6. Conditional Least Squares Adjustment

- The general form of the conditional mathematical model was given before as:

$$\mathbf{f}_{r,1}\left(\hat{\mathbf{l}}_{n,1}\right)=\mathbf{0}_{r,1} \text{ where } \hat{\mathrm{l}}=\mathrm{l}+\hat{\mathrm{v}}$$

$\hat{\mathbf{l}}$ is the vector of adjusted observations
$\hat{\mathbf{v}}$ is the vector of adjusted residuals
r (degrees of freedom) = n –u

- Two basic properties must be satisfied for the conditional model:
  1) Number of equations = number of degrees of freedoms. This means that each redundant observation provides one independent condition equation.
  2) The equations describe the functional relationship among the observations only. This obviously indicates that the unknown parameters **x** will not be among the direct output of the conditional adjustment. Thus the adjusted parameters $\hat{\mathbf{x}}$ and their covariance matrix $\mathbf{C}_{\hat{\mathbf{x}}}$, have to be computed after the adjustment using the direct model ($\hat{\mathbf{x}}=\mathbf{f}\left(\hat{\mathbf{l}}\right)$) and the law of propagation of covariances (this is a disadvantage when comparing the conditional and the parametric adjustment).

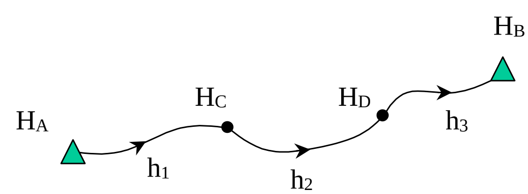
6.1. Examples of the Conditional Math Model:

1. Levelling networks – two types of conditions:

i) Levelling line with fixed end-points

$x_{1,2}^{T}=\left[H_C \;\; H_D\right] \quad L_{1,3}^{T}=\left[h_1 \;\; h_2 \;\; h_3\right] \qquad r=1$

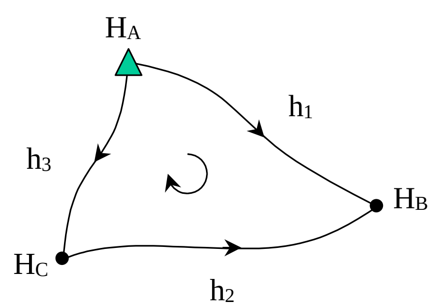
$$h_1+h_2+h_3+\left(H_A-H_B\right)=0$$



ii) Closed loop

$x_{1,2}^{T}=\left[H_B \;\; H_C\right] \; L_{1,3}^{T}=\left[h_1 \;\; h_2 \;\; h_3\right]$
$r=1$

$$h_1-h_2-h_3=0$$

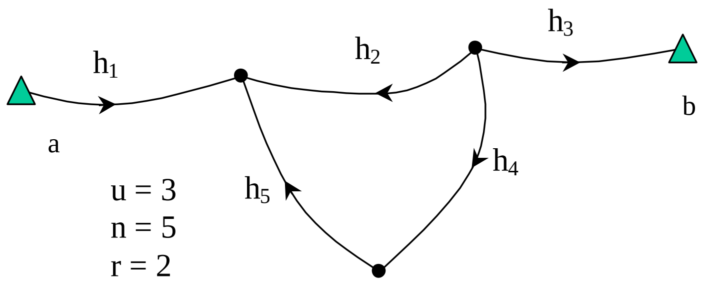


- How to choose the condition equations:

  i) Make sure that each observation appears at least once in the condition equations.

ii) Ensure that the equations are ***linearly independent***. If you ensure that the number of equations = *r*, this by default will give you linearly independent equations.

- Example:



$$\begin{array}{r}row(1):\\row(2):\\row(3):\end{array}\begin{bmatrix}1 & -1 & 1 & 0 & 0\\0 & -1 & 0 & 1 & 1\\1 & 0 & 1 & -1 & -1\end{bmatrix}\begin{bmatrix}h_1\\h_2\\h_3\\h_4\\h_5\end{bmatrix}+\begin{bmatrix}H_a-H_b\\0\\H_a-H_b\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix}$$

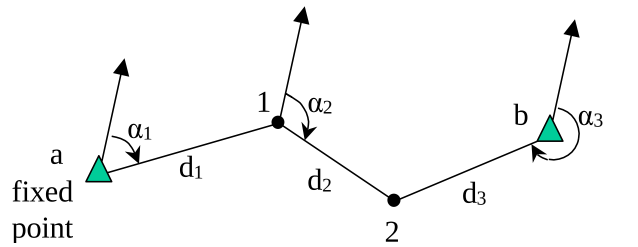
- Note:

