Photogrammetry & Robotics Lab

3D Coordinate Systems

(Bsc Geodesy & Geoinformation)

9. Estimating Parameters
of Spatial Transformations

Wolfgang Förstner

The slides have been created by Wolfgang Förstner.

1

5. Estimating parameters of transformations
5.1 Introduction

2

Goal

- Estimate transformation parameters
- Characterize the uncertainty of the estimates
- Handle outliers in the given data

For this video see Förstner/Wrobel (2016), 75-84

3

5.1.1 Task

- Given: pairs of corresponding points

$$(\mathcal{X}_i, \mathcal{X}'_i), i = 1, \ldots, I$$

- Assumption: related by a similarity

$$\mathcal{X}'_i = \mathcal{A}(\mathcal{X}_i)$$

- Task: Helmert-transformation *
  - Estimate the parameters of the similarity
  - Provide quality measures for the estimation

Friedrich Robert Helmert, geodesist 1843 -- 1917

4

Task

Overconstraint problem

- Dimension $d=2$: $2I$ constrains, 4 parameters

  Redundancy = number of remaining constraints

  $$R = 2I - 4$$

  → minimum number of correspondences: $I=2$
- Dimension $d=3$: $3I$ constrains, 7 parameters

  Redundancy

  $$R = 3I - 7$$

  → minimum number of correspondences: $I=3$

5

Task

Inconsistencies due to measurement deviations

$$\boldsymbol{x}_i' \neq \lambda\, R\, \boldsymbol{x}_i + \boldsymbol{t}, i=1,...,I$$

Here: assumption: only χ_i' are perturbated

Corrections* $\boldsymbol{v}_{x_i'}$ at coordinates $\boldsymbol{x}_i'$ → constraints

$$\boldsymbol{x}_i' + \boldsymbol{v}_{x_i'} = \lambda\, R\, \boldsymbol{x}_i + \boldsymbol{t}, i=1,...,I$$

→Underconstrained system:

unknown parameters, unknown corrections

* German: "**V**erbesserungen"

6

Solution

Use principle of least squares (LS, regularization) following Carl Friedrich Gauss (1777—1855)

Determine corrections and parameters such that corrections are smallest → Minimize sum of squared corrections

$$\Omega = \sum_{i=1}^{I} \boldsymbol{v}'^{\mathsf{T}}_{x_i} \boldsymbol{v}_{x_i'} = \sum_{i=1}^{I} v'^2_{x_i} + v'^2_{y_i} + v'^2_{z_i}$$

7

Solution?

There exists a direct LS solution for estimating

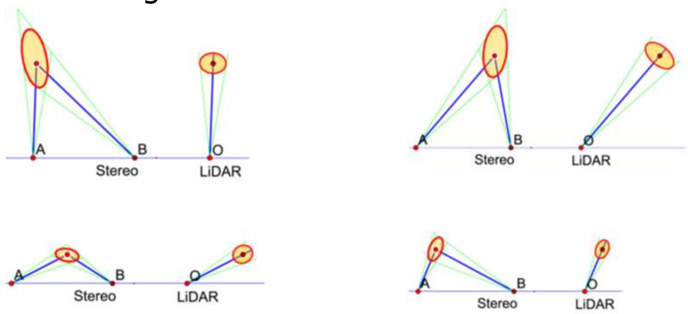
- 7 parameters of 3D similarity
- 6 parameters of 3D motion

Why is this not sufficient?

8

Point uncertainty highly varying

- Stereo images and LiDAR



- GPS: 3D accuracy = f(satellite configuration)

Solution?

- No uncertainty of observations is taken into account
  Given points may
  - have different accuracies
  - may have correlated coordinates
  - may be correlated, e.g. due to uncertainties of A, B and O
- No covariance matrix for resultant parameters

→ Take uncertainty of given coordinates into account
→ Derive uncertainty of estimated parameters

On notation

World of applications
e.g. coordinates $\boldsymbol{x}$, angles $\boldsymbol{\alpha}$
transformations M , parameters $\boldsymbol{p}$, etc.

