22. EVGA Meeting 2015

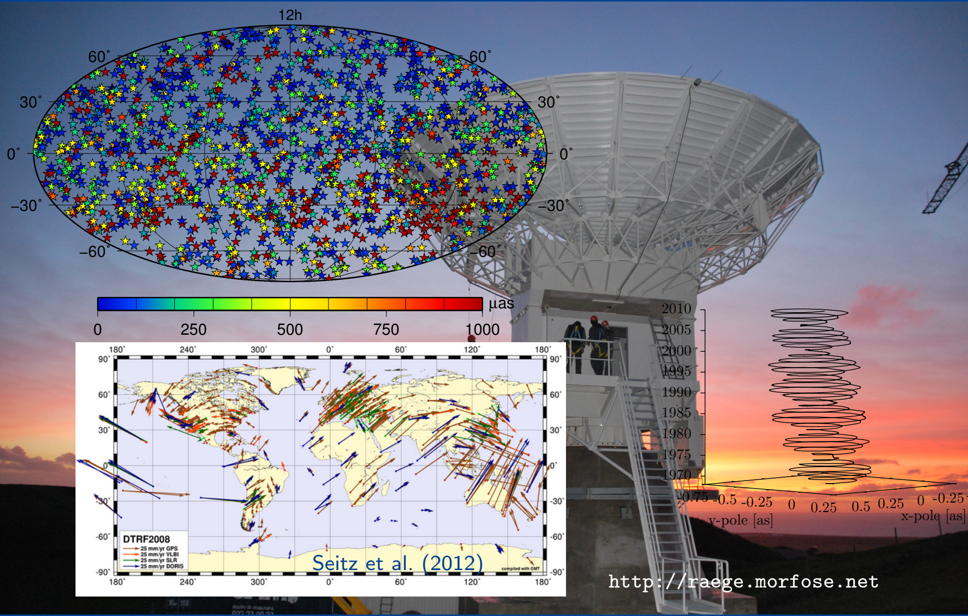
Numerical Issues of VLBI Data Analysis

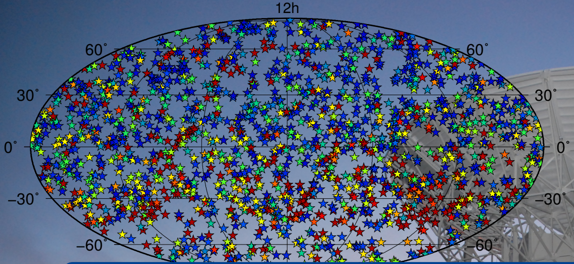
T. Artz S. Halsig A. Iddink A. Nothnagel

Institut für Geodäsie und Geoinformation, Universität Bonn

2015/05/19

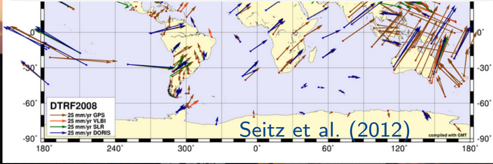
VLBI

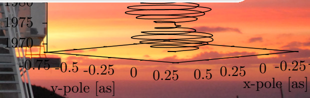

12h
60°
30°
0°
−30°
−60°
µas
0
250
500
750
1000
DTRF2008
25 mm/yr GPS
25 mm/yr VLBI
25 mm/yr SLR
25 mm/yr DORIS
Seitz et al. (2012)
2010
2005
2000
1995
1990
1985
1980
1975
1970
y-pole [as]
x-pole [as]
http://raege.morfose.net



Parameter Estimation

Least squares adjustment:
gradient of $\|\mathbf{Ax}-\mathbf{b}\|_2^2 = \langle \mathbf{Ax}-\mathbf{b}\rangle$ vanishes
$\Rightarrow 0 = \mathbf{A}^T\mathbf{r} = \mathbf{A}^T\mathbf{Ax} - \mathbf{A}^T\mathbf{b}$

