C_v \end{bmatrix} \quad (6.19)$$

and

$$B = [0 \mid B_s \mid B_v] \quad (6.20)$$

for  $C_r$  and  $B$ , respectively, in formula (3.23).

These substitutions yield:

$$\begin{aligned} \hat{r}'' &= \begin{bmatrix} \hat{s}_p \\ \hat{s} \\ \hat{v} \end{bmatrix} = \begin{bmatrix} C_{s_p} & C_{s_p s} & 0 \\ C_{ss_p} & C_s & 0 \\ 0 & 0 & C_v \end{bmatrix} \begin{bmatrix} 0 \\ B_s^T \\ B_v^T \end{bmatrix} Lw = \\ &= \begin{bmatrix} C_{s_p} & C_{s_p s} & 0 \\ C_{ss_p} & C_s & 0 \\ 0 & 0 & C_v \end{bmatrix} \begin{bmatrix} 0 \\ B_s^T Lw \\ B_v^T Lw \end{bmatrix} \end{aligned} \quad (6.21)$$

or,

$$\hat{r}'' = \begin{bmatrix} \hat{s}_p \\ \hat{s} \\ \hat{v} \end{bmatrix} = \begin{bmatrix} C_{ss_p} B_s^T Lw \\ C_s B_s^T Lw \\ C_v B_v^T Lw \end{bmatrix} \quad (6.22)$$

Therefore,

$$\hat{s}_p = -C_{s_p s_p} B_s^T L w, \quad (6.23)$$

$$\hat{s} = -C_s B_s^T L w, \quad (6.24)$$

$$\hat{v} = -C_v B_v^T L w. \quad (6.25)$$

Also, it can easily be verified that

$$C_{s_p s_p} = C_{s_p s_p} B_s^T L B_s C_{s s_p} \quad (6.26)$$

(see Vanicek and Krakiwsky, [1981]). It should be noted that the formulae (6.23), (6.24) and (6.25) can be written as

$$\hat{s}_p = -C_{s_p s_p} B_s^T M (A\hat{\delta} + w) \quad (6.27)$$

$$\hat{s} = -C_s B_s^T M (A\hat{\delta} + w) \quad (6.28)$$

$$\hat{v} = -C_v B_v^T M (A\hat{\delta} + w) \quad (6.29)$$

This will become clear if we compare the equations (3.22) and (3.23).

Note that in both formulations (the two component adjustment and least-squares collocation) the covariance matrices of the results could be obtained, but because there is no need for them in this work, these formulae were not reviewed here.

### 6.5 Diagram For The Least-Squares Collocation

In this section a diagram for the least-squares collocation is presented. This diagram is for a combined case linear model;

$$A\hat{\delta} + B_s\hat{s} + B_v\hat{v} + w = 0 \quad (6.30)$$

(see Figure 6.4).

In drawing this diagram, the same mathematical structure, as was already described in the previous chapter for ordinary adjustment models, has been applied. This time, however, instead of  $C_r$  as a metric for the contravariant observation space and also instead of matrix  $B$ , we are dealing with  $C_r''$  and  $B''$ , (equations (6.19) and (6.20) respectively).

The main idea for drawing these matrices came from "graph theory", where the differential neighbourhoods of some corresponding points in both covariant and contravariant statistically dependent, statistically independent and prediction spaces serve as vertices, and covariances serve as edges (see Deo, [1974]).

The cross-covariances  $C_{s_p s}$  and  $C_{ss_p}$  for the reasons which are explained in the next section of this chapter, have been used as the media for relating the prediction and statistically dependent observation spaces.

Recall that  $s_p$  is a sample of an effect in prediction space  $P$ ; in other words, a differential neighbourhood of a point  $p$ , in  $P$ , manifests itself to the observation space via these cross-covariances, which are supposed to be known.

