ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 2000
Seria: GÓRNICTWO z. 246 Nr kol.1480

Silverio GARCÍA CORTÉS
Wydział Górniczy, Oviedo
Violetta SOKOŁA-SZEWIOŁA, Ai PHAM QUANG
Politechnika Śląska, Gliwice

CONSIDERACIONES GENERALES SOBRE EL AJUSTE DE OBSERVACIONES TOPOGRÁFICAS MEDIANTE MÍNIMOS CUADRADOS

Resumen. Este artículo constituye una introducción teórica donde se analizan las alternativas existentes en el planteamiento del problema del ajuste de observaciones. Se trata de justificar los planteamientos teóricos que se han implementado en el programa DART (Diseño y Ajuste de Redes Topográficas) desarrollado en el Área de Ingeniería Cartográfica Geodesia y Fotogrametría del Dpto. de Explotación de Minas de la Universidad de Oviedo. Este artículo por tanto debe considerarse en conexión con un artículo posterior en el que describe en detalle la funcionalidad del programa mencionado.

OGÓLNE ROZWAŻANIA NA TEMAT WYRÓWNANIA OBSERWACJI TOPOGRAFICZNYCH METODĄ NAJMNIEJSZYCH KWADRATÓW

Streszczenie. Artykuł jest teoretycznym wprowadzeniem do analizy istniejących alternatywnych rozwiązań związanych z zagadnieniem wyrównania obserwacji. Jego celem jest przedstawienie założeń teoretycznych zastosowanych w programie DART (Projektowanie i wyrównanie sieci topograficznych) opracowanym w Zakładzie Inżynierii Kartograficznej, Geodezji i Fotogrametrii Wydziału Górniczego w Oviedo.

GENERAL CONSIDERATIONS ON TOPOGRAPHIC OBSERVATIONS USING THE LEAST SQUARES METHOD

Summary. The paper is a theoretical introduction to analysis of existing alternative solutions connected to problem of observation equalisation. Its purpose is to present the theoretical assumptions used in DART programme (Designing and equalisation of topographic networks) developed in Zakład Inżynierii Kartograficznej, Geodezji i Fotogrametrii Wydziału Górniczego (Cartography Engineering, Geodesy and Photogrammetry Plant of Mining Faculty) in Oviedo.

Introducción

Durante los trabajos topográficos de implantación de una nueva estructura, replanteo o simplemente registro y control de características naturales o artificiales, se realiza la toma de medidas en campo mediante observaciones topográficas. Dada la naturaleza de los métodos e instrumentos de medida, el producto de este trabajo, es un conjunto de observaciones con una inexactitud asociada, que obliga al planteamiento de un método para la reducción de los errores y aumento de la precisión. El método fundamental es la toma de observaciones redundantes que permitan el cálculo de unos parámetros lo más libres de error que sea posible. Surge entonces el concepto de "ajuste" o "estimación" como el proceso de cálculo que permite, de un lado el aprovechamiento de la totalidad de la información contenida en las observaciones y de otro la obtención de una solución única que pueda ser considerada como la mejor según algún criterio. El criterio más frecuentemente empleado es el conocido como método de mínimos cuadrados.

Sistema de referencia

Las observaciones que se toman en el campo están referidas a un sistema local definido por el instrumento de observación y su posición sobre el punto de estación. Con cada tipo de instrumento e incluso con distintos estacionamientos del mismo instrumento, el sistema de referencia empleado es distinto (sistema geodésico local). Es necesario referir las observaciones tomadas, a un sistema concreto y único en el que se va a realizar el ajuste. Si el trabajo se trata de una nivelación el sistema que debe definirse será unidimensional, bidimensional si se trata de un ajuste planimétrico y tridimensional en el caso de un ajuste GPS. Es posible también, realizar simultánemente un ajuste altimétrico y planimétrico usando como referencia un sistema tridimensional, pero en la práctica dada las distintas características de instrumentos, precisiones y de correcciones a realizar en los trabajos planimétricos y de nivelación, es común efectuar los ajustes separadamente.

