A General Framework f~~F~~or Error Analysis i~~I~~n Measurement-B~~b~~ased GIS, Part 1: The Basic Measurement-Error Model a~~A~~nd Related Concepts

Yee Leung
Department of Geography and Resource Management, Center for Environmental Policy
and Resource Management, and Joint Laboratory for Geoinformation Science,
The Chinese University of Hong Kong, Hong Kong
E-mail: yeeleung@cuhk.edu.hk

Jiang-Hong Ma
Faculty of Science, Xi'an Jiaotong University and Chang'an University, Xi'an, P.R. China
E-mail: jhmath@pub.xaonline.com

Michael F. Goodchild
Department of Geography, University of California, Santa Barbara, California, U.S.A.
E-mail: good@ncgia.ucsb.edu

Abstract. This is the first part of our four-part series of papers which proposes a general framework for error analysis in measurement-based geographical information systems (MBGIS). The purpose of the series is to investigate the fundamental issues involved in measurement error (ME) analysis in MBGIS, and to provide a unified and effective treatment of errors and their propagations in various interrelated GIS and spatial operations. Part 1 deals with the formulation of the basic ME model together with the law of error propagation. Part 2 investigates the classic point-in-polygon problem under ME. Continued onto Part 3 is the analysis of ME in intersections and polygon overlays. In Part 4, error analyses in length and area measurements are made. In the present part, a simple but general model for ME in MBGIS is introduced. An approximate law of error propagation is then formulated. A simple, unified, and effective treatment of error bands for a line segment is made under the name of "covariance-based error band". A new concept, called "maximal allowable limit", which guarantees invariance in topology or geometric-property of a polygon under ME is also advanced. To handle errors in indirect measurements, a geodetic model for MBGIS is proposed and its error distribution problem is studied on the basis of the basic ME model as well as the approximate law of error propagation. Simulation experiments all substantiate the effectiveness of the proposed theoretical construct.

Comment [v1]: "Propagation" may be more suitable

Keywords: Covariance-based error band, error propagation, geodetic model, geographical information systems, maximal allowable limit, measurement error

1. Introduction

Geographical information science in general and geographical information system (GIS) in particular have developed into a major discipline that fosters theoretical investigations and practical tools for the management, analysis and display of spatial information. With the ever increasing volume of geo-referenced data being generated, transferred, and utilized, the amount of uncertainty embedded in spatial databases has become a major issue of crucial theoretical importance and practical consideration. It is apparent that the oblivious use of error-laden spatial data without considering the intrinsic uncertainty involved will lead to serious consequences in concepts and practices.

Uncertainty (in attribute values and in positions) in spatial databases generally involves accuracy, statistical precision, and bias in initial values or in the estimated coefficients in statistically calibrated equations. Most importantly, spatial uncertainty includes the estimation of errors (both in position and in attribute) in the final output that result from the propagation of external (initial value) uncertainty and internal (model) uncertainty.

Therefore, research on uncertainty in spatial data intends to investigate how uncertainties arise and distribute through GIS operations, and to assess the plausible effects on subsequent decision-making. It is thus important to be able to track the occurrence and propagation of uncertainties (Goodchild, 1991; UCGIS, 1996). Apparently, map accuracy is closely related to spatial uncertainty, and its assessment and provision are essential to end users (Keefer et al., 1991). Research on accuracy must be associated with errors in GIS. Thus the study of errors in GIS has been extensive and diverse (see for example, Goodchild and Gopal, 1989; Leung and Yan, 1998; Mowrer and Congalton, 2000; Stanislawski et al., 1996; Veregin, 1989; Wolf and Ghilani, 1997; Zhang and Goodchild, 2002). In a broader context, a study in spatial statistical analysis is given ~~in~~ by Cressie (1993). The error taxonomy recognizes that different classes of spatial data exhibit different types of errors, and errors may be introduced and propagated in various stages of data manipulation and spatial processing.

Errors in spatial databases are complex and multivariate. They can generally be reckoned as the *inherent error* and the *operational error*. Inherent error is the error present in source documents, including the accumulated error in the map used as input to a GIS. Operational error, categorized as positional and identification errors, is produced through the data capture and manipulation functions of a GIS or is introduced during the process of data entry and occurs throughout data manipulation and spatial modeling. Digitizing, for example, is one mechanism of data input that introduces error into maps. Its magnitude and impact upon digital map accuracy has not been well studied (Chrisman, 1982). In fact, it is often overlooked or assumed to be negligible.

Since GIS databases, at the most general level, are based on a model of geographical data, errors can be classified into spatial, temporal and thematic error. From the modeling point of view, they can also be classified as either systematic or random. Systematic errors usually follow physical laws and they can be removed by model modification. It is however impossible to avoid random errors in measurements entirely (Wolf and Ghilani, 1997). Such random error is also called measurement error (ME), which is one of the most important problems in the use of geo-referenced data.

It should be noted that a feature of conventional vector-based data is the representation of position by derived coordinates, rather than by original measurements. In such coordinate-based GIS, an important problem is the determination of error structures of location coordinates. It is a basis of error analysis. If it ~~is unable to~~cannot be solved, it will then be impossible to perform error propagation or to estimate uncertainties in derived products. To overcome the difficulties in uncertainty management intrinsic to conventional GIS, the concept of a measurement-based GIS (MBGIS) has been proposed ~~in~~ by Goodchild (1999). Under this concept, a MBGIS is a system that provides access to measurements used to determine the locations of objects, to the geographical procedures (transformation functions) that link measurements to quantities to be measured, and to the rules used to determine interpolated positions. It also provides access to the locations which may either be stored, or derived on the fly. The basic idea is to retain details of measurements so that error analysis can be made possible, and corrections to positions can be appropriately propagated through the database.

To make MBGIS a reality, it is essential to develop a general framework within which a rigorous approach, particularly the statistical approach, to measurement error analysis and error propagation can be formulated. By extending~~Extended on~~ the theory of ME (Neuilly and CETAMA, 1999) and its analysis in the context of GIS (Heuvelink, 1998; Heuvelink et al., 1989), we propose in the present study a formal basis for estimating how errors in~~ a~~ location can be propagated through various GIS operations, and how errors in the corresponding products can be estimated.

Our research work is organized into four parts (see Fig. 1.1). The purpose of this four-part series of papers is to explore the fundamental issues involved in error analysis in MBGIS, and to render a consistent and effective treatment of errors and their propagations in a variety of interrelated GIS and spatial operations. We formulate in Part 1 a basic measurement error model on which formal analysis of error propagation can be made, and ME models for various GIS operations can be constructed. On the basis of the basic ME model, we also scrutinize the concepts of error band for a line segment, and propose a new error-band construct, ~~and~~ together with a concept of maximal allowable limit for positional error. To be able to analyze geodetic problems, a geodetic model for MBGIS is also constructed and substantiated by numerical examples. In Part 2, we investigate the classic point-in-polygon problem with a new perspective. We propose conditions under which the problem should be addressed and re-open the issue of the error band for a line segment with a new view~~-~~point. To make point-in-polygon analysis under ME operable, we formulate an algebra-based probability model that can compute probabilities by circumventing the difficulties and complexities of the geometric methods that deal with polygons directly. In Part 3, we examine the problems surrounding line-in-polygon and polygon-on-polygon overlays. We first formulate an approximate law for error propagation in intersection coordinates which forms the basis for the analysis of error propagation in polygon overlays. To further extend the scope of our analysis, we analyze in Part 4 the relative errors in length and area measurements in MBGIS. In addition to the substantiation by numerical examples, we derive the approximate law of error propagation in length measurements and the exact law of error propagation in area measurements.

Comment [v2]: we have not done this

In what remains in this part of the four-part series, we first formulate the basic measurement error model in MBGIS, and examine the concept of the error band for a line segment in Section 2. A geodetic model for MBGIS is then proposed in Section 3, with a substantiation ~~by~~ through three applications. We then conclude Part 1 with a summary and a prelude to Part 2.

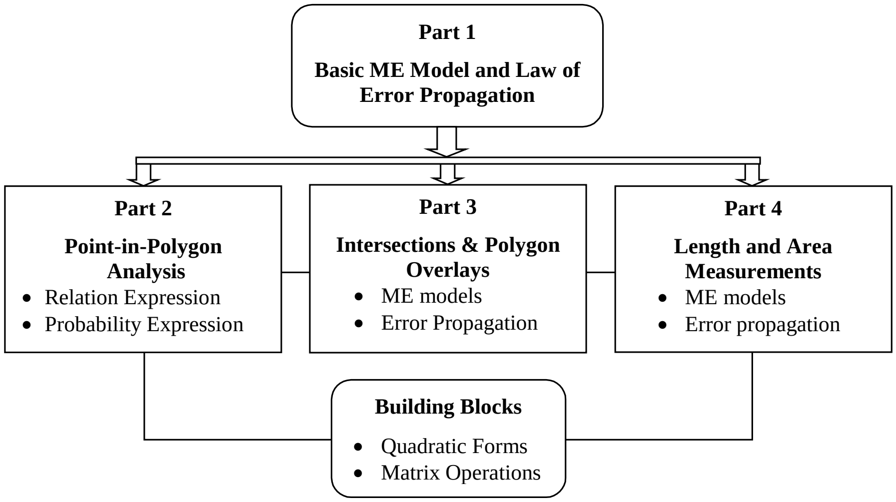


Fig. 1.1 A summary of the research plan

2. Basic Model for Measurement Error (ME) in MBGIS

In this section, we formulate a basic ME model in MBGIS. It is a general model for error analysis and propagation in vector-based databases. Our discussion is restricted to spatial objects in the two-dimensional problem.

2.1 Basic ME model

Measurements are defined as observations made to determine unknown quantities. They may basically be classified as either direct or indirect. Direct measurements are made by applying an instrument (or geometric techniques) directly to the unknown quantity and observing its value, e.g., by reading it directly from graduated scale on the device~~, its value~~. Examples of direct measurement in GI~~P~~S are determining the distance between two points by making a direct measurement using a graduated tape, or measuring an angle by making a direct observation from the graduated circle of a theodolite or total station instrument.

The ME model for direct measurement can usually be expressed as

$$\mathbf{x} = \boldsymbol{\mu} + \boldsymbol{\varepsilon}, \tag{2.1}$$

where $\mathbf{x}$ is the observed (measurement) value vector, $\boldsymbol{\mu}$ is its true value vector and $\boldsymbol{\varepsilon}$ the vector of errors in measurement, called the ME vector henceforth. In GIS, $\boldsymbol{\mu}$ is usually a true coordinate representation of a location, while $\mathbf{x}$ is the measured coordinates of the location. Indirect measurements are obtained when it is difficult or impossible to make direct measurements. Under such a situation, the quantity desired is determined from its mathematical (functional) relationship to direct measurements. During this procedure, the errors that were present in the original direct observations are distributed by the computational process into the indirect values. As a result, indirect measurements contain errors that are functions of the original errors. This distribution of errors from direct measurements is called *error propagation* (Wolf and Ghilani, 1997).

Interpolation of locations of spatial objects can be viewed as indirect measurement. A location $\mathbf{y}$ may be interpolated between measured locations. Indirect measurements occur, for example, when the position of some feature recognizable on an aerial photograph is established with respect to registered tics or control points. They also occur when a surveyor establishes the location of a boundary by linking surveyed monuments with a mathematical straight line.

In general, we can eliminate systematic errors which follow physical laws. So, it is reasonable to assume that only random errors remain in the direct observations. In MBGIS, we essentially need to derive the basic propagation equation for random errors.

Let $\mathbf{x} \in R^p$ be the set of p measurements which can be measured directly, $\mathbf{y} \in R^q$ be q required quantities to be measured, and f (a well-defined geographical procedure, called the *operation function*, or, more generally, the *transformation function*) be the function linking these measurements $\mathbf{x}$ to the quantities $\mathbf{y}$ so that

$$\mathbf{y} = f(\mathbf{x}).$$

Once $\mathbf{x}$ is measured and f is known, $\mathbf{y}$ and the associated error distribution are uniquely and automatically determined. The inverse of f is denoted by f^{-1}, that is, the function that allows measurements $\mathbf{x}$ to be determined from the quantities $\mathbf{y}$. This reciprocal relationship between $\mathbf{x}$ and $\mathbf{y}$ will prove to be instrumental in error analysis and database maintenance. The basic ME model for indirect measurement can simply be expressed as

$$(\mathrm{I}): \quad \begin{cases} \mathbf{Y} = f(\mathbf{X}), & (2.2) \\ \mathbf{X} = \boldsymbol{\mu}_x + \boldsymbol{\varepsilon}_x, \quad \boldsymbol{\varepsilon}_x \sim (\mathbf{0}, \boldsymbol{\Sigma}_x), & (2.3) \end{cases}$$

Formatted

where $\boldsymbol{\mu}_x$ is the true-value vector, $\mathbf{X}$ the random measurement-value vector, $\mathbf{Y}$ the indirect measurement-value vector obtained by f, and $\boldsymbol{\varepsilon}_x$ the random ME vector with zero mean $\mathbf{0}$ and the variance-covariance matrix $\boldsymbol{\Sigma}_x$. To simplify, the variance-covariance matrix will be called the covariance matrix henceforth. According to (2.2), $\mathbf{Y}$ is random and its error covariance matrix $\boldsymbol{\Sigma}_y$ is propagated from $\boldsymbol{\Sigma}_x$. However, it should be noted that in general $\boldsymbol{\Sigma}_y$ also depends on the true-value vector $\boldsymbol{\mu}_x$. To see it, model (I) can be expressed as

$$\mathbf{Y} = f(\mathbf{X}) = f(\boldsymbol{\mu}_x + \boldsymbol{\varepsilon}_x) = E[f(\boldsymbol{\mu}_x + \boldsymbol{\varepsilon}_x)] + \boldsymbol{\varepsilon}_y(\boldsymbol{\mu}_x + \boldsymbol{\varepsilon}_x) = \boldsymbol{\mu}_y + \boldsymbol{\varepsilon}_y ,$$

where $\boldsymbol{\mu}_y \equiv E[f(\boldsymbol{\mu}_x + \boldsymbol{\varepsilon}_x)]$ is dependent on $\boldsymbol{\varepsilon}_x$ through the expectation, and $\boldsymbol{\varepsilon}_y \equiv \boldsymbol{\varepsilon}_y(\boldsymbol{\mu}_x + \boldsymbol{\varepsilon}_x)$ is dependent on $\boldsymbol{\varepsilon}_x$ through a function $\boldsymbol{\varepsilon}_y(\cdot)$. Therefore, the dependence of $\boldsymbol{\mu}_y$ on $\boldsymbol{\varepsilon}_x$ is weaker than that of $\boldsymbol{\varepsilon}_y$. Although such dependence does not occur in most cases, it should not be omitted in the general sense. Since $\boldsymbol{\Sigma}_y \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_y) = \mathrm{cov}[\boldsymbol{\varepsilon}_y(\boldsymbol{\mu}_x + \boldsymbol{\varepsilon}_x)]$, it is often a general function of $\boldsymbol{\Sigma}_x$ and $\boldsymbol{\mu}_x$, i.e., $\boldsymbol{\Sigma}_y = F(\boldsymbol{\Sigma}_x; \boldsymbol{\mu}_x)$ and is dependent on $\boldsymbol{\mu}_x$. In particular, when $f(\mathbf{x})$ is linear in $\mathbf{x}$, i.e., $f(\mathbf{x}) = \mathbf{a} + \mathbf{B}\,\mathbf{x}$, where $\mathbf{a}$ and $\mathbf{B}$ are a constant vector and matrix respectively, we have $\boldsymbol{\mu}_y = \mathbf{a} + \mathbf{B}\,\boldsymbol{\mu}_x$, $\boldsymbol{\varepsilon}_y = \mathbf{B}\,\boldsymbol{\varepsilon}_x$, and $\boldsymbol{\Sigma}_y = \mathbf{B}\,\boldsymbol{\Sigma}_x \mathbf{B}^{\mathrm{T}}$.

