

Análise da Distribuição e Correlação das Observações GPS

Mestrando Alberto Luis da Silva ¹
Prof.^a Dr.^a Verônica Maria Costa Romão ²
Prof. Dr. Tarcísio Ferreira da Silva ³
Prof. Dr. Antônio Simões Silva ⁴
Prof. Dr. Robert Wayne Samohyl ⁵

UFPE – Depto. de Engenharia Cartográfica
50740-530 Recife PE

¹ silvabeto@zipmail.com.br

² vcosta@ufpe.br

³ tarcisiofs@yahoo.de

⁴ UFV – Depto. de Engenharia Civil
36571-000 Viçosa MG

asimoes@ufv.br

⁵ UFSC – Depto. de Engenharia de Produção
88010-970 Florianópolis SC
samohyl@deps.ufsc.br

Resumo: Com a utilização cada vez maior do Sistema de Posicionamento Global (GPS), na determinação precisa de posições, torna-se de fundamental importância a consideração de alguns fatores que, ao serem negligenciados, podem prejudicar a qualidade dos resultados. No processamento e ajustamento das observações GPS utilizando o método dos mínimos-quadrados, considera-se que as observações têm propriedade de ruído branco, ou seja, são não-correlacionadas e que são normalmente distribuídas. Nem sempre essas propriedades são satisfeitas, principalmente quando as observações são obtidas com intervalos de gravação muito pequenos, no caso da autocorrelação; ou através do processamento de longas linhas de base, no caso da distribuição normal.

O objetivo deste artigo, é verificar a presença de autocorrelação nos resíduos do processamento da linha de base, com o método dos mínimos-quadrados, e analisar seu efeito na precisão das coordenadas, quando ela passa a ser considerada. Será realizado também, através dos momentos empíricos e testes estatísticos, a verificação da distribuição das observações, no que diz respeito a sua normalidade.

Palavras-chaves: GPS, autocorrelação, não-normalidade.

Abstract: With the increasing use of the Global Positioning System (GPS), in the determination of precise positions, it becomes of fundamental importance the consideration of some factors, that when they are neglected, the quality of the results can be damaged. In the processing and adjustment of the GPS observations based on the method of the least-squares, it is considered that the observations have the property of white noise, in other words, they are non-correlated and that they are normally distributed. Not always those properties are satisfied, mainly when the observations are obtained with very small recording intervals, in the case of the autocorrelation; or through the processing of long baselines, in the case of the normal distribution.

The objective of this work, is to verify the presence of autocorrelation in the residuals of the processing of the baseline, based on the least-squares method, and to analyze its effect in the precision of the coordinates, when it is considered. The normality of the residuals is also investigated through the empirical moments and statistical tests.

Keywords: GPS, autocorrelation, non-normality

1 Introdução

A utilização do Sistema de Posicionamento Global (GPS) tornou-se a principal técnica empregada no posicionamento geodésico graças a sua facilidade de manuseio e a não necessidade de intervisibilidade entre os pontos, principalmente no caso de implantação de redes geodésicas. Mas, será que a precisão das coordenadas obtidas com o sistema GPS são realmente confiáveis ao nível de precisão que se tem anunciado?

Alguns fatores são negligenciados no processamento das observações GPS, colocando a qualidade dos resultados sob suspeita. Um desses fatores é a autocorrelação das observações, que podem estar presentes devido, entre outros, ao processamento interno das observáveis no receptor (Borre & Tiberius, 2000), e/ou devido a diferentes condições atmosféricas percorridas pelo sinal GPS. Uma outra característica que é assumida no processamento e ajustamento das observações GPS, é quanto à estocasticidade dos resíduos, o que, às vezes, pode não ser satisfeita, principalmente quando se processam longas linhas de base. Essas considerações podem influenciar na qualidade do resultado, ou seja, na qualidade das coordenadas finais.

As observações GPS, assim como a maioria das observações geodésicas, são processadas utilizando o algoritmo dos mínimos-quadrados, em que normalmente este algoritmo não considera a autocorrelação das observações de ruído. No modelo estocástico das observáveis, ou seja, na definição de sua matriz variância-covariância, normalmente se supõe que todas as observáveis apresentam iguais características de ruído e que elas não variam com o tempo, ou seja, o processo randômico é considerado estacionário. Além disso, consideram-se também que os ruídos são brancos, ou seja, duas observações realizadas em duas épocas consecutivas são não correlacionadas entre si.

O objetivo deste artigo é, no contexto de posicionamento relativo preciso GPS, verificar a presença de autocorrelação nos resíduos oriundos do processamento das linhas de base realizado pelo método dos mínimos quadrados, e analisar este efeito na precisão das coordenadas obtidas, quando considerado. Além disso, será feita uma verificação quanto à normalidade das observações GPS, utilizando o terceiro e o quarto momentos amostrais, e aplicando os testes estatísticos de Komogorov-Smirnov e Anderson-Darling.

