Introduction to Least Squares Adjustment for geodetic VLBI

Matthias Schartner[a], David Mayer[b]

[a]TU Wien, Department of Geodesy and Geoinformation
[b]BEV, Federal Office of Metrology and Surveying

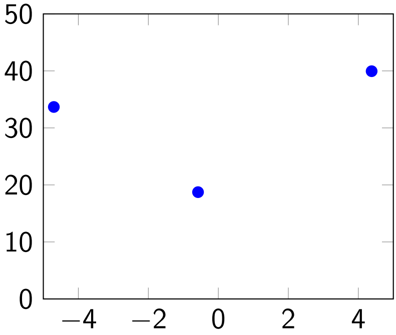
Least Squares Adjustment why?

- observation is τ (baseline)
- unknown parameters are:
  - station positions
  - EOP
  - source positions
  - troposphere...

How do we link our observations to our unknown parameters?

A very simple example

- three sample points
- data fitting
- polynome of order 2:
  $y = ax^2 + bx + c$
- observations: (x_i, y_i)
- unknowns: a, b, c



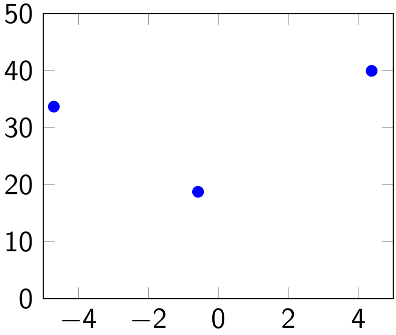
A very simple example

$(x_1, y_1)\ (x_2, y_2)\ (x_3, y_3)$

$$
\begin{aligned}
y_1 &= ax_1^2 + bx_1 + c \\
y_2 &= ax_2^2 + bx_2 + c \\
y_3 &= ax_3^2 + bx_3 + c
\end{aligned}
$$

Or in matrix notation:

$$
\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}
$$



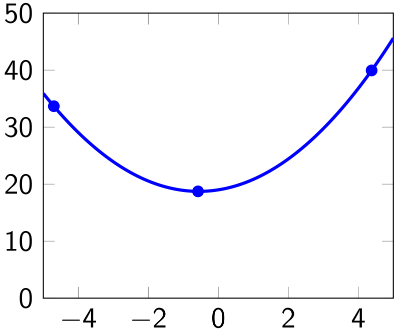
A very simple example

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

solve the equations

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{pmatrix}^{-1} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

coefficients for
$y = ax^2 + bx + c$



A very simple example

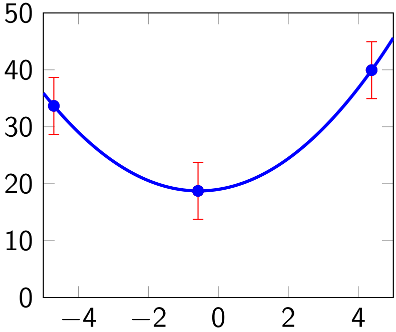
Warning!

Every observation has an error

How accurate are our coefficients a, b and c ?

What can we do?

If we have more observations we can average out the error



A very simple example

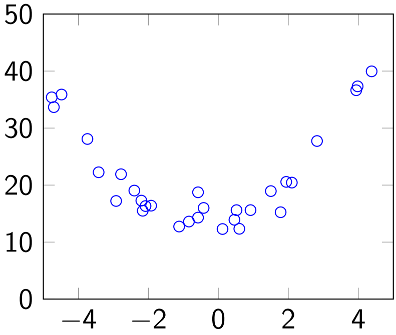
What happens if I have more observations?

$y = ax^2 + bx + c$

- three unknowns a, b, c (n_{unk})
- many observations (n_{obs})

$$n_{obs} > n_{unk}$$

How do we get our unknown parameters?



