3.2 The datum problem

So far we have disregarded the fact that the matrix $A^\mathsf{T}A$ might not be invertible because it is rank deficient. From matrix algebra it is known that the rank of the normal equation matrix $N := A^\mathsf{T}A$, $\operatorname{rank} N$, equals the the rank of A, $\operatorname{rank} A$. If it should happen now that – for some reason – matrix A is rank deficient, then the normal equation matrix $N = A^\mathsf{T}A$ cannot be inverted. The following statements are equivalent:

- Matrix $\underset{m\times n}{A}$ rank deficient ($\operatorname{rank} A < n$),
- A has linear dependent columns,
- $Ax = 0$ has non-trivial solution $x_{\text{hom}} \neq 0$, i.e. the null space $\mathcal{N}(A)$ of A is not empty,
- $\det(A^\mathsf{T}A) = 0$,
- $A^\mathsf{T}A$ has zero eigenvalues.

Let us investigate this problem of rank deficiency of A and N using levelling observations between points P_1, P_2 and P_3 of the height network shown in fig. 3.2.

$$\left.\begin{aligned} h_{12} &= H_2 - H_1 \\ h_{13} &= H_3 - H_1 \\ h_{32} &= H_2 - H_3 \end{aligned}\right\} \implies \begin{pmatrix} h_{12} \\ h_{13} \\ h_{32} \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} H_1 \\ H_2 \\ H_3 \end{pmatrix}$$

$$\implies \underset{3\times 1}{y} = \underset{3\times 3}{A}\ \underset{3\times 1}{x}$$

- $m = 3, n = 3, \operatorname{rank} A = 2 \implies d = n - \operatorname{rank} A = 1 \implies r = m - (n - d) = 1$,
- $\det A = -1 \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix}, -(-1) \begin{vmatrix} 1 & 0 \\ 1 & -1 \end{vmatrix} = 1 + (-1) = 0$,
- $\implies A$ and $N = A^\mathsf{T}A$ are not invertible,
- $d := \dim \mathcal{N}(A) > 0$,
- $Ax = 0$ has a nontrivial solution $\implies$ homogeneous solution $x_{\text{hom}} \neq 0$.

$\implies x + \lambda x_{\text{hom}}$ is a solution of $y = Ax$ because

$$A\,(x + \lambda x_{\text{hom}}) = Ax + \lambda \underbrace{Ax_{\text{hom}}}_{=0} = Ax = y$$

is fullfilled.

Interpretation:

- Unknown heights can be changed by an arbitrary constant height shift without affecting the observations.
- Observed height differences are not sensitive to the null space $\mathcal{N}(A)$.

Solution approach 1: reduce solution space

- Fix $d = \dim \mathcal{N}(A)$ unknowns and eliminate corresponding columns in A so that the rank of A, $\operatorname{rank} A = n - d$, is full.
- Move fixed unknowns to the observation vector, e.g. fix H_1:

$$\Longrightarrow \quad \begin{pmatrix} h_{12} + H_1 \\ h_{13} + H_1 \\ h_{32} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} H_2 \\ H_3 \end{pmatrix}$$

Solution approach 2: augment solution space

Augment solution space by adding $d = \dim \mathcal{N}(A)$ constraints, e.g.

$$H_1 = 0 \quad \Longrightarrow \quad \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} H_1 \\ H_2 \\ H_3 \end{pmatrix} = 0 \qquad \sim \quad \underset{d\times n}{D^\mathsf{T}} \underset{n\times 1}{x} = \underset{d\times 1}{c}$$

In order to remove the rank deficiency of A, matrix D^T must be chosen in such a way that

$$\operatorname{rank}\left([\underset{n\times m}{A^\mathsf{T}} \mid \underset{n\times d}{D}] \right) = n\,.$$

