8. Special topics in least squares estimation

8.1. Combinations of Models

- Assume that observations are made in two groups, with the second group consisting of one or several observations. Both groups have a common set of parameters, i.e.

- **Given:** 2 sets of observations collected at different times for the same parameters

i.e. $\overbrace{\boldsymbol{l}_1 \ and\ \boldsymbol{C}_{\boldsymbol{l}_1}}^{n_1} \quad and \quad \overbrace{\boldsymbol{l}_2 \ and\ \boldsymbol{C}_{\boldsymbol{l}_2}}^{n_2}$

- **Required:** $\mathbf{x}_{u,1}$ The best estimate for a group of parameters from a group of measurements that have been captured at two different times (e.g. $\mathbf{l_1}$ at t_1 and $\mathbf{l_2}$ at t_2)

- **Functional model:**

$$f_1\left(\hat{x}, \hat{l}_1\right) = 0 \qquad\qquad A_{1_{m1,u}} \hat{\delta}_{u,1} + B_{11_{m1,n1}} \hat{v}_{1_{n1,1}} + w_{1_{m1,1}} = \underline{0}$$

$$f_2\left(\hat{x}, \hat{l}_2\right) = 0 \qquad\qquad A_{2_{m2,u}} \hat{\delta}_{u,1} + B_{22_{m2,n2}} \hat{v}_{2_{n2,1}} + w_{2_{m2,1}} = \underline{0}$$

- Variation function using Lagrange multipliers

$$\varphi = \hat{v}_1^T P_1 \hat{v}_1 + \hat{v}_2^T P_2 \hat{v}_2 + 2\hat{k}_1^T \left(A_1\hat{\delta} + B_{11}\hat{v}_1 + w_1\right) + 2\hat{k}_2^T \left(A_2\hat{\delta} + B_{22}\hat{v}_2 + w_2\right) = \min$$

- Note: For each group of observations, there's a quadratic form and a Lagrange multiplier

- To minimize φ, differentiate φ w.r.t. all variables ($\hat{\boldsymbol{\delta}}, \hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2, \hat{\boldsymbol{k}}_1, \hat{\boldsymbol{k}}_2$) and equate to zero.

- Then arrange in hyper-matrix notation and solve by elimination.

$$\frac{\partial \boldsymbol{\varphi}}{\partial \boldsymbol{\delta}} = \mathbf{2}\ \hat{\mathbf{k}}_1^{\mathbf{T}}\mathbf{A}_1 + \mathbf{2}\ \hat{\mathbf{k}}_2^{\mathbf{T}}\mathbf{A}_2 \qquad = \mathbf{0}$$

$$\frac{\partial \boldsymbol{\varphi}}{\partial \mathbf{v}_1} = \mathbf{2}\ \hat{\mathbf{v}}_1^{\mathbf{T}}\mathbf{P}_1 + \mathbf{2}\ \hat{\mathbf{k}}_1^{\mathbf{T}}\mathbf{B}_{11} \qquad = \mathbf{0}$$

$$\frac{\partial \boldsymbol{\varphi}}{\partial \mathbf{v}_2} = \mathbf{2}\ \hat{\mathbf{v}}_2^{\mathbf{T}}\mathbf{P}_2 + \mathbf{2}\ \hat{\mathbf{k}}_2^{\mathbf{T}}\mathbf{B}_{22} \qquad = \mathbf{0}$$

$$\frac{\partial \boldsymbol{\varphi}}{\partial \mathbf{k}_1} = \mathbf{2}\ \hat{\boldsymbol{\delta}}^{\mathbf{T}}\mathbf{A}_1^{\mathbf{T}} + \mathbf{2}\hat{\mathbf{v}}_1^{\mathbf{T}}\mathbf{B}_{11}^{\mathbf{T}} + \mathbf{2}\mathbf{w}_1^{\mathbf{T}} \quad = \mathbf{0}$$

$$\frac{\partial \boldsymbol{\varphi}}{\partial \mathbf{k}_2} = \mathbf{2}\hat{\boldsymbol{\delta}}^{\mathbf{T}}\mathbf{A}_2^{\mathbf{T}} + \mathbf{2}\hat{\mathbf{v}}_2^{\mathbf{T}}\mathbf{B}_{22}^{\mathbf{T}} + \mathbf{2}\mathbf{w}_2^{\mathbf{T}} \quad = \mathbf{0}$$

- Transpose all equations and divide by 2 and arrange in hyper matrix

$$\begin{pmatrix} P_1 & 0 & B_{11}^T & 0 & 0 \\ 0 & P_2 & 0 & B_{22}^T & 0 \\ B_{11} & 0 & 0 & 0 & A_1 \\ 0 & B_{22} & 0 & 0 & A_2 \\ 0 & 0 & A_1^T & A_2^T & 0 \end{pmatrix} \begin{pmatrix} \hat{v}_1 \\ \hat{v}_2 \\ \hat{k}_1 \\ \hat{k}_2 \\ \hat{\delta} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ w_1 \\ w_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

- To solve this system algebraically, partition and eliminate $\hat{\boldsymbol{v}}_1$
- Then partition and eliminate $\hat{\boldsymbol{v}}_2$
- Perform the elimination until a solution for $\hat{\boldsymbol{\delta}}$ is reached
- Then perform back substitution to get expressions for all other variables ($\hat{\boldsymbol{v}}_1$, $\hat{\boldsymbol{v}}_2$, $\hat{\boldsymbol{k}}_1$, and $\hat{\boldsymbol{k}}_2$)
- Apply the law of propagation of covariance to estimate the V-C matrices

