

AJUSTAMENTO DE OBSERVAÇÕES: UMA INTERPRETAÇÃO GEOMÉTRICA PARA O MÉTODO DOS MÍNIMOS QUADRADOS

Adjustment of observations: A geometric interpretation for the least squares method

IVANDRO KLEIN¹
MARCELO TOMIO MATSUOKA¹
SERGIO FLORENCIO DE SOUZA¹
MAURICIO ROBERTO VERONEZ²

¹Laboratório de Pesquisas em Geodésia (LAGEO)/Departamento de
Geodésia/Instituto de Geociências – UFRGS
Av. Bento Gonçalves, 9500, CEP 91501-970,
Cx. Postal 15001, Porto Alegre, RS, Brasil
ivandroklein@gmail.com; tomiomatsuoka@gmail.com;
sergioflorenciodesouza@gmail.com

²Laboratório de Sensoriamento Remoto e Cartografia Digital – LASERCA
Universidade do Vale do Rio dos Sinos - UNISINOS
Av. Unisinos, 950, CEP 93.022-000,
São Leopoldo, RS, Brasil
veronez@unisinos.br

RESUMO

O Método dos Mínimos Quadrados (MMQ) é um dos critérios mais utilizados para o ajustamento de informações onde o número de observações é superabundante e o sistema de equações, devido à presença de erros no processo experimental de medições, inconsistente. Este método adota como solução única para tais problemas aquela que minimiza a soma do quadrado dos erros aleatórios, e é largamente empregado em aplicações geodésicas. A fundamentação teórica envolvida no MMQ, contendo o desenvolvimento matemático do método, é amplamente difundida na comunidade geodésica. Porém, neste artigo, o objetivo é apresentar uma revisão teórica sobre a interpretação geométrica do MMQ, e uma solução alternativa de cálculo para este, advinda desta interpretação. Dois exemplos também são apresentados, onde a interpretação geométrica foi demonstrada numericamente

para problemas envolvendo duas observações e um parâmetro e também quatro observações e dois parâmetros.

Palavras-chave: Ajustamento de Observações; Método dos Mínimos Quadrados; Interpretação Geométrica do MMQ.

ABSTRACT

The Least Squares Method (LSM) is one of the most useful criteria for the adjustment of data where the number of observations is redundant and the system of equations, due to the presence of errors in experimental measurements, inconsistent. This method takes as a single solution that which minimizes the sum of the squared random errors, and is widely used in geodetic applications. The theoretical foundation involved in the LSM, containing the mathematical development of the method, is widespread in the geodetic community. However, in this paper, the goal is to present a theoretical review of the geometric interpretation of the LSM, and an alternative computation solution, derived from this interpretation. Two examples are also shown, where this geometric interpretation has been demonstrated numerically for problems involving two observations and a parameter, and also four observations and two parameters.

Keywords: Adjustment of Observations; Least Squares Method; LSM Geometric Interpretation.

1. INTRODUÇÃO

O Método dos Mínimos Quadrados (MMQ) é um dos critérios mais utilizados para o ajustamento de dados onde o número de observações é superabundante e o sistema de equações, devido à presença de erros no processo experimental de medições, inconsistente. Segundo Dalmolin (2002), “o ajustamento é um ramo da matemática aplicada que tem por objetivo a solução única para problemas onde o número de observações é superabundante e o sistema de equações lineares é inconsistente”. Consiste numa expansão do método dos mínimos quadrados desenvolvidos independentemente por Gauss (1795) e Legendre (1805). Não faz sentido falar em ajustamento para problemas onde os dados (observações ou medidas) não excedem o mínimo requerido para solução.

As observações possuem uma propriedade inerente a elas, conhecida como flutuações probabilísticas, pois, quando se repete “n” vezes a medida de uma grandeza, os n valores não são idênticos, mas estão dispersos numa certa região ou intervalo. Essas flutuações, tradicionalmente, são classificadas como erros aleatórios e resultam na inconsistência do sistema de equações (Gemael, 1994). Dentre os diversos critérios possíveis para o ajustamento das observações, o mais difundido e aceito em Geodésia é aquele que segue o princípio dos mínimos quadrados, ou seja, aquele que minimiza a soma do quadrado dos erros aleatórios, ponderados pelos respectivos pesos das observações (Ghilani & Wolf, 2006).

O método dos mínimos quadrados já é bastante consolidado e utilizado em aplicações geodésicas, com diversas publicações abordando o assunto, como por exemplo: Rainsford (1957), Mikhail (1976), Gemael (1994), Strang e Borre (1997), Koch (1999), Dalmolin (2002), Teunissen (2003) e Ghilani e Wolf (2006).

No cenário nacional, pode-se afirmar que o desenvolvimento matemático do MMQ é amplamente difundido e conhecido pela comunidade geodésica. Porém, neste cenário, ainda é incipiente uma revisão detalhada a respeito da interpretação geométrica do MMQ.

Neste sentido, este artigo pretende contribuir com o entendimento teórico do MMQ, apresentando uma revisão da interpretação geométrica do método, obtida em Teunissen (2003), bem como uma solução alternativa de cálculo, visando comprovar a interpretação geométrica apresentada. É importante ressaltar que se pretendeu uma abordagem mais didática, com um maior detalhamento na revisão apresentada.

O artigo encerra apresentando dois exemplos numéricos, demonstrando a solução alternativa de cálculo obtida por meio da interpretação geométrica, comprovando a mesma. Cabe ressaltar que este artigo não tem a pretensão de propor a utilização desta solução alternativa em substituição da tradicional. Ela apenas é demonstrada e aplicada para comprovar a interpretação geométrica do MMQ, que é o foco deste artigo/revisão.

2. O MÉTODO DOS MÍNIMOS QUADRADOS

Considerando um sistema linear de equações, expresso na forma matricial:

$$\mathbf{y}_{nx1} = \mathbf{A}_{nxu} \mathbf{x}_{ux1} + \mathbf{e}_{nx1} \quad (1)$$

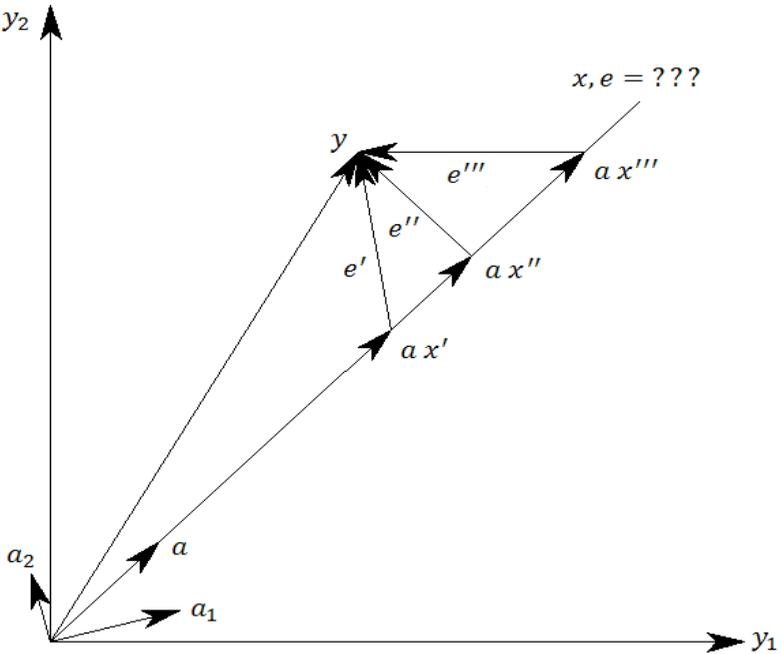
onde \mathbf{y}_{nx1} é o vetor das n observações ou medidas realizadas, \mathbf{A}_{nxu} é a matriz dos coeficientes das equações lineares (matriz design), \mathbf{x}_{ux1} é o vetor dos u parâmetros incógnitos a serem estimados e \mathbf{e}_{nx1} é o vetor dos n erros aleatórios.