$AD = 0$, however is not required. As an example, $D^\mathsf{T} = [1, -1, 0]$ is not permitted. The approach of augmenting the solution space is far more flexible as compared to approach 1: no changes of original quantities y, A are necessary. Even curious constraints are allowed as long as datum deficiency is resolved. However, we are faced with the constrained Lagrangian

$$\begin{aligned} \mathcal{L}_D(x, \lambda) &= \frac{1}{2} e^\mathsf{T} e + \lambda(D^\mathsf{T} x - c) \\ &= \frac{1}{2} y^\mathsf{T} y - y^\mathsf{T} A x + \frac{1}{2} x^\mathsf{T} A^\mathsf{T} A x + \lambda(D^\mathsf{T} x - c) \\ \frac{\partial \mathcal{L}_D}{\partial x} &= -A^\mathsf{T} y + A^\mathsf{T} A x + D\lambda = 0 \\ \frac{\partial \mathcal{L}_D}{\partial \lambda} &= D^\mathsf{T} x - c = 0 \end{aligned}$$

$$\Longrightarrow \quad \underset{(n+d)\times(n+d)}{\begin{pmatrix} A^\mathsf{T} A & D \\ D^\mathsf{T} & 0 \end{pmatrix}} \underset{(n+d)\times 1}{\begin{pmatrix} \hat{x} \\ \hat{\lambda} \end{pmatrix}} = \begin{pmatrix} A^\mathsf{T} y \\ c \end{pmatrix} \quad \Longrightarrow \quad M\hat{z} = v$$

E.g.

$$A = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \quad \Longrightarrow \quad A^\mathsf{T} A = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

$$M = \begin{pmatrix} 2 & 1 & -1 & 1 \\ 1 & 2 & -1 & 0 \\ -1 & -1 & 2 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\det M = -1 \cdot \det \begin{pmatrix} 1 & 2 & -1 \\ -1 & -1 & 2 \\ 1 & 0 & 0 \end{pmatrix} = -1 \cdot 1 \cdot \det \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = -3$$

$$\Longrightarrow M \text{ regular} \quad \Longrightarrow \quad \hat{z} = M^{-1} v$$

$$\hat{x} = N^{-1} \left\langle A^\mathsf{T} y + Dc - \left\{ D(D^\mathsf{T} N^{-1} D)^{-1} \left[D^\mathsf{T} N^{-1} A^\mathsf{T} y + (D^\mathsf{T} N^{-1} D - I) c \right] \right\} \right\rangle$$

$$N := A^\mathsf{T} A + D D^\mathsf{T}$$

3.3 Linearization of non-linear observation equations

General 1-D-formulation

The functional model

$$y = f(x)\,,$$

expressed by Taylor's theorem, becomes

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \\ &= f(x_0) + \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} (x - x_0) + \underbrace{\frac{1}{2} \left.\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right|_{x_0} (x - x_0)^2 + \ldots}_{\text{negligible if } x - x_0 \text{ small}} \end{aligned}$$

Substracting $f(x_0)$ yields

$$f(x) - f(x_0) = y - y_0 = \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} (x - x_0) + \ldots$$

$$\underbrace{\Delta y = \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{0} (\Delta x) +}_{\text{linear model}} \quad \underbrace{O(\Delta x^2)}_{\substack{\text{terms of higher order} \\ = \text{ model errors}}}$$

with $\Delta x := x - x_0$ and $\Delta y := y - y_0$.

General multi-D formulation

$$y_i = f_i(x_j), \quad i = 1, \ldots, m; \; j = 1, \ldots, n$$
$$x_{j,0} \longrightarrow y_{i,0} = f_i(x_{j,0})$$

$$\begin{aligned}
\Delta y_1 &= \left.\frac{\partial f_1}{\partial x_1}\right|_0 \Delta x_1 + \left.\frac{\partial f_1}{\partial x_2}\right|_0 \Delta x_2 + \cdots + \left.\frac{\partial f_1}{\partial x_n}\right|_0 \Delta x_n \\
\Delta y_2 &= \left.\frac{\partial f_2}{\partial x_1}\right|_0 \Delta x_1 + \left.\frac{\partial f_2}{\partial x_2}\right|_0 \Delta x_2 + \cdots + \left.\frac{\partial f_2}{\partial x_n}\right|_0 \Delta x_n \\
&\vdots \\
\Delta y_m &= \left.\frac{\partial f_m}{\partial x_1}\right|_0 \Delta x_1 + \left.\frac{\partial f_m}{\partial x_2}\right|_0 \Delta x_2 + \cdots + \left.\frac{\partial f_m}{\partial x_n}\right|_0 \Delta x_n.
\end{aligned}$$

Terms of second order and higher have been neglected.