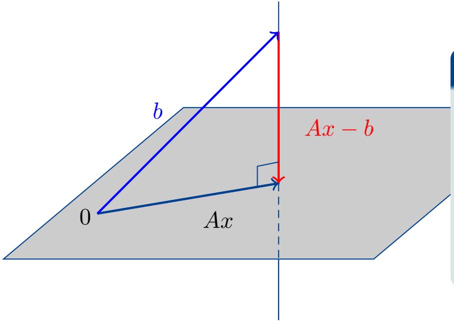




http://raege.morfose.net

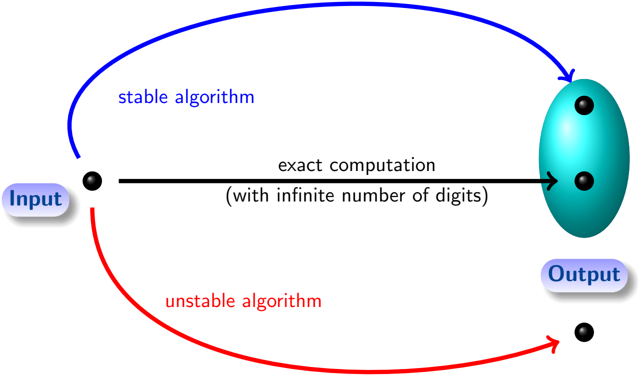
Parameter Estimation

Least squares adjustment:
gradient of $\|\mathbf{Ax}-\mathbf{b}\|_2^2 = \langle \mathbf{Ax}-\mathbf{b}\rangle$ vanishes
$\Rightarrow 0 = \mathbf{A}^T\mathbf{r} = \mathbf{A}^T\mathbf{Ax} - \mathbf{A}^T\mathbf{b}$



Numerical Issues

- Condition of the Jacobian matrix
  $\Rightarrow$ orientation of the plane
- stability of the algorithm
  $\Rightarrow$ precision of mapping



Stability of the Algorithm

If an algorithm is numerically unstable, at a given point, the errors do not remain bounded and tend to grow up in an uncontrolled way corrupting completely the final result.

LS Solution Algorithms

Normal Equations

$$\mathbf{N} = \mathbf{A}^T\mathbf{A}, \;\; \mathbf{n} = \mathbf{A}^T\mathbf{b}$$

$$\mathbf{x} = \mathbf{N}^{-1}\mathbf{n}$$

$\mathbf{N}$: symmetric & positive definite

$\Rightarrow$ Cholesky decomposition

$$\mathbf{A} = \mathbf{L}^T\mathbf{L}$$

and back solution to determine $\mathbf{x}$.

QR Decomposition

Decompose $\mathbf{A}$ in orthogonal matrix $\mathbf{Q}$ and upper tri-angular matrix $\mathbf{R}$

$$\mathbf{A} = \mathbf{QR} \;\; \Rightarrow \mathbf{x} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b}$$

SVD

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T.$$

$\Rightarrow$ solution of $min_{\mathbf{x}}\|\mathbf{A}\cdot\mathbf{x} - \mathbf{b}\|_2$:

$$\mathbf{x} = \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{U}^T\mathbf{b}.$$

LS Solution Algorithms

Normal Equations

$$\mathbf{N} = \mathbf{A}^T\mathbf{A}, \quad \mathbf{n} = \mathbf{A}^T\mathbf{b}$$

fast, but least accurate

$$\mathbf{x} = \mathbf{N}^{-1}\mathbf{n}$$

$\mathbf{N}$: symmetric & positive definite
$\Rightarrow$ Cholesky decomposition

$$\mathbf{A} = \mathbf{L}^T\mathbf{L}$$

and back solution to determine $\mathbf{x}$.

QR Decomposition

Decompose $\mathbf{A}$ in orthogonal matrix $\mathbf{Q}$ and upper tri-angular matrix $\mathbf{R}$

costs up to twice as NEQ

$$\mathbf{A} = \mathbf{QR} \quad \Rightarrow \mathbf{x} = \mathbf{R}^{-1}\mathbf{Q}^T\mathbf{b}$$

SVD

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$$

for ill-conditioned systems;
$\mathbf{A}$ is singular

$\Rightarrow$ solution of $min_{\mathbf{x}} \|\mathbf{A}\cdot\mathbf{x} - \mathbf{b}\|_2$:

$$\mathbf{x} = \mathbf{V}\Sigma^{-1}\mathbf{U}^T\mathbf{b}.$$

Weighted Solution

$$\mathbf{N} = \mathbf{A}^T\mathbf{W}\mathbf{A} = \mathbf{A}\Sigma_{bb}^{-1}\mathbf{A}^T$$

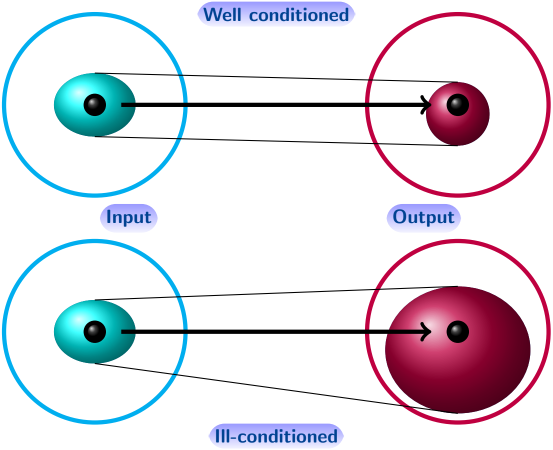
$$\mathbf{n} = \mathbf{A}^T\mathbf{W}\mathbf{b}$$