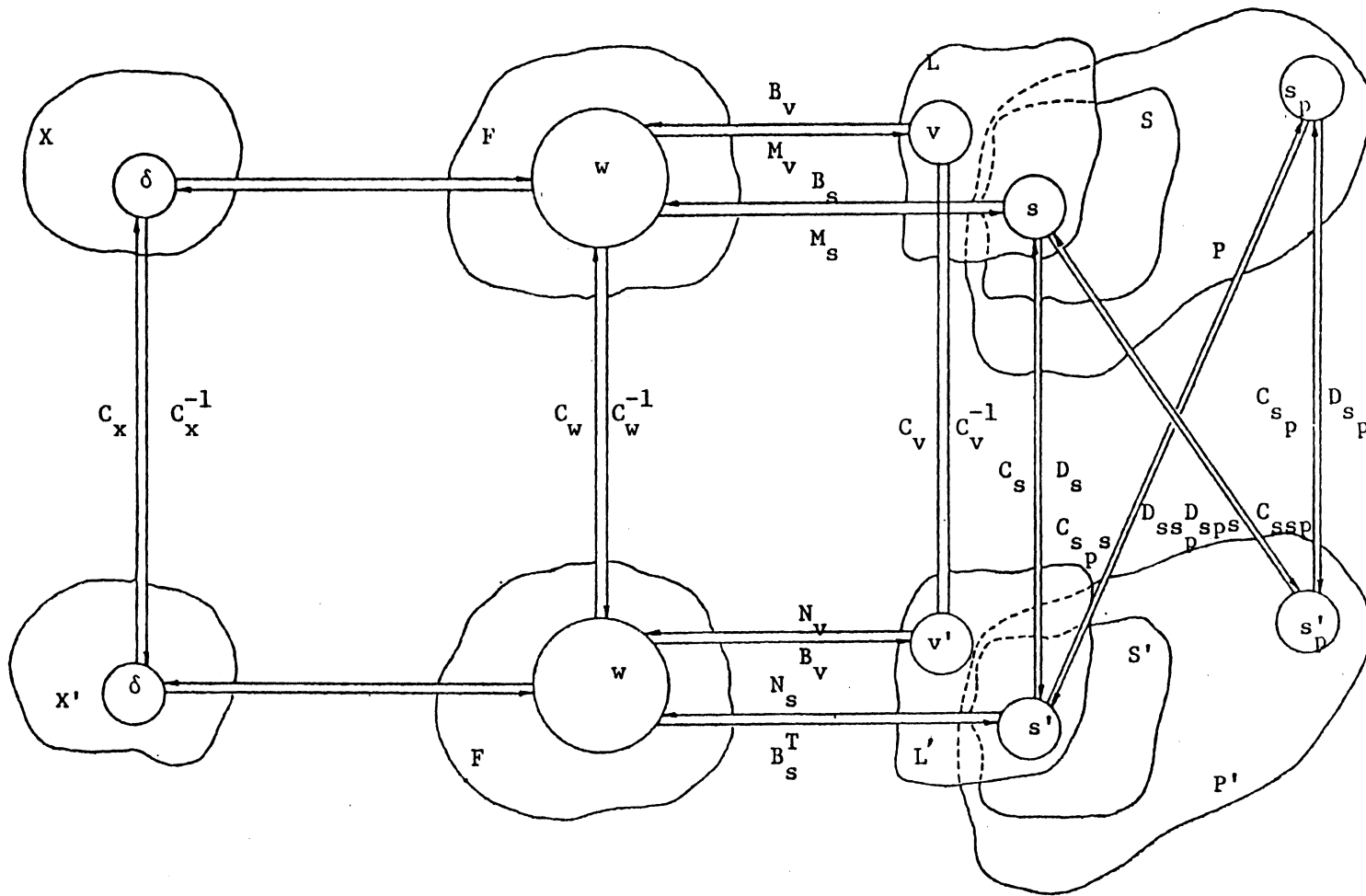


Figure 6.4. Diagram for the Least-Squares Collocation

It should be noted here that cross-covariances  $C_{s_p s}$  and  $C_{ss_p}$  play the same role in the connection between covariant and contravariant counterparts in the graph.

Note that two elements  $v$  and  $s$  in the covariant part of this diagram and two elements  $v'$  and  $s'$  in its contravariant counterpart, are the components of two residuals,  $r$  and  $r'$ , respectively. (In this diagram, for the reasons which are explained in the next section, some arrows from the covariant space to the contravariant space have been denoted by notation  $D$ ).