$$\Longrightarrow \begin{pmatrix} \Delta y_1 \\ \Delta y_2 \\ \vdots \\ \Delta y_m \end{pmatrix} = \underbrace{\left.\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}\right|_0}_{\text{Jacobian matrix } A} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \end{pmatrix} \qquad \sim \qquad \Delta y = A(x_0)\, \Delta x$$

Planar distance observation:

$$s_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad \overset{?}{\longrightarrow} \quad y = Ax$$

answer: linearize, Taylor series expansion

Linearization of planar distance observation equation (given Taylor point of expansion is x_i^0, y_i^0, x_j^0, y_j^0 = approximate values of unknown point coordinates); explicit differentiation

$$\underset{\text{“measured”}}{s_{ij}} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} = \sqrt{x_{ij}^2 + y_{ij}^2}$$

$$x_i = x_i^0 + \Delta x_i, \quad y_i = y_i^0 + \Delta y_i,$$
$$x_j = x_j^0 + \Delta x_j, \quad y_j = y_j^0 + \Delta y_j$$

$$\begin{aligned}
s_{ij} &= \sqrt{\left(x_j^0 + \Delta x_j - \left(x_i^0 + \Delta x_i\right)\right)^2 + \left(y_j^0 + \Delta y_j - \left(y_i^0 + \Delta y_i\right)\right)^2} \\
&= \underbrace{\sqrt{\left(x_j^0 - x_i^0\right)^2 + \left(y_j^0 - y_i^0\right)^2}}_{= s_{ij}^0 \text{ (distance from approximate coordinates)}} + \left.\frac{\partial s_{ij}}{\partial x_i}\right|_0 \Delta x_i + \left.\frac{\partial s_{ij}}{\partial x_j}\right|_0 \Delta x_j + \left.\frac{\partial s_{ij}}{\partial y_i}\right|_0 \Delta y_i + \left.\frac{\partial s_{ij}}{\partial y_j}\right|_0 \Delta y_j
\end{aligned}$$

$$\frac{\partial s_{ij}}{\partial x_i} = \frac{\partial s_{ij}}{\partial x_{ij}}\frac{\partial x_{ij}}{\partial x_i} = \frac{1}{2}\frac{1}{\sqrt{x_{ij}^2 + y_{ij}^2}} 2x_{ij}\,(-1) = -\frac{x_j - x_i}{s_{ij}}$$

$$\frac{\partial s_{ij}}{\partial x_j} = +\frac{x_j - x_i}{s_{ij}}, \qquad \frac{\partial s_{ij}}{\partial y_i} = -\frac{y_j - y_i}{s_{ij}}, \qquad \frac{\partial s_{ij}}{\partial y_j} = +\frac{y_j - y_i}{s_{ij}}$$

$$\Longrightarrow \Delta s_{ij} := \underbrace{s_{ij} - s_{ij}^0}_{\text{“reduced observation”}} = \begin{pmatrix} -\frac{x_j^0 - x_i^0}{s_{ij}^0} & -\frac{y_j^0 - y_i^0}{s_{ij}^0} & \frac{x_j^0 - x_i^0}{s_{ij}^0} & \frac{y_j^0 - y_i^0}{s_{ij}^0} \end{pmatrix} \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ \Delta x_j \\ \Delta y_j \end{pmatrix}$$

$$\Delta y = A(x_0)\,\Delta x$$

Sometimes it is more convenient to use implicit differentiation within the linearization of observation equations.

Depart from $s_{ij}^2 = \left(x_j - x_i\right)^2 + \left(y_j - y_i\right)^2$ instead from s_{ij} and calculate the total differential:

$$\not{2} s_{ij}\,\mathrm{d}s_{ij} = \not{2}\left(x_j - x_i\right)\left(\mathrm{d}x_j - \mathrm{d}x_i\right) + \not{2}\left(y_j - y_i\right)\left(\mathrm{d}y_j - \mathrm{d}y_i\right)$$

Solve for $\mathrm{d}s_{ij}$, introduce approximate value and switch from $\mathrm{d} \longrightarrow \Delta$:

$$\Delta s_{ij} := s_{ij} - s_{ij}^0 = \frac{x_j^0 - x_i^0}{s_{ij}^0}\left(\Delta x_j - \Delta x_i\right) + \frac{y_j^0 - y_i^0}{s_{ij}^0}\left(\Delta y_j - \Delta y_i\right)$$