De-correlation

$$\mathbf{W} = \mathbf{L}^T\mathbf{L}$$

$$\widetilde{\mathbf{A}} = \mathbf{L}\mathbf{A}, \widetilde{\mathbf{b}} = \mathbf{L}\mathbf{b}$$

$$\begin{aligned}\mathbf{x} &= \left(\widetilde{\mathbf{A}}^T\mathbf{E}\widetilde{\mathbf{A}}\right)^{-1}\widetilde{\mathbf{A}}^T\mathbf{E}\widetilde{\mathbf{b}} \\ &= \left(\mathbf{A}^T\mathbf{L}^T\mathbf{L}\mathbf{A}^T\right)^{-1}\mathbf{A}^T\mathbf{L}^T\mathbf{L}\mathbf{b} \\ &= \left(\mathbf{A}^T\mathbf{W}\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{W}\mathbf{b}\end{aligned}$$

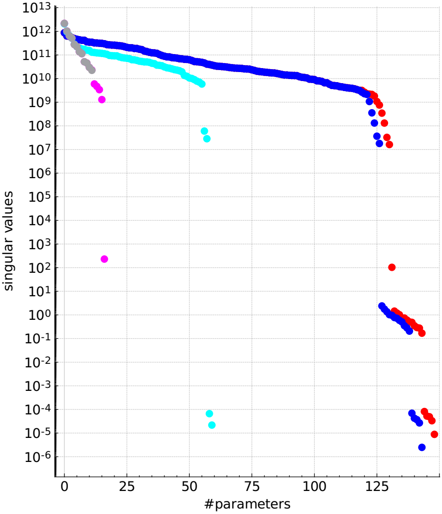


Well Conditioned

a small perturbation of the input data leads to small variations of the results

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

$$\kappa(\mathbf{A}) = \frac{\sigma_{max}}{\sigma_{min}}$$



CL	ZWD	GR	ERP	NUT
d2	d1	-	-	-
d2	d1	-	d0	d0
d2 c1	d1	-	-	-
d2 c1	c1	c24	-	-
d2 c1	c1	c24	d0	d0

d: polynomial of degree;

c: CPWLF with interval length [h]

$\kappa(\mathbf{A})$	rank
$9.7 \cdot 10^{1}$	
$9.4 \cdot 10^{9}$	
$9.6 \cdot 10^{16}$	
$3.5 \cdot 10^{17}$	
$9.8 \cdot 10^{16}$	

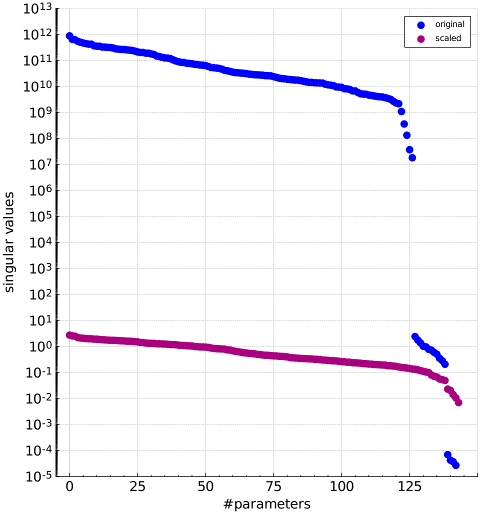
Scaling

$$\hat{\mathbf{A}} = \widetilde{\mathbf{A}}\mathbf{F}$$

$$\mathbf{F} = diag\left(\frac{1}{\sqrt{diag\left(\widetilde{\mathbf{A}}^T\widetilde{\mathbf{A}}\right)}}\right)$$

