A General Framework fFor Error Analysis iIn Measurement-Bbased GIS, Part 3: Error Analysis iIn Intersections aAnd Overlays

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 paper is Part 3 of our four-part research series on the development of a general framework for error analysis in measurement-based geographic information systems (MBGIS). In this paper, we study the characteristics of error structures in intersections and polygon overlays. When locations of the endpoints of two line segments are in errors, we analyze errors of the intersection point and obtain its error covariance matrix through the propagation of the error covariance matrices of the endpoints. An approximate law of error propagation for the intersection point is formulated within the MBGIS framework. From simulation experiments, it appears that both of the relative positioning of two line segments and the error characteristic of the endpoints can affect the error characteristics of the intersection. Nevertheless, the approximate law of error propagation captures nicely the error characteristics under various situations. Based on the derived results, error analysis in polygon-on-polygon overlay operation is also performed. The relationship between the error covariance matrices of the original polygons and the overlaid polygons is approximately established.

Keywords: Error analysis, line-in-polygon overlay, polygon-on-polygon overlay, intersection point, approximate law of error propagation

1. Introduction

Overlay is a common but important operation in GIS applications (Rigaux et al., 2002). For example, map overlay can be used for the purpose of resource exploitation and environmental risk

assessment. Since vector-based topological map overlay operations involve overlaying point, line, or polygon features of one layer on thosehat of another, point-in-polygon, line-in-polygon and polygon-on-polygon operations are thus of fundamental importance (Goodchild, 1978). The ability to integrate a variety of data sources using overlay operations is a key analytical capability of a GIS. Overlay operations can be performed both in vector- and raster-based GIS. Vector overlays are methodologically and technically more complex than raster overlays. They usually result in a more complex output layer with more nodes, lines and polygons than that of the original files. In comparison with raster-based overlay operations, they are more time-consuming, complex and computationally expensive, and must update the topological tables of spatial relationships between points, lines and polygons. In fact, vector data overlay, especially vector polygon overlay, is among the most computationally intensive group of GIS operations and frequently proves to be a bottleneck in 'production-line' processing (Harding et al., 1998).

Error analyses for overlay operations focus mainly on raster-based data; . Ssimilar study research on vector-based data is, however, limitedseldom made. For new geographic objects created from overlay, their geometry is computed by applying the intersection operation to the geometry of the involved geographic objects. Polygon overlay involves the intersection of the boundaries of one set of polygons with the boundaries of a second set of polygons to produce a third set. Each polygon in the output set is related to one polygon in each of the input sets. The attributes and positions of the output polygons can therefore be derived directly from the attributes and positions of the corresponding input polygons.

In GIS applications, error propagation models for overlay operations focus on either the spatial or the thematic data component. Griffith et al. (1999) provides a study on map error and its propagation. Lanter and Veregin (1992) describe a research paradigm for propagating error in layer-based GIS. A general discussion on error analysis and propagation in GIS can be found in Lunetta et al. (1991). Veregin (1995), in particular, examines the issue of error propagation in the context of GIS overlay operations and proposes a model which is based on the propagation of the entire classification error matrix (CEM).

In MBGIS, uncertainty about positions rather than attributes is of primary concern~~s~~. The position of a polygon is determined by its boundaries, which are made of line segments. When a polygon is convex, the line-in-polygon operation can be reduced to the point-in-polygon operation since a line segment is inside a polygon as long as its endpoints are inside the polygon. When a polygon is non-convex, it becomes a question of whether or not the line segment intersects the polygon. Naturally the study of intersection becomes a key step in line-in-polygon and polygon-on-polygon operations. Under the effect of measurement error (ME), point-in-polygon analysis has been studied in Leung et al. (2003b). We continue in this part to analyze line-in-polygon and polygon-on-polygon analyses under ME in MBGIS.

Although there exists a variety of algorithms for the intersection of line segments in computational geometry and GIS (Berg et al., 2000; Rigaux et al., 2002; Wise, 2002), error analysis in intersections has seldom been made in the literature and the analytic expression of error propagation from the endpoints to the intersections has not been developed. The purpose of this paper is to give a formal analysis of these problems, with a substantiation by simulation experiments.

We first perform error analysis in intersections and establish an approximate law of error propagation for intersections in Section 2. Based on the derived results, error analysis in polygon-on-polygon overlay ~~are~~ is then performed and the corresponding approximate law of error propagation is derived in Section 3. Theoretical results are substantiated by simulation experiments. We then conclude the paper with a summary in Section 4.

2. Error analysis on position measurement of intersection points

In this section, we first give an analytic expression for the intersection point of two line segments and then derive the approximate law of error propagation from the endpoints of the line segments to the intersection point.

2.1 Identification and analytic expressions for an intersection point

Suppose that $V_i(\mathbf{x}_i)$, $i=1,2,3,4$, are the endpoints of two line segments V_1V_2 and V_3V_4 (For simplicity and without confusion, we henceforth use V_i for both the singular and plural form of V_i

(i.e., rather than using V_i's for plural), and the same applies to all other relevant symbols), and $V_c(\mathbf{x}_c)$ is the intersection point of line segments V_1V_2 and V_3V_4, where $\mathbf{x}_i \equiv (x_{i1}, x_{i2})^{\mathrm{T}}$ and $\mathbf{x}_c \equiv (x_{c1}, x_{c2})^{\mathrm{T}}$ are column vectors of the corresponding coordinates. To establish the relationship between $\mathbf{x}_i$ and $\mathbf{x}_c$, let $\mathbf{x}_{(4)}$ be the joint (augmented) vector of $\mathbf{x}_i$, i.e.,

$$\mathbf{x}_{(4)}^{\mathrm{T}} \equiv (\mathbf{x}_1^{\mathrm{T}} \ \mathbf{x}_2^{\mathrm{T}} \ \mathbf{x}_3^{\mathrm{T}} \ \mathbf{x}_4^{\mathrm{T}}), \tag{2.1}$$

where the subscript (4) indicates that there are four points.

Now we derive the analytic expressions of $\mathbf{x}_c$ by $\mathbf{x}_{(4)}$. It is well known that if the line segments V_1V_2 and V_3V_4 intersect, $\mathbf{x}_c$ can then be expressed as

$$\mathbf{x}_c = \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2, \ \ \lambda_1 + \lambda_2 = 1, \ \ \lambda_i \geq 0 \ , \ i = 1, 2, \tag{2.2}$$

and

$$\mathbf{x}_c = \lambda_3\mathbf{x}_3 + \lambda_4\mathbf{x}_4, \ \ \lambda_3 + \lambda_4 = 1, \ \ \lambda_i \geq 0 \ , \ i = 3, 4, \tag{2.3}$$

where λ_i are real numbers. From (2.2) and (2.3), we can obtain an equation with respect to λ_1 and λ_3: $\lambda_1(\mathbf{x}_1 - \mathbf{x}_2) + \lambda_3(\mathbf{x}_4 - \mathbf{x}_3) = \mathbf{x}_4 - \mathbf{x}_2$, i.e.,

$$\begin{pmatrix} \mathbf{x}_1 - \mathbf{x}_2 & \mathbf{x}_4 - \mathbf{x}_3 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_3 \end{pmatrix} = \mathbf{x}_4 - \mathbf{x}_2 .$$

By solving it, we obtain

$$\begin{cases} \lambda_1^* = \dfrac{f_{\det}(\mathbf{x}_4 - \mathbf{x}_2, \mathbf{x}_3 - \mathbf{x}_4)}{f_{\det}(\mathbf{x}_1 - \mathbf{x}_2, \mathbf{x}_3 - \mathbf{x}_4)}, & (2.4) \\ \lambda_3^* = \dfrac{f_{\det}(\mathbf{x}_1 - \mathbf{x}_2, \mathbf{x}_2 - \mathbf{x}_4)}{f_{\det}(\mathbf{x}_1 - \mathbf{x}_2, \mathbf{x}_3 - \mathbf{x}_4)}, & (2.5) \end{cases}$$

where $f_{\det}$ is defined in Leung et al. (2003b) as

$$f_{\det}(\mathbf{x}, \mathbf{y}) \equiv \det(\mathbf{x}, \mathbf{y}) = \mathbf{x}^{\mathrm{T}}\mathbf{H}_0\mathbf{y}, \quad \mathbf{H}_0 \equiv \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.6}$$

To denote $\mathbf{x}_i$ by $\mathbf{x}_{(4)}$, let $\mathbf{e}_{n,i}$ be a $n \times 1$ unit column vector whose ith component takes on the value 1, and 0 elsewhere, i.e.,

$$\underset{n\times 1}{\mathbf{e}_{n,i}} \equiv (\underset{1}{0} \ \underset{\cdots}{\cdots} \ \underset{i-1}{0} \ \underset{i}{1} \ \underset{i+1}{0} \ \underset{\cdots}{\cdots} \ \underset{n}{0})^{\mathrm{T}} . \tag{2.7}$$

Then we have

$$\mathbf{x}_i = \mathbf{D}_i\mathbf{x}_{(4)}, \quad \mathbf{D}_i \equiv \mathbf{e}_{4,i}^{\mathrm{T}} \otimes \mathbf{I}_2, \quad i = 1, 2, 3, 4, \tag{2.8}$$

where the symbol "$\otimes$" denotes the Kronecker product of matrices, and $\mathbf{I}_2$ is a 2×2 identity matrix. Clearly, all $\mathbf{D}_i$ are 2×8 matrices. By simple derivation, the following conclusion can be obtained:

Proposition 2.1 λ_i^* ($i = 1, 2, 3, 4$) determined by (2.2)~(2.5) can be expressed as

$$\lambda_i^* = \lambda_i^*(\mathbf{x}_{(4)}) \equiv \frac{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_i\mathbf{x}_{(4)}}{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,\mathbf{x}_{(4)}}, \quad i = 1, 2, 3, 4 \tag{2.9}$$

where

$$\overline{\mathbf{H}}_i \equiv \overline{\mathbf{\Delta}}_i \otimes \mathbf{H}_0\,,\quad \overline{\mathbf{H}} \equiv \overline{\mathbf{\Delta}} \otimes \mathbf{H}_0\,, \tag{2.10}$$

$$\overline{\mathbf{\Delta}}_1 \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix},\quad \overline{\mathbf{\Delta}}_2 \equiv \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 \end{pmatrix},\quad \overline{\mathbf{\Delta}}_3 \equiv \begin{pmatrix} 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{pmatrix},\quad \overline{\mathbf{\Delta}}_4 \equiv \begin{pmatrix} 0 & 1 & -1 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{2.11}$$

$$\overline{\mathbf{\Delta}} \equiv \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{pmatrix}. \tag{2.12}$$

The proof is given in Appendix 1.

Remark 1 All of $\overline{\mathbf{\Delta}}_i$ in (2.11) and $\overline{\mathbf{\Delta}}$ in (2.12) are skew-symmetric matrices and $\overline{\mathbf{\Delta}}_i$ can be produced from $\overline{\mathbf{\Delta}}$ according to a certain rule. Furthermore, all of $\overline{\mathbf{H}}_i$ and $\overline{\mathbf{H}}$ in (2.10) are symmetric matrices and $\overline{\mathbf{H}}_1 + \overline{\mathbf{H}}_2 = \overline{\mathbf{H}} = \overline{\mathbf{H}}_3 + \overline{\mathbf{H}}_4$. Therefore, the numerators and denominators of all λ_i^* ($i = 1, 2, 3, 4$) are quadratic forms in the joint vector $\mathbf{x}_{(4)}$. This conclusion has double significance: (1) it provides a possible approach to analyze the statistical distribution of the coordinate vector of the intersection (see Leung et al. (2003d)), and (2) it renders an easier derivation of the approximate law of error propagation for the coordinate vector of the intersection (as shown in the following subsection).

It is interesting to note that the matrices $\overline{\mathbf{\Delta}}_i$ and $\mathbf{\Delta}_i$ defined in Leung et al. (2003b) are the same except $\overline{\mathbf{\Delta}}_3 = -\mathbf{\Delta}_3$, while $\mathbf{\Delta}_i$ are used to determine the quadratic forms for the identification of the relationship between a point and a triangle. Although $\overline{\mathbf{\Delta}}_i$ and $\mathbf{\Delta}_i$ are for different purposes, there exists indeed a natural link between them.

