2. THE GENERAL LEAST SQUARES ADJUSTMENT TECHNIQUE

A common treatment of the least squares technique of estimation starts with simple linear mathematical models having observations (or measurements) as explicit functions of parameters with non-linear models developed as extensions. This adjustment technique is generally described as adjustment of indirect observations (also called parametric least squares). Cases where the mathematical models contain only measurements are usually treated separately and this technique is often described as adjustment of observations only. Both techniques are of course particular cases of a general adjustment model, the solution of which is set out below. The general adjustment technique also assumes that the parameters, if any, can be treated as "observables" ie, they have an a priori covariance matrix. This concept allows the general technique to be adapted to sequential processing of data where parameters are updated by the addition of new observations.

In general, least squares solutions require iteration, since a non-linear model is assumed. The iterative process is explained below. In addition, a proper treatment of covariance propagation is presented and cofactor matrices given for all the computed and derived quantities in the adjustment process. Finally, the particular cases of the general least squares technique are described.

2.1. The General Least Squares Adjustment Model

Consider the following set of non-linear equations representing the mathematical model in an adjustment

$$F\left(\hat{\mathbf{l}}, \hat{\mathbf{x}}\right) = \mathbf{0} \tag{2.1}$$

where $\mathbf{l}$ is a vector of n observations and $\mathbf{x}$ is a vector of u parameters; $\hat{\mathbf{l}}$ and $\hat{\mathbf{x}}$ referring to estimates derived from the least squares process such that

$$\hat{\mathbf{l}} = \mathbf{l} + \mathbf{v} \tag{2.2}$$

$$\hat{\mathbf{x}} = \mathbf{x} + \delta\mathbf{x} \tag{2.3}$$

where $\mathbf{v}$ is a vector of residuals or small corrections and $\delta\mathbf{x}$ is a vector of small corrections.

As is usual, the independent observations $\mathbf{l}$ have an a priori diagonal cofactor matrix $\mathbf{Q}_{ll}$ containing estimates of the variances of the observations, and in this general adjustment, the parameters $\mathbf{x}$ are treated as "observables" with a full a priori cofactor matrix $\mathbf{Q}_{xx}$. The diagonal elements of $\mathbf{Q}_{xx}$ contain estimates of variances of the parameters and the off-diagonal elements contain estimates of the covariances between parameters. Cofactor matrices $\mathbf{Q}_{ll}$ and $\mathbf{Q}_{xx}$ are related to the covariance matrices Σ_{ll} and Σ_{xx} by the variance factor σ_0^2

$$\Sigma_{ll} = \sigma_0^2\, \mathbf{Q}_{ll} \tag{2.4}$$

$$\Sigma_{xx} = \sigma_0^2\, \mathbf{Q}_{xx} \tag{2.5}$$

Also, weight matrices $\mathbf{W}$ are useful and are defined, in general, as the inverse of the cofactor matrices

$$\mathbf{W} = \mathbf{Q}^{-1} \tag{2.6}$$

and covariance, cofactor and weight matrices are all symmetric, hence $\mathbf{Q}^T = \mathbf{Q}$ and $\mathbf{W}^T = \mathbf{W}$ where the superscript T denotes the transpose of the matrix.

Note also, that in this development where $\mathbf{Q}$ and $\mathbf{W}$ are written without subscripts they refer to the observations, i.e., $\mathbf{Q}_{ll} = \mathbf{Q}$ and $\mathbf{W}_{ll} = \mathbf{W}$

Linearizing (2.1) using Taylor's theorem and ignoring 2nd and higher order terms, gives

$$F\left(\hat{\mathbf{l}}, \hat{\mathbf{x}}\right) = F(\mathbf{l}, \mathbf{x}) + \left.\frac{\partial F}{\partial \hat{\mathbf{l}}}\right|_{l,x} \left(\hat{\mathbf{l}} - \mathbf{l}\right) + \left.\frac{\partial F}{\partial \hat{\mathbf{x}}}\right|_{l,x} \left(\hat{\mathbf{x}} - \mathbf{x}\right) = \mathbf{0} \tag{2.7}$$

and with $\mathbf{v} = \hat{\mathbf{l}} - \mathbf{l}$ and $\delta\mathbf{x} = \hat{\mathbf{x}} - \mathbf{x}$ from (2.2) and (2.3), we may write the linearized model in symbolic form as

$$\mathbf{A}\mathbf{v} + \mathbf{B}\,\delta\mathbf{x} = \mathbf{f} \tag{2.8}$$

Equation (2.8) represents a system of m equations that will be used to estimate the u parameters from n observations. It is assumed that this is a redundant system where

$$n \geq m \geq u \tag{2.9}$$

The redundancy or degrees of freedom is

$$r = m - u \tag{2.10}$$

In equation (2.8) the coefficient matrices **A** and **B** are design matrices containing partial derivatives of the function evaluated using the observations **l** and the "observed" parameters **x**.

$$\mathbf{A}_{m,n} = \left.\frac{\partial F}{\partial \hat{\mathbf{l}}}\right|_{l,x} \tag{2.11}$$

$$\mathbf{B}_{m,u} = \left.\frac{\partial F}{\partial \hat{\mathbf{x}}}\right|_{l,x} \tag{2.12}$$

The vector **f** contains m numeric terms calculated from the functional model using **l** and **x**.

$$\mathbf{f}_{m,1} = -\left\{F(\mathbf{l}, \mathbf{x})\right\} \tag{2.13}$$

2.2. The Least Squares Solution of the General Adjustment Model

The least squares solution of (2.8), ie, the solution which makes the sums of the squares of the weighted residuals a minimum, is obtained by minimising the scalar function φ

$$\varphi = \mathbf{v}^T\mathbf{W}\,\mathbf{v} + \delta\mathbf{x}^T\mathbf{W}_{xx}\delta\mathbf{x} - 2\mathbf{k}^T\left(\mathbf{A}\mathbf{v} + \mathbf{B}\delta\mathbf{x} - \mathbf{f}\right) \tag{2.14}$$

where **k** is a vector of m Lagrange multipliers. φ is a minimum when its derivatives with respect to **v** and $\delta\mathbf{x}$ are equated to zero, ie.

$$\frac{\partial\varphi}{\partial\mathbf{v}} = 2\mathbf{v}^T\mathbf{W} \qquad - 2\mathbf{k}^T\mathbf{A} = \mathbf{0}^T$$
$$\frac{\partial\varphi}{\partial\delta\mathbf{x}} = 2\delta\mathbf{x}^T\mathbf{W}_{xx} - 2\mathbf{k}^T\mathbf{B} = \mathbf{0}^T$$

These equations can be simplified by dividing both sides by two, transposing and changing signs to give

$$-\mathbf{W}\mathbf{v} + \mathbf{A}^T\mathbf{k} = \mathbf{0} \tag{2.15}$$

$$-\mathbf{W}_{xx}\delta\mathbf{x} + \mathbf{B}^T\mathbf{k} = \mathbf{0} \tag{2.16}$$