### 6.6 Obtaining the Least-Squares Collocation Formulae from the Graph

The formulae (6.27), (6.28) and (6.29) as well as those described in Chapter 5, for the determination of the unknown parameters  $x$ , can be obtained by this diagram.

Note that to obtain an element of prediction space  $P$  as an unknown, one has to start from an arrow which connects that element to an element of a known space, which is the statistically dependent observation space.

To get  $S_p$  we have to follow the following steps:

- (1) Start from  $F$  by  $C_w^{-1} = M$  and proceed towards  $F'$
- (2) From  $F'$  to  $S'$  by  $B_s^T$
- (3) From  $S'$  to  $P$  by  $C_{s_p s}$  we get:

$$\hat{s}_p = -C_{s_p s} B_s^T M (\hat{A} + w) . \quad (6.31)$$

6.7 Some Remarks About the Mathematical Structure of the Least-Squares Collocation

As we can see from the formula (6.12), as long as there are no cross-covariance components in the covariance matrix of the observations, we cannot predict anything. That is, in such a case we cannot have any point of prediction space, since the prediction space is a null space. Where cross covariances in the covariance matrix exist, we can predict some elements of the prediction space. In other words, the prediction in such a case is not a null space.

Mathematically, we can say that if we multiply the elements of contravariant space  $S'$ , see Figure (6.4), by its metric tensor  $C_s$  we obtain the elements of the covariant space  $S$ . Now if we multiply the elements of  $S'$  by something else, say  $C_{s_p}$  we would definitely obtain a new space, which we call a prediction space. This space can be metricized by its own covariance matrix. These, somewhat show the role of the cross-covariance elements of the covariance matrix of the observation space.

To clarify the problem let us have a closer look at the formulation of the least-squares collocation and the geometric meaning of  $C_r''$  in formula (6.19).

For the ordinary adjustment case we are dealing with an observation space. In two component adjustment and collocation, however, we are dealing with two and three spaces, respectively. These spaces are statistically independent, statistically dependent and prediction spaces. In each of these spaces a set of vectors



is selected as a basis or coordinate axes.

Together, these spaces constitute a unique space. This space can be considered as a three dimensional space, if we imagine each set of base vectors or coordinate axes in those three spaces as a coordinate axis in this three dimensional space. We take this space as a contravariant space, which can be metricized by  $C''_r$  (6.19). All elements of this matrix serve for the transformation of vectors in contravariant space to the covariant space.

With respect to the above remarks, and the form of  $C''_r$  as a metric tensor, we can imagine that the contravariant space can be characterized by a three dimensional coordinate system, with two inclined coordinate axes and the third axis orthogonal to the plane formed by the two inclined axes. These axes are labelled  $X_1, X_2$  and  $X_3$  in Figure 6.5.

A vector  $r'$  in the contravariant space has components on each of the three axes and can be written as:

$$r' = \begin{bmatrix} s'_p \\ s'_s \\ v' \end{bmatrix}. \quad (6.32)$$

The covariant counterpart of  $r$  is obtained by

$$r = \begin{bmatrix} s_p \\ s \\ v \end{bmatrix} = \begin{bmatrix} C_s & C_{s_p s} & 0 \\ C_{ssp} & C_s & 0 \\ 0 & 0 & C_v \end{bmatrix} \begin{bmatrix} s'_p \\ s'_s \\ v' \end{bmatrix} = \begin{bmatrix} C_{s_p s} s'_p + C_s s'_s \\ C_{ssp} s'_p + C_s s'_s \\ C_v v' \end{bmatrix}. \quad (6.33)$$

In this formula we can see the effect of the presence of off-diagonal elements in the metric tensor.