En el caso del ajuste GPS, las observaciones nos vendrán dadas como fruto de un proceso de cálculo previo que tiene como objetivo la eliminación de errores de distinta naturaleza y que como producto final, arrojan unas observaciones constituidas por las denominadas baselíneas ΔX, ΔY, ΔZ y una matriz de varianza y covarianzas asociada, que expresa la precisión de las componentes de esta baselínea y la correlación existente entre esas componentes. Por la estructura constructiva del GPS el sistema de referencia es un sistema cartesiano geocéntrico válido a nivel global y que tiene como elipsoide de referencia el WGS84, como es bien conocido.

Para el caso del ajuste planimétrico es frecuente trabajar bien con un elipsoide como superficie de referencia o bien con un plano (como el definido por un sistema de representación cartográfico por ejemplo el plano conforme UTM). La primera opción permite un trabajo de mayor precisión y con mayor flexibilidad a la hora de abordar especialmente trabajos geodésicos de gran extensión, aunque a costa de un mayor esfuerzo de cálculo y complicación en la formulación de las observaciones sobre dicha superficie. La segunda opción, aunque menos rigurosa, puesto que asume los errores inherentes al sistema cartográfico que representa el elipsoide sobre una superficie plana, permite una formulación de las ecuaciones de observación en un plano cartesiano mediante expresiones muy sencillas.

En el caso del ajuste altimétrico es suficiente con la definición de un "nivel de referencia" unidimensional. Este puede ser la superficie de elipsoide, la superficie del geoide u otro que arbitrariamente se desee establecer. Ha de tenerse en cuenta que en la mayoría de las ocasiones en ingeniería no es tan impoprtante el posicionamiento absoluto como el relativo, y por tanto es frecuente el trabajo (tanto en los casos altimétricos como en los planimétricos) con sistemas de referencia locales arbitrariamente definidos.

Métodos de estimación mínimo cuadrática

En una estimación mínimo cuadrática se ven involucrados dos tipos de modelos, el matemático y el estadístico. El modelo estadístico hace referencia a la naturaleza de las observaciones tomadas en cuanto a su calidad y a las relaciones mutuas entre observaciones (correlación). Por simplicidad, es frecuente considerar que cada observación (que es considerada como una variable aleatoria) es independiente estadísticamente de las demás. La materialización concreta del modelo estadístico se produce a través de la denominada matriz de pesos, que se construye en base a las varianzas y covarianzas de las observaciones implicadas en el ajuste.

Ecuaciones de Observación

Como se dijo antes, las ecuaciones de observación que se plantean para los observables topográficos varían su forma, en función del sistema de referencia que se quiere establecer. En lo que sigue, se trabajará con el modelo de un sistema de referencia correspondiente a un plano cartesiano, para las observaciones que se consideran en el ajuste planimétrico y en un sistema geocéntrico para el caso de las baselíneas GPS. De modo que el sistema de referencia para el ajuste planimétrico será bidimensional, unidimensional para el ajuste altimétrico y tridimensional para el caso de las baselíneas GPS.

A consecuencia de esto para llevar las observaciones planimétricas originales hasta dicho plano serán necesarias correcciones meteorológicas, instrumentales, geométricas y en su caso, las correspondientes al sistema cartográfico empleado. Es posible plantear cualquier observación en un sistema tridimensional, bien en coordenadas cartesianas o en coordenadas curvilíneas aunque de forma general aumenta la complicación de la expresión de las ecuaciones de observación.

Restricciones y observaciones

Las ecuaciones de observación enumeradas anteriormente conforman el grupo de ecuaciones que habrán de establecerse en cada tipo de ajuste para construir, la denominada matriz de diseño o matriz A en el modelo de ecuaciones de observación $A \cdot X + L = V$. Ésta matriz recibe el sobrenombre de matriz de diseño, porque es posible formularla sin haber realizado observación alguna en campo. Una vez que se conozca la configuración geométrica aproximada de la red basta con asignar unos valores iniciales aproximados a los nodos de la misma y definir que observaciones se tratarán de medir para que sea posible plantear la matriz A. Ésta peculiaridad es empleada en el diseño y estudio de las características así como en la optimización de las redes topográficas.