Admitindo a inconsistência do sistema, o problema que se deseja solucionar é: Dado o vetor das observações \mathbf{y}_{nx1} e a matriz design \mathbf{A}_{nxu} que relaciona linearmente as observações com os parâmetros, qual a melhor estimativa para o vetor dos parâmetros \mathbf{x}_{ux1} e para o vetor dos erros aleatórios \mathbf{e}_{nx1} ?

Para o caso em que se têm apenas duas observações y_1 e y_2 e uma incógnita x , a matriz design \mathbf{A}_{nxu} se torna um vetor \mathbf{a}_{2x1} e uma visualização do problema é obtida por meio da Figura 1:

Analisando a Figura 1 e a equação (1), nota-se que, para cada escolha distinta do parâmetro x , têm-se um correspondente vetor dos erros aleatórios \mathbf{e}_{2x1} associado. O vetor dos erros aleatórios, por meio de manipulações algébricas em (1), pode ser expresso como:

Figura 1 – Diferentes estimadores para o parâmetro desconhecido x e o vetor e .



$$e_{nx1} = y_{nx1} - A_{nxu} x_{ux1}$$

(2)

No método dos mínimos quadrados, busca-se como estimativa para o vetor dos parâmetros x_{ux1} aquela cuja soma do quadrado dos erros aleatórios apresentar valor mínimo, ou seja:

$$(y_{nx1} - A_{nxu} x_{ux1})^T (y_{nx1} - A_{nxu} x_{ux1}) = \text{Mínimo}$$

(3)

Onde o sobrescrito T denota a operação de transposição, e o produto interno entre o transposto e o vetor denota o quadrado de sua norma, ou seja, a soma do quadrado das componentes deste vetor (Lay, 1997).

Considerando uma matriz peso das observações, simétrica e positiva-definida, denotada por W_{nxn} , o critério expresso em (3) se torna:

$$(y_{1x1} - A_{nxu} x_{ux1})^T W_{nxn} (y_{1x1} - A_{nxu} x_{ux1})_{nx1} = \textit{Mínimo} \quad (4)$$

Onde os elementos da matriz W_{nxn} podem ser escolhidos para ponderar a influência de observações específicas sobre a estimativa do vetor dos parâmetros x_{ux1} . Observações com peso maior terão maior importância no processo de ajustamento (Gemael, 1994).

Por exemplo, para um caso bidimensional ($n=2$ e $u=1$), onde

$$y_{2x1} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{2x1} ;$$

$$a_{2x1} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}_{2x1} ;$$

$$W_{2x2} = \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix}_{2x2}$$

e

$$x_{1x1} = x$$

a equação (4) torna-se:

$$w_{11}(y_1 - a_1x)^2 + 2w_{12}(y_1 - a_1x)(y_2 - a_2x) + w_{22}(y_2 - a_2x)^2 = \textit{Mínimo} \quad (5)$$

Na expressão (5) é utilizada a propriedade de simetria da matriz W_{2x2} , de forma que $w_{12} = w_{21}$ (Lay, 1997).

Voltando ao caso geral da matriz W com dimensão nxn , realizando as operações algébricas em (4) e denotando esta soma ponderada do quadrado dos erros aleatórios por $F(x)$, ou seja, como função do vetor dos parâmetros x_{ux1} estimado, tem-se:

$$F(x) = y^T W y - 2 y^T W A x + A^T W A x^T x \quad (6)$$

Pela teoria do Cálculo vetorial, sabe-se que esta função apresentará um valor $F(x) = \textit{Mínimo}$ quando o vetor dos parâmetros x_{ux1} for um vetor ...

$$\hat{x}_{ux1}$$

tal que (Koch, 1999):

$$\begin{cases} \frac{dF}{dx} \Big|_{x=\hat{x}} = 0 \\ \frac{d^2F}{dx^2} \Big|_{x=\hat{x}} > 0 \end{cases}$$

(7)

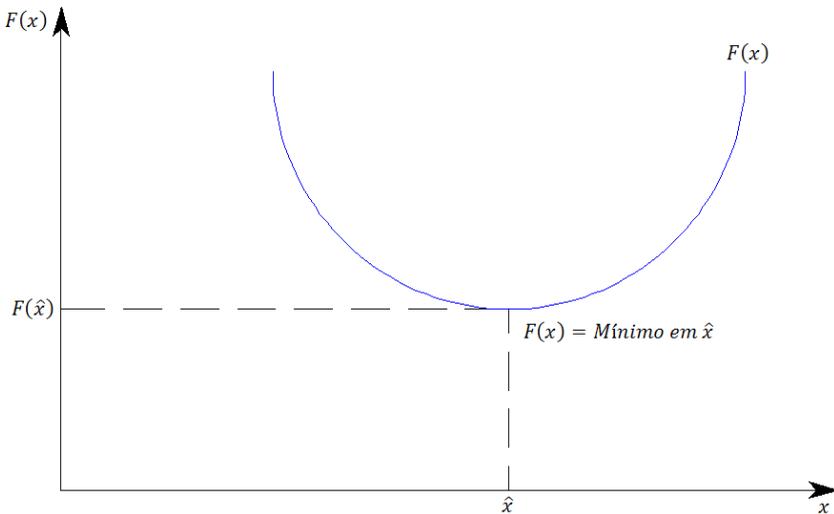
Aplicando as condições de (7) na expressão (6), resulta em:

$$\begin{cases} \frac{dF}{dx} \Big|_{x=\hat{x}} = -2A_{uxm}^T W_{n \times n} y_{n \times 1} + 2A_{uxm}^T W_{n \times n} A_{n \times u} \hat{x}_{u \times 1} = 0 \\ \frac{d^2F}{dx^2} \Big|_{x=\hat{x}} = 2A_{uxm}^T W_{n \times n} A_{n \times u} > 0 \end{cases}$$

(8)

A Figura 2 ilustra esta condição de $F(x) = \text{Mínimo}$ para uma função F de apenas uma variável x ($u=1$).

Figura 2 – Ponto \hat{x} no qual a função $F(x)$ apresenta valor mínimo $F(\hat{x})$.



Como a matriz peso $W_{n \times n}$ é assumida ser positiva-definida e simétrica, a segunda condição de (8), ou seja, a garantia de ser ponto de mínimo, é sempre obedecida (Lay, 1997). Para a primeira condição de (8), realizando as operações algébricas, têm-se (Strang e Borre, 1997):

$$\hat{x}_{ux1} = (A_{uxn}^T W_{n \times n} A_{n \times u})^{-1} A_{uxn}^T W_{n \times n} y_{nx1} \quad (9)$$

Onde \hat{x}_{ux1} é a estimativa por mínimos quadrados do vetor dos parâmetros (é o vetor dos parâmetros ajustados). O vetor das observações ajustadas \hat{y}_{nx1} é obtido aplicando o vetor dos parâmetros ajustados ao modelo linear, resultando em:

$$\hat{y}_{nx1} = A_{n \times u} \hat{x}_{ux1} = [A_{n \times u} (A_{uxn}^T W_{n \times n} A_{n \times u})^{-1} A_{uxn}^T W_{n \times n}] y_{nx1} \quad (10)$$

Finalmente, o vetor dos erros aleatórios ajustados pode ser obtido combinando as expressões (2) com (10), resultando em:

$$\hat{e}_{nx1} = [I_{n \times n} - A_{n \times u} (A_{uxn}^T W_{n \times n} A_{n \times u})^{-1} A_{uxn}^T W_{n \times n}] y_{nx1} \quad (11)$$

Onde \hat{e}_{nx1} é a estimativa por mínimos quadrados do vetor dos erros aleatórios (ou simplesmente vetor dos erros aleatórios) e $I_{n \times n}$ é a matriz identidade $n \times n$. Os sobrescritos “*” em (9), (10) e (11) denotam o fato dos referidos vetores \hat{x}_{ux1} , \hat{y}_{nx1} e \hat{e}_{nx1} se tratarem de estimativas. Nos casos em que a matriz design $A_{n \times u}$ é não linear em relação aos parâmetros x_{ux1} , primeiro deve ser feita uma linearização do modelo via série de Taylor, em torno de um vetor dos parâmetros aproximados (detalhes em Gemael, 1994; e em Dalmolin, 2002).