$\Rightarrow$ only 1 on main diagonal of $\hat{\mathbf{N}}$

$\kappa(\mathbf{A})$	rank
$3.5 \cdot 10^{17}$	
$3.9 \cdot 10^{2}$	



Change $\kappa(\mathbf{A})$ – Scaling

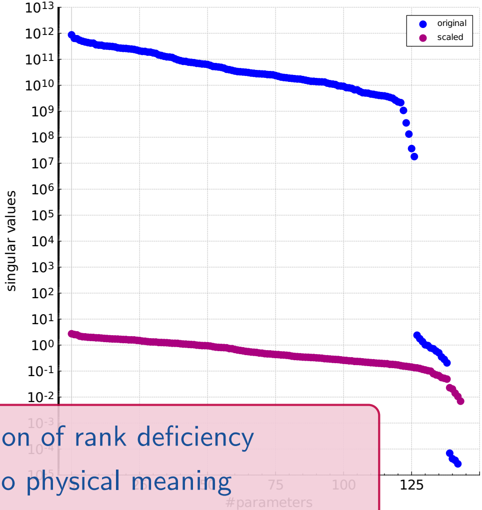
Scaling

$$\hat{\mathbf{A}} = \widetilde{\mathbf{A}}\mathbf{F}$$

$$\mathbf{F} = diag\left(\frac{1}{\sqrt{diag\left(\widetilde{\mathbf{A}}^T\widetilde{\mathbf{A}}\right)}}\right)$$

$\Rightarrow$ only 1 on main diagonal of $\hat{\mathbf{N}}$

$\kappa(\mathbf{A})$	rank
$3.5 \cdot 10^{17}$	
$3.9 \cdot 10^{2}$	

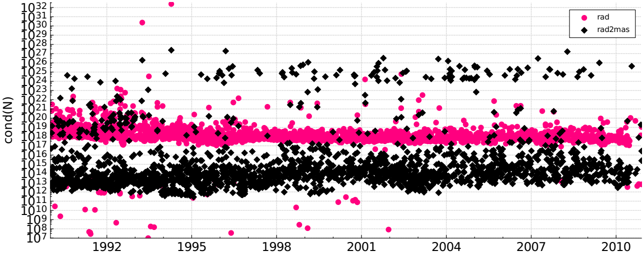


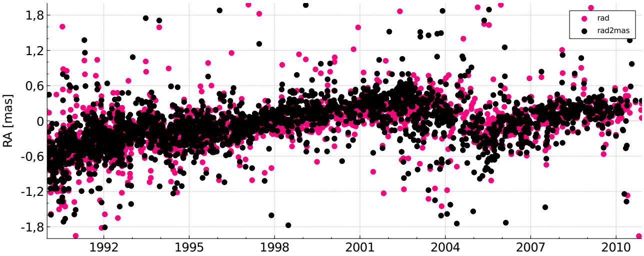
- elimination of rank deficiency
- units w/o physical meaning
- estimates have to be re-scaled: $\mathbf{x} = \hat{\mathbf{x}}\mathbf{F}$

Change $\kappa(\mathbf{A})$ – Units

Scaling

- gsf2014a
- SINEX: rad
- changed to mas





- $\kappa(\mathbf{N})$ typically improves
- different results
- same noise level

Constraining

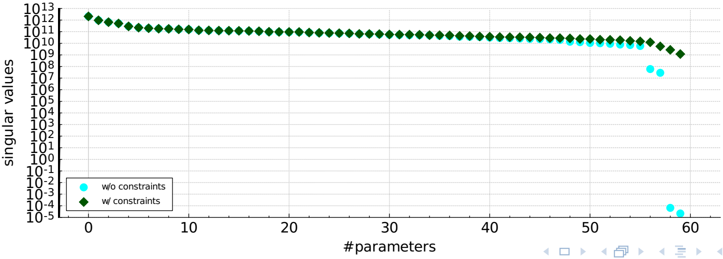
Add information, e.g., rate between two subsequent ZWD is zero with a given σ_{Bi}

$$\mathbf{B} = \begin{pmatrix} 0 & \dots & 1 & -1 & 0 & \dots \end{pmatrix},$$

$$\mathbf{W}_B = diag(1/{\sigma_{Bi}}^2)$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}, \qquad \mathbf{W} = \begin{pmatrix} \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_B \end{pmatrix}$$

$\kappa(\mathbf{A})$	rank
$9.6 \cdot 10^{16}$	
$1.7 \cdot 10^{3}$	



Constraining

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

$$\mathbf{\Sigma} = diag\left(\sigma_1 \quad \sigma_2 \quad \dots \quad \sigma_r \quad 0 \quad \dots \quad 0\right)$$

$$\mathbf{U} = \Big(\underbrace{u_1 \quad u_2 \quad \dots \quad u_r}_{\mathbf{U}_r} \quad \underbrace{u_{r+1} \quad \dots \quad u_m}_{\mathbf{U}_0} \Big)$$

$$\mathbf{V} = \Big(\underbrace{v_1 \quad v_2 \quad \dots \quad v_r}_{\mathbf{V}_r} \quad \underbrace{v_{r+1} \quad \dots \quad v_n}_{\mathbf{V}_0} \Big)$$

