CONSTRAINED LEAST SQUARES

Least Squares is used extensively in the analysis and adjustment of survey network measurements. In the majority of applications the measurements (say, directions, distances, height differences, etc) are connected to the unknowns (say coordinates and heights of points) by properly posed models, i.e., there are sufficient "fixed" points in the network to properly define the coordinate origin, the orientation of the network and the datum for heights. In such cases it may be desirable for certain unknowns to accord with geometric conditions, say for instance, the Reduced Levels (RL) of two unknown points in a height network are to be held at a fixed height difference. We may call this a constraint on the adjusted RL's. Or, in a traverse network, we may wish to hold a particular line of the traverse to a fixed bearing. Here we may say that the adjustment has been constrained by the fixed bearing. In these cases, the geometric conditions can be expressed as constraint equations linking the adjusted quantities to certain values. These constrain equations can be added to the normal equations and the combined system solved. The addition of constraint equations lends flexibility to survey network adjustment.

In cases where the adjustment model is not properly posed, i.e., the coordinate datum is not fixed, the network orientation is unknown or the height datum is not fixed, constraint equations can be used to correct for these datum defects. In such cases a minimum number of constraints are required to obtain a solution of the network problem. For example, in a 2-Dimensional survey network of directions and distances the minimum number of constraints would be one fixed point (two coordinates held fixed) and a fixed bearing of a line in the network, or a single fixed point and one other coordinate in the network. Also, constraint equations can be used in "free net" adjustments where every point in the network is regarded as floating and there are no fixed points. In free net adjustments, the constraint equations must take a certain form that will be discussed in subsequent notes.

These notes follow closely the techniques and notation in *Observations and Least Squares* by E.M. Mikhail (Mikhail 1976).

Consider a set of n observation equations in u unknowns $(n > u)$ in the matrix form

$$\mathbf{v} + \mathbf{Bx} = \mathbf{f} \tag{1}$$

$\mathbf{v}$ is the $n \times 1$ vector of residuals

$\mathbf{x}$ is the $u \times 1$ vector of unknowns (or parameters)

$\mathbf{B}$ is the $n \times u$ matrix of coefficients (design matrix)

$\mathbf{f}$ is the $n \times 1$ vector of numeric terms (constants)

n is the number of equations (or observations)

u is the number of unknowns

Also, suppose that the unknowns $\mathbf{x}$ must satisfy certain CONSTRAINTS expressed as equations and written in matrix form as

$$\mathbf{Cx} = \mathbf{g} \tag{2}$$

$\mathbf{C}$ is the $c \times u$ matrix of coefficients

$\mathbf{g}$ is the $c \times 1$ vector of numeric terms (constants)

c is the number of constraint equations

The *Least Squares* condition is enforced by minimizing the function

$$\varphi = \mathbf{v}^T\mathbf{Wv} - 2\mathbf{k}^T\left(\mathbf{Cx} - \mathbf{g}\right) \;\Rightarrow\; \text{minimum} \tag{3}$$

$\mathbf{W}$ is the $n \times n$ weight matrix

$\mathbf{k}$ is the $c \times 1$ vector of *Lagrange multipliers*

To find an expression for $\mathbf{v}^T\mathbf{Wv}$ consider equation (1) which can be rearranged as

$$\mathbf{v} = \mathbf{f} - \mathbf{Bx}$$

and using the matrix rules for transposition

$$\begin{aligned}\mathbf{v}^T &= \left(\mathbf{f} - \mathbf{Bx}\right)^T \\ &= \mathbf{f}^T - \left(\mathbf{Bx}\right)^T \\ &= \mathbf{f}^T - \mathbf{x}^T\mathbf{B}^T\end{aligned}$$

hence

$$\begin{aligned}\mathbf{v}^T\mathbf{W}\mathbf{v} &= \left(\mathbf{f}^T - \mathbf{x}^T\mathbf{B}^T\right)\mathbf{W}\left(\mathbf{f} - \mathbf{B}\mathbf{x}\right) \\ &= \left(\mathbf{f}^T - \mathbf{x}^T\mathbf{B}^T\right)\left(\mathbf{W}\mathbf{f} - \mathbf{W}\mathbf{B}\mathbf{x}\right) \\ &= \mathbf{f}^T\mathbf{W}\mathbf{f} - \mathbf{f}^T\mathbf{W}\mathbf{B}\mathbf{x} - \mathbf{x}^T\mathbf{B}^T\mathbf{W}\mathbf{f} + \mathbf{x}^T\mathbf{B}^T\mathbf{W}\mathbf{B}\mathbf{x}\end{aligned}$$