A very simple example

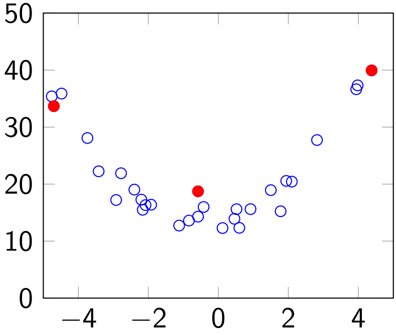
simply pick three:

$$(x_1, y_1)\ (x_2, y_2)\ (x_3, y_3)$$

$$\begin{aligned} y_1 &= ax_1^2 + bx_1 + c \\ y_2 &= ax_2^2 + bx_2 + c \\ y_3 &= ax_3^2 + bx_3 + c \end{aligned}$$

Or in matrix notation:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$



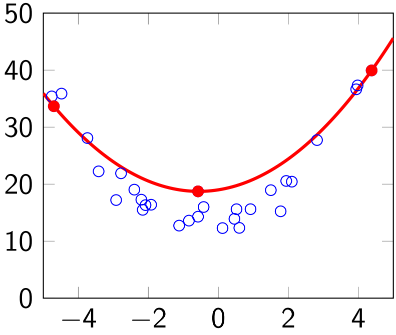
A very simple example

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

solve the equations

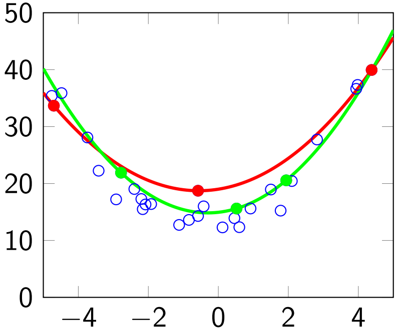
$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{pmatrix}^{-1} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

coefficients for
$y = ax^2 + bx + c$



A very simple example

Why not this three?

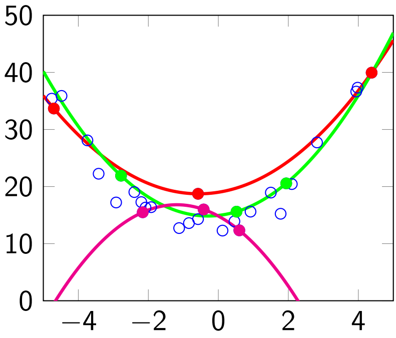


A very simple example

or this three:

Our goal:

we want to use every available observations

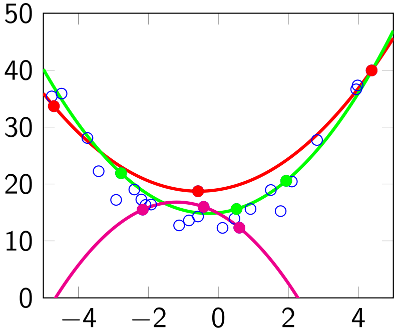


A very simple example

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{\vec{l}} = \underbrace{\begin{pmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} a \\ b \\ c \end{pmatrix}}_{\vec{x}}$$

$A \in \mathbb{R}^{n_{obs} \times n_{unk}}$
$\vec{x} \in \mathbb{R}^{n_{unk}}$
$\vec{l} \in \mathbb{R}^{n_{obs}}$

$$A\vec{x} = \vec{l} \qquad (1)$$



A very simple example

trick and solution

$$A^\top A\vec{x} = A^\top\vec{l} \quad (2)$$

$$\vec{x} = \left(A^\top A\right)^{-1} A^\top\vec{l} \quad (3)$$

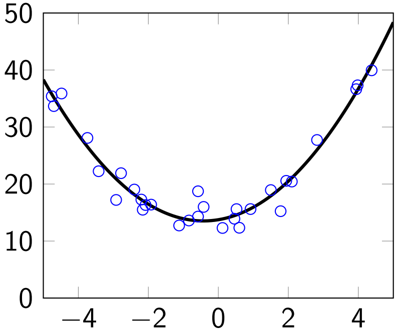
residuals v:

$$v_i = ax_i^2 + bx_i + c - y_i$$

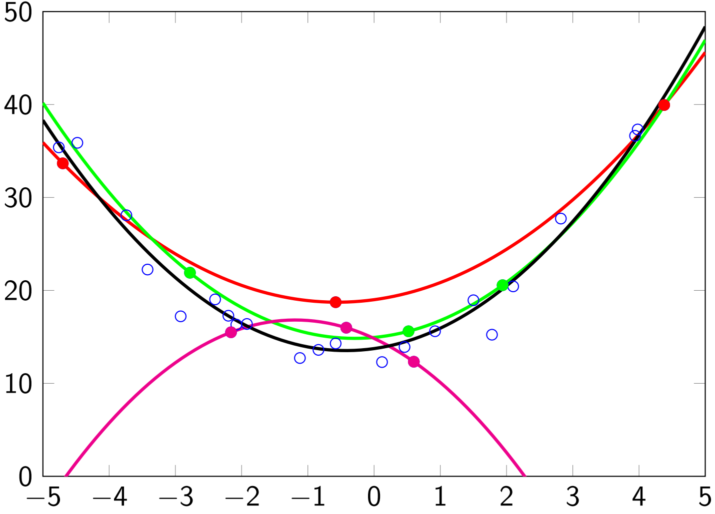
$$\vec{v} = A\vec{x} - \vec{l} \quad (4)$$

