5. Estimating parameters of transformations

5.1 Introduction

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

5.1.1 Task
- Given: pairs of corresponding points \((x_i, x'_i), i = 1, \ldots, I\)
- Assumption: related by a similarity
  \(x'_i = A(x_i)\)
- Task: Helmert-transformation *
  - Estimate the parameters of the similarity
  - Provide quality measures for the estimation

Friedrich Robert Helmert, geodesist 1843 -- 1917
Task
Overconstraint problem
- Dimension $d = 2$: 2l constrains, 4 parameters
  Redundancy = number of remaining constraints $R = 2l - 4$
  → minimum number of correspondences: $I = 2$
- Dimension $d = 3$: 3l constrains, 7 parameters
  Redundancy $R = 3l - 7$
  → minimum number of correspondences: $I = 3$

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
$$\Omega = \sum_{i=1}^{l} v_{x_i}^T v_{x_i}' = \sum_{i=1}^{l} v_{x_i}^2 + v_{y_i}^2 + v_{z_i}^2$$

Task
Inconsistencies due to measurement deviations
$$x_i' \neq \lambda R x_i + t, i = 1, ..., I$$
Here: assumption: only $x_i'$ are perturbed
Corrections* $v_{x_i}'$ at coordinates $x_i' \rightarrow$ constraints
$$x_i' + v_{x_i}' = \lambda R x_i + t, i = 1, ..., I$$
→ Underconstrained system:
  unknown parameters, unknown corrections
* German: “Verbesserungen”

Solution?
There exists a direct LS solution for estimating
- 7 parameters of 3D similarity
- 6 parameters of 3D motion
Why is this not sufficient?
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 \(x\), angles \(\alpha\)
transformations \(M\), parameters \(p\), etc.

World of estimation
(measurements, unknown parameters)
\((l, x)\) or \((y, \beta)\) or \((x, \theta)\)
(depending on community !)

→ Estimation theory

On notation

→ Leave notation in both worlds and assign
e.g. using estimation notation \((l, x)\)

1. observed coordinates, unknown parameters
\[ l := [x_i] \]
\[ x := p \]

2. After estimating parameters
\[ \hat{p} := \hat{x} \]
Estimation principles

Given:
- Unknown parameters with some prior uncertainty \( p(x) \)
- Observations with uncertainty given the parameters \( p(l \mid x) \)

Task:
estimate parameters best explaining the observations \( \hat{x} = \arg\max_x p(x \mid l) \)

Bayesian estimation

Bayesian theorem

\[
p(x \mid l) = \frac{p(l \mid x) p(x)}{p(l)}
\]

For given \( l \), fixed
- Prior distribution \( p(x) \)
- Likelihood \( L(x) = p(l \mid x) \)
- Posterior distribution \( p(x \mid l) \)
- Probability to observe \( l \), fixed

MAP and ML estimation

1. Maximize posterior probability (MAP)
   - Maximum posterior estimate (also Bayesian estimate)
     \( \hat{x}_{MAP} = \arg\max_x p(l \mid x) p(x) = \arg\max_x L(x) p(x) \)

2. Maximum likelihood (ML) estimate, \( p(x) = \text{const.} \)
   \( \hat{x}_{ML} = \arg\max_x p(l \mid x) = \arg\max_x L(x) \)

Best linear unbiased estimates

Assume Gaussian distribution, linear dependency

\[
p(l \mid x) \propto \exp \left( -\frac{1}{2} (l - Ax)^T \Sigma_l^{-1} (l - Ax) \right)
\]

3. Best linear unbiased estimator
   a) Leads to smallest CovM of estimates
   b) Estimate is a linear function of observations
   c) Estimate is unbiased
     \( \hat{x}_{BLUE} = \arg\min_x (l - Ax)^T \Sigma_l^{-1} (l - Ax) \)
Least squares estimates
▪ Ordinary least squares
  \[ \hat{x}_{LS} = \arg\min_x (l - Ax)^T(l - Ax) \]
▪ Generalized least squares with weight matrix \(W_{ll}\)
  \[ \hat{x}_{GLS} = \arg\min_x (l - Ax)^TW_{ll}(l - Ax) \]

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 \(l\), fixed sample from
\[ l \sim p(l \mid x) = \mathcal{N}(E(l \mid x), D(l \mid x)) \]

Notation: stochastical variables are underscored
Mathematical model

Functional model:
  generative model for \( l \), given parameters \( x \)
  model is non-linear
  \[ \mathbb{E}(l \mid x) = f(x) \]