Furthermore, if a vector $\mathbf{z} \in R^r$ can be obtained by $\mathbf{y}$ and another known function $g(\cdot)$, model (I) can then be utilized repeatedly and the corresponding errors are propagated throughout the operation process as depicted in Fig. 2.1.

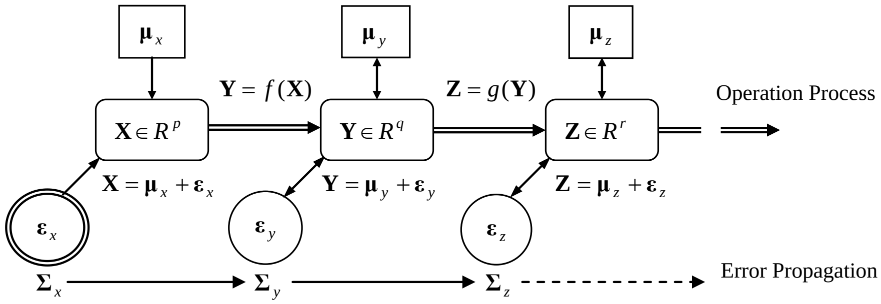


Fig. 2.1 A flowchart for error propagation based on the basic ME model (I)

2.2 Approximate law of error propagation

When the transformation function $\mathbf{y} = f(\mathbf{x})$ is completely known, error propagation can theoretically be determined in terms of this functional relationship. In this case, we can ~~divide it into~~consider two situations: (1) the error distribution of $\mathbf{Y}$ can be derived exactly from the relationship $\mathbf{Y} = f(\mathbf{X})$, e.g., the area measurement of a triangle discussed in Leung et al. (2003d); and (2) the error distribution cannot be exactly obtained or ~~it~~ is too complex to be derived because of nonlinearity of *f*. In practice, the latter situation often occurs. The focus of this subsection is to derive an approximate law of error propagation.

In case of nonlinearity, the basic idea generally adopted is to seek a linear (or second-order) approximation to $\mathbf{y} = f(\mathbf{x})$. First and second order Taylor methods, for example, may be analytical approaches to solve problems in GIS. Rosenblueth's method, on the other hand, may be an alternative to the first-~~-~~order Taylor method. If analytical expressions are not a concern, a Monte Carlo method may be employed for easy implementation. (see Heuvelink, 1998 for a discussion).

Comment [M F3]: Shouldn't linear be first-order (constant is zero-order)? Or are you saying "linear (first-order) or second-order"?

For its analytical clarity and general applicability, we first employ the Taylor method to derive a general approximate equation for error propagation. Consider a transformation function $y = f(\mathbf{x})$ in p variables $\mathbf{x} = (x_1, \cdots, x_p)^{\mathrm{T}} \in R^p$, where $y \in R$. If this function has continuous derivatives up to (n+1)-order, then it can be expanded into a Taylor series about a point $\mathbf{a} = (a_1, \cdots, a_p)^{\mathrm{T}} \in R^p$ as:

$$f(\mathbf{x}) = f(x_1, \cdots, x_p) = \sum_{j=0}^{n} \left\{ \frac{1}{j!} \left[\sum_{k=1}^{p} (x_k - a_k) \frac{\partial}{\partial x'_k} \right]^j f(x'_1, \cdots, x'_p) \right\}_{x'_1 = a_1, \cdots, x'_p = a_p} + R_n, \tag{2.4}$$

where R_n is~~,~~ called the *remainder* after n+1 terms. So the first-order approximation of $f(\mathbf{x})$ is

$$\tilde{f}(\mathbf{x}) = \tilde{f}(x_1, \cdots, x_p) = f(a_1, \cdots, a_p) + \sum_{k=1}^{p} (x_k - a_k) \frac{\partial f(x'_1, \cdots, x'_p)}{\partial x'_k} \bigg|_{x'_1 = a_1, \cdots, x'_p = a_p},$$

or

$$\tilde{f}(\mathbf{x}) = c_a + \mathbf{b}_a \mathbf{x}, \tag{2.5}$$

where

$$c_a \equiv f(a_1, \cdots, a_p) - \sum_{k=1}^{m} a_k \frac{\partial f(x'_1, \cdots, x'_p)}{\partial x'_k} \bigg|_{x'_1 = a_1, \cdots, x'_p = a_p}, \tag{2.6}$$

$$\underset{1\times p}{\mathbf{b}_a} \equiv \left(\frac{\partial f(x_1',\cdots,x_p')}{\partial x_1'},\cdots,\frac{\partial f(x_1',\cdots,x_p')}{\partial x_p'} \right)_{x_1'=a_1,\cdots,x_p'=a_p} . \tag{2.7}$$

I~~i~~t should be noted that c_a in (2.6) is a constant and $\mathbf{b}_a$ in (2.7) is a constant matrix, and they are independent of **x**. Since (2.5) is an approximation to $y = f(\mathbf{x})$, we have

$$y \approx c_a + \mathbf{b}_a \mathbf{x} . \tag{2.8}$$

If we consider $\mathbf{y} = f(\mathbf{x})$, i.e., q functions $y_i = f_i(\mathbf{x})$ in $\mathbf{x}$ $(i = 1,\cdots,q)$ which are approximated by the first-order Taylor expansion (2.5), then we can express them into a vector-matrix form as:

$$y_i \approx c_{a,i} + \mathbf{b}_{a,i}\mathbf{x} ,\quad i = 1,\cdots,q ,\quad \text{i.e.,} \begin{pmatrix} y_1 \\ \vdots \\ y_q \end{pmatrix} \approx \begin{pmatrix} c_{a,1} \\ \vdots \\ c_{a,q} \end{pmatrix} + \begin{pmatrix} \mathbf{b}_{a,1} \\ \vdots \\ \mathbf{b}_{a,q} \end{pmatrix} \mathbf{x} , \tag{2.9}$$

or

$$\mathbf{y} \approx \mathbf{c}_a + \mathbf{B}_a \mathbf{x} , \tag{2.10}$$

where $\mathbf{y} = (y_1,\cdots,y_q)^{\mathrm{T}}$, $\mathbf{c}_a = (c_{a,1},\cdots,c_{a,q})^{\mathrm{T}}$, and the coefficient matrix

$$\underset{q\times p}{\mathbf{B}_a} \equiv (\mathbf{b}_{a,1}^{\mathrm{T}},\cdots,\mathbf{b}_{a,q}^{\mathrm{T}})^{\mathrm{T}} = \left(\frac{\partial f_i(x_1',\cdots,x_p')}{\partial x_j'} \right)_{x_1'=a_1,\cdots,x_p'=a_p} \tag{2.11}$$

is called a *Jacobian matrix*, which is a matrix of partial derivatives with respect to each of the components (For simplicity and without confusion, we henceforth use y_i for both the singular and plural form of y_i, ~~(~~i.e., rather than using y_i's for plural~~)~~, and the same applies to all other relevant symbols).

Now we give the approximate law of error propagation in the basic ME model (I). Choose $\mathbf{a} = \boldsymbol{\mu}_x$ in (2.10). Then for the random measurement vector, we have

$$\mathbf{Y} \approx \mathbf{c}_{\mu_x} + \mathbf{B}_{\mu_x} \mathbf{X} = (\mathbf{c}_{\mu_x} + \mathbf{B}_{\mu_x} \boldsymbol{\mu}_x) + \mathbf{B}_{\mu_x} \boldsymbol{\varepsilon}_x .$$

Thus,

$$E(\mathbf{Y}) \approx \mathbf{c}_a + \mathbf{B}_a \boldsymbol{\mu}_x ,$$

$$\mathrm{Cov}(\mathbf{Y}) = E(\mathbf{Y} - E\,\mathbf{Y})(\mathbf{Y} - E\,\mathbf{Y})^{\mathrm{T}} \approx \mathbf{B}_{\mu_x} [E(\mathbf{X} - E\,\mathbf{X})(\mathbf{X} - E\,\mathbf{X})^{\mathrm{T}}] \mathbf{B}_{\mu_x}^{\mathrm{T}} = \mathbf{B}_{\mu_x} \boldsymbol{\Sigma}_x \mathbf{B}_{\mu_x}^{\mathrm{T}} .$$

Therefore,

$$\boldsymbol{\Sigma}_y \approx \tilde{\boldsymbol{\Sigma}}_y \equiv \mathbf{B}_{\mu_x} \boldsymbol{\Sigma}_x \mathbf{B}_{\mu_x}^{\mathrm{T}} . \tag{2.12}$$

That is, the covariance matrix, $\boldsymbol{\Sigma}_y$, for the functions **Y** can be approximately expressed by the covariance matrix $\boldsymbol{\Sigma}_x$ for the measurements **X** and the Jacobian matrix $\mathbf{B}_{\mu_x}$, as shown in Fig. 2.1. Therefore, the error in X is propagated to the error in Y, or more specifically $\boldsymbol{\Sigma}_x \longrightarrow \boldsymbol{\Sigma}_y$.

Without confusion, we denote $\mathbf{B}_{\mu_x}$ by $\mathbf{B}_\mu$ henceforth.

Remark 1. The term "error propagation law" appears in many places in the literature with some authors giving its derivation. For example, Heuvelink (1998) gives a component-wise derivation; while no approximate notation is given in the relation derived by Wolf and Ghilani (1997). ~~From~~ By using matrices~~x~~ in our derivation, the approximate law of error propagation takes on a very simple form. The emphasis here is on "approximation" and "local properties~~y~~" of the law for the nonlinear transformation function *f*. It is an "approximation" because the first-order Taylor approximation is applied in the derivation. It is local because such an approximation is effective only in the local neighborhood of $\boldsymbol{\mu}_x$. If $f(\mathbf{x})$ is nonlinear in **x**, (2.8) or (2.12) has a good approximation to the left-hand side only when **x** is very close to $\boldsymbol{\mu}_x$. In other words, the approximation is effective only in the local neighborhood of $\boldsymbol{\mu}_x$. Different from the notation adopted in Wolf and Ghilani (1997) and others, the Jacobian matrix $\mathbf{B}_\mu$ here is appended with the subscript μ in order to reflect its dependence on the local point $\boldsymbol{\mu}_x$. Therefore, to be precise, (2.12) should be called the *approximate law of error propagation* for nonlinear f.

In general, we have the following law of error propagation:

$$(\mathrm{II}): \qquad \boldsymbol{\Sigma}_y \begin{cases} \approx \tilde{\tilde{\boldsymbol{\Sigma}}}_y \equiv \mathbf{B}_\mu \boldsymbol{\Sigma}_x \mathbf{B}_\mu^{\mathrm{T}}, & \text{if } f(\mathbf{x}) \text{ is nonlinear in } \mathbf{x}, \\ = \mathbf{B}\,\boldsymbol{\Sigma}_x \mathbf{B}^{\mathrm{T}}, & \text{if } f(\mathbf{x}) \text{ is linear in } \mathbf{x}, \text{ i.e., } f(\mathbf{x}) = \mathbf{a} + \mathbf{B}\,\mathbf{x}, \\ & \text{where a and B are constants.} \end{cases}$$

Besides the first-order Taylor approximation, the second-order Taylor method in (2.4) can also be applied. It will give a more accurate approximation in, however, a more complex form (see Heuvelink, 1998; and Heuvelink et al., 1989). Though the above matrix approach may be helpful in reducing the complexity of the derivations using second- or higher-order Taylor methods, it is still difficult if not impossible to express the law of error propagation due to the complexity obtained from the expansion. Since the first-order Taylor method usually gives a good approximation (as evident in the simulation experiments in our studies), and a simple mathematical form, it is employed in our analysis wherever necessary.

If the measurements, $x_1,\ldots,x_p$, are unrelated, that is, if they are statistically independent (i.e., $\sigma_{x_i x_i} = 0$, $i \neq j$), then the covariance terms (off-diagonal elements) are equal to zero. Furthermore, if there is only one function y (q = 1), then (2.12) can be reduced into

$$\sigma_y^2 \approx \tilde{\sigma}_y^2 = \left(\sigma_{x_1}\frac{\partial y}{\partial x_1}\right)^2 + \left(\sigma_{x_2}\frac{\partial y}{\partial x_2}\right)^2 + \cdots + \left(\sigma_{x_p}\frac{\partial y}{\partial x_p}\right)^2. \tag{2.13}$$

In this equation, each term $(\partial y/\partial x_i)\sigma_{x_i}$ represents the individual contributions to the total error resulting from ME in each of the independent variables. For large error, inspection of these individual terms will indicate from where the largest contributions are coming. Then the most efficient method to reduce the overall error in the function is to examine closely ways to reduce those largest terms in (2.13).

When the transformation function $\mathbf{y} = f(\mathbf{x})$ is known up to an unknown parameter vector, e.g. $y = f(\mathbf{x}, \boldsymbol{\theta})$, we can consider estimating this parameter vector $\boldsymbol{\theta}$ by the ME model and obtain the estimated transformation function $y = \hat{f}(\mathbf{x}, \hat{\boldsymbol{\theta}})$. Because of the difference between f and $\hat{f}$, new errors in addition to ME may be introduced in subsequent analysis. This, however, is not a subject of analysis in the present paper.

2.3 Covariance-based error band

The epsilon band model of digitizing accuracy has been used to make estimates of the level of positional uncertainty and ME that is due to digitizing polygon outlines (Dunn et al., 1990). Although the paper did not develop a comprehensive analytical approach to error for the conversion of map data to digital form, the empirical results on positional accuracy and ME suggest that the interaction of the scale and quality of source documents with the unique process of digitizing may introduce unexpectedly large amounts of error. The main form of this uncertainty is positional error, because it is the location of polygon boundaries which is uncertain. Nevertheless, this in turn leads to ME since the estimation of area is then subjected to a large degree of uncertainty. There is thus an urgent need to develop methods for assessing the accuracy of vector-based GIS. Furthermore, these accuracy assessments need to be incorporated into the spatial analysis procedures, e.g., polygon overlay (the sliver polygon problem) and point-in-polygon operations, used in GIS. Compared to raster-based GIS, error analysis in vector-based GIS is much more complex.

