



## UMA REFLEXÃO SOBRE AS REGIÕES DE ERRO

João CASACA

*Núcleo de Geodesia Aplicada do Laboratório Nacional de Engenharia Civil*

**Resumo:** A comunicação apresenta algumas considerações sobre as regiões de erro (hiper-elipsóide de erro e elipses de erro) usadas no planeamento das redes de triangulação que servem para a monitorização de deslocamentos em grandes barragens. Vinte e quatro observações de uma destas redes de triangulação são usadas como um caso de estudo. É introduzido um factor de variância para quantificar a qualidade das observações.

**Palavras-chave:** distribuição gama / elipse de erro / factor de variância / hiper-elipsóide de erro

### 1. INTRODUÇÃO

As redes de triangulação usadas na monitorização de deslocamentos horizontais de grandes barragens de betão são observadas periodicamente, durante a vida útil das obras, em geral ao longo de dezenas de anos, e permitem medir os deslocamentos de pontos notáveis, com incertezas de medição que se pretendem submilimétricas. Dado o tempo de vida útil das redes e a incerteza de medição que se pretende, as redes devem ser cuidadosamente planeadas (Casaca, 2001).

Desde a década de 1980, o LNEC tem usado as regiões de erro das redes (o hiper-elipsóide e as elipses de erro) como critério de aperfeiçoamento (para não usar o termo optimização) da configuração das redes. As regiões de erro têm como premissas (hipótese nula) modelos estocásticos dos erros de observação que, na prática são muitas vezes ultrapassados (hipótese alternativa) por erros sistemáticos instrumentais (calibragem do equipamento) e ambientais, devidos especialmente ao efeito nefasto da refacção atmosférica nas observações angulares e lineares (Casaca, 2004). As regiões de erro, que caracterizam o desempenho das redes sob a hipótese nula, devem pois ser confrontadas com os resultados das observações e, se necessário, redimensionadas.

Esta comunicação apresenta algumas reflexões sobre o desempenho das regiões de erro estabelecidas a priori, ilustradas com um caso de estudo: vinte e quatro observações de uma rede de triangulação usada na monitorização de deslocamentos de uma grande barragem. A qualidade das observações é quantificada por um factor de variância, cuja distribuição é modelada com a distribuição gama, cujo valor médio e quantil de probabilidade 0.95 são propostos como parâmetros, respectivamente, de regiões de erro médias e regiões de erro máximo.

### 2. AS REGIÕES DE ERRO

O modelo usado no LNEC, no ajustamento das variações das grandezas observáveis das redes de triangulação (ângulos, distâncias, bases GNSS, etc.) tem uma componente funcional, que consiste numa relação linear:

$$A\theta = \mu \quad (01)$$

onde  $A(m,n)$  é a matriz da configuração de primeira ordem da rede, que resulta de uma Taylorização de primeira ordem das relações entre as grandezas observáveis (ângulos e distâncias) e o vector (desconhecido) dos deslo-



camentos dos vértices da rede  $\theta(n,1)$ , e onde  $\mu(m,1)$  é o vector das variações (exactas) das grandezas observáveis entre duas épocas (Casaca, 2001).

A componente estocástica do modelo consiste em tomar o vector  $Y(m,1)$  das variações das grandezas observáveis como uma realização de um vector aleatório normal com vector médio  $\mu = A\theta$  desconhecido e com matriz de variância  $\Sigma(m,m)$  parcialmente desconhecida, simbolicamente:

$$Y \in N(\mu = A\theta, \Sigma = \omega Q) \quad (02)$$

onde  $\omega$  é um escalar positivo desconhecido (factor de variância) e  $Q(m,m)$  é a matriz da configuração de segunda ordem da rede, supostamente conhecida. O factor de variância é um indicador da qualidade (dispersão) das observações: i) Igual ou melhor ao esperado, se ( $\omega \leq 1$ ); ii) Pior do que o esperado, se ( $\omega > 1$ ).

