Weighting observations from multi-station coordinate measuring systems

A Forbes, B Hughes, W Sun
National Physical Laboratory, Teddington, TW11 0LW, UK
alistair.forbes@npl.co.uk

MathMet 2010, 20–21 June, PTB Berlin

Overview

Multi-station coordinate measuring systems

Posterior distribution for noise parameters

Maximising the posterior distribution

Application to laser tracker calibration

Large scale coordinate metrology

Target located at $\mathbf{x}$, station located at $\mathbf{p}$, orientation defined by rotation matrix $R(\boldsymbol{\alpha})$.

Distance to target given by $\|\mathbf{x}-\mathbf{p}\|$.

Azimuth and elevation angles: if $\mathbf{n} = \|\mathbf{x}-\mathbf{p}\|/d^*$ then there are $\theta^* = \theta^*(\mathbf{x},\mathbf{p},\boldsymbol{\alpha})$ and $\phi^* = \phi^*(\mathbf{x},\mathbf{p},\boldsymbol{\alpha})$ such that

$$\begin{bmatrix} \cos\theta^*\cos\phi^* \\ \sin\theta^*\cos\phi^* \\ \sin\phi^* \end{bmatrix} = R(\boldsymbol{\alpha})\mathbf{n}.$$

Observations

$$d = d^* + \epsilon_D, \quad , d^* = (1+\mu)\|\mathbf{x}-\mathbf{p}\| + \lambda, \quad \theta = \theta^* + \epsilon_A, \quad \phi = \phi^* + \epsilon_E,$$

involving random effects

$$\epsilon_D \in \mathrm{N}(0,\sigma_D^2), \quad \epsilon_A \in \mathrm{N}(0,\sigma_A^2), \quad \epsilon_E \in \mathrm{N}(0,\sigma_E^2),$$

due to refractive index changes, etc.

Multi-station systems

Target parameters $\mathbf{x}_j$, coordinates of the location of the jth target

Configuration parameters $\mathbf{b}$: locations $\mathbf{p}_k$, orientation angles $\boldsymbol{\alpha}_k$ associated with the measuring stations, other systematic effects λ_l, μ_r, etc.

Nonlinear least squares problem (MLE):

$$\min_{\{\mathbf{x}_j\},\mathbf{b}} \sum_i \left\{ w_{i,D}^2 (d_i - d_i^*(\mathbf{x},\mathbf{b}))^2 + w_{i,A}^2 (\theta_i - \theta_i^*(\mathbf{x},\mathbf{b}))^2 + w_{i,E}^2 (\phi_i - \phi_i^*(\mathbf{x},\mathbf{b}))^2 \right\}.$$

Weights $w_{i,D} = 1/\sigma_{i,D}$, etc.

Need to weight angle measurements relative to distance measurements.

Adjusting noise parameters: single sensor case

Suppose

$$\mathbf{y}|\boldsymbol{\alpha} \sim \mathsf{N}(C\boldsymbol{\alpha}, \sigma^2 I), \quad C, \quad m \times n, \quad y_i = \alpha_1 + \alpha_2 x_i + \epsilon_i, \quad \epsilon_i \in \mathsf{N}(0, \sigma^2).$$

Least squares solution:

$$\mathbf{a} = \left(C^\top C\right)^{-1} C^\top \mathbf{y} = R_1^{-1} Q_1^\top \mathbf{y}, \quad C = QR = [Q_1 \; Q_2] \left[\begin{array}{c} R_1 \\ \mathbf{0} \end{array} \right].$$

Also

$$\hat{\mathbf{y}} = C\mathbf{a} = Q_1 Q_1^\top \mathbf{y}, \quad \mathbf{r} = \mathbf{y} - \hat{\mathbf{y}} = (I - Q_1 Q_1^\top)\mathbf{y} = Q_2 Q_2^\top \mathbf{y},$$

so that

$$\mathbf{a}|\boldsymbol{\alpha}, \sigma \sim \mathsf{N}(\boldsymbol{\alpha}, \sigma^2 \left(C^\top C\right)^{-1}), \quad \mathbf{r}|\boldsymbol{\alpha}, \sigma \sim \mathsf{N}(\mathbf{0}, \sigma^2 Q_2 Q_2^\top).$$