minimization

$$\sum v^2 = v^\top v \quad (5)$$



A very simple example



A very simple example

conclusion / goal

- get the most reasonable values for our unknown parameters
- take all available information

→ more observations then unknown parameters

- get some error margins

How can we do this?

Warning!

every observation has an error

Least Squares Methode

pros

- ☺ easy while powerful
- ☺ minimize residuals
- ☺ rigorous statistical approach
- ☺ individual weighting (28)
- ☺ accuracy information (32)
- ☺ manipulate geodetic datum (37)
- ☺ combine results of various LSMs (42)

cons

- ☹ only for linear functions
  - liniarization via Taylor polynome (22)
  - → a priori values
  - → iterations (31)
- ☹ we need to model our observations $o - c$
- ☹ outliers

requirements for LSM

- partial derivatives (24)

What do we want to estimate?

- station coordinates
- source coordinates
- earth orientation parameters
- tropospheric parameters
- clock behavior
- and many more ...

Many unknown parameters...

What do we want to estimate?

Example: Scan with 5 stations $\rightarrow \frac{5 \cdot 4}{2} = 10$ observations

Estimates:

- Troposphere: (15p.)
    - zenith wet delay (5p.)
    - north gradient (5p.)
    - east gradient (5p.)
- clock (5p.)
- ...

$\sum$ 20+ Parameters

Way more unknown parameters then observations ☹

Requirement for LSM

We need more observations then unknowns! $n_{obs} > n_{unk}$

What do we want to estimate?

Solution

piecewise linear offsets (34)

- we estimate parameters only every x minutes
- linear interpolation in between

→ reduces number of estimated parameters!

Improvement

constraints (35)

- $zwd_{sta1,03:00} = zwd_{sta1,04:00} \underbrace{\pm 1.5\ [\mathrm{cm}]}_{constraint}$

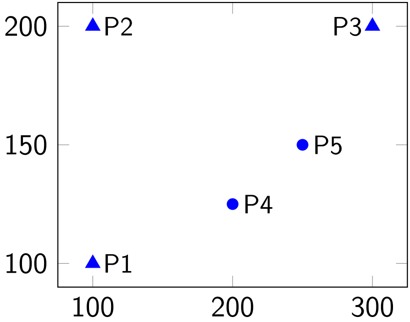
Conclusion

What should you remeber?

- we calculate our solutions using Least Squares Adjustment
- we minimize (squared sum of) residuals
- we need to model our observations! → theoretical delay $o - c$
- we need partial derivatives for our Jacobian matrix A
- we need a priori values X_0 (usually not a big deal)
- we estimate additions x to a priori values X_0
- we need to eliminate outliers

Example 2

- 5 stations
- 3 with known coordinates
- 2 with unknown coordinates



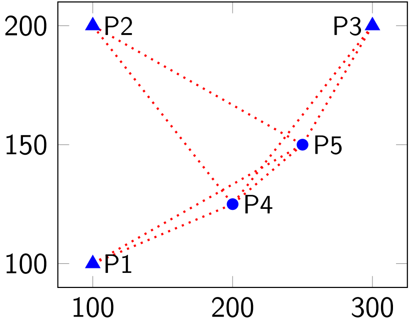
Example 2

measured distance s_{ij}

- unknowns:
  $(x4, y4)$ $(x5, y5)$
  $\rightarrow n_{unk} = 4$
- observations: s_{ij}
  $\rightarrow n_{obs} = 7$

$$n_{obs} > n_{unk}$$

$A \in \mathbb{R}^{7\times4}$
$\vec{x} \in \mathbb{R}^4$
$\vec{l} \in \mathbb{R}^7$



Some basics about LSM linearization

We need to model our system:

$$s_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$$

not linear → Taylor expansion:

$$\begin{aligned} T_{\infty,x_0}(x) &= f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots \\ &= \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \end{aligned}$$

→ liniarized observation equation

What do we need?