Thus, according to (2.2), (2.8) and (2.9), we can obtain the analytic expressions of $\mathbf{x}_c$ by $\mathbf{x}_{(4)}$ or the transformation function from $\mathbf{x}_{(4)}$ to $\mathbf{x}_c$ as follows:

$$\mathbf{x}_c = f(\mathbf{x}_{(4)}) \equiv \lambda_1^*(\mathbf{x}_{(4)})\,\mathbf{x}_1 + \lambda_2^*(\mathbf{x}_{(4)})\,\mathbf{x}_2 = [\lambda_1^*(\mathbf{x}_{(4)})\,\mathbf{D}_1 + \lambda_2^*(\mathbf{x}_{(4)})\,\mathbf{D}_2]\,\mathbf{x}_{(4)}\,. \tag{2.13}$$

From (2.2), (2.3) and (2.9), we have immediately:

Proposition 2.2 Two line segments V_1V_2 and V_3V_4 intersect if and only if all of $\lambda_i^*(\mathbf{x}_{(4)})$ are nonnegative, i.e., $\lambda_i^*(\mathbf{x}_{(4)}) \geq 0$, where $\lambda_i^*(\mathbf{x}_{(4)})$ is defined by (2.9), $i = 1, 2, 3, 4$.

Fig. 2.1 gives a geometrical interpretation of Proposition 2.2, by which the meanings of the sign of $\lambda_i^*(\cdot)$ can be understood from (2.2) or (2.3). For example, for λ_1, we have $\mathbf{x}_c = \mathbf{x}_2 + \lambda_1(\mathbf{x}_1 - \mathbf{x}_2)$.

When $\lambda_1 \geq 0$, the equation $\mathbf{x} = \mathbf{x}_2 + \lambda_1(\mathbf{x}_1 - \mathbf{x}_2)$ represents the ray line which starts from the point $V_2(\mathbf{x}_2)$ and points toward the point $V_1(\mathbf{x}_1)$. If the line segments V_1V_2 and V_3V_4 intersect, the intersection point $V_c(\mathbf{x}_c)$ should be in the ray line and in this case $\lambda_1^*(\mathbf{x}_{(4)}) \geq 0$. Similarly, λ_2, λ_3 and λ_4 can be interpreted likewise.

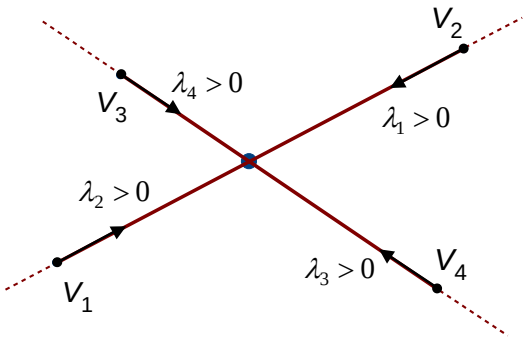


Fig. 2.1 A geometrical interpretation for Proposition 2.2

2.2 Approximate law of error propagation in the intersection coordinates

Suppose that $V_1^0(\boldsymbol{\mu}_1)$ and $V_2^0(\boldsymbol{\mu}_2)$ are the endpoints of a true line segment, $V_3^0(\boldsymbol{\mu}_3)$ and $V_4^0(\boldsymbol{\mu}_4)$ are the endpoints of the other line segment, and $V_c^0(\boldsymbol{\mu}_c)$ is their intersection point, where $\boldsymbol{\mu}_i \equiv (\mu_{i1}, \mu_{i2})^{\mathrm{T}}$ and $\boldsymbol{\mu}_c \equiv (\mu_{c1}, \mu_{c2})^{\mathrm{T}}$ are the corresponding coordinate vectors, $i = 1, 2, 3, 4$. Under ME, what we really observed are the random points $V_i(\mathbf{X}_i)$, $\mathbf{X}_i \equiv (X_{i1}, X_{i2})^{\mathrm{T}}$, $i = 1, 2, 3, 4$, and their coordinate vectors can be represented by the following ME model:

$$\mathbf{X}_i = \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_i, \ i = 1, 2, 3, 4, \tag{2.14}$$

where the ME vectors $\boldsymbol{\varepsilon}_i \equiv (\varepsilon_{i1}, \ \varepsilon_{i2})^{\mathrm{T}} \sim (\mathbf{0}, \boldsymbol{\Sigma}_i)$, $i = 1, 2, 3, 4$. Since randomness of the endpoints of the two line segments determines the randomness of their intersection point V_c, its positional coordinate vector $\mathbf{X}_c \equiv (X_{c1}, X_{c2})^{\mathrm{T}}$ is random and can also be written as $\mathbf{X}_c = \boldsymbol{\mu}_c + \boldsymbol{\varepsilon}_c$, where the error vector $\boldsymbol{\varepsilon}_c$ is determined by the ME vectors $\boldsymbol{\varepsilon}_i$, $i = 1, 2, 3, 4$, and its covariance matrix is denoted by $\boldsymbol{\Sigma}_c \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_c)$. Our purpose is to establish the relationship between $\boldsymbol{\Sigma}_c$ and $\boldsymbol{\Sigma}_i$ ($i = 1, 2, 3, 4$).

The expression (2.14) is readily available and convenient when all endpoints are independent. However, it is usual that the ME vectors of the endpoint coordinates are not independent, especially for the overlaid polygons. To study the general case, we need to consider the joint (augmented) positional vector $\mathbf{X}_{(4)}$:

$$\mathbf{X}_{(4)}^{\mathrm{T}} \equiv (\mathbf{X}_1^{\mathrm{T}} \ \mathbf{X}_2^{\mathrm{T}} \ \mathbf{X}_3^{\mathrm{T}} \ \mathbf{X}_4^{\mathrm{T}}), \tag{2.15}$$

and let $\mathbf{\Sigma}_{(4)} \equiv \mathrm{cov}(\mathbf{X}_{(4)})$ be the covariance matrix of $\mathbf{X}_{(4)}$, and $\boldsymbol{\mu}_{(4)}$ and $\boldsymbol{\varepsilon}_{(4)}$ be the corresponding joint true coordinate and ME vectors, with $\boldsymbol{\mu}_{(4)}^{\mathrm{T}} \equiv (\boldsymbol{\mu}_1^{\mathrm{T}} \ \boldsymbol{\mu}_2^{\mathrm{T}} \ \boldsymbol{\mu}_3^{\mathrm{T}} \ \boldsymbol{\mu}_4^{\mathrm{T}})$, and $\boldsymbol{\varepsilon}_{(4)}^{\mathrm{T}} \equiv (\boldsymbol{\varepsilon}_1^{\mathrm{T}} \ \boldsymbol{\varepsilon}_2^{\mathrm{T}} \ \boldsymbol{\varepsilon}_3^{\mathrm{T}} \ \boldsymbol{\varepsilon}_4^{\mathrm{T}})$. Then the corresponding ME model can be formulated in line with the basic ME model in Leung et al. (2003a) as:

$$(\mathrm{I_c}): \quad \begin{cases} \mathbf{X}_c = f(\mathbf{X}_{(4)}) = [\lambda_1^*(\mathbf{X}_{(4)})\,\mathbf{D}_1 + \lambda_2^*(\mathbf{X}_{(4)})\,\mathbf{D}_2]\,\mathbf{X}_{(4)}, & (2.16) \\ \mathbf{X}_{(4)} = \boldsymbol{\mu}_{(4)} + \boldsymbol{\varepsilon}_{(4)}, \ \boldsymbol{\varepsilon}_{(4)} \sim (\mathbf{0}, \ \mathbf{\Sigma}_{(4)}), & (2.17) \end{cases}$$

where $\mathbf{\Sigma}_{(4)} = \mathrm{cov}(\mathbf{X}_{(4)}) = \mathrm{cov}(\boldsymbol{\varepsilon}_{(4)})$. Our problem is thus to search for the relationship between $\mathbf{\Sigma}_c$ and $\mathbf{\Sigma}_{(4)}$.

Remark 2. Whether or not the ME vectors of the endpoints of the two line segments are independent, model ($\mathrm{I_c}$) can always be utilized since it is based on the general joint ME covariance matrix $\mathbf{\Sigma}_{(4)}$. When $(\boldsymbol{\varepsilon}_1^{\mathrm{T}} \ \boldsymbol{\varepsilon}_2^{\mathrm{T}})^{\mathrm{T}}$ and $(\boldsymbol{\varepsilon}_3^{\mathrm{T}} \ \boldsymbol{\varepsilon}_4^{\mathrm{T}})^{\mathrm{T}}$ are independent, it becomes a two-block-diagonal matrix. Furthermore, if the ME vectors of all endpoints (i.e., all $\boldsymbol{\varepsilon}_i$) are independent, it becomes a four-block-diagonal matrix. Especially, if the ME vectors of all endpoints are independent and the MEs of the coordinates of each endpoint are independent, it becomes a diagonal matrix. All of these situations are special cases of the model ($\mathrm{I_c}$).

Since the transformation function *f* in (2.13) is nonlinear in $\mathbf{x}_{(4)}$, it may be difficult to give the exact law of error propagation for $\mathbf{X}_c$ when its statistical distribution cannot be given yet. Indeed, although $\lambda_i^*(\mathbf{X}_{(4)})$ can be expressed by the quotient of two quadratic forms in the joint vector $\mathbf{X}_{(4)}$, we cannot at present obtained the distribution of $\mathbf{X}_c = f(\mathbf{X}_{(4)})$. However, using (2.16) and the approximate law of error propagation for nonlinear function in Leung et al. (2003a), we can derive an approximate law of error propagation for $\mathbf{X}_c = f(\mathbf{X}_{(4)})$.

To get the Jacobian matrix $\mathbf{B}_{\mu_{(4)}}$ of f at $\boldsymbol{\mu}_{(4)}$, the two sides in (2.13) are differentiated with respect to $\mathbf{x}_{(4)}$. Then we have

$$\mathrm{d}\,\mathbf{x}_c = \mathrm{d}\,f(\mathbf{x}_{(4)}) = [\lambda_1^*(\mathbf{x}_{(4)})\mathbf{D}_1 + \lambda_2^*(\mathbf{x}_{(4)})\mathbf{D}_2](\mathrm{d}\,\mathbf{x}_{(4)}) + \mathbf{D}_1\mathbf{x}_{(4)}\,[\mathrm{d}\,\lambda_1^*(\mathbf{x}_{(4)})] + \mathbf{D}_2\mathbf{x}_{(4)}\,[\mathrm{d}\,\lambda_2^*(\mathbf{x}_{(4)})].$$

Since for any matrix **A** and any column vector **x**,

$$\mathrm{d}(\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}) = (\mathrm{d}\,\mathbf{x})^{\mathrm{T}}\mathbf{A}\mathbf{x} + \mathbf{x}^{\mathrm{T}}\mathbf{A}(\mathrm{d}\,\mathbf{x}) = [(\mathrm{d}\,\mathbf{x})^{\mathrm{T}}\mathbf{A}\mathbf{x}]^{\mathrm{T}} + \mathbf{x}^{\mathrm{T}}\mathbf{A}(\mathrm{d}\,\mathbf{x}) = \mathbf{x}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}} + \mathbf{A})(\mathrm{d}\,\mathbf{x}),$$

$$\begin{aligned}\mathrm{d}\,\lambda_i^*(\mathbf{x}_{(4)}) &= \mathrm{d}\,\frac{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_i\mathbf{x}_{(4)}}{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,\mathbf{x}_{(4)}} = -\frac{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_i\mathbf{x}_{(4)}}{(\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,\mathbf{x}_{(4)})^2}\mathrm{d}\,(\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\mathbf{x}_{(4)}) + \frac{1}{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,\mathbf{x}_{(4)}}\mathrm{d}\,(\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_i\mathbf{x}_{(4)}) \\ &= \frac{2}{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,\mathbf{x}_{(4)}}\mathbf{x}_{(4)}^{\mathrm{T}}[\overline{\mathbf{H}}_i - \lambda_i^*(\mathbf{x}_{(4)})\,\overline{\mathbf{H}}](\mathrm{d}\,\mathbf{x}_{(4)}).\end{aligned}$$