If $\tilde{\mathbf{r}} = Q_2^\top \mathbf{r}$,

$$\tilde{\mathbf{r}}|\boldsymbol{\alpha}, \sigma \sim \mathsf{N}(\mathbf{0}, \sigma^2 I_{m-n}), \quad \frac{1}{\sigma^2}\mathbf{r}^\top \mathbf{r} = \frac{1}{\sigma^2}\tilde{\mathbf{r}}^\top \tilde{\mathbf{r}} \sim \chi^2_{m-n}.$$

Choose $\hat{\sigma}$ such the observed $\mathbf{r}^\top \mathbf{r}$ is $\mathsf{E}(\mathbf{r}^\top \mathbf{r}|\hat{\sigma})$:

$$\hat{\sigma}^2 = \frac{\mathbf{r}^\top \mathbf{r}}{m-n}, \quad \mathsf{E}(\hat{\sigma}|\sigma) = \sigma.$$

National Physical Laboratory

Partial information about σ

Classical estimation caters for σ known or completely unknown. Suppose there is partial information about σ, e.g., from a previous experiment.

Suitable prior distribution for $\phi = 1/\sigma^2$:

$$m_0\sigma_0^2\phi \sim \chi^2_{m_0}, \quad p(\phi) \propto \phi^{m_0/2-1} \exp\left\{-\frac{\phi}{2}m_0\sigma_0^2\right\},$$

where σ_0 is the prior estimate and m_0 encodes degree of belief in σ_0. This prior corresponds to information gained about σ by observing a standard deviation of σ_0 of m_0 samples $z_i \in \mathrm{N}(0, \sigma^2)$.

Applying Bayes' theorem

$$p(\boldsymbol{\alpha}, \sigma|\mathbf{y}) \propto \phi^{(m_0+m)/2-1} \exp\left\{-\frac{\phi}{2}\left[m_0\sigma_0^2 + ms^2 + (\boldsymbol{\alpha} - \mathbf{a})^\top C^\top C(\boldsymbol{\alpha} - \mathbf{a})\right]\right\},$$

where

$$\mathbf{a} = (C^\top C)^{-1}C^\top\mathbf{y}, \quad \mathbf{r} = \mathbf{y} - C\mathbf{a}, \qquad ms^2 = \mathbf{r}^\top\mathbf{r}.$$

Marginal distributions

Model parameters:

$$p(\boldsymbol{\alpha}|\mathbf{y}) = \int_S p(\boldsymbol{\alpha}, \sigma|\mathbf{y}) d\sigma \Rightarrow \boldsymbol{\alpha} \sim t_\nu(\mathbf{a}, \bar{V}),$$

where

$$\nu = m_0 + m - n, \quad \bar{V} = \bar{\sigma}^2 (C^\top C)^{-1}, \quad \bar{\sigma}^2 = \frac{m_0 \sigma_0^2 + m s^2}{m_0 + m - n}.$$

If $m_0 \mapsto \infty$, $\bar{\sigma}^2 \mapsto \sigma_0^2$, if $m_0 \mapsto 0$, $\bar{\sigma}^2 \mapsto \widehat{\sigma}^2$.