8.2. Summation of Normals – Parametric Model

- Given: two group of measurements, ($\mathbf{l_1}$ and $\mathbf{C}_{\mathbf{l}_1}$) and ($\mathbf{l_2}$ and $\mathbf{C}_{\mathbf{l}_2}$), that could have been collected either at different times or by different instruments, or by different observers. Furthermore, there is no correlation between $\mathbf{l_1}$ and $\mathbf{l_2}$
- Required: $\hat{x}$ by using the two sets of observation in one solution
- Functional models:

$$\hat{l}_1 = f_1(\hat{x}) \xrightarrow{\text{with the linearized model}} A_{1_{n1,u}} \hat{\delta}_{u,1} + w_{1_{n1,1}} = \hat{v}_{1_{n1,1}}$$

$$\hat{l}_2 = f_2(\hat{x}) \xrightarrow{\text{with the linearized model}} A_{2_{n2,u}} \hat{\delta}_{u,1} + w_{2_{n2,1}} = \hat{v}_{2_{n2,1}}$$

- Variation function

$$\varphi = \hat{v}_1^T P_1 \hat{v}_1 + \hat{v}_2^T P_2 \hat{v}_2 = min$$

$$= \left(\hat{\delta}^T A_1^T + w_1^T\right) P_1 \left(A_1 \hat{\delta} + w_1\right) + \left(\hat{\delta}^T A_2^T + w_2^T\right) P_2 \left(A_2 \hat{\delta} + w_2\right)$$

$$= \hat{\delta}^T A_1^T P_1 A_1 \hat{\delta} + \underbrace{\hat{\delta}^T A_1^T P_1 w_1}_{\substack{\text{equivelant}\\ \text{scalars}}} + \underbrace{w_1^T P_1 A_1 \hat{\delta}}_{\substack{\text{equivelant}\\ \text{scalars}}} + w_1^T P_1 w_1$$

$$+ \hat{\delta}_T A_2 P_2 A_1 \hat{\delta} + \underbrace{\hat{\delta}^T A_2 P_2 w_2}_{\substack{\text{equivelant}\\ \text{scalars}\\ \longrightarrow}} + \underbrace{w_1^T P_1 A_1 \hat{\delta}}_{\substack{\text{equivelant}\\ \text{scalars}\\ \longleftarrow}} + w_2^T P_2 w_2 = min$$

$$\varphi = \hat{\delta}^T A_1^T P_1 A_1 \hat{\delta} + 2\hat{\delta}^T A_1^T P_1 w_1 + w_1^T P_1 w_1$$

$$+ \hat{\delta}^T A_2^T P_2 A_2 \hat{\delta} + 2\hat{\delta}^T A_2 P_2 w_2 + w_2^T P_2 w_2 = min$$

- Now minimize φ (φ is a function of only $\hat{\delta}$)

$$\varphi = \hat{\delta}^T A_1^T P_1 A_1 \hat{\delta} + 2\hat{\delta}^T A_1^T P_1 w_1 + \underbrace{w_1^T P_1 w_1}_{\text{Constant}}$$

$$+ \hat{\delta}^T A_2^T P_2 A_2 \hat{\delta} + 2\hat{\delta}^T A_2 P_2 w_2 + \underbrace{w_2^T P_2 w_2}_{\text{Constant}} = min$$

$$\frac{\partial \varphi}{\partial \delta} = 2\ \hat{\delta}^T A_1^T P_1\ A_1 + 2\ w_1^T P_1 A_1 \qquad + \ 0$$

$$2\ \hat{\delta}^T A_2^T P_2\ A_2 + 2\ w_2^T P_2 A_2 \qquad + 0 \qquad = 0$$

- Transpose and divide by 2

$$\left(\underbrace{\underbrace{A_1^T P_1 A_1}_{u,u}}_{N_1} + \underbrace{\underbrace{A_2^T P_2 A_2}_{u,u}}_{N_2} \right) \hat{\delta} + \left(\underbrace{\underbrace{A_1^T P_1 w_1}_{u,1}}_{u_1} + \underbrace{\underbrace{A_2^T P_2 w_2}_{u,1}}_{u_2} \right) = 0$$

$$\therefore \hat{\delta} = -\left(N_1 + N_2\right)^{-1}\left(u_1 + u_2\right)$$

- Note: The $\hat{\delta}$ vector involves addition of the normal equation matrices and vectors corresponding to each set of observations.
- The same procedure can be applied for 3 (or more) set of observations
- For n groups of observations, the problem of summation of normal can be formulated as:

$$\hat{\delta} = -\left(\sum_{i=1}^{n} N_i\right)^{-1}\left(\sum_{i=1}^{n} u_i\right)$$

Variance propagation to estimate C_δ and $C_{\hat{x}}$ (assume $C_l^{-1} = P$)

$$\hat{\delta} = -\left(N_1 + N_2\right)^{-1}\left(\underbrace{A_1^T C_{l_1}^{-1}\left(f\left(x^o\right) - l_1\right)}_{u_1} + \underbrace{A_2^T C_{l_2}^{-1}\left(f\left(x^o\right) - l_2\right)}_{u_2}\right)$$

x^o: Non-stochastic and therefore, $\hat{\delta}$ is a function of l

$$\therefore C_\delta = J\ \ C_l\ \ J^T$$

$$\therefore C_\delta = \frac{\partial \delta}{\partial l} C_l \left(\frac{\partial \delta}{\partial l}\right)^T$$