Las observaciones anteriores son funciones matemáticas que definen relaciones **relativas** entre los parámetros que constituyen las incógnitas del ajuste. En realidad, si disponemos de una red entre cuyos nodos se han realizado una serie de observaciones, (por ejemplo distancias y ángulos) no es posible extraer un información **absoluta** como son los valores de los parámetros (coordenadas de los nodos de la red) sin añadir alguna otra fuente de información al conjunto de observaciones realizado. Matemáticamente, ésta imposibilidad se concreta en una deficiencia de rango en la matriz de coeficientes N del sistema de ecuaciones normales. Esto quiere decir que dicha matriz es singular, y por tanto no existe solución única al sistema normal.

La información adicional que es necesaria constituye lo que se denomina el **datum** de la red. La cantidad de información necesaria varía según el sistema de referencia elegido en cada caso. En la tabla siguiente se consignan el número de parámetros necesarios para lograr la definición de distintos tipos de sistemas de referencia.

TIPO DE RED	*Parámetros que definen el DATUM*	*Nº Parámetros necesarios*
Red altimétrica	*1 traslación Tz*	*1*
Red planimétrica Triangulada	*2 traslaciones Tx, Ty y 1 rotación* κ *y 1 factor de escala S*	*4*
Red planimétrica Trilaterada	*2 traslaciones Tx, Ty 1 rotación* κ	*5*
Red tridimensional (con distancias)	*3 traslaciones Tx, Ty, Tz 3 rotaciones* ω, φ, κ	*6*
Red tridimensional (sin distancias)	*3 traslaciones Tx, Ty, Tz 3 rotaciones* ω, φ, κ *y un factor de escala S*	*7*

Incluídas dentro de las observaciones existe cierta cantidad de información que resulta válida para la definición del datum en una red con un sistema de referencia dado. De manera que es posible, que parte de la información imprescindible para definir el datum venga incluída ya con las observaciones. Por ejemplo, en una red planimétrica triangular en la que sólo se hayan medido ángulos, es claro que cualquier triángulo que conserve ángulos entre sus lados iguales a los medidos, sea cual sea su tamaño y su posición en el plano, cumple las ecuaciones de observación y puede constitutir una solución al ajuste. Existen entonces 4 grados de libertad en el sistema que sería necesario fijar para definir una única solución, dos componentes de una traslación plana sobre la red, un factor de escala y una rotación que aún permanecen indeterminados. Sin embargo, si se ha medido una distancia entre dos de los vértices de dicha red, el tamaño del triángulo ya queda definido de modo que, al añadir la nueva observación de la distancia, el número de grados de libertad del sistema se reduce a 3, puesto que el factor de escala queda definido por la información presente en la observación distancia. De un modo semejante puede intuirse como en un sistema de referencia tridimensional, como es el sistema geocéntrico que se emplea en el ajuste de baselíneas GPS, el número de grados de libertad es de 7, (3 componentes de una traslación en el espacio, un factor de escala y tres rotaciones alrededor de los tres ejes).

Distintos tipos de observaciones aportan diferente información sobre el datum de la red. En la tabla siguiente puede apreciarse este hecho:

Observación	*Parámetros del Datum*						
	Tx	*Ty*	*Tz*	ω	φ	κ	*S*
Angulos y lecturas angulares	—	—	—	—	—	—	—
Distancias	—	—	—	—	—	—	√
Acimuts	—	—	—	—	—	√	—
Angulos cenitales	—	—	—	√	√	—	—
Desniveles	—	—	—	—	√	√	√
Posiciones GPS	√	√	√	√	√	√	√
Baselíneas GPS	—	—	—	√	√	√	√

En el caso más frecuente, la información que sobre el datum pueda extraerse, a partir de las observaciones disponibles, no será suficiente y por tanto la matriz *N* seguirá siendo singular.

Para eliminar esta singularidad es necesario añadir una serie de nuevas relaciones de los parámetros, denominadas **restricciones**, que siendo independientes entre sí (dichas relaciones) y con respecto las ecuaciones de observación permitan construir un sistema de solución única que será la solución mínimo cuadrática. Este proceso realiza lo que se denomina la definición del DATUM de la red.