Noise parameter:

$$p(\sigma|\mathbf{y}) = \int_A p(\boldsymbol{\alpha}, \sigma|\mathbf{y}) d\boldsymbol{\alpha} \Rightarrow \nu \bar{\sigma}^2 \phi \sim \chi_\nu^2.$$

National Physical Laboratory

Multiple sensors

Suppose

$$\mathbf{y}_k|\boldsymbol{\alpha},\sigma_k \sim \mathsf{N}(C_k\boldsymbol{\alpha},\sigma_k^2 I), \quad k=1,\ldots,n_K,$$

where C_k is $m_k \times n$. Let

$$V_{\boldsymbol{\sigma}} = \begin{bmatrix} \sigma_1^2 I_{m_1} & & \\ & \ddots & \\ & & \sigma_{n_K}^2 I_{m_{n_K}} \end{bmatrix}, \quad C = \begin{bmatrix} C_1 \\ \vdots \\ C_{n_K} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{n_K} \end{bmatrix}, \quad m = \sum_{k=1}^{n_K} m_k.$$

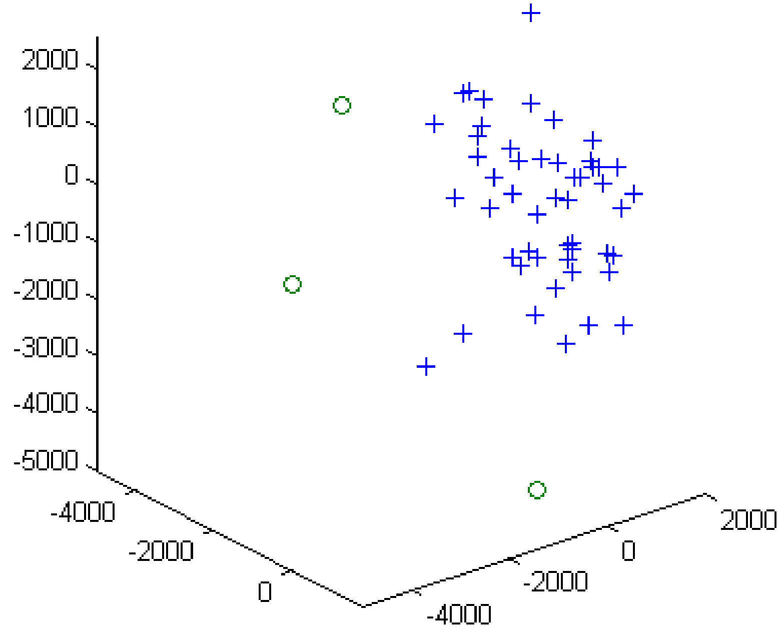
Naive posterior estimate of σ_k if $m_k > n$: given estimates $\sigma_{0,k}$, determine the least squares solution $\mathbf{a}$ of

$$\min_{\boldsymbol{\alpha}} (\mathbf{y} - C\boldsymbol{\alpha})^\top V_{\boldsymbol{\sigma}}^{-1}(\mathbf{y} - C\boldsymbol{\alpha}),$$