Where:

$$\mathrm{l}_{(n_1+n_2)} = \begin{pmatrix} \mathrm{l}_1 \\ \mathrm{l}_2 \end{pmatrix} \qquad \mathrm{C}_{\mathrm{l}_{(n_1+n_2)(n_1+n_2)}} = \begin{pmatrix} \mathrm{C}_{\mathrm{l}_1} & 0 \\ 0 & \mathrm{C}_{\mathrm{l}_2} \end{pmatrix}$$

$$\frac{\partial\delta}{\partial \mathrm{l}} = \left(\frac{\partial\delta}{\partial \mathrm{l}_1} \frac{\partial\delta}{\partial \mathrm{l}_2} \right)$$

$$\frac{\partial\boldsymbol{\delta}}{\partial\boldsymbol{l_1}} = (\boldsymbol{N_1}+\boldsymbol{N_2})^{-1}\boldsymbol{A_1^T C_{l_1}^{-1}} = \boldsymbol{N^{-1} A_1^T C_{l_1}^{-1}}$$

$$\frac{\partial\boldsymbol{\delta}}{\partial\boldsymbol{l_2}} = (\boldsymbol{N_1}+\boldsymbol{N_2})^{-1}\boldsymbol{A_2^T C_{l_2}^{-2}} = \boldsymbol{N^{-1} A_2^T C_{l_2}^{-1}}$$

$$\therefore \boldsymbol{C_\delta} = \underbrace{\begin{pmatrix} \boldsymbol{N^{-1}A_1^T C_{l_1}^{-1}} & \boldsymbol{N^{-1}A_2^T C_{l_2}^{-1}} \end{pmatrix}}_{\downarrow} \underbrace{\begin{pmatrix} \boldsymbol{C_{l_1}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C_{l_2}} \end{pmatrix} \begin{pmatrix} \boldsymbol{C_{l_1}^{-1}} & \boldsymbol{A_1} & \boldsymbol{N^{-1}} \\ \boldsymbol{C_{l_2}^{-1}} & \boldsymbol{A_2} & \boldsymbol{N^{-1}} \end{pmatrix}}_{\begin{pmatrix} \boldsymbol{C_{l_1}C_{l_1}^{-1}} & \boldsymbol{A_1N^{-1}} \\ \underbrace{\boldsymbol{C_{l_2}C_{l_2}^{-1}}}_{\boldsymbol{I}} & \boldsymbol{A_2N^{-1}} \end{pmatrix}}$$

$$= \boldsymbol{N^{-1}A_1^T C_{l_1}^{-1} A_1 N^{-1}} + \boldsymbol{N^{-1}A_2^T C_{l_2}^{-1} A_2 N^{-1}}$$

$$= \boldsymbol{N^{-1}}\left(\boldsymbol{A_1^T C_{l_1}^{-1} A_1} + \boldsymbol{A_2^T C_{l_2}^{-1} A_2}\right)\boldsymbol{N^{-1}}$$

$$= \boldsymbol{N^{-1}}(\boldsymbol{N_1}+\boldsymbol{N_2})\boldsymbol{N^{-1}} = \boldsymbol{N^{-1}NN^{-1}}$$

$$\therefore \boldsymbol{C_\delta} = \boldsymbol{N^{-1}} = (\boldsymbol{N_1}+\boldsymbol{N_2})^{-1}$$

$$\therefore \hat{\boldsymbol{x}} = \boldsymbol{x^o} + \boldsymbol{\delta}$$

$$\therefore \boldsymbol{C_{\hat{x}}} = \boldsymbol{C_\delta}$$

(Ignore that the inverse of the normal equations matrix must be scaled by the a-posteriori variance factor for the time being.)

8.3. Sequential Least Squares – Parametric Model

- In the previous section (summation of normals), it was shown that two (or more) sets of observations (l_1 and l_2) for the same set of parameters can be combined to get a new solution

$$\hat{\boldsymbol{\delta}} = -(\boldsymbol{N_1} + \boldsymbol{N_2})^{-1}(\boldsymbol{u_1} + \boldsymbol{u_2}), \text{ where}$$

$$\boldsymbol{N_1} = \underbrace{\boldsymbol{A}^T_{1_{u,n1}} \boldsymbol{P}_{1_{n1,n1}} \boldsymbol{A}_{1_{n1,u}}}_{u,u}, \text{ and} \quad \boldsymbol{N_2} = \underbrace{\boldsymbol{A}^T_{2_{un2}} \boldsymbol{P}_{2_{n2,n2}} \boldsymbol{A}_{2_{n2,u}}}_{u,u}$$

- What if $n_2 << u$ (e.g. $n_2 = 1$ and $u= 4$)? With the summation of normals method, a (u,u) matrix must be inverted again to add the new (single) observation (the assumption here the solution has been obtained already for the 1st group of observations).

- This can be a significant computational burden especially when observations are being added at a regular interval in real time (as for example in GPS positioning).

- Sequential least squares provides a method where only (n_2, n_2) inversion is required for updating the solution with new n_2 observations.