World of estimation
(measurements, unknown parameters)

$$(\boldsymbol{l}, \boldsymbol{x}) \quad \text{or} \quad (\boldsymbol{y}, \boldsymbol{\beta}) \quad \text{or} \quad (\boldsymbol{x}, \boldsymbol{\theta})$$

(depending on community !)
→ Estimation theory

On notation

→Leave notation in both worlds and **assign**

e.g. using estimation notation $(\boldsymbol{l}, \boldsymbol{x})$
1. observed coordinates, unknown parameters

$$\boldsymbol{l} := [\boldsymbol{x}_i]$$
$$\boldsymbol{x} := \boldsymbol{p}$$

2. After estimating parameters

$$\widehat{\boldsymbol{p}} := \widehat{\boldsymbol{x}}$$

Estimation principles

Given:

- Unknown parameters with some prior uncertainty

$$p(\boldsymbol{x})$$

- Observations with uncertainty given the parameters

$$p(\boldsymbol{l} \mid \boldsymbol{x})$$

Task:

estimate parameters best explaining the observations

$$\widehat{\boldsymbol{x}} = \operatorname{argmax}_x \, p(\boldsymbol{x} \mid \boldsymbol{l})$$

Bayesian estimation

Law of conditional probability

$p(\boldsymbol{l}, \boldsymbol{x}) = p(\boldsymbol{x} \mid \boldsymbol{l}) \, p(\boldsymbol{l}) = p(\boldsymbol{l} \mid \boldsymbol{x}) \, p(\boldsymbol{x})$

Bayesian theorem

$$p(\boldsymbol{x} \mid \boldsymbol{l}) = \frac{p(\boldsymbol{l} \mid \boldsymbol{x}) \, p(\boldsymbol{x})}{p(\boldsymbol{l})}$$

For given $\boldsymbol{l}$, fixed

- Prior distribution $p(\boldsymbol{x})$
- Likelihood $L(\boldsymbol{x}) = p(\boldsymbol{l} \mid \boldsymbol{x})$
- Posterior distribution $p(\boldsymbol{x} \mid \boldsymbol{l})$
- Probability to observe $\boldsymbol{l}$, fixed $p(\boldsymbol{l})$

MAP and ML estimation

1. Maximize posterior probability (**MAP**)
   → Maximum posterior estimate
   (also Bayesian estimate)

$$\widehat{\boldsymbol{x}}_{\mathrm{MAP}} = \operatorname{argmax}_x \, p(\boldsymbol{l} \mid \boldsymbol{x}) \, p(\boldsymbol{x}) = \operatorname{argmax}_x \, L(\boldsymbol{x}) \, p(\boldsymbol{x})$$

2. Maximum likelihood (**ML**) estimate, $p(\boldsymbol{x}) = \mathrm{const.}$

$$\widehat{\boldsymbol{x}}_{\mathrm{ML}} = \operatorname{argmax}_x \, p(\boldsymbol{l} \mid \boldsymbol{x}) = \operatorname{argmax}_x \, L(\boldsymbol{x})$$

Best linear unbiased estimates

Assume Gaussian distribution, linear dependency

$$p(\boldsymbol{l} \mid \boldsymbol{x}) \;\propto\; \exp\left(-\frac{1}{2}(\boldsymbol{l} - A\boldsymbol{x})^{\mathsf{T}} \Sigma_{ll}^{-1} (\boldsymbol{l} - A\boldsymbol{x})\right)$$

3. Best linear unbiased estimator
   a) Leads to smallest CovM of estimates
   b) Estimate is a linear function of observations
   c) Estimate is unbiased

$$\widehat{\boldsymbol{x}}_{BLUE} = \operatorname{argmin}_x \, (\boldsymbol{l} - A\boldsymbol{x})^{\mathsf{T}} \Sigma_{ll}^{-1} (\boldsymbol{l} - A\boldsymbol{x})$$

Least squares estimates

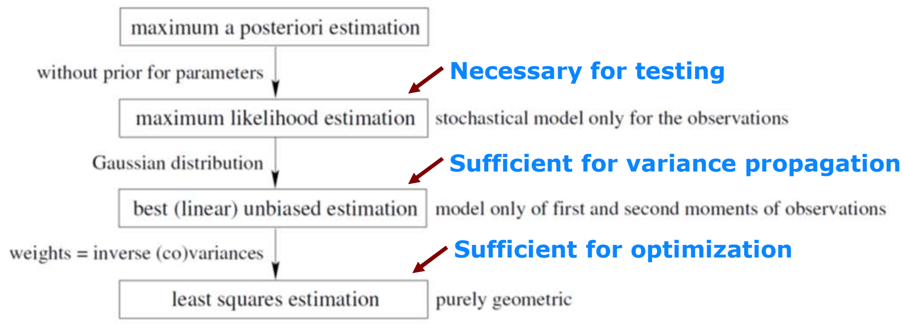
- Ordinary least squares

$$\widehat{\boldsymbol{x}}_{LS} = \text{argmin}_x \; (\boldsymbol{l} - A\boldsymbol{x})^{\mathsf{T}}(\boldsymbol{l} - A\boldsymbol{x})$$