Thus,

$$\begin{aligned}\mathrm{d}\,\mathbf{x}_c &= [\lambda_1^*(\mathbf{x}_{(4)})\mathbf{D}_1 + \lambda_2^*(\mathbf{x}_{(4)})\mathbf{D}_2](\mathrm{d}\,\mathbf{x}_{(4)}) + \frac{2}{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,\mathbf{x}_{(4)}}[\mathbf{D}_1\mathbf{x}_{(4)}\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_1 + \mathbf{D}_2\mathbf{x}_{(4)}\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_2](\mathrm{d}\,\mathbf{x}_{(4)}) \\ &\quad - \frac{2}{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,\mathbf{x}_{(4)}}[\lambda_1^*(\mathbf{x}_{(4)})\,\mathbf{D}_1 + \lambda_2^*(\mathbf{x}_{(4)})\,\mathbf{D}_2]\,\mathbf{x}_{(4)}\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,(\mathrm{d}\,\mathbf{x}_{(4)}),\end{aligned}$$

$$\begin{aligned}\mathbf{B}_{\mu_{(4)}} &= \frac{2}{\boldsymbol{\mu}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,\boldsymbol{\mu}_{(4)}}[\mathbf{D}_1\boldsymbol{\mu}_{(4)}\boldsymbol{\mu}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_1 + \mathbf{D}_2\boldsymbol{\mu}_{(4)}\boldsymbol{\mu}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_2] \\ &\quad + [\lambda_1^*(\boldsymbol{\mu}_{(4)})\,\mathbf{D}_1 + \lambda_2^*(\boldsymbol{\mu}_{(4)})\,\mathbf{D}_2]\left(\mathbf{I}_8 - \frac{2}{\boldsymbol{\mu}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,\boldsymbol{\mu}_{(4)}}\boldsymbol{\mu}_{(4)}\boldsymbol{\mu}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\right).\end{aligned} \tag{2.18}$$

Then from the approximate law of error propagation in Leung et al. (2003a), we obtain

Proposition 2.3 Under the ME model ($\mathrm{I_c}$), the approximate law of error propagation for the intersection point $V_c(\mathbf{X}_c)$ of two line segments $V_1(\mathbf{X}_1)V_2(\mathbf{X}_2)$ and $V_3(\mathbf{X}_3)V_4(\mathbf{X}_4)$ is

$$\boldsymbol{\Sigma}_c \approx \widetilde{\boldsymbol{\Sigma}}_c = \widetilde{F}(\boldsymbol{\Sigma}_{(4)}; \boldsymbol{\mu}_{(4)}) \equiv \mathbf{B}_{\mu_{(4)}}\boldsymbol{\Sigma}_{(4)}\mathbf{B}_{\mu_{(4)}}^{\mathrm{T}}, \tag{2.19}$$

where $\mathbf{B}_{\mu_{(4)}}$ is a 2×8 matrix given by (2.18), $\boldsymbol{\Sigma}_{(4)} = \mathrm{cov}(\mathbf{X}_{(4)}) = \mathrm{cov}(\boldsymbol{\varepsilon}_{(4)})$ is the covariance matrix of the joint vector $\mathbf{X}_{(4)}$ (or the joint ME vector $\boldsymbol{\varepsilon}_{(4)}$), and $\boldsymbol{\Sigma}_c$ is the covariance matrix of the coordinate vector $\mathbf{X}_c$ (or error vector $\boldsymbol{\varepsilon}_c$) of the intersection.

In particular, when the ME vectors $\boldsymbol{\varepsilon}_i$ of the endpoints V_i are independent, $\boldsymbol{\Sigma}_{(4)}$ is a block-diagonal matrix and it can be expressed as $\boldsymbol{\Sigma}_{(4)} = \mathrm{diag}(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \boldsymbol{\Sigma}_3, \boldsymbol{\Sigma}_4)$, where $\boldsymbol{\Sigma}_i$ is the covariance matrix of $\boldsymbol{\varepsilon}_i$. However, when a polygon overlay operation is performed, the ME vectors of the endpoints of some edges on the overlaid polygon will be dependent. It results in the study of $\boldsymbol{\Sigma}_c$ in the general case $\boldsymbol{\Sigma}_{(4)}$.

A similar conclusion to (2.19) based on $\lambda_3^*(\mathbf{x}_{(4)})$ and $\lambda_4^*(\mathbf{x}_{(4)})$ can be obtained using (2.3). A natural question is whether these two conclusions are consistent since $\mathbf{B}_{\mu_{(4)}}$ in (2.18) is different if $\mathbf{X}_{(4)}$ is replaced by $\mathbf{X}'_{(4)}$, where $\mathbf{X}'^{\mathrm{T}}_{(4)} \equiv (\mathbf{X}_3^{\mathrm{T}}\ \mathbf{X}_4^{\mathrm{T}}\ \mathbf{X}_2^{\mathrm{T}}\ \mathbf{X}_1^{\mathrm{T}})$. In fact, although $\mathbf{B}_{\mu_{(4)}}$ is different, the conclusions are the same, i.e., $\tilde{\boldsymbol{\Sigma}}_c = \tilde{\boldsymbol{\Sigma}}'_c$. The theoretical argument for the general cases is given in the following proposition whose proof is given in Appendix 2.

Proposition 2.4 Under the condition that the pair relation of $\mathbf{X}_i$ in $\mathbf{X}_{(4)}$ (see (2.15)) is unchanged, for any permutation of $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4)$ in $\mathbf{X}_{(4)}$, the corresponding approximate covariance matrix $\tilde{\boldsymbol{\Sigma}}_c$ is invariant. That is, if $\mathbf{X}_{(4)}$ is changed into $\mathbf{X}'_{(4)}$ under the said condition, we still have $\tilde{\boldsymbol{\Sigma}}_c = \tilde{\boldsymbol{\Sigma}}'_c$.

Remark 3 This proposition shows that $\tilde{\boldsymbol{\Sigma}}_c$ in (2.19) is uniquely determined as long as the given partitioning of four points is formed into two line segments. In other words, $\tilde{\boldsymbol{\Sigma}}_c$ only depends on the grouping and is independent of the arranged order of $\mathbf{X}_i$ in $\mathbf{X}_{(4)}$.

2.3 Simulations

We employ several simulation experiments under different ME conditions to show the effect of endpoints ME on the coordinates of the intersection point and the effectiveness of the approximate law of error propagation (2.19). Example 2.1 shows the situation when all endpoints have the same error structures (i.e., circular covariance matrix). Example 2.2 demonstrates the situation when locations of endpoints are varied and the endpoints have different and ~~same~~ equal error structures.

Example 2.1 (Endpoints ME with the same error structure) Assume that four true endpoints of two line segments $V_1^0V_2^0$ and $V_3^0V_4^0$ are V_1^0 (0, 0), V_2^0 (3, –1), V_3^0 (1, 2), V_4^0 (2, –2) (Fig. 2.1(a)). Thus the true joint vector is $\boldsymbol{\mu}_{(4)} = (0\ 0\ 3\ -1\ 1\ 2\ 2\ -2)^{\mathrm{T}}$. By calculation, the true intersection point is V_c^0 (1.636364, –0.5454545). MEs of the four endpoints are assumed to be independently and identically distributed as a bivariate normal distribution, $\boldsymbol{\varepsilon}_i \sim N_2(\mathbf{0}, \boldsymbol{\Sigma}_i)$, with the same circular covariance matrix $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}_\sigma$, $i = 1, 2, 3, 4$, where

$$\boldsymbol{\Sigma}_{\sigma} \equiv \sigma^2 \mathbf{I}_2 , \tag{2.20}$$

Thus $\boldsymbol{\Sigma}_{(4)} = \mathrm{diag}(\boldsymbol{\Sigma}_{\sigma}, \boldsymbol{\Sigma}_{\sigma}, \boldsymbol{\Sigma}_{\sigma}, \boldsymbol{\Sigma}_{\sigma}) = \sigma^2 \mathbf{I}_8$.

By 1000 simulations in case (1), we obtain the sample covariance matrices $\hat{\boldsymbol{\Sigma}}_c$ of the coordinates of the intersection point for different σ^2 (see Table 2.1). Since $\hat{\boldsymbol{\Sigma}}_c$ is an unbiased estimator of $\boldsymbol{\Sigma}_c$, it can be viewed as a good approximation to $\boldsymbol{\Sigma}_c$. From Table 2.1, it can be observed that $\tilde{\boldsymbol{\Sigma}}_c$ given by (2.19) is also a good approximation to $\boldsymbol{\Sigma}_c$, especially when the error variance σ^2 of the endpoints is smaller. The simulation line segments and the corresponding intersections of 100 samples with $\sigma^2 = 0.01$ are shown in Fig. 2.1(b). When $\sigma^2 = 0.01$ and $\sigma^2 = 0.1$, the scatter plots of 1000 sample endpoints and intersection points are depicted in Fig.2.1(c) and (d) respectively.

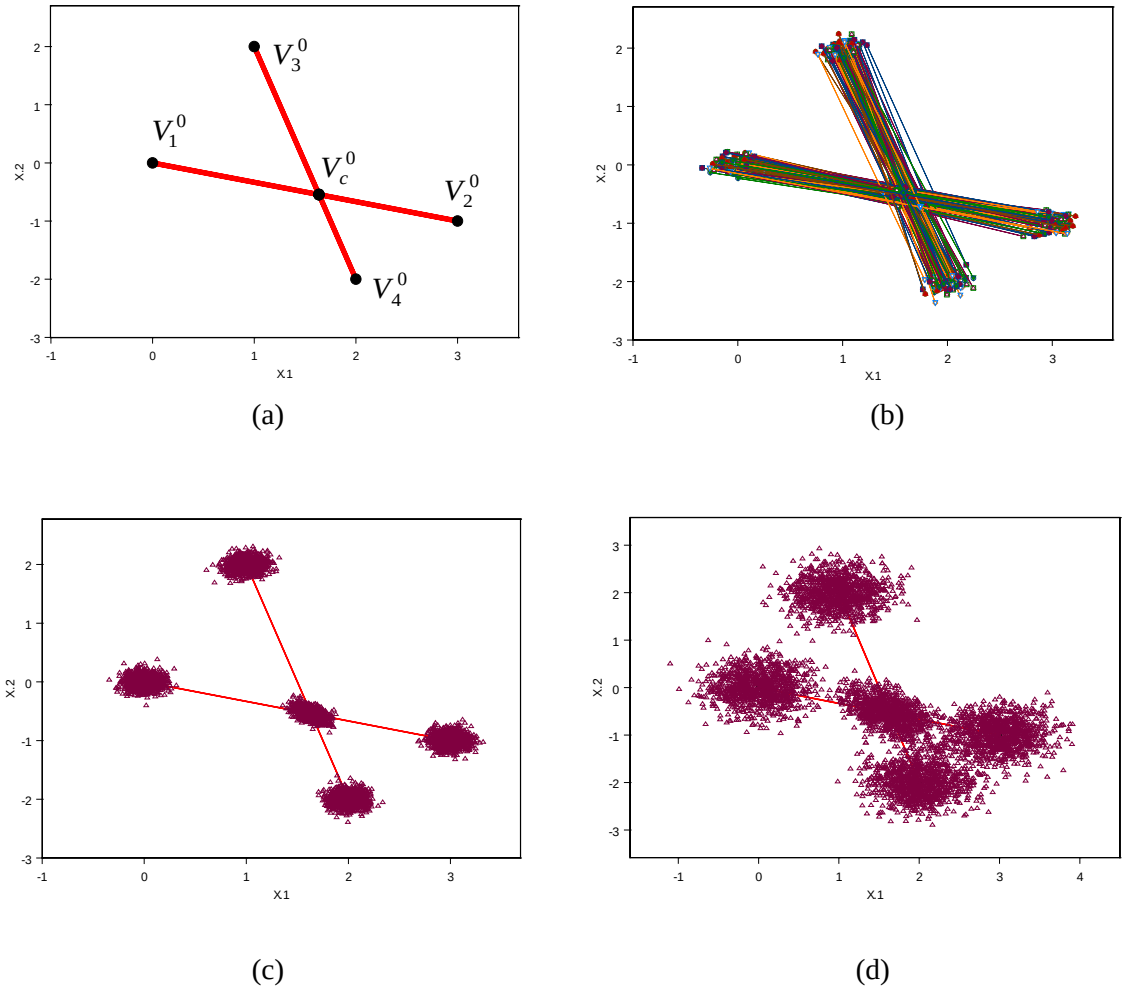


Fig. 2.1 Simulation results of intersection points and endpoints

Table 2.1 Sample covariance matrices and the approximate law of error propagation (2.19)