  Row (3) = Row (1) – Row (2)

  (i.e. Row (3) is linearly dependent on Row (1) and Row (2))

2. Traverse networks

i) Traverse connecting two fixed points



$$\mathbf{x}=\begin{bmatrix}x_1\\y_1\\x_2\\y_2\end{bmatrix}\quad \mathbf{l}=\begin{bmatrix}\alpha_1\\d_1\\\alpha_2\\d_2\\\alpha_3\\d_3\end{bmatrix}$$

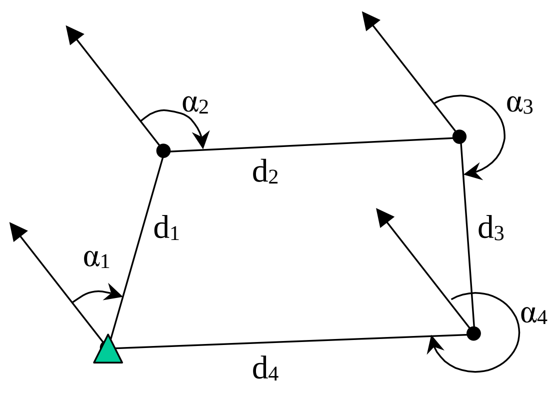
$$\sum_{i=1}^{3}\Delta x_i-\left(x_b-x_a\right)=0 \quad \text{where } \Delta x_i=d_i\sin\alpha_i$$

$$\sum_{i=1}^{3}\Delta y_i-\left(y_b-y_a\right)=0 \qquad \Delta y_i=d_i\cos\alpha_i$$

u = 4, n = 6

$r=6-4=2$

ii) Closed traverse



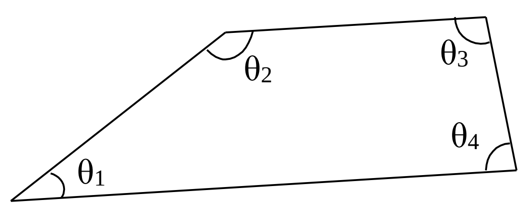
n = 8 u = 6

r = 8 – 6 = 2

$$\sum\Delta x_i=0$$

$$\sum\Delta y_i=0$$

iii) Closed traverse with observed internal angles will add an additional condition



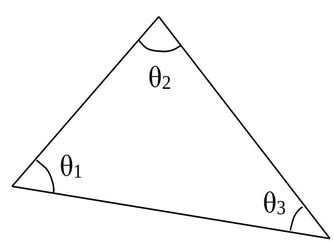
$$\sum_{i=1}^{4}\theta_i-K=0$$

where $K=(2S-4)\cdot 90^0$

and S = no of stations

e.g. if s = 4 → K = 360^0

e.g. triangle $\theta_1+\theta_2+\theta_3-180=0$



6.2. Linearized (or reduced) form of the condition equation:

- As stated before, the least squares adjustment requires a linear model.
- As in the parametric case, Taylor series expansion is used.
- The point of expansion (POE) here is defined by the values of the observation vector.
- The linearized functional model

$$\mathbf{f}_{\mathbf{r,1}}\left(\hat{\mathbf{l}}_{\mathbf{n,1}}\right) = 0_{\mathbf{r,1}}$$

$$POE \quad :\mathbf{l}^{\mathbf{obs}}$$

$$\mathbf{f}\left(\mathbf{l}^{\mathbf{obs}}\right) + \frac{\partial \mathbf{f}}{\partial \, \mathrm{l}}\hat{\mathbf{v}} = \mathbf{0}$$

r equations in n unknowns (r<n)

$$\mathbf{w}_{\mathbf{r,1}} \quad + \quad \mathbf{B}_{\mathbf{r,n}}\hat{\mathbf{v}}_{\mathbf{n,1}} = \mathbf{0}_{\mathbf{r,1}}$$

  - where $\mathbf{B} = \frac{\partial \mathbf{f}}{\partial \mathbf{l}}$ and $\mathbf{w} = \mathbf{f}(\mathbf{l}^{\mathbf{obs}})$

- Stochastic model:

$$\underset{\mathbf{n,n}}{\mathbf{C}_{\mathbf{l}}} \quad \propto \textbf{ or } = \quad \underset{\mathbf{n,n}}{\mathbf{P}^{-1}}$$

- Notes:
  - After adjustment, the misclosure vector (w) must become zero.
  - Both matrices B and w are numerically known, the only unknown here is the residual (**v**) vector.
- Variation function ϕ in matrix form:

$$\boldsymbol{\phi} = \hat{\boldsymbol{v}}^{\mathbf{T}}\, \mathbf{P}\, \hat{\boldsymbol{v}} = \mathbf{min}$$

- In the parametric model derivation, an explicit substitution was made for **v**. This cannot be done here since **v** is not isolated in the functional model.