There are a few uncertainty models for a line segment. They are for example the epsilon error band (the commonly used buffer zone) (Perkal, 1956; Perkal, 1966; Blakemore, 1984), the error band model (Shi, 1994; Shi et al, 1999), the ε_σ error band and ε_m error band (Tong et al., 1999), and the positional uncertainty model of line segments (Alesheikh and Li, 1996; Alesheikh et al., 1999). Some of these concepts are either too simple for analysis or too complex in derivation. Moreover, they usually depict a particular rather than the general version of the error band for a line segment. In fact, the basic problems around the issue are: (a) What exactly is an error band for a line segment? (b) Can probability be assigned to an error band? (c) What should the error band for a line segment be? We intend to give an answer to these questions in here and Part 2 of the present series of studies.

Actually, a simple and unified error band model can be strictly established from the concept of covariance. We call it the *covariance-based error band* and give its derivation and discussion as follows:

Uncertainty of a point can be derived from the covariance matrix associated with it and can be presented by an error ellipse (in 2-D) or an error ellipsoid (in 3-D). Since a line consists of points, uncertainty (i.e. the covariance matrix) of any point on the line can naturally reflect that of the line. Let $\mathbf{X}_i \equiv (X_{i1}, X_{i2})^{\mathrm{T}}$, $\boldsymbol{\mu}_i \equiv (\mu_{i1}, \mu_{i2})^{\mathrm{T}}$, and $\boldsymbol{\varepsilon}_i \equiv (\varepsilon_{i1}, \varepsilon_{i2})^{\mathrm{T}}$, $i = 1, 2$, be respectively the random, true, and ME vectors of the endpoint~~s~~ coordinates of the line segment V_1V_2 satisfying the relation (2.3). To make our discussion more general, we consider the 4×1 joint ME vector $\boldsymbol{\varepsilon}_{(2)} \equiv (\boldsymbol{\varepsilon}_1^{\mathrm{T}}, \boldsymbol{\varepsilon}_2^{\mathrm{T}})^{\mathrm{T}}$ and let its covariance matrix be $\boldsymbol{\Sigma}_{(2)}$, i.e.,

$$\boldsymbol{\Sigma}_{(2)} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_{(2)}) = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}, \ \boldsymbol{\Sigma}_{ij} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_j), \ i, j = 1, 2, \tag{2.15}$$

where the subscript (2) indicates that there are two points. Similarly, we have the 4×1 joint vectors: $\mathbf{X}_{(2)} \equiv (\mathbf{X}_1^{\mathrm{T}}, \mathbf{X}_2^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{\mu}_{(2)} \equiv (\boldsymbol{\mu}_1^{\mathrm{T}}, \boldsymbol{\mu}_2^{\mathrm{T}})^{\mathrm{T}}$, which satisfy

$$\mathbf{X}_{(2)} = \boldsymbol{\mu}_{(2)} + \boldsymbol{\varepsilon}_{(2)}, \ \boldsymbol{\varepsilon}_{(2)} \sim (\mathbf{0}, \boldsymbol{\Sigma}_{(2)}). \tag{2.16}$$

Comment [M F4]: An ellipsoid is normally defined as the result of rotating an ellipse about its minor axis – in which case two of the variances must be equal, and one of the covariances must be zero – so is the ellipsoid general enough – or do you define it some other way?

Now, for any point V' on the line segment V_1V_2, its coordinate vector $\mathbf{X}'$ can be represented via the joint coordinate vector $\mathbf{X}_{(2)}$ as:

$$\mathbf{X}' = t\,\mathbf{X}_1 + (1-t)\mathbf{X}_2 = \mathbf{D}_t\mathbf{X}_{(2)}, \quad 0 \le t \le 1, \tag{2.17}$$

where $\mathbf{D}_t \equiv \left(t\mathbf{I}_2 \;\; (1-t)\mathbf{I}_2\right)$ is a 2×4 constant matrix. Furthermore, we have

$$E\,\mathbf{X}' = t\,E\,\mathbf{X}_1 + (1-t)E\,\mathbf{X}_2 = t\,\boldsymbol{\mu}_1 + (1-t)\,\boldsymbol{\mu}_2 \equiv \boldsymbol{\mu}_t, \tag{2.18}$$

$$\operatorname{cov}(\mathbf{X}') = \mathbf{D}_t\operatorname{cov}(\mathbf{X}_{(2)})\,\mathbf{D}_t^{\mathrm{T}} = \left(t\mathbf{I}_2 \;\; (1-t)\mathbf{I}_2\right)\begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix}\begin{pmatrix}t\mathbf{I}_2\\ (1-t)\mathbf{I}_2\end{pmatrix}$$

$$= t^2\boldsymbol{\Sigma}_{11} + (1-t)^2\boldsymbol{\Sigma}_{22} + t\,(1-t)(\boldsymbol{\Sigma}_{12} + \boldsymbol{\Sigma}_{21}) \equiv \boldsymbol{\Sigma}_t. \tag{2.19}$$

To give an interpretation of the expression in (2.19), the first and second terms, i.e., $t^2\boldsymbol{\Sigma}_{11}$ and $(1-t)^2\boldsymbol{\Sigma}_{22}$, indicate respectively the contributions of the covariance matrices of $\mathbf{X}_1$ and $\mathbf{X}_2$ to the covariance matrix of $\mathbf{X}'$, the closer $\mathbf{X}'$ is to $\mathbf{X}_1$ (or $\mathbf{X}_2$), the bigger is the influence of the variation of $\mathbf{X}_1$ (or $\mathbf{X}_2$) on that of $\mathbf{X}'$. The third term, $t\,(1-t)(\boldsymbol{\Sigma}_{12} + \boldsymbol{\Sigma}_{21})$, on the other hand, indicates the contribution of the interrelationship (covariation) between the coordinates of $\mathbf{X}_1$ and $\mathbf{X}_2$. In particular, if $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent (i.e., the ME vectors $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ for the endpoints are independent), this term will drop out automatically and the expression in (2.19) can be reduced into

$$\operatorname{cov}(\mathbf{X}') = t^2\boldsymbol{\Sigma}_{11} + (1-t)^2\boldsymbol{\Sigma}_{22}.$$

Remark 2. Although ~~discussion on~~ the covariance matrix for arbitrary points along a line segment has been ~~made~~ discussed (Alesheikh and Li, 1996; Alesheikh, 1999; Shi, 1994; Dai et al., 1999; Tong et al., 2000), a matrix-form analytic expression, such as (2.19), of the joint ME covariance matrix has not been derived. However, expressions appearin~~ed~~ in component form can be found in a few studies (Shi and Liu, 2000; Liu and Hua, 1998). It should be noted that the derivation of (2.19) is independent of any assumption, such as the normality assumption made in many of the discussions in the literature. However, the corresponding regions will have geometric manifestations if error structures of the endpoints take on specific forms. Thus, our derivation does not involve any predetermined geometric arguments, axis rotation or component-wise computation. It greatly simplifies the derivation and gives a concise and natural expression of the result. Such a matrix method is particularly important if error analysis is to be carried out in high~~-~~dimensional space, e.g., errors in coordinates and attributes.

Component-wise derivation will get to be too tedious and complicated to derive a simple and general form for the error structure and the associated error propagation.

If we assume that $\boldsymbol{\varepsilon}_{(2)}$ is distributed as a normal distribution, i.e., $\boldsymbol{\varepsilon}_{(2)} \sim N_4(0, \boldsymbol{\Sigma}_{(2)})$, both of $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ are then normal and $\boldsymbol{\varepsilon}_i^{\mathrm{T}} \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\varepsilon}_i \sim \chi_2^2$, $i = 1, 2$, where the notation χ_p^2 denotes the chi-square distribution with p degrees of freedom. The confidence region for $\boldsymbol{\mu}_i$ with confidence probability $(1-\alpha)$ can be constructed as:

$$R_i^{(\alpha)} \equiv \{V(\mathbf{x}): \ \mathbf{x} \equiv (x_1, x_2)^{\mathrm{T}}, \ (\mathbf{x} - \boldsymbol{\mu}_i)^{\mathrm{T}} \boldsymbol{\Sigma}_i^{-1} (\mathbf{x} - \boldsymbol{\mu}_i) \le \chi_{2,\alpha}^2 \}, \ i = 1, 2, \tag{2.20}$$

since $P[V(\mathbf{X}_i) \in R_i^{(\alpha)}] = P[(\mathbf{X}_i - \boldsymbol{\mu}_i)^{\mathrm{T}} \boldsymbol{\Sigma}_i^{-1} (\mathbf{X}_i - \boldsymbol{\mu}_i) \le \chi_{2,\alpha}^2] = P[\boldsymbol{\varepsilon}_i^{\mathrm{T}} \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\varepsilon}_i \le \chi_{2,\alpha}^2] = 1 - \alpha$, where $\chi_{2,\alpha}^2$ is the upper α-quantile of the chi-square distribution χ_2^2 with 2 degrees of freedom. The shape of the confidence region is elliptical. When $\alpha = e^{-1/2}$, this ellipse is called *standard error ellipse* (Alesheikh and Li, 1996) or *error ellipse* (Wolf and Ghilani, 1997), which has confidence probability 0.393469. Based on $R_1^{(\alpha)}$ and $R_2^{(\alpha)}$, we can construct a larger region as follows:

$$R_{\{1,2\}}^{(\alpha)} \equiv \{V(\mathbf{x}): \text{there is a real number } t \ (0 \le t \le 1) \text{ such that } (\mathbf{x} - \boldsymbol{\mu}_t)^{\mathrm{T}} \boldsymbol{\Sigma}_t^{-1} (\mathbf{x} - \boldsymbol{\mu}_t) \le \chi_{2,\alpha}^2 \}\}. \tag{2.21}$$

It is obvious that $R_{\{1,2\}}^{(\alpha)}$ is the union of confidence regions (ellipses) for all points on the line segment. On this basis, we give the following definition for the error band of a line segment:

Definition 1. Assume that the joint ME vector $\boldsymbol{\varepsilon}_{(2)}$ of the endpoints coordinates of a line segment is normal. We call the region $R_{\{1,2\}}^{(\alpha)}$ in (2.21) the *covariance-based* α *-error band* for the line segment, henceforth notated as *Cov-error band* for short.

Such a band can be viewed as the confidence region for the line segment. Although its confidence probability cannot be exactly described, simulation experiments show that it is appropriate and effective to characterize uncertainties of a line segment (see Example 2.1). Moreover, it renders a unified treatment of uncertainty in line segments. For example, when $\boldsymbol{\Sigma}_{11} = \boldsymbol{\Sigma}_{22} = \boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21} = \varepsilon^2 \mathbf{I}_2$, $R_{\{1,2\}}^{(\alpha)}$ becomes the classical epsilon band; when $\boldsymbol{\Sigma}_{11} = \boldsymbol{\Sigma}_{22} = \sigma^2 \mathbf{I}_2$ and $\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21} = \mathbf{0}$, $R_{\{1,2\}}^{(\alpha)}$ becomes the error band model. With varying $\boldsymbol{\Sigma}_{(2)}$, the Cov-error band forms varying shapes. In other words, the Cov-error band has no fixed shape. Its shape depends on the structure of the joint ME covariance matrix. This is why we can appropriately call it the Cov-error band. Therefore, the Cov-error band in (2.21) is our answer to the question: "What exactly is an error band for a line segment? ".

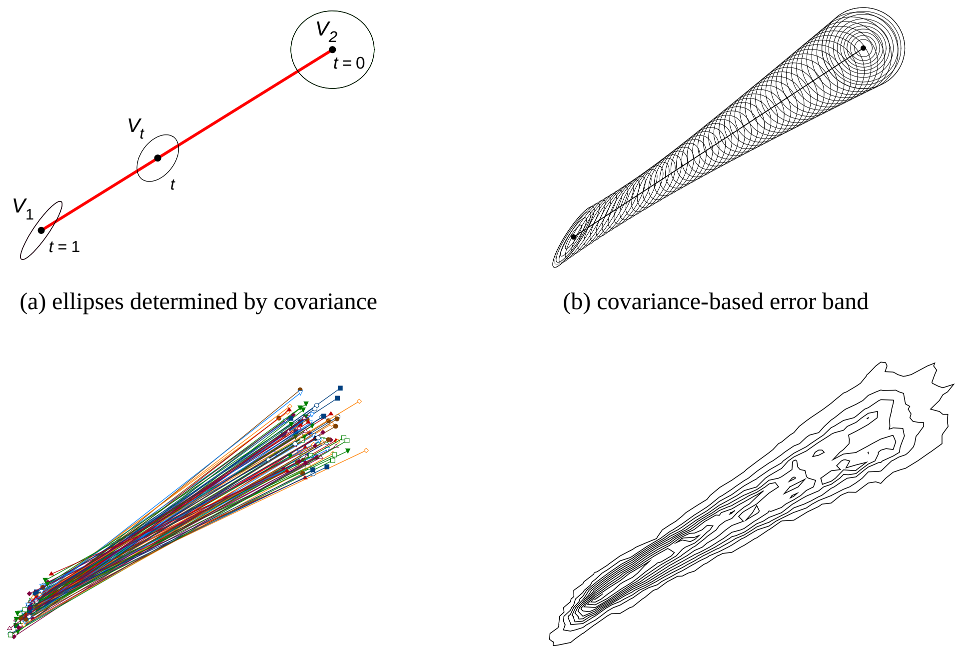
It should be noted that the confidence level of the Cov-error band is not $1-\alpha$ nor $(1-\alpha)^2$ although both of the confidence region $R_1^{(\alpha)}$ and $R_2^{(\alpha)}$ for the endpoints $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ have confidence level $1-\alpha$. The reason is that the random event $\{V_1(\mathbf{X}_1)\in R_1^{(\alpha)}$ and $V_2(\mathbf{X}_2)\in R_2^{(\alpha)}\}$ is not equal to the random event $\{L(\mathbf{X}_1,\mathbf{X}_2):\subseteq R_{\{1,2\}}^{(\alpha)}\}$ since $\{V_1(\mathbf{X}_1)\in R_1^{(\alpha)}$ and $V_2(\mathbf{X}_2)\in R_2^{(\alpha)}\}$ does not imply $\{V_1V_2\subseteq R_{\{1,2\}}^{(\alpha)}\}$ and vice versa, where $L(\mathbf{X}_1,\mathbf{X}_2)$ is the set of points on the random line segment V_1V_2. Thus, our answer to the question "Can probability be assigned to an error band? " is "no" unless a certain relaxation is made in the delimitation of the Cov-error band (to be discussed in Part 2), which in turn becomes our answer to "What should the error band for a line segment be? ".