Grid bearings:

$$T_{ij} = \arctan\frac{x_j - x_i}{y_j - y_i}$$

$\Longrightarrow$ Linearized grid bearing observation equation:

$$\begin{aligned} T_{ij} &= T_{ij}^0 + \frac{1}{1 + \left(\frac{x_j^0 - x_i^0}{y_j^0 - y_i^0}\right)^2}\left(-\frac{1}{y_j^0 - y_i^0}\Delta x_i + \frac{x_j^0 - x_i^0}{(y_j^0 - y_i^0)^2}\Delta y_i + \frac{1}{y_j^0 - y_i^0}\Delta x_j - \frac{x_j^0 - x_i^0}{(y_j^0 - y_i^0)^2}\Delta y_j\right) \\ &= T_{ij}^0 + \frac{(y_j^0 - y_i^0)^2}{(s_{ij}^0)^2}\left(-\frac{1}{y_j^0 - y_i^0}\Delta x_i + \frac{x_j^0 - x_i^0}{(y_j^0 - y_i^0)^2}\Delta y_i + \frac{1}{y_j^0 - y_i^0}\Delta x_j - \frac{x_j^0 - x_i^0}{(y_j^0 - y_i^0)^2}\Delta y_j\right) \\ &= T_{ij}^0 - \frac{y_j^0 - y_i^0}{(s_{ij}^0)^2}\Delta x_i + \frac{x_j^0 - x_i^0}{(s_{ij}^0)^2}\Delta y_i + \frac{y_j^0 - y_i^0}{(s_{ij}^0)^2}\Delta x_j - \frac{x_j^0 - x_i^0}{(s_{ij}^0)^2}\Delta y_j \end{aligned}$$

Directions:

$$r_{ij} = T_{ij} - \omega_i \qquad (\omega_i \text{ additional unknown})$$

$\Longrightarrow$ linearization of bearing observation equation (see also Fig. 3.3)

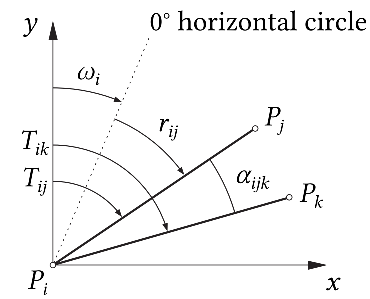


Figure 3.3: Linearization of bearing observation equation, bearing r_{ij}, orientation unknown ω_i.

$$
\begin{aligned}
r_{ij} &= T_{ij} - \omega_i \\
&= \arctan \frac{x_j - x_i}{y_j - y_i} - \omega_i \\
&= r_{ij}^0 - \frac{y_j^0 - y_i^0}{(s_{ij}^0)^2}\Delta x_i + \frac{x_j^0 - x_i^0}{(s_{ij}^0)^2}\Delta y_i + \frac{y_j^0 - y_i^0}{(s_{ij}^0)^2}\Delta x_j - \frac{x_j^0 - x_i^0}{(s_{ij}^0)^2}\Delta y_j - \omega_i
\end{aligned}
$$

Angles:

$$
\begin{aligned}
\alpha_{ijk} &= T_{ik} - T_{ij} \\
&= \arctan \frac{x_k - x_i}{y_k - y_i} - \arctan \frac{x_j - x_i}{y_j - y_i}
\end{aligned}
$$

$\Longrightarrow$ Linearized angle observation equation:

$$
\begin{aligned}
\alpha_{ijk} &= T_{ik}^0 - T_{ij}^0 + \left(-\frac{y_k^0 - y_i^0}{(s_{ik}^0)^2} + \frac{y_j^0 - y_i^0}{(s_{ij}^0)^2} \right) \Delta x_i + \left(\frac{x_k^0 - x_i^0}{(s_{ik}^0)^2} - \frac{x_j^0 - x_i^0}{(s_{ij}^0)^2} \right) \Delta y_i \\
&\quad + \frac{y_k^0 - y_i^0}{(s_{ik}^0)^2}\Delta x_k - \frac{x_k^0 - x_i^0}{(s_{ik}^0)^2}\Delta y_k - \frac{y_j^0 - y_i^0}{(s_{ij}^0)^2}\Delta x_j + \frac{x_j^0 - x_i^0}{(s_{ij}^0)^2}\Delta y_j \\
&= \alpha_{ijk}^0 + \dots
\end{aligned}
$$

3D intersection with additional vertical angles

3D distances:

$$
d_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2 + (z_j - z_i)^2} \qquad (i = 1, \dots, 4;\ j \equiv P)
$$

… linearization as usual.