	$\sigma^2 = 0.01$	$\sigma^2 = 0.02$	$\sigma^2 = 0.05$	$\sigma^2 = 0.1$
$\tilde{\tilde{\boldsymbol{\Sigma}}}_c$	$\begin{pmatrix} 0.0072 & -0.0039 \\ -0.0039 & 0.0074 \end{pmatrix}$	$\begin{pmatrix} 0.0144 & -0.0079 \\ -0.0079 & 0.0148 \end{pmatrix}$	$\begin{pmatrix} 0.0360 & -0.0197 \\ -0.0197 & 0.0371 \end{pmatrix}$	$\begin{pmatrix} 0.0721 & -0.0393 \\ -0.0393 & 0.0742 \end{pmatrix}$
$\hat{\boldsymbol{\Sigma}}_c$	$\begin{pmatrix} 0.0074 & -0.0039 \\ -0.0039 & 0.0075 \end{pmatrix}$	$\begin{pmatrix} 0.0146 & -0.0086 \\ -0.0086 & 0.0153 \end{pmatrix}$	$\begin{pmatrix} 0.0405 & -0.0232 \\ -0.0232 & 0.0388 \end{pmatrix}$	$\begin{pmatrix} 0.0801 & -0.0442 \\ -0.0442 & 0.0815 \end{pmatrix}$

Example 2.2 (Endpoints ME with different error structures) To investigate the effect of varying location on the error structure of the intersection point, we consider two cases as shown in Fig. 2.2(a) and (b). The difference between these two cases is that endpoint V_4^0 (2, –2) in Fig. 2.2(a) is changed into $V_4'^0$ (1.5, –3) in Fig. 2.2(b), while the other endpoints remain the same. Assume that the ME vectors of all endpoints are independent. The results of 1000 simulation runs are depicted in Fig. 2.2(c) and (d) when the ME distribution of the coordinate vectors of endpoints V_1^0 and V_2^0 are (2.20), where $\sigma^2 = 0.05$, and the ME distribution of coordinate vectors of endpoints V_3^0 and V_4^0 (or $V_4'^0$) are the bivariate normal distribution with elliptical covariance matrices

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \tag{2.21}$$

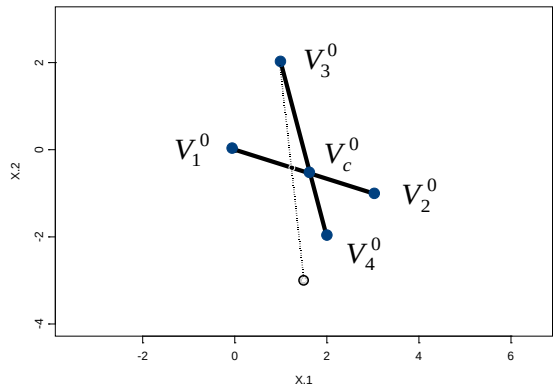
where $\rho = 0.6$, $\sigma_1 = 0.1$, $\sigma_2 = 0.3$ (in this case, $\boldsymbol{\Sigma}_{(4)} = \mathrm{diag}(\boldsymbol{\Sigma}, \boldsymbol{\Sigma}, \boldsymbol{\Sigma}, \boldsymbol{\Sigma}) = \mathbf{I}_4 \otimes \boldsymbol{\Sigma}$).

Table 2.2 Sample covariance matrices and covariance matrices obtained by the approximate law of error propagation for Fig. 2.2

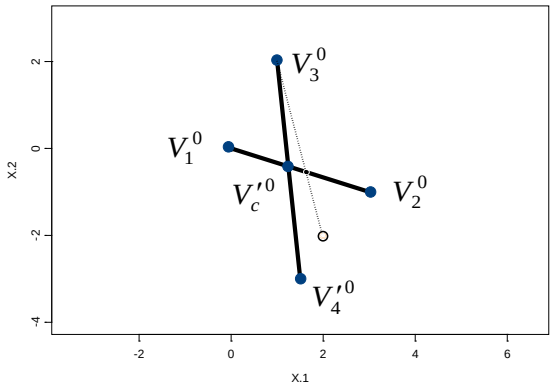
	(c)	(d)	(e)	(f)
$\tilde{\tilde{\boldsymbol{\Sigma}}}_c$	$\begin{pmatrix} 0.0178 & -0.0136 \\ -0.0136 & 0.0351 \end{pmatrix}$	$\begin{pmatrix} 0.0081 & -0.0057 \\ -0.0057 & 0.0315 \end{pmatrix}$	$\begin{pmatrix} 0.0196 & -0.0207 \\ -0.0207 & 0.0636 \end{pmatrix}$	$\begin{pmatrix} 0.0083 & -0.0083 \\ -0.0083 & 0.0577 \end{pmatrix}$
$\hat{\boldsymbol{\Sigma}}_c$	$\begin{pmatrix} 0.0191 & -0.0143 \\ -0.0143 & 0.0369 \end{pmatrix}$	$\begin{pmatrix} 0.0081 & -0.0052 \\ -0.0052 & 0.0340 \end{pmatrix}$	$\begin{pmatrix} 0.0217 & -0.0217 \\ -0.0217 & 0.0659 \end{pmatrix}$	$\begin{pmatrix} 0.0079 & -0.0084 \\ -0.0084 & 0.0576 \end{pmatrix}$

As a comparison, we also give in Fig. 2.2(e) and (f) the 1000 simulation results when the ME distributions of all endpoints are (2.21). All numerical estimate results for the covariance matrices of the intersection points are tabulated in Table 2.2, where the first row represents the covariance matrices

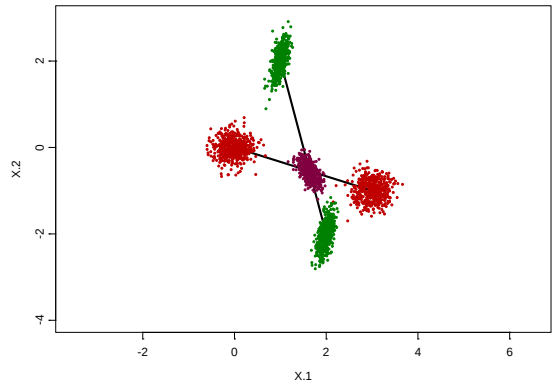
of the intersection points given by the approximate law of error propagation (2.19) and the second row consists of the sample covariance matrix estimates.



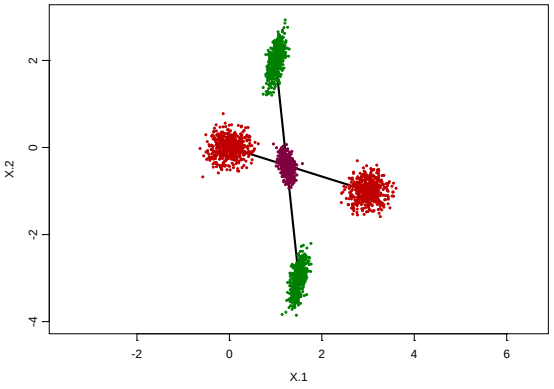
(a) with V_4^0 (2, –2)



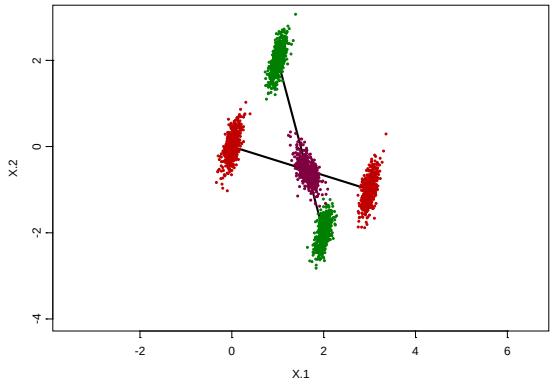
(b) with $V_4'^0$ (1.5, –3)



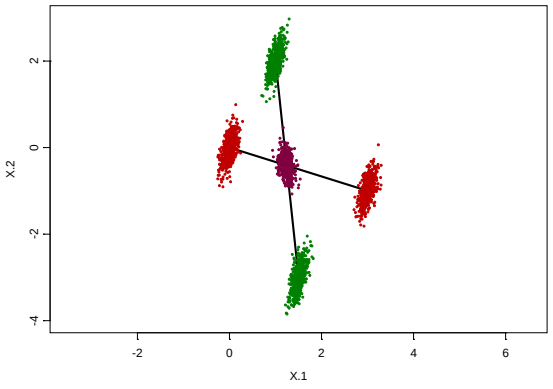
(c) with circular covariance matrices at (V_1^0 , V_2^0) and elliptical covariance matrices at (V_3^0 , V_4^0)



(d) with circular covariance matrices at (V_1^0 , V_2^0) and elliptical covariance matrices at (V_3^0 , $V_4'^0$)



(e) with elliptical covariance matrices at all endpoints



(f) with elliptical covariance matrices at all endpoints

Fig. 2.2 Simulation results of varying locations of endpoints

As shown in Example 2.1, it is obvious from Table 2.2 that $\tilde{\boldsymbol{\Sigma}}_c$ and $\hat{\boldsymbol{\Sigma}}_c$ are very similar. We can view $\tilde{\boldsymbol{\Sigma}}_c$ as a good approximation to $\boldsymbol{\Sigma}_c$. It can be observed that the angle $\angle V_2^0 V_c^0 V_4^0$ of the two line segments $V_1^0 V_2^0$ and $V_3^0 V_4^0$ is bigger than the angle $\angle V_2^0 V_c'^0 V_4'^0$ of the two line segments $V_1^0 V_2^0$ and $V_3^0 V_4'^0$. In Table 2.2, columns (c) and (d) correspond respectively to the same error structures of the endpoints in Fig. 2.2 (c) and (d). The resulting difference of the covariance matrix estimates of the intersection points are apparent, especially for $\tilde{\sigma}_{c1}^2$, from 0.0178 to 0.0081. For columns (e) and (f) corresponding to Fig. 2.2(e) and (f), similar results are obtained. Furthermore, the positions of the true line segments corresponding to columns (c) and (e) are the same but the error covariance structures of the endpoints are different. We can see that the difference between the estimated covariance matrices of (c) and (e) is not bigger than (c) and (d). In fact, for $\tilde{\sigma}_{c1}^2$ the former is 0.0196–0.0178 = 0.0018, while the latter is 0.0178–0.0081 = 0.0097 and is about 5 times that of the former. However, for the same position of the true line segments with the error structure (2.20) (where $\sigma^2 = 0.05$), the value of $\tilde{\sigma}_{c1}^2$ is 0.0360 in Table 2.1 and the difference between this result and (c) in Table 2.2 is 0.0360–0.0178 = 0.0182, which is about 2 times the difference of (c) and (d). Therefore, in general, we cannot say which (the position or error structure of the endpoints) has a bigger effect on the covariance structure of the intersection point.

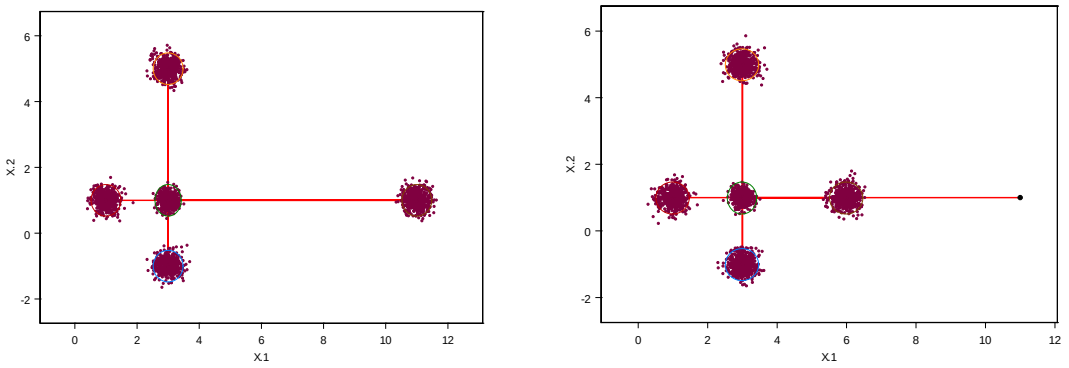
To further investigate the effects of different positions and error covariance matrices of the endpoints on the intersection point when two line segments are perpendicular, a series of simulations have been performed under the assumption that all endpoints are independent and have one of the following ME structures: $\boldsymbol{\varepsilon}_i \sim N_2(\mathbf{0}, \boldsymbol{\Sigma})$, $i = 1, 2, 3, 4$,

(i) $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_\sigma = \sigma^2 \mathbf{I}_2$, $\sigma^2 = 0.05$;