O melhor (de variância mínima) estimador linear cêntrico (BLUE – *best linear unbiased estimator*) do vector dos deslocamentos  $\theta$ , dado por (i):

$$i) \bar{\theta} = (A^T Q^{-1} A)^{-1} A^T Q^{-1} Y \quad ii) \varepsilon = \bar{\theta} - \theta \quad (03)$$

é independente do factor de variância. Um estimador cêntrico do factor de variância  $\omega$ , desconhecido, é:

$$w = \frac{1}{m-n} (Y - A\bar{\theta})^T Q^{-1} (Y - A\bar{\theta}) \quad (04)$$

A estratégia utilizada no LNEC no planeamento das redes (Casaca, 2001) baseia-se na adopção de uma função densidade de probabilidade (FDP) *a priori*, no sentido Bayesiano do termo, para modelar o vector ( $\varepsilon$ ) dos erros que afectam o vector dos deslocamentos estimados (03.ii). A FDP anterior dos erros de observação é uma FDP multinormal cêntrica (vector médio nulo), cuja matriz de variância é dada por:

$$\Sigma_\varepsilon = (A^T \Sigma^{-1} A)^{-1} = \omega \underbrace{(A^T Q^{-1} A)^{-1}}_C = \omega C^{-1} \quad (05)$$

onde o factor de variância ( $\omega$ ) é fixado de acordo com informação disponível *a priori*. Uma consequência deste modelo é que a variável aleatória:

$$\omega^{-1} s = \omega^{-1} \underbrace{(\varepsilon^T C \varepsilon)}_s = \omega^{-1} \varepsilon^T \underbrace{(A^T Q^{-1} A)}_C \varepsilon = \varepsilon^T \Sigma_\varepsilon^{-1} \varepsilon \quad (06)$$

se distribui como uma variável aleatória qui-quadrado central com  $n$  graus de liberdade. Esta propriedade pode ser aproveitada para estabelecer a inequação (em ordem ao vector  $X$ ) que define o hiper-elipsóide de erro *a priori* da rede, para o nível de probabilidade  $(1 - \alpha)$ :

$$X^T C X \leq \omega q(n, 1 - \alpha) \quad (07)$$

onde  $q(n; 1 - \alpha)$  é o quantil de probabilidade  $(1 - \alpha)$  de uma variável aleatória com uma distribuição qui-quadrado central com  $n$  graus de liberdade. A forma quadrática aleatória  $s = \varepsilon^T C \varepsilon$  (06), que não é observável, é o quadrado da norma-C (norma induzida pela matriz simétrica definida positiva  $C$ ) do vector dos erros dos deslocamentos.

O raciocínio anterior é válido para qualquer subvector do vector erro (Morrison, 1990), pelo que as elipses de erro dos vértices da rede são obtidas de modo semelhante, considerando apenas os pares de componentes do vector erro e os blocos diagonais de ordem dois da matriz de variância (05) correspondentes aos vértices.



### 3. CASO DE ESTUDO

O caso de estudo é uma pequena rede de triangulação homogénea (somente com observações azimutais) destinada à monitorização dos deslocamentos horizontais de pontos objecto materializados no paramento de jusante de uma grande barragem de betão, que tem vindo a ser observada, pelo LNEC, desde 1981. As observações angulares têm sido realizadas, pelo método dos giros de horizonte, com teodolitos electrónicos cuja incerteza de medição é, de acordo com o fabricante, caracterizada por um desvio padrão  $\sigma = 1.5$  dmgon. Este desvio padrão foi usado na construção da matriz de variância ( $\Sigma$ ) das observações angulares (02).