Stochastical model
  \[ D(l \mid x) = D(l - f(x)) = \Sigma_{ll} \]

---

Maximum likelihood estimation

Maximize likelihood

\[ L(x) = p(l \mid x) \]

Here

\[ p(l \mid x) \propto \exp \left( -\frac{1}{2} (l - f(x))^\top \Sigma_{ll}^{-1} (l - f(x)) \right) \]

Hence minimize

\[ \Omega = (l - f(x))^\top \Sigma_{ll}^{-1} (l - f(x)) \]

or

\[ \Omega = v^\top \Sigma_{ll}^{-1} v \]

\( \Rightarrow \) Best estimate, generalizes LS with \( W_{ll} = \Sigma_{ll}^{-1} \)

---

Mathematical model

- Generative model with perturbations \( e \)
  \[ l = f(x) + e \quad \text{with} \quad D(e) = \Sigma_{ee} \]
  
or with corrections \( v = -e \)
  \[ l + v = f(x) \quad \text{with} \quad D(v) = \Sigma_{vv} = \Sigma_{ee} \]

- Matrix \( \Sigma_{ll} \) refers to corrections or perturbations
  \[ \Sigma_{ll} = D(l - f(x)) = D(e) = D(v) \]
  
  \( \Rightarrow \) the mean values \( \mathbb{E}(l \mid x) = f(x) \) are fixed

---

5.1.3 Model for Helmert transformation

- Functional model
- Stochastical model
  - General
  - Simplified
Observations and corrections

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

\[
\begin{bmatrix}
 l_1 \\
\vdots \\
 l_n \\
\vdots \\
 l_N \\
\end{bmatrix}
= I_{N \times 1}
\begin{bmatrix}
 x'_1 \\
\vdots \\
 x'_i \\
\vdots \\
 x'_f \\
\end{bmatrix}
= v_{x'_i}
\]

Parameters and functional model

Unknown parameters, collected in $U$-vector

\[
\begin{bmatrix}
 x \\
 t \\
 \lambda \\
\end{bmatrix}
= \begin{bmatrix}
 r \\
 t \\
 \lambda \\
\end{bmatrix}
\]

Functional model

\[
E(l) = f(x) \quad \text{or} \quad E\begin{bmatrix}
 x'_1 \\
\vdots \\
 x'_i \\
\vdots \\
 x'_f \\
\end{bmatrix}
= \begin{bmatrix}
 \lambda R x_1 + t \\
\vdots \\
 \lambda R x_i + t \\
\vdots \\
 \lambda R x_f + t \\
\end{bmatrix}
\]

Representation of rotation: open

Stochastical model

Stochastical model:
- All points may be mutually dependent
- Coordinates of one point may be dependent

\[
\Sigma_{li} := \begin{bmatrix}
 \Sigma_{x'_1x'_1} & \cdots & \Sigma_{x'_1x'_i} & \cdots & \Sigma_{x'_1x'_f} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \Sigma_{x'_i'x'_i} & \cdots & \Sigma_{x'_i'x'_f} \\
\vdots & \cdots & \cdots & \Sigma_{x'_f'x'_f} & \cdots \\
\Sigma_{x'_f'x'_1} & \cdots & \Sigma_{x'_f'x'_i} & \cdots & \Sigma_{x'_f'x'_f} \\
\end{bmatrix}
\]

Stochastical model

Simplified stochastical model

\[
\Sigma_{li} := \text{Diag}\left(\sigma_{x_i}^2 I_3\right)
\]

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

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

\[
\begin{align*}
\sigma_i^2 & \quad \rightarrow \quad \sigma_i^2 \\
\rightarrow & \quad \rightarrow \\
\end{align*}
\]

**Simplified optimization problem**

If uncertainty is isotropic and points are independent estimates are

\[
(\hat{R}, \hat{t}, \hat{\lambda}) = \arg\min_{R, t, \lambda} \Omega(R, t, \lambda)
\]

with

\[
\Omega(R, t, \lambda) := \sum_{i=1}^{I} (x_i' - \lambda Rx_i - t)^T (x_i' - \lambda Rx_i - t) w_i
\]

and weights

\[
w_i := \frac{1}{\sigma_{x_i}^2}
\]

Non-linear least squares problem

**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

**Next lecture**

5.2 Direct solution for spatial similarity
References of video series