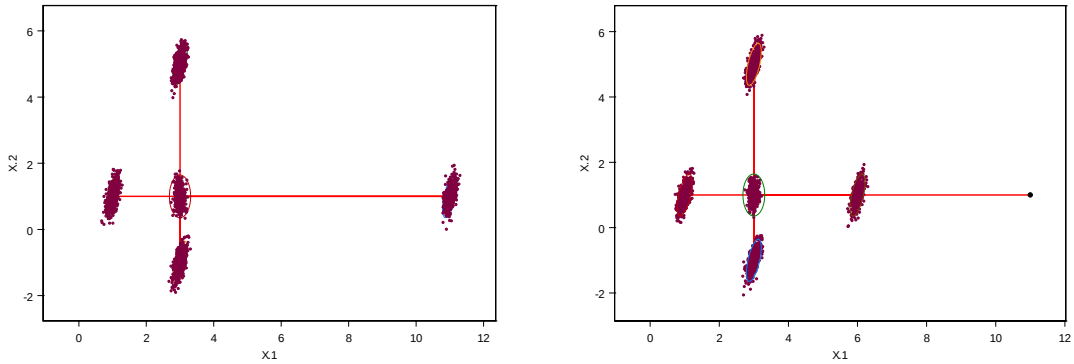
(ii) $\boldsymbol{\Sigma}$ is given by (2.21), where $\rho = 0.6$, $\sigma_1 = 0.1$, $\sigma_2 = 0.3$;

(iii) $\boldsymbol{\Sigma}$ is given by (2.21), where $\rho = 0.9$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$;

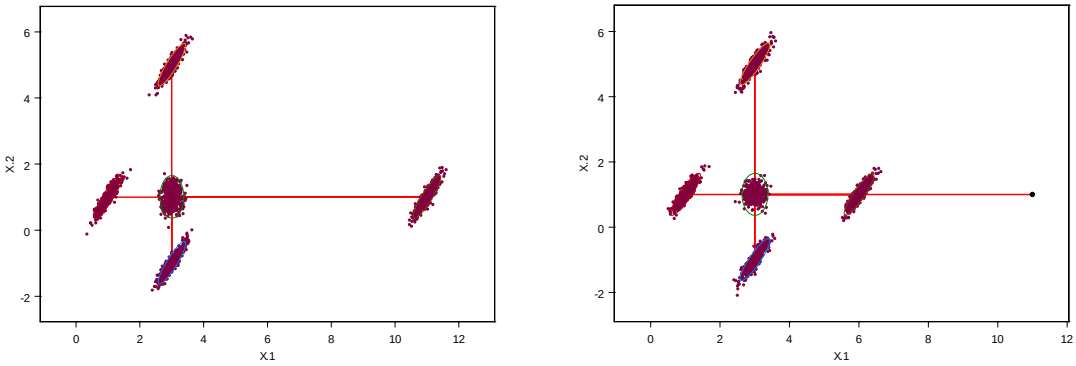
(iv) $\boldsymbol{\Sigma}$ is given by (2.21), where $\rho = -0.9$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$.



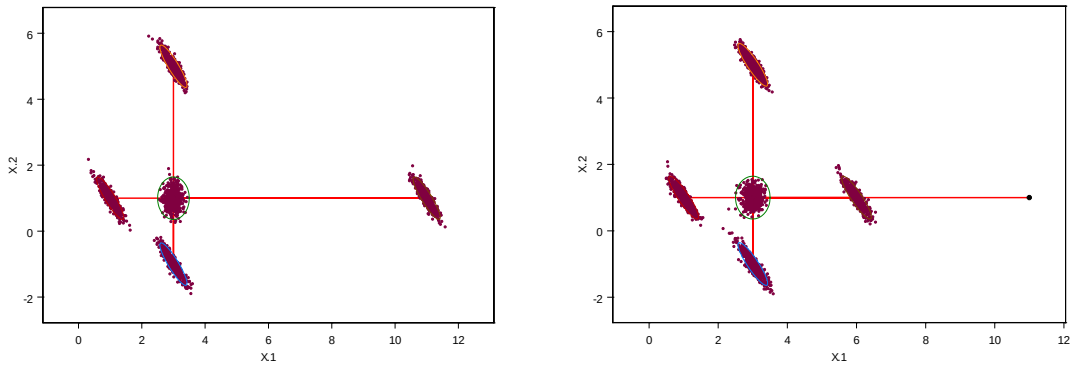
(1) Circular ME (i)



(2) Elliptical ME (ii)



(3) Elliptical ME (iii)



(4) Elliptical ME (iv)

Fig. 2.3 Simulation experiments with different error structures and positionings of the endpoints

The simulation results are depicted in Fig. 2.3. In each case of the simulations, the sample size is 500, the level of the confidence ellipse of the intersection point is 0.9, and its covariance matrix is given by the approximate law of error propagation (2.19). It can be observed that in this experiment, effect of the endpoint error structures on the covariance matrix of the intersection point seems to be bigger than that of the positions of the endpoints.

This series of simulation experiments thus reveals the complexity of error propagation from the endpoints of the line segments to the intersection point. It is perhaps a reason for our not being able to derive an exact law of error propagation for line intersection at the present moment. It is however apparent that although the positions and error structures of the endpoints are different, the approximate law of error propagation (2.19) can always capture the error characteristics of the corresponding intersection points.

3. Error propagation in polygon-on-polygon analysis

Due to its methodological and technical complexities, developing error propagation models for overlay operations, especially on vector-based data, has seldom been attempted. Error propagation in the polygon-on-polygon operation is a typical example. Based on the results of Section 2, we discuss in this section the error propagation problem in polygon-on-polygon overlay and construct a formal model.

We consider directed polygons P_i with n_i vertices $V_j^{(i)}$ listed in a counter-clockwise manner: $V_1^{(i)}, \cdots, V_j^{(i)}, V_{j+1}^{(i)}, \cdots, V_{n_i}^{(i)}, V_{n_i+1}^{(i)} = V_1^{(i)}$, $i = 1, 2$ (see Fig. 3.1), where the superscript (i) denotes the *i*th polygon and the subscript j denotes the *j*th vertex. Assume that vertices $V_j^{(i)}$ have the coordinates vectors $\mathbf{X}_j^{(i)}$ under ME, their true coordinates vectors are $\boldsymbol{\mu}_j^{(i)}$, and the corresponding ME vectors are $\boldsymbol{\varepsilon}_j^{(i)}$, $j = 1, \ldots, n_i$, $i = 1, 2$. Then

$$\mathbf{X}_j^{(i)} = \boldsymbol{\mu}_j^{(i)} + \boldsymbol{\varepsilon}_j^{(i)}, \quad j = 1, \ldots, n_i, \; i = 1, 2. \tag{3.1}$$

It is obvious that the first task is to identify the intersection points of these two polygons. Computing the intersection of two polygons is a most useful operation, not only in spatial databases but in computer graphics and CAD/CAM applications. In the literature, usual methods for detecting an intersection in a set of edges are mostly sweep-line techniques (Rigaux, 2002; Berg, 2000), which are a

Boolean operation in GIS. One of ~~its~~ their advantages is that we do not need to test each edge against all edges of the other polygon. It may be fast in time, but its disadvantage is that we do not have analytic expressions for error analysis.

In our study, a natural way is to use Proposition 2.2 for each pair of edges of the polygons concerned. Such a method may have higher computational cost when the number of vertices is large. However, we can improve it by a simpler test. The key step to improve it is to avoid testing all pairs of edges for intersection. Intuitively, unlike edges that are far apart, edges that are close together are candidates for intersection. By the method in Berg (2000), we can define the *y*-intervals of an edge to be its orthogonal projection onto the *y*-axis (see Fig. 3.2). Then project all edges onto the *y*-axis and observe whether the *y*-intervals of a pair of edges overlap. If they do not overlap, the pair of edges does not intersect. Hence, we only need to test pairs of edges whose *y*-intervals overlap. Another method is to use a bounding box or minimum enclosing rectangle (MER)(Wise, 2002), which is a box around each edge just neatly enclosing it (see Fig. 3.3). The MER is defined by just four numbers, i.e., the values of the *x*- and *y*-coordinates of the endpoints of the edges. If the MERs for a pair of edges do not intersect, then the pair of edges cannot intersect and there is no need to do any further test. These two methods are also suitable for detecting whether two lines, each made up of a series of segments, intersect.

When the edge $V_i^{(1)}V_{i+1}^{(1)}$ of the polygon P_1 and the edge $V_j^{(2)}V_{j+1}^{(2)}$ of the polygon P_2 intersect, their intersection point is denoted by $V_{i,j}$. For simplicity, we restrict our analysis to the most usual case in which there are only two intersection points. Let the intersection points be $V_{i,j}$ and $V_{k,l}$. First, consider $V_{i,j}$ and test whether its former point $V_i^{(1)}$ on P_1 is inside P_2. If no, then $V_{i+1}^{(1)}$ is inside P_2, consequently $V_k^{(1)}$ is inside P_2 and $V_{k+1}^{(1)}$ is outside P_2 ; elsewise, $V_{i+1}^{(1)}$ is outside P_2, and $V_k^{(1)}$ is outside P_2 and $V_{k+1}^{(1)}$ is inside P_2. Second, test likewise whether its former point $V_j^{(2)}$ on P_2 is inside P_1. Finally, we can obtain the vertex sets for the overlaid polygons $P_1 - P_2$, $P_1 \cap P_2$, and $P_2 - P_1$. For example,

$$P_1 - P_2: \quad V_1^{(1)}, \cdots, V_i^{(1)}, V_{i,j}, V_j^{(2)}, V_{j-1}^{(2)}, \cdots, V_{l+1}^{(2)}, V_{k,l}, V_{k+1}^{(1)}, \cdots, V_{n_1}^{(1)}, \tag{3.2}$$

$$P_1 \cap P_2: \quad V_{i,j}, V_{i+1}^{(1)}, \cdots, V_k^{(1)}, V_{k,l}, V_{l+1}^{(2)}, \cdots, V_j^{(2)}, \tag{3.3}$$

$$P_2 - P_1: \quad V_1^{(2)}, \cdots, V_l^{(2)}, V_{k,l}, V_k^{(1)}, V_{k-1}^{(1)}, \cdots, V_{i+1}^{(1)}, V_{i,j}, V_{j+1}^{(2)}, \cdots, V_{n_2}^{(2)}. \tag{3.4}$$

Now we investigate the ME of vertices in these overlaid polygons and study error propagation of polygon overlay within our framework. Each polygon consists of three classes of points coming from P_1, P_2, and the intersection points respectively. The approximate covariance matrix of the ME of the intersection points can be obtained by (3.1) and (2.19). Thus, the ME of each overlaid polygon will be heterogeneous if the MEs of the original polygons are homogeneous. Furthermore, the ME of vertices of each overlaid polygon will be dependent statistically even if the MEs of vertices of the original polygons are independent statistically.

Assume that the overlaid polygons are represented by (3.2)~(3.4). Let $\boldsymbol{\varepsilon}_{\{1-2\}}$, $\boldsymbol{\varepsilon}_{\{2-1\}}$, $\boldsymbol{\varepsilon}_{\{1,2\}}$, $\boldsymbol{\varepsilon}_{\{1\}}$, and $\boldsymbol{\varepsilon}_{\{2\}}$ be respectively the joint ME vectors of the vertices coordinates (counter-clockwisely listed) of the overlaid polygons $P_1 - P_2$, $P_2 - P_1$, $P_1 \cap P_2$, P_1, and P_2, where the elements of $\boldsymbol{\varepsilon}_{\{1-2\}}$, $\boldsymbol{\varepsilon}_{\{2-1\}}$, and $\boldsymbol{\varepsilon}_{\{1,2\}}$ are arranged by the order (3.2)~(3.4); $n_{\{1-2\}}$, $n_{\{2-1\}}$, and $n_{\{1,2\}}$ are the numbers of vertices of $P_1 - P_2$, $P_2 - P_1$, $P_1 \cap P_2$. Then by (3.2)~(3.4)

$$n_{\{1-2\}} = i + 1 + (j - l) + 1 + (n_1 - k) = n_1 + 2 + (i + j) - (k + l),$$

$$n_{\{1,2\}} = 1 + (k - i) + 1 + (j - l) = 2 + (j + k) - (i + l),$$

$$n_{\{2-1\}} = l + 1 + (k - i) + 1 + (n_2 - j) = n_2 + 2 + (k + l) - (i + j).$$

Let

$$\boldsymbol{\Sigma}_{\{1\}} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_{\{1\}}), \quad \boldsymbol{\Sigma}_{\{2\}} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_{\{2\}}), \tag{3.5}$$

$$\boldsymbol{\Sigma} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}) = \begin{pmatrix} \boldsymbol{\Sigma}_{\{1\}} & \boldsymbol{\Sigma}_{\{1\}\{2\}} \\ \boldsymbol{\Sigma}_{\{1\}\{2\}} & \boldsymbol{\Sigma}_{\{2\}} \end{pmatrix}, \ \boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_{\{1\}}^{\mathrm{T}}, \boldsymbol{\varepsilon}_{\{2\}}^{\mathrm{T}})^{\mathrm{T}}, \tag{3.6}$$

$$\boldsymbol{\Sigma}_{\{1-2\}} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_{\{1-2\}}), \ \boldsymbol{\Sigma}_{\{2-1\}} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_{\{2-1\}}), \ \boldsymbol{\Sigma}_{\{1,2\}} \equiv \mathrm{cov}(\boldsymbol{\varepsilon}_{\{1,2\}}). \tag{3.7}$$

The purpose of error analysis is to investigate the relation of (3.6) and each one in (3.7) in order to analyze the error propagation problem from the joint ME covariance matrix $\boldsymbol{\Sigma}$ of the original polygons to the error covariance matrices of the overlaid polygons.