- We need partial derivative
- We need a priori values $\vec{x}_0$

Some basics about LSM definitions

- observations: L
- (computed) approximated observations: L_0
- reduced observation vector (observed minus computed): l

$$l = L - L_0 \tag{6}$$

- estimated unknowns: $\hat{X}$
- approximated unknowns: X_0
- reduced unknowns: x

$$x = \hat{X} - X_0 \tag{7}$$

- Designmatrix or Jacobian matrix: A (partials)

Back to Example 2

unknown: $x_4,\ y_4,\ x_5,\ y_5$
known: $x_1,\ y_1,\ x_2,\ y_2,\ x_3,\ y_3$
approximate values: $x_{4,0},\ y_{4,0},\ x_{5,0},\ y_{5,0}$

$$s_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$$

$$A = \begin{pmatrix} \left(\frac{\partial s_{14}}{\partial x_4}\right) & \left(\frac{\partial s_{14}}{\partial y_4}\right) & \left(\frac{\partial s_{14}}{\partial x_5}\right) & \left(\frac{\partial s_{14}}{\partial y_5}\right) \\ \left(\frac{\partial s_{15}}{\partial x_4}\right) & \left(\frac{\partial s_{15}}{\partial y_4}\right) & \left(\frac{\partial s_{15}}{\partial x_5}\right) & \left(\frac{\partial s_{15}}{\partial y_5}\right) \\ \vdots & \vdots & \vdots & \vdots \\ \left(\frac{\partial s_{45}}{\partial x_4}\right) & \left(\frac{\partial s_{45}}{\partial y_4}\right) & \left(\frac{\partial s_{45}}{\partial x_5}\right) & \left(\frac{\partial s_{45}}{\partial y_5}\right) \end{pmatrix}$$

$$\frac{\partial s_{14}}{\partial x_4} = \frac{x_{4,0} - x_1}{\sqrt{\left(x_{4,0} - x_1\right)^2 + \left(y_{4,0} - y_1\right)^2}} \quad \frac{\partial s_{14}}{\partial y_4} = \frac{y_{4,0} - y_1}{\sqrt{\left(x_{4,0} - x_1\right)^2 + \left(y_{4,0} - y_1\right)^2}} \quad \frac{\partial s_{14}}{\partial x_5} = \frac{\partial s_{14}}{\partial y_5} = 0$$

Example 2

unknown: $x_4, \; y_4, \; x_5, \; y_5$
known: $x_1, \; y_1, \; x_2, \; y_2, \; x_3, \; y_3$
approximate values: $x_{4,0}, \; y_{4,0}, \; x_{5,0}, \; y_{5,0}$

reduced observation vector (observed minus computed)
$l = L - L_0$

$$l = \begin{pmatrix} s_{14} - s_{14,0} \\ s_{15} - s_{15,0} \\ \vdots \\ s_{45} - s_{45,0} \end{pmatrix}$$

$$s_{14,0} = \sqrt{(x_{4,0} - x_1)^2 + (y_{4,0} - y_1)^2}$$

Example 2

reduced observation vector (observed minus computed)
$l = L - L_0$
Jacobian matrix (partials): A
reduced unknowns: $x = X - X_0$

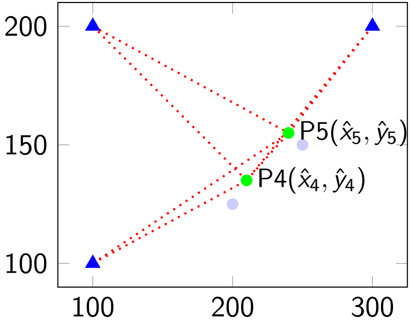
$$Ax = l \xrightarrow{trick} A^\top Ax = A^\top l \rightarrow$$
$$x = \left(A^\top A\right)^{-1} A^\top l \tag{8}$$

What do we estimate?