Remark 3. The idea of using covariance to formulate error bands is indeed not new. Most error band models in the literature are in one way or the other involved with the covariance matrix, particularly Alesheikh and Li (1996) and Alesheikh et al. (1999). However, ~~there lacks~~ a strict and indepth treatment that can take us through the error band problem on the basis of the covariance matrices of the endpoints and the covariation of the endpoints is lacking.

Example 2.1 We use a numerical example to substantiate the concept of Cov-error band. Let the true coordinate vector of the endpoints of a line segment be $\boldsymbol{\mu}_1=(0,0)^{\mathrm{T}}$ and $\boldsymbol{\mu}_2=(6,4)^{\mathrm{T}}$, and the corresponding ME covariance matrices be respectively

$$\boldsymbol{\Sigma}_{11}=\begin{pmatrix}0.04 & 0.054\\ 0.054 & 0.09\end{pmatrix},\ \text{and}\ \boldsymbol{\Sigma}_{22}=\begin{pmatrix}0.16 & 0\\ 0 & 0.16\end{pmatrix},$$

that is, $\sigma_{1,}=0.2$, $\sigma_{2,}=0.3$, $\sigma_{,1}=\sigma_{,2}=0.4$, $\rho_{12,}=0.9$, and $\rho_{,12}=0$. Here and henceforth, "," in the subscript is used to separate our reference to the endpoint~~s~~ coordinate vectors $\mathbf{X}_1$ and $\mathbf{X}_2$. The subscript to the left of "," refers to the coordinate(s) of $\mathbf{X}_1$ and subscript to the right of "," refers to the coordinate(s) of $\mathbf{X}_2$. For example, $\sigma_{1,}$ means the standard deviation of the first coordinate of $\mathbf{X}_1$, $\sigma_{,2}$ means the standard deviation of the second coordinate of $\mathbf{X}_2$, and $\rho_{12,}$ means the correlation coefficient of the first and second coordinates of $\mathbf{X}_2$. Fig. 2.2(a) shows the ellipses of the endpoints and arbitrary point along a line segment. Its Cov-error band is plotted in Fig. 2.2(b), and the result of 100 random simulations is shown in Fig. 2.2(c). It can be observed that the shapes formed by these random lines is very similar to that described by the Cov-error band in Fig. 2.2(b). In other words, the Cov-error band can effectively capture the uncertainty of the line segment. To further substantiate this observation, we can investigate the distribution of random points on these random line segments by the contour plot of the empirical density function. First, 50 points are selected with equi-spacing on each line segment and 5000 points in total are then generated. Second, a region $\boldsymbol{R} = \{(x_1, x_2): -2 \le x_1 \le 8, -2 \le x_2 \le 6\}$ that contains all random line segments is partitioned into 2000 squared subregions $\boldsymbol{R}_1$, …, $\boldsymbol{R}_{2000}$, with edge length 0.2. Based on 5000 points, the empirical probability of $\boldsymbol{R}_i$ is defined by



(a) ellipses determined by covariance (b) covariance-based error band

(c) the result for random simulations 100 times (d) empirical probability contour plot

Fig. 2.2 The covariance-based error band for a line segment*
* The coordinate system is omitted for generality

$$P_i = \#\{ j \in \{1,2,\ldots,5000\} : \mathbf{z}_j \in \boldsymbol{R}_i \} / 5000 , \quad i = 1, \ldots, 2000$$

i.e., P_i is the relative frequency of the 5000 points falling into $\boldsymbol{R}_i$. We can then plot the contour lines of the empirical probability indicating the spatial distribution of a certain probability level (Fig. 2.2(d)). Obviously, the plot is consistent with the Cov-error band (Fig.2.2 (b)). So, the Cov-error bands are effective in capturing the uncertainty of a line segment.

It is well known that the correlation matrix plays an important role in statistical characterizations. To describe the correlation between endpoints rigorously, we can express the covariance matrix by the correlation matrix. The correlation matrix consists of relevant correlation coefficients with each taking values between 1.0 and –1.0. If there is a perfect positive linear relationship between the two variables, the correlation will be 1.0. If there is a perfect negative linear relationship between the two variables, the correlation coefficient is –1.0. A correlation coefficient of zero means that there is no linear relationship between the variables.

According to the relation that $\mathrm{cov}(X,Y)=\sigma_{xy}=\rho_{xy}\sigma_x\sigma_y$, where ρ_{xy} is the correlation coefficient between *X* and *Y*, we have

$$\boldsymbol{\Sigma}_{(2)}=\begin{pmatrix}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{pmatrix}=\begin{pmatrix}\sigma_{1,}^2 & \sigma_{12,} & \sigma_{1,1} & \sigma_{1,2}\\ \sigma_{12,} & \sigma_{2,}^2 & \sigma_{2,1} & \sigma_{2,2}\\ \sigma_{1,1} & \sigma_{2,1} & \sigma_{,1}^2 & \sigma_{,12}\\ \sigma_{1,2} & \sigma_{2,2} & \sigma_{,12} & \sigma_{,2}^2\end{pmatrix}$$

$$=\begin{pmatrix}\sigma_{1,} & 0 & 0 & 0\\ 0 & \sigma_{2,} & 0 & 0\\ 0 & 0 & \sigma_{,1} & 0\\ 0 & 0 & 0 & \sigma_{,2}\end{pmatrix}\begin{pmatrix}1 & \rho_{12,} & \rho_{1,1} & \rho_{1,2}\\ \rho_{12,} & 1 & \rho_{2,1} & \rho_{2,2}\\ \rho_{1,1} & \rho_{2,1} & 1 & \rho_{,12}\\ \rho_{1,2} & \rho_{2,2} & \rho_{,12} & 1\end{pmatrix}\begin{pmatrix}\sigma_{1,} & 0 & 0 & 0\\ 0 & \sigma_{2,} & 0 & 0\\ 0 & 0 & \sigma_{,1} & 0\\ 0 & 0 & 0 & \sigma_{,2}\end{pmatrix},$$

i.e.,

$$\boldsymbol{\Sigma}_{(2)}=\boldsymbol{\Sigma}_{\sigma}\boldsymbol{\rho}\,\boldsymbol{\Sigma}_{\sigma}, \tag{2.22}$$

$$\boldsymbol{\Sigma}_{\sigma}\equiv\begin{pmatrix}\sigma_{1,} & 0 & 0 & 0\\ 0 & \sigma_{2,} & 0 & 0\\ 0 & 0 & \sigma_{,1} & 0\\ 0 & 0 & 0 & \sigma_{,2}\end{pmatrix},\quad \boldsymbol{\rho}\equiv\begin{pmatrix}\boldsymbol{\rho}_{11} & \boldsymbol{\rho}_{12}\\ \boldsymbol{\rho}_{21} & \boldsymbol{\rho}_{22}\end{pmatrix}=\begin{pmatrix}1 & \rho_{12,} & \rho_{1,1} & \rho_{1,2}\\ \rho_{12,} & 1 & \rho_{2,1} & \rho_{2,2}\\ \rho_{1,1} & \rho_{2,1} & 1 & \rho_{,12}\\ \rho_{1,2} & \rho_{2,2} & \rho_{,12} & 1\end{pmatrix}, \tag{2.23}$$

where $\boldsymbol{\Sigma}_{\sigma}$ is a diagonal matrix whose diagonal elements are the standard deviations of the coordinates ME, $\boldsymbol{\rho}$ is a *correlation matrix* among the coordinates ME, $\sigma_{i,j}$ denotes the covariance of the *i*th coordinate ME of $\mathbf{X}_1$ and the *j*th coordinate ME of $\mathbf{X}_2$; $\sigma_{ij,}$ the covariance of the *i*th coordinate and the *j*th coordinate of $\mathbf{X}_1$; $\sigma_{,ij}$ the covariance of the *i*th coordinate and the *j*th coordinate of $\mathbf{X}_2$; and $\sigma_{i,}^2$

and $\sigma_{,i}^2$ are variances of the *i*th coordinates of $\mathbf{X}_1$ and $\mathbf{X}_2$ respectively. For the correlation coefficients ρ, similar notations are likewise defined.

It should be noted that ~~both of~~ the correlation coefficients $\rho_{1,2}$ and $\rho_{2,1}$ are in general not equal since $\rho_{1,2} = \mathrm{cor}(X_{11}, X_{22})$ and $\rho_{2,1} = \mathrm{cor}(X_{12}, X_{21})$. Thus the correlation matrix $\boldsymbol{\rho}_{12}$ between $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ is itself not symmetric. Moreover, the correlation matrix $\boldsymbol{\rho}$ is positive semidefinite, especially, positive definite in most cases that it is nonsingular. Although each element in $\boldsymbol{\rho}$ is a correlation coefficient that may theoretically vary between 1 and -1, we still need to observe a restriction: the positive semidefiniteness of $\boldsymbol{\rho}$ must be guaranteed. This case is different from that of a single correlation coefficient. Here is an example.

If we have a correlation matrix as follows:

$$\boldsymbol{\rho} = \begin{pmatrix} 1 & 0.9 & \rho_{1,1} & 0 \\ 0.9 & 1 & 0 & \rho_{2,2} \\ \rho_{1,1} & 0 & 1 & 0 \\ 0 & \rho_{2,2} & 0 & 1 \end{pmatrix},$$

then how can we determine the ranges of $\rho_{1,1}$ and $\rho_{2,2}$? We can use such a conclusion from matrix theory: a symmetric matrix is positive semidefinite if and only if all principal minors are non-negative. According to this conclusion, we should have

$$\det\begin{pmatrix} 1 & 0.9 & \rho_{1,1} \\ 0.9 & 1 & 0 \\ \rho_{1,1} & 0 & 1 \end{pmatrix} = 0.19 - \rho_{1,1}^2 \geq 0,\ \det\begin{pmatrix} 1 & 0.9 & \rho_{1,1} & 0 \\ 0.9 & 1 & 0 & \rho_{2,2} \\ \rho_{1,1} & 0 & 1 & 0 \\ 0 & \rho_{2,2} & 0 & 1 \end{pmatrix} = 0.19 - \rho_{1,1}^2 - \rho_{2,2}^2 + \rho_{1,1}^2\rho_{2,2}^2 \geq 0.$$

From the first condition, we can choose an approximate maximum $\hat{\rho}_{1,1} = 0.43$ for $\rho_{1,1}$ and replace it in the second condition. Then we get an approximate maximum $\hat{\rho}_{2,2} = 0.079$ for $\rho_{2,2}$.

If we consider another form of

$$\boldsymbol{\rho} = \begin{pmatrix} 1 & 0.9 & \rho & 0 \\ 0.9 & 1 & 0 & \rho \\ \rho & 0 & 1 & 0 \\ 0 & \rho & 0 & 1 \end{pmatrix},$$

then the approximate maximum for ρ that satisfies the restriction is $\hat{\rho} = 0.316$. However, if ρ is bigger than $\hat{\rho}$, e.g., $\rho = 0.32$, then $\det(\boldsymbol{\rho}) = -0.0043 < 0$ and $\boldsymbol{\rho}$ is not a correlation matrix.

Then we consider the different correlation cases for two endpoints in order to investigate the effect of such a correlation on the error band. For simplicity, we choose the following three cases for $\boldsymbol{\rho}_{12}$:

(a) $\rho_{1,1} = 0.43$, $\rho_{2,2} = 0.079$, $\rho_{1,2} = 0$, and $\rho_{2,1} = 0$;
(b) $\boldsymbol{\rho}_{12} = \mathbf{0}$;
(c) $\rho_{1,1} = -0.43$, $\rho_{2,2} = -0.079$, $\rho_{1,2} = 0$, and $\rho_{2,1} = 0$,

that is, $\boldsymbol{\Sigma}_{(2)}$ for these three cases are respectively

$$\begin{pmatrix} 0.040 & 0.054 & 0.0344 & 0 \\ 0.054 & 0.090 & 0 & 0.00948 \\ 0.0344 & 0 & 0.16 & 0 \\ 0 & 0.00948 & 0 & 0.16 \end{pmatrix}, \begin{pmatrix} 0.040 & 0.054 & 0 & 0 \\ 0.054 & 0.090 & 0 & 0 \\ 0 & 0 & 0.16 & 0 \\ 0 & 0 & 0 & 0.16 \end{pmatrix}, \begin{pmatrix} 0.040 & 0.054 & -0.0344 & 0 \\ 0.054 & 0.090 & 0 & -0.00948 \\ -0.0344 & 0 & 0.16 & 0 \\ 0 & -0.00948 & 0 & 0.16 \end{pmatrix}.$$

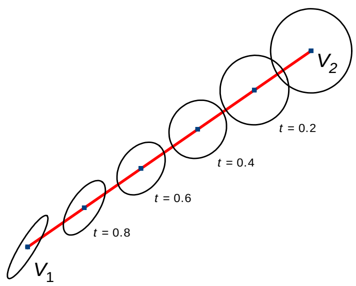
The results are depicted in Fig. 2.3 and the changes in covariance with respect to t for each case are listed in Table 2.1. It can be observed that the shapes of the Cov-error bands of the three cases are not the same. The difference is especially apparent at t = 0.8 where we can witness the biggest dispersion in (a) and the smallest dispersion in (c). It is made more revealing with reference to the corresponding correlation coefficients.

