Investigation of non-linear least squares problems using the example of circle fitting

Anastasia Pasioti, Frank Neitzel, Svetozar Petrovic

Institute of Geodesy and Geoinformation Science, TU Berlin

Geodätische Woche 15.-17.09.2015, Stuttgart

Motivation

Non-linear least squares problems:

- occur more frequently in Geodesy than linear problems do
- direct solution usually is not possible

Techniques for solving non-linear least squares problems:

- In the geodetic literature: Gauss-Newton (GN)
- Alternative techniques:
  - Levenberg-Marquardt (LM)
  - Newton-Raphson (NR)
  - Lapaine-Neitzel-Petrovic (LNP)

All techniques:
Iterative computation using appropriate initial values

Some questions arise:

- How sensitive are the techniques with respect to the initial values?
- Does the distribution of the points in a 2D plane affect the solution?

Gauss-Newton (GN) with condition equations

Technische Universität Berlin

Condition equations

$$\underbrace{\sqrt{((x_i + v_{x_i}) - x_0)^2 + ((y_i + v_{y_i}) - y_0)^2} - R}_{\psi_i(\mathbf{L}+\mathbf{v},\mathbf{X})} = 0 \qquad i = 1, 2, \ldots, n$$

Linearization and rigorous solution within the Gauss-Helmert model

$$\begin{bmatrix} \mathbf{k} \\ (\mathbf{X} - \mathbf{X}^0) \end{bmatrix} = \begin{bmatrix} \mathbf{B}\mathbf{Q}_{\mathbf{LL}}\mathbf{B}^{\mathrm{T}} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & 0 \end{bmatrix}^{-1} \begin{bmatrix} -\mathbf{w} \\ 0 \end{bmatrix}$$

Unknowns from

$$\hat{\mathbf{X}} = \mathbf{X}^0 + \hat{\mathbf{x}}$$

Residuals from

$$\mathbf{v} = \mathbf{Q}_{\mathbf{LL}}\mathbf{B}^{\mathrm{T}}\mathbf{k}$$

Gauss-Newton (GN) with observation equations

Observation equations

$$\underbrace{\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} - R}_{\psi_i(\mathbf{X})} = 0 + v_i \qquad i = 1, 2, \ldots, n$$

Linearization and solution within the Gauss-Markov model

Normal equations

$$\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}$$

$$\hat{\mathbf{x}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}$$

Unknowns from

$$\hat{\mathbf{X}} = \mathbf{X}^0 + \hat{\mathbf{x}}$$

Levenberg-Marquardt (LM)

Technische Universität Berlin

Now: Normal equations with damping parameter $\mu \geq 0$

$$\left(\underbrace{\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}}_{\mathbf{N}} + \mu \mathbf{I} \right) \hat{\mathbf{x}} = \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{l}$$

Initial value (MADSEN et al. 2004)

$$\mu^0 = \tau \cdot \max_i \left\{ n_{ii}^0 \right\}$$

procedure UPDATE μ
    if $\rho > 0$ **then**
        $\mu = \mu \cdot \max\left\{ \frac{1}{3}; 1 - (2\rho - 1)^3 \right\}$
    else
        $\mu = \mu \cdot 2$
    end if
end procedure

Gain ratio (REUSKEN 2015)

$$\rho = \frac{\left\| \mathbf{l}\left(\mathbf{X}^0\right) \right\|_2^2 - \left\| \mathbf{l}\left(\mathbf{X}^0 + \hat{\mathbf{x}}\right) \right\|_2^2}{\left\| \mathbf{l}\left(\mathbf{X}^0\right) \right\|_2^2 - \left\| \mathbf{l}\left(\mathbf{X}^0\right) + \mathbf{A}\left(\mathbf{X}^0\right)\hat{\mathbf{x}} \right\|_2^2}$$

Newton-Raphson (NR)

Observation equations

$$\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} - R = 0 + v_i$$

System of equations/non-linear normal equations

$$\Omega'\left(x_j\right) = 0 \qquad j = 1,2,\ldots,u$$

Linearization of system of equations/normal equations

$$\mathbf{\Omega}'(\mathbf{X}) = \mathbf{\Omega}'\left(\mathbf{X}^0\right) + \mathbf{\Omega}''\left(\mathbf{X}^0\right)\hat{\mathbf{x}} = 0 \quad \text{and} \quad \hat{\mathbf{x}} = \mathbf{X} - \mathbf{X}^0$$

with $\mathbf{H}\left(\mathbf{X}^0\right) = \mathbf{\Omega}''\left(\mathbf{X}^0\right)$