En realidad las restricciones no son ecuaciones especiales (o pueden no serlo) distintas de las ecuaciones de observación, es más las ecuaciones de observación sirven para establecer restricciones. Lo que realmente determina la diferencia entre una ecuación de observación y una restricción es "la intensidad" con la que se exige el cumplimiento de esta ecuación. Por ejemplo, un acimut entre dos puntos de una red planimétrica puede constituir una observación o una restricción. En el segundo caso se exigirá que las coordenadas de los dos puntos que están relacionados por ese valor acimutal cumplan estrictamente la relación que representa la ecuación del acimut, de modo que no existirá residuo alguno para ese acimut considerado como restricción, mientras que si interviene en el ajuste como una ecuación de observación la relación se cumplirá, pero no exactamente y el valor del acimut calculado a partir de las coordenadas ajustadas será ligeramente diferente al observado (es decir existirá un residuo no nulo para esa observación). En realidad ni siquiera esta diferencia a la que se alude en el párrafo anterior, constituye verdaderamente una diferencia en algunos métodos de resolución, puesto que como se verá existen métodos, en los que es posible imponer una restricción ponderadamente sin exigir su cumplimiento estricto.

Cuando el número de restricciones independientes añadido al sistema de ecuaciones de observación es igual al necesario para definir el datum se dice que el ajuste es de restricción mínima, si el número de restricciones es mayor, el ajuste se denomina sobreconstreñido o sobrerrestringido. Este último tipo de ajuste debe aplicarse con cuidado, puesto que los errores que pudieran existir en las restricciones que definirán el datum, se transmitirán a la nueva red ajustada. A continuación se describen algunas de las restricciones que pueden imponerse con más frecuencia en un ajuste topográfico.

Métodos de imposición de restricciones

Fundamentalmente existen tres métodos para la imposición de las restricciones en el ajuste de una red y obtener por tanto la definición del datum con la consiguiente eliminación de la singularidad en el sistema normal. Aunque las denominaciones varían según diversos autores, los diferentes métodos, pueden nombrarse del modo que sigue: Método de eliminación de parámetros, Ajuste con funciones de parámetros y Método ponderado.

Medidas de Precisión y Fiabilidad

Una vez resuelto el ajuste de una red topográfica es necesario estudiar su calidad. Para ello es necesario describir los aspectos relativos a la precisión, fiabilidad y economía. Por precisión puede entenderse la descripción del modo en que la calidad de las observaciones afecta a los resultados del ajuste a través de la geometría de la red. La fiabilidad, de una red, sin embargo hace referencia a como reacciona la misma ante pequeñas desviaciones de las observaciones, se refiere a la robustez de la red, es decir a la capacidad de resistir errores groseros indetectables en las observaciones. O lo que es lo mismo "hace referencia a la controlabilidad de las observaciones, es decir la capacidad de detectar errores y de estimar los efectos de errores no detectados sobre la solución". [LEI95] . Para valorar estos aspectos es posible emplear múltiples indicadores, tanto globales como locales. Los más frecuentes son las elipses de error locales y relativas como indicadores de precisión y los números de redundancia interna y los parámetros de fiabilidad externa para la fiabilidad.

Análisis estadístico posterior al ajuste

El análisis estadístico posterior al ajuste se concentra en la detección de errores groseros de pequeña magnitud. Los errores de magnitudes grandes son fácilmente detectables puesto que producen grandes residuos en las observaciones de una zona concreta de la red e incluso en casos de resolución de problemas linealizados pueden ocasionar la no convergencia del proceso. Este análisis se basa en la realización de tests estadísticos sobre los residuos de las observaciones.

Test de bondad del ajuste

También denominado "Test global" sirve para determinar si la varianza de referencia a posteriori es compatible con la varianza de referencia a priori. El estadístico empleado en el test es:

$$y = \frac{r \cdot \hat{\sigma}_0^2}{\sigma_0^2}$$

donde r es el número de grados de libertad del ajuste. Bajo la hipótesis nula puede demostrarse que el estadístico sigue una distribución $\chi^2(\ )$ con r grados de libertad, de modo que como la esperanza matemática de $\chi^2(\ )$ es r, podemos poner que:

$$E\{y|H_0\} = r \Rightarrow E\left\{\frac{\hat{\sigma}_0^2}{\sigma_0^2}\middle|H_0\right\} = 1 \Rightarrow E\{\hat{\sigma}_0^2|H_0\} = \sigma_0^2$$

Por tanto la hipótesis nula del test afirma que la igualdad estadística entre las varianzas de referencia a priori y a posteriori. Si fijamos un nivel de significación α para el test de hipótesis de dos colas, la hipótesis nula será aceptada si el estadístico cumple que:

$$\chi^2_{\alpha/2}(r) \leq y \leq \chi^2_{1-(\alpha/2)}(r)$$

Como causas del fallo del test anterior pueden considerarse una distribución incorrecta de los pesos de las observaciones o bien la existencia de errores groseros en las observaciones.