We don't estimate whole unknown parameter $\hat{X}$ but only additions x to some a priori values X_0

Example 2

The coordinates of our unknown points change



Some more basics about LSM covariance matrix

usually not every observations is equally accurate:
Covariance matrix $Q_{ll} \in \mathbb{R}^{n_{obs} \times n_{obs}}$, Weight matrix P:

$$Q_{ll} = \begin{pmatrix} \sigma_{l_1}^2 & \sigma_{l_1,l_2} & \dots & \sigma_{l_1,l_n} \\ \sigma_{l_1,l_2} & \sigma_{l_2}^2 & \dots & \sigma_{l_2,l_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{l_1,l_n} & \sigma_{l_2,l_n} & \dots & \sigma_{l_n}^2 \end{pmatrix} \tag{9}$$

if observations are independent $\sigma_{l_i,l_j} = 0$

$$Q_{ll} = \begin{pmatrix} \sigma_{l_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{l_2}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{l_n}^2 \end{pmatrix} \quad P = \underbrace{\frac{1}{s_0^2}}_{usually 1} Q_{ll}^{-1}$$

Some more basics about LSM covariance matrix

Variance-Covariance matrix Q_{ll}, Weight matrix $P \in \mathbb{R}^{n_{obs} \times n_{obs}}$:

$$P = \underbrace{\frac{1}{s_0^2}}_{usually 1} Q_{ll}^{-1} = \begin{pmatrix} \frac{1}{\sigma_{l_1}^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_{l_2}^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sigma_{l_n}^2} \end{pmatrix} \quad (10)$$

good observation $\rightarrow$ small standard deviation σ_{l_i} and variance $\sigma_{l_i}^2$ $\rightarrow$ high weight in P matrix

Some more basics about LSM equations

Final formulas:

$$\underbrace{A^\top PA}_{N} x = \underbrace{A^\top Pl}_{b} \tag{11}$$

$$x = N^{-1} b \tag{12}$$

Important for LSM

- Normal equation matrix $N = \left(A^\top PA\right) \quad N \in \mathbb{R}^{n_{unk} \times n_{unk}}$
- right hand side $b = A^\top Pl \quad b \in \mathbb{R}^{n_{unk}}$
- we minimize $v^\top Pv$

$$\hat{X} = X_0 + \left(A^\top PA\right)^{-1} A^\top Pl = X_0 + N^{-1} b \tag{13}$$

Some more basics about LSM iteration

Sometimes a iteration is necessary.

- inaccurate a priori values X_0
- numerical reasons

You can use $\hat{X}$ again as new a priori values X_0 and redo everything

$$X_{0,n+1} = \hat{X}_n = X_{0,n} + N_n^{-1} b_n \tag{14}$$

How good are our estimates? covariance matrix

a posteriori variance factor σ_0^2:

$$\sigma_0^2 = \frac{v^\top P v}{n_{obs} - n_{unk}} \tag{15}$$

remember our residuals v are: $v = Ax - l$

Variance-Covariance Matrix $Q_{xx} \in \mathbb{R}^{n_{unk} \times n_{unk}}$

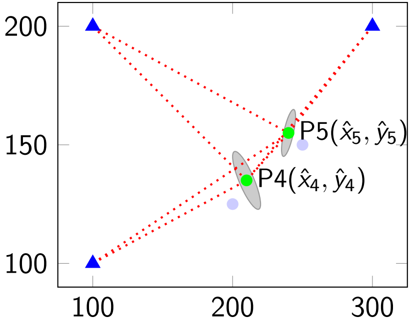
$$Q_{xx} = \sigma_0^2 N^{-1} = \begin{pmatrix} \sigma_{x_1}^2 & \sigma_{x_1,x_2} & \dots & \sigma_{x_1,x_n} \\ \sigma_{x_1,x_2} & \sigma_{x_2}^2 & \dots & \sigma_{x_2,x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{x_1,x_n} & \sigma_{x_2,x_n} & \dots & \sigma_{x_n}^2 \end{pmatrix} \tag{16}$$

$\sigma_{x_i}^2$ is variance of unknown parameter x_i

σ_{x_i} is standard deviation

Back to example 2

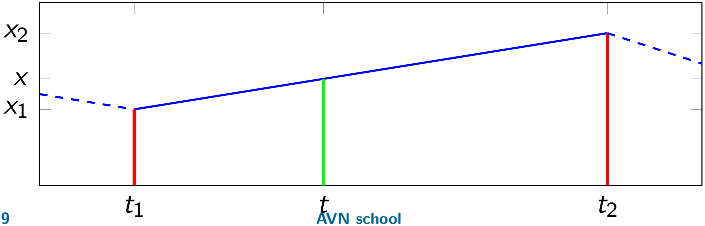
Now we know how accurate our coordinates are