Solution from $\hat{\mathbf{x}} = -\mathbf{H}\left(\mathbf{X}^0\right)^{-1}\mathbf{\Omega}'\left(\mathbf{X}^0\right)$

Unknowns from $\hat{\mathbf{X}} = \mathbf{X}^0 + \hat{\mathbf{x}}$

Lapaine-Neitzel-Petrovic (LNP)

Technische Universität Berlin

Observation equations

$$r_i - R = 0 + v_i \qquad i = 1, 2, \ldots, n$$

with $$r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$$

Normal equations

$$\sum_{i=1}^{n} (r_i - R) \frac{(x_i - x_0)}{r_i} = 0, \quad \sum_{i=1}^{n} (r_i - R) \frac{(y_i - y_0)}{r_i} = 0, \quad \sum_{i=1}^{n} (r_i - R) = 0$$

$$\rightarrow R = \frac{1}{n} \sum_{i=1}^{n} r_i$$

Solution from

$$x_0 = \frac{1}{n} \left(\sum_{i=1}^{n} x_i - \frac{1}{n} \sum_{i=1}^{n} r_j \sum_{i=1}^{n} \frac{(x_i - x_0)}{r_i} \right), \quad y_0 = \frac{1}{n} \left(\sum_{i=1}^{n} y_i - \frac{1}{n} \sum_{i=1}^{n} r_j \sum_{i=1}^{n} \frac{(y_i - y_0)}{r_i} \right)$$

Calculating initial values

Given by SPÄTH (1996), with objective function

$$\tilde{\Omega} = \sum_{i=1}^{n} \tilde{v}_i^2 = \sum_{i=1}^{n} \left(r_i^2 - R^2 \right)^2 \rightarrow \min .$$

$$\left(2\sum_{i=1}^{n} x_i^{\,2} \right) x_0 + \left(2\sum_{i=1}^{n} x_i y_i \right) y_0 + \left(\sum_{i=1}^{n} x_i \right) z_0 = \sum_{i=1}^{n} x_i \left(x_i^{\,2} + y_i^{\,2} \right)$$

$$\left(2\sum_{i=1}^{n} x_i y_i \right) x_0 + \left(2\sum_{i=1}^{n} y_i^{\,2} \right) y_0 + \left(\sum_{i=1}^{n} y_i \right) z_0 = \sum_{i=1}^{n} y_i \left(x_i^{\,2} + y_i^{\,2} \right)$$

$$\left(2\sum_{i=1}^{n} x_i \right) x_0 + \left(2\sum_{i=1}^{n} y_i \right) y_0 + n \cdot z_0 = \sum_{i=1}^{n} \left(x_i^{\,2} + y_i^{\,2} \right)$$

with $z_0 = R^2 - x_0^{\,2} - y_0^{\,2}$

Solution of system of equations:

$$x_0^0 = x_0, \quad y_0^0 = y_0, \quad R^0 = \sqrt{z_0 + x_0^2 + y_0^2}$$

Numerical investigations

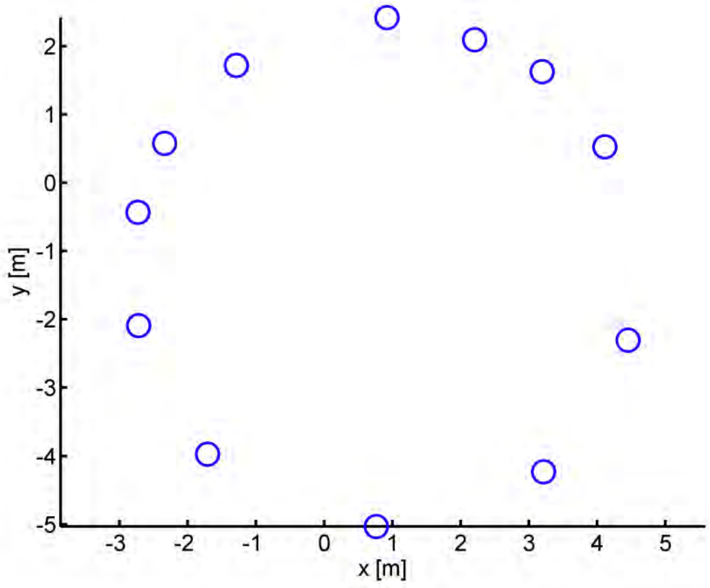
Technische Universität Berlin

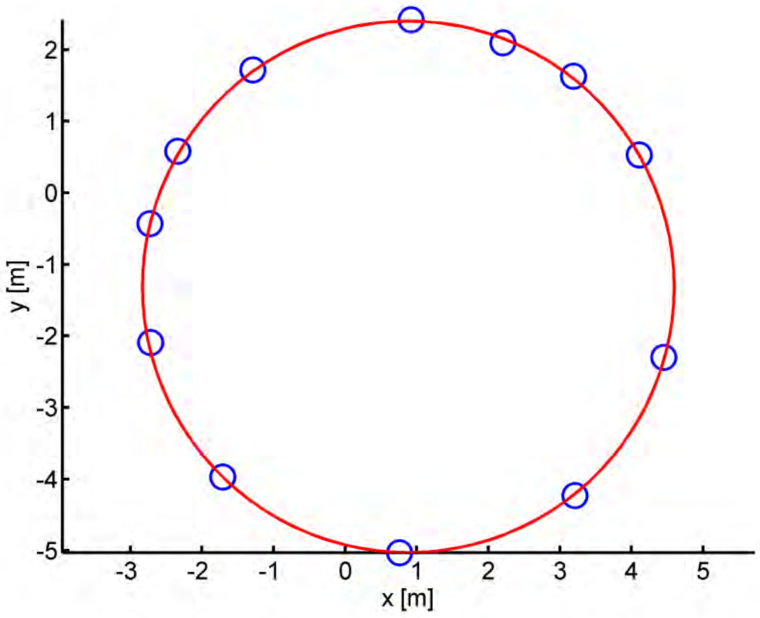
Comparison of the techniques:

- 3 different numerical examples chosen from the literature
- all the techniques are iterative → numerical value for the break-off condition 10^{-10}
- initial values calculated according to SPÄTH (1996)
- sensitivity with respect to falsified initial values → systematic approach: new initial values = old initial values + 10% ... up to 100%
- iteration steps of each technique are counted and compared