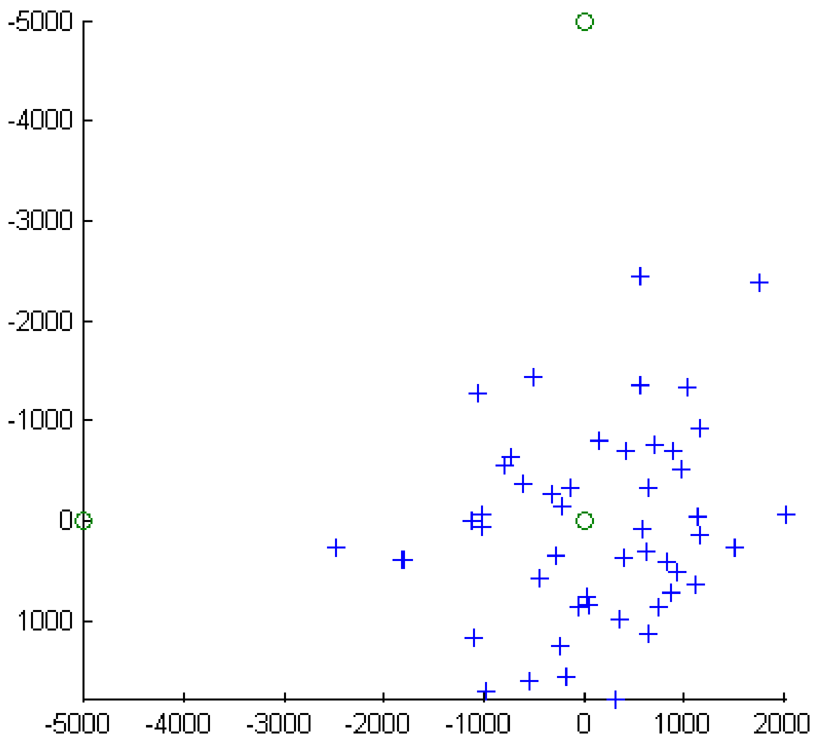
and set

$$\mathbf{r}_k = \mathbf{y}_k - C\mathbf{a}, \quad \hat{\sigma}_k^2 = \frac{m_k}{m}\frac{\mathbf{r}_k^\top \mathbf{r}_k}{m-n}.$$

Example: three laser trackers



Example: three laser trackers



Example: three laser trackers

$$\widehat{\sigma}_k^2 = \frac{m_k}{m}\frac{\mathbf{r}_k^\top \mathbf{r}_k}{m-n}$$

$$\widehat{\sigma}_D = 0.39, \quad \widehat{\sigma}_A = \widehat{\sigma}_E = 1.17.$$

Use these estimates to reweight, and recalculate $\widehat{\sigma}_k$:

$$\widehat{\sigma}_D = 0.13, \quad \widehat{\sigma}_A = \widehat{\sigma}_E = 1.05.$$

Better to compare $\mathbf{r}$ with $u(\mathbf{r})$ derived from $\mathbf{r}|\boldsymbol{\alpha}, \sigma \sim \mathrm{N}(\mathbf{0}, \sigma^2 Q_2 Q_2^\top)$, to take into account leverage (but ignoring correlation).

(ILCs: $C = (5,1,1,1,1,1,1,1)^\top$, $u(\mathbf{r}) = (0.26,1,1,1,1,1,1,1)^\top$. [Remember, $\mathbf{r}$ is the *weighted* residual.])

Multiple sensors, Bayesian approach

Prior information about $\phi_k = 1/\sigma_k^2$ given by

$$m_{0,k}\sigma_{0,k}^2\phi_k \sim \chi^2_{m_{0,k}}.$$

Assuming prior independence of $p(\phi_k)$,

$$p(\boldsymbol{\phi}) \propto \left(\prod_{k=1}^{n_K} \phi_k^{n_{0,k}/2-1}\right) \exp\left\{-\frac{1}{2}\sum_{k=1}^{n_K} n_{0,k}\sigma_{0,k}^2\phi_k\right\}.$$

Let

$$V_{\boldsymbol{\phi}} = \begin{bmatrix} \phi_1^{-1} I_{m_1} & & \\ & \ddots & \\ & & \phi_{n_K}^{-1} I_{m_{n_K}} \end{bmatrix}, \quad C = \begin{bmatrix} C_1 \\ \vdots \\ C_{n_K} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{n_K} \end{bmatrix}, \quad m = \sum_{k=1}^{n_K} m_k.$$

If $p(\boldsymbol{\alpha}) \propto 1$, then the posterior distribution for $\boldsymbol{\alpha}$ and $\boldsymbol{\phi}$ is such that

$$p(\boldsymbol{\alpha}, \boldsymbol{\phi}|\mathbf{y}) \propto p(\boldsymbol{\phi})|V_{\boldsymbol{\phi}}|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{y} - C\boldsymbol{\alpha})^\top V_{\boldsymbol{\phi}}^{-1}(\mathbf{y} - C\boldsymbol{\alpha})\right\}.$$