It can be observed that if we let $\boldsymbol{\varepsilon}_{i,j}^{(4)} \equiv (\boldsymbol{\varepsilon}_i^{(1)\,\mathrm{T}}, \boldsymbol{\varepsilon}_{i+1}^{(1)\,\mathrm{T}}, \boldsymbol{\varepsilon}_j^{(2)\,\mathrm{T}}, \boldsymbol{\varepsilon}_{j+1}^{(2)\,\mathrm{T}})^{\mathrm{T}}$, then

$$\boldsymbol{\varepsilon}_{i,j}^{(4)} = \mathbf{D}_{i,j}\boldsymbol{\varepsilon}, \ \mathbf{D}_{i,j} \equiv \begin{pmatrix} \mathbf{e}_{n_1+n_2,i}^{\mathrm{T}} \\ \mathbf{e}_{n_1+n_2,i+1}^{\mathrm{T}} \\ \mathbf{e}_{n_1+n_2,n_1+j}^{\mathrm{T}} \\ \mathbf{e}_{n_1+n_2,n_1+j+1}^{\mathrm{T}} \end{pmatrix} \otimes \mathbf{I}_2, \tag{3.8}$$

where $\mathbf{e}_{n,i}$ is given by (2.7). According to (2.18) and the derivations in Section 2, we have

$$\tilde{\boldsymbol{\varepsilon}}_{i,j} = \mathbf{B}_{\mu_{i,j}^{(4)}} \boldsymbol{\varepsilon}_{i.j}^{(4)} = \mathbf{B}_{\mu_{i,j}^{(4)}} \mathbf{D}_{i,j} \boldsymbol{\varepsilon}, \tag{3.9}$$

where $\boldsymbol{\mu}_{i,j}^{(4)} = \mathbf{D}_{i,j}\boldsymbol{\mu}$, $\boldsymbol{\mu}$ is the true coordinate vector corresponding to $\boldsymbol{\varepsilon}$, and $\mathbf{B}_{\mu}$ is given by (2.18). Thus,

$$\tilde{\boldsymbol{\varepsilon}}_{\{1-2\}} = (\boldsymbol{\varepsilon}_1^{(1)\,\mathrm{T}}, \cdots, \boldsymbol{\varepsilon}_i^{(1)\,\mathrm{T}}, \tilde{\boldsymbol{\varepsilon}}_{i,j}^{\mathrm{T}}, \boldsymbol{\varepsilon}_j^{(2)\,\mathrm{T}}, \cdots, \boldsymbol{\varepsilon}_{l+1}^{(2)\,\mathrm{T}}, \tilde{\boldsymbol{\varepsilon}}_{k,l}^{\mathrm{T}}, \boldsymbol{\varepsilon}_{k+1}^{(1)\,\mathrm{T}}, \cdots, \boldsymbol{\varepsilon}_{n_1}^{(1)\,\mathrm{T}})^{\mathrm{T}} = \mathbf{D}_{\{1-2\}}\boldsymbol{\varepsilon}, \tag{3.10}$$

$$\mathbf{D}_{\{1-2\}} \equiv \begin{pmatrix} \mathbf{I}_{2i} & & \mathbf{0} \\ & \mathbf{B}_{\mu_{i,j}^{(4)}} \mathbf{D}_{i,j} & \\ \mathbf{0} & (\mathbf{J}_{j-l} \otimes \mathbf{I}_2) & \mathbf{0} \\ & \mathbf{B}_{\mu_{k,l}^{(4)}} \mathbf{D}_{k,l} & \\ \mathbf{0} & \mathbf{I}_{2(n_1-k)} & \mathbf{0} \end{pmatrix}, \text{ where } \mathbf{J}_i \equiv \begin{pmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \iddots & 1 & 0 \\ 0 & \iddots & \iddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix}_{i\times i}. \tag{3.11}$$

Similarly, we have

$$\tilde{\boldsymbol{\varepsilon}}_{\{1,2\}}^{\mathrm{T}} = \mathbf{D}_{\{1,2\}}\boldsymbol{\varepsilon}, \quad \tilde{\boldsymbol{\varepsilon}}_{\{2-1\}}^{\mathrm{T}} = \mathbf{D}_{\{2-1\}}\boldsymbol{\varepsilon}, \tag{3.12}$$

$$\mathbf{D}_{\{1,2\}} \equiv \begin{pmatrix} \mathbf{B}_{\mu_{i,j}^{(4)}} \mathbf{D}_{i,j} & & \\ \mathbf{0} & \mathbf{I}_{2(k-i)} & \mathbf{0} \\ \mathbf{B}_{\mu_{k,l}^{(4)}} \mathbf{D}_{k,l} & & \\ \mathbf{0} & \mathbf{I}_{2(j-l)} & \mathbf{0} \end{pmatrix}, \quad \mathbf{D}_{\{2-1\}} \equiv \begin{pmatrix} \mathbf{0} & \mathbf{I}_{2l} & \mathbf{0} \\ & \mathbf{B}_{\mu_{k,l}^{(4)}} \mathbf{D}_{k,l} & \\ \mathbf{0} & (\mathbf{J}_{k-i} \otimes \mathbf{I}_2) & \mathbf{0} \\ & \mathbf{B}_{\mu_{i,j}^{(4)}} \mathbf{D}_{i,j} & \\ \mathbf{0} & \mathbf{I}_{2(n_2-j)} & \end{pmatrix}. \tag{3.13}$$

Therefore, we can obtain the following approximate laws of error propagation in polygon-on-polygon operation:

$$\tilde{\boldsymbol{\Sigma}}_{\{1-2\}} = \mathbf{D}_{\{1-2\}} \boldsymbol{\Sigma}\, \mathbf{D}_{\{1-2\}}^{\mathrm{T}}, \quad \tilde{\boldsymbol{\Sigma}}_{\{1,2\}} = \mathbf{D}_{\{1,2\}} \boldsymbol{\Sigma}\, \mathbf{D}_{\{1,2\}}^{\mathrm{T}}, \quad \tilde{\boldsymbol{\Sigma}}_{\{2-1\}} = \mathbf{D}_{\{2-1\}} \boldsymbol{\Sigma}\, \mathbf{D}_{\{2-1\}}^{\mathrm{T}}. \tag{3.14}$$

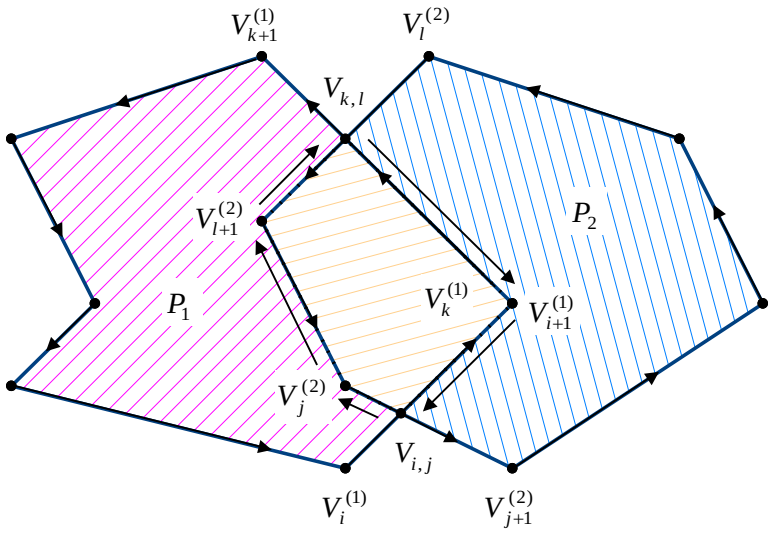


Fig. 3.1 Directed polygons in polygon-on-polygon analysis

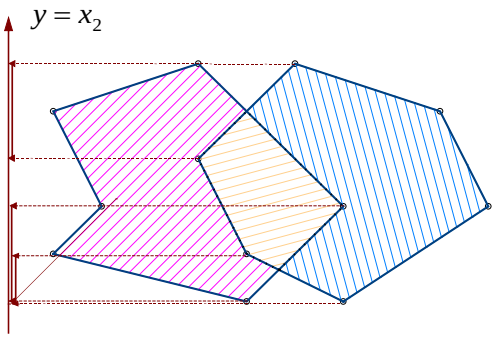


Fig. 3.2 The *y*-intervals test for intersection

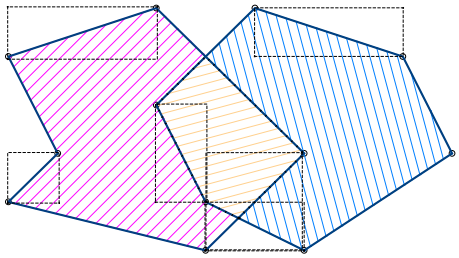


Fig. 3.3 The MER test for intersection

Example 3.1 Consider the polygons shown in Fig. 3.1, where the vertices of P_1 ($n_1 = 6$) are $V_1^{(1)}(0,1)$, $V_2^{(1)}(4,0)$, $V_3^{(1)}(6,2)$, $V_4^{(1)}(3,5)$, $V_5^{(1)}(0,4)$, $V_6^{(1)}(1,2)$, and the vertices of P_2 ($n_2 = 6$) are $V_1^{(2)}(6,0)$, $V_2^{(2)}(9,2)$, $V_3^{(2)}(8,4)$, $V_4^{(2)}(5,5)$, $V_5^{(2)}(3,3)$, $V_2^{(2)}(4,1)$. The true coordinate vectors of intersection points $V_{2,6}$ ($i = 2$, $j = 6$) and $V_{3,4}$ ($k = 3$, $l = 4$) are $\boldsymbol{\mu}_{2,6} = (\frac{14}{3}, \frac{2}{3})^{\mathrm{T}}$ and $\boldsymbol{\mu}_{3,4} = (4,4)^{\mathrm{T}}$. Choose the structures of the ME covariance matrix be (i), (ii), and (iii) in Example 2.2. The overlaid polygons are

$$P_1 - P_2: \quad V_1^{(1)}, V_2^{(1)}, V_{2,6}, V_6^{(2)}, V_5^{(2)}, V_{3,4}, V_4^{(1)}, V_5^{(1)}, V_6^{(1)}, (n_{\{1-2\}} = 9)$$

$$P_1 \cap P_2: \quad V_{2,6}, V_3^{(1)}, V_{3,4}, V_5^{(2)}, V_6^{(2)}, \ (n_{\{1,2\}} = 5)$$

$$P_2 - P_1: \quad V_1^{(2)}, V_2^{(2)}, V_3^{(2)}, V_4^{(2)}, V_{3,4}, V_3^{(1)}, V_{2,6}. \ (n_{\{2-1\}} = 7)$$

The 500 simulation results of P_1 with the ME (i) and P_2 with the ME (iii) are plotted in Fig. 3.4(a). The simulation results of overlaid polygons $P_1 - P_2$, and $P_2 - P_1$, $P_1 \cap P_2$, are shown in Fig. 3.4(b), (c) and (d) respectively. The simulation results of P_1 with the ME (iii) and P_2 with the ME (iv) are plotted in Fig. 3.4(e). The simulation results of overlaid polygons $P_1 - P_2$, and $P_2 - P_1$, $P_1 \cap P_2$, are shown in Fig. 3.4(f), (g) and (h) respectively.