2.1 Interpretação Geométrica do MMQ

Em Teunissen (2003) é demonstrada uma interpretação geométrica para o método dos mínimos quadrados. Esta interpretação geométrica será apresentada para o caso bidimensional já definido anteriormente ($n=2$ e $u=1$).

Visando auxiliar nesta interpretação, inicialmente, considera-se a seguinte equação geral, envolvendo um vetor genérico z de dimensão n e uma matriz simétrica W de dimensão $n \times n$:

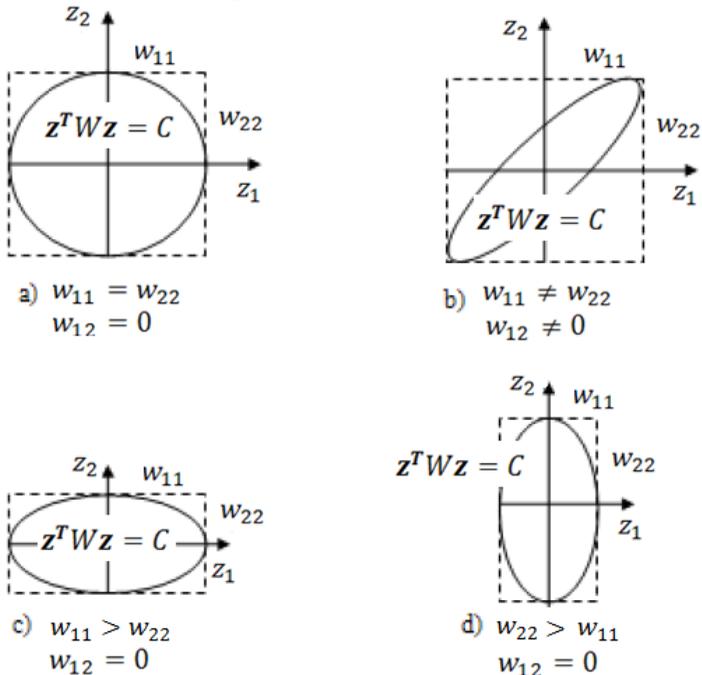
$$z_{1 \times n}^T W_{n \times n} z_{n \times 1} = \text{Constante} = C \quad (12)$$

No caso bidimensional ($n=2$), a expressão (12) se torna:

$$w_{11} z_1^2 + 2 w_{12} (z_1 z_2) + w_{22} z_2^2 = C \quad (13)$$

A equação (13) descreve uma elipse no plano. Se $w_{11} = w_{22}$ e $w_{12} = 0$ a elipse é um círculo. Se $w_{11} \neq w_{22}$ e $w_{12} = 0$ a elipse possui seus semi-eixos paralelos aos eixos z_1 e z_2 . E se $w_{12} \neq 0$ a elipse pode ter uma orientação arbitrária (Figura 3).

Figura 3 – Diferentes orientações possíveis para uma elipse no plano
 Fonte: Adaptada de Mikhail e Gracie (1981).



Definindo uma função $F(z)$ para o vetor $z_{2 \times 1}$ como $F(z) = z^T W z = C$, sabe-se do Cálculo que o gradiente de $F(z)$, calculado em um ponto z_0 , é um vetor que é normal à elipse $F(z) = C$ em z_0 . Em outras palavras, o vetor gradiente em z_0 é perpendicular à linha tangente da função que descreve a elipse neste ponto z_0 . O vetor gradiente da função $F(z)$, calculado em um ponto z_0 , é dado por:

$$\left(\begin{matrix} \frac{\partial F}{\partial z_1} \\ \frac{\partial F}{\partial z_2} \end{matrix} \right) \Big|_{z_0} = \left(\begin{matrix} \frac{\partial (w_{11} z_1^2 + 2 w_{12} (z_1 z_2) + w_{22} z_2^2)}{\partial z_1} \\ \frac{\partial (w_{11} z_1^2 + 2 w_{12} (z_1 z_2) + w_{22} z_2^2)}{\partial z_2} \end{matrix} \right) \Big|_{z_0} \tag{14}$$

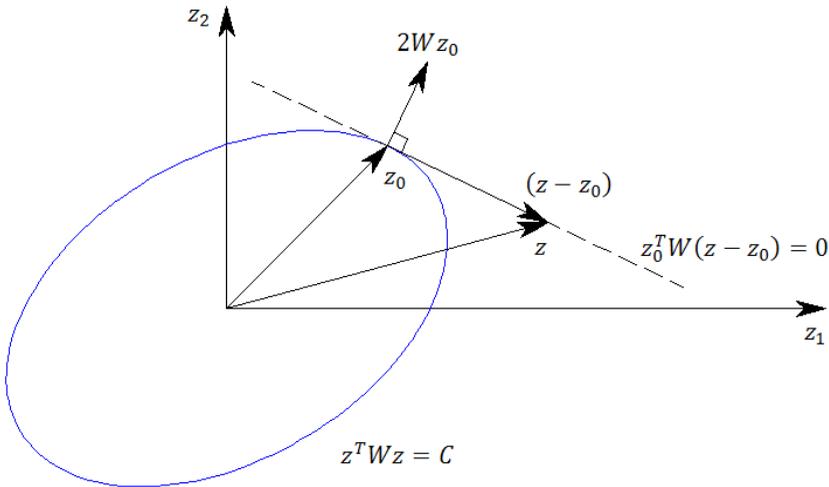
Aplicando as derivadas parciais em (14), resulta em (Teunissen, 2003):

$$\left. \begin{pmatrix} \frac{\partial F}{\partial z_1} \\ \frac{\partial F}{\partial z_2} \end{pmatrix} \right|_{z_0} = \left. \begin{pmatrix} 2w_{11}z_1 + 2w_{12}z_2 \\ 2w_{12}z_1 + 2w_{22}z_2 \end{pmatrix} \right|_{z_0} = 2W_{2 \times 2} z_{0 \ 2 \times 1} \quad (15)$$

Este gradiente é normal à elipse $F(z) = C$ no ponto z_0 (Figura 4).

Figura 4: Elipse $F(z) = C$ e o gradiente desta em um ponto z_0 .

Fonte: Adaptada de Teunissen (2003)



Analisando a Figura 4, nota-se que, para um dado vetor z que se estende da origem até um ponto sobre a linha tangente a elipse em z_0 , têm-se a relação de ortogonalidade:

$$z_{0 \ 1 \times 2}^T W_{2 \times 2} (z_{2 \times 1} - z_{0 \ 1 \times 2}) = 0 \quad (16)$$

Ou seja, o vetor Wz_0 é perpendicular ao vetor $(z - z_0)$, pois o produto interno entre estes dois vetores é nulo (Strang e Borre, 1997). Nota-se também que o escalar 2 que multiplica o vetor Wz_0 não interfere nesta condição de ortogonalidade (não altera a direção e o sentido, apenas o módulo do vetor Wz_0).