Vertical angles:

$$\beta_{ij} = \operatorname{arccot} \frac{\sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}}{z_j - z_i} \qquad \text{other trigonometric relations applicable}$$

$$= \operatorname{arccot} \frac{s_{ij}}{z_j - z_i}$$

$$= \beta_{ij}^0 - \frac{1}{1 + \left(\frac{s_{ij}}{z_j - z_i}\right)^2} \cdot \ldots \Delta x_i + \ldots \Delta y_i + \quad \ldots \quad + \ldots \Delta z_j$$

Attention: physical units!

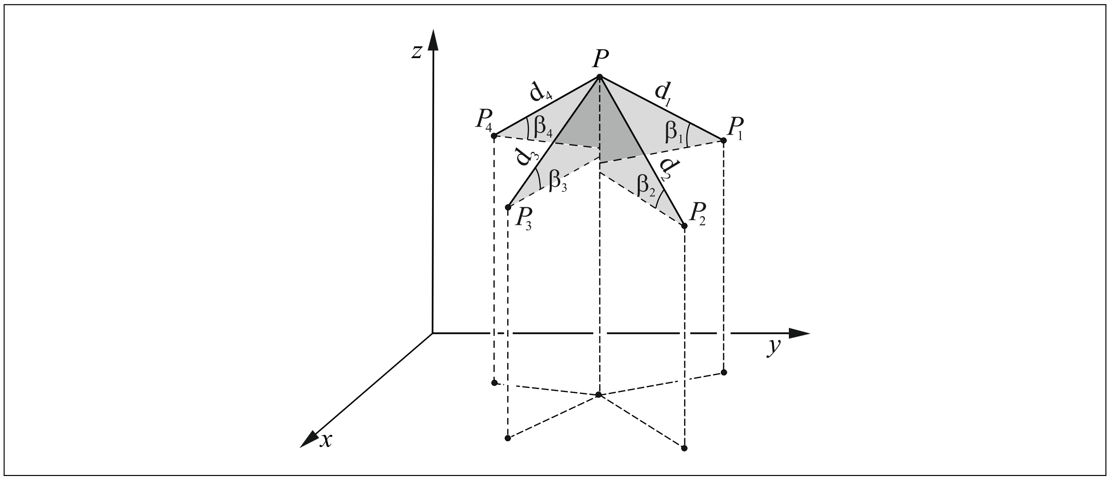


Figure 3.4: 3D intersection and vertical angles.

Iteration (see fig. 3.5)

Linearization (see 3.3) of the functional model $y = f(x)$ yields the linear model:

$$\Delta y = \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} \Delta x + e = A(x_0)\,\Delta x + e\,.$$

The datum problem again

- Matrix $\underset{m\times n}{A}$ is rank deficient (rank $A < n$),
- A has linear dependent columns,
- $Ax = 0$ has non-trivial solution $x_{\text{hom}} \neq 0$, i.e. the null space $\mathcal{N}(A)$ of A is not empty,
- $\det(A^\mathsf{T} A) = 0$,
- $A^\mathsf{T} A$ has zero eigenvalues.

