

Ricerca automatica di variazioni locali tra configurazioni spaziali puntuali mediante analisi procustiana robusta

Fabio CROSILLA, Alberto BEINAT

Università di Udine, Dipartimento di Georisorse e Territorio, via Cottonificio 114 - 33100 UDINE
tel. 0432 558704, fax. 0432 558700, *email* fabio.crosilla@uniud.it; alberto.beinat@uniud.it.

Riassunto

L'articolo descrive un metodo originale per eseguire la ricerca automatica di variazioni locali tra configurazioni spaziali puntuali. Il metodo, basato sull'analisi procustiana generalizzata robusta, è stato recentemente sviluppata dagli autori nell'ambito dello studio di innovativi strumenti statistici di ricerca degli errori grossolani, ed applicato positivamente per eseguire l'allineamento globale di scene tridimensionali non stazionarie di punti laser. I modelli procustiani sono stati implementati secondo una sequenza progressiva ispirata alla tecnica statistica denominata "forward search". A partire da un nucleo comune iniziale delle configurazioni in esame, che soddisfa il metodo statistico "least median of squares" (LMS) per gli errori di misura, il nucleo originario viene iterativamente aggiornato e allargato con i punti rimanenti, fintanto che non interviene una variazione statisticamente significativa nei parametri di allineamento, la quale segnala la presenza di punti non stazionari tra quelli appena considerati. La descrizione del metodo è completata da esempi numerici relativi ad applicazioni in ambito cartografico.

Abstract

The paper describes an original method to execute the automatic research of local variations among discrete spatial configurations. The method, based on robust generalized procrustes analysis, has been recently developed by the authors as innovative statistical search tools of outliers, and positively applied to execute the global alignment of non stationary 3D scenes of laser points. The procrustes models have been implemented according to a progressive sequence inspired to the statistical technique known as "forward search". Starting from an initial common point subset of the studied configuration, satisfying the LMS method for the measurement errors, the original subset is iteratively updated and enlarged with the remaining points till a significant variation of the alignment parameters occurs. This signals the presence of non stationary points among those ones just considered. The method description is completed by some numerical examples related to cartographic applications.

1. Introduzione

In cartografia numerica, nelle applicazioni GIS e CAD, e in particolare nella pratica topografica relativa al monitoraggio di fenomeni deformativi che interessano il territorio e le strutture, si affronta frequentemente il problema di identificare e analizzare le variazioni che intercorrono nelle configurazioni geometriche di una stessa serie di punti, definite in epoche o in condizioni strumentali e ambientali diverse, contaminate anche dalla presenza di errori accidentali nelle misure. Accanto agli approcci assiomatici, non sempre agevoli, nei quali è l'operatore stesso che, sulla base dell'esperienza o della conoscenza del fenomeno, interviene per riconoscere le componenti invarianti delle varie configurazioni, vi sono altre soluzioni automatiche sviluppate nel campo della morfometria, o derivabili dall'analisi delle forme implementata nella robotica. Schematicamente, il problema che queste affrontano è quello di individuare in modo automatico quelle parti delle configurazioni in esame che risultano stazionarie, e che si mantengono conformi

nelle diverse realizzazioni, permettendo di identificare rispetto ad esse, gli elementi per i quali sono intervenute variazioni o deformazioni statisticamente significative.

L'articolo propone un metodo originale per eseguire la ricerca automatica di variazioni locali tra configurazioni spaziali puntuali. Il metodo, basato sull'analisi procustiana generalizzata robusta, è stato implementato secondo una sequenza progressiva ispirata alla tecnica nota con il termine di "forward search" (Cerioli e Riani 2003). A partire da un nucleo comune iniziale delle configurazioni in esame, che soddisfa il metodo statistico "least median of squares" (LMS) per gli errori di misura, il nucleo originario viene iterativamente aggiornato e allargato con i punti rimanenti, fintanto che non interviene una variazione statisticamente significativa nei parametri di trasformazione, la quale segnala la presenza di punti non stazionari tra quelli appena considerati. L'articolo descrive innanzitutto il metodo statistico proposto; la trattazione è quindi completata da un esempio numerico relativo ad una applicazione in ambito cartografico.