Equations (2.15) and (2.16) can be combined with (2.8) and arranged in matrix form as

$$\begin{bmatrix} -\mathbf{W} & \mathbf{A}^T & \mathbf{0} \\ \mathbf{A} & \mathbf{0} & \mathbf{B} \\ \mathbf{0} & \mathbf{B}^T & -\mathbf{W}_{xx} \end{bmatrix} \begin{bmatrix} \mathbf{v} \\ \mathbf{k} \\ \delta\mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{f} \\ \mathbf{0} \end{bmatrix} \tag{2.17}$$

Equation (2.17) can be solved by the following reduction process given by Cross (1992, pp. 22-23). Consider the partitioned matrix equation $\mathbf{P}\,\mathbf{y} = \mathbf{u}$ given as

$$\left[\begin{array}{c|c} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \hline \mathbf{P}_{21} & \mathbf{P}_{22} \end{array}\right] \left[\begin{array}{c} \mathbf{y}_1 \\ \hline \mathbf{y}_2 \end{array}\right] = \left[\begin{array}{c} \mathbf{u}_1 \\ \hline \mathbf{u}_2 \end{array}\right] \tag{2.18}$$

which can be expanded to give

$$\mathbf{P}_{11}\,\mathbf{y}_1 + \mathbf{P}_{12}\,\mathbf{y}_2 = \mathbf{u}_1$$

or

$$\mathbf{y}_1 = \mathbf{P}_{11}^{-1}\left(\mathbf{u}_1 - \mathbf{P}_{12}\,\mathbf{y}_2\right) \tag{2.19}$$

Eliminating $\mathbf{y}_1$ by substituting (2.19) into (2.18) gives

$$\left[\begin{array}{c|c} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \hline \mathbf{P}_{21} & \mathbf{P}_{22} \end{array}\right] \left[\begin{array}{c} \mathbf{P}_{11}^{-1}\left(\mathbf{u}_1 - \mathbf{P}_{12}\,\mathbf{y}_2\right) \\ \hline \mathbf{y}_2 \end{array}\right] = \left[\begin{array}{c} \mathbf{u}_1 \\ \hline \mathbf{u}_2 \end{array}\right]$$

Expanding the matrix equation gives

$$\begin{aligned} \mathbf{P}_{21}\mathbf{P}_{11}^{-1}\left(\mathbf{u}_1 - \mathbf{P}_{12}\mathbf{y}_2\right) + \mathbf{P}_{22}\mathbf{y}_2 &= \mathbf{u}_2 \\ \mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{u}_1 - \mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\mathbf{y}_2 + \mathbf{P}_{22}\mathbf{y}_2 &= \mathbf{u}_2 \end{aligned}$$

and an expression for $\mathbf{y}_2$ is given by

$$\left(\mathbf{P}_{22} - \mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\right)\mathbf{y}_2 = \mathbf{u}_2 - \mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{u}_1 \tag{2.20}$$

Now partitioning (2.17) in the same way as (2.18)

$$\left[\begin{array}{c|cc} -\mathbf{W} & \mathbf{A}^T & \mathbf{0} \\ \hline \mathbf{A} & \mathbf{0} & \mathbf{B} \\ \mathbf{0} & \mathbf{B}^T & -\mathbf{W}_{xx} \end{array}\right] \left[\begin{array}{c} \mathbf{v} \\ \hline \mathbf{k} \\ \delta\mathbf{x} \end{array}\right] = \left[\begin{array}{c} \mathbf{0} \\ \hline \mathbf{f} \\ \mathbf{0} \end{array}\right] \tag{2.21}$$

$\mathbf{v}$ can be eliminated by applying (2.20)

$$\left[\begin{bmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{B}^T & -\mathbf{W}_{xx} \end{bmatrix} - \begin{bmatrix} \mathbf{A} \\ \mathbf{0} \end{bmatrix}\left[\mathbf{W}^{-1}\right]\left[\mathbf{A}^T \quad \mathbf{0}\right]\right]\begin{bmatrix} \mathbf{k} \\ \delta\mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{A} \\ \mathbf{0} \end{bmatrix}\left[-\mathbf{W}^{-1}\right]\left[\mathbf{0}\right]$$

Remembering that $\mathbf{Q} = \mathbf{W}^{-1}$ the equation can be simplified as

$$\left[\begin{array}{c|c} \mathbf{AQA}^T & \mathbf{B} \\ \hline \mathbf{B}^T & -\mathbf{W}_{xx} \end{array}\right]\left[\begin{array}{c} \mathbf{k} \\ \hline \delta\mathbf{x} \end{array}\right] = \left[\begin{array}{c} \mathbf{f} \\ \hline \mathbf{0} \end{array}\right] \tag{2.22}$$

Again, applying (2.20) to the partitioned equation (2.22) gives

$$\left(-\mathbf{W}_{xx} - \mathbf{B}^T\left(\mathbf{AQA}^T\right)^{-1}\mathbf{B}\right)\delta\mathbf{x} = \mathbf{0} - \mathbf{B}^T\left(\mathbf{AQA}^T\right)^{-1}\mathbf{f}$$

and re-arranging gives the normal equations

$$\left(\mathbf{B}^T\left(\mathbf{AQA}^T\right)^{-1}\mathbf{B} + \mathbf{W}_{xx}\right)\delta\mathbf{x} = \mathbf{B}^T\left(\mathbf{AQA}^T\right)^{-1}\mathbf{f} \tag{2.23}$$

Mikhail (1976, p. 114) simplifies (2.23) by introducing *equivalent observations* $\mathbf{l}_e$ where

$$\mathbf{l}_e = \mathbf{A}\,\mathbf{l} \tag{2.24}$$

Applying the matrix rule for cofactor propagation (Mikhail 1976, pp. 76-79) gives the cofactor matrix of the equivalent observations as

$$\mathbf{Q}_e = \mathbf{AQA}^T \tag{2.25}$$

With the usual relationship between weight matrices and cofactor matrices, see (2.6), we may write

$$\mathbf{W}_e = \mathbf{Q}_e^{-1} = \left(\mathbf{AQA}^T\right)^{-1} \tag{2.26}$$

Using (2.26) in (2.23) gives the normal equations as

$$\boxed{\left(\mathbf{B}^T\mathbf{W}_e\mathbf{B} + \mathbf{W}_{xx}\right)\delta\mathbf{x} = \mathbf{B}^T\mathbf{W}_e\mathbf{f}} \tag{2.27}$$

With the auxiliaries $\mathbf{N}$ and $\mathbf{t}$

$$\mathbf{N} = \mathbf{B}^T\mathbf{W}_e\mathbf{B} \tag{2.28}$$

$$\mathbf{t} = \mathbf{B}^T\mathbf{W}_e\mathbf{f} \tag{2.29}$$

the vector of corrections $\delta\mathbf{x}$ is given by

$$\boxed{\delta\mathbf{x} = \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{t}} \tag{2.30}$$

The vector of Lagrange multipliers **k** are obtained from (2.22) by applying (2.19) to give

$$\mathbf{k} = \left(\mathbf{AQA}^T\right)^{-1}(\mathbf{f} - \mathbf{B}\,\delta\mathbf{x}) = \mathbf{W}_e(\mathbf{f} - \mathbf{B}\,\delta\mathbf{x}) \tag{2.31}$$

and the vector of residuals **v** is obtained from (2.21) as

$$-\mathbf{Wv} + \mathbf{A}^T\mathbf{k} = \mathbf{0}$$

giving

$$\mathbf{v} = \mathbf{W}^{-1}\mathbf{A}^T\mathbf{k} = \mathbf{QA}^T\mathbf{k} \tag{2.32}$$

2.3. The Iterative Process of Solution

Remembering that $\hat{\mathbf{x}} = \mathbf{x} + \delta\mathbf{x}$, see (2.3), where **x** is the vector of a priori estimates of the parameters, $\delta\mathbf{x}$ is a vector of corrections and $\hat{\mathbf{x}}$ is the least squares estimate of the parameters.