According to (2.18), we obtain

$$\mathbf{B}_{\mu_{2,6}^{(4)}} = \frac{1}{9}\begin{pmatrix} 4 & -4 & 2 & -2 & 2 & 4 & 1 & 2 \\ -2 & 2 & -1 & 1 & 2 & 4 & 1 & 2 \end{pmatrix}, \quad \mathbf{B}_{\mu_{3,4}^{(4)}} = \frac{1}{12}\begin{pmatrix} 2 & 2 & 4 & 4 & 3 & -3 & 3 & -3 \\ 2 & 2 & 4 & 4 & -3 & 3 & -3 & 3 \end{pmatrix}.$$

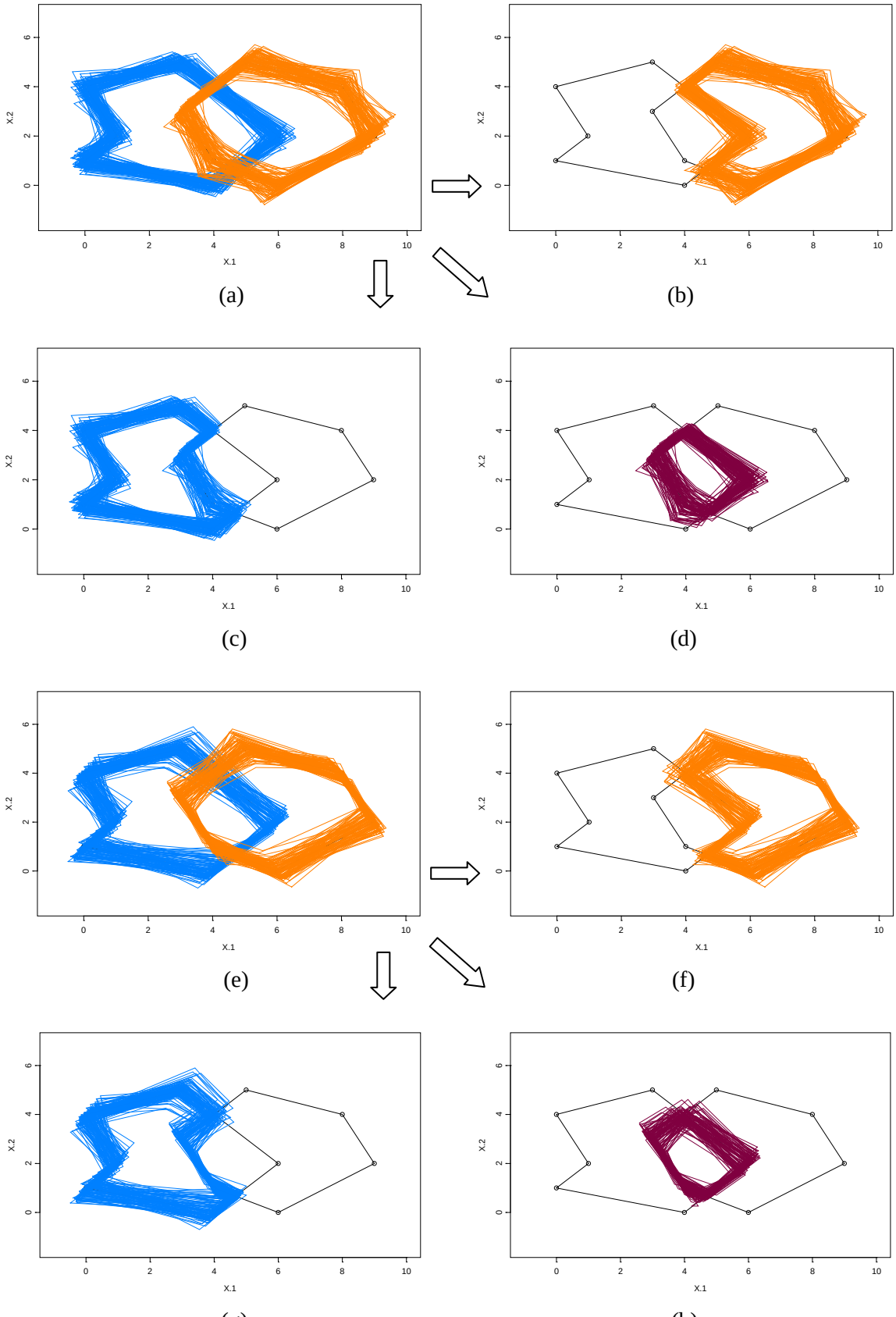


Fig. 3.4 Overlaid polygons from polygon overlay operation

Thus the 18×24 matrix $\mathbf{D}_{\{1-2\}}$, 10×24 matrix $\mathbf{D}_{\{1,2\}}$ and 14×24 matrix $\mathbf{D}_{\{2-1\}}$ can respectively be formed from (3.11) and (3.13). Accordingly, for the overlaid polygons, the approximate laws of error propagation in the polygon-on-polygon operation can in general be obtained from (3.14) when the joint ME covariance matrix $\mathbf{\Sigma}$ for the original polygons are given. For simplicity, we assume that the ME vectors of all vertices of the original polygons are independent. Then $\mathbf{\Sigma}$ is block-diagonal and the approximate laws of error propagation for overlay can be reduced since only the ME covariance matrices of the intersection points need to be computed. For the polygons P_1 with the ME (i) and P_2 with the ME (iii) in Fig. 3.4(a), by simple calculations the approximate ME covariance matrices of intersection points $V_{2,6}$ and $V_{3,4}$ are respectively

$$\tilde{\mathbf{\Sigma}}_{V_{2,6}}=\begin{pmatrix}0.0627 & 0.0257\\ 0.0257 & 0.0442\end{pmatrix},\quad \tilde{\mathbf{\Sigma}}_{V_{3,4}}=\begin{pmatrix}0.0166 & 0.0111\\ 0.0111 & 0.0166\end{pmatrix}.$$

For the polygons P_1 with the ME (iii) and P_2 with the ME (iv) in Fig. 3.4(e), by simple calculations the approximate ME covariance matrices of intersection points $V_{2,6}$ and $V_{3,4}$ are respectively

$$\tilde{\mathbf{\Sigma}}_{V_{2,6}}=\begin{pmatrix}0.0168 & 0.0086\\ 0.0086 & 0.0127\end{pmatrix},\quad \tilde{\mathbf{\Sigma}}_{V_{3,4}}=\begin{pmatrix}0.0628 & 0.0033\\ 0.0033 & 0.0628\end{pmatrix}.$$

Finally, for the polygons P_1 with the ME (iii) and P_2 with the ME (iv) in Fig. 3.4(e), we plot in Fig. 3.5 the confidence regions of the two original polygons formed by the covariance-based error bands with $1-\alpha = 0.9$. Obviously, it is very consistent with the simulation result in Fig. 3.4(e). The effectiveness of the covariance-based error bands is substantiated once again.

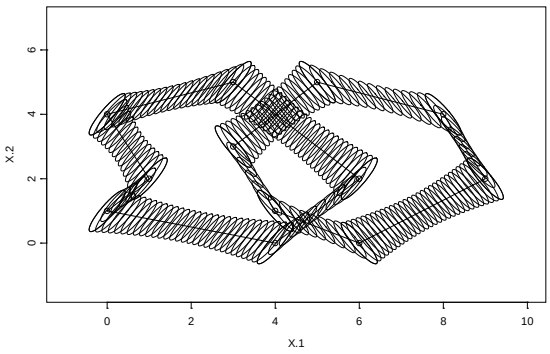


Fig. 3.5 Confidence regions of two original polygons

4. Conclusion

We have established in this part of the series the approximate law of error propagation for the intersection point of two random line segments and the approximate law of error propagation for polygon-on-polygon overlay. As a key to error analysis of vector-data overlay in GIS, the error covariance matrix of the intersection point of two line segments is approximately given in a simple expression, which is based on the joint ME covariance matrix of all endpoints whose ME vectors may be dependent or independent. This approximate law of error propagation provides us with valuable information about the characteristics of error of the intersection point. One basic observation is that the error of an intersection point is influenced by the positional relations of two line segments and the error characteristics of their endpoints. Theoretical results on error propagation for line intersections in turn forms a basis for error analysis in the polygon-on-polygon operation. The error covariance matrices of the overlaid polygons have thus been approximately obtained. Error propagation from two original polygons (may be correlated) to the three overlaid polygons have also been approximately described and performed. The validity and effectiveness of the proposed model have also been demonstrated with simulation experiments.

An outstanding problem for further study is to obtain the analytical error distribution of the coordinate vectors of the intersection point so that the exact law of error propagation for the intersection point of two random line segments can be derived. In the final part of the present series of studies, we will investigate error propagation in length and area measurements in MBGIS.

Appendix 1

Proof of Proposition 2.1

Since

$$
\begin{aligned}
f_{\det}(\mathbf{x}_4-\mathbf{x}_2,\mathbf{x}_3-\mathbf{x}_4) &= (\mathbf{x}_4-\mathbf{x}_2)^{\mathrm{T}}\mathbf{H}_0(\mathbf{x}_3-\mathbf{x}_4) = \mathbf{x}_{(4)}^{\mathrm{T}}(\mathbf{D}_4-\mathbf{D}_2)^{\mathrm{T}}\mathbf{H}_0(\mathbf{D}_3-\mathbf{D}_4)\mathbf{x}_{(4)} \\
&= \mathbf{x}_{(4)}^{\mathrm{T}}[(\mathbf{e}_{4,4}-\mathbf{e}_{4,2})\otimes\mathbf{I}_2]\mathbf{H}_0[(\mathbf{e}_{4,3}-\mathbf{e}_{4,4})^{\mathrm{T}}\otimes\mathbf{I}_2]\mathbf{x}_{(4)} \\
&= \mathbf{x}_{(4)}^{\mathrm{T}}\{[(\mathbf{e}_{4,4}-\mathbf{e}_{4,2})(\mathbf{e}_{4,3}-\mathbf{e}_{4,4})^{\mathrm{T}}]\otimes\mathbf{H}_0\}\mathbf{x}_{(4)} \\
&= \tfrac{1}{2}\mathbf{x}_{(4)}^{\mathrm{T}}\left([(\mathbf{e}_{4,4}-\mathbf{e}_{4,2})(\mathbf{e}_{4,3}-\mathbf{e}_{4,4})^{\mathrm{T}}]\otimes\mathbf{H}_0+\{[(\mathbf{e}_{4,4}-\mathbf{e}_{4,2})(\mathbf{e}_{4,3}-\mathbf{e}_{4,4})^{\mathrm{T}}]\otimes\mathbf{H}_0\}^{\mathrm{T}}\right)\mathbf{x}_{(4)} \\
&= \tfrac{1}{2}\mathbf{x}_{(4)}^{\mathrm{T}}\{[(\mathbf{e}_{4,4}-\mathbf{e}_{4,2})(\mathbf{e}_{4,3}-\mathbf{e}_{4,4})^{\mathrm{T}}-(\mathbf{e}_{4,3}-\mathbf{e}_{4,4})(\mathbf{e}_{4,4}-\mathbf{e}_{4,2})^{\mathrm{T}}]\otimes\mathbf{H}_0\}\mathbf{x}_{(4)}
\end{aligned}
$$

$$= -\tfrac{1}{2}\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_1\mathbf{x}_{(4)},$$

where $\overline{\mathbf{H}}_1 \equiv \overline{\boldsymbol{\Delta}}_1 \otimes \mathbf{H}_0$,

$$\overline{\boldsymbol{\Delta}}_1 \equiv (\mathbf{e}_{4,3}-\mathbf{e}_{4,4})(\mathbf{e}_{4,4}-\mathbf{e}_{4,2})^{\mathrm{T}} - (\mathbf{e}_{4,4}-\mathbf{e}_{4,2})(\mathbf{e}_{4,3}-\mathbf{e}_{4,4})^{\mathrm{T}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix}.$$

Similarly, we can get

$$f_{\det}(\mathbf{x}_1-\mathbf{x}_2,\mathbf{x}_3-\mathbf{x}_4) = -\tfrac{1}{2}\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,\mathbf{x}_{(4)}, \quad f_{\det}(\mathbf{x}_1-\mathbf{x}_2,\mathbf{x}_2-\mathbf{x}_4) = -\tfrac{1}{2}\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_3\mathbf{x}_{(4)}.$$

Thus, from (2.4) and (2.5),

$$\lambda_1^* = \lambda_1^*(\mathbf{x}_{(4)}) \equiv \frac{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_1\mathbf{x}_{(4)}}{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,\mathbf{x}_{(4)}}, \quad \lambda_3^* = \lambda_3^*(\mathbf{x}_{(4)}) \equiv \frac{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_3\mathbf{x}_{(4)}}{\mathbf{x}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}\,\mathbf{x}_{(4)}}.$$

According to $\lambda_2^* = 1-\lambda_1^*$ and $\lambda_4^* = 1-\lambda_3^*$, the proposition can thus be obtained. ڤ

Appendix 2

Proof of Proposition 2.4

Since every permutation corresponds to a unique permutation matrix, we first introduce the following permutation matrices:

$$\mathbf{P}_{1,2} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \ \mathbf{P}_{3,4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \text{ and } \mathbf{P}_{13,24} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

They are obtained from a permutation of the corresponding columns or rows of an identity matrix $\mathbf{I}_4$. Their effect is that when they premultiply (postmultiply) a given matrix **A**, the result is a matrix whose rows (columns) are obtained from a permutation of the corresponding rows (columns) of **A**. They are symmetric and $\mathbf{P}_{1,2}^2 = \mathbf{P}_{3,4}^2 = \mathbf{I}_4$.