Table 2.1 The change process of covariance of some points on the line segment

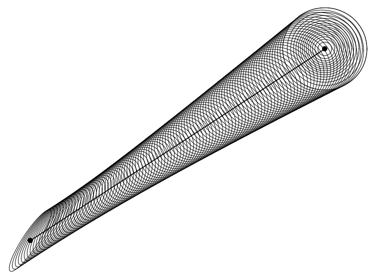
t	0	0.2	0.4	0.6	0.8	1
(a)		$\begin{pmatrix} 0.1150 & 0.0022 \\ 0.0022 & 0.1090 \end{pmatrix}$ $\rho = 0.0193$	$\begin{pmatrix} 0.0805 & 0.0086 \\ 0.0086 & 0.0766 \end{pmatrix}$ $\rho = 0.1101$	$\begin{pmatrix} 0.0565 & 0.0194 \\ 0.0194 & 0.0626 \end{pmatrix}$ $\rho = 0.3207$	$\begin{pmatrix} 0.0430 & 0.0346 \\ 0.0346 & 0.0670 \end{pmatrix}$ $\rho = 0.6437$	
(b)	$\begin{pmatrix} 0.16 & 0 \\ 0 & 0.16 \end{pmatrix}$	$\begin{pmatrix} 0.1040 & 0.0022 \\ 0.0022 & 0.1060 \end{pmatrix}$ $\rho = 0.0206$	$\begin{pmatrix} 0.0640 & 0.0086 \\ 0.0086 & 0.0720 \end{pmatrix}$ $\rho = 0.1273$	$\begin{pmatrix} 0.0400 & 0.0194 \\ 0.0194 & 0.0580 \end{pmatrix}$ $\rho = 0.4036$	$\begin{pmatrix} 0.0320 & 0.0346 \\ 0.0346 & 0.0640 \end{pmatrix}$ $\rho = 0.7637$	$\begin{pmatrix} 0.040 & 0.054 \\ 0.054 & 0.090 \end{pmatrix}$
(c)	$\rho = 0$	$\begin{pmatrix} 0.0930 & 0.0022 \\ 0.0022 & 0.1030 \end{pmatrix}$ $\rho = 0.0221$	$\begin{pmatrix} 0.0475 & 0.0086 \\ 0.0086 & 0.0674 \end{pmatrix}$ $\rho = 0.1527$	$\begin{pmatrix} 0.0235 & 0.0194 \\ 0.0194 & 0.0534 \end{pmatrix}$ $\rho = 0.5487$	$\begin{pmatrix} 0.0210 & 0.0346 \\ 0.0346 & 0.0610 \end{pmatrix}$ $\rho = 0.9661$	$\rho = 0.9$

In particular, when $\boldsymbol{\Sigma}_{11} = \boldsymbol{\Sigma}_{22} = \boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21} = \sigma^2 \mathbf{I}_2$, the Cov-error band with probability 0.39347 (i.e. formed by the error ellipses) becomes the epsilon ($\varepsilon = \sigma$) band model (see Fig. 2.4(a)), in this case, there is a positive linear relationship ($\rho_{i,i} = 1$) between two ME ε_{1i} and ε_{2i} of the endpoints, $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2})^{\mathrm{T}}$, $i = 1, 2$; when $\boldsymbol{\Sigma}_{11} = \boldsymbol{\Sigma}_{22} = -\boldsymbol{\Sigma}_{12} = -\boldsymbol{\Sigma}_{21} = \sigma^2 \mathbf{I}_2$, the Cov-error band with the same probability becomes that shown in Fig. 2.4(b) since there is perfect negative linear relationship between the corresponding MEs, $\rho_{1,1} = \rho_{2,2} = -1$; when $\boldsymbol{\Sigma}_{11} = \boldsymbol{\Sigma}_{22} = \sigma^2 \mathbf{I}_2$ and $\boldsymbol{\Sigma}_{12}$ = $\boldsymbol{\Sigma}_{21} = \mathbf{0}$, the band becomes the error band model (see Fig. 2.4(c)); and when $\boldsymbol{\Sigma}_{11} = \boldsymbol{\Sigma}_{22}$ and $\boldsymbol{\Sigma}_{12}$ =

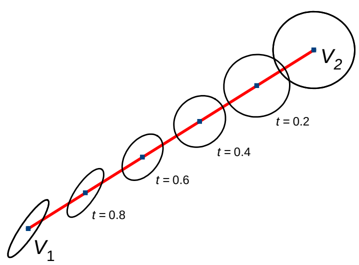
$\mathbf{\Sigma}_{21} = \mathbf{0}$, the band has the general shape as shown in Fig. 2.4(d). Note that for Fig.2.4(a) and Fig.2.4(b), their corresponding $\mathbf{\Sigma}_{(2)}$ are singular since $\det(\mathbf{\Sigma}_{(2)}) = 0$ (in fact, $\mathrm{rank}(\mathbf{\Sigma}_{(2)}) = 2$).



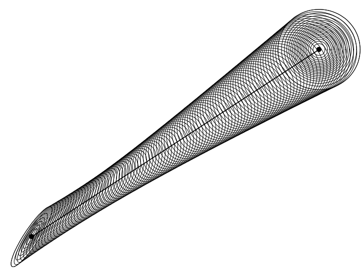
(1) Change of covariance matrices for (a)



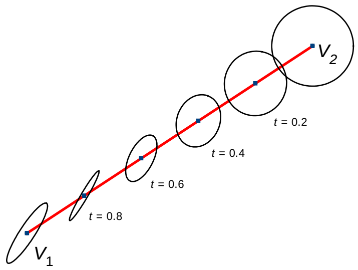
(2) Covariance-based error band for (a)



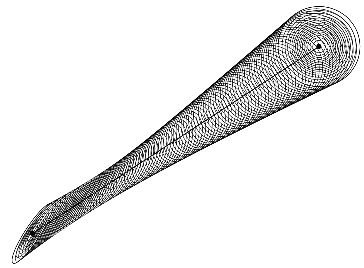
(3) Change of covariance matrices for (b)



(4) Covariance-based error band for (b)



(5) Change of covariance matrices for (c)



(6) Covariance-based error band for (c)

Fig. 2. 3 Changes in the covariance ellipses of points on the line segment under different correlation coefficients

In a word, Cov-error band can take on different shapes according to the joint ME covariance matrix $\mathbf{\Sigma}_{(2)}$. Different $\mathbf{\Sigma}_{(2)}$ yield error bands of different geometric appearances. Therefore, the error band for a line actually has no fixed shape and all of this depends on the structures of the relevant ME covariance matrices.

Fig. 2.4 (a), (b), (c) and (d) and figures below them are the Cov-error bands and 100 simulation experiment results for the following four cases:

(a) $\mathbf{\Sigma}_{11} = \mathbf{\Sigma}_{22} = \mathbf{\Sigma}_{12} = \mathbf{\Sigma}_{21} = \varepsilon^2 \mathbf{I}_2$, $\varepsilon = 0.1$; (b) $\mathbf{\Sigma}_{11} = \mathbf{\Sigma}_{22} = -\mathbf{\Sigma}_{12} = -\mathbf{\Sigma}_{21} = \sigma^2 \mathbf{I}_2$, $\sigma = 0.1$;

(c) $\mathbf{\Sigma}_{11} = \mathbf{\Sigma}_{22} = \begin{pmatrix} 0.1^2 & 0 \\ 0 & 0.1^2 \end{pmatrix}$, $\mathbf{\Sigma}_{12} = \mathbf{\Sigma}_{21} = \mathbf{0}$; (d) $\mathbf{\Sigma}_{11} = \mathbf{\Sigma}_{22} = \begin{pmatrix} 0.1^2 & 0.018 \\ 0.018 & 0.2^2 \end{pmatrix}$, $\mathbf{\Sigma}_{12} = \mathbf{\Sigma}_{21} = \mathbf{0}$.

The length of the line segment is 6.

For the examples on the Cov-error bands of polygons, see Leung et al (2003b).

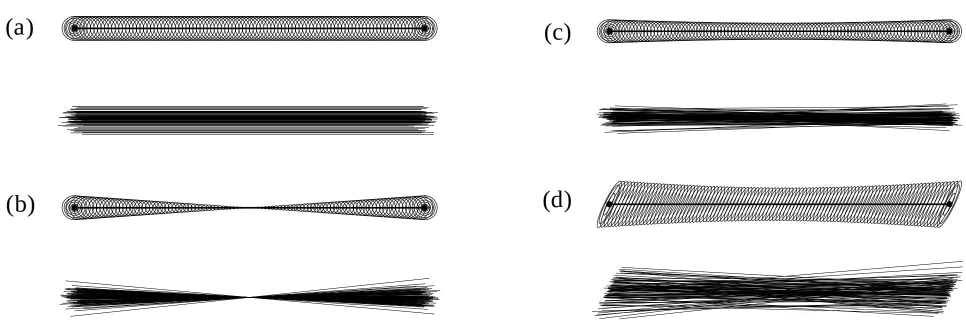
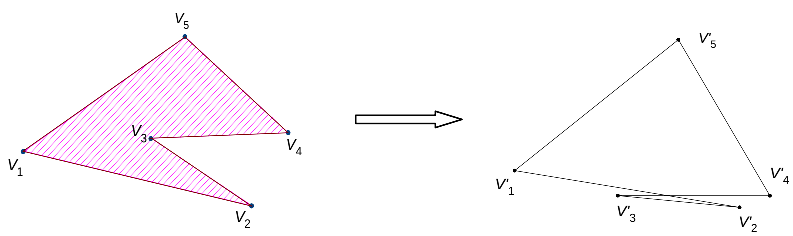


Fig. 2.4 Cov-error bands for different cases ($1-\alpha = 0.9$)

2.4 Maximal Allowable Limit (MAL) for positional error

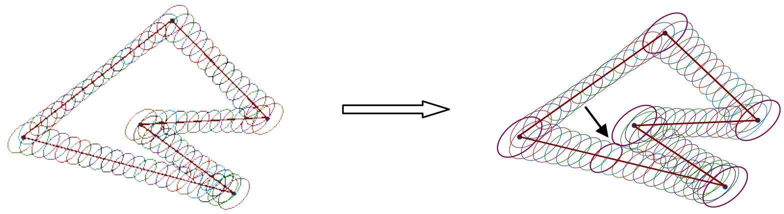
Positional accuracy can be separated into two components: absolute and relative. Relative positional accuracy reflects the consistency of any position on a map with respect to any other. While absolute positional accuracy of a map may directly influence relative accuracy, research ~~proposed~~ on this relationship is scanty. In Stanislawski et al. (1996) a technique has been proposed for quantifying absolute and relative positional accuracy estimated through error propagation from a covariance matrix for affine transformation parameters. That is, absolute positional accuracy is represented with a certainty region (propagated error ellipses) for transformed points, and relative accuracy is represented with confidence intervals for distance and azimuth values computed between transformed points.

Since these two classes of accuracy are related, when ~~a~~ one class of error is investigated the other class should be considered. It is well known that one of the important characteristics of a spatial object is its topological (or geometrical) property which involves relative error. Although we can tolerate a certain degree of ME in locations, ME must be restricted in a certain range so that it will not distort the original topology. That is, it is necessary to put a limit on ME in order to guarantee that the topology of a spatial object is invariant under ME. We call such limit the *maximal allowable limit* (MAL) for ME. For example, if two points *A* and *B* are in a straight line and have a relative locational relationship (e.g., *A* is the left of *B*), their measurements *A*' and *B*' may no longer maintain the original (true) locational relationship or may even enter into an opposite relationship (e.g., *A*' is to the right of *B*') due to exceedingly large ME. Such measurements are thus reality distorting. Therefore, large ME may change the topology and geometric property of a spatial object. Fig.2.5 depicts the situation under which the observed polygon is very different from the true polygon not only in shape but also in nature under the effect of large ME. The original polygon is simple (does not intersect itself), while the observed polygon is not simple. And the relative locations of the vertices of the original polygon are changed in the topological sense.



(a) original polygon (b) distorted polygon

Fig. 2.5 Effect of too big a ME on the topology of spatial objects



(a) smaller elliptical ME (b) MAL for elliptical ME

Fig. 2.6 The process of determination of MAL $\breve{\sigma}^2(R^{(\alpha)})$ for elliptical ME ($\rho = 0.5$, $\sigma_1^2 = \sigma_2^2 = \sigma^2$)

Comment [M F5]: It looks like there's a spurious (in the first expression – but I see it shows up in subsequent expressions as well. In my version of Word it looks like a left parenthesis above the sigma

S~~ince s~~patial objects such as lines and polygons in vector-based data are formed by points. For each point, its MALs relative to the confidence region of an object can conceptually be defined as the maximums of the variances of its coordinates so that its confidence region does not intersect the confidence region of the object. However, in practice, it is difficult and not practical to define and compute the MALs for each point. A more feasible method is to define the MALs for each object under the condition that its confidence region is delimited up to the confidence level. As a common spatial object, polygon is our main concern in this paper.

Unless specified otherwise, our discussion mostly centers~~ed~~ around simple polygons. A polygon is *simple* if it is described by a single, non-intersecting boundary; otherwise it is called *complex*. That is, a simple polygon satisfies the following two conditions: (1) all adjacent edges have only a single shared point; and (2) all non-adjacent edges do not intersect.

As discussed in subsection 2.3, when we give the structures of ME on the vertices coordinates of a polygon, the Cov-error bands of its boundaries can be constructed. More generally, MAL of a polygon with the vertices of the same error structure can be defined as follows:

Definition 2. Assume σ_1^2 and σ_2^2 be respectively the common variances of the first and second coordinates of vertices of a given simple polygon. Let $R_{1-\alpha}$ be the confidence region of the boundaries of the polygon at confidence level $(1-\alpha)$. Then the maximal allowable limits (MALs) $\breve{\sigma}_1^2$ and $\breve{\sigma}_2^2$ of the polygon are defined as the maximal allowable values of σ_1^2 and σ_2^2 with which the confidence region (given by $R_{1-\alpha}$) of each vert~~ices~~ex does not intersect the confidence regions (bands) of its disjoint edges, and are denoted by $\breve{\sigma}_1^2(R^{(\alpha)})$ and $\breve{\sigma}_2^2(R^{(\alpha)})$.

Since a confidence level is attached to the confidence regions, we may consider it as the confidence level of MAL.

Thus as long as the ME variances σ_1^2 and σ_2^2 are not greater than $\breve{\sigma}_1^2$ and $\breve{\sigma}_2^2$ respectively, the observed polygon in the presence of ME will not change the topological property of the original (true) polygon (i.e., the situation in Fig.2.5 will not occur) at the confidence level $(1-\alpha)$ and their difference results only from errors in position of the vertices. In this definition, the confidence region of the boundaries (line segments) of a polygon can be formed by any error bands of its edges. According to the advantages of the Cov-error bands, we will use the Cov-error bands as a basis for the formulation of the confidence regions of the boundaries of a polygon. Unfortunately, the confidence level of a Cov-error band cannot be determined analytically at present. An approximation to the Cov-error bands proposed in Leung et al (2003b) can however enables us to give a lower bound of the confidence level.

In general, it may not be easy to determine directly the MAL. Since the shape of the confidence region of a polygon depends on the structure of ME at the vertices, the ME structure becomes a key to obtain MALs more easily. Under certain particular situations of ME, the determination of MALs can indeed be simplified. For example, if all of the MEs of the vertices have positive linear relationships and have circular ME, i.e., $\sigma_{i1}^2 = \sigma_{i2}^2 \equiv \sigma^2$, the well-known epsilon-band is chosen (see Fig. 2.4(a)) and the corresponding MAL can be computed by distance. When $\sigma_{i1}^2 = \sigma_{i2}^2 \equiv \sigma^2$ and the correlation coefficient ρ between ε_{i1} and ε_{i2} is fixed at a certain value, e.g., 0.5, the determination of MAL using the Cov-error band is illustrated in Fig.2.6. It can be observed that the larger ~~is~~ the confidence level, the wider are the Cov-error bands and the smaller are the corresponding MALs (see Leung et al (2003b)).

MAL is thus a useful concept, particularly for ME simulation and modeling in order to control the allowable range of ME. However, we may not need to consider MAL in practice if ME is sufficiently small.

3. A geodetic model for MBGIS

Based on the basic ME model (I) and error propagation law (II), we establish in this section a geodetic model for MBGIS and substantiate it with three simple applications in geodesy.

