Robot Mapping

Robust Least Squares for SLAM

Gian Diego Tipaldi, Wolfram Burgard

Courtesy for most images: Pratik Agarwal
Least Squares in General

- Minimizes the **sum of the squared errors**
- ML estimation in the Gaussian case

**Problems:**
- Sensitive to outliers
- Only Gaussians (single modes)
Data Association Is Ambiguous And Not Always Perfect

- Places that look identical
- Similar rooms in the same building
- Cluttered scenes
- GPS multi pass (signal reflections)
- ...
Example

3D world

belief about the robot’s pose

Courtesy: E. Olson
Such Situations Occur In Reality

Courtesy: E. Olson, P. Agarwal
Committing To The Wrong Mode Can Lead to Mapping Failures

Courtesy: E. Olson, P. Agarwal
Data Association Is Ambiguous And Not Always Perfect

- Places that look identical
- Similar rooms in the same building
- Cluttered scenes
- GPS multi pass (signal reflections)
- ...

How to incorporate that into graph-based SLAM?
Data Association Is Ambiguous And Not Always Perfect

- Places that look identical
- Similar rooms in the same building
- Cluttered scenes
- GPS multi pass (signal reflections)
- ...

How to incorporate that into graph-based SLAM?
Mathematical Model

- We can express a multi-modal belief by a sum of Gaussians

\[ p(z \mid x) = \eta \exp\left( -\frac{1}{2} e_{ij}^T \Omega_{ij} e_{ij} \right) \]

\[ p(z \mid x) = \sum_{k} \omega_k \eta_k \exp\left( -\frac{1}{2} e_{ijk}^T \Omega_{ijk} e_{ijk} \right) \]

Sum of Gaussians with k modes
Problem

- During error minimization, we consider the negative log likelihood

\[- \log p(z \mid x) = \frac{1}{2} e^{T} \Omega e - \log \eta\]

\[- \log p(z \mid x) = - \log \sum_{k} w_{k} \eta_{k} \exp \left( - \frac{1}{2} e^{T}_{ijk} \Omega e_{rijk} \right)\]

The log cannot be moved inside the sum!
Max-Mixture Approximation

- Instead of computing the sum of Gaussians at $x$, compute the maximum of the Gaussians

$$p(z \mid x) = \sum_k w_k \eta_k \exp\left(-\frac{1}{2}e^T_{ijk} \Omega_{ijk} e_{ijk}\right)$$

$$\simeq \max_k w_k \eta_k \exp\left(-\frac{1}{2}e^T_{ijk} \Omega_{ijk} e_{ijk}\right)$$
Max-Mixture Approximation

Original bi-modal mixture

Max-mixture

Sum-mixture

approximation error
Log Likelihood Of The Max-Mixture Formulation

- The log can be moved inside the max operator

\[ p(z \mid x) \simeq \max_k w_k \eta_k \exp\left( -\frac{1}{2} e_{ijk}^T \Omega_{ijk} e_{ijk} \right) \]

\[ \log p(z \mid x) \simeq \max_k -\frac{1}{2} e_{ijk}^T \Omega_{ijk} e_{ijk} + \log(w_k \eta_k) \]

or:

\[ -\log p(z \mid x) \simeq \min_k \frac{1}{2} e_{ijk}^T \Omega_{ijk} e_{ijk} - \log(w_k \eta_k) \]
Integration

- With the max-mixture formulation, the log likelihood again results in local quadratic forms
- Easy to integrate in the optimizer:
  1. Evaluate all \( k \) components
  2. Select the component with the maximum log likelihood
  3. Perform the optimization as before using only the max components (as a single Gaussian)
Performance (Gauss vs. MM)

[Images of network diagrams]

Courtesy: E. Olson, P. Agarwal
Runtime

Run time analysis for Intel Dataset

- Cholesky-MM
- Normal Cholesky

Node processed vs. processing time (s)
Max-Mixture and Outliers

- MM formulation is useful for multi-modal measurements (D.A. ambiguities)
- MM is also a handy tool for outliers (D.A. failures)
- Here, one mode represents the edge and a second model uses a flat Gaussian for the outlier hypothesis
Max-Mixture and Outliers

Bi-modal false loop closure

Multi-modal with null-hypothesis

Bi-modal odometry slippage

Courtesy: E. Olson, P. Agarwal
Performance (1 outlier)

Gauss-Newton

MM Gauss-Newton
Performance (10 outliers)

Gauss-Newton

MM Gauss-Newton
Performance (100 outliers)