- Generalized least squares with weight matrix W_{ll}

$$\widehat{\boldsymbol{x}}_{GLS} = \text{argmin}_x \; (\boldsymbol{l} - A\boldsymbol{x})^{\mathsf{T}} W_{ll} (\boldsymbol{l} - A\boldsymbol{x})$$

Hierarchy of estimates



Hierarchy of estimates

- ML estimate = MAP estimate if no prior
- Best linear unbiased estimate = ML estimate if
  - Distribution is Gaussian
  - Model is linear
- GLS estimate = best linear unbiased estimate if

$$W_{ll} = \Sigma_{ll}^{-1}$$

- LS estimate = GLS estimate if

$$W_{ll} = I$$

→ e. g. **LSE = MAP if model linear, Gaussian** $\Sigma_{ll} = I$

5.1.2 Non-linear Gauss-Markov Model

Basis: Maximum likelihood estimation (ML-estimation)

Given: observation vector $\boldsymbol{l}$, fixed
sample from

$$\underline{\boldsymbol{l}} \sim p(\boldsymbol{l} \mid \boldsymbol{x}) = \mathcal{N}(\mathbb{E}(\underline{\boldsymbol{l}} \mid \boldsymbol{x}), \mathbb{D}(\underline{\boldsymbol{l}} \mid \boldsymbol{x}))$$

Notation: stochastical variables are underscored

Mathematical model

Functional model:

generative model for $\boldsymbol{l}$, given parameters $\boldsymbol{x}$

model is non-linear

$$\mathbb{E}(\underline{\boldsymbol{l}} \mid \boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})$$

Stochastical model

$$\mathbb{D}(\underline{\boldsymbol{l}} \mid \boldsymbol{x}) = \mathbb{D}(\underline{\boldsymbol{l}} - \boldsymbol{f}(\boldsymbol{x})) = \Sigma_{ll}$$

21

Mathematical model

- Generative model with perturbations $\boldsymbol{e}$

$$\underline{\boldsymbol{l}} = \boldsymbol{f}(\boldsymbol{x}) + \underline{\boldsymbol{e}} \quad \text{with} \quad \mathbb{D}(\underline{\boldsymbol{e}}) = \Sigma_{ee}$$

or with corrections $\boldsymbol{v} = -\boldsymbol{e}$

$$\underline{\boldsymbol{l}} + \underline{\boldsymbol{v}} = \boldsymbol{f}(\boldsymbol{x}) \quad \text{with} \quad \mathbb{D}(\underline{\boldsymbol{v}}) = \Sigma_{vv} = \Sigma_{ee}$$

- Matrix Σ_{ll} refers to corrections or perturbations

$$\Sigma_{ll} = \mathbb{D}(\underline{\boldsymbol{l}} - \boldsymbol{f}(\boldsymbol{x})) = \mathbb{D}(\underline{\boldsymbol{e}}) = \mathbb{D}(\underline{\boldsymbol{v}})$$

→ the mean values $\mathbb{E}(\underline{\boldsymbol{l}} \mid \boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})$ are fixed

22

Maximum likelihood estimation

Maximize likelihood

$$L(\boldsymbol{x}) = p(\boldsymbol{l} \mid \boldsymbol{x})$$

Here

$$p(\boldsymbol{l} \mid \boldsymbol{x}) \propto \exp\left(-\frac{1}{2}(\boldsymbol{l} - \boldsymbol{f}(\boldsymbol{x}))^{\mathsf{T}}\Sigma_{ll}^{-1}(\boldsymbol{l} - \boldsymbol{f}(\boldsymbol{x}))\right)$$