2. Analisi Procustiana Generalizzata Robusta

L'Analisi Procustiana Generalizzata (GPA) è una tecnica di analisi statistica multivariata usata per realizzare direttamente trasformazioni di similitudine multiple e simultanee di $M \geq 2$ serie di dati, ciascuna costituita dalle coordinate D-dimensionali di P punti corrispondenti, definite rispetto a $M \geq 2$ diversi sistemi di riferimento e caratterizzate da errore di misura. Per tale scopo, si richiede il soddisfacimento della seguente funzione obiettivo:

$$S = \text{tr} \sum_{i < j}^M \left[(c_i \mathbf{X}_i \mathbf{R}_i + \mathbf{1t}'_i) - (c_j \mathbf{X}_j \mathbf{R}_j + \mathbf{1t}'_j) \right]' \cdot \left[(c_i \mathbf{X}_i \mathbf{R}_i + \mathbf{1t}'_i) - (c_j \mathbf{X}_j \mathbf{R}_j + \mathbf{1t}'_j) \right] = \min \quad (1)$$

con la condizione di ortogonalità per la matrice \mathbf{R} , cioè $\mathbf{R} \mathbf{R}' = \mathbf{I}$, e dove $\mathbf{X}_1 \dots \mathbf{X}_M$ sono $M \geq 2$ matrici di dati di dimensione (PxD), ciascuna contenente le coordinate della stessa serie di P punti relative a M diversi sistemi di riferimento; $\mathbf{1}$ è un vettore ausiliario (Px1); \mathbf{t}_j è il vettore incognito di traslazione (Dx1), \mathbf{R}_j è la matrice incognita di rotazione (DxD) e c_j è il fattore di scala isotropo, con j che varia fra 1 e M.

Il soddisfacimento della (1) rappresenta la soluzione del problema GPA descritta da Kristof e Wingersky (1971), Gower (1975), ten Berge (1977), e Goodall (1991).

Questo problema ha una formulazione alternativa. Definito $\mathbf{X}_i^p = c_i \mathbf{X}_i \mathbf{R}_i + \mathbf{1t}'_i$, le seguenti misure:

$$\sum_{i < j}^M \|\mathbf{X}_i^p - \mathbf{X}_j^p\|^2 = \sum_{i < j}^M \text{tr} (\mathbf{X}_i^p - \mathbf{X}_j^p)' (\mathbf{X}_i^p - \mathbf{X}_j^p) \quad (2)$$

$$M \sum_i^M \|\mathbf{X}_i^p - \mathbf{H}\|^2 = M \sum_i^M \text{tr} (\mathbf{X}_i^p - \mathbf{H})' (\mathbf{X}_i^p - \mathbf{H}) \quad (3)$$

sono perfettamente equivalenti (vedi: Borg e Groenen, 1997), dove \mathbf{H} corrisponde al centroide incognito. Pertanto l'equazione (3), invece della (2) o della (1), può essere minimizzata in modo da determinare le incognite $\{c, \mathbf{R}, \mathbf{t}\}_j$ ($j = 1 \dots M$) che consentono di calcolare iterativamente la soluzione finale \mathbf{x}_i^p ($i=1 \dots M$).

La matrice $\hat{\mathbf{H}} = \frac{1}{M} \sum_{i=1}^M \mathbf{X}_i^p$ rappresenta la stima ai minimi quadrati di \mathbf{H} . Si noti che $\mathbf{H} + \mathbf{E}_i = \mathbf{X}_i^p$,

dove $\text{vec}(\mathbf{E}_i): N \{0, \Sigma = \sigma^2 (\mathbf{Q}_n \otimes \mathbf{Q}_k)\}$ e Σ ha una struttura fattoriale.

Nell'attuale implementazione dell'algoritmo (Crosilla e Beinat, 2006; 2007), la procedura inizia da una configurazione parziale di punti contenente solo quelli stazionari. A ciascuna iterazione il set iniziale di dati viene allargato con uno o più punti, fintanto che si verifica una variazione significativa dei parametri di trasformazione.