and each term of this equation is a scalar quantity. Now, since $\left(\mathbf{x}^T\mathbf{B}^T\mathbf{W}\mathbf{f}\right)^T = \mathbf{f}^T\mathbf{W}\mathbf{B}\mathbf{x}$, noting that $\mathbf{W}$ is symmetric, hence $\mathbf{W}^T = \mathbf{W}$, then

$$\mathbf{v}^T\mathbf{W}\mathbf{v} = \mathbf{f}^T\mathbf{W}\mathbf{f} - 2\mathbf{f}^T\mathbf{W}\mathbf{B}\mathbf{x} + \mathbf{x}^T\mathbf{B}^T\mathbf{W}\mathbf{B}\mathbf{x} \tag{4}$$

Making the substitutions

$$\mathbf{N} = \mathbf{B}^T\mathbf{W}\mathbf{B} \tag{5}$$

$$\mathbf{t} = \mathbf{B}^T\mathbf{W}\mathbf{f} \tag{6}$$

and noting that $\mathbf{t}^T = \left(\mathbf{B}^T\mathbf{W}\mathbf{f}\right)^T = \mathbf{f}^T\mathbf{W}\mathbf{B}$, equation (4) becomes

$$\mathbf{v}^T\mathbf{W}\mathbf{v} = \mathbf{f}^T\mathbf{W}\mathbf{f} - 2\mathbf{t}^T\mathbf{x} + \mathbf{x}^T\mathbf{N}\mathbf{x} \tag{7}$$

Substituting equation (7) into equation (3) gives

$$\varphi = \mathbf{f}^T\mathbf{W}\mathbf{f} - 2\mathbf{t}^T\mathbf{x} + \mathbf{x}^T\mathbf{N}\mathbf{x} - 2\mathbf{k}^T\left(\mathbf{C}\mathbf{x} - \mathbf{g}\right) \tag{8}$$

Minimizing φ by equating the derivative $\dfrac{\partial\varphi}{\partial\mathbf{x}}$ to zero gives

$$\frac{\partial\varphi}{\partial\mathbf{x}} = -2\mathbf{t}^T + 2\mathbf{x}^T\mathbf{N} - 2\mathbf{k}^T\mathbf{C} = \mathbf{0}$$

Dividing by 2, transposing and rearranging gives

$$\mathbf{N}\mathbf{x} - \mathbf{C}^T\mathbf{k} - \mathbf{t} = \mathbf{0} \tag{9}$$

$\mathbf{N}$ is the $u \times u$ coefficient matrix of the least squares normal equations

$\mathbf{t}$ is the $u \times 1$ vector of numeric terms

Equations (2) and (9) can be expressed as

$$\begin{aligned}-\mathbf{N}\mathbf{x} + \mathbf{C}^T\mathbf{k} + \mathbf{t} &= \mathbf{0} \\ \mathbf{C}\mathbf{x} + \mathbf{0}\mathbf{k} \quad - \mathbf{g} &= \mathbf{0}\end{aligned}$$

or in partitioned matrix form

$$\begin{bmatrix} -\mathbf{N} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{k} \end{bmatrix} = \begin{bmatrix} -\mathbf{t} \\ \mathbf{g} \end{bmatrix} \tag{10}$$

The orders of the sub-matrices and matrices are

$$\begin{bmatrix} -\mathbf{N}_{u\times u} & \mathbf{C}^T_{u\times c} \\ \mathbf{C}_{c\times u} & \mathbf{0}_{c\times c} \end{bmatrix}_{(u+c)\times(u+c)} \quad ; \quad \begin{bmatrix} \mathbf{x}_{u\times 1} \\ \mathbf{k}_{c\times 1} \end{bmatrix}_{(u+c)\times 1} \quad ; \quad \begin{bmatrix} -\mathbf{t}_{u\times 1} \\ \mathbf{g}_{c\times 1} \end{bmatrix}_{(u+c)\times 1}$$

Equation (10) can be solved directly for **x** and **k** by

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{k} \end{bmatrix} = \begin{bmatrix} -\mathbf{N} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{t} \\ \mathbf{g} \end{bmatrix} \tag{11}$$

Note that in the case where **N** is singular (usually because of "datum" problems leading to $rank(\mathbf{N}) < u$), the coefficient matrix in equation (10) will be non-singular provided the constraint equations $\mathbf{Cx} = \mathbf{g}$ have been properly chosen.