2 Ajustamento das observações pelo Método dos Mínimos Quadrados

Sabemos que todas as observações conduzidas pelo homem, contêm erro. Daí a necessidade de se realizar mais de uma medida, com o objetivo de, a partir de uma pluralidade de dados, obtermos um valor que mais se aproxima do valor real. É a partir deste pressuposto que se utiliza o método dos mínimos quadrados que tem como princípio fundamental a minimização da soma dos quadrados dos resíduos, ou seja:

$$\sum_{i=1}^n v_i^2 = \min \quad \text{onde } v_i = x - l_i$$

onde v_i são os resíduos que quando somados com os valores observados l_i , reproduzem o valor estimado x .

Como as observações GPS, assim como todas as outras medições, possuem erros, é preciso realizar um ajustamento a partir das observações redundantes, com o objetivo de se extrair um resultado único e consistente. O modelo de ajustamento utilizado no processamento de observações GPS, é o ajustamento pelo método dos mínimos quadrados, baseado sobre equações onde as observações são expressas como uma função dos parâmetros. Este modelo é dito paramétrico e é definido a partir da seguinte expressão:

$$L_a = F(X_a)$$

onde os valores observados ajustados podem ser expressos como função dos parâmetros ajustados. A partir desta expressão acima, obtemos o modelo matemático linearizado de Gauss-Markov do método dos parâmetros:

$$V = AX + L \quad ,$$

sendo A a matriz design (ou matriz dos coeficientes) formada pelas derivadas parciais

$$A = \frac{\partial F}{\partial X_a} \Big|_{X_o, L_b}$$

e L a diferença entre as observações aproximadas e as observações brutas

$$L = L_o - L_b$$

Minimizando a forma quadrática fundamental, obtemos o valor de X

$$X = -(A^T P A)^{-1} A^T P L$$

onde P corresponde a matriz dos pesos e X corresponde a correção que devemos realizar sobre os parâmetros aproximados X_o .

$$X_a = X_o + X$$

Após a determinação dos parâmetros de interesse, utiliza-se um modelo estocástico para determinar a precisão destes, através da matriz variância-covariância, \sum_{X_a} , a partir da seguinte expressão:

$$\sum_{X_a} = \sigma_o^2 (A^T P A)^{-1}$$

onde σ_o^2 é a variância a posteriori da unidade de peso (Gemael, 1994).

3 Análise da distribuição dos dados

3.1 Momento amostral

Momentos são características da função densidade de probabilidade, sendo que os dois primeiros, média (m_1) e desvio padrão (s) especificam completamente a distribuição normal. Além destes dois, existem o terceiro momento e o quarto momento, denominados assimetria (skewness) e curtose (kurtosis) respectivamente, que são utilizados para analisar a normalidade da distribuição, e são obtidos a partir da razão de momentos, e por isso, definido como momentos padronizados (Tiberius & Borre, 1999).

3.1.1 Coeficiente de assimetria (skewness)

O coeficiente de assimetria ($\sqrt{b_1}$) é uma medida de simetria, ou seja, ela indica o quanto a distribuição é assimétrica. A distribuição normal é simétrica em torno de sua média, e portanto o coeficiente de assimetria, neste caso, é aproximadamente zero. Para um nível de significância desejado, é obtido através de tabelas, o valor crítico do coeficiente, definindo uma região onde se julga ser normal, caso o coeficiente de assimetria esteja dentro desta.

A equação que define o coeficiente de assimetria é a seguinte (Tiberius & Borre, 1999):

$$\sqrt{b_1} = \frac{\frac{1}{N-1} \sum_{i=1}^N (x_i - m_1)^3}{\left(\frac{1}{N-1} \sum_{i=1}^N (x_i - m_1)^2 \right)^{\frac{3}{2}}} \text{ onde } N \text{ é o número de resíduos.}$$

3.1.2 Coeficiente de curtose (kurtosis)

O coeficiente de curtose (b_2) mede a relação entre a cauda e o pico de uma distribuição, e assume o valor 3 para uma distribuição normal. Assim como o coeficiente anterior, ele também possui uma região crítica, definido a partir de um nível de significância desejado e obtido através de tabela. Caso o coeficiente de curtose esteja localizado nesta região, considera-se que a distribuição estudada segue uma distribuição normal.

A equação que define o coeficiente de curtose é a seguinte (Tiberius & Borre, 1999):

$$b_2 = \frac{\frac{1}{N-1} \sum_{i=1}^N (x_i - m_1)^4}{\left(\frac{1}{N-1} \sum_{i=1}^N (x_i - m_1)^2 \right)^2}$$

3.2 Função Distribuição Empírica

Será discutido neste artigo, dois testes estatísticos que usam a função de distribuição empírica e compara com a distribuição normal, verificando o quanto uma se ajusta com a outra, baseados em medidas de divergências. Dentre vários testes encontrados na literatura, os dois que serão discutidos aqui, são os testes de Kolmogorov-Smirnov e Anderson-Darling.