- Could we do the following in order to come up with an expression for **(v)**?
  1. Starting with $\mathbf{B}\ \hat{\mathbf{v}}\ + \mathbf{w} = \mathbf{0}$
     - pre-multiply by $\mathbf{B}^{-1}$

$$\mathrm{B}^{-1}\,\mathrm{B}\ \hat{\mathbf{v}}\ = -\ \mathrm{B}^{-1}\,\mathrm{w}$$

$$\mathrm{I}\ \hat{\mathbf{v}}\ = -\ \mathrm{B}^{-1}\,\mathrm{w}$$

- This is not possible since **B** is not square matrix ($r < n$ always)

2. Starting with $\mathbf{B}\,\hat{\mathbf{v}} + \mathbf{w} = \mathbf{0}$
   - pre-multiply by $\mathbf{B}^{\mathbf{T}}$

     $B^T B\, \hat{\mathbf{v}} = - B^T w$

     $\hat{\mathbf{v}} = - (B^T B)^{-1} B^T w$

   - To examine the validity of this possibility, look at the rank and dimension of **B** and $\mathbf{B}^{\mathbf{T}}\,\mathbf{B}$.
   - Provided that all the condition equations are independent, rank (**B**) = r (its smallest dimension).
   - Since rank ($\mathbf{B}^{\mathbf{T}}\,\mathbf{B}$) = rank (**B**) = $r < n$ and dimension of ($\mathbf{B}^{\mathbf{T}}\,\mathbf{B}$) = n, n
   - Therefore, $\mathbf{B}^{\mathbf{T}}\,\mathbf{B}$ is singular (i.e., $(\mathbf{B}^{\mathbf{T}}\,\mathbf{B})^{-1}$ does not exist)

- The way to minimize ϕ is to use the Lagrange multiplier technique.
- Lagrange multipliers allow the minimization of $\phi = \hat{\mathbf{v}}^{\mathbf{T}}\,\mathbf{P}\,\hat{\mathbf{v}}$ subject to the condition $\mathbf{B}\,\hat{\mathbf{v}} + \mathbf{w} = \mathbf{0}$.
- The variation function is:

$$\varphi = \hat{\mathbf{v}}^{\mathbf{T}}_{\mathbf{1,n}}\mathbf{P}_{\mathbf{n,n}}\hat{\mathbf{v}}_{\mathbf{n,1}} + 2\hat{\mathbf{k}}^{\mathbf{T}}_{\mathbf{1,r}}\left(\mathbf{B}_{\mathbf{r,n}}\hat{\mathbf{v}}_{\mathbf{n,1}} + \mathbf{w}_{\mathbf{r,l}}\right) = \min$$

  - Where k is the vector of Lagrange multipliers (one-multiplier/condition) (The factor 2 is introduced for convenience only)
  - Note: the second term equals zero ($\mathbf{B}\,\hat{\mathbf{v}} + \mathbf{w} = \mathbf{0}$) and therefore does not change the value of ϕ

- To minimise ϕ, differentiate with respect to $\hat{\mathbf{v}}$ and $\hat{\mathbf{k}}$

$$\phi = \hat{\mathbf{v}}^{\mathbf{T}}\,\mathbf{P}\,\hat{\mathbf{v}} + 2\hat{\mathbf{k}}^{\mathbf{T}}\,(\mathbf{B}\,\hat{\mathbf{v}} + \mathbf{w}) = \min$$

$$\phi = \hat{\mathbf{v}}^{\mathbf{T}}\,\mathbf{P}\,\hat{\mathbf{v}} + 2\hat{\mathbf{k}}^{\mathbf{T}}\,\mathbf{B}\,\hat{\mathbf{v}} + 2\hat{\mathbf{k}}^{\mathbf{T}}\,\mathbf{w} = \min$$

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

$$\frac{\partial \varphi}{\partial \mathbf{k}} = \mathbf{2}\hat{\mathbf{v}}^{\mathbf{T}}\mathbf{B}^{T} + \mathbf{2}\mathbf{w}^{\mathbf{T}} = \mathbf{0}$$