Al fine di definire la configurazione iniziale di punti X^i di X , cioè quella contenente punti stazionari, è necessario calcolare la stima ai minimi quadrati del centroide corrispondente \hat{H}^i , e conseguentemente determinare i parametri di trasformazione di similitudine per tutti i punti contenuti nelle $j = 1 \dots M$ sotto matrici X_j^i :

$$\hat{H}^i = \frac{1}{M} \sum_{j=1}^M \left(c_j^i X_j^i \mathbf{R}_j^i + \mathbf{1} t_j^{i'} \right) \quad (4)$$

dove \hat{H}^i corrisponde alla stima ai minimi quadrati della matrice incognita H^i . Tale procedura viene ripetuta per ogni $i = 1 \dots \binom{P}{S}$ possibile sottoconfigurazione X^i , dove S è il numero di punti che formano il sottoinsieme.

A questo punto lo pseudo-centroide globale viene calcolato applicando i parametri della trasformazione, relativi alla i -esima sotto-matrice di dati X_j^i , alla completa matrice corrispondente X_j , ottenendo così $X_j^{P(i)}$:

$$\tilde{H}^i = \frac{1}{M} \sum_{j=1}^M \left(c_j^i X_j^i \mathbf{R}_j^i + \mathbf{1} t_j^{i'} \right) = \frac{1}{M} \sum_{j=1}^M X_j^{P(i)} \quad (5)$$

Per definire il sottoinsieme iniziale X^i contenente punti stazionari, si applica il così detto principio "least median of squares" (LMS) (Rousseauw, 1984). Come è noto dalla letteratura, tale metodo di regressione può raggiungere una soglia *break down* pari al 50%: fra tutte le possibili configurazioni dei sottoinsiemi X^i , si sceglie quale configurazione iniziale quella che soddisfa il seguente principio LMS:

$$\text{med diag} \sum_{j=1}^M \left(X_j^{P(i)} - \tilde{H}^i \right) \left(X_j^{P(i)} - \tilde{H}^i \right)' = \min \quad (6)$$

Il sottoinsieme iniziale viene allargato considerando il punto per cui

$$\text{diag} \sum_{j=1}^M \left(X_j^{P(i)} - \tilde{H}^i \right) \left(X_j^{P(i)} - \tilde{H}^i \right)' = \min \quad (7)$$

selezionato fra i rimanenti (P-S) punti della configurazione non appartenenti al sottoinsieme iniziale. La stima ai minimi quadrati del centroide parziale allargato $\hat{H}^{i(+1)}$, e i parametri delle trasformazioni per le M sottomatrici di similitudine $X_j^{i(+1)}$, sono nuovamente calcolati dalla:

$$\hat{H}^{i(+1)} = \frac{1}{M} \sum_{j=1}^M \left(c_j^{i(+1)} X_j^{i(+1)} \mathbf{R}_j^{i(+1)} + \mathbf{1} t_j^{i(+1)'} \right) = \frac{1}{M} \sum_{j=1}^M X_j^{i,P[i(+1)]} \quad (8)$$

A questo punto viene applicata la cosiddetta "statistica procustiana" (Sibson, 1979; Langron e Collins, 1985) per verificare una eventuale variazione significativa dei parametri delle trasformazioni di similitudine dopo aver allargato l'originale sottoinsieme di punti. A tal fine si procede al calcolo della distanza totale fra il centroide parziale \hat{H}^i e le M sottomatrici $X_j^{i,P[i(+1)]}$ ottenute applicando all'originale X_j i parametri della trasformazione di similitudine relativi alla $i(+1)$ serie di dati:

$$\mathbf{G} = \sum_{j=1}^M \text{tr} \left(X_j^{i,P[i(+1)]} - \hat{H}^i \right) \left(X_j^{i,P[i(+1)]} - \hat{H}^i \right)' \quad (9)$$

Vengono anche calcolate le seguenti distanze:

$$\mathbf{G}_t = \sum_{j=1}^M \text{tr} \left(X_j^{i,P[i(+1)]} + \mathbf{1} dt_j' - \hat{H}^i \right) \left(X_j^{i,P[i(+1)]} + \mathbf{1} dt_j' - \hat{H}^i \right)' \quad (10a)$$

$$\mathbf{G}_{tR} = \sum_{j=1}^M \text{tr} \left(X_j^{i,P[i(+1)]} d\mathbf{R}_j + \mathbf{1} dt'_j - \hat{H}^i \right)' \left(X_j^{i,P[i(+1)]} d\mathbf{R}_j + \mathbf{1} dt'_j - \hat{H}^i \right) \quad (10b)$$