3.2.1 Teste Kolmogorov-Smirnov

O teste estatístico (D) de Kolmogorov-Smirnov (KS), é uma medida simples definida pelo valor máximo da diferença absoluta entre a função distribuição acumulativa empírica, $F_N(x)$, e a distribuição teórica (distribuição normal, $F_0(x)$)

$$D = \max_{-\infty < x < +\infty} |F_N(x) - F_0(x)|$$

O valor de D assumirá valores entre 0 e 1; se for igual a 0, há uma igualdade perfeita entre os dados empíricos e os dados teóricos (Tiberius & Borre, 1999). Assim como nos momentos empíricos, este teste também possui uma região crítica, que é definida a partir de um valor D_k que é tabelado e que depende do tamanho da amostra, além do nível de significância. Caso $D > D_k$ rejeita-se a hipótese de distribuição normal.

3.2.2 Teste Anderson-Darling

O teste estatístico (A) de Anderson-Darling (AD), em vez de utilizar uma simples medida de distância (diferença) como realizado no teste (KS), emprega uma maior quantidade de diferenças, fornecendo a diferença quadrática integrada (Tiberius & Borre, 1999)

$$A^2 = -N - \frac{1}{N} \sum_{i=1}^N \{(2i-1)(\log(u_i) + \log(1-u_{N+1-i}))\}$$

onde $u_i = F_0(x_i)$, é a variável uniforme no intervalo $[0,1]$, transformada a partir da variável amostral x_i .

Este teste também possui uma região crítica, que é definida a partir de um valor A_{mod}^2 tabelado e que depende da quantidade de observações, além do nível de significância. Caso $A^2 > A_{mod}^2$ rejeita-se a hipótese de distribuição normal.

4 Correlação Serial

4.1 Série estacionária

Quando uma série de dados temporal, não apresentar tendências ou essas já estiverem sido removidas, dizemos que se trata de uma série estacionária. Para melhorar o conceito de estacionariedade, consideremos a série Z_1, Z_2, \dots, Z_N com N variáveis aleatórias tendo distribuição conjunta $f(Z_1, Z_2, \dots, Z_N)$ e distribuições marginais $f(Z_1), \dots, f(Z_N)$. Daí, a série é dita estacionária se para qualquer instante de tempo t e qualquer m , tivermos (Morettin & Pedro, 1987)

$$f(z_t) = f(z_{t+m}) \text{ ou}$$

$$f(z_t, z_{t+1}) = f(z_{t+m}, z_{t+m+1}), m = \pm 1, \pm 2, \dots$$

ou seja, as distribuições de Z_1, Z_2, \dots, Z_N , são invariantes por translações no tempo e conseqüentemente a média e a variância são constantes, independente do instante t .

4.2 Autocovariância e autocorrelação

Consideremos as observações Z_1, Z_2, \dots, Z_N de uma série temporal. A função de autocovariância, γ_j , entre os pares de observações de lag j (ou período j ou ordem j) da série temporal é :

$$\gamma_j = \text{cov}(Z_t, Z_{t+j}) = \text{cov}(Z_{t+m}, Z_{t+m+j})$$

onde m é um valor inteiro. Supondo que a série seja estacionária, isto é, não há tendência presente ou essa tendência já foi removida preliminarmente, a função de autocovariância (γ_j) torna-se (Borre & Tiberius, 1999):

$$\gamma_j = E\{(Z_t - \mu)(Z_{t+j} - \mu)\}$$

A autocovariância depende somente da diferença de tempo entre o instante $t+j$ e t , e essa diferença é chamada lag.

Uma importante ferramenta para analisar as propriedades de uma série temporal é fornecida por uma série de quantidades chamada coeficientes de autocorrelação amostral, a qual mede a correlação entre observações separadas por diferentes distâncias (lags). A função de autocorrelação é definida como (Borre & Tiberius, 1999):

$$\rho(j) = \frac{\gamma(j)}{\gamma(0)}$$

e representa o coeficiente de correlação entre pares de $Z(t)$ e $Z(t+j)$, separados por um intervalo j . A autocovariância para lag zero é:

$$\gamma(0) = \sigma^2$$

é a variância da série Z_i .