An alternative solution for **x** can be obtained from equation (10), <u>but only in the case where **N** is non-singular</u> (the usual case if there are no datum problems), using a reduction process given by Cross (1992, pp. 22-23).

Consider the partitioned matrix equation $\mathbf{P}\,\mathbf{y} = \mathbf{u}$ given as

$$\begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} \tag{12}$$

which can be expanded to give

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

or

$$\mathbf{y}_1 = \mathbf{P}_{11}^{-1}\left(\mathbf{u}_1 - \mathbf{P}_{12}\,\mathbf{y}_2\right) \tag{13}$$

Eliminating $\mathbf{y}_1$ by substituting (13) into (12) 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

$$\mathbf{y}_2 = \left(\mathbf{P}_{22} - \mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\right)^{-1}\left(\mathbf{u}_2 - \mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{u}_1\right) \tag{14}$$

Applying equations (13) and (14) to equation (10) gives

$$\begin{aligned}\mathbf{x} &= \left(-\mathbf{N}\right)^{-1}\left(-\mathbf{t} - \mathbf{C}^T\mathbf{k}\right) \\ &= \mathbf{N}^{-1}\mathbf{t} + \mathbf{N}^{-1}\mathbf{C}^T\mathbf{k}\end{aligned} \tag{15}$$

$$\begin{aligned}\mathbf{k} &= \left(\mathbf{0} - \mathbf{C}\left(-\mathbf{N}\right)^{-1}\mathbf{C}^T\right)^{-1}\left(\mathbf{g} - \mathbf{C}\left(-\mathbf{N}\right)^{-1}\left(-\mathbf{t}\right)\right) \\ &= \left(\mathbf{C}\mathbf{N}^{-1}\mathbf{C}^T\right)^{-1}\left(\mathbf{g} - \mathbf{C}\mathbf{N}^{-1}\mathbf{t}\right) \\ &= \left(\mathbf{C}\mathbf{N}^{-1}\mathbf{C}^T\right)^{-1}\mathbf{g} - \left(\mathbf{C}\mathbf{N}^{-1}\mathbf{C}^T\right)^{-1}\mathbf{C}\mathbf{N}^{-1}\mathbf{t}\end{aligned} \tag{16}$$

Note that these solutions for **x** and **k** are only possible when **N** is non-singular, i.e., $\left|\mathbf{N}\right| \neq \mathbf{0}$ and $\mathbf{N}^{-1}$ exists. Substituting equation (16) into equation (15) gives a solution for **x** as

$$\mathbf{x} = \mathbf{N}^{-1}\mathbf{t} + \mathbf{N}^{-1}\mathbf{C}^T\left(\mathbf{C}\mathbf{N}^{-1}\mathbf{C}^T\right)^{-1}\mathbf{g} - \mathbf{N}^{-1}\mathbf{C}^T\left(\mathbf{C}\mathbf{N}^{-1}\mathbf{C}^T\right)^{-1}\mathbf{C}\mathbf{N}^{-1}\mathbf{t}$$

Making the substitutions

$$\mathbf{M} = \left(\mathbf{C}\mathbf{N}^{-1}\mathbf{C}^T\right)^{-1} \quad \text{and} \quad \mathbf{x}' = \mathbf{N}^{-1}\mathbf{t} \tag{17}$$

gives

$$\mathbf{x} = \mathbf{x}' + \mathbf{N}^{-1}\mathbf{C}^T\mathbf{M}\left(\mathbf{g} - \mathbf{C}\mathbf{x}'\right) \tag{18}$$

COVARIANCE PROPAGATION TO FIND $\mathbf{Q}_{xx}$

The solution for **x** and **k**, given by equation (11), may be written as

$$\mathbf{z} = \mathbf{D}\mathbf{w} \tag{19}$$

and using the general law of propagation of variances (linear functions)

$$\mathbf{Q}_{zz} = \mathbf{D}\mathbf{Q}_{ww}\mathbf{D}^T \tag{20}$$

where

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{k} \end{bmatrix} \text{ and } \mathbf{Q}_{zz} = \begin{bmatrix} \mathbf{Q}_{xx} & \mathbf{Q}_{xk} \\ \mathbf{Q}_{kx} & \mathbf{Q}_{kk} \end{bmatrix} \tag{21}$$

$$\mathbf{D} = \begin{bmatrix} -\mathbf{N} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{\alpha} & \boldsymbol{\beta}^T \\ \boldsymbol{\beta} & \boldsymbol{\gamma} \end{bmatrix} \tag{22}$$

$$\mathbf{w} = \begin{bmatrix} -\mathbf{t} \\ \mathbf{g} \end{bmatrix} \text{ and } \mathbf{Q}_{ww} = \begin{bmatrix} \mathbf{Q}_{tt} & \mathbf{Q}_{tg} \\ \mathbf{Q}_{gt} & \mathbf{Q}_{gg} \end{bmatrix} = \begin{bmatrix} \mathbf{N} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{23}$$