Marginalised distribution

The marginalised distribution $p(\phi|\mathbf{y})$ is given by

$$p(\phi|\mathbf{y}) = \int p(\boldsymbol{\alpha}, \phi|\mathbf{y})\mathrm{d}\boldsymbol{\alpha}.$$

Let $\mathbf{a} = \mathbf{a}(\phi)$ solve

$$\min_{\boldsymbol{\alpha}} \quad (\mathbf{y} - C\boldsymbol{\alpha})^{\top} V_{\boldsymbol{\phi}}^{-1}(\mathbf{y} - C\boldsymbol{\alpha})$$

and set

$$(ms^2 = \mathbf{r}^{\top}\mathbf{r} =) \quad F(\phi) = (\mathbf{y} - C\mathbf{a})V_{\boldsymbol{\phi}}^{-1}(\mathbf{y} - C\mathbf{a}).$$

Then

$$p(\phi|\mathbf{y}) \propto p(\phi)|V_{\boldsymbol{\phi}}|^{-1/2}|C^{\top}V_{\boldsymbol{\phi}}^{-1}C|^{-1/2} \exp\left\{-\frac{1}{2}F(\phi)\right\}.$$

The ϕ that maximises $p(\phi|\mathbf{y})$ can be found by minimising $-\log p(\phi|\mathbf{y})$ or equivalently, minimising

$$E(\phi|\mathbf{y}) = \frac{1}{2}\left(F(\phi) + \log|V_{\boldsymbol{\phi}}| + \log|C^{\top}V_{\boldsymbol{\phi}}^{-1}C|\right) - \log p(\phi).$$

Evaluating $E(\boldsymbol{\phi}|\mathbf{y}) = \frac{1}{2}\left(F(\boldsymbol{\phi}) + \log|V_{\boldsymbol{\phi}}| + \log|C^\top V_{\boldsymbol{\phi}}^{-1}C|\right) - \log p(\boldsymbol{\phi})$

Firstly,

$$\log|V_{\boldsymbol{\phi}}| = -\sum_{k=1}^{n_k} m_k \log\phi_k.$$

Let

$$V_{\boldsymbol{\phi}}^{-1} = W_{\boldsymbol{\phi}}^2, \quad C_{\boldsymbol{\phi}} = W_{\boldsymbol{\phi}}C, \quad \mathbf{y}_{\boldsymbol{\phi}} = W_{\boldsymbol{\phi}}\mathbf{y}.$$

If $C_{\boldsymbol{\phi}}$ has QR factorisation

$$C_{\boldsymbol{\phi}} = [Q_1 \ Q_2]\begin{bmatrix} R_1 \\ \mathbf{0} \end{bmatrix} = Q_1R_1,$$

then

$$\log|C^\top V_{\boldsymbol{\phi}}^{-1}C| = \sum_{i=1}^{n} \log r_{ii}^2$$

where r_{ii} is the ith diagonal element of R_1, and

$$F(\boldsymbol{\phi}) = \mathbf{q}_2^\top\mathbf{q}_2, \quad \mathbf{q}_2 = Q_2^\top\mathbf{y}.$$

Evaluating $\partial E(\boldsymbol{\phi}|\mathbf{y})/\partial\phi_k$

Matrix determinant lemma: if A is square, full rank and a function of ϕ, then

$$\frac{d}{d\phi}\log|A| = \mathrm{Tr}\left[A^{-1}\frac{dA}{d\phi}\right].$$

Let

$$W_k = \frac{\partial W_{\boldsymbol{\phi}}}{\partial\phi_k} = \mathrm{diag}(\mathbf{0}, \mathbf{1}_k, \mathbf{0})^{\top}.$$

We have

$$\frac{\partial}{\partial\phi_k}\log|V_{\boldsymbol{\phi}}| = -\frac{m_k}{\phi_k},$$

$$\frac{\partial}{\partial\phi_k}\log|C^{\top}V_{\boldsymbol{\phi}}^{-1}C| = 2\mathrm{Tr}(S_k), \quad R_1 S_k = Q_1^{\top}W_k C$$