Para poder discernir cual de éstas dos posibles causas provoca el fallo del test, es conveniente analizar el vector de los residuos. La existencia de un error grosero en las observaciones se traducirá en unos residuos de magnitudes elevadas y con media diferente de cero. Si por el contrario, los residuos parecen de magnitudes razonables debe pensarse en la primera de las causas.

El estadístico y, depende exclusivamente del vector residuos y de la matriz varianza-covarianza de las observaciones, tal y como puede verse al recordar cual es la expresión del cálculo de la varianza de referencia a posteriori:

$$\left.\begin{aligned} y &= \frac{r\cdot\hat{\sigma}_0^2}{\sigma_0^2} \\ \hat{\sigma}_0^2 &= \frac{v^T\cdot P\cdot v}{r} = \frac{v^T\cdot\sigma_0^2\cdot\Sigma^{-1}\cdot v}{r} \end{aligned}\right\} \Rightarrow y = v^T\cdot\Sigma^{-1}\cdot v$$

Por tanto el estadístico es inversamente proporcional a la matriz varianza-covarianza de las observaciones. Un valor reducido del estadístico indicará que la varianza asociada a alguna de las observaciones es demasiado grande de modo que se ha considerado alguna de las observaciones menos precisa de lo que lo es en realidad. Por el contrario un valor del estadístico elevado indicará una varianza demasiado pequeña en alguna observación, habiéndose por tanto sobreestimado la precisión de alguna de las observaciones. En este caso se recomienda una reconstrucción de la matriz varianza-covarianza de las observaciones.[LEI95].

"Data Snooping". Test de Baarda

Bajo esta denominación se encuentra una técnica que combina la detección de los residuos anormalmente grandes bajo un cierto criterio estadístico (outlier) y la localización del error grosero y su eliminación. La existencia de errores groseros en los datos, es aún una hipótesis alternativa a la hipótesis nula considerada en el test de bondad del ajuste, por ello Baarda trabaja bajo el supuesto de que solamente una observación de cada vez está afectada por un error grosero. El estadístico elegido para el test de hipótesis es:

$$w_i = \frac{-v_i}{\sqrt{r_i}\cdot\sigma_{l_i}} = \frac{v_i}{\sigma_{w_i}} = \frac{v_i}{\sigma_0\cdot\sqrt{q_{w_i}}}$$

donde $_w$ es el elemento de la diagonal de la matriz cofactor de los residuos correspondiente a la observación i-ésima, σ_w es la desviación típica del residuo de la observación i-ésima, es el número de redundancia de la observación y σ es la desviación típica de referencia a priori [LEI95]. se puede demostrar que bajo la hipótesis nula el estadístico w_i está normalmente distribuido con media cero y varianza unidad, de modo que:

$$w_i | H_0 \in n(0,1)$$

Si realizamos un test de dos colas para un nivel de significación α tendremos que H_0, es aceptada si el estadístico se encuentra en el intervalo siguiente:

$$N_{\alpha/2}(0,1) \leq w_i \leq N_{1-(\alpha/2)}(0,1)$$

Varios autores [HAR91] , [WOL97] afirman que un valor de 3.29 funciona bien como criterio de rechazo de residuos con observaciones asociadas conteniendo errores groseros. Este valor corresponde a un nivel de significación de $\alpha = 0.001$. En la práctica, por tanto, una observación deberá ser marcada como sospechosa de contener un error grosero cuando $|w_i| > 3,29$.