Para ajudar a interpretar um conjunto de coeficientes de autocorrelação, utiliza-se um gráfico chamado correlograma mostrado na figura 1 (Borre & Tiberius, 2000), no qual $\rho(\tau)$ está em função do lag (s):

Repare que a autocorrelação para lag (0) é igual a 1, ou seja, quando o intervalo entre duas observações é igual a zero, considera-se a mesma observação, e portanto, autocorrela-

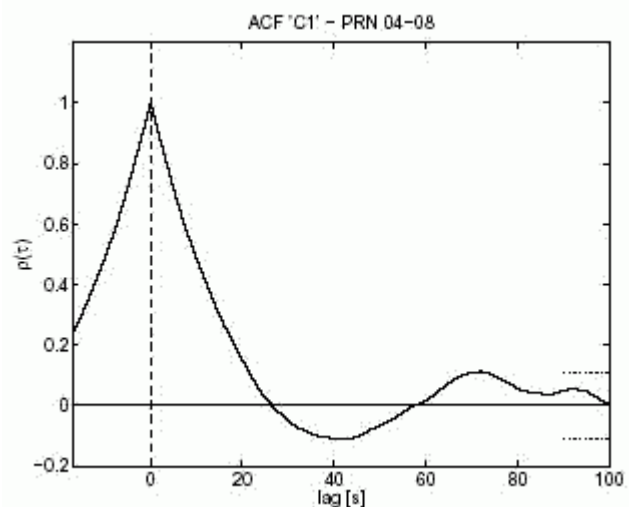


Figura 1: Coeficientes de autocorrelação amostral para o código C1 dos satélites PRN04-08

ção 1. No caso acima, a medida que aumenta o intervalo entre observações, a autocorrelação diminui tendendo para zero. Isso é uma característica de séries estacionárias.

5 Experimento

Foi realizado no campus da Universidade Federal de Pernambuco (UFPE) em diferentes dias, observações GPS, utilizando o receptor de dupla frequência TRIMBLE 4000SST, com taxa de gravação de 15 segundos e duração de aproximadamente 4 horas. Essas observações foram realizadas em um mesmo ponto e considerando a mesma configuração geométrica dos satélites (DOP), com o objetivo de processar as observações GPS sob as mesmas condições. Utilizou-se também, as observações do ponto RECF pertencente a Rede Brasileira de Monitoramento Contínuo (RBMC), localizado também no campus da UFPE, e que possui o receptor de dupla frequência TRIMBLE 4000SSI, com taxa de gravação de 15 segundos. Processando estes dois pontos simultaneamente, determina-se uma linha de base com aproximadamente 400 metros para cada dia de observação.

As análises serão feitas sobre os resíduos dos mínimos-quadrados, obtidos a partir de um ajustamento realizado sobre os dados do processamento de uma linha de base. Como as observações GPS foram realizadas sob as mesmas condições, utilizando os mesmos receptores e antenas, podem-se comparar os resultados verificando se eles se mantêm para dias diferentes sob condições iguais.

Tanto as análises estatísticas quanto os cálculos dos momentos empíricos, assim como o cálculo dos coeficientes de autocorrelação, serão realizados utilizando rotinas desenvolvidas no MATLAB e ferramentas do EXCEL. Além disso, para a obtenção dos resíduos do ajustamento realizado no processamento da linha de base, será utilizado também um programa desenvolvido no MATLAB.

6 Conclusão

Após a realização do experimento, conclusões sobre a distribuição das observações GPS serão feitas, com o intuito de validar ou não a suposição de normalidade dos dados, para as diversas condições. Além disso, será feito um estudo sobre a sua influência nos resultados do processamento das observações GPS.

Conclusões também serão feitas quanto à correlação das observações GPS, verificando se elas obedecem à lei de ruído branco ou são autocorrelacionadas, quantificando, caso exista, a variação na precisão das coordenadas quando se considera e quando não se considera esse fator.

7 Referências Bibliográficas

Borre, K.; Tiberius, C.: *Time Series Analysis of GPS Observables*, Proc. of the 13th International Technical Meeting of the Satellite Division of the Institute of Navigation, Salt Lake City, Utah, September 19-22, 2000, pp. 1885-1894 (2001).

Chatfield, C.: *The analysis of time series, an introduction*, CHAPMAN & HALL, 1996.

Gemael, C. *Introdução ao ajustamento de observações: aplicações geodésicas*, Editora da UFPR, Curitiba, 1994.

Monico, J.F.G.: *Posicionamento pelo NAVSTAR-GPS: Descrição, fundamentos e aplicações*, UNESP, Presidente Prudente, 2000.

Morettin, P.A.; Tolo, C.M.: *Séries temporais*, Editora Atual, São Paulo, 1987.

Strang, G.; Borre, k.: *Linear algebra, geodesy and GPS*. Wellesley-Cambridge Press. 1997.

Tiberius, C.C.J.M.; Borre, K.: *Are GPS data normally distributed?*, International Association of Geodesy Symposia; 121, pp. 243-248, Berlin, 2000.

Tiberius, C.C.J.M.; Borre, K.: *Probability distribution of GPS code and phase data*. Zeitschrift für Vermessungswesen, 1999. 124(8): 264-273.