Concepts What are piecewise linear offsets?

time dependent parameters
Example: Troposphere (mf = mapping function)

$$\Delta L_w(t) = mf_w(t)x_1 + mf_w(t)\frac{t-t_1}{t_2-t_1}(x_2-x_1)$$
$$\frac{\partial \Delta L_w}{\partial x_1} = mf_w(t) - mf_w(t)\frac{t-t_1}{t_2-t_1}$$
$$\frac{\partial \Delta L_w}{\partial x_2} = mf_w(t)\frac{t-t_1}{t_2-t_1}$$



Concepts
What are constraints?

- pseudo observations
- help us with singularities
- useful concept

A simple example: (think of troposphere between two time epochs)

$$\begin{aligned} x_j - x_i &= 0 \pm 5 \\ x_{j-1} - x_{i-1} &= 0 \pm 3 \end{aligned}$$

($\pm$ means standard deviation)

$A_c \in \mathbb{R}^{n_{const} \times n_{unk}}$

$$A_c = \begin{pmatrix} \dots & 0 & -1 & 0 & \dots & 0 & 1 & 0 & \dots \\ \dots & -1 & 0 & 0 & \dots & 1 & 0 & 0 & \dots \end{pmatrix}$$

$l_c \in \mathbb{R}^{n_{const}}$

$P_c \in \mathbb{R}^{n_{const} \times n_{const}}$

$$l_c = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$P_c = \begin{pmatrix} \frac{1}{5^2} & 0 \\ 0 & \frac{1}{3^2} \end{pmatrix}$$

Concepts

What are constraints?

$N_c \in \mathbb{R}^{n_{unk} \times n_{unk}}$

$$N_c = A_c^\top P_c A_c \tag{17}$$

$b_c \in \mathbb{R}^{n_{const}}$

$$b_c = A_c^\top P_c l_c \tag{18}$$

Combination with the real observations:

$$N_{total} = A^\top PA + A_c^\top P_c A_c = N + N_c \tag{19}$$

$$b_{total} = A^\top Pl + A_c^\top P_c l_c = b + b_c \tag{20}$$

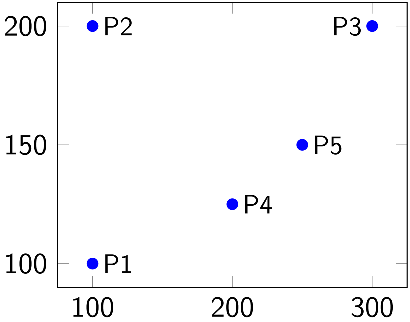
The only thing that changes is:

$$\sigma^2_{0,const} = \frac{v^\top Pv + v_c^\top P_c v_c}{n_{obs} + n_{const} - n_{unk}} \tag{21}$$

Concepts

What is a geodetic datum? - Example 2

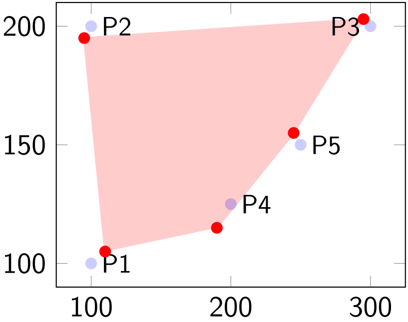
- Think of example 2
- This time no station has fixed coordinates
- unknowns: all coordinates
- observations: s_{ij}
- we have a priori coordinates



Concepts What is a geodetic datum?

- we know inner geometry very well
- we can not estimate absolute coordinates
- N matrix is singular

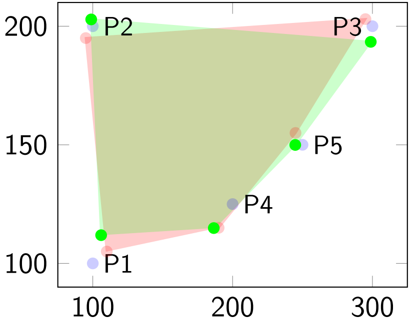
is this correct?



Concepts

What is a geodetic datum? - Example 2

- we know inner geometry very well
- we can not estimate absolute coordinates
- N matrix is singular



or this?