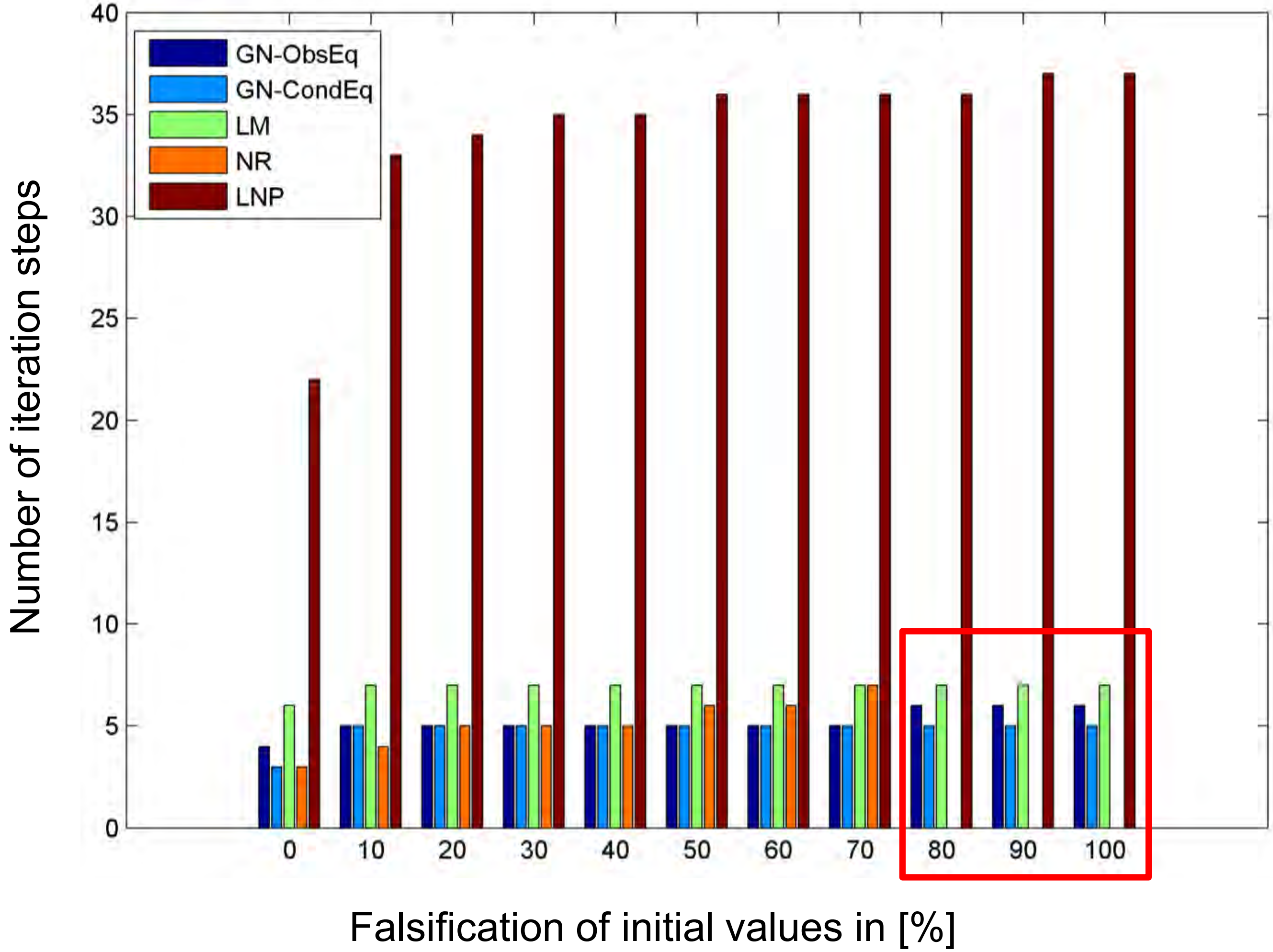
Example 1

BARTOLINI (1994):





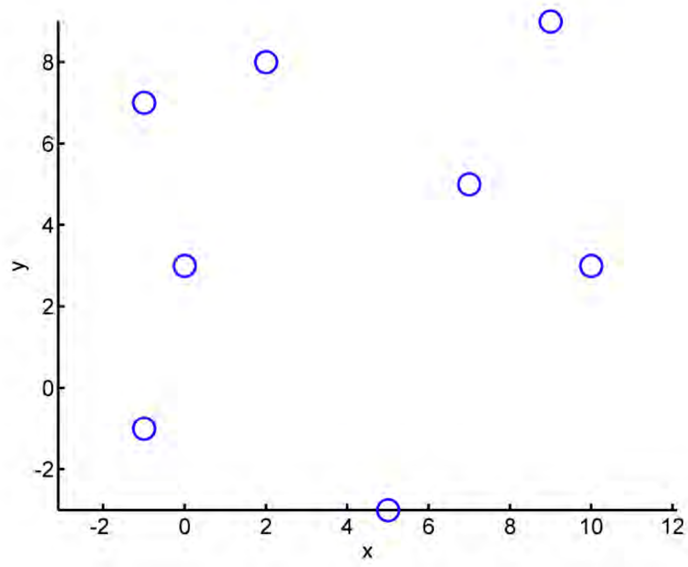
Example 1

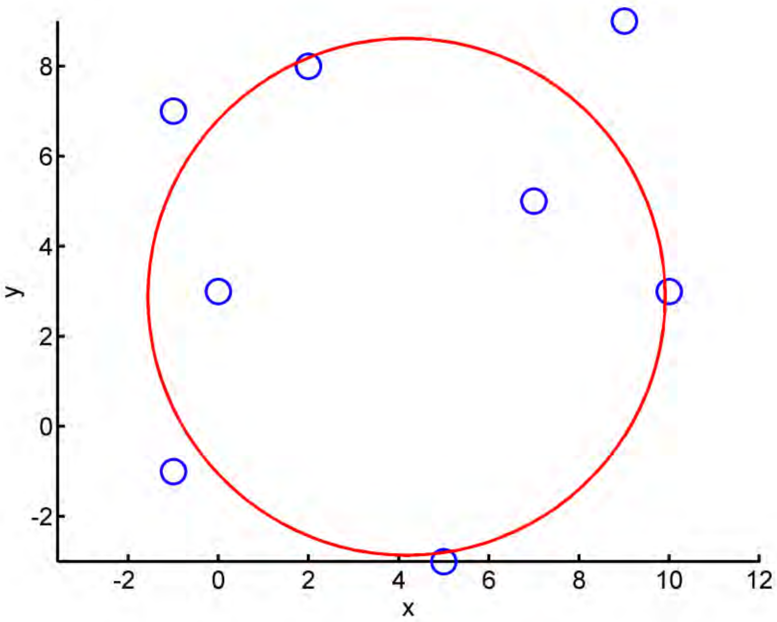


GN-ObsEq
GN-CondEq
LM
NR
LNP
Number of iteration steps
0
5
10
15
20
25
30
35
40
0
10
20
30
40
50
60
70
80
90
100
Falsification of initial values in [%]