- Transpose each and divide by 2

$$\mathbf{P}_{\mathbf{n,n}}\hat{\mathbf{v}}_{\mathbf{n,1}} + \mathbf{B}^{\mathbf{T}}_{\mathbf{n,r}}\hat{\mathbf{k}}_{\mathbf{r,1}} = \mathbf{0} \qquad \text{– } n \text{ equations in } (n + r) \text{ unknowns}$$

$$\mathbf{B}_{r,n}\hat{\mathbf{v}}_{\mathbf{n,1}} + \mathbf{w}_{\mathbf{r,1}} = \mathbf{0} \qquad \text{– } r \text{ equations in } (n) \text{ unknowns}$$

- Express in hyper-matrix form

$$\begin{pmatrix} \mathbf{P} & \mathbf{B}^{\mathbf{T}} \\ \mathbf{B} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{v}} \\ \hat{\mathbf{k}} \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \mathbf{w} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}$$

- Solution possibilities to the hyper-matrix system:
  1. Invert hyper-matrix

$$\begin{pmatrix} \hat{\mathbf{v}} \\ \hat{\mathbf{k}} \end{pmatrix} = -\begin{pmatrix} \mathbf{P} & \mathbf{B}^{\mathbf{T}} \\ \mathbf{B} & \mathbf{0} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0} \\ \mathbf{w} \end{pmatrix}$$

     - This solution requires inversion of a large matrix
  2. Elimination
     - For the hyper-matrix system

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} + \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}$$

$$\mathbf{A\,x + B\,y + u = 0} \qquad \textbf{(i)}$$

$$\mathbf{C\,x + D\,y + v = 0} \qquad \textbf{(ii)}$$

     - Eliminate x from (i)

$$\mathbf{x = -\,A^{-1}\,(B\,y + u)} \ \textit{(provided } A^{-1} \textit{ exists!)} \qquad \textbf{(iii)}$$

     - Now substitute (iii) into (ii)

$$\mathbf{-\,C\,A^{-1}\,(B\,y + u) + D\,y + v = 0}$$

$$\mathbf{-\,C\,A^{-1}\,B\,y - C\,A^{-1}\,u + D\,y + v = 0}$$

$$\mathbf{(D - C\,A^{-1}\,B)\,y + (v - C\,A^{-1}\,u) = 0}$$

     - In this case

$$\mathbf{A = P,\ B = B^T,\ C = B,\ D = 0,\ x = \hat{v},\ y = \hat{k},\ u = 0,\ v = w}$$

     - substitute into

$$\mathbf{(D - C\,A^{-1}\,B)\,y + (v - C\,A^{-1}\,u) = 0}$$

$$\mathbf{(0 - B\,P^{-1}\,B^T)\,\hat{k} + w = 0}$$

$$\mathbf{(B\,P^{-1}\,B^T)\,\hat{k} = w}$$

$$\boxed{\therefore\ \hat{\boldsymbol{k}} = \mathbf{(B\,P^{-1}\,B^T)^{-1}\,w = (B\,C_l\,B^T)^{-1}\,w}}$$

Note: Since the P matrix is not included in the computation of ***k*** *or the* ***v*** *vectors, it is usually the case to use* $\sigma_o^2 = 1$

     - Is this a valid solution? Yes

$$rank\ (B) = r \quad rank\ (P) = n > r \quad rank\ (\mathbf{BP^{-1}B^T})^{-1} = r$$

- To solve for $\hat{\mathbf{v}}$, substitute $\hat{\mathbf{k}}$ into

  $\mathbf{P}\ \hat{\mathbf{v}} + \mathbf{B}^{\mathbf{T}}\ \hat{\mathbf{k}} = \mathbf{0}$ (first group of hyper-matrix)

  $$\boxed{\hat{\mathbf{v}} = -\mathbf{P}^{-1}\mathbf{B}^{\mathbf{T}}\,\hat{\boldsymbol{k}} = -\mathbf{C}_{\mathbf{l}}\,\mathbf{B}^{\mathbf{T}}\,(\mathbf{B}\,\mathbf{C}_{\mathbf{l}}\,\mathbf{B}^{\mathbf{T}})^{-1}\,\mathbf{w}}$$

♦ Adjusted observation

$$\hat{\mathbf{l}} = \mathbf{l} + \hat{\mathbf{v}}$$

♦ Estimated variance factor

$$\hat{\sigma}_{\mathbf{0}}^{\mathbf{2}} = \frac{\hat{\mathbf{v}}^{\mathbf{T}}\mathbf{P}\hat{\mathbf{v}}}{\mathbf{r}}$$