Derivation

To derive the sequential expressions, the summation of normals solution is re-written as:

$$\boldsymbol{\delta} = -\left(\boldsymbol{N_1} + \boldsymbol{A_2^T C_{l_2}^{-1} A_2}\right)^{-1}\left(\boldsymbol{u_1} + \boldsymbol{A_2^T C_{l_2}^{-1} w_2}\right)$$

$$\boldsymbol{\delta} = -\left(\boldsymbol{N_1} + \boldsymbol{A_2^T C_{l_2}^{-1} A_2}\right)^{-1}\boldsymbol{u_1} - \left(\boldsymbol{N_1} + \boldsymbol{A_2^T C_{l_2}^{-1} A_2}\right)^{-1}\boldsymbol{A_2^T C_{l_2}^{-1} w_2}$$

To proceed, two-matrix inversion lemmas are utilized.

$$\left(S^{-1} + T^T R^{-1} T\right)^{-1} = S - ST^T\left(R + TST^T\right)^{-1}TS \ldots\ldots\ldots(i)$$

$$\left(S^{-1} + T^T R^{-1} T\right)^{-1} T^T R^{-1} = ST^T\left(R + TST^T\right)^{-1} \ldots\ldots\ldots(ii)$$

Apply lemma (i) to the 1st term and lemma (ii) to the 2nd term of the δ equation with:

$\mathrm{S} = \mathrm{N}_1^{-1}$ $\quad$ $\mathrm{T} = \mathrm{A}_2$ $\quad$ $\mathrm{R} = \mathrm{C}_{l_2}$

1st term:

$$-\underbrace{\left(N_1 + A_2^T C_{l_2}^{-1} A_2\right)^{-1}}_{LHS\ of\ Lemma(i)} \quad u_1 = -\underbrace{\left[N_1^{-1} - N_1^{-1} A_2^T \left(C_{l_2} + A_2 N_1^{-1} A_2^T\right)^{-1} A_2 N_1^{-1}\right]}_{RHS\ of\ Lemma(i)} \quad u_1$$

2nd term:

$$-\underbrace{\left(N_1 + A_2^T C_{l_2}^{-1} A_2\right)^{-1} A_2^T C_{l_2}^{-1}}_{LHS\ of\ Lemma(ii)} \quad w_2 = -\underbrace{N_1^{-1} A_2^T \left(C_{l_2} + A_2 N_1^{-1} A_2^T\right)^{-1}}_{RHS\ of\ Lemma(ii)} \quad w_2$$

Combining terms 1 and 2

$$\delta = -\left[N_1^{-1} - N_1^{-1} A_2^T \left(C_{l_2} + A_2 N_1^{-1} A_2^T\right)^{-1} A_2 N_1^{-1}\right] u_1 - N_1 A_2^T \left(C_{l_2} + A_2 N_1^{-1} A_2^T\right)^{-1} w_2$$

$$\delta = -N_1^{-1} u_1 + N_1^{-1} A_2^T \left(C_{l_2} + A_2 N_1^{-1} A_2^T\right)^{-1} A_2 N_1^{-1} u_1 \quad - N_1^{-1} A_2^T \left(C_{l_2} + A_2 N_1^{-1} A_2^T\right)^{-1} w_2$$

$$\delta = \underbrace{-N_1^{-1} u_1}_{\delta(-)} + N_1^{-1} A_2^T \left(C_{l_2} + A_2 N_1^{-1} A_2^T\right)^{-1} \left(A_2 \underbrace{N_1^{-1} u_1}_{-\delta(-)} - w_2\right)$$

Now set:

$-N_1^{-1} u_1 = \delta(-)$ Solution before update with new observation (s)

$\delta \quad = \delta(+)$ Updated solution with new observation(s)

$$\therefore \underset{u,1}{\delta(+)} = \underset{u,1}{\delta(-)} - \underset{u,u}{N_1^{-1}} \underset{u,n2}{A_2^T} \underbrace{\left(\underset{n_2,n_2}{C_{l_2}} + \underset{n_2,u}{A_2} \underset{u,u}{N_1^{-1}} \underset{u,n_2}{A_2^T}\right)^{-1}}_{n_2,n_2} \left(\underset{n_2,u}{A_2} \underset{u,1}{\delta(-)} + \underset{n_2 1}{w_2}\right)$$

$$\boxed{\begin{aligned}&\delta(+)=\delta_{u,1}(-)-K_{u,n_2}\left(A_2\delta(-)+w_2\right)\\&K=N_1^{-1}A_2^T\left(C_{l_2}+A_2N_1^{-1}A_2^T\right)^{-1}\end{aligned}}$$

K is known as the gain matrix, which quantifies how much each new observation will contribute to the corrections to the parameters

Note: 1.) Only an (n_2 x n_2) inversion is required
2.) The parameters before update are

$$\hat{\boldsymbol{x}}(-)=\boldsymbol{x^o}+\hat{\boldsymbol{\delta}}(-)$$

3.) The updated parameters are

$$\hat{\boldsymbol{x}}(+)=\boldsymbol{x^o}+\hat{\boldsymbol{\delta}}(+)$$

Note: that $\mathbf{x^0}$ (The POE) is the same in both cases

<u>Understanding the Gain Matrix:</u>

$$K=N_1^{-1}A_2^T\left(C_{l_2}+A_2N_1^{-1}A_2^T\right)^{-1}$$

- If the 2nd set of observations, l_2, is imprecise, i.e. C_{l_2} has large elements, the gain matrix will generally have small elements. Thus the new observations will not greatly contribute to the solution update.
- As the precision of l_2 increases, C_{l_2} element decreases, l_2 tends to contribute more to the solution update.
- $\boldsymbol{\delta}(+)=\boldsymbol{\delta}(-)-\boldsymbol{K}\underbrace{\left(\boldsymbol{A_2\delta}(-)+\boldsymbol{w_2}\right)}_{\text{of the form } A\delta+w=v}$

That is:

- $v_2(-)=A_2\delta(-)+w_2$ The predicted residuals of the innovations vector. Hence, the estimated residuals after the update is given by
- $v_2(+)=A_2\delta(+)+w_2$

Covariance Matrices

$$C_{\hat{x}(-)} = C_{\delta(-)} = N_1^{-1} = \left(A_1^T C_{l_1}^{-1} A_1\right)^{-1}$$
$$C_{\hat{x}(+)} = C_{\delta(+)} = C_{\delta(-)} - KA_2 C_{\delta(-)}$$

Note: $C_{\delta(+)} < C_{\delta(-)}$ due to the subtraction of $KA_2C_{\delta(-)}$ and thus the covariance matrix is improved by adding observations.