At the beginning of the iterative solution, it can be assumed that $\hat{\mathbf{x}}$ equals the a priori estimates $\mathbf{x}_1$ and a set of corrections $\delta\mathbf{x}_1$ computed. These are added to $\mathbf{x}_1$ to give an updated set $\mathbf{x}_2$. **A** and **B** are recalculated and a new weight matrix $\mathbf{W}_{xx}$ computed by cofactor propagation. The corrections are computed again, and the whole process cycles through until the corrections reach some predetermined value, which terminates the process.

$$\hat{\mathbf{x}}_{n+1} = \mathbf{x}_n + \delta\mathbf{x}_n \tag{2.33}$$

2.4. Derivation of Cofactor Matrices

In this section, the cofactor matrices of the vectors $\hat{\mathbf{x}}$, $\delta\mathbf{x}$, **v** and $\hat{\mathbf{l}}$ will be derived. The law of cofactor propagation will be used and is defined as follows (Mikhail 1976, pp. 76-89).

Given a functional relationship

$$\mathbf{z} = F(\mathbf{x}) \tag{2.34}$$

between two random vectors **z** and **x** and the variance-covariance matrix Σ_{xx}, the variance-covariance matrix of **z** is given by

$$\mathbf{\Sigma}_{zz} = \mathbf{J}_{zx}\mathbf{\Sigma}_{xx}\mathbf{J}_{zx}^T \tag{2.35}$$

where $\mathbf{J}_{zx}$ is a matrix of partial derivatives

$$\mathbf{J}_{zx} = \frac{\partial F}{\partial \mathbf{x}} = \begin{bmatrix} \dfrac{\partial z_1}{\partial x_1} & \dfrac{\partial z_1}{\partial x_2} & \cdots & \dfrac{\partial z_1}{\partial x_n} \\ \dfrac{\partial z_2}{\partial x_1} & \dfrac{\partial z_2}{\partial x_2} & \cdots & \dfrac{\partial z_2}{\partial x_n} \\ \vdots & & & \vdots \\ \dfrac{\partial z_m}{\partial x_1} & \dfrac{\partial z_m}{\partial x_2} & \cdots & \dfrac{\partial z_m}{\partial x_n} \end{bmatrix}$$

Using the relationship between variance-covariance matrices and cofactor matrices, see (2.5), the law of cofactor propagation may be obtained from (2.35), as

$$\mathbf{Q}_{zz} = \mathbf{J}_{zx}\mathbf{Q}_{xx}\mathbf{J}_{zx}^T \tag{2.36}$$

For a function $\mathbf{z}$ containing two independent random variables $\mathbf{x}$ and $\mathbf{y}$ with cofactor matrices $\mathbf{Q}_{xx}$ and $\mathbf{Q}_{yy}$

$$\mathbf{z} = F(\mathbf{x}, \mathbf{y}) \tag{2.37}$$

the law of cofactor propagation gives the cofactor matrix $\mathbf{Q}_{zz}$ as

$$\mathbf{Q}_{zz} = \mathbf{J}_{zx}\mathbf{Q}_{xx}\mathbf{J}_{zx}^T + \mathbf{J}_{zy}\mathbf{Q}_{yy}\mathbf{J}_{zy}^T \tag{2.38}$$

2.4.1. Cofactor Matrix for $\hat{\mathbf{x}}$

According to equations (2.33) and (2.30) with (2.29) the least squares estimate $\hat{\mathbf{x}}$ is

$$\hat{\mathbf{x}} = \mathbf{x} + \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{B}^T\mathbf{W}_e\mathbf{f} \tag{2.39}$$

and $\hat{\mathbf{x}}$ is a function of the a priori parameters $\mathbf{x}$ (the "observables") and the observations $\mathbf{l}$ since the vector of numeric terms $\mathbf{f}$ contains functions of both. Applying the law of propagation of cofactors gives

$$\mathbf{Q}_{\hat{x}\hat{x}} = \left(\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}}\right)\mathbf{Q}_{xx}\left(\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}}\right)^T + \left(\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{l}}\right)\mathbf{Q}\left(\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{l}}\right)^T \tag{2.40}$$

The partial derivatives of (2.39) are

$$\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} = \mathbf{I} + \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{B}^T\mathbf{W}_e\frac{\partial \mathbf{f}}{\partial \mathbf{x}} \tag{2.41}$$

$$\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{l}} = \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} \mathbf{B}^T \mathbf{W}_e \frac{\partial \mathbf{f}}{\partial \mathbf{l}} \tag{2.42}$$

From (2.13), $\mathbf{f} = -F(\mathbf{x}, \mathbf{l})$ the partial derivatives $\dfrac{\partial \mathbf{f}}{\partial \mathbf{x}}$ and $\dfrac{\partial \mathbf{f}}{\partial \mathbf{l}}$, are the design matrices **A** and **B** given by (2.11) and (2.12)

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = -\mathbf{B} \tag{2.43}$$

$$\frac{\partial \mathbf{f}}{\partial \mathbf{l}} = -\mathbf{A} \tag{2.44}$$

Substituting (2.43) and (2.44) into (2.41) and (2.42) with the auxiliary $\mathbf{N} = \mathbf{B}^T \mathbf{W}_e \mathbf{B}$ gives

$$\begin{aligned} \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} &= \mathbf{I} - \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} \mathbf{B}^T \mathbf{W}_e \mathbf{B} \\ &= \mathbf{I} - \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} \mathbf{N} \end{aligned} \tag{2.45}$$

$$\frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{l}} = -\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} \mathbf{B}^T \mathbf{W}_e \mathbf{A} \tag{2.46}$$

Substituting (2.45) and (2.46) into (2.40) gives

$$\begin{aligned} \mathbf{Q}_{\hat{x}\hat{x}} &= \left\{\mathbf{I} - \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} \mathbf{N}\right\} \mathbf{Q}_{xx} \left\{\mathbf{I} - \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} \mathbf{N}\right\}^T \\ &+ \left\{-\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} \mathbf{B}^T \mathbf{W}_e \mathbf{A}\right\} \mathbf{Q} \left\{-\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} \mathbf{B}^T \mathbf{W}_e \mathbf{A}\right\}^T \end{aligned} \tag{2.47}$$

With the auxiliary $$\dot{\mathbf{N}} = \left(\mathbf{N} + \mathbf{W}_{xx}\right) \tag{2.48}$$

and noting that the matrices $\mathbf{I}, \mathbf{N}, \dot{\mathbf{N}}$ and $\mathbf{W}_{xx}$ are all symmetric, (2.47) may be simplified as

$$\mathbf{Q}_{\hat{x}\hat{x}} = \left(\mathbf{I} - \dot{\mathbf{N}}^{-1} \mathbf{N}\right) \mathbf{Q}_{xx} \left(\mathbf{I} - \mathbf{N} \dot{\mathbf{N}}^{-1}\right) + \left(\dot{\mathbf{N}}^{-1} \mathbf{B}^T \mathbf{W}_e \mathbf{A}\right) \mathbf{Q} \left(\mathbf{A}^T \mathbf{W}_e \mathbf{B} \dot{\mathbf{N}}^{-1}\right)$$