Ainda considerando o caso bidimensional ($n=2$ e $u=1$), o estimador por mínimos quadrados ponderados do parâmetro x é dado por um escalar \hat{x} tal que o

vetor dos erros aleatórios $\hat{e} = (y - A\hat{x})$ é paralelo a linha tangente da elipse $z^T W z = a^T W a$ em a , (Figura 5), pois realizando operações algébricas em (9), pode-se escrever:

$$\hat{x} = (a^T W a)^{-1} a^T W y \Leftrightarrow a^T W (y - a\hat{x}) = 0 \tag{17}$$

Ou seja, o vetor dos erros aleatórios \hat{e} deve ser perpendicular ao vetor $W a$ e, portanto, paralelo a linha tangente da elipse $a^T W a = C$ no vetor a .

Nota-se ainda que, quando $W_{2 \times 2} = I_{2 \times 2}$, resulta em $a^T (y - A\hat{x}) = 0$, ou seja, a elipse se torna um círculo ($w_{11} = w_{22} = 1$ e $w_{12} = w_{21} = 0$), e a estimativa por mínimos quadrados não ponderados é obtida a partir da relação de ortogonalidade entre os vetores

$$\hat{y} = a\hat{x}$$

Ver Figura 6.

Figura 5 – Interpretação geométrica da estimação por mínimos quadrados ponderados no plano.

Fonte: Adaptada de Teunissen (2003)

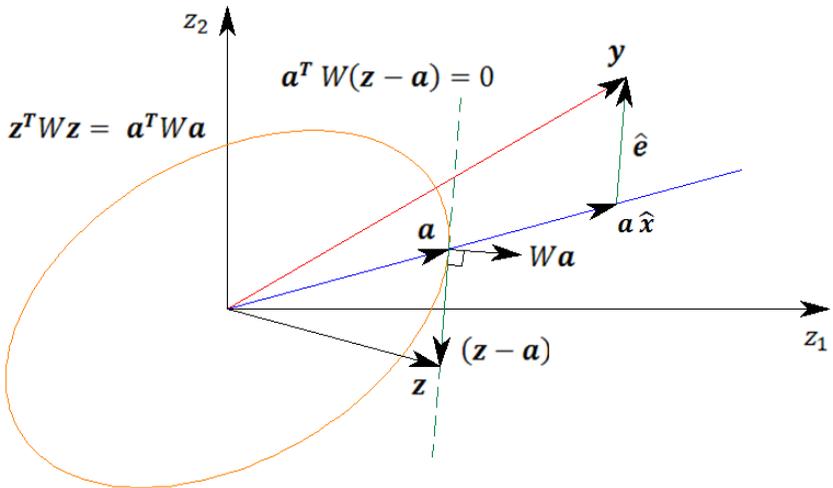
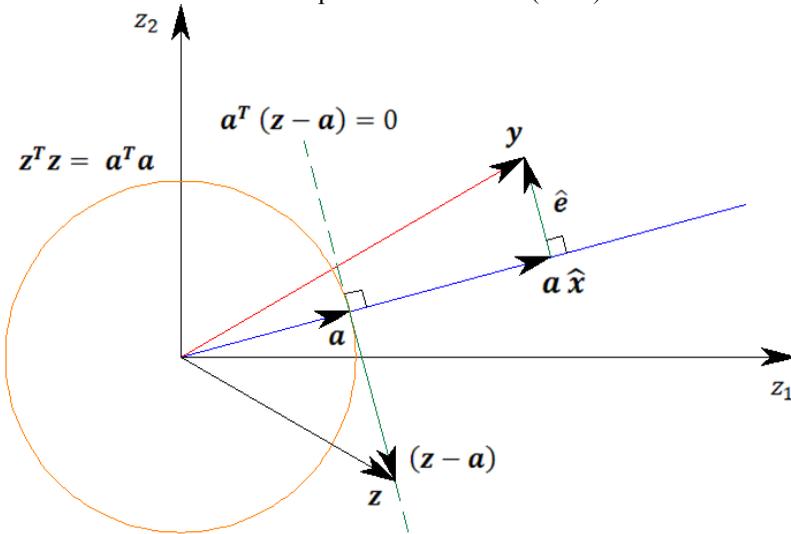


Figura 6 – Interpretação geométrica da estimação por mínimos quadrados não ponderados no plano.

Fonte: Adaptada de Teunissen (2003)



De acordo com a expressão (10), $\hat{y} = [A(A^T W A)^{-1} A^T W] y$, e, segundo a expressão (11), $\hat{e} = [I - A(A^T W A)^{-1} A^T W] y$.

Analisando estas expressões e a Figura 5, nota-se que os vetores \hat{y} e \hat{e} são projeções oblíquas do vetor y , ou seja, as matrizes $A(A^T W A)^{-1} A^T W$ e $[I - A(A^T W A)^{-1} A^T W]$ são projetores oblíquos, denotados por (Teunissen, 2003):

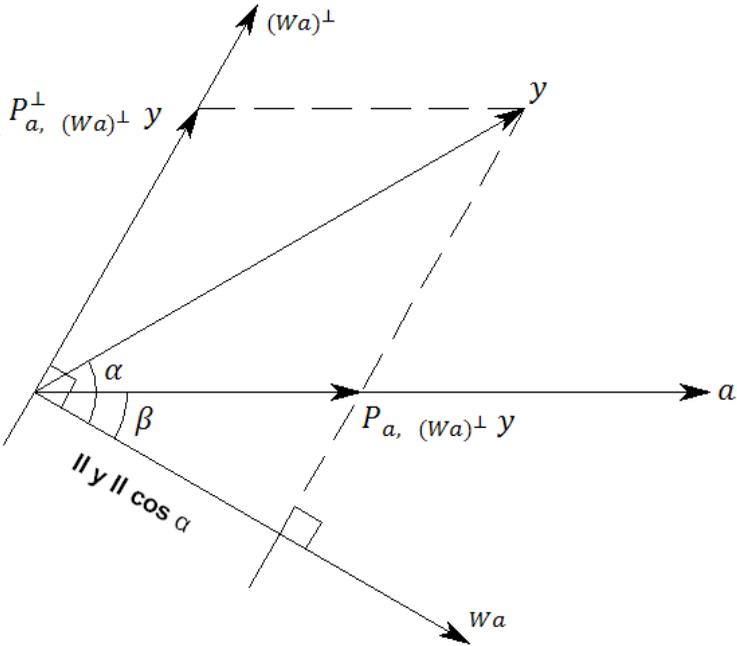
$$\begin{cases} P_{a, (W a)^\perp} = A_{n \times u} (A_{u \times n}^T W_{n \times n} A_{n \times u})^{-1} A_{u \times n}^T W_{n \times n} \\ P_{a, (W a)^\perp}^\perp = I_{n \times n} - A_{n \times u} (A_{u \times n}^T W_{n \times n} A_{n \times u})^{-1} A_{u \times n}^T W_{n \times n} \end{cases} \quad (18)$$

Em outras palavras, o vetor das observações ajustadas \hat{y} e o vetor dos erros aleatórios ajustados \hat{e} são projeções oblíquas do vetor das observações y , sendo estas duas projeções definidas pelos projetores oblíquos $P_{a, (W a)^\perp}$ e $P_{a, (W a)^\perp}^\perp$ em (18), respectivamente.