$$\mathbf{G}_{tRc} = \sum_{j=1}^M \text{tr} \left(dc_j X_j^{i,P[i(+1)]} d\mathbf{R}_j + \mathbf{1} dt'_j - \hat{H}^i \right)' \left(dc_j X_j^{i,P[i(+1)]} d\mathbf{R}_j + \mathbf{1} dt'_j - \hat{H}^i \right) \quad (10c)$$

dopo aver considerato il fatto che le componenti di traslazione relative al $i(+1)$ sottoinsieme devono essere preliminarmente ridotte per la differenza di distanza fra i centroidi $\hat{H}^{i(+1)}$ e \hat{H}^i . Le distanze riportate in (10 a, b, c) corrispondono a quelle residue dopo aver applicato una trasformazione Procustiana. In particolare G_t corrisponde alla distanza residua successivamente ad una traslazione, G_{tR} è la distanza residua dopo una traslazione e una rotazione e G_{tRc} è la distanza residua dopo una traslazione, una rotazione e una variazione di scala isotropa. Assumendo un valore per l'errore di primo tipo α e considerando gli appropriati gradi di libertà df_1 e df_2 il rifiuto dell'ipotesi nulla per i test successivi (Langron e Collins 1985):

$$\left\{ \frac{G - G_t}{G_{tRc}}, \frac{G_t - G_{tR}}{G_{tRc}}, \frac{G_{tR} - G_{tRc}}{G_{tRc}} \right\} > F_{1-\alpha, df_1, df_2} \quad (11)$$

indica a questo passo del processo una variazione significativa di uno o di tutti i parametri della trasformazione, dovuta al possibile ingresso di dati non stazionari nel set $X_j^{i(+1)}$. Se invece viene accettata l'ipotesi nulla per tutti i test applicati, il processo iterativo continua con l'inserimento di un ulteriore nuovo punto $X_j^{i(+2)}$, che soddisfa l'equazione (7) fra i rimanenti punti analizzati.

3. Applicazioni cartografiche

Tra le applicazioni più significative, si descrive la soluzione automatica di un problema cartografico ricorrente, nel quale si deve individuare, tra diversi rilievi di un edificio o di una particella catastale, il sottoinsieme comune dei vertici congruenti, per procedere quindi alla sovrapposizione ottimale delle corrispondenti entità. Operazioni simili si incontrano, ad esempio, nell'inserimento degli aggiornamenti della mappa catastale mediante il programma Pregeo.

Per sottolineare le caratteristiche della procedura proposta si analizza il caso di un oggetto cartografico, comprendente due edifici e la particella entro cui sorgono, contenuti in tre diverse configurazioni di 24 vertici ciascuna (Figura 1), prodotte da rilievi effettuati in epoche diverse, ed espresse in un proprio indipendente sistema di riferimento.

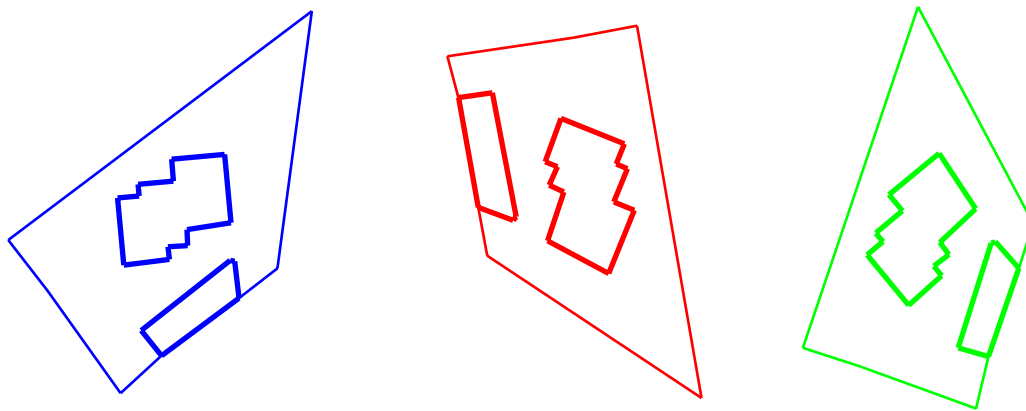


Figura 1 - Tre rilievi di uno stesso oggetto cartografico, effettuati in diverse epoche, espressi in sistemi di riferimento indipendenti. La procedura consente di identificare le differenze significative e di sovrapporre rigorosamente le geometrie.