♦ Covariance matrix for the adjusted residuals: $\mathbf{C}_{\hat{\mathbf{v}}}$

- Functional model

$$\begin{aligned}\hat{v} &= -\ C_l B^T (B C_l B^T)^{-1} w \\ &= -\ C_l B^T (B C_l B^T)^{-1} f(l) = -C_l B^T k f(l)\end{aligned}$$

- Apply the law of variance-covariance propagation

$$\mathbf{C}_{\hat{\mathbf{v}}} = \left(\frac{\partial \hat{\mathbf{v}}}{\partial \mathbf{l}}\right)\mathbf{C}_{\mathbf{l}}\left(\frac{\partial \hat{\mathbf{v}}}{\partial \mathbf{l}}\right)^{\mathbf{T}}$$

$$\frac{\partial \hat{v}}{\partial l} = -\ \frac{\partial (C_l B^T k f(l))}{\partial l} = -C_l B^T k \frac{\partial (f(l))}{\partial l} = -C_l B^T k B = -C_l B^T (B C_l B^T)^{-1} B$$

$$\begin{aligned}\mathbf{C}_{\hat{\mathbf{v}}} &= [\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}]\mathbf{C}_{\mathbf{l}}[\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}]^{\mathbf{T}} \\ &= \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}\mathbf{C}_{\mathbf{l}}\end{aligned}$$

$$\boxed{\therefore \mathbf{C}_{\hat{\mathbf{v}}} = \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}\mathbf{C}_{\mathbf{l}}}$$

♦ Covariance matrix for the adjusted observations: $\mathbf{C}_{\hat{\mathbf{l}}}$

- Functional model

$$\begin{aligned}\hat{\mathbf{l}} &= \mathbf{l} + \hat{\mathbf{v}} \\ &= \mathbf{l} - \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{w} \\ &= \mathbf{l} - \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{f}(\mathbf{l})\end{aligned}$$

- Covariance propagation

$$\mathbf{C}_{\hat{\mathbf{l}}} = \left(\frac{\partial \hat{\mathbf{l}}}{\partial \mathbf{l}}\right)\mathbf{C}_{\mathbf{l}}\left(\frac{\partial \hat{\mathbf{l}}}{\partial \mathbf{l}}\right)^{\mathbf{T}}$$

$$\frac{\partial \hat{\mathbf{l}}}{\partial \mathbf{l}} = \mathbf{I} - \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}$$

$$\begin{aligned}
\mathbf{C}_{\hat{\mathbf{l}}} &= [\mathbf{I} - \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}]\mathbf{C}_{\mathbf{l}}[\mathbf{I} - \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}]^{\mathbf{T}} \\
&= [\mathbf{C}_{\mathbf{l}} - \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}\mathbf{C}_{\mathbf{l}}][\mathbf{I} - \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}]^{\mathbf{T}} \\
&= \mathbf{C}_{\mathbf{l}} - \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}\mathbf{C}_{\mathbf{l}} - \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}\mathbf{C}_{\mathbf{l}} \\
&\quad + \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}\mathbf{C}_{\mathbf{l}}
\end{aligned}$$

$$\boxed{\therefore \mathbf{C}_{\hat{\mathbf{l}}} = \mathbf{C}_{\mathbf{l}} - \mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}}(\mathbf{B}\mathbf{C}_{\mathbf{l}}\mathbf{B}^{\mathbf{T}})^{-1}\mathbf{B}\mathbf{C}_{\mathbf{l}}}$$

$$\therefore \mathbf{C}_{\hat{\mathbf{l}}} = \mathbf{C}_{\mathbf{l}} - \mathbf{C}_{\hat{\mathbf{v}}}$$

6.3. Iterative solution of the conditional model

- Linearized model

  $\mathbf{B}\,\hat{\boldsymbol{v}} + \mathbf{w} = \mathbf{0}$

- Initial computation

$$\mathbf{B}_{(1)} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{l}}\right|_{\mathbf{l}^{\mathbf{obs}}}$$

$$\mathbf{w}_{(1)} = \mathbf{f}(\mathbf{l}^{\mathbf{obs}})$$

- Iteration (1):
  - Solve for $\hat{\mathbf{v}}_{(1)}$,
  - Correct observations $\hat{\mathrm{l}}_{(1)} = \mathrm{l}^{\mathrm{obs}} + \hat{\mathrm{v}}_{(1)}$
- Iteration (2): needed if the model is non-linear