Una vez que se ha detectado un residuo anormal (outlier) según el criterio de Baarda, es necesario localizar la observación que contiene esos errores. En principio debe suponerse que al menos el número máximo de errores groseros en las observaciones no excede el número de grados de libertad del sistema. La detección del error va a depender fuertemente de la geometría del sistema más teniendo en cuenta que se trata de localizar errores groseros de pequeña magnitud.

A causa de la correlación entre los residuos estimados un residuo que no cumple el test de Baarda puede indicar un error grosero en la observación que le corresponde si tiene un valor dominante su número de redundancia en la fila correspondiente dentro de la matriz $R = Q_V \cdot P$.

τ -test

El test de Baarda requiere que se conozca la varianza de referencia a priori o lo que es lo mismo que las varianzas de las observaciones estén bien escaladas. Si este valor a priori no es bien conocido o no se quiere depender de él entonces se emplea el valor de la varianza a posteriori el proceso de "Data Snooping" es modificado con el nuevo estadístico propuesto por Pope en 1976. La expresión del estadístico es:

$$\tau_i = \frac{v_i'}{\sigma_{w_i}} = \frac{v_i}{\hat{\sigma}_0 \cdot \sqrt{q_{w_i}}}$$

Este estadístico sigue una distribución τ (tau) con r grados de libertad, de modo que un residuo de una observación será marcado si el estadístico supera el valor crítico del test para el grado de significación elegido, es decir:

$$|\tau| > \tau_{\alpha/2}$$

donde:

$$\tau_{\alpha/2}(r) = \frac{\sqrt{r} \cdot t_{\alpha/2}(r-1)}{\sqrt{r-1+t_{\alpha/2}^2(r-1)}}$$

El problema con este test es que como la varianza de referencia está afectada por la presencia de errores groseros entre los datos cuanto mayor es esta varianza menor es el valor del estadístico y es mayor la probabilidad de no detectar errores groseros.

La distribución τ (tau) converge hacia la t de Student o la normal cuando el número de grados de libertad tiende a ∞. [LEI95]. Leick presenta varios gráficos con valores límite para el test con distintos grados de libertad en los que escoge como nivel de significación del test $\alpha = 0.05$.

REFERENCIAS

1. [BJÖ96] Björck, Ake. (1996). 'Numerical Methods for Least squares Problems'. Society for Industrial and Applied Mathematics.
2. [CAS88] Caspary, W.F.(1988). 'Concepts of network and deformation analysis'. The University of New South Wales.Monograph 11.
3. [COO87] Cooper, M.A.R. (1987). 'Control surveys in civil engineering'. Collins Profesional and Technical Books.
4. [CRO83] Cross, P.A. (1983). 'Advanced least squares applied to position fixing'. North East London Polytechnic. Working papers.
5. [CHU96c] Chueca Pazos, Manuel; Herráez Boquera, José; Berné Valero, José Luis. (1996). 'Redes Topográficas y locales. Microgeodesia'. 'Tratado de Topografía. Tomo III '. Paraninfo.
6. [HAR91] Harvey, Bruce R. (1991). 'Practical Least Squares and statistics for surveyors'. The University of New South Wales. Monograph 13.
7. [LEI95] Leick, Alfred. (1995). 'GPS Satellite surveying'. Second edition .John Willey & Sons.
8. [MIK76] Mikhail Edward M., Ackermann F. (1976).'Observations and Least Squares'.Harper & Row Publishers.
9. [STR97] Strang, Gilbert; Borre Kai. (1997). 'Linear Algebra, Geodesy and GPS'. Wellesley Cambridge Press.

10. [WOL97] Wolf Paul R., Ghilani Charles D. (1997). 'Adjustment Computations'. Wiley Interscience.

Recenzent: doc. dr inż. Tadeusz Dziura

Abstract

During measurement of terrain topography, due to character of methods used as well as measurement equipment, one obtains inaccurate observations complex; this fact forces to develop a method which allows to reduce errors. In the paper, general considerations on equalisation the topographic observations by means of the least square method are presented. Some theoretical assumptions are presented, i.e. such ones which were used in programme (Designing and equalisation of topographic networks) developed in Zakład Inżynierii Kartograficznej, Geodezji i Fotogrametrii Wydziału Górniczego (Cartography Engineering, Geodesy and Photogrammetry Plant of Mining Faculty) in Oviedo.