3.1 A geodetic model for MBGIS

Comment [M F6]: In the original paper I was careful to stress that the use of the term "geodetic" did not imply that geodesy still uses this approach – it reflects a highly simplified model of traditional surveying – it might be better to use another word, perhaps hierarchical?

A MBGIS may be defined as one that provides access to the measurements **m** used to determine the locations of objects, to the function *f*, and to the rules used to determine interpolated positions. It also provides access to the locations, which may either be stored, or derived on the fly from measurements.

We consider the cases that these positions are indirectly measured by an instrument and some control points or *monuments* that are established with great accuracy by geodetic survey. Actually, a MBGIS can be constructed with reference to a geodetic model, which is formulated in such a way that locations are arranged in a hierarchy (see Fig. 3.1). At the top of the hierarchy is a small number of control points. From these a much larger number of locations are established by measurements, through a process of *densification*. Since these measurements are not as accurate as those used to establish the monuments, the second tier of locations is also less accurately known. Further measurements using even less accurate instruments are used to register aerial photographs, lay out boundary lines, and determine the contents of geographic databases.

Since there will be strong correlations in errors between any locations whose lineages share part or the entire tree, all points inherit the errors present in the monuments, but distance between points that share the same monument are not affected by errors in the location of the monument itself. Through the hierarchical structure of the data and the measurement methods, we can derive the law or approximate law for error propagation at every location. Thus it is possible to analyze uncertainties in the results of some GIS operations such as overlay and area measurement because of inaccuracies in positioning.

Based on the above discussion, a MBGIS can be formally structured as a hierarchy. As depicted in Fig. 3.1, let $\mathbf{x}^{(i)}$ denote the location at level *i* in the hierarchy. Then locations at level *i*+1 are derived from locations at levels less than or equal to *i* through equations of the form:

$$\mathbf{x}^{(i+1)} = f_i(\mathbf{m}^{(i)}, \mathbf{x}^{(i)}). \tag{3.1}$$

The top (or root) of the tree are locations $\mathbf{x}^{(0)}$ which are monuments that *anchor* the tree. At each level the measurements $\mathbf{m}^{(i)}$ constitute the trunks, the locations to be measured are the corresponding leaves, ~~and~~ the functions f_i are stored, and the locations $\mathbf{x}^{(i)}$ are either stored or derived as needed.

For a multi-level MBGIS with different measurements and functions in various levels, the propagation of error from the first level (can be from any level less than or equal to *i* through the tree) can be carried out via the relation between $\mathbf{x}^{(i+1)}$ and $\mathbf{x}^{(i)}$ depicted in (3.1).

Now we establish the ME model from the root $\mathbf{x}^{(0)}$ to any location $\mathbf{x}^{(i)}$ and discuss its error propagation problems. Let $\mathbf{M}^{(i)}$, $\boldsymbol{\varepsilon}_m^{(i)}$ and $\boldsymbol{\mu}_m^{(i)}$ be respectively the random measurement vector, the ME vector and the true value vector of $\mathbf{m}^{(i)}$, and let $\mathbf{X}^{(i)}$ be the random measurement vector of $\mathbf{x}^{(i)}$, $\boldsymbol{\Sigma}_m^{(i)} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_m^{(i)})$. Then, $\mathbf{M}^{(i)}$ can be expressed as

$$\mathbf{M}^{(i)} = \boldsymbol{\mu}_m^{(i)} + \boldsymbol{\varepsilon}_m^{(i)},\ \boldsymbol{\varepsilon}_m^{(i)} \sim (\mathbf{0},\ \boldsymbol{\Sigma}_m^{(i)}). \tag{3.2}$$

Since $\mathbf{x}^{(0)}$ is the monument, it can be viewed as constant without ME. For the sake of having unified notations, we let $\mathbf{M}_{(1)} \equiv \mathbf{M}^{(0)}$, $\boldsymbol{\mu}_{(1)} \equiv \boldsymbol{\mu}_m^{(0)}$, $\boldsymbol{\varepsilon}_{(1)} \equiv \boldsymbol{\varepsilon}_m^{(0)}$, and $\boldsymbol{\Sigma}_{(1)} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_{(1)}) = \boldsymbol{\Sigma}_m^{(0)}$. Thus we have the ME model for $\mathbf{x}^{(1)}$:

$$(\mathrm{M.1}):\ \begin{cases} \mathbf{X}^{(1)} = g_1(\mathbf{M}_{(1)}) \equiv f_0(\mathbf{M}^{(0)}, \mathbf{x}^{(0)}), & (3.3) \\ \mathbf{M}_{(1)} = \boldsymbol{\mu}_{(1)} + \boldsymbol{\varepsilon}_{(1)},\ \boldsymbol{\varepsilon}_{(1)} \sim (\mathbf{0},\ \boldsymbol{\Sigma}_{(1)}), & (3.4) \end{cases}$$

which is just the basic ME model (I). So, by similar procedure, the corresponding approximate law for error propagation can be established as:

$$\boldsymbol{\Sigma}_x^{(1)} \equiv \mathrm{cov}(\mathbf{X}^{(1)}) \approx \tilde{\boldsymbol{\Sigma}}_x^{(1)} \equiv \mathbf{B}_{\mu_{(1)}} \boldsymbol{\Sigma}_{(1)} \mathbf{B}_{\mu_{(1)}}^{\mathrm{T}},$$

where $\mathbf{B}_{\mu_{(1)}}$ is the Jacobian matrix of $g_1(\mathbf{M}_{(1)})$ at $\mathbf{M}_{(1)} = \boldsymbol{\mu}_{(1)}$.

According to (3.1) and (3.2), we can obtain the composite function g_2 as follows:

$$\mathbf{X}^{(2)} = f_1(\mathbf{M}^{(1)}, \mathbf{X}^{(1)}) = f_1(\mathbf{M}^{(1)}, f_0(\mathbf{M}^{(0)}, \mathbf{x}^{(0)})) = g_2(\mathbf{M}_{(2)}),$$

where $g_2(\mathbf{M}_{(2)}) \equiv f_1(\mathbf{M}^{(1)}, f_0(\mathbf{M}^{(0)}, \mathbf{x}^{(0)}))$ defined by f_0 and f_1 is a function of the joint ME vector $\mathbf{M}_{(2)} \equiv (\mathbf{M}^{(0)^{\mathrm{T}}}, \mathbf{M}^{(1)^{\mathrm{T}}})^{\mathrm{T}}$. Similarly, we can define the symbols $\boldsymbol{\mu}_{(2)}$ and $\boldsymbol{\varepsilon}_{(2)}$, and obtain the ME model, like (I), for $\mathbf{x}^{(2)}$:

$$(\mathrm{M.2}): \begin{cases} \mathbf{X}^{(2)} = g_2(\mathbf{M}_{(2)}), & (3.5) \\ \mathbf{M}_{(2)} = \boldsymbol{\mu}_{(2)} + \boldsymbol{\varepsilon}_{(2)},\ \boldsymbol{\varepsilon}_{(2)} \sim (\mathbf{0},\ \boldsymbol{\Sigma}_{(2)}), & (3.6) \end{cases}$$

and the approximate law of error propagation is:

$$\boldsymbol{\Sigma}_x^{(2)} \equiv \mathrm{cov}(\mathbf{X}^{(2)}) \approx \tilde{\boldsymbol{\Sigma}}_x^{(2)} \equiv \mathbf{B}_{\mu_{(2)}} \boldsymbol{\Sigma}_{(2)} \mathbf{B}_{\mu_{(2)}}^{\mathrm{T}},$$

where $\mathbf{B}_{\mu_{(2)}}$ is the Jacobian matrix of $g_2(\mathbf{M}_{(2)})$ at $\mathbf{M}_{(2)} = \boldsymbol{\mu}_{(2)}$, $\boldsymbol{\Sigma}_{(2)} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_{(2)})$, $\boldsymbol{\varepsilon}_{(2)} \equiv (\boldsymbol{\varepsilon}_m^{(0)^{\mathrm{T}}}, \boldsymbol{\varepsilon}_m^{(1)^{\mathrm{T}}})^{\mathrm{T}}$, and $\boldsymbol{\varepsilon}_m^{(i)}$ ($i = 0, 1$) are defined by (3.2). The propagation relation between the error covariance matrices $\boldsymbol{\Sigma}_{(1)} = \boldsymbol{\Sigma}_m^{(0)}$ and $\boldsymbol{\Sigma}_{(2)}$ is given by

$$\boldsymbol{\Sigma}_{(2)} = \begin{pmatrix} \boldsymbol{\Sigma}_{(1)} & \boldsymbol{\Sigma}_{(1),m(1)} \\ \boldsymbol{\Sigma}_{(1),m(1)}^{\mathrm{T}} & \boldsymbol{\Sigma}_m^{(1)} \end{pmatrix}.$$

Recursively, we have

$$\begin{aligned} \mathbf{X}^{(i)} &= f_{i-1}(\mathbf{M}^{(i-1)}, \mathbf{X}^{(i-1)}) = f_{i-1}(\mathbf{M}^{(i-1)}, f_{i-2}(\mathbf{M}^{(i-2)}, \mathbf{X}^{(i-2)})) \\ &= \cdots = f_{i-1}(\mathbf{M}^{(i-1)}, f_{i-2}(\mathbf{M}^{(i-2)}, \cdots, f_0(\mathbf{M}^{(0)}, \mathbf{x}^{(0)}) \cdots)) = g_i(\mathbf{M}_{(i)}), \end{aligned} \tag{3.7}$$

which results in the general ME model:

$$(\mathrm{M.}i): \begin{cases} \mathbf{X}^{(i)} = g_i(\mathbf{M}_{(i)}), & (3.8) \\ \mathbf{M}_{(i)} = \boldsymbol{\mu}_{(i)} + \boldsymbol{\varepsilon}_{(i)},\ \boldsymbol{\varepsilon}_{(i)} \sim (\mathbf{0},\ \boldsymbol{\Sigma}_{(i)}), & (3.9) \end{cases}$$

and

$$\boldsymbol{\Sigma}_x^{(i)} \equiv \mathrm{cov}(\mathbf{X}^{(i)}) \approx \tilde{\boldsymbol{\Sigma}}_x^{(i)} \equiv \mathbf{B}_{\mu_{(i)}} \boldsymbol{\Sigma}_{(i)} \mathbf{B}_{\mu_{(i)}}^{\mathrm{T}}, \tag{3.10}$$

where $\mathbf{B}_{\mu_{(i)}}$ is the Jacobian matrix of $g_i(\mathbf{M}_{(i)})$ at $\mathbf{M}_{(i)} = \boldsymbol{\mu}_{(i)}$, $\boldsymbol{\Sigma}_{(i)} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_{(i)})$, $\boldsymbol{\varepsilon}_{(i)} \equiv (\boldsymbol{\varepsilon}_{(i-1)}^{\mathrm{T}}, \boldsymbol{\varepsilon}_m^{(i-1)^{\mathrm{T}}})^{\mathrm{T}}$ $= (\boldsymbol{\varepsilon}_m^{(0)^{\mathrm{T}}}, \cdots, \boldsymbol{\varepsilon}_m^{(i-2)^{\mathrm{T}}}, \boldsymbol{\varepsilon}_m^{(i-1)^{\mathrm{T}}})^{\mathrm{T}}$, and $\boldsymbol{\varepsilon}_m^{(i)}$ ($i = 0, \ldots, i$) are defined by (3.2). The propagation relation between the successive error covariance matrices is given by

$$\boldsymbol{\Sigma}_{(i)} = \begin{pmatrix} \boldsymbol{\Sigma}_{(i-1)} & \boldsymbol{\Sigma}_{(i-1),m(i-1)} \\ \boldsymbol{\Sigma}_{(i-1),m(i-1)}^{\mathrm{T}} & \boldsymbol{\Sigma}_m^{(i-1)} \end{pmatrix}.$$

Obviously, both of the ME model (M.1) and (M.2) can be unified into (M.i). The corresponding approximate laws of error propagation can also be unified into (3.10). The models (M.i) and (3.10) provide an effective tool for error propagation from the root $\mathbf{x}^{(0)}$ to any location $\mathbf{x}^{(i)}$ at level i. They form the basis for error analysis in the geodetic model for MBGIS. Moreover, using a similar approach, we can formulate the ME model and the approximate law of error propagation from any location $\mathbf{x}^{(i)}$ at level i to the location derived from $\mathbf{x}^{(i)}$ at level higher than i.

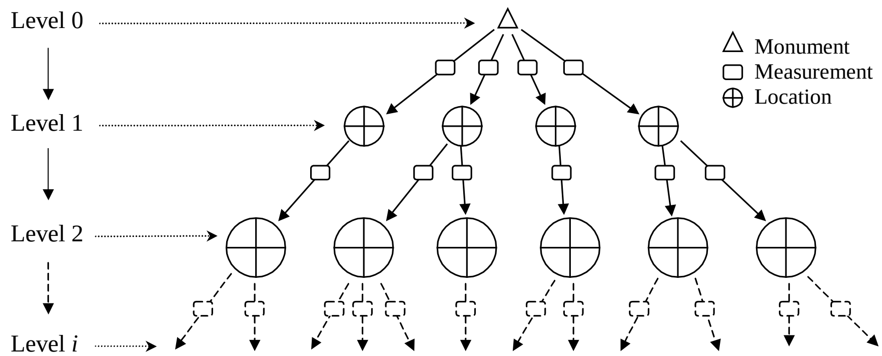


Fig. 3.1 Hierarchy tree of a simple geodetic model

Employing the results in the above discussion, we investigate error propagation in three simple measurement problems in geodesy to substantiate the geodetic model for MBGIS.