Per brevità, le corrispondenze tra punti omologhi si assumono già definite. Le coordinate dei vertici che compongono ogni figura sono affette da errori casuali di entità limitata, con media nulla e scarti massimi dell'ordine di ± 25 mm. Sei vertici risultano significativamente spostati, tra 50 e 150 cm, per effetto di errori grossolani commessi nel rilevamento o di modifiche reali intervenute nel corso del tempo. La Figura 2 evidenzia il risultato della sovrapposizione erronea che si otterrebbe considerando tutti i punti delle varie configurazioni e ripartendo gli scarti tra tutti i vertici stazionari e non: malgrado la piccola scala del disegno, le frecce evidenziano gli effetti prodotti dalle componenti sistematiche presenti nelle coordinate dei vertici.

Il metodo proposto è designato proprio per evitare questo inconveniente. Come descritto nel testo, la procedura identifica inizialmente il sottoinsieme di massima concordanza tra le varie configurazioni, ovvero quello che soddisfa la (6) secondo il principio LMS. La dimensione di partenza è arbitraria; sebbene il nucleo minimo iniziale sia un triangolo, nell'esempio si è imposta la scelta di un quadrilatero. Il calcolo della trasformazione di similitudine tra figure triangolari infatti, se non assoggettata ad alcuna forma di vincolo, può generare matrici di rotazione ortogonali che producono un ribaltamento di uno dei due triangoli. Con quattro vertici questo inconveniente diviene del tutto eccezionale, e con cinque vertici tale comportamento non si è mai osservato.

Il nucleo iniziale viene quindi allargato inglobando, uno alla volta, i vertici successivi che soddisfano la (7), e sottoponendo la configurazione così estesa, e quella precedente, a test (11), assumendo in genere un livello di significatività del 95%. Quando la verifica fallisce, segnalando una variazione significativa di traslazione, di rotazione o di fattore di scala tra le due configurazioni in esame, il punto candidato a far parte del nucleo iniziale viene respinto e il processo di allargamento termina. Tutti i punti non inglobati sono conseguentemente classificati come *outlayer* o non-stazionari.

La Figura 3, per confronto con la Figura 2, dimostra l'ottima congruenza ottenuta allineando i rilievi sui soli punti stazionari; inoltre evidenzia con dei cerchietti i vertici che hanno subito uno spostamento o sono risultati affetti da errori grossolani di misura, così come sono stati correttamente riconosciuti dalla procedura.

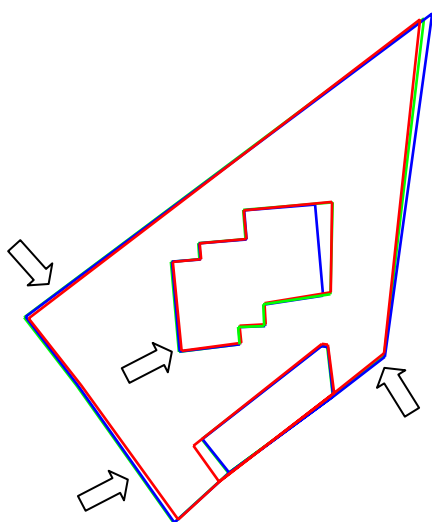


Figura 2 - Allineamento effettuato con tutti i punti: le frecce evidenziano alcuni disallineamenti causati da vertici non stazionari o affetti da errori grossolani.

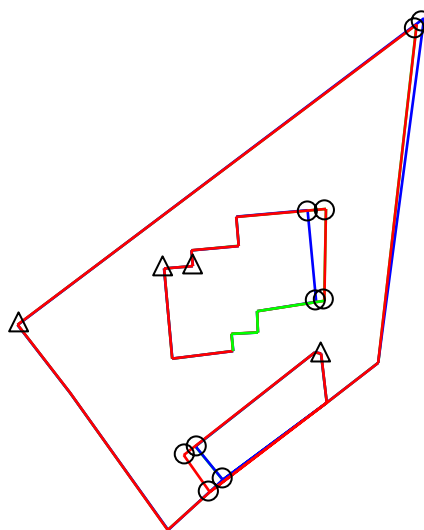


Figura 3 - Allineamento realizzato sui soli punti stazionari. I triangoli indicano i punti del nucleo iniziale ed i cerchietti localizzano i soli vertici non congruenti. Tutti sono individuati automaticamente dalla procedura.