NPL National Physical Laboratory

Derivative calculations for least squares with matrices of functions

Let $m \times n$ matrix $C = C(\phi)$ be a function of a parameter ϕ and let $\dot{C} = dC/d\phi$ and

$$C = [Q_1 \; Q_2] \left[\begin{array}{c} R_1 \\ \mathbf{0} \end{array} \right], \quad \mathbf{a} = (C^\top C)^{-1} C^\top \mathbf{y} \quad \mathbf{r} = \mathbf{y} - C\mathbf{a}.$$

Then

$$\frac{d}{d\phi} \left[C(C^\top C)^{-1} C^\top \right] = F + F^\top, \quad F = Q_2 Q_2^\top \dot{C} R_1^{-1} Q_1^\top,$$

and

$$\frac{d}{d\phi}\mathbf{a} = (F + F^\top)\mathbf{y}, \quad \frac{d}{d\phi}\mathbf{r}^\top\mathbf{r} = -2\mathbf{r}^\top \dot{C}\mathbf{a}.$$

If $\mathbf{y}$ is also a function of ϕ, then

$$\frac{d}{d\phi}\mathbf{r}^\top\mathbf{r} = 2\mathbf{r}^\top \left(\dot{\mathbf{y}} - \dot{C}\mathbf{a} \right).$$

Application to $F(\phi)$

$$\frac{\partial F}{\partial \phi_k}(\phi) = 2(\mathbf{y}_{\boldsymbol{\phi}} - C_{\boldsymbol{\phi}}\mathbf{a})^{\top} W_k(\mathbf{y} - C\mathbf{a}) = 2\mathbf{q}_2^{\top} W_k W_{\boldsymbol{\phi}}^{-1}(Q_2\mathbf{q}_2).$$