**Quadro 1** – Estimativas ( $w$ ) do factor de variância ( $\omega$ ), resultantes dos 24 ajustamentos das variações angulares observadas

Ano	w	Ano	w	Ano	w	Ano	w
1982	0.26	1988	0.41	1994	2.43	2004	0.47
1983	0.17	1989	0.68	1995	0.46	2005	1.08
1984	0.67	1990	1.27	1996	0.67	2006	0.25
1985	1.00	1991	0.70	1998	0.63	2007	0.44
1986	0.84	1992	1.37	2002	0.44	2008	0.53
1987	0.42	1993	0.25	2003	0.50	2009	1.03

O ajustamento das variações angulares verificadas entre cada observação e a observação inicial (1981), pelo método da variação de coordenadas (Casaca, 2001), fornece directamente os deslocamentos dos pontos objecto e uma estimativa do factor de variância anterior ( $\omega = 1$ ). As estimativas ( $w$ ), apresentadas no Quadro 1, constituem uma amostra, de tamanho 24, cuja média empírica é 0.71 e cuja variância empírica é 0.24.

**Quadro 2** – Estimativas do quadrado da norma-C do vector erro ( $s$ ), inferidas dos factores de variância ( $w$ ), resultantes dos 24 ajustamentos das variações angulares observadas

Ano	s	Ano	s	Ano	s	Ano	s
1982	3.63*	1988	5.75*	1994	34.09**	2004	6.63*
1983	2.35*	1989	9.53	1995	6.45*	2005	15.06
1984	9.31	1990	17.77	1996	9.31	2006	3.50*
1985	13.98	1991	9.75	1998	8.89	2007	6.10*
1986	11.71	1992	19.13	2002	6.19*	2008	7.38*
1987	5.84*	1993	3.43*	2003	7.00*	2009	14.38

Tendo em atenção as regras para a transformação de variáveis aleatórias e a expressão analítica da FDP de uma variável aleatória qui-quadrado central, com  $n$  graus de liberdade, facilmente se verifica que a FDP da variável  $s$  (quadrado da norma-C do vector erro), condicionada pelo factor de variância  $\omega$ , é dada por:

$$f(s | \omega) = \frac{s^{nv-1}}{2^{nv} \omega^{nv} \Gamma(nv)} \exp\left(-\frac{s}{2\omega}\right) \quad (08)$$

onde  $\Gamma$  é a função gama e onde  $nv (= n/2)$  é o número de vértices da rede. O valor de  $s$  que maximiza a FDP condicionada (08) é:

$$s_0 = 2(nv - 1) \omega \quad (09)$$



No Quadro 2, são apresentadas as estimativas ( $s$ ) do quadrado da norma-C do vector erro, decorrentes da relação (09) aplicada aos factores de variância estimados nas campanhas de medição realizadas. Tendo em atenção que os quantis de probabilidade 0.05 e 0.95 de um qui-quadrado central com ( $n = 16$ ) graus de liberdade são, respectivamente 8.0 e 26.3, constata-se, pela análise do Quadro 2, que: i) As doze (50%) observações assinaladas com um asterisco são de qualidade “excessivamente” boa; ii) A observação assinalada com dois asteriscos é de má qualidade. O facto de ocorrerem 50% de observações com qualidade “excessivamente” boa pode ter várias explicações, como: i) A informação usada na construção da matriz de variância das observações ter sido demasiado conservadora; ii) A rede de triangulação ser pouco redundante; etc.

#### 4. A DISTRIBUIÇÃO DO FACTOR DE VARIÂNCIA

Os valores apresentados no Quadro 1 suscitam, pela sua variabilidade, diversas questões, tais como: i) Será que as estimativas ( $w$ ) são compatíveis com o factor de variância adoptado ( $\omega = 1$ )? Não deverá o parâmetro de variância ( $\omega$ ) ser encarado como uma ocorrência de uma variável aleatória, que assume um valor diferente em cada campanha de observação? Após vinte e quatro observações, continuarão as regiões de erros anteriores iniciais (hiper-elipsóide e elipses de de erro) a ser os melhores indicadores sobre o desempenho da rede? Procuraremos responder seguidamente a estas questões.