Gauss-Newton

MM Gauss-Newton
Standard Gaussian Least Squares

\[ X^* = \arg \min_X \sum_{ij} e_{ij}(X)^T \Omega_{ij} e_{ij}(X) \]

\[ \chi^2_{ij} \]
Dynamic Covariance Scaling

\[ X^* = \arg\min_X \sum_{ij} e_{ij}(X)^T \Omega_{ij} e_{ij}(X) \]

\[ X^* = \arg\min_X \sum_{ij} e_{ij}(X)^T \left( s_{ij}^2 \Omega_{ij} \right) e_{ij}(X) \]
Scaling Parameter

\[ X^* = \arg\min_X \sum_{ij} e_{ij}(X)^T \left( s_{ij}^2 \Omega_{ij} \right) e_{ij}(X) \]

\[ s_{ij} = \min \left( 1, \frac{2\Phi}{\Phi + \chi^2_{ij}} \right) \]
Dynamic Covariance Scaling

![Graph showing dynamic covariance scaling]
Dynamic Covariance Scaling

Both have squared error
Dynamic Covariance Scaling

Original error

Scaled error

Squared error
Scaling function (s)
Scaled error for different s
Dynamic Covariance Scaling

![Graph showing Dynamic Covariance Scaling](image_url)

- Squared error
- Scaling function \( (s) \)
- Scaled error for different \( s \)

**Linearization**
Dynamic Covariance Scaling
Optimizing With Outliers

- Assuming a Gaussian error in the measurement is not always realistic
- Large errors are problematic
Robust M-Estimators

- Assume non-normally-distributed noise
- Intuitively: PDF with “heavy tails”
- $\rho(e)$ function used to define the PDF

$$p(e) = \exp(-\rho(e))$$

- Minimizing the neg. log likelihood

$$x^* = \arg\min_x \sum_i \rho(e_i(x))$$
Different Rho Functions

- **Gaussian:** \( \rho(e) = e^2 \)
- **Absolute values (L1 norm):** \( \rho(e) = |e| \)
- **Huber M-estimator**

\[
\rho(e) = \begin{cases} 
\frac{e^2}{2} & \text{if } |e| < c \\
\frac{c}{2}(|e| - \frac{c}{2}) & \text{otherwise}
\end{cases}
\]

- **Several others** (Tukey, Cauchy, Blake-Zisserman, Corrupted Gaussian, ...)
Huber

- Mixture of a quadratic and a linear function

\[ \rho(e) = \begin{cases} 
\frac{e^2}{2} & \text{if } |e| < c \\
\frac{e}{c}(|e| - \frac{c}{2}) & \text{otherwise}
\end{cases} \]
Different Rho Functions

L1 norm
Huber
Tukey

Cauchy
Blake-Zisserman
Corrupted G.
MM Cost Function For Outliers

- Max Mixture
- Corrupted Gaussian
Robust Estimation

- Choice of the rho function depends on the problem at hand
- Huber function is often used
- MM for outlier handling is similar to a corrupted Gaussian
- MM additionally supports multi-model measurements
- Dynamic Covariance Scaling is a robust M-estimator
Conclusions

- Sum of Gaussians cannot be used easily in the optimization framework
- Max-Mixture formulation approximates the sum by the max operator
- This allows for handling data association ambiguities and failures
- Minimal performance overhead
- Minimal code changes for integration
Literature

Max-Mixture Approach:

- Olson, Agarwal: “Inference on Networks of Mixtures for Robust Robot Mapping”

Dynamic Covariance Scaling:

- Agarwal, Tipaldi, Spinello, Stachniss, Burgard: “Robust Map Optimization Using Dynamic Covariance Scaling”
Slide Information

- These slides have been created by Cyrill Stachniss as part of the robot mapping course taught in 2012/13 and 2013/14. I created this set of slides partially extending existing material of Edwin Olson, Pratik Agarwal, and myself.
- I tried to acknowledge all people that contributed image or video material. In case I missed something, please let me know. If you adapt this course material, please make sure you keep the acknowledgements.
- Feel free to use and change the slides. If you use them, I would appreciate an acknowledgement as well. To satisfy my own curiosity, I appreciate a short email notice in case you use the material in your course.
- My video recordings are available through YouTube: http://www.youtube.com/playlist?list=PLgnQPqTFOGQrZ4O5QzbIHgJ3b1JHimN_&feature=g-list

Cyrill Stachniss, 2014
cyrill.stachniss@igg.uniconn.de