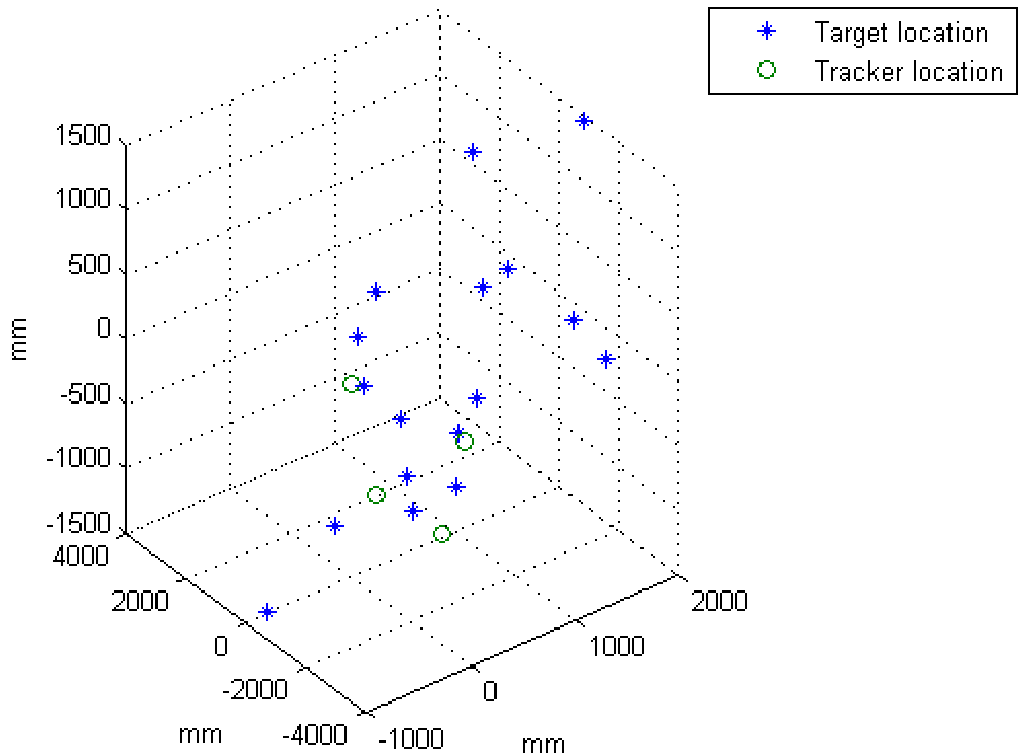
Partial upper-triangularisation

Let Q_k be such that

$$Q_k^{\top} \left[\begin{array}{cc} C_k & \mathbf{y}_k \end{array} \right] = \left[\begin{array}{cc} R_k & \mathbf{q}_k \\ & q_{0,k} \end{array} \right].$$

Then the least squares problem defined by $\{C_k, \mathbf{y}_k\}$ is equivalent to that defined by $\{R_k, \mathbf{q}_k, q_{0,k}\}$.

Tracker calibration experiments



Results

	sigma_0	sigma(Bayes)
distance	5.00e-3 mm	2.86e-3 mm
angle	5.00e-6 rad	4.31e-6 rad

Conclusions

Modal estimates of noise parameters obtained by maximising the marginalised posterior distribution $p(\boldsymbol{\phi}|\mathbf{y})$.

Applied to large scale coordinate metrology using laser trackers.

Calculations tractable using matrix factorisations, etc.

Posterior sampling algorithms: Markov chain Monte Carlo (Metropolis-Hastings with i) independent sampling, ii) random walks, iii) Hamiltonian (hybrid) Monte Carlo.

Acknowledgements

Work supported by the EMRP JRP T3.J2.2 Metrology for New Industrial Measurement Technologies (PTB, NPL, INRIM, CMI, University of Maribor)

Support from MathMet 2010 gratefully acknowledged.

Selected references

A. B. Forbes. Surface fitting taking into account uncertainty structure in coordinate data. *Measurement Science and Technology*, 17:553–558, 2006.

A. B. Forbes. Uncertainty evaluation associated with fitting geometric surfaces to coordinate data. *Metrologia*, 43(4):S282–S290, August 2006.

A. B. Forbes. Nonlinear least squares and Bayesian inference. In F. Pavese, M. Bär, A. B. Forbes, J.-M. Linares, C. Perruchet, and N.-F. Zhang, editors, *Advanced Mathematical and Computational Tools in Metrology VIII*, pages 103–111, Singapore, 2009. World Scientific.

A. B. Forbes. Parameter estimation based on least squares methods. In F. Pavese and A. B. Forbes, editors, *Data modeling for metrology and testing in measurement science*, pages 147–176, New York, 2009. Birkhäuser-Boston.

A. B. Forbes and P. M. Harris. Uncertainty associated with co-ordinate measurements. In P. Shore, editor, *Laser Metrology and Machine Performance VII*, pages 30–39, Bedford, 2005. Euspen.

A. B. Forbes, E. B. Hughes, and W. Sun. Comparison of measurements in coordinate metrology. *Measurement*, pages 1473–1477. 2009.

A. B. Forbes and C. E. Matthews. Least squares, outliers and coordinate metrology. In K. Cheng and A. B. Forbes, editors, *Laser Metrology and Machine Performance IX*, pages 253–262, Bedford, 2009. Euspen.

A. Gelman, J. B. Carlin, H. S. Stern and D B Rubin, *Bayesian Data Analysis*, Chapman&Hall/CRC, Boca Raton, 2004.

G. N. Peggs, P. G. Maropoulos, E. B. Hughes, A. B. Forbes, S. Robson, M. Ziebart, and B. Muralikrishnan. Recent developments in large-scale dimensional metrology. *Journal of Engineering Manufacture: Part B*, 223:571–595, 2009.