

ESTADO DA ARTE NO CONTROLE DE QUALIDADE DO AJUSTAMENTO DE OBSERVAÇÕES: Confiabilidade Interna e Externa no contexto dos múltiplos *outliers*

IVANDRO KLEIN (1)
MARCELO TOMIO MATSUOKA (1)
SERGIO FLORENCIO DE SOUZA (1)

(1) Universidade Federal do Rio Grande do Sul
Instituto de Geociências
Departamento de Geodésia, Porto Alegre – RS
LAGEO – Laboratório de Pesquisas em Geodésia
ivandroklein@gmail.com
tomiomatsuoka@gmail.com
sergiovlorenciodesouza@gmail.com

Na estimação por mínimos quadrados em modelos lineares, o vetor das n -observações (ou medidas) realizadas se relaciona com o vetor dos u -parâmetros incógnitos que se deseja estimar por meio de um modelo matemático da forma $y_{nx1} = A_{nxu}x_{ux1}$, onde y_{nx1} é o vetor das n -observações realizadas, x_{ux1} é o vetor dos u -parâmetros incógnitos do modelo, e A_{nxu} é a matriz *design* que relaciona linearmente as observações com os parâmetros que se deseja estimar. O sistema deve ser redundante, isto é, o número de observações deve ser maior do que o número de incógnitas ($n > u$), para não existência de solução única, possibilitando o controle de qualidade das medidas. Admitindo a inconsistência do sistema de equações, isto é, a presença de inevitáveis erros aleatórios, inerentes ao processo experimental de medições, adiciona-se um vetor de resíduos e_{nx1} ao modelo linear, resultando em $y_{nx1} = A_{nxu}x_{ux1} + e_{nx1}$. Com a adição dos n resíduos e_{nx1} ao sistema original, este se torna um sistema consistente, porém com infinitas soluções, pois cada estimativa distinta do vetor dos parâmetros $\hat{x}_{ux1}$ resulta em um vetor dos resíduos $\hat{e}_{nx1}$ diferente. Desta forma, a estimação por mínimos quadrados adota como solução única para o vetor dos parâmetros, aquele vetor $\hat{x}_{ux1}$ que minimiza a soma ponderada do quadrado dos resíduos, isto é, aquele vetor $\hat{x}_{ux1}$ que apresenta $\hat{e}_{nx1}^T W_{nxn} \hat{e}_{nx1} = \text{Mínimo}$, sendo W_{nxn} a matriz peso que foi definida para as observações (observações com peso maior, terão maior importância no processo de estimação). Usualmente, a matriz peso W_{nxn} é tomada como sendo o inverso da matriz de covariâncias das observações y_{nx1} , ou seja, quanto maior a precisão (menor variância), maior o peso de uma observação y_i de y_{nx1} . Após a estimação por mínimos quadrados ter sido realizada, é possível a detecção, e até mesmo a identificação de erros grosseiros (isto é, não aleatórios), no vetor das observações y_{nx1} , por meio de testes de hipóteses. Os testes estatísticos usados para este fim são formulados de acordo com a razão de verossimilhança generalizada, e baseiam-se essencialmente na magnitude dos resíduos $\hat{e}_{nx1}$. Entretanto, pelo fato da verdadeira magnitude dos erros que contaminam as observações ser sempre desconhecida, além dos testes estatísticos possuírem níveis de probabilidade associados, pode haver a presença de possíveis erros grosseiros nas observações, de pequena magnitude, não detectados pelo procedimento de teste utilizado. Desta maneira, uma forma de se obter um controle de qualidade dos resultados dos testes é por meio da teoria de confiabilidade. A confiabilidade interna é uma medida do MDB (*minimal detectable bias* – menor erro detectável) em uma observação y_i de y_{nx1} ($i = 1, 2, \dots, n$), segundo o procedimento de teste utilizado, e a confiabilidade externa reflete a influência de possíveis erros não detectados (MDB) no resultado final do processo de estimação, isto é, no vetor dos parâmetros estimados $\hat{x}_{ux1}$. A teoria convencional de confiabilidade, desenvolvida no final da década de 1960, foi formulada pressumindo-se que apenas um erro grosseiro (*outlier*) contamina o vetor das observações y_{nx1} , um pressuposto que nem sempre está correto e dificilmente ocorre na prática. Atualmente, diversos trabalhos vêm sendo publicados, apresentando extensões da teoria da confiabilidade, para que esta possa ser aplicada presumindo-se a existência de múltiplos *outliers* (simultaneamente) no vetor das observações y_{nx1} , por meio da derivação de medidas adequadas de confiabilidade interna (MDB) e confiabilidade externa para estes casos. O menor erro detectável (MDB) em uma observação y_i depende do parâmetro de não centralidade do modelo (separação entre a hipótese nula – ausência de erros grosseiros nas observações e a hipótese alternativa – presença de erros grosseiros nas observações), que por sua vez depende do número de *outliers* considerado, do poder e do nível de significância arbitrado para os testes de detecção e identificação de erros, da matriz *design* A_{nxu} (da geometria/configuração do problema), e da matriz peso W_{nxn} que foi definida para as observações y_{nx1} . Dentro deste contexto, o objetivo do presente trabalho é realizar uma pesquisa na área de controle de qualidade aplicado ao ajustamento de redes geodésicas,

por meio da simulação de diferentes cenários em uma rede GPS (*global positioning system*) tridimensional, alterando a geometria da rede (número de observações e parâmetros), a precisão das observações (matriz de covariâncias das observações), o número de injunções no sistema, bem como a inserção de erros propositais, verificando a influência destas questões na detecção e identificação de erros no vetor das observações y_{nx1} , nas medidas de confiabilidade interna e externa, bem como na qualidade final dos parâmetros estimados da rede (coordenadas dos pontos desconhecidos). Nestas diferentes simulações, serão aplicadas medidas de confiabilidade interna e externa utilizando a teoria clássica de confiabilidade (existência de um único *outlier* por vez em y_{nx1}), bem como extensões alternativas da teoria de confiabilidade, presumindo-se a existência simultânea de múltiplos *outliers* em y_{nx1} , confrontando os resultados obtidos em busca de conclusões a respeito de como a geometria da rede, a precisão das observações, o número de injunções, e o número de *outliers* considerado nos testes de hipóteses podem influenciar na detecção e identificação de possíveis erros não aleatórios no vetor das observações y_{nx1} , bem como verificar as vantagens e desvantagens de se utilizar as abordagens para múltiplos *outliers* em relação à teoria convencional de confiabilidade.