Concepts

What is a geodetic datum? - Formulas

We need datum stations with additional conditions
for example in this case we could say:

translation in x	$\sum dx_i = 0$
translation in y	$\sum dy_i = 0$
rotation around z	$\sum (y_i dx_i - x_i dy_i) = 0$
scale	$\sum (x_i dx_i - y_i dy_i) = 0$

This would lead to the following matrix: $G \in \mathbb{R}^{n_{datum} \times n_{unk}}$

$$G = \begin{pmatrix} 1 & 0 & \dots & 1 & 0 \\ 0 & 1 & \dots & 0 & 1 \\ y_1 & -x_1 & \dots & y_n & -x_n \\ x_1 & y_1 & \dots & x_n & y_n \end{pmatrix} \tag{22}$$

Concepts — What is a geodetic datum? - Solution

The solution is now:

$$\begin{pmatrix} A^\top PA & G \\ G^\top & 0 \end{pmatrix} \begin{pmatrix} x \\ k \end{pmatrix} = \begin{pmatrix} A^\top Pl \\ 0 \end{pmatrix} \tag{23}$$

What else is different:

$$\sigma_0^2 = \frac{v^\top Pv}{n_{obs} + n_{datum} - n_{unk}}; \qquad \begin{pmatrix} Q_{xx} & Q_{xk} \\ Q_{xk} & Q_{kk} \end{pmatrix} = \sigma_0^2 \begin{pmatrix} A^\top PA & G \\ G^\top & 0 \end{pmatrix}^{-1} \tag{24}$$

Geodetic datum in VLBI

usually we use NNT/NNR - No Net Rotation and No Net Translation conditions (3 NNT and 3 NNR) scale is fixed by observations

Concepts

What is a global solution?

- combination of large number of sessions to estimate accurate common parameters
- combination is done at the normal equation level
- you need to make sure N matrix has same structure
  $\rightarrow$ you want to get rid of some parameters

$$Nx = b \quad \rightarrow \quad \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

$$\underbrace{\left(N_{11} - N_{12} N_{22}^{-1} N_{21}\right)}_{N_{reduc}} x_1 = \underbrace{b_1 - N_{12} N_{22}^{-1} b_2}_{b_{reduc}} \qquad (25)$$

$$N_{REDUC} = N_{reduc_1} + N_{reduc_2} + \ldots + N_{reduc_n} \qquad (26)$$

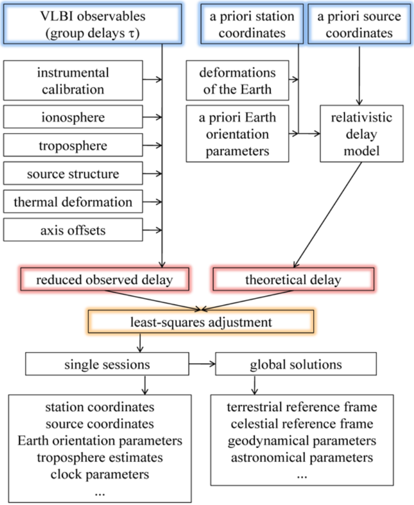
$$b_{REDUC} = b_{reduc_1} + b_{reduc_2} + \ldots + b_{reduc_n} \qquad (27)$$

Conclusion

What is to remeber?

- we calculate our solutions using Least Squares Adjustment
- we minimize $v^\top P v$
- we need to model our observations! $\rightarrow$ theoretical delay
- we need partial derivatives for our Jacobian matrix A
- we need a priori values X_0 (usually not a big deal)
- we estimate additions x to a priori values X_0

Link to VLBI


VLBI observables
(group delays τ)
a priori station coordinates
a priori source coordinates
instrumental calibration
ionosphere
troposphere
source structure
thermal deformation
axis offsets
deformations of the Earth
a priori Earth orientation parameters
relativistic delay model
reduced observed delay
theoretical delay
least-squares adjustment
single sessions
global solutions
station coordinates
source coordinates
Earth orientation parameters
troposphere estimates
clock parameters
...
terrestrial reference frame
celestial reference frame
geodynamical parameters
astronomical parameters
...

Lecture LSM

Matthias Schartner[a], matthias.schartner@geo.tuwien.ac.at
David Mayer[b], david.mayer@bev.gv.at

[a]TU Wien, Department of Geodesy and Geoinformation
[b]BEV, Federal Office of Metrology and Surveying