Remembering that $\mathbf{Q}_e = \mathbf{A}\mathbf{Q}\mathbf{A}^T$ and $\mathbf{W}_e = \mathbf{Q}_e^{-1}$

$$\mathbf{Q}_{\hat{x}\hat{x}} = \mathbf{Q}_{xx} - \mathbf{Q}_{xx} \mathbf{N} \dot{\mathbf{N}}^{-1} - \dot{\mathbf{N}}^{-1} \mathbf{N} \mathbf{Q}_{xx} + \dot{\mathbf{N}}^{-1} \mathbf{N} \mathbf{Q}_{xx} \mathbf{N} \dot{\mathbf{N}}^{-1} + \dot{\mathbf{N}}^{-1} \mathbf{N} \dot{\mathbf{N}}^{-1} \tag{2.49}$$

The last two terms of (2.49) can be simplified as follows

$$\dot{\mathbf{N}}^{-1}\,\mathbf{N}\mathbf{Q}_{xx}\mathbf{N}\,\dot{\mathbf{N}}^{-1}+\dot{\mathbf{N}}^{-1}\,\mathbf{N}\,\dot{\mathbf{N}}^{-1} = \dot{\mathbf{N}}^{-1}\,\mathbf{N}\mathbf{Q}_{xx}\left(\mathbf{N}\,\dot{\mathbf{N}}^{-1}+\mathbf{W}_{xx}\,\dot{\mathbf{N}}^{-1}\right)$$
$$= \dot{\mathbf{N}}^{-1}\,\mathbf{N}\mathbf{Q}_{xx}\left(\mathbf{N}+\mathbf{W}_{xx}\right)\dot{\mathbf{N}}^{-1}$$
$$= \dot{\mathbf{N}}^{-1}\,\mathbf{N}\mathbf{Q}_{xx}\,\dot{\mathbf{N}}\,\dot{\mathbf{N}}^{-1}$$
$$= \dot{\mathbf{N}}^{-1}\,\mathbf{N}\mathbf{Q}_{xx}$$

and substituting this result into (2.49) gives

$$\begin{aligned}\mathbf{Q}_{\hat{x}\hat{x}} &= \mathbf{Q}_{xx} - \mathbf{Q}_{xx}\mathbf{N}\,\dot{\mathbf{N}}^{-1} - \dot{\mathbf{N}}^{-1}\,\mathbf{N}\mathbf{Q}_{xx} + \dot{\mathbf{N}}^{-1}\,\mathbf{N}\mathbf{Q}_{xx}\\ &= \mathbf{Q}_{xx} - \mathbf{Q}_{xx}\mathbf{N}\,\dot{\mathbf{N}}^{-1}\end{aligned} \tag{2.50}$$

Further simplification gives

$$\begin{aligned}\mathbf{Q}_{\hat{x}\hat{x}} &= \mathbf{Q}_{xx}\left(\mathbf{I} - \mathbf{N}\,\dot{\mathbf{N}}^{-1}\right)\\ &= \mathbf{Q}_{xx}\left(\dot{\mathbf{N}} - \mathbf{N}\right)\dot{\mathbf{N}}^{-1}\\ &= \mathbf{Q}_{xx}\left(\mathbf{N} + \mathbf{W}_{xx} - \mathbf{N}\right)\dot{\mathbf{N}}^{-1}\\ &= \mathbf{Q}_{xx}\mathbf{W}_{xx}\,\dot{\mathbf{N}}^{-1}\end{aligned} \tag{2.51}$$

and since $\mathbf{Q}_{xx}\mathbf{W}_{xx} = \mathbf{I}$ the cofactor matrix of the least squares estimates $\hat{\mathbf{x}}$ is

$$\boxed{\mathbf{Q}_{\hat{x}\hat{x}} = \dot{\mathbf{N}}^{-1} = \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}} \tag{2.52}$$

2.4.2. Cofactor Matrix for $\hat{\mathbf{l}}$

Beginning with the final adjusted observations given by (2.2)

$$\hat{\mathbf{l}} = \mathbf{l} + \mathbf{v} \tag{2.53}$$

and using (2.32) and (2.31) we have

$$\begin{aligned}\hat{\mathbf{l}} &= \mathbf{l} + \mathbf{Q}\mathbf{A}^T\mathbf{k}\\ &= \mathbf{l} + \mathbf{Q}\mathbf{A}^T\mathbf{W}_e\left(\mathbf{f} - \mathbf{B}\delta\mathbf{x}\right)\\ &= \mathbf{l} + \mathbf{Q}\mathbf{A}^T\mathbf{W}_e\mathbf{f} - \mathbf{Q}\mathbf{A}^T\mathbf{W}_e\mathbf{B}\delta\mathbf{x}\end{aligned}$$

Substituting the expression for $\delta\mathbf{x}$ given by (2.30) with the auxiliaries $\mathbf{t}$ and $\dot{\mathbf{N}}$ given by (2.29) and (2.48) respectively gives

$$\begin{aligned}\hat{\mathbf{l}} &= \mathbf{l} + \mathbf{QA}^T\mathbf{W}_e\mathbf{f} - \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{t} \\ &= \mathbf{l} + \mathbf{QA}^T\mathbf{W}_e\mathbf{f} - \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{B}^T\mathbf{W}_e\mathbf{f} \\ &= \mathbf{l} + \mathbf{QA}^T\mathbf{W}_e\mathbf{f} - \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{f}\end{aligned} \tag{2.54}$$

and $\hat{\mathbf{l}}$ is function of the observables $\mathbf{x}$ and the observations $\mathbf{l}$ since $\mathbf{f} = -F(\mathbf{x},\mathbf{l})$. Applying the law of propagation of cofactors to (2.54) gives

$$\mathbf{Q}_{\hat{l}\hat{l}} = \left(\frac{\partial\hat{\mathbf{l}}}{\partial\mathbf{x}}\right)\mathbf{Q}_{xx}\left(\frac{\partial\hat{\mathbf{l}}}{\partial\mathbf{x}}\right)^T + \left(\frac{\partial\hat{\mathbf{l}}}{\partial\mathbf{l}}\right)\mathbf{Q}\left(\frac{\partial\hat{\mathbf{l}}}{\partial\mathbf{l}}\right)^T \tag{2.55}$$

and the partial derivatives are obtained from (2.54) as

$$\frac{\partial\hat{\mathbf{l}}}{\partial\mathbf{x}} = \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\frac{\partial\mathbf{f}}{\partial\mathbf{x}} - \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\frac{\partial\mathbf{f}}{\partial\mathbf{x}}$$

$$\frac{\partial\hat{\mathbf{l}}}{\partial\mathbf{l}} = \mathbf{I} + \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\frac{\partial\mathbf{f}}{\partial\mathbf{l}} - \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\frac{\partial\mathbf{f}}{\partial\mathbf{l}}$$