8.4. Sequential Solution of Linear Parametric Models

Step 1:
Model: $\mathbf{l_1 = f_1(x)}$

- Linearized Model:

$$\underset{n_1,u}{A_1}\ \underset{u,1}{\hat{\delta}} + \underset{n_1,1}{w_1} = \underset{n_1,1}{\hat{v}_1}$$

  - P.O.E.: x^0
  - Initial solution with n_1 observations $(n_1 \geq u)$,

$$\delta(-) = -\left(A_1^T P_1 A_1\right)^{-1} A_1^T P_1 w_1$$
$$= -(N_1)^{-1} u_1$$

where

$$P = \sigma_0^2 C_{l_1}^{-1}$$
$$C_{\delta(-)} = N_1^{-1}$$

Step 2:

Model: $\mathbf{l_2} = \mathbf{f_2}(\mathbf{x})$

Linearized model: $$\underset{n_2,u}{A_2}\ \underset{u,1}{\hat{\delta}} + \underset{n_2,1}{w_2} = \underset{n_2,1}{\hat{v}_2}$$

P.O.E.: x^0 (same as 1st model)

Update solution with n_2 observations, $(n_2 \geq 1)$

$$\delta(+) = \delta(-) - K\left(A_2\delta_{(-)} + w_2\right)$$

Note: $\delta(-)$ is the solution from step 1 (i.e. from the 1st group of observations)
Final parameter estimates

$$\hat{x} = x^0 + \delta(+)$$

Updated covariance matrix

$$C_{\delta(+)} = C_{\hat{x}} = C_{\delta(-)} - KA_2C_{\delta(-)}$$

Step 3:

Addition of a 3rd set of observations

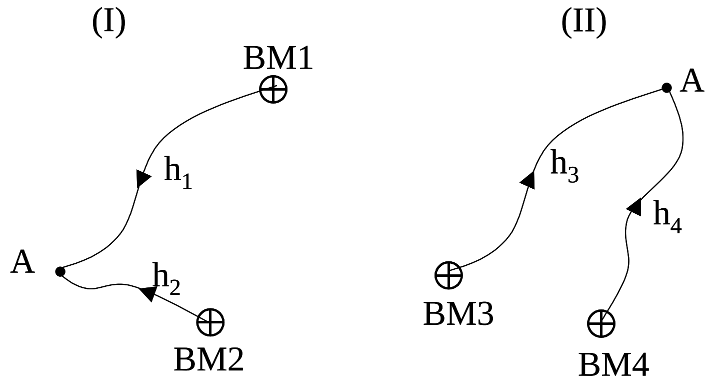
$$\mathbf{l_3} = \mathbf{f_3}(\mathbf{x})$$

$\delta(+)$ and $C_{\delta(+)}$ from step (2) become $\delta(-)$ and $C_{\delta(-)}$ respectively, for step (3).

Note: If n_2 >> u, summation of normals is the preferred method.

Example from an old final exam

The opposite figure shows a layout of two levelling networks (I and II) established for the determination of the height of point (A) from four different BM. You are given the following information:



For levelling network (I):

$$l_I = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} 3.2 \\ 0.9 \end{bmatrix} m \quad \text{and the elevation of} \quad \begin{bmatrix} BM1 \\ BM2 \end{bmatrix} = \begin{bmatrix} 7.0 \\ 9.0 \end{bmatrix} m \qquad C_{l(I)} = I \quad cm^2$$

For levelling network (II):

$$l_{II} = \begin{bmatrix} h_3 \\ h_4 \end{bmatrix} = \begin{bmatrix} 4.1 \\ 1.8 \end{bmatrix} m \quad \text{and the elevation of} \quad \begin{bmatrix} BM3 \\ BM4 \end{bmatrix} = \begin{bmatrix} 6.0 \\ 8.0 \end{bmatrix} m \qquad C_{l(II)} = I \quad cm^2$$

Calculate the elevation of point A (call it H) for the following cases:

a) Using all observations from levelling network I (i.e. l_I) and levelling network II (i.e. l_{II}) together, and using the method of parametric least squares
b) Summation of normals using all observations from levelling network I (i.e. l_I) as first solution followed by using all observations from levelling network II (i.e. l_{II})
c) Using all observations from levelling network I (i.e. l_I) followed by a sequential solution using all observations from levelling network II (i.e. l_{II})
d) Calculate the variance of the final solution [i.e. C_δ for questions (a) and (b) and $C_{\delta(+)}$ for question (c)] obtained in all the above cases
e) Are the results in (a), (b), (c) and (d) all the same? Why or why not?

a) All observations – Parametric LSA

$l = [3.2 \quad 0.9 \quad 4.1 \quad 1.8]^T$ m $\qquad C_l = I\ cm^2$ or $P = I\ cm^{-2}$ $\qquad n = 4 \qquad u = 1$