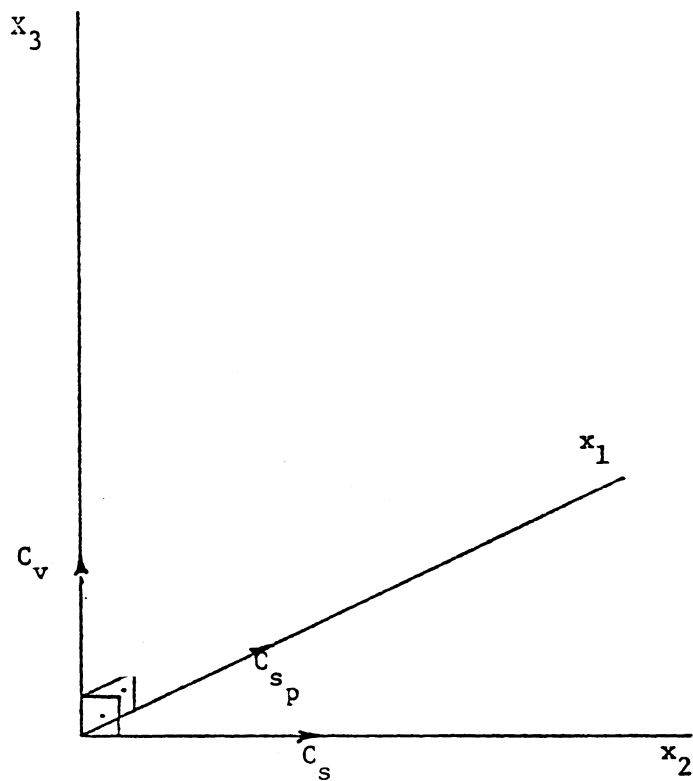


Figure 6.5. Illustration of the Coordinate System in the Observation and Prediction Spaces

According to Figure (6.5) and formula (6.33), the covariant components of a vector in the  $X_2X_3$  plane can be obtained by

$$r = \begin{bmatrix} s_p \\ s \\ v \end{bmatrix} = \begin{bmatrix} C_{s_p} & C_{ss_p} & 0 \\ C_{ss_p} & C_s & 0 \\ 0 & 0 & C_v \end{bmatrix} \begin{bmatrix} 0 \\ s' \\ v' \end{bmatrix} = \begin{bmatrix} C_{s_p s'} \\ C_s s' \\ C_v v' \end{bmatrix} \quad (6.34)$$

The inclined coordinate axes result in three non zero covariant components, even when only two of the three contravariant components of a vector are nonzero.

This is the case in the least squares collocation problem, where we do not have any idea about the prediction space and we want to get some information about that space from the observation space. Recall that even though for the solution of the least-squares collocation problems we have to have the covariance matrix of prediction space ( $C_{s_p}$ ), it is not present in the equation for the prediction signal  $s_p$ .

Therefore, mathematically, in a collocation problem we are always looking for the covariant components of the vectors whose contravariant components lay in the  $X_2X_3$  plane. This is clearer if we compare formulae (6.21), (6.33) and (6.34).

With respect to the above remarks and explanations, we can conclude that the presence of the off-diagonal elements in the metric tensor causes the image of a point in one of the component subspaces not to be completely into the corresponding subspace of

the dual. This is clear from Figure 6.4, where, for a point in the contravariant space with coordinates  $(v', s', 0)$  we have a corresponding point in the covariant space with coordinates  $(v, s, s)$ .

It follows algebraically that, unlike the case of a diagonal metric tensor where the elements of the inverse of this metric tensor are the inverse of each of the individual elements of the primal metric tensor, the presence of the off-diagonal elements in the metric tensor makes the individual elements of the inverse of the metric tensor to not correspond to the inverse of the individual elements of the primal metric tensor. Therefore, if  $C''_I$  in formula (6.19) is the metric tensor for the contravariant vector space (see Figure 6.4) then the metric tensor of the corresponding covariant vector space is of the following form:

$$\begin{bmatrix} C_{s_p} & C_{ss_p} & 0 \\ C_{s_p s} & C_s & 0 \\ 0 & 0 & C_v \end{bmatrix}^{-1} = \begin{bmatrix} D_{s_p} & D_{ss_p} & 0 \\ D_{s_p s} & D_s & 0 \\ 0 & 0 & C_v^{-1} \end{bmatrix}$$

## CHAPTER 7

### CONCLUSION

Some related concepts of functional analysis and ordinary least-squares adjustment have been reviewed.

By considering the original diagram for a simple parametric model from [Vaníček, 1979], diagrams for the least-squares adjustment, in the conditional and combined cases, and least-squares collocation have been constructed. These diagrams were constructed on the basis of the parallelism between the concept of a covariance matrix and a metric tensor and also on the basis of duality property of covariant and contravariant spaces, as mentioned in [Vanicek, 1979]. In this thesis, however, the mathematical structure of the diagrams from both functional analysis and tensor operations points of view have been thoroughly explained.