With $\dfrac{\partial\mathbf{f}}{\partial\mathbf{x}} = -\mathbf{B}$ and $\dfrac{\partial\mathbf{f}}{\partial\mathbf{l}} = -\mathbf{A}$, and with the auxiliary $\mathbf{N} = \mathbf{B}^T\mathbf{W}_e\mathbf{B}$ the partial derivatives become

$$\begin{aligned}\frac{\partial\hat{\mathbf{l}}}{\partial\mathbf{x}} &= \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e - \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{B} \\ &= \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{N} - \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{B}\end{aligned} \tag{2.56}$$

$$\frac{\partial\hat{\mathbf{l}}}{\partial\mathbf{l}} = \mathbf{I} + \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{A} - \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{A} \tag{2.57}$$

Substituting (2.56) and (2.57) into (2.55) gives

$$\mathbf{Q}_{\hat{l}\hat{l}} = \left\{1^{\text{st}}\ \text{term}\right\} + \left\{2^{\text{nd}}\ \text{term}\right\} \tag{2.58}$$

where

$$\left\{1^{\text{st}}\text{ term}\right\} = \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{NQ}_{xx}\mathbf{N}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} - \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{NQ}_{xx}\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$

$$- \mathbf{QA}^T\mathbf{W}_e\mathbf{BQ}_{xx}\mathbf{N}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} + \mathbf{QA}^T\mathbf{W}_e\mathbf{BQ}_{xx}\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$

$$\left\{2^{\text{nd}}\text{ term}\right\} = \mathbf{Q} + \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} - \mathbf{QA}^T\mathbf{W}_e\mathbf{AQ}$$

$$+ \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$

$$+ \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$

$$- \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQA}^T\mathbf{W}_e\mathbf{AQ} - \mathbf{QA}^T\mathbf{W}_e\mathbf{AQ}$$

$$- \mathbf{QA}^T\mathbf{W}_e\mathbf{AQA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} + \mathbf{QA}^T\mathbf{W}_e\mathbf{AQA}^T\mathbf{W}_e\mathbf{AQ}$$

The 1st term can be simplified as

$$\left\{1^{\text{st}}\text{ term}\right\} = \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\left(\dot{\mathbf{N}}^{-1}\,\mathbf{NQ}_{xx}\mathbf{N}\,\dot{\mathbf{N}}^{-1} - \dot{\mathbf{N}}^{-1}\,\mathbf{NQ}_{xx} - \mathbf{Q}_{xx}\,\dot{\mathbf{N}}^{-1}\,\mathbf{N} + \mathbf{Q}_{xx}\right)\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$

$$= \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\left(\dot{\mathbf{N}}^{-1}\,\mathbf{N}\left(\mathbf{Q}_{xx}\mathbf{N}\,\dot{\mathbf{N}}^{-1} - \mathbf{Q}_{xx}\right) - \mathbf{Q}_{xx}\,\dot{\mathbf{N}}^{-1}\,\mathbf{N} + \mathbf{Q}_{xx}\right)\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$

but we know from (2.50) that $\mathbf{Q}_{\hat{x}\hat{x}} = \mathbf{Q}_{xx} - \mathbf{Q}_{xx}\mathbf{N}\,\dot{\mathbf{N}}^{-1}$, and from (2.52) that $\mathbf{Q}_{\hat{x}\hat{x}} = \dot{\mathbf{N}}^{-1}$ so

$$\left\{1^{\text{st}}\text{ term}\right\} = \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\left(\mathbf{Q}_{\hat{x}\hat{x}} - \dot{\mathbf{N}}^{-1}\,\mathbf{NQ}_{\hat{x}\hat{x}}\right)\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$

$$= \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\left(\dot{\mathbf{N}}^{-1} - \dot{\mathbf{N}}^{-1}\,\mathbf{N}\,\dot{\mathbf{N}}^{-1}\right)\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$

$$= \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\left(\mathbf{I} - \mathbf{N}\,\dot{\mathbf{N}}^{-1}\right)\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$

The term in brackets has been simplified in (2.51) as $\mathbf{W}_{xx}\,\dot{\mathbf{N}}^{-1}$ which gives the 1st term as

$$\left\{1^{\text{st}}\text{ term}\right\} = \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{W}_{xx}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} \tag{2.59}$$

The 2nd term of (2.58) can be simplified by remembering that $\mathbf{AQA}^T = \mathbf{Q}_e = \mathbf{W}_e^{-1}$ so that after some cancellation of terms we have

$$\left\{2^{\text{nd}}\text{ term}\right\} = \mathbf{Q} + \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{N}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} - \mathbf{QA}^T\mathbf{W}_e\mathbf{AQ} \tag{2.60}$$

Substituting (2.59) and (2.60) into (2.58) gives the cofactor matrix of the adjusted observations as

$$\boxed{\mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{Q} + \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} - \mathbf{QA}^T\mathbf{W}_e\mathbf{AQ}} \tag{2.61}$$

2.4.3. Cofactor Matrix for $\delta\mathbf{x}$

From (2.30) and (2.29)

$$\begin{aligned}\delta\mathbf{x} &= \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{B}^T\mathbf{W}_e\mathbf{f} \\ &= \dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{f}\end{aligned} \tag{2.62}$$

and applying the law of propagation of cofactors gives

$$\mathbf{Q}_{\delta\mathbf{x}\delta\mathbf{x}} = \left(\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\right)\mathbf{Q}_{ff}\left(\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\right)^T \tag{2.63}$$

The cofactor matrix $\mathbf{Q}_{ff}$ is obtained from $\mathbf{f} = -F(\mathbf{x},\mathbf{l})$ as

$$\begin{aligned}\mathbf{Q}_{ff} &= \left(\frac{\partial\mathbf{f}}{\partial\mathbf{x}}\right)\mathbf{Q}_{xx}\left(\frac{\partial\mathbf{f}}{\partial\mathbf{x}}\right)^T + \left(\frac{\partial\mathbf{f}}{\partial\mathbf{l}}\right)\mathbf{Q}\left(\frac{\partial\mathbf{f}}{\partial\mathbf{l}}\right)^T \\ &= (-\mathbf{B})\,\mathbf{Q}_{xx}\,(-\mathbf{B})^T + (-\mathbf{A})\,\mathbf{Q}\,(-\mathbf{A})^T \\ &= \mathbf{B}\mathbf{Q}_{xx}\mathbf{B}^T + \mathbf{A}\mathbf{Q}\mathbf{A}^T \\ &= \mathbf{B}\mathbf{Q}_{xx}\mathbf{B}^T + \mathbf{Q}_e\end{aligned} \tag{2.64}$$

Substituting (2.64) into (2.63) and simplifying gives

$$\mathbf{Q}_{\delta\mathbf{x}\delta\mathbf{x}} = \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{N}\mathbf{Q}_{xx}\mathbf{N}\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} + \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{N}\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} \tag{2.65}$$