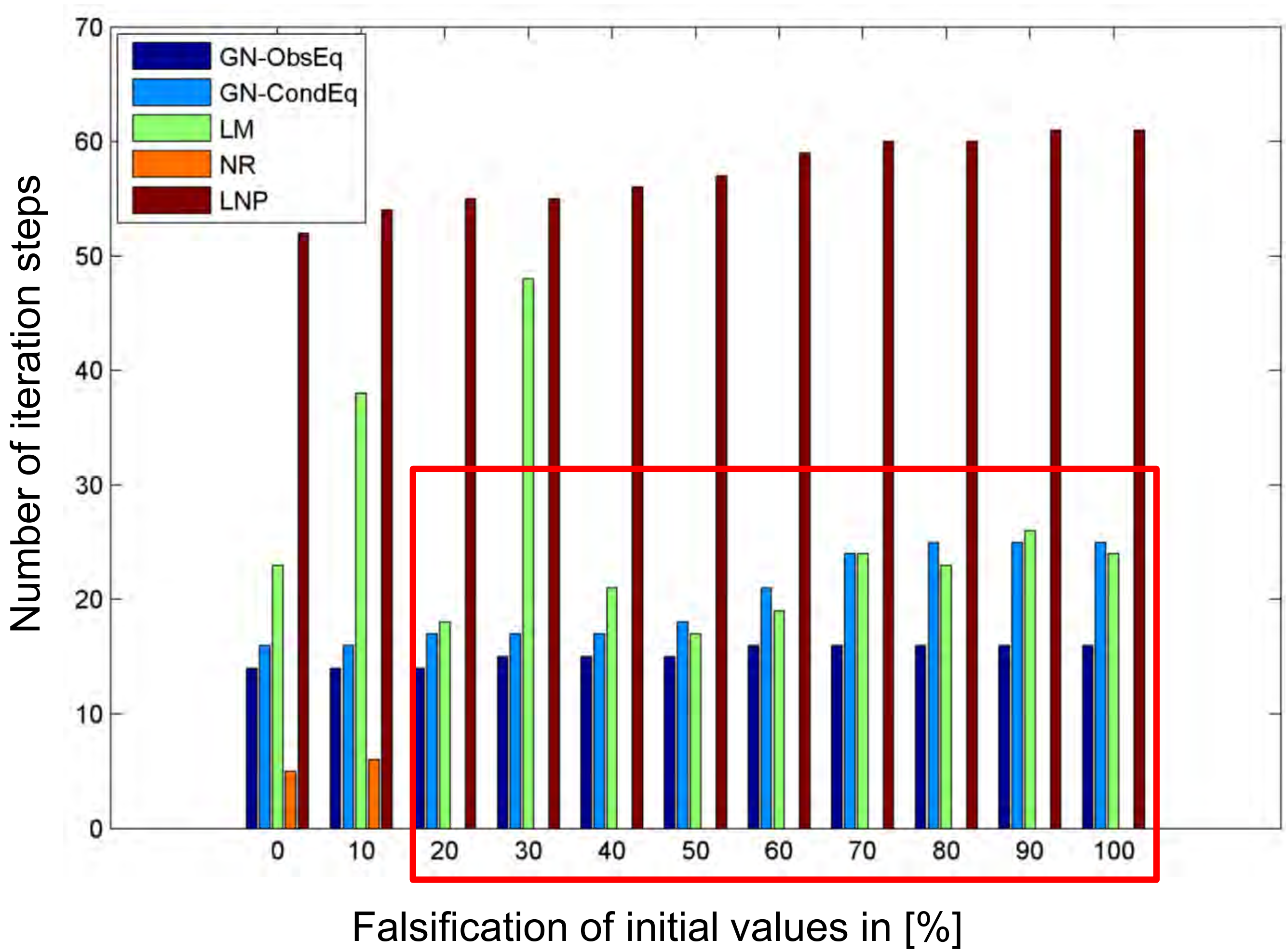
Example 2

SPÄTH (1996):