Constants = BM elevations = [7 9 6 8] m

Functional model: $\hat{l} = f(\hat{x}, c)$ $\qquad$ Linearized model $\rightarrow$ $A_{4,1}\, \hat{\delta}_{1,1} + w_{4,1} = \hat{v}_{4,1}$

$$A = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad w = f(x^0) - l$$

Observation equations, $h_1 = H - BM_1$

$h_2 = H - BM_2$

$h_3 = H - BM_3$

$h_4 = H - BM_4$

assume $x^0 = 0$ (i.e. H = 0 m)

$$\therefore \qquad w = -\begin{bmatrix} BM_1 + h_1 \\ BM_2 + h_2 \\ BM_3 + h_3 \\ BM_4 + h_4 \end{bmatrix} = -\begin{bmatrix} 10.2 \\ 9.9 \\ 10.1 \\ 9.8 \end{bmatrix} \text{m}$$

$$\therefore \qquad \hat{\delta} = -(A^T P A)^{-1} A^T P w = -\frac{\Sigma w_i}{4} = 10 \text{ m} \qquad \rightarrow \qquad \hat{x} = x^0 + \hat{\delta} = 10 \text{ m}$$

b) Parametric LSA with Summation of Normals

$\hat{\delta}$ = - $(N_1 + N_2)^{-1}$ $(u_1 + u_2)$ = $(2 + 2)^{-1}$ (20.1 + 19.9) = 10 m

where

$$N_1 = A_1^T A_1 = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2 = N_2 \quad \text{cm}^{-2}$$

$$u_1 = A_1^T w_1 = -\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 10.2 \\ 9.9 \end{bmatrix} = -20.1 \quad \text{m}\cdot\text{cm}^{-2}$$

$$u_2 = A_2^T w_2 = -\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 10.1 \\ 9.8 \end{bmatrix} = -19.9 \quad \text{m}\cdot\text{cm}^{-2}$$

$$\hat{x} = x^0 + \hat{\delta} = 10 \text{ m}$$

c) Sequential LSA

$\delta(-) = -\ N_1^{-1}u_1 = \frac{1}{2}\left[10.2+9.9\right] = 10.05\ \text{m}$ and $C_{\delta(-)} = N_1^{-1} = \frac{1}{2}$ cm^2

$$
\begin{aligned}
K &= N_1^{-1} A_2^T \left(C_{l_2} + A_2 N_1^{-1} A_2^T\right)^{-1} \\
&= \frac{1}{2}\begin{bmatrix}1 & 1\end{bmatrix}\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \begin{bmatrix}1\\ 1\end{bmatrix}\frac{1}{2}\begin{bmatrix}1 & 1\end{bmatrix}\right)^{-1} \\
&= \frac{1}{2}\begin{bmatrix}1 & 1\end{bmatrix}\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\right)^{-1} \\
&= \frac{1}{2}\begin{bmatrix}1 & 1\end{bmatrix}\begin{pmatrix}1.5 & 0.5\\ 0.5 & 1.5\end{pmatrix}^{-1} \\
&= \frac{1}{2}\begin{bmatrix}1 & 1\end{bmatrix}\frac{1}{2}\begin{pmatrix}1.5 & -0.5\\ -0.5 & 1.5\end{pmatrix} \\
&= \frac{1}{4}\begin{bmatrix}1 & 1\end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\therefore\ \delta(+) &= \delta(-) - K\left(A_2\,\delta(-) + w_2\right) \\
&= 10.05 - \frac{1}{4}\begin{bmatrix}1 & 1\end{bmatrix}\left(\begin{bmatrix}1\\ 1\end{bmatrix}10.05 + \begin{bmatrix}-10.1\\ -9.8\end{bmatrix}\right) \\
&= 10.05 - \frac{1}{4}\begin{bmatrix}1 & 1\end{bmatrix}\begin{pmatrix}10.05 & -10.1\\ 10.05 & -9.8\end{pmatrix} \\
&= 10.05 - \frac{1}{4}\begin{bmatrix}1 & 1\end{bmatrix}\begin{pmatrix}-0.05\\ 0.25\end{pmatrix} \\
&= 10.05 - \frac{1}{4}\left[0.2\right] = 10\quad \text{m}
\end{aligned}
$$

d)

for a) $C_\delta = N^{-1} = ¼$ cm^2

for b) $C_\delta = (N_1 + N_2)^{-1} = ¼$ cm^2

for c) $C_{\delta(+)} = C_{\delta(-)} - KA_2\, C_{\delta(-)}$

$$= \frac{1}{2} - \frac{1}{4}\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1 \\ 1\end{bmatrix} \cdot \frac{1}{2}$$

$$= \frac{1}{2} - \frac{1}{4} \cdot 2 \cdot \frac{1}{2} = \frac{1}{4} \quad \text{cm}^2$$

e) The final results should be all the same since all methods use the same information

8.5. Summation of Normals and Sequential LS for the Implicit Models:

The combinations of models discussed so far have mainly been derived for the parametric model.

The derivation of the summation of normals and sequential LS for the implicit model follows the same scheme as the parametric one.

PARAMETRIC	IMPLICIT
$\delta = -(N_1 + N_2)^{-1}(u_1 + u_2)$ $N_i = A_i^T P_i A_i$ $u_i = A_i^T P_i w_i$ $C_\delta = C_{\hat{x}} = (N_1 + N_2)^{-1}$	$\delta = -(N_1' + N_2')^{-1}\left(u_1' + u_2'\right)$ $N_i' = A_i^T \left(B_i P_i^{-1} B_i^T\right)^{-1} A_i$ $u_i' = A_i^T \underbrace{\left(B_i P_i^{-1} B_i^T\right)^{-1}}_{M_i} w_i$ $P_i = \sigma_0^2 C_{l_i}^{-1} \; \& \; C_{l_i} = \sigma_0^2 P_i^{-1}$ $C_\delta = C_{\hat{x}} = \left(N_1' + N_2'\right)^{-1}$
$\delta(+) = \delta(-) - K(A_2\delta(-) + w_2)$ $\delta(-) = -\underbrace{\left(A_1^T P_1 A_1\right)^{-1}}_{N_1}\underbrace{\left(A_1^T P_1 w_1\right)}_{u_1}$ $K = N_1^{-1} A_2^T \left(C_{l_2} + A_2 N_1^{-1} A_2^T\right)^{-1}$	Change the following: P_1 by $\left(B_1 P_1^{-1} B_1^T\right)^{-1}$ or $\left(B_1 C_{l_1} B_1^T\right)^{-1}$ and $C_{l_2} \rightarrow B_2 C_{l_2} B_2^T$