3.2 Measurement with a distance and a direction

Let V_0 be the point with known coordinates vector $\mathbf{x}_0$ (e.g., a control point), V the measured point with unknown coordinate vector $\mathbf{x}$, l the distance between V_0 and V, and θ the angle formed by the line segment V_0V and the reference line (see Fig. 3.2). Our concern is to see how the ME is propagated when $\mathbf{x}$ is measured by the measurement $\mathbf{m} \equiv (l, \theta)^{\mathrm{T}}$ with the ME vector $\boldsymbol{\varepsilon}_m \equiv (\varepsilon_l, \varepsilon_\theta)^{\mathrm{T}}$. To this end, we first need to establish the transformation function $\mathbf{x} = f(\mathbf{m})$ from $\mathbf{m}$ to $\mathbf{x}$. Obviously, we have the following expression

$$\begin{cases} x_1 = x_{01} + l\cos(\theta + \varphi_0), \\ x_2 = x_{02} + l\sin(\theta + \varphi_0), \end{cases}$$

where φ_0 is a suitably selected constant such that $\theta + \varphi_0$ becomes the angle with the x_1-coodinate axis. Thus by $\mathbf{x} = (x_1, x_2)^{\mathrm{T}}$,

$$\mathbf{x} = f(\mathbf{m}) \equiv \mathbf{x}_0 + l \cdot \mathbf{r}(\theta), \quad \mathbf{m} \equiv (l, \theta)^{\mathrm{T}}, \ \mathbf{r}(\theta) = (\cos(\theta + \varphi_0), \sin(\theta + \varphi_0))^{\mathrm{T}}.$$

Therefore, under the effect of ME $\boldsymbol{\varepsilon}_m$, the observed coordinate vector $\mathbf{X}$ of the point V can be represented as a positional ME model like (2.2) and (2.3):

$$\begin{cases} \mathbf{X} = f(\mathbf{M}) \equiv \mathbf{x}_0 + L \cdot \mathbf{r}(\Theta), & (3.11) \\ \mathbf{M} = \boldsymbol{\mu}_m + \boldsymbol{\varepsilon}_m, \ \boldsymbol{\varepsilon}_m \sim (\mathbf{0}, \boldsymbol{\Sigma}_m), & (3.12) \end{cases}$$

where $\boldsymbol{\mu}_m \equiv (\mu_l, \mu_\theta)^{\mathrm{T}}$ are the true value vector of l and θ, $\mathbf{M} \equiv (L, \Theta)^{\mathrm{T}}$ their measured value vector in the presence of the ME $\boldsymbol{\varepsilon}_m$, and $\boldsymbol{\Sigma}_m \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_m)$. Since

$$\begin{aligned}
\mathrm{d}\,\mathbf{x} &= \mathrm{d}\, f(\mathbf{m}) = (\mathrm{d}\, l)\cdot \mathbf{r}(\theta) + l\cdot \mathrm{d}\,\mathbf{r}(\theta) = \mathbf{r}(\theta)(\mathbf{e}_{2,1}^{\mathrm{T}}\mathrm{d}\,\mathbf{m}) + l\cdot(-\sin(\theta+\varphi_0), \cos(\theta+\varphi_0))^{\mathrm{T}}\,\mathrm{d}\,\theta \\
&= [\mathbf{r}(\theta)\,\mathbf{e}_{2,1}^{\mathrm{T}} + l\cdot(-\sin(\theta+\varphi_0), \cos(\theta+\varphi_0))^{\mathrm{T}}\,\mathbf{e}_{2,2}^{\mathrm{T}}](\mathrm{d}\,\mathbf{m}), \\
\mathbf{B}_{\mu_m} &= [\mathbf{r}(\theta)\,\mathbf{e}_{2,1}^{\mathrm{T}} + l\cdot(-\sin(\theta+\varphi_0), \cos(\theta+\varphi_0))^{\mathrm{T}}\,\mathbf{e}_{2,2}^{\mathrm{T}}]_{m=\mu_m} \\
&= \begin{pmatrix} \cos(\mu_\theta+\varphi_0) & -\mu_l \sin(\mu_\theta+\varphi_0) \\ \sin(\mu_\theta+\varphi_0) & \mu_l \cos(\mu_\theta+\varphi_0) \end{pmatrix}.
\end{aligned} \tag{3.13}$$

Note that $\mathbf{X} = f(\mathbf{M})$ is nonlinear in **M**. Accordingly, the approximate law of error propagation for the point V can be obtained as:

$$\tilde{\boldsymbol{\Sigma}}_x = \mathbf{B}_{\mu_m}\boldsymbol{\Sigma}_m\mathbf{B}_{\mu_m}^{\mathrm{T}}. \tag{3.14}$$

That is, the covariance matrix $\boldsymbol{\Sigma}_x$ of ME of coordinates of the position V is approximately given by (3.14).

It should be noted that if the known point V_0 (e.g., location at a certain level) is also involved with random error (the corresponding ME has the covariance matrix $\boldsymbol{\Sigma}_0$) and is independent of the measurements of l and θ, then from (3.11), (3.14) becomes

$$\tilde{\boldsymbol{\Sigma}}_x = \boldsymbol{\Sigma}_0 + \mathbf{B}_{\mu_m}\boldsymbol{\Sigma}_m\mathbf{B}_{\mu_m}^{\mathrm{T}}. \tag{3.15}$$

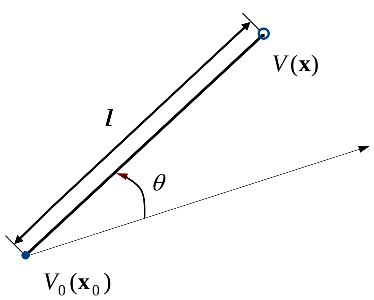


Fig. 3.2 Measurement with a distance and a direction

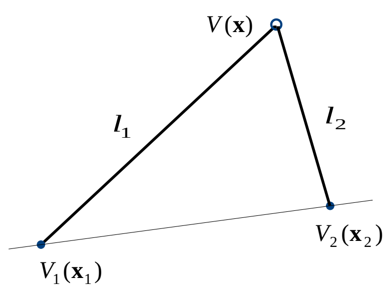


Fig. 3.3 Measurement with two distances

3.3 Measurement with two distances

Denote respectively by l_1 and l_2 the distances measured from two points V_1 and V_2 with known coordinates $\mathbf{x}_1$ and $\mathbf{x}_2$ to a common unknown point V. We should know on which side of the line $V_1 V_2$ is V located. The purpose is to measure the coordinate vector **x** of V by the measurements $\mathbf{m}_l \equiv (l_1, l_2)^{\mathrm{T}}$, and study the corresponding error propagation problem (see Fig. 3.3). Let $\boldsymbol{\varepsilon}_l \equiv (\varepsilon_{l,1}, \varepsilon_{l,2})^{\mathrm{T}}$ be the ME vector. On these conditions, we have

$$l_i^2 = (\mathbf{x} - \mathbf{x}_i)^{\mathrm{T}} (\mathbf{x} - \mathbf{x}_i)\,,\ i = 1, 2\,,$$

which can be viewed as an equation $F_l(\mathbf{x}, \mathbf{m}_l) = 0$. According to the implicit function theorem, we can obtain the required transformation function $\mathbf{x} = f_l(\mathbf{m}_l)$. Thus, the positional ME model is

$$\begin{cases} \mathbf{X} = f_l(\mathbf{M}_l)\,, & (3.16) \\ \mathbf{M}_l = \boldsymbol{\mu}_l + \boldsymbol{\varepsilon}_l\,,\quad \boldsymbol{\varepsilon}_l \sim (\mathbf{0}, \boldsymbol{\Sigma}_l)\,, & (3.17) \end{cases}$$

where $\boldsymbol{\mu}_l \equiv (\mu_{l,1}, \mu_{l,2})^{\mathrm{T}}$ is the true vector of $\mathbf{m}_l$, $\mathbf{M}_l \equiv (L_1, L_2)^{\mathrm{T}}$ the observed vector of $\mathbf{m}_l$ under the ME.

Although the explicit expression of $f_l(\cdot)$ is not given, the Jacobian matrix can still be obtained. Indeed, by differentiating the above equalities, we have

$$l_i\,(\mathrm{d}\,l_i) = (\mathbf{x} - \mathbf{x}_i)^{\mathrm{T}} (\mathrm{d}\,\mathbf{x})\,,\ i = 1, 2\,,$$

which can be rewritten in matrix form as

$$\begin{pmatrix} (\mathbf{x}-\mathbf{x}_1)^{\mathrm{T}} \\ (\mathbf{x}-\mathbf{x}_2)^{\mathrm{T}} \end{pmatrix} (\mathrm{d}\,\mathbf{x}) = \begin{pmatrix} l_1 & 0 \\ 0 & l_2 \end{pmatrix} (\mathrm{d}\,\mathbf{m}_l)\text{, i.e., } \mathrm{d}\,\mathbf{x} = \begin{pmatrix} (\mathbf{x}-\mathbf{x}_1)^{\mathrm{T}} \\ (\mathbf{x}-\mathbf{x}_2)^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} l_1 & 0 \\ 0 & l_2 \end{pmatrix} (\mathrm{d}\,\mathbf{m}_l)\,. \tag{3.18}$$

Therefore,

$$\mathbf{B}_{\mu_l} = \begin{pmatrix} (\mathbf{x}-\mathbf{x}_1)^{\mathrm{T}} \\ (\mathbf{x}-\mathbf{x}_2)^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} l_1 & 0 \\ 0 & l_2 \end{pmatrix}_{\boldsymbol{\mu}_l} = \begin{pmatrix} (\boldsymbol{\mu}_x-\mathbf{x}_1)^{\mathrm{T}} \\ (\boldsymbol{\mu}_x-\mathbf{x}_2)^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} \mu_{l,1} & 0 \\ 0 & \mu_{l,2} \end{pmatrix}. \tag{3.19}$$

It should be noted that $\boldsymbol{\mu}_x \equiv (\mu_{x,1}, \mu_{x,2})^{\mathrm{T}}$ in (3.19) is determined by $\mu_{l,i}^2 = (\boldsymbol{\mu}_x - \mathbf{x}_i)^{\mathrm{T}} (\boldsymbol{\mu}_x - \mathbf{x}_i)$, $i = 1, 2$, i.e., $\boldsymbol{\mu}_x$ is the true coordinate vector of the unknown point V. Accordingly, $\mathbf{B}_{\mu_l}$ is only dependent on $\boldsymbol{\mu}_l$. The approximate law for error propagation is

$$\tilde{\boldsymbol{\Sigma}}_x = \mathbf{B}_{\mu_l} \boldsymbol{\Sigma}_l \mathbf{B}_{\mu_l}^{\mathrm{T}}\,. \tag{3.20}$$

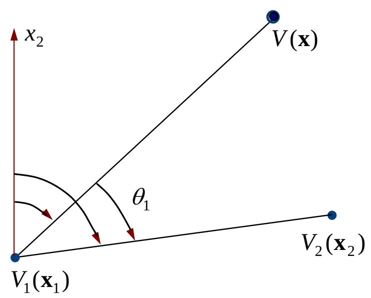


Fig. 3.4 An angle and two azimuths

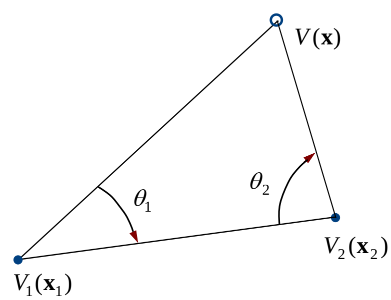


Fig. 3.5 Measurement with two angles

3.4 Measurement with two angles

In surveying, a useful concept is azimuths or angles. For example, the positions of widely spaced stations can be computed from measured angles and a minimal number of measured distances called baselines (Wolf and Ghilani, 1997). The azimuth of an object is the angular distance along the horizon to the location of the object. By convention, azimuth is measured from north towards the east along the horizon (see Fig. 3.4). We will utilize such a concept in order to make our conclusions suitable to such a task.

First, we introduce an equation for angle observation. In Fig. 3.4, the azimuths of $V(\mathbf{x})$ and $V_2(\mathbf{x}_2)$ with respect to $V_1(\mathbf{x}_1)$ can be used to represent an angle θ_1 between two line segments:

$$\theta_1 \equiv \angle VV_1V_2 = \tan^{-1}\left(\frac{x_{2,1}-x_{1,1}}{x_{2,2}-x_{1,2}}\right) - \tan^{-1}\left(\frac{x_1-x_{1,1}}{x_2-x_{1,2}}\right) + D, \tag{3.21}$$

where D is a constant that depends on the quadrants in which V and V_2 occur (for details, see Wolf and Ghilani (1997)), here V and V_2 correspond to the backsight and foresight station respectively, and V_1 the instrument station, $\mathbf{x}_i = (x_{i,1}, x_{i,2})^{\mathrm{T}}$, $i = 1, 2$, $\mathbf{x} = (x_1, x_2)^{\mathrm{T}}$.

Now assume that the unknown point V is measured by two angles θ_1 and θ_2 in V_1 and V_2, respectively. Then for θ_2 we have also the corresponding equation

$$\theta_2 \equiv \angle V_1V_2V = \tan^{-1}\left(\frac{x_1-x_{2,1}}{x_2-x_{2,2}}\right) - \tan^{-1}\left(\frac{x_{1,1}-x_{2,1}}{x_{1,2}-x_{2,2}}\right) + D. \tag{3.22}$$

They can be formed into a new multivariate equation $F_\theta(\mathbf{x}, \mathbf{m}_\theta) = 0$, where $\mathbf{m}_\theta \equiv (\theta_1, \theta_2)^{\mathrm{T}}$ is the measurement. Once again, the implicit function theorem can be applied. The transformation function $\mathbf{x} = f_\theta(\mathbf{m}_\theta)$ is formally generated. In the presence of the ME vector $\boldsymbol{\varepsilon}_\theta \equiv (\varepsilon_{\theta,1}, \varepsilon_{\theta,2})^{\mathrm{T}}$, the positional ME model is

$$\begin{cases} \mathbf{X} = f_\theta(\mathbf{M}_\theta), & (3.23) \\ \mathbf{M}_\theta = \boldsymbol{\mu}_\theta + \boldsymbol{\varepsilon}_\theta, \quad \boldsymbol{\varepsilon}_\theta \sim (\mathbf{0}, \boldsymbol{\Sigma}_\theta), & (3.24) \end{cases}$$

where $\boldsymbol{\mu}_\theta \equiv (\mu_{\theta,1}, \mu_{\theta,2})^{\mathrm{T}}$ is the true vector of $\mathbf{m}_\theta$, $\mathbf{M}_\theta \equiv (\Theta_1, \Theta_2)^{\mathrm{T}}$ the observed vector of $\mathbf{m}_\theta$ under the ME $\boldsymbol{\varepsilon}_\theta$.