No caso bidimensional em questão ($n=2$ e $u=1$), o vetor $P_{a, (W a)^\perp}$ projeta sobre uma linha gerada pelo vetor $a_{2 \times 1}$, e ao longo de uma linha gerada pelo vetor $(W a)^\perp$, que é ortogonal ao vetor $W_{2 \times 2} a_{2 \times 1}$ (Figura 7), e o vetor $P_{a, (W a)^\perp}^\perp$ projeta sobre uma linha ortogonal ao vetor $W_{2 \times 2} a_{2 \times 1}$, e paralelamente à direção do vetor $a_{2 \times 1}$ (Figura 7).

Figura 7 – Projetores oblíquos $P_a, (Wa)^\perp$ e $P^\perp_a, (Wa)^\perp$ do vetor das observações y .

Fonte: Adaptada de Teunissen (2003)



Como Wa e $(Wa)^\perp$ são ortogonais entre si, o produto interno destes vetores resulta em $[(Wa)^\perp]^T Wa = 0$.

Com o auxílio de um vetor $b_{2 \times 1}$, definido de forma que seja ortogonal ao vetor $a_{2 \times 1}$, ou seja, $(b_{1 \times 2})^T a_{2 \times 1} = 0$, pode-se escrever $(Wa)^\perp$ como $W^T b$.

Portanto, $(W^T b)^T (Wa) = 0$, pois a inversa W^{-1} de W existe, desde que W seja positiva-definida (Lay, 1997).

Desta forma, com o auxílio do vetor $W^{-1}b$, pode-se representar os dois projetores definidos em (18) alternativamente como (Teunissen, 2003):

$$\begin{cases} P_{a, (Wa)^\perp} = P_{a, W^{-1}b} = I_{2 \times 2} - W_{2 \times 2}^{-1} b_{2 \times 1} (b_{1 \times 2}^T W_{2 \times 2}^{-1} b_{2 \times 1})^{-1} b_{1 \times 2}^T \\ P^\perp_{a, (Wa)^\perp} = P_{W^{-1}b, a} = W_{2 \times 2}^{-1} b_{2 \times 1} (b_{1 \times 2}^T W_{2 \times 2}^{-1} b_{2 \times 1})^{-1} b_{1 \times 2}^T \end{cases} \quad (19)$$

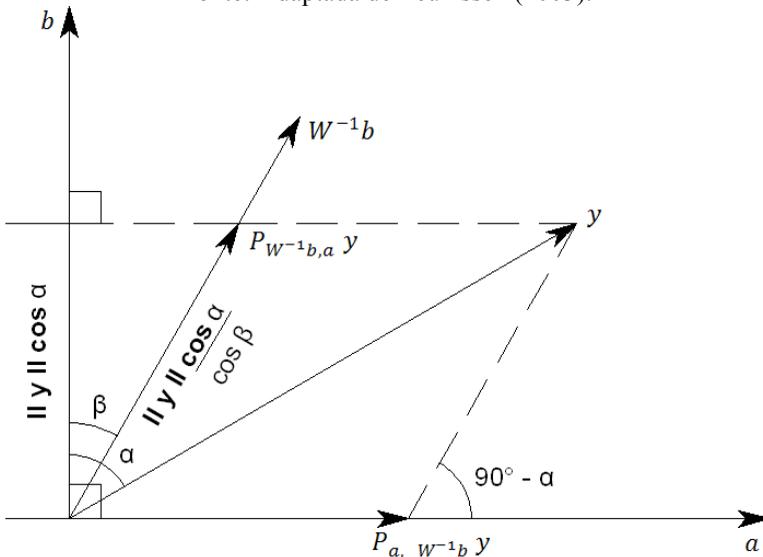
A expressão (19) é melhor compreendida com o auxílio da Figura 8, que se utiliza de relações trigonométricas elementares, e das expressões (18) e (11), podendo-se escrever, para o vetor dos erros aleatórios $\hat{\epsilon}_{n \times 1}$ (Teunissen, 2003):

$$\hat{e} = P_{a, (wa)}^\perp y = P_{W^{-1}b, a} y = \|y\| \frac{\cos \alpha}{\cos \beta} \frac{W^{-1}b}{\|W^{-1}b\|} \quad (20)$$

Onde o escalar $\|y\| \frac{\cos \alpha}{\cos \beta}$ define o módulo do vetor resultante $\hat{e} = P_{W^{-1}b, a} y$, e o vetor normalizado (unitário) $\frac{W^{-1}b}{\|W^{-1}b\|}$ define a direção deste vetor resultante $\hat{e} = P_{W^{-1}b, a} y$, lembrando que o vetor dos erros aleatórios ajustados \hat{e} é uma projeção oblíqua do vetor das observações y e a matriz $P_{W^{-1}b, a}$ é o seu projetor.

Figura 8 – Representação alternativa dos projetores oblíquos do vetor das observações y .

Fonte: Adaptada de Teunissen (2003).



Rearranjando a expressão (20), de forma a agrupar os termos escalares (unidimensionais) à direita, pode-se escrever:

$$P_{W^{-1}b,a} y = W^{-1}{}_{2 \times 2} b_{2 \times 1} \left[\frac{\|y\| \cos \alpha}{\|W^{-1}b\| \cos \beta} \right]_{1 \times 1} \quad (21)$$

Analisando a Figura (8), o escalar $\|y\| \cos \alpha$ é igual ao produto interno entre os vetores $b_{2 \times 1}$ e $y_{2 \times 1}$ (ver Lay, 1997; e também Strang e Borre, 1997), ou seja, $\|y\| \cos \alpha = b^T y$, e de maneira análoga, $\|W^{-1}b\| \cos \beta = (W^{-1}b)^T b$.

Desta forma, pode-se finalmente escrever (Teunissen, 2003):

$$P_{W^{-1}b,a} y = W^{-1}b \frac{b^T y}{(W^{-1}b)^T b} = W^{-1}b (W^{-1}b (b^T W^{-1}b)^{-1} b^T) y \quad (22)$$

Ou seja, a segunda expressão de (19) é demonstrada geometricamente. A primeira expressão de (19), relativa ao projetor oblíquo do vetor das observações y em a (sendo sua projeção resultante o vetor \hat{y}), pode ser demonstrada de maneira similar. Dando continuidade a interpretação geométrica do MMQ, lembrando o critério dos mínimos quadrados ponderados, que busca o vetor dos parâmetros x que minimiza a seguinte expressão:

$$\min_x (y_{n \times 1} - A_{n \times n} x_{n \times 1})_{1 \times n}^T W_{n \times n} (y_{n \times 1} - A_{n \times n} x_{n \times 1})_{n \times 1} \quad (23)$$

Definindo um vetor z por $z = Ax$, a expressão (23) é equivalente à:

$$\min_z (y_{n \times 1} - z_{n \times 1})_{1 \times n}^T W_{n \times n} (y_{n \times 1} - z_{n \times 1})_{n \times 1} \quad (24)$$

Como $e = y - Ax$, logo $z = y - e$. Para o caso bidimensional considerado, o vetor b é ortogonal ao vetor a , ou seja, $b^T a = 0$. Com este pressuposto, resulta em:

$$b^T z = b^T (ax) = 0 \Leftrightarrow b^T (y - e) = 0 \quad (25)$$

A expressão (25) é satisfeita quando $b^T y = b^T e$. Mas se $e = y - Ax$, a expressão (23) é equivalente a:

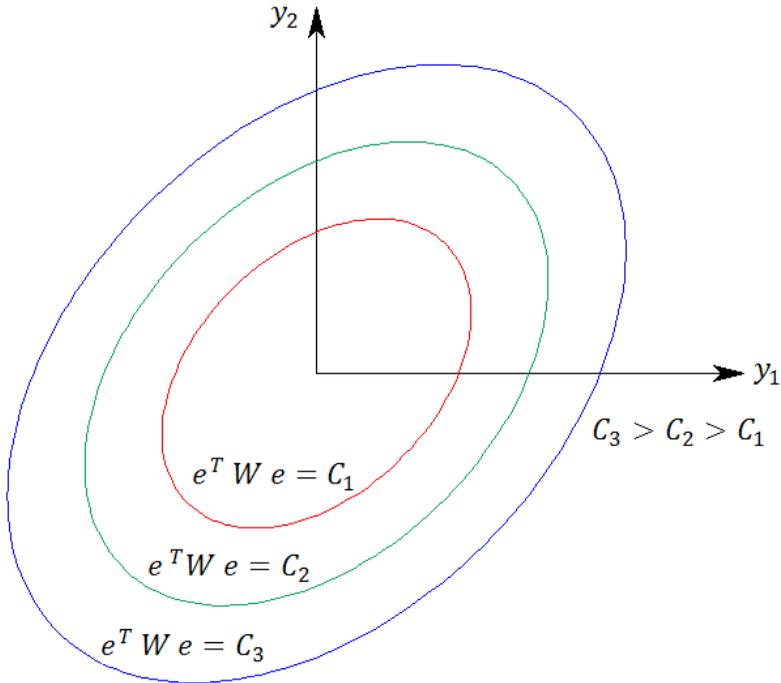
$$\min_e e_{1 \times n}^T W_{n \times n} e_{n \times 1} \quad (26)$$

Ou seja, no caso bidimensional considerado, o método dos mínimos quadrados ponderados consiste em estimar um vetor dos erros aleatórios \hat{e} que minimiza a expressão $e^T W e$ e obedece à relação $b^T y = b^T \hat{e}$.

Fazendo $e^T W e = \text{Constante} = C$, têm-se a equação de uma elipse, conforme já fora visto. Para diferentes valores da constante C têm-se diferentes elipses, e quanto maior o valor da constante C , maior o tamanho da elipse (Figura 9).

Figura 9 – Diferentes elipses associadas a diferentes valores para a constante C .

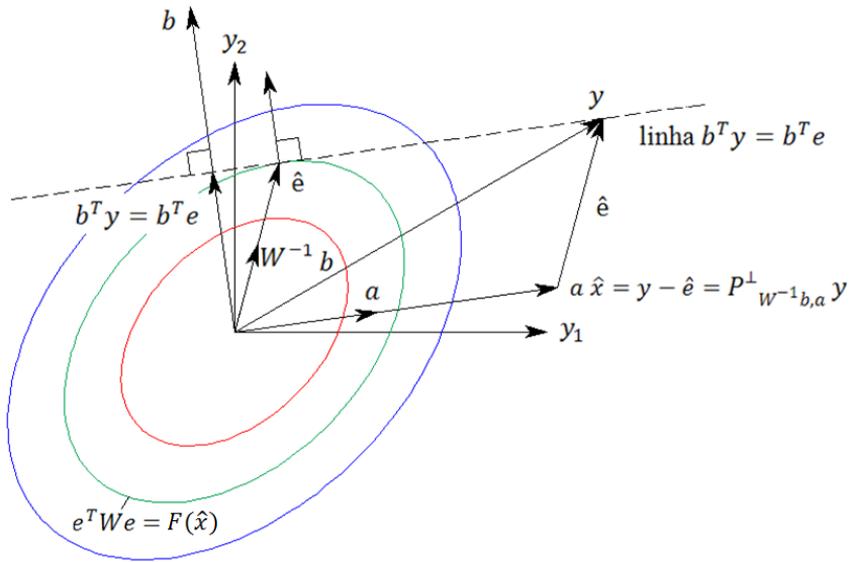
Fonte: Adaptada de Teunissen (2003)



A estimativa \hat{e} do vetor dos erros aleatórios é dada pelo ponto de tangência da linha $b^T y = b^T \hat{e}$ com a elipse $e^T W e = \text{Mínimo} = F(\hat{x})$, (Figura 10):

Figura 10 – Interpretação geométrica para a estimação do vetor \hat{e} por Mínimos Quadrados Ponderados.

Fonte: Adaptada de Teunissen (2003)



Neste ponto de tangência \hat{e} , a normal da elipse é paralela a normal da linha $b^T y = b^T e$. Mas como a normal da linha $b^T y = b^T e$ é dada por b , a normal da elipse em \hat{e} também é dada por b (para maiores detalhes ver produto interno de vetores em Lay, 1997). Portanto, o vetor \hat{e} é paralelo a $W^{-1}b$. Logo, pode-se expressar o vetor \hat{e} por um produto entre $W^{-1}b$ e um escalar α (Teunissen, 2003):

$$\hat{e} = W^{-1}b \alpha \tag{27}$$

O escalar desconhecido α pode ser determinado pelo fato de que o vetor \hat{e} tem que se estender até a linha $b^T y = b^T e$. Desta forma, pré-multiplicando (27) por b^T e realizando operações algébricas, resulta em:

$$\alpha = (b^T W^{-1}b)^{-1} b^T \hat{e} = (b^T W^{-1}b)^{-1} b^T y \tag{28}$$

Substituindo (28) em (27), os estimadores \hat{e} e \hat{y} por mínimos quadrados ponderados tornam-se (Teunissen, 2003):

$$\begin{cases} \hat{\boldsymbol{\varepsilon}} = P_{W^{-1}\mathbf{b}, \mathbf{a}} \mathbf{y} = [W^{-1} \mathbf{b} (\mathbf{b}^T W^{-1} \mathbf{b})^{-1} \mathbf{b}^T] \mathbf{y} \\ \hat{\mathbf{y}} = \mathbf{y} - \hat{\boldsymbol{\varepsilon}} = P_{W^{-1}\mathbf{b}, \mathbf{a}}^\perp \mathbf{y} = [I - W^{-1} \mathbf{b} (\mathbf{b}^T W^{-1} \mathbf{b})^{-1} \mathbf{b}^T] \mathbf{y} \end{cases} \quad (29)$$

Quando $W_{n \times n} = I_{n \times n}$, têm-se:

$$\begin{cases} \hat{\boldsymbol{\varepsilon}} = [\mathbf{b} (\mathbf{b}^T \mathbf{b})^{-1} \mathbf{b}^T] \mathbf{y} \\ \hat{\mathbf{y}} = \mathbf{y} - \hat{\boldsymbol{\varepsilon}} = [I - \mathbf{b} (\mathbf{b}^T \mathbf{b})^{-1} \mathbf{b}^T] \mathbf{y} \end{cases} \quad (30)$$

Analisando as expressões (29) e (30), nota-se que com as relações geométricas estabelecidas, para os casos em que o vetor dos parâmetros $\mathbf{x}_{n \times 1}$ se torna um escalar $x_{1 \times 1}$ e conseqüentemente, a matriz design $A_{n \times u}$ se torna um vetor $\mathbf{a}_{n \times 1}$, é possível estimar o vetor das observações ajustadas $\hat{\mathbf{y}}_{n \times 1}$ e o vetor dos erros aleatórios $\hat{\boldsymbol{\varepsilon}}_{n \times 1}$ sem o conhecimento prévio do parâmetro ajustado \hat{x} , definindo um vetor $\mathbf{b}_{n \times 1}$ que seja ortogonal ao vetor $\mathbf{a}_{n \times 1}$ e lembrando que $\hat{\mathbf{y}}_{n \times 1} = \mathbf{a}_{n \times 1} \hat{x}$.

Os vetores $\hat{\mathbf{y}}_{n \times 1}$ e $\hat{\boldsymbol{\varepsilon}}_{n \times 1}$ também poderiam ser obtidos diretamente em função de $\mathbf{y}_{n \times 1}$, $W_{n \times n}$ e $A_{n \times u}$, conforme mostram as expressões (10) e (11).