Uma distribuição adequada à modelação estocástica do factor de variância (que assume somente valores positivos) é a distribuição gama, cuja FDP depende de dois parâmetros positivos, o parâmetro de forma  $\alpha$  e o parâmetro de escala  $\beta$ :

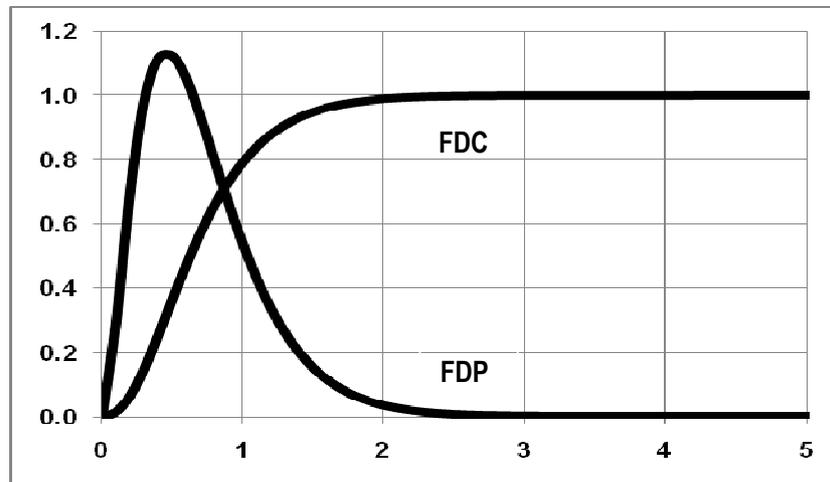
$$f(\omega | \alpha, \beta) = \frac{\omega^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} \exp\left(-\frac{\omega}{\beta}\right) \quad (10)$$

onde  $\Gamma$  é a função gama (Casaca *et al.*, 2009). As estimativas de máxima verosimilhança dos parâmetros de forma e de escala baseadas na amostra de factores de variância apresentada no Quadro 1 são  $\alpha = 2.83$  e  $\beta = 0.25$  (cf. Casaca *et al.*, 2009). A partir destes parâmetros de forma e de escala, obtêm-se a moda (i), o valor médio (ii) e a variância (iii):

$$\text{i) } \mu_0 = (\alpha - 1)\beta = 0.46, \quad \text{ii) } \mu = \alpha\beta = 0.71, \quad \text{iii) } \sigma^2 = \alpha\beta^2 = 0.18 \quad (11)$$

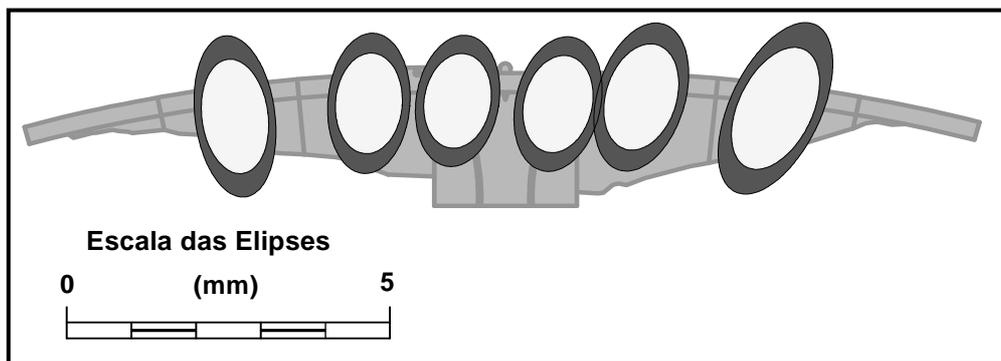
Na Figura 1, apresentam-se os gráficos da FDP e da função de distribuição cumulativa (FDC) gama com parâmetros  $\alpha = 2.83$  e  $\beta = 0.25$ . A FDC gama, permite determinar os quantis de probabilidade da distribuição do factor de variância. O quantil de probabilidade 0.95 é  $q(0.95) = 1.56$ , isto é, de acordo com a FDC gama, a probabilidade da ocorrência de um factor de variância igual ou inferior a 1.56 é 0.95.