8.6. Parameter Observations

- Parameter observations is a method which can be used in cases where a-priori information about the parameters is available.
- For example, station coordinates (or elevations) and their covariance matrix $\left(\hat{x} \text{ and } C_{\hat{x}}\right)$ may be available from a previous adjustment.
- In this case $\hat{x}$ can be considered as direct observations (with $C_{\hat{x}}$) along with an observation vector to better estimate $\hat{x}$.

Functional Model:

$$\hat{x}^{obs}_{u,1} = \hat{x}_{u,1}$$

Linearized Functional Model:

$$x^{obs} + \hat{v}_x = x^0 + \hat{\delta}$$

$$\hat{v}_x = \hat{\delta} + \left(x^0 - x^{obs}\right)$$

$$\hat{v}_{x_{u,1}} = \hat{\delta}_{u,1} + w_{x_{u,1}} \qquad \left(\hat{v} = A\hat{\delta} + w\right)$$

with the stochastic model $C_{x_{u,u}} = P^{-1}_{x_{u,u}}$

- Note: a variable (observation or parameter) that has infinite variance, $\sigma^2 \rightarrow \infty$, has a corresponding weight of $P = 1/\sigma^2 = 0$. In this case, the variable becomes an unknown parameter.
- A variable with zero variance $\sigma^2 = 0$, has infinite weight, $P \rightarrow \infty$, and therefore is regarded as a constant.
- In between these two extreme cases there are an infinite number of possibilities for weighting parameters.

Functional Models (Linearized)

$$A_{n,u}\hat{\delta}_{u,1} + w_{n,1} = \hat{v}_{n,1}$$

$$I_{u,u}\hat{\delta}_{u,1} + w_{x_{u,1}} = \hat{v}_{u,1}$$

with the stochastic models:

$$C_l = P^{-1} \qquad \left(P = C_l^{-1}\right)$$

$$C_x = P_x^{-1} \qquad \left(P_x = C_x^{-1}\right)$$

Variation function:

$$\varphi = \hat{v}^T P \hat{v} + \hat{v}_x^T P_x \hat{v}_x \qquad = \min$$

$$= \left(\hat{\delta}^T A^T + w^T\right) P \left(A\hat{\delta} + w\right) + \left(\hat{\delta}^T + w_x^T\right) P_x \left(\hat{\delta} + w_x\right) \qquad = \min$$

$$\begin{aligned} &= \hat{\delta}^T A^T P A \hat{\delta} + \hat{\delta}^T A^T P w + w^T P A \hat{\delta} + w^T P w \\ &+ \hat{\delta}^T P_x \hat{\delta} + \hat{\delta}^T P_x w_x + w_x^T P_x \hat{\delta} + w_x^T P_x w_x \qquad = \min \end{aligned}$$

$$\begin{aligned} &= \hat{\delta}^T A^T P A \hat{\delta} + 2 \hat{\delta}^T A^T P w + w^T P w \\ &+ \hat{\delta}^T P_x \hat{\delta} + 2 \hat{\delta}^T P_x w_x + w_x^T P_x w_x \qquad = \min \end{aligned}$$

Minimize φ

$$\frac{\partial \varphi}{\partial \delta} = 2\hat{\delta}^T A^T P A + 2 w^T P A + 2\hat{\delta}^T P_x + 2 w_x^T P_x = 0$$

Transpose and divide by 2

$$A^T P A \hat{\delta} + P_x \hat{\delta} + A^T P w + P_x w_x = 0$$

$$\left(A^T P A + P_x\right)\hat{\delta} + \left(A^T P w + P_x w_x\right) = 0$$

$$\boxed{\hat{\delta}_{u,1} = -\left(A_{u,n}^T P_{n,n} A_{n,u} + P_{x_{u,u}}\right)^{-1} \left(A_{u,n}^T P_{n,n} w_{n,1} + P_{x_{u,u}} w_{x_{u,1}}\right)}$$

Variance-Covariance Matrix of Estimated Correction C_δ:

Functional model

$$\hat{\delta} = -\left(A^T P A + P_x\right)^{-1}\left(A^T P\left(f\left(x^0\right) - l^{obs}\right) + P_x\left(x^0 - x^{obs}\right)\right)$$

$$\therefore C_\delta = \left(\frac{\partial \delta}{\partial l}\right) C_l \left(\frac{\partial \delta}{\partial l}\right)^T + \left(\frac{\partial \delta}{\partial x^{obs}}\right) C_x \left(\frac{\partial \delta}{\partial x^{obs}}\right)^T$$

$$C_\delta = \left(\underbrace{A^T P A}_{N} + P_x\right)^{-1}$$

Analysis of the equation

$$\hat{\delta} = -\left(A^T C_l^{-1} A + C_x^{-1}\right)^{-1}\left(A^T C_l^{-1} w + C_x^{-1} w_x\right)$$

- If the observed parameters are highly precise, i.e. $\mathbf{C_x}$ has small elements, then $\mathbf{C_x^{-1}}$ will have large elements, and therefore in $\left(\mathbf{A^T C_l^{-1} A} + \mathbf{C_x^{-1}}\right)$ its contribution will be large. Finally, the values in $\hat{\delta}$ will be smaller than if it was calculated without the additional parameter observations.