Differentiating (3.21) and (3.22), and by straightforward calculation, we obtain

$$\begin{cases} (x_1 - x_{1,1})(\mathrm{d}x_2) - (x_2 - x_{1,2})(\mathrm{d}x_1) = l_{x,1}^2(\mathrm{d}\theta_1), \\ (x_2 - x_{2,2})(\mathrm{d}x_1) - (x_1 - x_{2,1})(\mathrm{d}x_2) = l_{x,2}^2(\mathrm{d}\theta_2), \end{cases}$$

where $l_{x,i}^2 = (\mathbf{x} - \mathbf{x}_i)^{\mathrm{T}}(\mathbf{x} - \mathbf{x}_i)$, $i = 1, 2$. Solving this system for $\mathrm{d}\mathbf{x} = (\mathrm{d}x_1, \mathrm{d}x_2)^{\mathrm{T}}$, we obtain

$$\mathrm{d}\,\mathbf{x} = \begin{pmatrix} -(x_2 - x_{1,2}) & x_1 - x_{1,1} \\ x_2 - x_{2,2} & -(x_1 - x_{2,1}) \end{pmatrix}^{-1} \begin{pmatrix} l_{x,1}^2 & 0 \\ 0 & l_{x,2}^2 \end{pmatrix} (\mathrm{d}\,\mathbf{m}_\theta)\,.$$

Thus,

$$\begin{aligned} \mathbf{B}_{\mu_\theta} &= \begin{pmatrix} -(x_2 - x_{1,2}) & x_1 - x_{1,1} \\ x_2 - x_{2,2} & -(x_1 - x_{2,1}) \end{pmatrix}^{-1} \begin{pmatrix} l_{x,1}^2 & 0 \\ 0 & l_{x,2}^2 \end{pmatrix}_{\boldsymbol{\theta}=\boldsymbol{\mu}_\theta} \\ &= \frac{1}{|\,\mathbf{x}-\mathbf{x}_1 \;\; \mathbf{x}-\mathbf{x}_2\,|} \Big(l_{x,1}^2(\mathbf{x}-\mathbf{x}_2) \quad l_{x,2}^2(\mathbf{x}-\mathbf{x}_1) \Big)_{\boldsymbol{\theta}=\boldsymbol{\mu}_\theta} \\ &= \frac{1}{|\,\boldsymbol{\mu}_x-\mathbf{x}_1 \;\; \boldsymbol{\mu}_x-\mathbf{x}_2\,|} \Big(\mu_{l,1}^2(\boldsymbol{\mu}_x-\mathbf{x}_2) \quad \mu_{l,2}^2(\boldsymbol{\mu}_x-\mathbf{x}_1) \Big), \end{aligned} \tag{3.25}$$

$$\tilde{\tilde{\boldsymbol{\Sigma}}}_x = \mathbf{B}_{\mu_\theta} \boldsymbol{\Sigma}_\theta \mathbf{B}_{\mu_\theta}^{\mathrm{T}}\,. \tag{3.26}$$

Similar to (3.19), $\boldsymbol{\mu}_x$ in (3.25) is determined by $\mu_{l,i}^2 = (\boldsymbol{\mu}_x - \mathbf{x}_i)^{\mathrm{T}}(\boldsymbol{\mu}_x - \mathbf{x}_i)$, $i = 1, 2$, where $\mu_{l,1}$ and $\mu_{l,2}$ are true edges V_1V and V_2V of the triangle formed by the edge V_1V_2 and the two true angles $\mu_{\theta,1}$ and $\mu_{\theta,2}$ (see Fig. 3.5).

Based on two angles θ_1 and θ_2 measured from known (or control) points, it has been determined that the position of the unknown point can be uniquely determined. The propagated ME is approximately given by (3.26). However, if additional measurement point (e.g., V_3) is available, the determination of the position of an unknown point can be strengthened by the adjustment of the least squares method (Wolf and Ghilani, 1997). The trilateration (distance measurements) and triangulation (angle measurements) adjustments in surveying are all performed by additional measurements.

To show the applicability of these models, we give the following simulation experiment as an example:

Example 3.1 Suppose that two known points are $V_1(\mathbf{x}_1)$ and $V_2(\mathbf{x}_2)$, $\mathbf{x}_1 = (1,\ 1)^{\mathrm{T}}$, $\mathbf{x}_2 = (5,\ 1)^{\mathrm{T}}$, an unknown point V to be measured is above the line segment V_1V_2, and the true position of the unknown point V is $\boldsymbol{\mu}_x = (2,\ 3)^{\mathrm{T}}$. The true distances are $\mu_{l,1} = \sqrt{(\boldsymbol{\mu}_x - \mathbf{x}_1)^{\mathrm{T}}(\boldsymbol{\mu}_x - \mathbf{x}_1)} = \sqrt{5}$, $\mu_{l,2} = \sqrt{(\boldsymbol{\mu}_x - \mathbf{x}_2)^{\mathrm{T}}(\boldsymbol{\mu}_x - \mathbf{x}_2)} = \sqrt{13}$, $\boldsymbol{\mu}_l = (\sqrt{5}, \sqrt{13})^{\mathrm{T}}$. From the relation

$$l_i^2 = (\mathbf{x}-\mathbf{x}_i)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}_i)\,,\ i = 1, 2\,,$$

we can get for $\mathbf{x} = (x_1,\ x_2)^{\mathrm{T}}$

$$\begin{cases} x_1 = 3 + \frac{1}{8}(l_1^2 - l_2^2)\,, \\ x_2 = 1 + \frac{1}{8}(-256 + 32\cdot l_1^2 - l_1^4 + 32\cdot l_2^2 + 2\cdot l_1^2\cdot l_2^2 - l_2^4)^{\frac{1}{2}}\,. \end{cases}$$

Thus we can obtained the transformation function $\mathbf{X} = f(\mathbf{L})$ and the ME model:

$$\begin{cases} \mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = f_l(\mathbf{M}_l) \equiv \begin{pmatrix} 3 + \frac{1}{8}(L_1^2 - L_1^2) \\ 1 + \frac{1}{8}(-256 + 32 \cdot L_1^2 - L_1^4 + 32 \cdot L_2^2 + 2 \cdot L_1^2 \cdot L_2^2 - L_2^4)^{\frac{1}{2}} \end{pmatrix}, \\ \mathbf{M}_l = \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} = \boldsymbol{\mu}_l + \boldsymbol{\varepsilon}_l = \begin{pmatrix} \sqrt{5} \\ \sqrt{13} \end{pmatrix} + \begin{pmatrix} \varepsilon_{l,1} \\ \varepsilon_{l,2} \end{pmatrix}, \end{cases}$$

Assume that $\boldsymbol{\varepsilon}_l \sim N_2(\mathbf{0}, \boldsymbol{\Sigma}_l)$, $\boldsymbol{\Sigma}_l = \sigma^2 \mathbf{I}_2$, and $\sigma = 0.2$. According to (3.19) and (3.20), we have

$$\mathbf{B}_{\mu_l} = \begin{pmatrix} (\mathbf{x} - \mathbf{x}_1)^{\mathrm{T}} \\ (\mathbf{x} - \mathbf{x}_2)^{\mathrm{T}} \end{pmatrix}^{-1} \begin{pmatrix} l_1 & 0 \\ 0 & l_2 \end{pmatrix}_{\boldsymbol{\mu}_l} = \begin{pmatrix} 1 & 2 \\ -3 & 2 \end{pmatrix}^{-1} \begin{pmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{13} \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 2\sqrt{5} & -2\sqrt{13} \\ 3\sqrt{5} & \sqrt{13} \end{pmatrix},$$

$$\tilde{\boldsymbol{\Sigma}}_x = \mathbf{B}_{\mu_l} \boldsymbol{\Sigma}_l \mathbf{B}_{\mu_l}^{\mathrm{T}} = \sigma^2 \begin{pmatrix} \frac{9}{8} & \frac{1}{16} \\ \frac{1}{16} & \frac{29}{32} \end{pmatrix} = \begin{pmatrix} 0.0450 & 0.00250 \\ 0.0025 & 0.03625 \end{pmatrix}.$$

To substantiate the effectiveness of $\tilde{\boldsymbol{\Sigma}}_x$, by 10000 random simulations we obtain the following sample mean and sample covariance matrix

$$\overline{\mathbf{X}} = \begin{pmatrix} 2.00108 \\ 2.99164 \end{pmatrix}, \quad \hat{\boldsymbol{\Sigma}}_x = \begin{pmatrix} 0.04483 & 0.00418 \\ 0.00418 & 0.03640 \end{pmatrix}.$$

Obviously, the difference between the simulated covariance matrix $\hat{\boldsymbol{\Sigma}}_x$ and the propagated covariance matrix $\tilde{\boldsymbol{\Sigma}}_x$ is very small. Therefore, $\tilde{\boldsymbol{\Sigma}}_x$ can approximately represent the covariance matrix of the ME vector of the position to be measured.

4. Conclusion

We have proposed in this part of the four-part series of papers a general framework for error analysis in measurement-based GIS within which the basic ME model has been constructed and the law of error propagation has been derived. A simple, strict and unified error band model for a line, called the “covariance-based error band”, has been formulated and its different shapes have been investigated under various situations for the joint ME covariance matrix. It has been demonstrated that the covariance-based error band has no fixed shape. Many of the existing error band models are just special cases of the covariance-based error band. A related concept, called the “ maximal allowable limit”, has been proposed to guarantee topology invariance under ME and to make error analysis logically consistent. Extending~~ed on~~ the basic ME model, we have also constructed a geodetic model for MBGIS and study its ME model and the corresponding law of error propagation. Three simple applications in geodesy with simulated data have been made to show their effectiveness.

On the basis of the theoretical and experimental results discussed in the present part of the series, we will investigate the point-in-polygon issue under ME in the second part, and by doing so will develop ~~along which~~ a new perspective on the error band for a line segment ~~will also be formed~~.

References

Alesheikh, A.A. and R. Li. 1996. Rigorous uncertainty models of line and polygon objects in GIS, *Proceedings of GIS/LIS '96*, Denver, CO, pp. 906-920.

Alesheikh, A.A., J.A.R. Blais, M.A. Chapman, and H. Karimi. 1999. Rigorous geospatial data uncertainty models for GISs. In *Spatial Accuracy Assessment: Land Information Uncertainty in Natural Resources*, K. Lowell and A. Jaton (eds), pp. 195-202, Chelsea, Michigan: Ann Arbor Press.

Blakemore, M. 1984. Generalization and error in spatial data bases, *Cartographica*, 21,131-139.

Chrisman, N.R. 1982. A theory of cartographic error and its measurement in digital data bases. *Proc. Aut-Carto* 5. Crystal City, Virginia, pp. 159-168.

Cressie, N.A.C. 1993. *Statistics for Spatial Data*, *Revised Edition*. John Wiley & Sons, New York.

Dai, H., W. Liu, and D. Du. 1999. A united model of visualizing positional uncertainties for spatial objects within vector GIS. In Shi, W.Z., Goodchild, M.F. and Fisher, P.F. (Eds), *Proceedings of the International Symosium on Spatial Data Quality* '99, Hong Kong: Hong Kong Polytechnic University, 1-9.

Dunn, R., A.R. Harrison, and J.C. White. 1990. Positional accuracy and measurement error in digital databases of land use: an empirical study. *Int. J. Geographical Information Systems*, 4(4), 385-398.

Goodchild, M.F. and S. Gopal. (Eds). 1989. *Accuracy of Spatial Databases*, London: Taylor & Francis.

Goodchild, M.F. 1991. Issues of quality and uncertainty, In Muller, J.C.(Ed), *Advances in Cartography*, London and New York: Elservier Science, 113-139.

Goodchild, M.F. 1999. Measurement-based GIS, In Shi, W.Z., Goodchild, M.F. and Fisher, P.F. (Eds), *Proceedings of the International Symosium on Spatial Data Quality* '99, Hong Kong: Hong Kong Polytechnic University, 1-9.

Heuvelink, G.B.M., P.A. Burrough, and A. Stein. 1989. *Propagation of errors in spatial modeling with GIS*. *Int. J. Geographical Information Systems*, 3, 303-322.

Heuvelink, G.B.M. 1998. *Error Propagation in Environmental Modelling with GIS*, London: Taylor & Francis.

Keefer, B.J., J.L. Smith, and T.G. Gregoire. 1991. Modeling and evaluating the effects of stream mode digitizing errors on map variables. *Photogrammetric Engineering and Remote Sensing*, 57(7), 957-963.

Leung, Y., and J.P. Yan. 1998. A locational error model for spatial features. *Int. J. Geographical Information Science*, 12, 607-620.

Leung, Y., and J.P. Yan. 1997. Point-in-polygon analysis under certainty and uncertainty. *GeoInformatica*, 1, 93-114.

Leung, Y., J. H. Ma, and M.F. Goodchild. 2003b. A general framework for error analysis in measurement-based GIS---Part 2: the algebraic-based probability model for point-in-polygon analysis. (unpublished paper)

Leung, Y., J. H. Ma, and M.F. Goodchild. 2003c. A general framework for error analysis in measurement-based GIS---Part 3: error analysis in intersections and overlays. (unpublished paper)

Leung, Y., J. H. Ma, and M.F. Goodchild. 2003d. A general framework for error analysis in measurement-based GIS---Part 4: error analysis in length and area measurements. (unpublished paper)

Liu, D., and Hua H. 1998. The more discussion to the modeling uncertainty of line primitives in GIS. *Acta Geodaetica et Cartographica Sinca*, 27(1): 45-49.

Mowrer, H. T. and Congalton, R. G. (eds). 2000. *Quantifying Spatial Uncertainty in Natural Resources: Theory and Applications for GIS and Remote Sensing*. Chelsea (Michigan): Ann Arbor Press.

Neuilly, M. and CETAMA. 1999. *Modelling and Estimation of Measurement Errors*, Hampshire (UK), New York: Intercept Ltd.

Perkal, J. 1956. On epsilon length, *Bulletin de i'Academie Polonaise des Sciences*, 4: 399-403.

Perkal, J. 1966. On the length of empirical curves, Discussion Paper Number 10, Michigan Inter-University Community of Mathematical Geography.

Shi, W. (1994). *Modeling Positional and Thematic Uncertainties in Intergration of Remote Sensing and GIS,* Ph.D. Thesis, ITC Publication No. 22. ITC, The Netherlands.

Shi, W., Ehlers, M., and Tempfli, K. 1999. Analytical modeling positional and thematic uncertainties in the intergration of remote sensing and geographical information systems, *Transactions in GIS*, 3(2): 119-136.

Shi, W., and W. Liu. 2000. A stochastic process-based model for the positional error of line segments in GIS. *Int. J. Geographical Information Science*, 14(1): 51-66.

Stanislawski, L.V., B.A. Dewitt & R. S. Shrestha 1996. Estimating positional accuracy of data layers within a GIS through error propagation. *Photogrammetric Engineering and Remote Sensing*, 62(4), 429-433.

Tong, X., Shi, W., and Liu, D. 1999. Development of error models for transition curves in GIS. In Shi, W.Z., Goodchild, M.F. and Fisher, P.F. (Eds), *Proceedings of the International Symosium on Spatial Data Quality* '99, Hong Kong: Hong Kong Polytechnic University, 299-307.

Tong, X., Shi, W., and Liu, D. 2000. Positional error of line primitives: incorporating the error of parameters in GIS. In Forer, P., Yeh, A.G.O. and He, J. (Eds), *Proceedings of 9th International Symosium on Spatial Data Quality* , Beijing: The Study Group on Geographical Information Science of the International Geographical Union, August 10-12, Errors 6a.3-12.

University Consortium for Geographic Information Science (UCGIS). 1996. Research priorities for geographic information science, *Cartography and Geographic Information Science*, 23, 115-127.

Veregin, H. 1989. A taxonomy of error in spatial databases. *Technical Paper 89-12*, *National Center for Geographic Information & Analysis*, Geography Department, University of California, Santa Barbara, California.

Veregin, H. 1989b. A review of error models for vector to raster conversion. *The Operational Geographer*, 7(1), 11-15.

Wolf, P.R. and Ghilani, C.D. 1997. *Adjustment Computations: Statistics and Least Squares in Surveying and GIS*, New York: John Wiley & Sons, Inc.