Equation (2.65) can be simplified further as

$$\begin{aligned}\mathbf{Q}_{\delta\mathbf{x}\delta\mathbf{x}} &= \dot{\mathbf{N}}^{-1}\,\mathbf{N}\mathbf{Q}_{xx}\mathbf{N}\,\dot{\mathbf{N}}^{-1} + \dot{\mathbf{N}}^{-1}\,\mathbf{N}\,\dot{\mathbf{N}}^{-1} \\ &= \dot{\mathbf{N}}^{-1}\,\mathbf{N}\mathbf{Q}_{xx}\left(\mathbf{N}\,\dot{\mathbf{N}}^{-1} + \mathbf{W}_{xx}\,\dot{\mathbf{N}}^{-1}\right) \\ &= \dot{\mathbf{N}}^{-1}\,\mathbf{N}\mathbf{Q}_{xx}\left(\mathbf{N} + \mathbf{W}_{xx}\right)\dot{\mathbf{N}}^{-1} \\ &= \dot{\mathbf{N}}^{-1}\,\mathbf{N}\mathbf{Q}_{xx}\,\dot{\mathbf{N}}\,\dot{\mathbf{N}}^{-1}\end{aligned}$$

or

$$\boxed{\mathbf{Q}_{\delta\mathbf{x}\delta\mathbf{x}} = \dot{\mathbf{N}}^{-1}\,\mathbf{N}\mathbf{Q}_{xx} = \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{N}\mathbf{Q}_{xx}} \tag{2.66}$$

2.4.4. Cofactor Matrix for v

From (2.32), (2.31) and (2.30) we may write the following

$$\begin{aligned}\mathbf{v} &= \mathbf{QA}^T\mathbf{k}\\ &= \mathbf{QA}^T\mathbf{W}_e(\mathbf{f} - \mathbf{B}\delta\mathbf{x})\\ &= \mathbf{QA}^T\mathbf{W}_e\mathbf{f} - \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\delta\mathbf{x}\\ &= \mathbf{QA}^T\mathbf{W}_e\mathbf{f} - \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{t}\end{aligned}$$

and with (2.29) and the auxiliary $\dot{\mathbf{N}}^{-1} = \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}$

$$\mathbf{v} = \mathbf{QA}^T\mathbf{W}_e\mathbf{f} - \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{f} \tag{2.67}$$

v is a function of the observables **x** and the observations **l** since $\mathbf{f} = -F(\mathbf{x},\mathbf{l})$ and applying the law of propagation of cofactors gives

$$\mathbf{Q}_{vv} = \left(\frac{\partial\mathbf{v}}{\partial\mathbf{x}}\right)\mathbf{Q}_{xx}\left(\frac{\partial\mathbf{v}}{\partial\mathbf{x}}\right)^T + \left(\frac{\partial\mathbf{v}}{\partial\mathbf{l}}\right)\mathbf{Q}\left(\frac{\partial\mathbf{v}}{\partial\mathbf{l}}\right)^T \tag{2.68}$$

The partial derivatives of (2.67) are

$$\frac{\partial\mathbf{v}}{\partial\mathbf{x}} = \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\frac{\partial\mathbf{f}}{\partial\mathbf{x}} - \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\frac{\partial\mathbf{f}}{\partial\mathbf{x}}$$

$$\frac{\partial\mathbf{v}}{\partial\mathbf{l}} = \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\frac{\partial\mathbf{f}}{\partial\mathbf{l}} - \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\frac{\partial\mathbf{f}}{\partial\mathbf{l}}$$

With $\dfrac{\partial\mathbf{f}}{\partial\mathbf{x}} = -\mathbf{B}$ and $\dfrac{\partial\mathbf{f}}{\partial\mathbf{l}} = -\mathbf{A}$, and with the auxiliary $\mathbf{N} = \mathbf{B}^T\mathbf{W}_e\mathbf{B}$ the partial derivatives become

$$\frac{\partial\mathbf{v}}{\partial\mathbf{x}} = \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{N} - \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{B} \tag{2.69}$$

$$\frac{\partial\mathbf{v}}{\partial\mathbf{l}} = \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{A} - \mathbf{Q}\,\mathbf{A}^T\mathbf{W}_e\mathbf{A} \tag{2.70}$$

Substituting (2.69) and (2.70) into (2.68) gives

$$\mathbf{Q}_{vv} = \left\{1^{\text{st}}\text{ term}\right\} + \left\{2^{\text{nd}}\text{ term}\right\} \tag{2.71}$$

where

$$\left\{1^{\text{st}}\ \text{term}\right\} = \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{NQ}_{xx}\mathbf{N}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} - \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{NQ}_{xx}\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$
$$- \mathbf{QA}^T\mathbf{W}_e\mathbf{BQ}_{xx}\mathbf{N}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} + \mathbf{QA}^T\mathbf{W}_e\mathbf{BQ}_{xx}\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$

$$\left\{2^{\text{nd}}\ \text{term}\right\} = \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$
$$- \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQA}^T\mathbf{W}_e\mathbf{AQ}$$
$$- \mathbf{QA}^T\mathbf{W}_e\mathbf{AQA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$
$$+ \mathbf{QA}^T\mathbf{W}_e\mathbf{AQA}^T\mathbf{W}_e\mathbf{AQ}$$

The 1$^{\text{st}}$ term above is identical to the 1$^{\text{st}}$ term of (2.58) which simplifies to (2.59) as

$$\left\{1^{\text{st}}\ \text{term}\right\} = \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{W}_{xx}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} \tag{2.72}$$

The 2$^{\text{nd}}$ term above can be simplified by remembering that $\mathbf{AQA}^T = \mathbf{Q}_e = \mathbf{W}_e^{-1}$ so that after some manipulation we have

$$\left\{2^{\text{nd}}\ \text{term}\right\} = \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\left(\dot{\mathbf{N}}^{-1}\,\mathbf{N}\,\dot{\mathbf{N}}^{-1} - \dot{\mathbf{N}}^{-1}\right)\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}$$
$$- \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} + \mathbf{QA}^T\mathbf{W}_e\mathbf{AQ}$$

The term in brackets can be expressed as

$$\begin{aligned}\dot{\mathbf{N}}^{-1}\,\mathbf{N}\,\dot{\mathbf{N}}^{-1} - \dot{\mathbf{N}}^{-1} &= \dot{\mathbf{N}}^{-1}\left(\mathbf{N} - \dot{\mathbf{N}}\right)\dot{\mathbf{N}}^{-1}\\ &= \dot{\mathbf{N}}^{-1}\left(\mathbf{N} - \left(\mathbf{N} + \mathbf{W}_{xx}\right)\right)\dot{\mathbf{N}}^{-1}\\ &= -\dot{\mathbf{N}}^{-1}\,\mathbf{W}_{xx}\,\dot{\mathbf{N}}^{-1}\end{aligned}$$

and the 2$^{\text{nd}}$ term becomes

$$\begin{aligned}\left\{2^{\text{nd}}\ \text{term}\right\} = &-\mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{W}_{xx}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ}\\ &- \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\dot{\mathbf{N}}^{-1}\,\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} + \mathbf{QA}^T\mathbf{W}_e\mathbf{AQ}\end{aligned} \tag{2.73}$$