If $\mathbf{X}_{(4)}'^{\mathrm{T}} = (\mathbf{X}_2^{\mathrm{T}}\ \mathbf{X}_1^{\mathrm{T}}\ \mathbf{X}_3^{\mathrm{T}}\ \mathbf{X}_4^{\mathrm{T}})$, we have $\mathbf{X}'_{(4)} = (\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\,\mathbf{X}_{(4)}$ and $\boldsymbol{\mu}'_{(4)} = (\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\,\boldsymbol{\mu}_{(4)}$. Since the following two equations hold:

$$(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\overline{\mathbf{H}}_1(\mathbf{P}_{1,2}\otimes\mathbf{I}_2) = (\mathbf{P}_{1,2}\otimes\mathbf{I}_2)(\overline{\boldsymbol{\Delta}}_1\otimes\mathbf{H}_0)(\mathbf{P}_{1,2}\otimes\mathbf{I}_2) = (\mathbf{P}_{1,2}\overline{\boldsymbol{\Delta}}_1\mathbf{P}_{1,2})\otimes\mathbf{H}_0 = -\overline{\mathbf{H}}_2,$$

$$(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\overline{\mathbf{H}}(\mathbf{P}_{1,2}\otimes\mathbf{I}_2) = (\mathbf{P}_{1,2}\otimes\mathbf{I}_2)(\overline{\boldsymbol{\Delta}}\otimes\mathbf{H}_0)(\mathbf{P}_{1,2}\otimes\mathbf{I}_2) = (\mathbf{P}_{1,2}\overline{\boldsymbol{\Delta}}\mathbf{P}_{1,2})\otimes\mathbf{H}_0 = -\overline{\mathbf{H}},$$

we can obtain

$$\boldsymbol{\mu}_{(4)}'^{\mathrm{T}}\overline{\mathbf{H}}_1\boldsymbol{\mu}'_{(4)} = \boldsymbol{\mu}_{(4)}^{\mathrm{T}}(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\overline{\mathbf{H}}_1(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\,\boldsymbol{\mu}_{(4)} = -\boldsymbol{\mu}_{(4)}^{\mathrm{T}}\overline{\mathbf{H}}_2\boldsymbol{\mu}_{(4)},$$

$$\boldsymbol{\mu}'^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}\boldsymbol{\mu}'_{(4)} = \boldsymbol{\mu}^{\mathrm{T}}_{(4)}(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\,\overline{\mathbf{H}}\,(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\,\boldsymbol{\mu}_{(4)} = -\boldsymbol{\mu}^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}\,\boldsymbol{\mu}_{(4)}\,.$$

Accordingly, $\lambda_1^*(\boldsymbol{\mu}'_{(4)}) = \lambda_2^*(\boldsymbol{\mu}_{(4)})$. Similarly, $\lambda_2^*(\boldsymbol{\mu}'_{(4)}) = \lambda_1^*(\boldsymbol{\mu}_{(4)})$. In addition,

$$\mathbf{D}_1(\mathbf{P}_{1,2}\otimes\mathbf{I}_2) = (\mathbf{e}^{\mathrm{T}}_{4,1}\otimes\mathbf{I}_2)(\mathbf{P}_{1,2}\otimes\mathbf{I}_2) = (\mathbf{e}^{\mathrm{T}}_{4,1}\mathbf{P}_{1,2})\otimes\mathbf{I}_2 = \mathbf{e}^{\mathrm{T}}_{4,2}\otimes\mathbf{I}_2 = \mathbf{D}_2\,.$$

By the same derivation, it can be implied that $\mathbf{D}_2(\mathbf{P}_{1,2}\otimes\mathbf{I}_2) = \mathbf{D}_1$. Thus from (2.18)

$$\begin{aligned}
\mathbf{B}_{\mu'_{(4)}} &= \frac{2}{\boldsymbol{\mu}'^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}\,\boldsymbol{\mu}'_{(4)}}[\mathbf{D}_1\boldsymbol{\mu}'_{(4)}\boldsymbol{\mu}'^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}_1 + \mathbf{D}_2\boldsymbol{\mu}'_{(4)}\boldsymbol{\mu}'^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}_2] \\
&\quad + [\lambda_1^*(\boldsymbol{\mu}'_{(4)})\,\mathbf{D}_1 + \lambda_2^*(\boldsymbol{\mu}'_{(4)})\,\mathbf{D}_2]\left(\mathbf{I}_8 - \frac{2}{\boldsymbol{\mu}'^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}\,\boldsymbol{\mu}'_{(4)}}\boldsymbol{\mu}'_{(4)}\boldsymbol{\mu}'^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}\right). \\
&= -\frac{2}{\boldsymbol{\mu}^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}\,\boldsymbol{\mu}_{(4)}}[\mathbf{D}_1(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\,\boldsymbol{\mu}_{(4)}\boldsymbol{\mu}^{\mathrm{T}}_{(4)}(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\overline{\mathbf{H}}_1 + \mathbf{D}_2(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\,\boldsymbol{\mu}_{(4)}\boldsymbol{\mu}^{\mathrm{T}}_{(4)}(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\overline{\mathbf{H}}_2] \\
&\quad + [\lambda_2^*(\boldsymbol{\mu}_{(4)})\mathbf{D}_1 + \lambda_1^*(\boldsymbol{\mu}_{(4)})\mathbf{D}_2]\left(\mathbf{I}_8 + \frac{2}{\boldsymbol{\mu}^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}\,\boldsymbol{\mu}_{(4)}}(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\,\boldsymbol{\mu}_{(4)}\,\boldsymbol{\mu}^{\mathrm{T}}_{(4)}(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\overline{\mathbf{H}}\right) \\
&= \frac{2}{\boldsymbol{\mu}^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}\,\boldsymbol{\mu}_{(4)}}[\mathbf{D}_2\boldsymbol{\mu}_{(4)}\boldsymbol{\mu}^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}_2 + \mathbf{D}_1\boldsymbol{\mu}_{(4)}\boldsymbol{\mu}^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}_1](\mathbf{P}_{1,2}\otimes\mathbf{I}_2) \\
&\quad + [\lambda_2^*(\boldsymbol{\mu}_{(4)})\mathbf{D}_2 + \lambda_1^*(\boldsymbol{\mu}_{(4)})\mathbf{D}_1]\left(\mathbf{I}_8 - \frac{2}{\boldsymbol{\mu}^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}\,\boldsymbol{\mu}_{(4)}}\boldsymbol{\mu}_{(4)}\boldsymbol{\mu}^{\mathrm{T}}_{(4)}\overline{\mathbf{H}}\right)(\mathbf{P}_{1,2}\otimes\mathbf{I}_2) \\
&= \mathbf{B}_{\mu_{(4)}}(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\ .
\end{aligned}$$

Combining the relationship

$$\begin{aligned}
\boldsymbol{\Sigma}'_{(4)} &= E(\mathbf{X}'_{(4)} - E\,\mathbf{X}'_{(4)})(\mathbf{X}'_{(4)} - E\,\mathbf{X}'_{(4)})^{\mathrm{T}} = (\mathbf{P}_{1,2}\otimes\mathbf{I}_2)E(\mathbf{X}_{(4)} - E\,\mathbf{X}_{(4)})(\mathbf{X}_{(4)} - E\,\mathbf{X}_{(4)})^{\mathrm{T}}(\mathbf{P}_{1,2}\otimes\mathbf{I}_2) \\
&= (\mathbf{P}_{1,2}\otimes\mathbf{I}_2)\boldsymbol{\Sigma}_{(4)}(\mathbf{P}_{1,2}\otimes\mathbf{I}_2),
\end{aligned}$$

we obtain

$$\widetilde{\boldsymbol{\Sigma}}'_c = \mathbf{B}_{\mu'_{(4)}}\boldsymbol{\Sigma}'_{(4)}\mathbf{B}^{\mathrm{T}}_{\mu'_{(4)}} = \mathbf{B}_{\mu_{(4)}}(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)^2\,\boldsymbol{\Sigma}_{(4)}(\mathbf{P}_{1,2}\otimes\mathbf{I}_2)^2\mathbf{B}^{\mathrm{T}}_{\mu_{(4)}} = \widetilde{\boldsymbol{\Sigma}}_c\,.$$

For the case $\mathbf{X}'^{\mathrm{T}}_{(4)} = (\mathbf{X}_1^{\mathrm{T}}\ \mathbf{X}_2^{\mathrm{T}}\ \mathbf{X}_4^{\mathrm{T}}\ \mathbf{X}_3^{\mathrm{T}})$, $\mathbf{X}'^{\mathrm{T}}_{(4)} = (\mathbf{X}_3^{\mathrm{T}}\ \mathbf{X}_4^{\mathrm{T}}\ \mathbf{X}_1^{\mathrm{T}}\ \mathbf{X}_2^{\mathrm{T}})$ and other cases, it can likewise be shown that $\widetilde{\boldsymbol{\Sigma}}'_c = \widetilde{\boldsymbol{\Sigma}}_c$. ف

References

Berg, M. de, M. van Kreveld, M. Overmars, and O. Schwarzkopf. 2000. *Computational Geometry: Algorithms and Applications* (2nd ed.), Berlin: Springer-Verlag.

Goodchild, M.F. 1978. Statistical aspects of the polygon overlay problem. In *Harvard Papers on Geographic Information Systems*, edited by G. Dutton. Cambridge, MA: Laboratory for Computer Graphics and Spatial Analysis, Harvard University, 6, pp.1-12.

Griffith, D.A., R.P. Haining, and G. Arbia. 1999. Unicertainty and error propagation in map analyses involving arithmetric and overlay operations: inventory and prospects. In *Spatial Accuracy*

Assessment: Land Information Uncertainty in Natural Resources, K. Lowell and A. Jaton (eds), pp. 11-25, Chelsea, Michigan: Ann Arbor Press.

Harding, T.J., R.G. Healey, S.Hopkins and S. Dowers. 1998. Vector polygon overlay. In Healey, R., S. Dowers, B. Gittings, and M. Mineter.(eds). *Parallel Processing Algorithms for GIS*. London: Taylor & Francis

Hunter, G.J., J. Qiu, and M.F. Goodchild. 1999. Application of a new model of vector data uncertainty. In K. Lowell and A. Jaton, editors, *Spatial Accuracy Assessment*: *Land Information Uncertainty in Natural Resources*. Chelsea, Michigan: Ann Arbor Press, 203-208.

Lantner, D. and H. Veregin. 1992. A research paradigm for propagating error in layer-based GIS, *Photogrammetric Eng. Remote Sensing*, 58, pp. 825-833.

Leung, Y., J. H. Ma, and M.F. Goodchild. 2003a. A general framework for error analysis in measurement-based GIS, Part 1: the basic measurement-error model and related concepts. (unpublished paper)

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

Lunetta, R.S., Congalton, R.G., Fenstermaker, L.K., Jensen, J.R., McGwire, K.C., and Tinney, L.R. 1991. Remote sensing and geographic information system data integration: error sources and research issues. *Photogrammetric Engineering and Remote Sensing*, 57(6), 677-687.

Rigaux, P., M. Scholl and A.Voisard, 2002. *Spatial Databases with Application to GIS*. San Francisco: Morgan Kaufmann Publishers.

Veregin, H. 1995. Developing and testing of an error propagation model for GIS overlay operations. *Int. J. Geographical Information Systems*, 9, 595-619.

Wise, S. 2002. *GIS Basics*. London: Taylor & Francis.