Para problemas envolvendo um maior número de parâmetros, esta solução alternativa se torna mais complexa, pois se deve definir uma matriz $\mathbf{B}_{n \times u}$ cujos vetores (colunas) sejam ortogonais a todos os vetores (colunas) da matriz design $A_{n \times u}$, sendo as matrizes \mathbf{A} e \mathbf{B} dois subespaços ortogonais entre si de \mathbf{u} vetores cada e dimensão n (Kock, 1999).

3. EXEMPLOS NUMÉRICOS

Para uma melhor compreensão e comprovação desta interpretação geométrica do MMQ, dois exemplos numéricos são apresentados a seguir.

Considerando que para determinar a distância entre dois pontos, duas medidas foram realizadas, resultando no vetor das observações,

$$\mathbf{y}_{2 \times 1} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{2 \times 1} = \begin{bmatrix} 99 \\ 101 \end{bmatrix}_{2 \times 1},$$

expresso em metros, e nas respectivas variâncias destas observações, dadas por $\sigma_{y_1}^2 = 5 \text{ m}^2$ e $\sigma_{y_2}^2 = 2.5 \text{ m}^2$.

Admitindo que as observações são independentes entre si, as covariâncias são nulas: $\sigma_{y_1 y_2} = \sigma_{y_2 y_1} = 0$. Desta forma, definindo a matriz peso das observações como sendo igual ao inverso de sua matriz de variâncias e covariâncias, resulta em:

$$W_{2 \times 2} = \begin{bmatrix} \sigma_{y_1}^2 & \sigma_{y_1 y_2} \\ \sigma_{y_2 y_1} & \sigma_{y_2}^2 \end{bmatrix}_{2 \times 2}^{-1} = \begin{bmatrix} \frac{1}{5} & 0 \\ 0 & \frac{1}{2.5} \end{bmatrix}_{2 \times 2}$$

O parâmetro desconhecido x que se deseja estimar é a própria distância entre os dois pontos, desta forma, o modelo matemático que relaciona as observações com o parâmetro x é dado por:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{2 \times 1} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}_{2 \times 1} x_{1 \times 1} + \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}_{2 \times 1} \Rightarrow \begin{bmatrix} 99 \\ 101 \end{bmatrix}_{2 \times 1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{2 \times 1} x + \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}_{2 \times 1}$$

Onde $\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}_{2 \times 1}$ é o vetor dos erros aleatórios, admitindo a inconsistência do sistema. Pelo método dos mínimos quadrados, a solução do problema é dada por:

$$\hat{x}_{1 \times 1} = (a_{1 \times 2}^T W_{2 \times 2} a_{2 \times 1})^{-1} a_{1 \times 2}^T W_{2 \times 2} y_{2 \times 1} = 100.333 \text{ m}$$

Para o vetor das observações ajustadas, tem-se:

$$\hat{y}_{2 \times 1} = a_{2 \times 1} \hat{x}_{1 \times 1} = \begin{bmatrix} 100.333 \\ 100.333 \end{bmatrix}_{2 \times 1},$$

e o vetor dos erros aleatórios resulta em

$$\hat{\theta}_{2 \times 1} = y_{2 \times 1} - \hat{y}_{2 \times 1} = \begin{bmatrix} -1.333 \\ 0.667 \end{bmatrix}_{2 \times 1}.$$

Analisando o vetor dos erros aleatórios e a estimativa por mínimos quadrados da distância desconhecida, nota-se que de fato a observação y_2 , de menor variância, e, portanto, maior peso, teve maior influência no processo de ajustamento.

Agora, o exemplo será resolvido pela solução de cálculo alternativa do MMQ, expressa em (29). Definindo um vetor $b_{2 \times 1}$ ortogonal ao vetor $a_{2 \times 1}$, como por exemplo,

$$b_{2 \times 1} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}_{2 \times 1},$$

tem-se:

$$a_{1 \times 2}^T b_{2 \times 1} = b_{1 \times 2}^T a_{2 \times 1} = 0$$

Utilizando as propriedades geométricas apresentadas, a estimativa por mínimos quadrados do vetor das observações ajustadas e do vetor dos erros aleatórios são dadas por:

$$\hat{y}_{2 \times 1} = [I - W^{-1} b(b^T W^{-1} b)^{-1} b^T]_{2 \times 2} y_{2 \times 1} = \begin{bmatrix} 100.333 \\ 100.333 \end{bmatrix}_{2 \times 1}$$

$$\hat{e}_{2 \times 1} = [W^{-1} b(b^T W^{-1} b)^{-1} b^T]_{2 \times 2} y_{2 \times 1} = \begin{bmatrix} -1.333 \\ 0.667 \end{bmatrix}_{2 \times 1}$$

Ou seja, de fato é possível estimar o vetor das observações ajustadas e o vetor dos erros aleatórios sem o conhecimento prévio do parâmetro ajustado \hat{x} .

Além disso, as demais relações geométricas demonstradas também são satisfeitas, como por exemplo:

$$b_{1 \times 2}^T y_{1 \times 1} = b_{1 \times 2}^T \hat{e}_{2 \times 1} = -2,$$

$$\alpha = (b^T W^{-1} b)^{-1}_{1 \times 1} b_{1 \times 2}^T y_{1 \times 1} = -0.02667$$

e

$$\hat{e}_{2 \times 1} = W_{2 \times 2}^{-1} b_{2 \times 1} \alpha_{1 \times 1} = \begin{bmatrix} -1.333 \\ 0.667 \end{bmatrix}_{2 \times 1}.$$

Portanto, os vetores $W^{-1}_{2 \times 2} b_{2 \times 1}$ e $\hat{e}_{2 \times 1}$ são de fato paralelos, diferindo apenas por um escalar $\alpha = -0.02667$.

Definindo outros vetores perpendiculares ao vetor

$$a_{2 \times 1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{2 \times 1},$$

como por exemplo,

$$b_{2 \times 1} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}_{2 \times 1}, \quad b_{2 \times 1} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}_{2 \times 1}, \quad b_{2 \times 1} = \begin{bmatrix} -\pi \\ \pi \end{bmatrix}_{2 \times 1} \text{ e etc.,}$$

as relações geométricas também são satisfeitas. A soma ponderada do quadrado dos erros, que independe da escolha do vetor b , resulta em $\hat{e}^T W \hat{e} = 0.533$.

O segundo exemplo numérico consiste em um problema de $n=4$ observações e $u=2$ incógnitas. Considerando que foram observadas as coordenadas planas de dois pontos, em dois sistemas de referência distintos (X, Y) e (x, y) , deseja-se determinar os parâmetros de transformação a e b do sistema (x, y) para o sistema (X, Y) , por meio de uma transformação de similaridade sem translações (detalhes em Lugnani, 1987).