Substituting (2.72) and (2.73) into (2.71) gives the cofactor matrix of the residuals **v** as

$$\boxed{\mathbf{Q}_{vv} = -\mathbf{QA}^T\mathbf{W}_e\mathbf{B}\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} + \mathbf{QA}^T\mathbf{W}_e\mathbf{AQ}} \tag{2.74}$$

and by inspection of (2.64) and (2.74)

$$\mathbf{Q}_{vv} = \mathbf{Q} - \mathbf{Q}_{\hat{l}\hat{l}} \tag{2.75}$$

2.4.5. Covariance Matrix $\Sigma_{\hat{x}\hat{x}}$

$$\Sigma_{\hat{x}\hat{x}} = \sigma_0^2\,\mathbf{Q}_{\hat{x}\hat{x}} \tag{2.76}$$

The estimated variance factor is

$$\sigma_0^2 = \frac{\mathbf{v}^T\mathbf{W}\mathbf{v} + \delta\mathbf{x}^T\mathbf{W}_{xx}\delta\mathbf{x}}{r} \tag{2.77}$$

and the degrees of freedom r are

$$r = m - u + u_x \tag{2.78}$$

where m is the number of equations used to estimate the u parameters from n observations. u_x is the number of weighted parameters. [Equation (2.78) is given by Krakiwsky (1975, p.17, eqn 2-62) who notes that it is an approximation only and directs the reader to Bossler (1972) for a complete and rigorous treatment.]

2.5. Generation of the Standard Least Squares Cases

Depending on the form of the design matrices **A** and **B**, and also on whether the parameters are treated as observables, ie, is $\mathbf{W}_{xx} = 0$, there are several different possibilities for the formulation and solution of least squares problems. The standard cases are listed below.

2.5.1. Combined Case with Weighted Parameters $(\mathbf{A},\ \mathbf{B},\ \mathbf{W},\ \mathbf{W}_{xx} \neq \mathbf{0})$ $\quad \mathbf{A}\mathbf{v} + \mathbf{B}\delta\mathbf{x} = \mathbf{f}$

The general case of a non-linear implicit model with weighted parameters treated as observables is known as the Combined Case with Weighted Parameters. It has a solution given by the following equations (2.30), (2.28), (2.29), (2.26), (2.3), (2.31), (2.32), (2.2), (2.65), (2.52), (2.74), (2.61), (2.64), (2.77) and (2.78).

$$\delta\mathbf{x} = \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{t} \tag{2.79}$$

$$\text{with} \quad \mathbf{N} = \mathbf{B}^T\mathbf{W}_e\mathbf{B} \tag{2.80}$$

$$\mathbf{t} = \mathbf{B}^T\mathbf{W}_e\mathbf{f} \tag{2.81}$$

$$\mathbf{W}_e = \mathbf{Q}_e^{-1} = \left(\mathbf{AQA}^T\right)^{-1} \tag{2.82}$$

$$\hat{\mathbf{x}} = \mathbf{x} + \delta\mathbf{x} \tag{2.83}$$

$$\mathbf{k} = \mathbf{W}_e(\mathbf{f} - \mathbf{B}\,\delta\mathbf{x}) \tag{2.84}$$

$$\mathbf{v} = \mathbf{W}^{-1}\mathbf{A}^T\mathbf{k} = \mathbf{QA}^T\mathbf{k} \tag{2.85}$$

$$\hat{\mathbf{l}} = \mathbf{l} + \mathbf{v} \tag{2.86}$$

$$\begin{aligned}\mathbf{Q}_{\delta\mathbf{x}\delta\mathbf{x}} &= \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{NQ}_{xx}\mathbf{N}\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} + \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{N}\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} \\ &= \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{NQ}_{xx}\end{aligned} \tag{2.87}$$

$$\mathbf{Q}_{\hat{x}\hat{x}} = \left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1} \tag{2.88}$$

$$\mathbf{Q}_{vv} = \mathbf{QA}^T\mathbf{W}_e\mathbf{AQ} - \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} \tag{2.89}$$

$$\mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{Q} + \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\left(\mathbf{N} + \mathbf{W}_{xx}\right)^{-1}\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} - \mathbf{QA}^T\mathbf{W}_e\mathbf{AQ} \tag{2.90}$$

$$\mathbf{Q}_{ff} = \mathbf{BQ}_{xx}\mathbf{B}^T + \mathbf{Q}_e \tag{2.91}$$

$$\sigma_0^2 = \frac{\mathbf{v}^T\mathbf{Wv} + \delta\mathbf{x}^T\mathbf{W}_{xx}\delta\mathbf{x}}{r} \tag{2.92}$$

$$r = m - u + u_x \tag{2.93}$$

$$\boldsymbol{\Sigma}_{\delta\mathbf{x}\,\delta\mathbf{x}} = \sigma_0^2\,\mathbf{Q}_{\delta\mathbf{x}\,\delta\mathbf{x}} \tag{2.94}$$

$$\boldsymbol{\Sigma}_{\hat{x}\hat{x}} = \sigma_0^2\,\mathbf{Q}_{\hat{x}\hat{x}} \tag{2.95}$$

$$\boldsymbol{\Sigma}_{vv} = \sigma_0^2\,\mathbf{Q}_{vv} \tag{2.96}$$

$$\boldsymbol{\Sigma}_{\hat{l}\hat{l}} = \sigma_0^2\,\mathbf{Q}_{\hat{l}\hat{l}} \tag{2.97}$$

$$\boldsymbol{\Sigma}_{ff} = \sigma_0^2\,\mathbf{Q}_{ff} \tag{2.98}$$

2.5.2. Combined Case $\left(\mathbf{A},\ \mathbf{B},\ \mathbf{W},\ \mathbf{W}_{xx} = \mathbf{0}\right)$ $\quad \mathbf{Av} + \mathbf{B}\delta\mathbf{x} = \mathbf{f}^0$

The Combined Case is a non-linear implicit mathematical model with no weights on the parameters. The set of equations for the solution is deduced from the Combined Case with Weighted Parameters by considering that if there are no weights then $\mathbf{W}_{xx} = \mathbf{0}$ and $\mathbf{Q}_{xx} = \mathbf{0}$. This implies that $\mathbf{x}$ is a constant vector (denoted by $\mathbf{x}^0$) of approximate values of the parameters, and partial derivatives with respect to $\mathbf{x}^0$ are undefined. Substituting these two null matrices and the constant vector $\mathbf{x} = \mathbf{x}^0$ into equations (2.1) to (2.78) gives the following results.