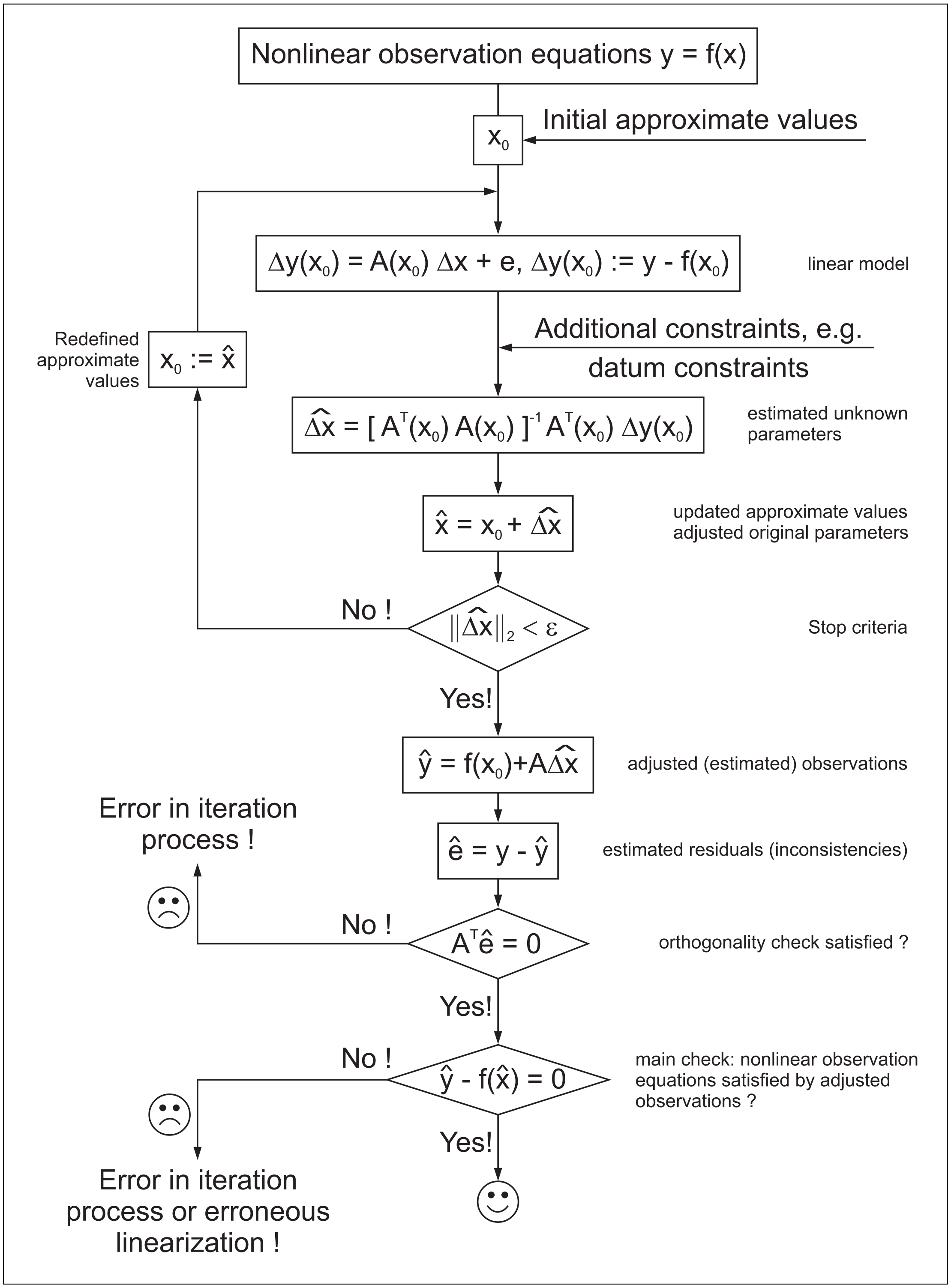


Figure 3.5: Iterative scheme

Example: planar distance network (fig. 3.6)

Rank defect:

- Translation $\longrightarrow$ 2 parameters (x-, y-direction),
- Rotation $\longrightarrow$ 1 parameter,

$\Longrightarrow$ total of $d = 3$ parameters,
$\Longrightarrow$ $\operatorname{rank} A = n - d = n - 3$,
9 points $\longrightarrow$ $n - d = 18 - 3 = 15$, $m = 19$, thus $r = 4$.
Conditional adjustment: How many conditions? Answer: r condition equations.

3.4 Higher dimensions: the B-model (Condition equations)

In the *ideal* case we had

$$h_{1B} - h_{1A} = (H_B - H_1) - (H_A - H_1) = H_B - H_A$$
$$h_{13} + h_{32} - h_{12} = (H_3 - H_1) + (H_2 - H_3) - (H_2 - H_1) = 0$$

or

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} h_{1B} \\ h_{13} \\ h_{12} \\ h_{32} \\ h_{1A} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} H_B \\ 0 \\ 0 \\ 0 \\ H_A \end{pmatrix} .$$