Para dois pontos i e j , tem-se as coordenadas destes nos dois sistemas de referência, ou seja, (X_i, Y_i, x_i, y_i) e (X_j, Y_j, x_j, y_j) , e o modelo matemático de ajustamento para determinar os parâmetros incógnitos a e b é dado por:

$$y_{4 \times 1} = A_{4 \times 2} x_{2 \times 1} + e_{4 \times 1} = \begin{bmatrix} X_i \\ Y_i \\ X_j \\ Y_j \end{bmatrix} = \begin{bmatrix} x_i & y_i \\ y_i & -x_i \\ x_j & y_j \\ y_j & -x_j \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

Considerando que as observações são independentes entre si e possuem a mesma precisão, pode-se considerar a matriz peso das observações como sendo igual à matriz identidade, ou seja, $W_{4 \times 4} = I_{4 \times 4}$. Atribuindo valores numéricos para as coordenadas dos pontos nos dois sistemas de referência, tem-se:

$$y_{4 \times 1} = \begin{bmatrix} X_i \\ Y_i \\ X_j \\ Y_j \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 10 \\ 10 \end{bmatrix}$$

e

$$A_{4 \times 2} = \begin{bmatrix} x_i & y_i \\ y_i & -x_i \\ x_j & y_j \\ y_j & -x_j \end{bmatrix} = \begin{bmatrix} 14 & 15 \\ 15 & -14 \\ 142 & 141 \\ 141 & -142 \end{bmatrix}$$

Pelo método dos mínimos quadrados, a solução do problema é dada por:

$$\hat{x}_{2 \times 1} = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (A_{2 \times 4}^T W_{4 \times 4} A_{4 \times 2})^{-1} A_{2 \times 4}^T W_{4 \times 4} y_{4 \times 1} = \begin{bmatrix} 0.0706519 \\ -0.0002224 \end{bmatrix}$$

Para o vetor das observações ajustadas, tem-se:

$$\hat{y}_{2 \times 1} = a_{2 \times 1} \hat{x}_{1 \times 1} = \begin{bmatrix} 0.9858 \\ 1.0629 \\ 10.0012 \\ 9.9935 \end{bmatrix}_{4 \times 1}$$

E o vetor dos erros aleatórios resulta em:

$$\hat{e}_{2 \times 1} = y_{2 \times 1} - \hat{y}_{2 \times 1} = \begin{bmatrix} 0.0142 \\ -0.0629 \\ -0.0012 \\ 0.0065 \end{bmatrix}_{4 \times 1}$$

Novamente, a solução alternativa do MMQ, desta vez expressa em (30), quando $W=I$, será utilizada para a resolução do exercício.

Definindo uma matriz $B_{4 \times 2}$ cujos vetores (colunas) sejam ortogonais a todos os vetores (colunas) da matriz $A_{4 \times 2}$, como por exemplo,

$$B_{4 \times 2} = \begin{bmatrix} 142 & -141 \\ -141 & -142 \\ -14 & 15 \\ 15 & 14 \end{bmatrix},$$

tem-se:

$$A_{2 \times 4}^T B_{4 \times 2} = B_{2 \times 4}^T A_{4 \times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Utilizando as propriedades geométricas apresentadas, a estimativa por mínimos quadrados do vetor das observações ajustadas e do vetor dos erros aleatórios são dadas por:

$$\hat{y}_{4 \times 1} = [I - b(b^T b)^{-1} b^T]_{4 \times 4} y_{4 \times 1} = \begin{bmatrix} 0.9858 \\ 1.0629 \\ 10.0012 \\ 9.9935 \end{bmatrix}_{4 \times 1}$$

$$\hat{e}_{4 \times 1} = [b(b^T b)^{-1} b^T]_{4 \times 4} y_{4 \times 1} = \begin{bmatrix} 0.0142 \\ -0.0629 \\ -0.0012 \\ 0.0065 \end{bmatrix}_{4 \times 1}$$

Ou seja, mais uma vez, foi possível estimar o vetor das observações ajustadas e o vetor dos erros aleatórios sem o conhecimento prévio do vetor dos parâmetros ajustados $\hat{x}_{2 \times 1}$. Além disso, novamente as demais relações geométricas demonstradas são satisfeitas, como por exemplo:

$$B_{2 \times 4}^T y_{4 \times 1} = B_{2 \times 4}^T \hat{e}_{4 \times 1} = \begin{bmatrix} 11 \\ 7 \end{bmatrix}_{2 \times 1}$$

Poderia ter-se escolhido outro subconjunto de vetores perpendiculares aos vetores que formam a matriz design $A_{4 \times 2}$, que as relações geométricas também seriam satisfeitas e os valores numéricos de \hat{y} , \hat{e} e \hat{x} seriam os mesmos.

Na prática, a maior dificuldade na aplicação desta solução alternativa é encontrar um subconjunto de vetores ortogonais a todos os vetores que formam as colunas da matriz A , sendo, usualmente, mais fácil empregar as expressões (9), (10) e (11) para a resolução dos problemas. Os dois exemplos apresentados tiveram por objetivo apenas demonstrar numericamente a interpretação geométrica do MMQ.

4. CONSIDERAÇÕES FINAIS

Neste artigo foi apresentada uma revisão teórica sobre o método dos mínimos quadrados, com auxílio de uma interpretação geométrica, obtida em Teunissen (2003).

Quando é utilizada uma matriz peso para as observações, o estimador por mínimos quadrados dos vetores dos erros aleatórios e das observações ajustadas são projeções oblíquas do vetor das observações. Já quando a matriz peso é tomada como sendo igual à matriz identidade, o estimador por mínimos quadrados dos vetores dos erros aleatórios e das observações ajustadas são projeções ortogonais do vetor das observações.

Além disso, quando o vetor dos parâmetros se tornar um escalar, e a matriz design A se tornar um vetor a , é possível estimar as observações ajustadas \hat{y} e o vetor dos erros aleatórios \hat{e} sem o conhecimento prévio do parâmetro ajustado \hat{x} , definindo um vetor b que seja ortogonal ao vetor a e usando as relações geométricas que foram apresentadas e discutidas.

Para problemas envolvendo um maior número de parâmetros, esta solução alternativa de cálculo se torna mais complexa, pois se deve definir uma matriz B cujos vetores (colunas) sejam ortogonais a todos os vetores (colunas) da matriz design A .

Visando uma melhor compreensão e comprovação da interpretação geométrica, dois exemplos numéricos foram simulados, para problemas envolvendo duas observações e um parâmetro e também quatro observações e dois parâmetros.

AGRADECIMENTOS

Os autores agradecem à CAPES pelo fornecimento da bolsa de Mestrado do primeiro autor, e ao CNPq pela Bolsa de Produtividade em Pesquisa (Proc.n. 307472/2009-4) do segundo autor.

REFERÊNCIAS BIBLIOGRÁFICAS

- DALMOLIN, Q. *Ajustamento por mínimos quadrados*. Edição Revisada. Curitiba: Curso de Pós-Graduação em Ciências Geodésicas - UFPR, 2002.
- GEMAEL, C. *Introdução ao ajustamento de observações: aplicações geodésicas*. 1.ed. Curitiba: Editora da UFPR, 1994. 319 p.
- GHILANI, C. D.; WOLF, P. R. *Adjustment Computations: Spatial Data Analysis*. 4º Ed. Editora: John Wiley & Sons, 2006.

- KOCH, K. R. *Parameter Estimation and Hypothesis Testing in Linear Models*. Springer, Berlin, New York, 1999.
- LAY, D. C.; *Linear Algebra And Its Applications*. Addison-Wesley, 2º Ed. 1997.
- LUGNANI, J. B. *Introdução à fototriangulação*. Curitiba: Imprensa Universitária - UFPR. 134 p. 1987.
- MIKHAIL, E. M. *Observations and Least Squares*. University Press of America, New York, 1976.
- MIKHAIL, E. M.; GRACIE, G. *Analysis and Adjustment of Survey Measurements*. New York, Van Nostrand Reinhold Company, 1981.
- RAINSFORD, H. F. *Survey Adjustment And Least Squares*. London: Constable and Co., 1957.
- STRANG, G.; BORRE, K. *Linear algebra, Geodesy and GPS*. Wellesley: Wellesley - Cambridge Press, 1997.
- TEUNISSEN, P.J.G.. *Adjustment Theory: an introduction*. Editora: Delft University Press. Delft, Holanda, 2003.193p.

(Recebido em março de 2011. Aceito em maio de 2011.)