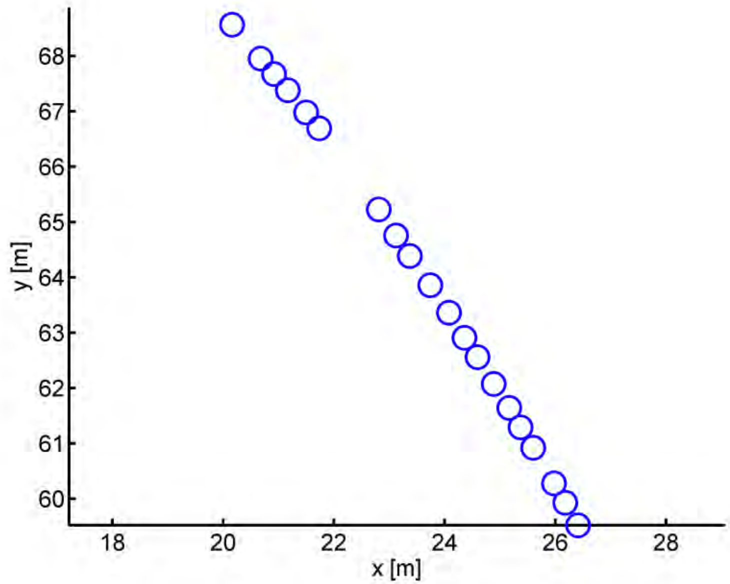
Example 2

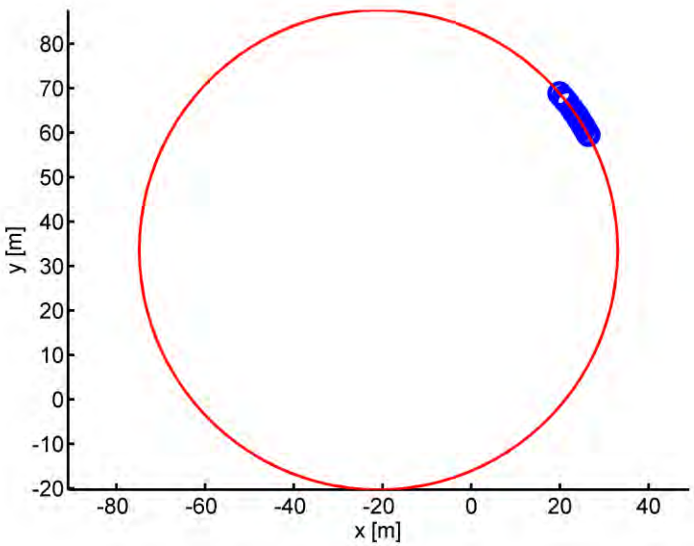

GN-ObsEq
GN-CondEq
LM
NR
LNP
Number of iteration steps
0
10
20
30
40
50
60
70
80
90
100
Falsification of initial values in [%]