Note that the cofactor matrices $\mathbf{Q}_{gg} = \mathbf{0}$, $\mathbf{Q}_{tg} = \mathbf{0}$ and $\mathbf{Q}_{gt} = \mathbf{0}$ since **g** is a vector of constants. The cofactor matrix $\mathbf{Q}_{tt} = \mathbf{N}$ can be obtained from propagation of variances understanding that $\mathbf{t} = \mathbf{B}^T\mathbf{W}\mathbf{f}$ and $\mathbf{Q}_{tt} = \left(\mathbf{B}^T\mathbf{W}\right)\mathbf{Q}_{ff}\left(\mathbf{B}^T\mathbf{W}\right)^T = \mathbf{N}$ since $\mathbf{Q}_{ff} = \mathbf{Q} = \mathbf{W}^{-1}$.

Multiplying the matrices above gives

$$\mathbf{Q}_{zz} = \begin{bmatrix} \mathbf{Q}_{xx} & \mathbf{Q}_{xk} \\ \mathbf{Q}_{kx} & \mathbf{Q}_{kk} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha} & \boldsymbol{\beta}^T \\ \boldsymbol{\beta} & \boldsymbol{\gamma} \end{bmatrix} \begin{bmatrix} \mathbf{N} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} & \boldsymbol{\beta}^T \\ \boldsymbol{\beta} & \boldsymbol{\gamma} \end{bmatrix}^T = \begin{bmatrix} \boldsymbol{\alpha}\mathbf{N}\boldsymbol{\alpha} & \boldsymbol{\alpha}\mathbf{N}\boldsymbol{\beta}^T \\ \boldsymbol{\beta}\mathbf{N}\boldsymbol{\alpha} & \boldsymbol{\beta}\mathbf{N}\boldsymbol{\beta}^T \end{bmatrix}$$

from which we obtain

$$\mathbf{Q}_{xx} = \boldsymbol{\alpha}\mathbf{N}\boldsymbol{\alpha} \tag{24}$$

But, from above, $\mathbf{D} = \begin{bmatrix} -\mathbf{N} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{\alpha} & \boldsymbol{\beta}^T \\ \boldsymbol{\beta} & \boldsymbol{\gamma} \end{bmatrix}$ and we may write $\begin{bmatrix} -\mathbf{N} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} & \boldsymbol{\beta}^T \\ \boldsymbol{\beta} & \boldsymbol{\gamma} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$

giving

$$\begin{aligned} -\mathbf{N}\boldsymbol{\alpha} + \mathbf{C}^T\boldsymbol{\beta} &= \mathbf{I} \\ \mathbf{N}\boldsymbol{\alpha} &= -\mathbf{I} + \mathbf{C}^T\boldsymbol{\beta} \\ \mathbf{C}\boldsymbol{\alpha} &= \mathbf{0} \\ \boldsymbol{\alpha}\mathbf{C}^T &= \mathbf{0} \qquad \text{(since } \boldsymbol{\alpha} \text{ is symmetric)} \end{aligned}$$

Substituting these results into equation (24) gives

$$\begin{aligned} \mathbf{Q}_{xx} &= \boldsymbol{\alpha}\left(-\mathbf{I} + \mathbf{C}^T\boldsymbol{\beta}\right) \\ &= -\boldsymbol{\alpha} + \boldsymbol{\alpha}\mathbf{C}^T\boldsymbol{\beta} \end{aligned}$$

but from above $\boldsymbol{\alpha}\mathbf{C}^T = \mathbf{0}$ hence

$$\mathbf{Q}_{xx} = -\boldsymbol{\alpha} \tag{25}$$

and the matrix $\boldsymbol{\alpha}$ of order $u \times u$ is the upper-left sub-matrix of **D** in equation (22).

REFERENCES

Cross, P.A. 1992, *Advanced Least Squares Applied to Position Fixing,* Working Paper No. 6, Department of Land Information, University of East London.

Mikhail, E.M., 1976, *Observations and Least Squares,* IEP—A Dun-Donnelley Publisher, New York.