4. Considerazioni conclusive

Numerose esperienze condotte su dati reali e simulati ci consentono di delineare l'affidabilità della procedura. In presenza di configurazioni caratterizzate da un basso rumore di fondo, il test è molto sensibile e individua tutte le componenti sistematiche presenti, anche se di lieve entità. In presenza di configurazioni con elevato sqm degli errori accidentali, la procedura tende invece a diventare permissiva. In linea di principio, la capacità del test di identificare componenti sistematiche o non stazionarie è proporzionale al rapporto tra il valore massimo dell'errore accidentale (l'errore a 3 sigma) e l'entità della componente sistematica o grossolana: è opportuno pertanto che il primo sia di almeno un ordine di grandezza inferiore al secondo. Questo limite operativo è principalmente legato alla sensitività del test di Fisher, assunto quale strumento di riconoscimento delle componenti sistematiche.

Nel prosieguo della ricerca appare opportuno approfondire la formulazione del test statistico allo scopo di rendere il metodo più sensibile nell'identificazione di componenti non stazionarie, permettendo di estendere il suo campo di applicazione anche ad altri settori del rilevamento e del monitoraggio.

Bibliografia

- Borg, I., Groenen, P.J.F., 1997. *Modern multidimensional scaling: Theory and applications*. New York, Springer.
- Cerioli, A., Riani, M., 2003. "Robust Methods for the Analysis of Spatially Autocorrelated Data", *Statistical Methods & Applications*, 11, pp. 335-358.
- Crosilla, F.; Beinat, A. 2006. "A forward search method for robust generalised Procrustes analysis", *Data Analysis, Classification and the Forward Search*, S. Zani, A. Cerioli, M. Riani and M. Vichi (Eds), Springer-Verlag, pp. 199-208.
- Crosilla, F. Beinat A. 2007. "Global Registration of Non Static 3D Lidar Point Clouds: SVD Factorisation and Robust Gpa Methods", The 5th International Symposium on Mobile Mapping Technology, Padova 29-31 maggio 2007.
- Goodall, C., 1991. "Procrustes methods in the statistical analysis of shape", *Journal Royal Stat. Soc.* Part B 53, 2, pp. 285-339.
- Gower, J. C., 1975. "Generalized Procrustes analysis", *Psychometrika*, 40(1), pp. 33-51.
- Langron, S. P.; Collins, A. J.; 1985. "Perturbation theory for Generalized Procrustes Analysis", *Journal Royal Statistical Society*, 47(2), pp. 277-284.
- Kristof, W.; Wingersky, B., 1971. "Generalization of the orthogonal Procrustes rotation procedure to more than two matrices", *Proc. of the 79-th Annual Conv. of the American Psychological Ass.*, 6, pp. 89-90
- Rousseeuw, P. J., 1984. "Least Median of Squares Regression", *J. of the American Statistical Association*, 79(388), pp. 871-880.
- Sibson, R., 1979. "Studies in the Robustness of Multidimensional Scaling: Perturbational Analysis of Classical Scaling", *Journ. R. Statis. Soc.*, B, 41, pp. 217-229.
- ten Berge, J. M. F., 1977. "Orthogonal Procrustes rotation for two or more matrices", *Psychometrika*, 42(2), pp. 267-276.

Ringraziamenti

Il presente lavoro di ricerca si svolge nell'ambito delle attività previste dal progetto EU INTERREG IIIA Phare/CBC Italia-Slovenia "Ricomposizione della cartografia catastale e integrazione della cartografia tecnica regionale numerica per i sistemi informativi territoriali degli enti locali mediante sperimentazione di nuove tecnologie di rilevamento" svolto dall'Università di Udine (<http://geomatica.uniud.it>) in collaborazione con il Geodetski Inštitut Slovenije di Lubiana.