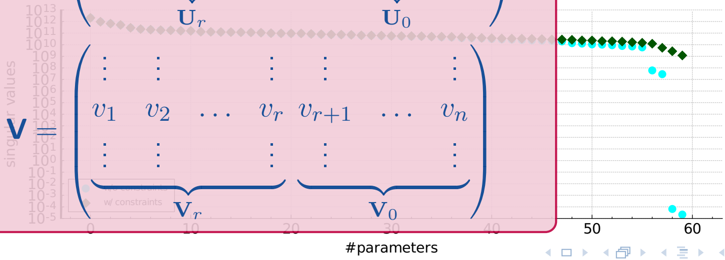
Add information, e.g., rate between two subsequent ZWD is zero with a given σ_{Bi}

$$\mathbf{B} = \left(0 \ \dots \ 1 \ -1 \ 0 \ \dots\right);$$

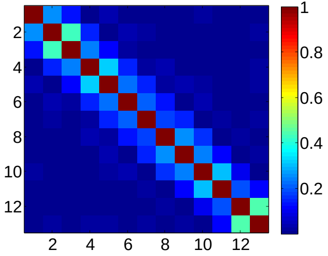
$$\mathbf{W}_B = diag(1/\sigma_{Bi}{}^2)$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix}, \quad \mathbf{W} = \begin{pmatrix} \mathbf{W} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_B \end{pmatrix}$$

$\kappa(\mathbf{A})$	rank
$9.6 \cdot 10^{16}$	
$1.7 \cdot 10^{3}$	

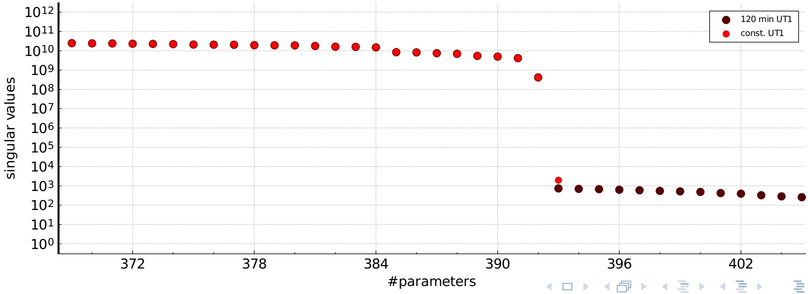


UT1 at Finer Intervals



Analysis Workshop

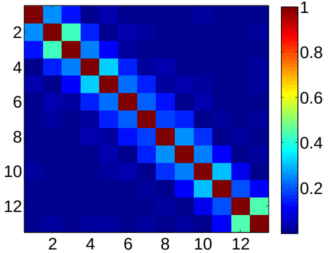
What happens if we solve for UT1 at finer intervals?

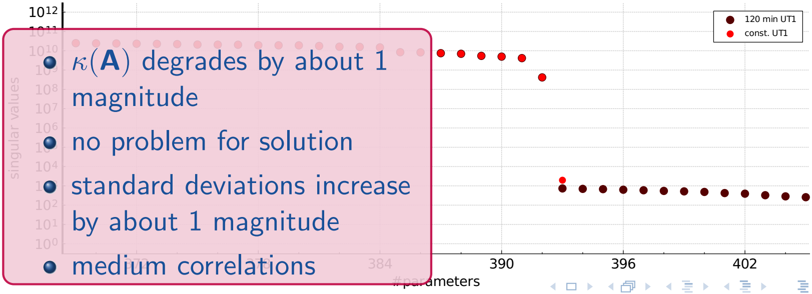


UT1 at Finer Intervals

Analysis Workshop

What happens if we solve for UT1 at finer intervals?





- $\kappa(\mathbf{A})$ degrades by about 1 magnitude
- no problem for solution
- standard deviations increase by about 1 magnitude
- medium correlations

Conclusions & Outlook

- equation system ill-conditioned
- simple options to improve condition
- different results with different modifications of $\kappa(\mathbf{A})$
- which is the correct approach?

Conclusions & Outlook

- equation system ill-conditioned
- simple options to improve condition
- different results with different modifications of $\kappa(\mathbf{A})$
- which is the correct approach?

possible next steps

- optimal selection of constraints
- investigation of parameterization options
- validation of algorithms

Thank you!

artz@igg.uni-bonn.de