Hence minimize

$$\Omega = (\boldsymbol{l} - \boldsymbol{f}(\boldsymbol{x}))^{\mathsf{T}}\Sigma_{ll}^{-1}(\boldsymbol{l} - \boldsymbol{f}(\boldsymbol{x}))$$

or

$$\Omega = \boldsymbol{v}^{\mathsf{T}}\Sigma_{ll}^{-1}\boldsymbol{v}$$

→ Best estimate, generalizes LS with $W_{ll} = \Sigma_{ll}^{-1}$

23

5.1.3 Model for Helmert transformation

- Functional model
- Stochastical model
  - General
  - Simplified

24

Observations and corrections

Collected in N-vectors,
for dimension d we have $N = dI$

$$\begin{bmatrix} l_1 \\ \dots \\ l_n \\ \dots \\ l_N \end{bmatrix} = \underset{N\times 1}{\boldsymbol{l}} := \begin{bmatrix} \boldsymbol{x}'_1 \\ \dots \\ \boldsymbol{x}'_i \\ \dots \\ \boldsymbol{x}'_I \end{bmatrix} \qquad \underset{N\times 1}{\boldsymbol{v}} := \begin{bmatrix} \boldsymbol{v}_{x'_1} \\ \dots \\ \boldsymbol{v}_{x'_i} \\ \dots \\ \boldsymbol{v}_{x'_I} \end{bmatrix}$$

25

Parameters and functional model

Unknown parameters, collected in U-vector

$$\underset{U\times 1}{\boldsymbol{x}} := \begin{bmatrix} \boldsymbol{r} \\ \boldsymbol{t} \\ \lambda \end{bmatrix}$$

Functional model

$$\mathbb{E}(\boldsymbol{l}) = \boldsymbol{f}(\boldsymbol{x}) \quad \text{or} \quad \mathbb{E}\begin{bmatrix} \boldsymbol{x}'_1 \\ \dots \\ \boldsymbol{x}'_i \\ \dots \\ \boldsymbol{x}'_I \end{bmatrix} = \begin{bmatrix} \lambda R\boldsymbol{x}_1 + \boldsymbol{t} \\ \dots \\ \lambda R\boldsymbol{x}_i + \boldsymbol{t} \\ \dots \\ \lambda R\boldsymbol{x}_I + \boldsymbol{t} \end{bmatrix}$$

Representation of rotation: open

26

Stochastical model

Stochastical model:

- All points may be mutually dependent
- Coordinates of one point may be dependent

$$\underset{N\times N}{\Sigma_{ll}} := \begin{bmatrix} \Sigma_{x'_1x'_1} & \dots & \Sigma_{x'_1x'_i} & \dots & \Sigma_{x'_1x'_j} & \dots & \Sigma_{x'_1x'_I} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \Sigma_{x'_ix'_1} & \dots & \Sigma_{x'_ix'_i} & \dots & \Sigma_{x'_ix'_j} & \dots & \Sigma_{x'_ix'_I} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \Sigma_{x'_jx'_1} & \dots & \Sigma_{x'_jx'_i} & \dots & \Sigma_{x'_jx'_j} & \dots & \Sigma_{x'_jx'_I} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \Sigma_{x'_Ix'_1} & \dots & \Sigma_{x'_Ix'_i} & \dots & \Sigma_{x'_Ix'_j} & \dots & \Sigma_{x'_ix'_I} \end{bmatrix}$$

27

Stochastical model

Simplified stochastical model

$$\Sigma_{ll} := \mathrm{Diag}\left(\sigma^2_{x_i} I_3\right) = \begin{bmatrix} \sigma^2_{x_1} I_3 & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \sigma^2_{x_i} I_3 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \sigma^2_{x_I} I_3 \end{bmatrix}$$

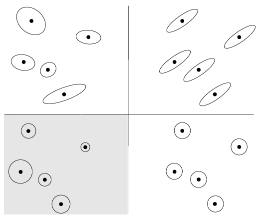
- Points are mutually independent
- Coordinates of one point are mutually independent
- Accuracy/weight may vary from point to point

28

Visualization of stochastical model

- Isotropic/anisotropic (lower/upper row):
  rotation invariant/variant
- Homogeneous/inhomogeneous (right/left column):
  translation invariant/variant