Example 3

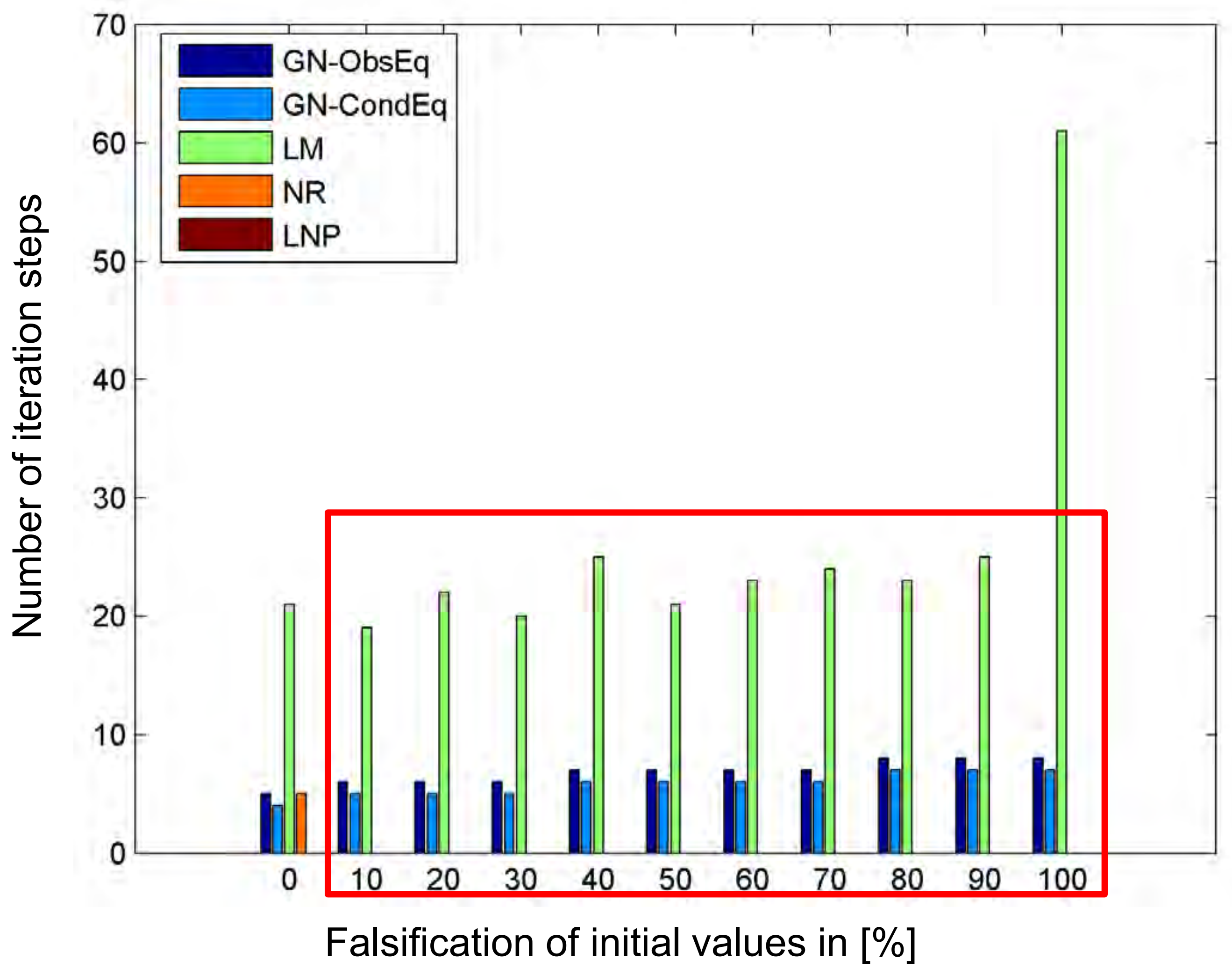
Technische Universität Berlin

RORRES AND GILMAN ROMANO (1997):





Example 3



Number of iteration steps
70
60
50
40
30
20
10
0
GN-ObsEq
GN-CondEq
LM
NR
LNP
0
10
20
30
40
50
60
70
80
90
100
Falsification of initial values in [%]

Conclusions

	GN		LM	NR	LNP
	Cond. Eq.	Obs. Eq.			
Sensitivity to initial values	+	+	++	--	++
Affected by the distribution of points in 2D plane	++	++	++	++	--

(++) very low, (+) low, (o) none, (-) high, (--) very high

Investigation of non-linear least squares problems using the example of circle fitting

Anastasia Pasioti, Frank Neitzel, Svetozar Petrovic
E-mail: **anastasia.pasioti@mailbox.tu-berlin.de**
Institute of Geodesy and Geoinformation Science, TU Berlin
Geodätische Woche 15.-17.09.2015, Stuttgart

References

BARTOLINI, S. (1994). Determinazione del centro di una struttura a pianta circolare. ANNO LIII - Bollettino di geodesia e scienze affini - n. 3, pages 315-324.

GHILANI, C. D. AND WOLF, P. R. (2006). Adjustment Computations-Spatial Data Analysis. John Wiley and Sons, Inc., Hoboken, New Jersey, 4th edition, ISBN: 978-0-471-69728-2, pages 178-179, 188-191, 546-554.

LEVENBERG, K. (1944). A method for the solution of certain non-linear problems in least-squares. Quarterly Journal of Applied Mathematics, Volume 2, Number 2, pages 164-168.

MADSEN, K. AND NIELSEN, H. B. AND TINGLEFF, O (2004). Lecture notes on Methods For Non-Linear Least Squares Problems, 2nd edition. Informatics and Mathematical Modelling, Department of Mathematical Modelling, Technical University of Denmark, pages 1-58.

MIKHAIL, E. M. AND ACKERMANN, F. (1976). Observations and Least Squares. IEP-Dun-Donnelley Harper and Row, Publishers New York Hagerstown San Francisco London, Thomas Y. Crowell Company, Inc., ISBN: 0-7002-2481-5, pages 3-8, 72-81, 108-109, 462-466.