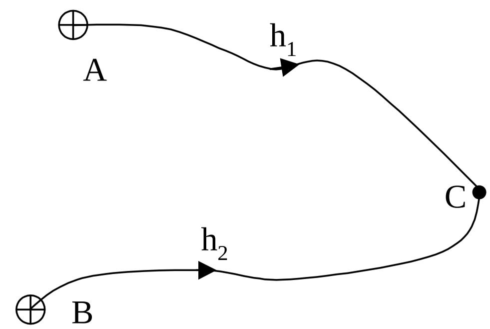
- If the parameter observations are not precise, $\mathbf{C_x}$ will have large elements, therefore $\mathbf{C_x^{-1}}$ will have small elements. Thus, the parameter observations will have little contribution to the solution vector.

Example of a LSA with parameter observations (from a previous final exam):

$$l = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} 1.74 \\ 2.76 \end{bmatrix} \text{ m}$$

and $\begin{bmatrix} H_A \\ H_B \end{bmatrix} = \begin{bmatrix} 5.0 \\ 4.0 \end{bmatrix}$ m

$$C_l = 10^{-4} I_{2,2} \quad \text{m}^2$$



Required: H_c (i.e. x = [H_C])

1) Using the parametric model: $l = f(x)$

$h_1 = H_C - H_A$ $\rightarrow H_C^0 = H_A + h_1 = 5 + 1.74 = 6.74$ m

$h_2 = H_C - H_B$

2) Linearized model: $\therefore A\delta + w = v$, where $A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$w = f(x^0) - l = \begin{bmatrix} H_C^0 & -H_A & -h_1 \\ H_C^0 & -H_B & -h_2 \end{bmatrix} \text{ m}$$

$$= \begin{bmatrix} 6.74 & -5 & -1.74 \\ 6.74 & -4 & -2.76 \end{bmatrix} = \begin{bmatrix} 0 \\ -0.02 \end{bmatrix} \text{ m}$$

3) Estimate the solution vector

$$\hat{\delta} = -(A^T P A)^{-1} A^T P w$$

$$P = C_l^{-1} = 10^4 I_{2,2} \quad \text{m}^{-2}$$

$$A^T P A = \begin{bmatrix} 1 & 1 \end{bmatrix} 10^4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2 \times 10^4 \quad \text{m}^{-2}$$

$$(A^T P A)^{-1} = (10^4 \times 2)^{-1} = 0.5 \times 10^{-4} \quad \text{m}^2$$

$$A^T P w = \begin{bmatrix} 1 & 1 \end{bmatrix} 10^4 \begin{bmatrix} 0 \\ -0.02 \end{bmatrix}$$
$$= -0.02 \times 10^4 \quad m^{-1}$$

$$\therefore \hat{\delta} = -\left[A^T P A\right]^{-1} A^T P w$$

$$= -\left[0.5 \times 10^{-4}\right] \left(-0.02 \times 10^4\right) = 0.01 \quad m$$

$$\therefore \hat{x} = x^0 + \hat{\delta} = 6.74 + 0.01 = 6.75 \quad m$$

Assume further that an a-priori value of the elevation of point C is 6.70 m with variance 0.01 m^2. Compute the elevation of point (C) considering this extra information.

This is a problem of parameter observations because we have new information about the parameter we are trying to estimate. Therefore, we have to use the equations for the parametric model with parameter observations, that is

$$\hat{\delta}' = -\left(A^T P A + P_X\right)^{-1} \left(A^T P w + P_X w_X\right)$$

where A^T, P, and w have been defined before, and

$$P_X = C_X^{-1} = \frac{1}{0.01} = 100 \quad m^{-2}$$
$$w_X = x^0 - x^{obs} = 6.74 - 6.70 = 0.04 \quad m$$

$$\left(A^T P A + P_X\right) = 2 \times 10^4 + 100 = 20100 = 2.01x10^4 \quad m^{-2}$$

$$\left(A^T P A + P_x\right)^{-1} = \frac{1}{2.01} \times 10^{-4} \quad m^2$$

$$\left(A^T P w + P_X w_X\right) = \left(-0.02 \times 10^4\right) + 100(0.04) = -196$$

$$= -0.0196 \times 10^4 \quad m^{-1}$$

$$\delta' = -\left(A^T P A + P_X\right)^{-1}\left(A^T P w + P_X w_X\right)$$

$$= -\left(\frac{1}{2.01} \text{ x } 10^{-4}\right)\left(-0.0196 \text{ x } 10^4\right) = 0.00975 \quad m$$

$$\therefore \hat{x}' = x^0 + \hat{\delta}' = 6.74 + 0.00975 = 6.74975 \approx 6.750 \quad m$$

Recall the variance-covariance matrix of the adjusted parameters (again, ignore the a-posteriori variance factor for the time being):

$$C_{\hat{x}} = (N)^{-1} = \left(A^T P A\right)^{-1}$$

$$= 0.5 \times 10^{-4} \quad m^2$$

$$C_{\hat{x}'} = \left(A^T P A + P_X\right)^{-1}$$

$$= \frac{1}{2.01} \times 10^{-4} \quad m^2$$

$$= 0.4975 \times 10^{-4} = 0.498 \times 10^{-4} \;\; m^2$$

Note that having extra information about the elevation of point C increases the precision of the final estimate.