Due to erroneous observations, a vector e of unknown inconsistencies must be introduced in order to make our linear model consistent.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} h_{1B} - e_{1B} \\ h_{13} - e_{13} \\ h_{12} - e_{12} \\ h_{32} - e_{32} \\ h_{1A} - e_{1A} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} H_B \\ 0 \\ 0 \\ 0 \\ H_A \end{pmatrix} .$$

or

$$\underset{2\times5}{B^{\mathsf{T}}} \left(\underset{5\times1}{\Delta h} - \underset{5\times1}{e} \right) = \underset{2\times1}{B^{\mathsf{T}} c} .$$

Connected with this example are the questions

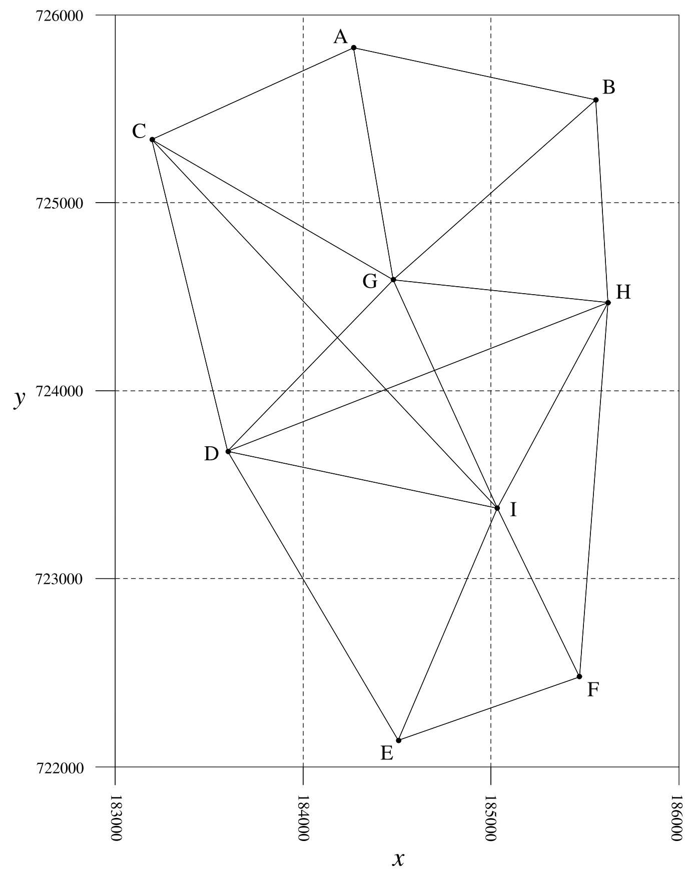
Q 1: How to handle constants like the vector c?

Q 2: How many conditions must be set up?

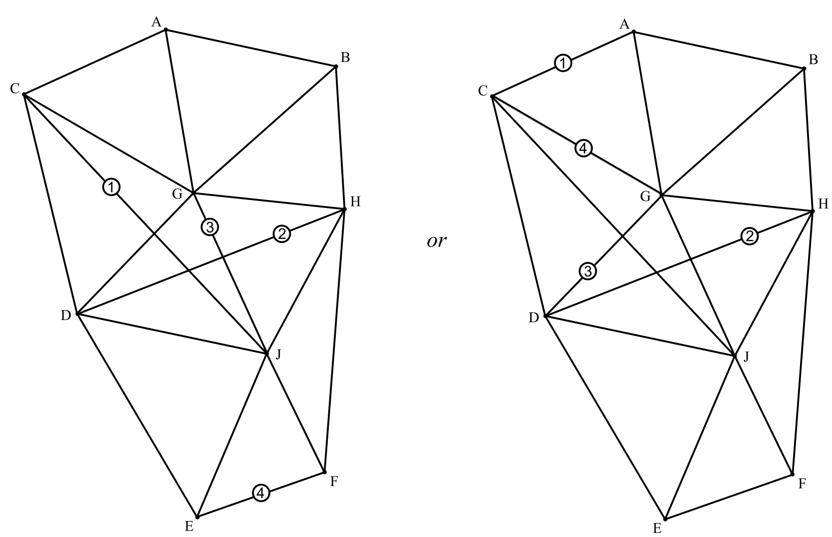
Q 3: Is the solution of the B-model identical to the one of the A-model?

A 1: Starting from

$$B^{\mathsf{T}}(\Delta h - e) = B^{\mathsf{T}} c ,$$



(a) distance network



(b) Four lines may be deleted without destabilizing the net.