$\mathbf{B}_{(2)} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{l}}\right|_{\hat{\mathbf{l}}_{(1)}}$, in general:

$$\boxed{\mathbf{B}_{(i+1)} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{l}}\right|_{\hat{\mathbf{l}}_{(i)}}}$$

$$\mathbf{w}_{(2)} = \mathbf{f}(\hat{\mathbf{l}}_{(1)}) + \mathbf{B}_{(2)}(\mathbf{l}^{obs} - \hat{\mathbf{l}}_{(1)})$$, in general

$$\boxed{\mathbf{w}_{(i+1)} = \mathbf{f}(\hat{\mathbf{l}}_{(i)}) + \mathbf{B}_{(i+1)}(\mathbf{l}^{obs} - \hat{\mathbf{l}}_{(i)})}$$

- Solve for $\hat{\mathbf{v}}_{(2)}$ where

$$\boxed{\hat{\mathbf{v}}_{(i+1)} = -\mathbf{C}_{\mathbf{l}}\,\mathbf{B}_{(i+1)}^{\mathbf{T}}\,(\mathbf{B}_{(i+1)}\,\mathbf{C}_{\mathbf{l}}\,\mathbf{B}_{(i+1)}^{\mathbf{T}})^{-1}\,\mathbf{w}_{(i+1)}}$$

- Correct observations $\hat{\mathbf{l}}_{(2)} = \mathbf{l}^{obs} + \hat{\mathbf{v}}_{(2)}$

- Iterate until

$$\hat{\mathbf{v}}_{(i+1)} - \hat{\mathbf{v}}_{(i)} \rightarrow \mathbf{0}$$ (for all elements of v)

- Calculate

$$\hat{l}_{(i+1)} = l^{obs} + \hat{v}_{(i+1)}$$

6.4. Direct model solution

- The conditional model adjustment will give the residual vector ($\hat{\mathbf{v}}$) and the adjusted observations along with their respective covariance matrices.
- The direct model, then, uses the adjusted observations from the conditional model solution to calculate the parameters (if any)
  - Functional model

$$\hat{\mathbf{x}} = \mathbf{g}(\hat{\mathbf{l}})$$

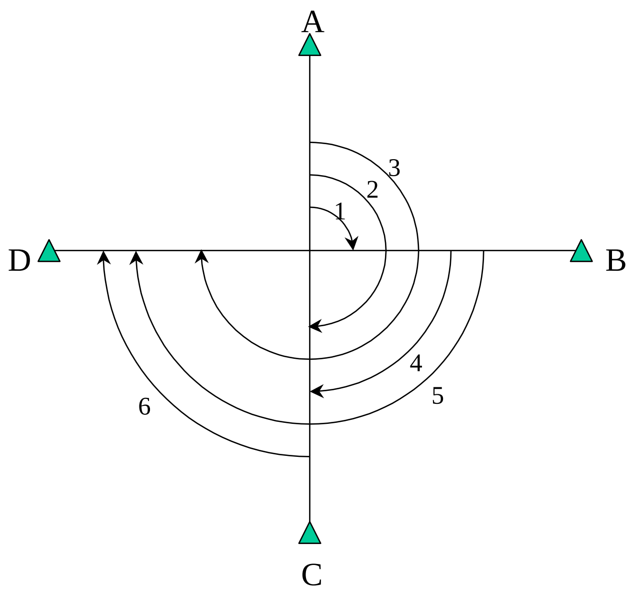
  - Variance propagation

$$\mathbf{C}_{\hat{\mathbf{x}}} = \left(\frac{\partial \hat{\mathbf{x}}}{\partial \hat{\mathbf{l}}}\right) \mathbf{C}_{\hat{\mathbf{l}}} \left(\frac{\partial \hat{\mathbf{x}}}{\partial \hat{\mathbf{l}}}\right)^{\mathbf{T}}$$

6.5. Example on conditional least squares

- Consider the case of station adjustment in triangulation network in which

$$\mathbf{x}_{3,1} = \begin{bmatrix} \alpha_1 \\ \alpha_4 \\ \alpha_6 \end{bmatrix} \quad \mathbf{l}_{6,1} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \end{bmatrix}$$



$n = 6, \; u = 3$

$r = 6\text{-}3 = 3$

Angle	Observed value
α_1	89° 59′ 58.3″
α_2	180° 00′ 01.4″
α_3	270° 00′ 00.2″
α_4	89° 59′ 59.8″
α_5	179° 59′ 57.0″
α_6	90° 00′ 03.1″

All the six angles have the same standard deviation σ = 1", that is $\boldsymbol{C_l} = I$ arcsec^2

Note: How many solutions can you have for angle a1?
a1 = a1
a1 = a2 - a4
a1 = a3 - a5
a1 = a3 - a4 - a6
a1 = a5 - a4
a1 = a6
and so on
But most likely none of them will be the same

- Solution

1. The three independent condition equations: $\mathbf{f}(\hat{\mathbf{l}}) = \mathbf{0}$

$$(\hat{\alpha}_1 + \hat{\alpha}_4) - \hat{\alpha}_2 = 0$$

$$(\hat{\alpha}_1 + \hat{\alpha}_5) - \hat{\alpha}_3 = 0$$

$$(\hat{\alpha}_2 + \hat{\alpha}_6) - \hat{\alpha}_3 = 0$$

Note: we wrote the condition equations in terms of the adjusted observations.