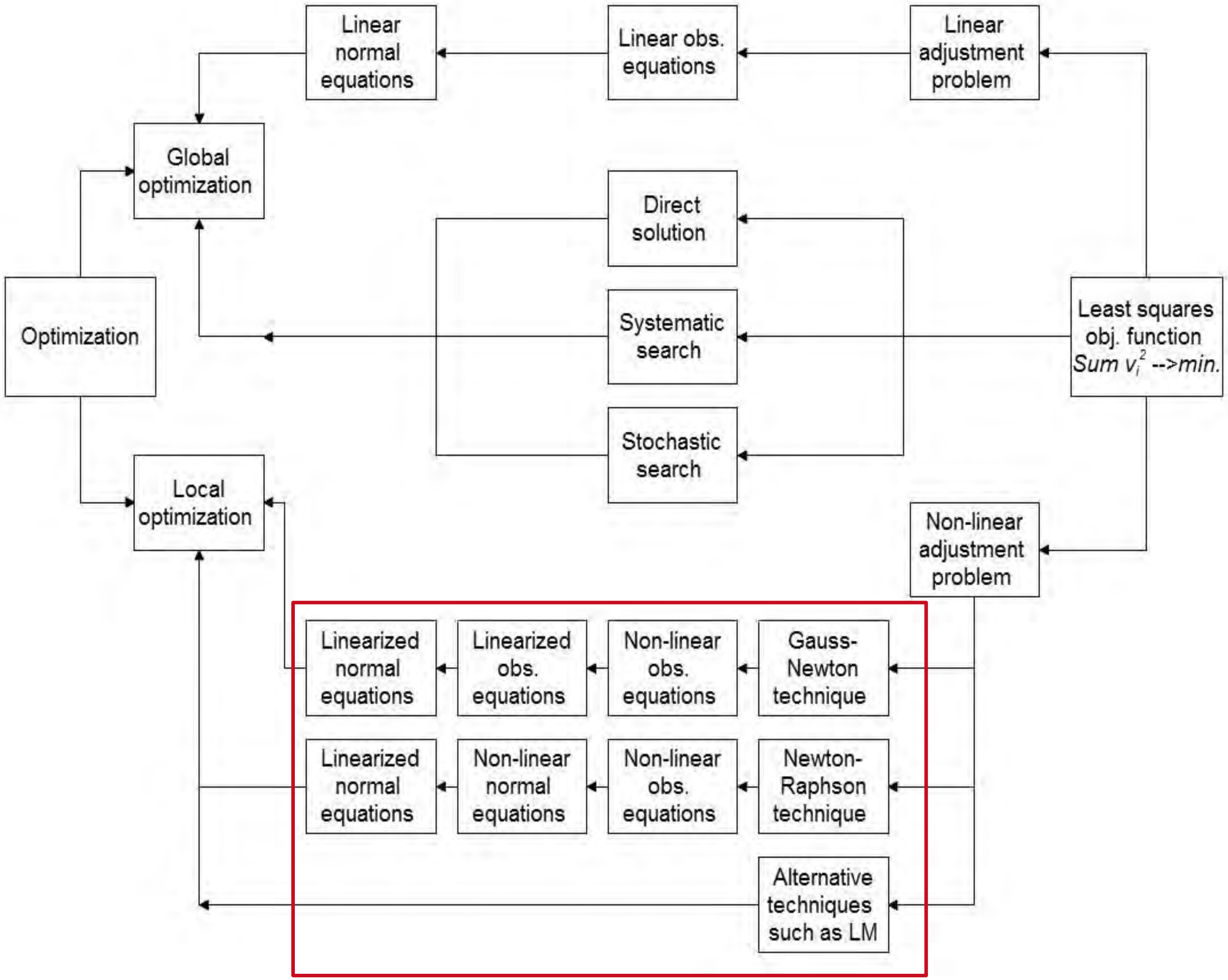
References

PASIOTI, A. (2015). Investigation of non-linear least squares problems using the example of circle fitting. Master thesis. Institute of Geodesy and Geoinformation Science, TU Berlin.

REUSKEN, A. (2015). Lecture notes on Numerik I für Maschinenbauer (SS 2010), Numerik für Maschinenbauer, Institut für Geometrie und Praktische Mathematik, RWTH Aachen University, http://www.igpm.rwth-aachen.de/Numa/NumaMB/SS10/LevMarq.pdf, Last visit: 18.03.2015

RORRES, C. AND GILMAN ROMANO, D. (1997). Finding the center of a circular starting line in an ancient Greek stadium. Society for Industrial and Applied Mathematics (SIAM), Volume 39, Number 4, pages 745-754.

SPÄTH, H. (1996). Least-Squares Fitting by Circles. Computing-Springer-Austria, Volume 57, Number 2, pages 179-185.



Linear normal equations
Linear obs. equations
Linear adjustment problem
Global optimization
Direct solution
Optimization
Systematic search
Least squares obj. function
Sum v_i^2 *-->min.*
Stochastic search
Local optimization
Non-linear adjustment problem
Linearized normal equations
Linearized obs. equations
Non-linear obs. equations
Gauss-Newton technique
Linearized normal equations
Non-linear normal equations
Non-linear obs. equations
Newton-Raphson technique
Alternative techniques such as LM