Figure 3.6

where solely e is unknown, we collect all unknown parts on the left and all known quantities on the right hand side

$$\begin{aligned}
\Longrightarrow B^{\mathsf{T}}\Delta h - B^{\mathsf{T}}e &= B^{\mathsf{T}}c \\
B^{\mathsf{T}}e &= B^{\mathsf{T}}\Delta h - B^{\mathsf{T}}c \\
\underset{r\times m}{B^{\mathsf{T}}}\ \underset{m\times 1}{e} &= \underset{r\times 1}{B^{\mathsf{T}}y} =: w \\
w &: \text{vector of misclosures } w := B^{\mathsf{T}}y \\
y &: \text{reduced vector of observations} \\
r &: \text{number of conditions}
\end{aligned}$$

A 2: The number of conditions equals the redundancy

$$r = m - n$$

Sometimes the number of conditions can hardly be determined without knowledge on the number n of unknowns in the A-model. This will be treated later in more detail together with the so-called datum problem.

A 3:

$$\mathcal{L}_B(e,\lambda) = \frac{1}{2}\underset{1\times m}{e^{\mathsf{T}}}\ \underset{m\times 1}{e} + \underbrace{\underset{1\times r}{\lambda^{\mathsf{T}}}\ (\underset{r\times m}{B^{\mathsf{T}}}\ \underset{m\times 1}{y} - \underset{r\times m}{B^{\mathsf{T}}}\ \underset{m\times 1}{e})}_{1\times 1} \longrightarrow \min_{e,\lambda}$$

$$\begin{aligned}
\frac{\partial \mathcal{L}_B}{\partial e}(\hat{e},\hat{\lambda}) &= \underset{m\times 1}{\hat{e}} - \underset{m\times r}{B}\ \underset{r\times 1}{\hat{\lambda}} = \underset{m\times 1}{0} \\
\frac{\partial \mathcal{L}_B}{\partial \lambda}(\hat{e},\hat{\lambda}) &= -\underset{r\times m}{B^{\mathsf{T}}}\ \underset{m\times 1}{\hat{e}} + \underset{r\times m}{B^{\mathsf{T}}}\ \underset{m\times 1}{y} = \underset{r\times 1}{0} \qquad (w = B^{\mathsf{T}}y) \\
&\Longrightarrow \begin{pmatrix} \underset{m\times m}{I} & \underset{m\times r}{-B} \\ \underset{r\times m}{-B^{\mathsf{T}}} & \underset{r\times r}{0} \end{pmatrix} \begin{pmatrix} \hat{e} \\ \hat{\lambda} \end{pmatrix} = \begin{pmatrix} \underset{m\times 1}{0} \\ \underset{m\times 1}{-w} \end{pmatrix} \\
\hat{e} = B\hat{\lambda} &\Longrightarrow B^{\mathsf{T}}B\hat{\lambda} = w \\
&\Longrightarrow \hat{\lambda} = (B^{\mathsf{T}}B)^{-1}w \qquad \operatorname{rank}(B^{\mathsf{T}}B) = r \\
\Longrightarrow \hat{e} &= B(B^{\mathsf{T}}B)^{-1}w \\
&= B(B^{\mathsf{T}}B)^{-1}B^{\mathsf{T}}y \\
&= P_B y \\
\hat{y} &= y - \hat{e} \\
&= \left[I - B(B^{\mathsf{T}}B)^{-1}B^{\mathsf{T}}\right]y \\
&= P_B^{\perp}y
\end{aligned}$$

For the transition

$$
\begin{array}{rcl}
\text{parametric model} & \longleftrightarrow & \text{model of condition equations} \\
y = Ax + e & \longleftrightarrow & B^{\mathsf{T}} e = B^{\mathsf{T}} y \,,
\end{array}
$$

left multiply $y = Ax + e$ by B^{T}:

$$
B^{\mathsf{T}} y = B^{\mathsf{T}} A x + B^{\mathsf{T}} e \quad \Longleftrightarrow \quad B^{\mathsf{T}} A = 0 \,.
$$

E.g.:

$$
\underset{2\times 5}{\begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 1 & 0 \end{pmatrix}}
\underset{5\times 3}{\begin{pmatrix} -1 & 0 & 0 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 0 \end{pmatrix}}
=
\underset{2\times 3}{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}} \quad .
$$