$$\delta\mathbf{x} = \mathbf{N}^{-1}\mathbf{t} \tag{2.99}$$

with
$$\mathbf{N} = \mathbf{B}^T\mathbf{W}_e\mathbf{B} \tag{2.100}$$

$$\mathbf{t} = \mathbf{B}^T\mathbf{W}_e\mathbf{f}^0 \tag{2.101}$$

$$\mathbf{f}^0 = -F\left(\mathbf{x}^0, \mathbf{l}\right) \tag{2.102}$$

$$\mathbf{W}_e = \mathbf{Q}_e^{-1} = \left(\mathbf{AQA}^T\right)^{-1} \tag{2.103}$$

$$\hat{\mathbf{x}} = \mathbf{x}^0 + \delta\mathbf{x} \tag{2.104}$$

$$\mathbf{k} = \mathbf{W}_e\left(\mathbf{f}^0 - \mathbf{B}\,\delta\mathbf{x}\right) \tag{2.105}$$

$$\mathbf{v} = \mathbf{W}^{-1}\mathbf{A}^T\mathbf{k} = \mathbf{QA}^T\mathbf{k} \tag{2.106}$$

$$\hat{\mathbf{l}} = \mathbf{l} + \mathbf{v} \tag{2.107}$$

$$\mathbf{Q}_{\delta\mathbf{x}\delta\mathbf{x}} = \mathbf{Q}_{\hat{x}\hat{x}} = \mathbf{N}^{-1} \tag{2.108}$$

$$\mathbf{Q}_{vv} = \mathbf{QA}^T\mathbf{W}_e\mathbf{AQ} - \mathbf{QA}^T\mathbf{W}_e\mathbf{BN}^{-1}\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} \tag{2.109}$$

$$\mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{Q} + \mathbf{QA}^T\mathbf{W}_e\mathbf{B}\,\mathbf{N}^{-1}\mathbf{B}^T\mathbf{W}_e\mathbf{AQ} - \mathbf{QA}^T\mathbf{W}_e\mathbf{AQ} \tag{2.110}$$

$$\mathbf{Q}_{f^0 f^0} = \mathbf{Q}_e \tag{2.111}$$

$$\sigma_0^2 = \frac{\mathbf{v}^T\mathbf{W}\mathbf{v}}{r} \tag{2.112}$$

$$r = m - u \tag{2.113}$$

$$\boldsymbol{\Sigma}_{\hat{x}\hat{x}} = \boldsymbol{\Sigma}_{\delta\mathbf{x}\delta\mathbf{x}} = \sigma_0^2\, \mathbf{Q}_{\hat{x}\hat{x}} \tag{2.114}$$

$$\boldsymbol{\Sigma}_{vv} = \sigma_0^2\, \mathbf{Q}_{vv} \tag{2.115}$$

$$\boldsymbol{\Sigma}_{\hat{l}\hat{l}} = \sigma_0^2\, \mathbf{Q}_{\hat{l}\hat{l}} \tag{2.116}$$

$$\boldsymbol{\Sigma}_{f^0 f^0} = \sigma_0^2\, \mathbf{Q}_{f^0 f^0} \tag{2.117}$$

2.5.3. Parametric Case $(\mathbf{A} = \mathbf{I},\ \mathbf{B},\ \mathbf{W},\ \mathbf{W}_{xx} = \mathbf{0})$ $\quad \mathbf{v} + \mathbf{B}\delta\mathbf{x} = \mathbf{f}^0$

The Parametric Case is a mathematical model with the observations $\mathbf{l}$ explicitly expressed by some non-linear function of the parameters $\mathbf{x}$ only. This implies that the design matrix $\mathbf{A}$ is equal to the identity matrix $\mathbf{I}$. Setting $\mathbf{A} = \mathbf{I}$ in the Combined Case (with no weights) leads to the following equations.

$$\delta\mathbf{x} = \mathbf{N}^{-1}\mathbf{t} \tag{2.118}$$

$$\text{with} \qquad \mathbf{N} = \mathbf{B}^T\mathbf{W}\mathbf{B} \tag{2.119}$$

$$\mathbf{t} = \mathbf{B}^T\mathbf{W}\mathbf{f}^0 \tag{2.120}$$

$$\mathbf{f}^0 = -F\left(\mathbf{x}^0, \mathbf{l}\right) \tag{2.121}$$

$$\hat{\mathbf{x}} = \mathbf{x}^0 + \delta\mathbf{x} \tag{2.122}$$

$$\mathbf{v} = \mathbf{f}^0 - \mathbf{B}\,\delta\mathbf{x} \tag{2.123}$$

$$\hat{\mathbf{l}} = \mathbf{l} + \mathbf{v} \tag{2.124}$$

$$\mathbf{Q}_{\delta\mathbf{x}\delta\mathbf{x}} = \mathbf{Q}_{\hat{x}\hat{x}} = \mathbf{N}^{-1} \tag{2.125}$$

$$\mathbf{Q}_{vv} = \mathbf{Q} - \mathbf{B}\mathbf{N}^{-1}\mathbf{B}^T \tag{2.126}$$

$$\mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{B}\,\mathbf{N}^{-1}\mathbf{B}^T \tag{2.127}$$

$$\mathbf{Q}_{f^0 f^0} = \mathbf{Q} \tag{2.128}$$

$$\sigma_0^2 = \frac{\mathbf{v}^T\mathbf{W}\mathbf{v}}{r} \tag{2.129}$$

$$r = n - u \tag{2.130}$$

$$\mathbf{\Sigma}_{\hat{x}\hat{x}} = \mathbf{\Sigma}_{\delta\mathbf{x}\delta\mathbf{x}} = \sigma_0^2\,\mathbf{Q}_{\hat{x}\hat{x}} \tag{2.131}$$

$$\mathbf{\Sigma}_{vv} = \sigma_0^2\,\mathbf{Q}_{vv} \tag{2.132}$$

$$\mathbf{\Sigma}_{\hat{l}\hat{l}} = \sigma_0^2\,\mathbf{Q}_{\hat{l}\hat{l}} \tag{2.133}$$

$$\mathbf{\Sigma}_{f^0 f^0} = \sigma_0^2\,\mathbf{Q}_{f^0 f^0} \tag{2.134}$$

2.5.4. Condition Case $\left(\mathbf{A},\ \mathbf{B} = \mathbf{0},\ \mathbf{W},\ \mathbf{W}_{xx} = \mathbf{0}\right)$ $\quad \mathbf{A}\mathbf{v} = \mathbf{f}$

The Condition Case is characterised by a non-linear model consisting of observations only. Setting $\mathbf{B} = \mathbf{0}$ in the Combined Case (with no weights) leads to the following equations.

$$\mathbf{k} = \mathbf{W}_e\mathbf{f} \tag{2.135}$$

$$\text{with} \qquad \mathbf{W}_e = \mathbf{Q}_e^{-1} = \left(\mathbf{A}\mathbf{Q}\mathbf{A}^T\right)^{-1} \tag{2.136}$$

$$\mathbf{f} = -F(\mathbf{l}) \tag{2.137}$$

$$\mathbf{v} = \mathbf{W}^{-1}\mathbf{A}^T\mathbf{k} = \mathbf{Q}\mathbf{A}^T\mathbf{k} \tag{2.138}$$

$$\hat{\mathbf{l}} = \mathbf{l} + \mathbf{v} \tag{2.139}$$

$$\mathbf{Q}_{vv} = \mathbf{Q}\mathbf{A}^T\mathbf{W}_e\mathbf{A}\mathbf{Q} \tag{2.140}$$

$$\mathbf{Q}_{\hat{l}\hat{l}} = \mathbf{Q} - \mathbf{Q}\mathbf{A}^T\mathbf{W}_e\mathbf{A}\mathbf{Q} \tag{2.141}$$

$$\sigma_0^2 = \frac{\mathbf{v}^T\mathbf{W}\mathbf{v}}{r} \tag{2.142}$$

$$r = m \tag{2.143}$$

$$\mathbf{\Sigma}_{vv} = \sigma_0^2 \, \mathbf{Q}_{vv} \tag{2.144}$$

$$\mathbf{\Sigma}_{\hat{l}\hat{l}} = \sigma_0^2 \, \mathbf{Q}_{\hat{l}\hat{l}} \tag{2.145}$$