$w_i := \frac{1}{\sigma^2_{x'_i}}$ →



← $w_i = w := \frac{1}{\sigma^2_{x'}}$

29

Simplified optimization problem

If uncertainty is isotropic and points are independent estimates are

$$(\widehat{R}, \widehat{\boldsymbol{t}}, \widehat{\lambda}) = \operatorname{argmin}_{R,t,\lambda}\, \Omega(R, \boldsymbol{t}, \lambda)$$

with

$$\Omega(R, \boldsymbol{t}, \lambda) := \sum_{i=1}^{I} (\boldsymbol{x}'_i - \lambda R \boldsymbol{x}_i - \boldsymbol{t})^{\mathsf{T}} (\boldsymbol{x}'_i - \lambda R \boldsymbol{x}_i - \boldsymbol{t}) w_i$$

and weights

$$w_i := \frac{1}{\sigma^2_{x'_i}}$$

Non-linear least squares problem

30

Solutions

- Optimal solution for general stochastical model
  - ☺ Yields covariance matrix of estimated parameters
  - ☺ Stochastical model can be simplified
  - ☹ Requires approximate values
  - ☹ Solution is iterative
- Direct solution for simplified stochastical model
- Fast direct solution for 3 points

31

Next lecture

5.2 Direct solution for spatial similarity

32

References of video series

Arun, K. S., T. S. Huang, and S. B. Blostein (1987). Least-Squares Fitting of Two 3D Point Sets. IEEE T-PAMI 9 (5), 698-700.

Bishop, C. (2006). Pattern Recognition and Machine Learning. Springer.

Boyd, S. and L. Vandenberghe (2004). Convex optimization. Cambridge University Press.

Cayley, A. (1846). Sur quelques proprétés des déterminants gauches. Journal für die reine und angewandte Mathematik 32, 119-123

Cover, T. and J. A. Thomas (1991). Elements of Information Theory. John Wiley & Sons.

H. S. M. Coxeter (1946): Quaternions and Reflections, The American Mathematical Monthly, 53(3), 136-146

Fallat, S. M. and M. J. Tsatsomeros (2002). On the Cayley Transform of Positivity Classes of Matrices. Electronic Journal of Linear Algebra 9, 190-196.

Fischler, M. A. and R. C. Bolles (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Communications of the ACM 24 (6), 381-395.

Förstner, W. and B. P. Wrobel (2016). Photogrammetric Computer Vision Statistics, Geometry, Orientation and Reconstruction. Springer.

Horn, B. K. B. (1987). Closed-form Solution of Absolute Orientation Using Unit Quaternions. Optical Soc. of America.

Howell, T. D. and J.-C. Lafon (1975). The Complexity of the Quaternion Product. Technical Report TR75-245, Cornell University.

Kanatani, K. (1990). Group Theoretical Methods in Image Understanding. New York: Springer.

Koch, K.-R. (1999). Parameter Estimation and Hypothesis Testing in Linear Models (2nd ed.). Springer.

Li, S. Z. (2000). Markov random field modeling in computer vision. Springer.

McGlone, C. J., E. M. Mikhail, and J. S. Bethel (2004). Manual of Photogrammetry (5th ed.). Maryland, USA: American Society of Photogrammetry and Remote Sensing.

Mikhail, E. M. and F. Ackermann (1976). Observations and Least Squares. University Press of America

Mikhail, E. M., J. S. Bethel, and J. C. McGlone (2001). Introduction to Modern Photogrammetry. Wiley.

Palais, B. and R. Palais (2007). Euler's fixed point theorem: The axis of a rotation. J. fixed point theory appl. 2, 215-220.

Raguram, R., O. Chum, M. Pollefeys, J. Matas, and J.-M. Frahm (2013). USAC: A Universal Framework for Random Sample Consensus. IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (8), 2022-2038.

Rao, R. C. (1973). Linear Statistical Inference and Its Applications. New York: Wiley.

Rodriguez, O. (1840). Des lois géometriques qui regissent les deplacements d'un système solide independament des causes qui peuvent les produire. Journal de mathématiques pures et appliquées 1 (5), 380440.

Sansò, F. (1973). An Exact Solution of the Roto-Translation Problem. Photogrammetria 29 (6), 203-206.

Vaseghi, S. V. (2000). Advanced Digital Signal Processing and Noise Reduction. Wiley.

Weber, M. (2003). Rotation between two vectors. Personal communication, Bonn.