2. The linearized condition equations: $\mathbf{B_{3,6}\hat{v}_{6,1} + w_{3,1} = 0}$

The B matrix is: $\mathbf{B_{3,6}} = \frac{\partial \mathbf{f(l)}}{\partial \mathbf{l}} = \begin{bmatrix} 1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{bmatrix}$ ***unitless***

Note: Since the observations are of the same physical quantities as the unknowns, the B matrix is unitless. Also, the problem is linear and therefore only one iteration is required.

The misclosure vector $\mathbf{w}_{3,1}$ is computed as:

$$\mathbf{w_{3,1}} = \mathbf{f}(\mathbf{l}^{obs}) = \begin{bmatrix} (\alpha_1 + \alpha_4) - \alpha_2 \\ (\alpha_1 + \alpha_5) - \alpha_3 \\ (\alpha_2 + \alpha_6) - \alpha_3 \end{bmatrix} = \begin{bmatrix} -3.3 \\ -4.9 \\ 4.3 \end{bmatrix} \textbf{\textit{arcsec}}$$

Note that $\hat{\boldsymbol{v}}$ and $\mathbf{w}$ will have the same units since B is unitless.

3. Covariance matrix of the observations: $\mathbf{C_l}$

Since all angles are of equal precision and uncorrelated, C_l will be a unit matrix, i.e.: $\mathbf{C_l}$ = I (in arcsec^2) and $\mathbf{P} = \boldsymbol{\sigma}_0^2\mathbf{C}_l^{-1}$ = I (in arcsec^{-2})

$\boldsymbol{\sigma}_0^2$ = apriori variance factor which could be assumed =1.

4. The least-squares estimated values of the vector of residuals: $\hat{\mathbf{v}}$

$$\hat{v} = -\underbrace{C_l}_{\text{arcsec}^2} \underbrace{B^T}_{\text{unitless}} \underbrace{(\underbrace{BC_lB^T}_{\text{arcsec}^2})^{-1}}_{\frac{1}{\text{arcsec}^2}} \underbrace{w}_{\text{arcsec}} \qquad \textit{arcsec}$$

$$\mathbf{BC_lB^T} = \mathbf{BB^T} = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix} \qquad \text{(B } C_l \text{ B}^T \text{ should be symmetric)}$$

$$(\mathbf{BB^T})^{-1}\mathbf{w} = \begin{bmatrix} 0.65 \\ -2.7 \\ 2.55 \end{bmatrix}$$

$$\therefore \hat{v} = -C_l B^T (B C_l B^T)^{-1} w = \begin{bmatrix} 2.05 \\ -1.90 \\ -0.15 \\ -0.65 \\ 2.70 \\ -2.55 \end{bmatrix} \text{arcsec}$$

5. The adjusted observations: $\hat{\mathbf{l}}$

$$\hat{\mathbf{l}} = \mathbf{l} + \hat{\mathbf{v}} = \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\alpha}_3 \\ \hat{\alpha}_4 \\ \hat{\alpha}_5 \\ \hat{\alpha}_6 \end{bmatrix} = \begin{bmatrix} 90 & 0 & 0.35 \\ 179 & 59 & 59.50 \\ 270 & 0 & 0.05 \\ 89 & 59 & 59.15 \\ 179 & 59 & 59.70 \\ 90 & 0 & 0.55 \end{bmatrix} \text{D°M'S''} \qquad \text{check } \mathbf{f}(\hat{\mathbf{l}}) = 0$$

Why does a1 not equal a6? Is this acceptable?
The answer is yes, because the difference is less than 0.5 of the standard deviations of the measurements

6. The variance-covariance matrix of the residual vector: $\mathbf{C}_{\hat{v}}$

$$\mathbf{C}_{\hat{v}} = \mathbf{C}_l \mathbf{B}^T (\mathbf{B} \mathbf{C}_l \mathbf{B}^T)^{-1} \mathbf{B} \mathbf{C}_l$$

Recall: These matrices have been computed before in the computation of $\hat{\mathbf{v}}$