A análise da amostra de factores de variância apresentada no Quadro 1, mostra que o factor de variância unitário adoptado com o desvio padrão  $\sigma_0 = 1.5$  dmgon, que foi usado na construção da matriz de variância ( $\Sigma$ ) das observações angulares deverá ser actualizado para 0.71 (média dos valores observados), o que equivale a construir a matriz de variância das observações com um desvio padrão  $\sigma_0 = 1.26$  dmgon. Por outro lado, a FDC da distribuição gama mostra que o factor de variância 1.56 é, com uma probabilidade elevada (0.95), um majorante das possíveis ocorrências dos factores de variância.



**Figura 1** – FDP e FDC gama, com parâmetro de forma  $\alpha = 2.83$  e parâmetro de escala  $\beta = 0.25$ .

Na Figura 2, apresentam-se as elipses de erro, para o nível de probabilidade 0.95, da rede de triangulação referida na Secção 4. São apresentadas duas classes de elipses: i) As elipses de erro médias construídas com o factor de variância 0.71, correspondente ao desvio padrão angular 1.26dmgon (tonalidade mais clara); ii) As elipses de erro majorantes construídas com o factor de variância 1.56, correspondente ao desvio padrão angular 1.87dmgon (tonalidade mais escura). As elipses de erro construídas com o factor de variância unitário ( $\omega = 1$ ) são pouco maiores do que as elipses de erro médias e, por isso, não são representadas na Figura.



**Figura 2** – Elipses de erro para o nível de probabilidade 0.95.

A representação conjunta das elipses de erro médias e majorantes (Figura 2) permite uma avaliação mais realista do desempenho da rede de triangulação, no que diz respeito aos erros de observação, do que a representação das elipses de erro *a priori*, sem menosprezar o importante papel que estas desempenharam inicialmente no planeamento da rede e nas primeiras observações.



## 5. CONCLUSÕES

O problema da caracterização do desempenho das redes de triangulação, em termos de qualidade dos resultados, é um problema complexo, que foi apenas esboçado nesta comunicação. Embora a introdução da distribuição do factor de variância ajude a modelar a variabilidade da qualidade das observações, qualquer solução satisfatória deverá passar pela abordagem do problema numa perspectiva Bayesiana, onde a distribuição do factor de variância seja adoptada como uma distribuição *a priori* informativa.

Numa eventual análise Bayesiana, será preferível usar uma distribuição gama invertida para modelar a distribuição do factor de variância (o que não apresenta grande dificuldade), dadas as propriedades teóricas desta distribuição como conjugada natural da distribuição multinormal (Casaca *et al*, 2008).

### Referências

- Casaca, J. (2001). O Método da Variação de Coordenadas na Observação Geodésica de Barragens. Lisboa: LNEC, Série ICT, ITB21.
- Casaca, J. (2004). A Refracção Atmosférica em Geodesia Aplicada. Lisboa: LNEC, Série ICT, INCB9.
- Casaca, J., Mateus, P. and Coelho, J. (2008). Bayesian Estimation in Dam Monitoring Networks. Proc. of the 4<sup>th</sup> IAG Int. Symp. on Geodesy for Geotechnical and Structural Engineering, Lisbon, May 12-15.
- Casaca, J., Charneca, V. e Mateus, P. (2009). A Análise Estatística de Amostras com Recurso à Distribuição Gama. Actas da VI Conferência Nacional de Cartografia e Geodesia, Caldas da Rainha, Maio 7-8.
- Morrison, D. F. (1990). Multivariate Statical Methods. New York: John Wiley and Son.

### Contactos

João CASACA

E-mail: [jmmcasaca@gmail.com](mailto:jmmcasaca@gmail.com)

Instituição: Laboratório Nacional de Engenharia Civil – Núcleo de Geodesia Aplicada

Endereço “web” da Instituição: [www.lnec.pt](http://www.lnec.pt)

País: Portugal