It is well known that the singular weight matrix for the observation space is a consequence of the dependence of some of the observations on the others. In this thesis, the observations have been considered as a system of coordinates in the observation space. The case of a singular weight matrix has been geometrically interpreted, and it is concluded that in such a case the dimension of the observation space is less than the number of observations. Mathematically, this obstacle can be

removed by reducing the number of the observations.

It has been proven that the diagrammatic approach can be considered as a method for obtaining the reflexive least-squares  $g$ -inverse of a singular matrix.

In the construction of the diagram for the least-squares collocation, in addition to the tensor structure of least-squares equation, which were used for the adjustment diagrams, some concepts from the Graph theory have been applied.

The value of this method lies in the ability to obtain with extreme ease the least-squares adjustment equations through an illustrative approach. This property makes this method more appealing than the standard least-squares adjustment and collocation. Therefore, it is recommended that the diagrammatic approach, as described, form part of the content of advanced least-squares courses.

REFERENCES

- Adam, J. (1980). Least-Squares In Hilbert Spaces, Paper Presented at the 4th International Symposium "Geodesy and Physics of the Earth", GDP, Karl-Marx-stadt.
- Balmino, G. (1977). Introduction to Least-Squares Collocation, Lecture delivered at the Second International Summer School in the Mountains on Mathematical Methods in Physical Geodesy, Ramsau, Austria.
- Deo, N. (1974). Graph Theory with Application to Engineering and Computer Science, Prentice-Hall.
- Halmos, R. (1958). Finite-Dimensional Vector Spaces, Van Nostrand Reinhold.
- Hotine, M. (1969). Mathematical Geodesy ESSA Monograph 2, U.S. Department of Commerce, Government Printing Office, Washington, D.C.
- Krakiwsky, E.J. (1975). A Synthesis of Recent Advances in the Method of Least-Squares, Lecture Notes No. 42, Department of Surveying Engineering, University of New Brunswick, N.B., Canada.
- Kreyszig, E. (1978). Introductory Functional Analysis With Applications, Wiley & Sons.
- Lipschutz, S. (1974). Theory and Problems of Linear Algebra, Schaum's Outline, McGraw-Hill.
- Luenberger, D.G. (1969). Optimization By Vector Space Method, Wiley & Sons.
- Meissl, P. (1977). Hilbert Spaces and Their Application to Geodetic Least-Squares Problems, 2nd International Summer School Ramsau, Austria.
- McConnell, A.J. (1931). Application of Tensor Analysis, Dover Reprint, (1957).
- Mikhail, E.M. (1976). Observation and Least-Squares, IEP - A Dun-Donneley.
- Pope, A.J. (1974). Two Approaches to Nonlinear Least-Squares Adjustments, Paper Presented at the International Symposium on Problems Related to the Redefinition of North American Geodetic Network held on the Campus of the University of

New Brunswick, reprinted from The Canadian Surveyor,  
Vol. 28, No. 5.

- Rao, C.R. & S.K. Mitra (1971). Generalization Inverse Matrix And Its Applications, Wiley & Sons.
- Spiegel, M.R. (1959). Vector Analysis and Introduction To Tensor Analysis, Schaum s outline, McGraw-Hill.
- Fienstra, P. (1956). Theory of the Adjustment of Normally Distributed Observations, edited by his friends Uitgeverij, N.V., Argus, Amsterdam.
- Vaníček, P. (1971). Introduction to Adjustment Calculus, Lecture Notes No. 20, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B., Canada.
- Vaníček, P. (1972). Tensors, Lecture Notes No. 27, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B., Canada.
- Vaníček, P. (1979). Tensor Structure and The Least-Squares. Bull. Geod. 53 (1979), pp. 221-225.
- Vaníček, P. and E.J. Krakiwsky (1981). Geodesy: The Concepts, North-Holland.
- Wells, D.E. (1974). Doppler Satellite Control, Technical Report No. 29, Department of Surveying Engineering, University of New Brunswick, Fredericton, N.B., Canada.