7. The variance-covariance matrix of the adjusted observations: $\mathbf{C}_{\hat{l}}$

$$\mathbf{C}_{\hat{l}} = \mathbf{C}_l - \mathbf{C}_{\hat{v}}$$

8. The unknown parameters $\hat{\mathbf{x}}$ and their variance-covariance matrix: $\mathbf{C}_{\hat{x}}$

   Since the unknown parameters are directly observed (i.e., they are among the $\mathbf{l}$ vector), we can simply extract the $\mathbf{C}_{\hat{x}}$ from the $\mathbf{C}_{\hat{l}}$ matrix. But the general methodology of doing that is through the following direct model: $\hat{x} = g(\hat{l})$ and by applying the variance-covariance law.

$$\mathbf{C}_{\hat{x}} = \mathbf{J}\mathbf{C}_{\hat{l}}\mathbf{J}^T \text{ where } \hat{\mathbf{x}} = \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_4 \\ \hat{\alpha}_6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\alpha}_3 \\ \hat{\alpha}_4 \\ \hat{\alpha}_5 \\ \hat{\alpha}_6 \end{bmatrix}$$

6.6. Summary of Parametric and Conditional Least Squares

Item	Parametric	Conditional	Comments
Mathematical model	$\hat{\mathbf{l}}_{n,1} = \mathbf{f}_{n,1}(\hat{\mathbf{x}}_{u,1})$ - No ambiguity since each observation provides an equation.	$\mathbf{f}_{r,1}(\hat{\mathbf{l}}_{n,1}) = \mathbf{0}$ $r = n - u$ - Some ambiguity may occur during the stage of selecting the condition equations.	If r < n, conditional is preferable.
Design matrices and linearization	$\hat{\boldsymbol{v}}_{n,1} = \mathbf{A}_{n,u}\,\hat{\boldsymbol{\delta}}_{u,1} + \mathbf{w}_{n,1}$ $\mathbf{A} = \left.\frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}}\right\|_{x=x^0}$ - Approximate values $\mathbf{x}^0$ are required $\mathbf{w} = \mathbf{f}(\mathbf{x}^o) - \mathbf{l}^{obs}$ - More computations	$\mathbf{B}_{r,n}\hat{\mathbf{v}}_{n,1} + \mathbf{w}_{r,1} = \mathbf{0}$ $\mathbf{B} = \left.\frac{\partial \mathbf{f}(\mathbf{l})}{\partial \mathbf{l}}\right\|_{\mathbf{l}}$ - No approximate values are needed for l - Misclosure **w = f(l)** is a direct calculation.	**(B, w)** takes less effort than **(A, w)**. Thus conditional is preferable.
Size of inversion	$\hat{\boldsymbol{\delta}} = -(\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}_{u,u}\mathbf{A}^T\mathbf{P}\mathbf{w}$ - size of inversion is (u×u)	$\hat{\mathbf{v}} = -\mathbf{C}_l\mathbf{B}^T(\mathbf{B}\mathbf{C}_l\mathbf{B}^T)^{-1}_{r,r}\mathbf{w}$ - size of inversion is (r×r)	If $u < r$, parametric is preferable. If $r < u$, conditional is preferable.
Final results	$\hat{\mathbf{x}} = \mathbf{x}^0 + \hat{\boldsymbol{\delta}}$ $\hat{\mathbf{l}} = \mathbf{l} + \hat{\mathbf{v}}$ $\mathbf{C}_{\hat{x}} = \hat{\sigma}_0^2(\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}$ $\mathbf{C}_{\hat{l}} = \mathbf{A}\mathbf{C}_{\hat{x}}\mathbf{A}^T$ where $\hat{\sigma}_0^2 = \frac{\hat{\mathbf{v}}^T\mathbf{P}\hat{\mathbf{v}}}{r}$ $\mathbf{C}_{\hat{v}} = \mathbf{C}_l - \mathbf{C}_{\hat{l}}$ (end of adjustment)	$\hat{\mathbf{l}} = \mathbf{l} + \hat{\mathbf{v}}$ $\mathbf{C}_{\hat{v}} = \mathbf{C}_l\mathbf{B}^T(\mathbf{B}\mathbf{C}_l\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{C}_l$ $\mathbf{C}_{\hat{l}} = \mathbf{C}_l - \mathbf{C}_{\hat{v}}$	Conditional results are incomplete
Additional computations	No additional computations are needed	$\hat{x} = g(\hat{l})$ $\mathbf{C}_{\hat{x}} = \mathbf{J}\mathbf{C}_{\hat{l}}\mathbf{J}^T$	More computations are necessary for the conditional least squares