Advanced Techniques for Mobile Robotics

Least Squares

Wolfram Burgard, Cyrill Stachniss,

Kai Arras, Maren Bennewitz

Problem

- Given a system described by a set of n observation functions $\{f_i(x)\}_{i=1:n}$
- Let
  - x be the state vector
  - z_i be a measurement of the state x
  - $z'_i=f_i(x)$ be a function which maps x to a predicted measurement z'_i
- Given n noisy measurements $z_{1:n}$ about the state x

Goal: Estimate the state x which bests explains the measurements $z_{1:n}$

Graphical Explanation

$$
\mathbf{x} \quad
\begin{array}{l}
\nearrow \; f_1(\mathbf{x}) = \mathbf{z}'_1 \qquad\qquad \mathbf{z}_1 \\
\rightarrow \; f_2(\mathbf{x}) = \mathbf{z}'_1 \qquad\qquad \mathbf{z}_1 \\
\qquad\qquad\qquad \cdots \\
\searrow \; f_n(\mathbf{x}) = \mathbf{z}'_n \qquad\qquad \mathbf{z}_n
\end{array}
$$

state (unknown)	predicted measurements	real measurements

Example:

- ***x***=position of a set of 3d features in space
- $\boldsymbol{z_i}$=coordinates of the 3d features projected on an image plane
- Estimate the most likely 3d position of the features in the scene given the images ***z*** (given the camera poses)

Error

- The error $\mathbf{e}_i$ is the difference between the predicted measurement and the actual one

$$\mathbf{e}_i(\mathbf{x}) \;=\; \mathbf{z}_i - f_i(\mathbf{x})$$

- We assume the error to be zero mean and normally distributed
- Gaussian error with an information matrix $\boldsymbol{\Omega}_i$
- The squared error of a measurement depends only on the state and it is a scalar

$$e_i(\mathbf{x}) \;=\; \mathbf{e}_i(\mathbf{x})^T \Omega_i \mathbf{e}_i(\mathbf{x})$$

Find the Minimum

- We want to find the state $\mathbf{x}^*$ which minimizes the error given all measurements

$$\begin{aligned}\mathbf{x}^* &= \operatorname*{argmin}_{\mathbf{x}} F(\mathbf{x}) && \text{global error (scalar)} \\ &= \operatorname*{argmin}_{\mathbf{x}} \sum_i e_i(\mathbf{x}) && \text{squared error terms (scalar)} \\ &= \operatorname*{argmin}_{\mathbf{x}} \sum_i \mathbf{e}_i^T(\mathbf{x})\boldsymbol{\Omega}_i\mathbf{e}_i(\mathbf{x}) && \text{error terms (vectors)}\end{aligned}$$

- A general solution is to derive the global error function and find its nulls
- In general, it is a complex problem which does not admit closed form solutions

➡ Numerical approaches

Approximation

- Assume
  - A “good” initial guess is available
  - The error functions are “smooth” in the neighborhood of the (hopefully global) minima
- Then, we can solve the problem by iterative local linearizations
  - Linearize the error terms around the current initial guess
  - Compute the first derivative the squared error
  - Set it to zero and solve the linear system
  - Determine the set of increments that can be summed to the previous estimate of the state to come closer to the minima

Linearizing the Error Function

- We can approximate the error functions around an initial guess ***x*** via Taylor expansion

$$\mathbf{e}_i(\mathbf{x}+\mathbf{\Delta x}) \simeq \mathbf{e}_i + \mathbf{J}_i(\mathbf{x})\mathbf{\Delta x}$$

- Reminder: Jacobian

$$\mathbf{J}_f(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \frac{\partial f_2(x)}{\partial x_1} & \frac{\partial f_2(x)}{\partial x_2} & \cdots & \frac{\partial f_2(x)}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m(x)}{\partial x_1} & \frac{\partial f_m(x)}{\partial x_2} & \cdots & \frac{\partial f_m(x)}{\partial x_n} \end{pmatrix}$$

Squared Error

- With the previous linearization, we can fix ***x*** and carry out the minimization in the increments ***Δx***
- We replace the Taylor expansion in the squared error terms:

$$e_i(\mathbf{x} + \boldsymbol{\Delta}\mathbf{x}) \;=\; \ldots$$

Squared Error

- With the previous linearization, we can fix ***x*** and carry out the minimization in the increments ***Δx***
- We replace the Taylor expansion in the squared error terms:

$$
\begin{aligned}
e_i(\mathbf{x}+\Delta\mathbf{x}) &= \mathbf{e}_i^T(\mathbf{x}+\Delta\mathbf{x})\mathbf{\Omega}_i\mathbf{e}_i(\mathbf{x}+\Delta\mathbf{x}) \\
&\simeq (\mathbf{e}_i+\mathbf{J}_i\Delta\mathbf{x})^T\mathbf{\Omega}_i(\mathbf{e}_i+\mathbf{J}_i\Delta\mathbf{x}) \\
&= \mathbf{e}_i^T\mathbf{\Omega}_i\mathbf{e}_i + \\
&\quad \mathbf{e}_i^T\mathbf{\Omega}_i\mathbf{J}_i\Delta\mathbf{x} + \Delta\mathbf{x}^T\mathbf{J}_i^T\mathbf{\Omega}_i\mathbf{e}_i + \\
&\quad \Delta\mathbf{x}^T\mathbf{J}_i^T\mathbf{\Omega}_i\mathbf{J}_i\Delta\mathbf{x}
\end{aligned}
$$

Squared Error (cont.)

- All summands are scalar so the transposition has no effect
- By grouping similar terms, we obtain:

$$
\begin{aligned}
& e_i(\mathbf{x} + \mathbf{\Delta x}) \\
\simeq \quad & \mathbf{e}_i^T \mathbf{\Omega}_i \mathbf{e}_i + \\
& \mathbf{e}_i^T \mathbf{\Omega}_i \mathbf{J}_i \mathbf{\Delta x} + \mathbf{\Delta x}^T \mathbf{J}_i^T \mathbf{\Omega}_i \mathbf{e}_i + \\
& \mathbf{\Delta x}^T \mathbf{J}_i^T \mathbf{\Omega}_i \mathbf{J}_i \mathbf{\Delta x} \\
= \quad & \underbrace{\mathbf{e}_i^T \mathbf{\Omega}_i \mathbf{e}_i}_{c_i} + 2 \underbrace{\mathbf{e}_i^T \mathbf{\Omega}_i \mathbf{J}_i}_{\mathbf{b}_i^T} \mathbf{\Delta x} + \mathbf{\Delta x}^T \underbrace{\mathbf{J}_i^T \mathbf{\Omega}_i \mathbf{J}_i}_{\mathbf{H}_i} \mathbf{\Delta x} \\
= \quad & c_i + 2\mathbf{b}_i^T \mathbf{\Delta x} + \mathbf{\Delta x}^T \mathbf{H}_i \mathbf{\Delta x}
\end{aligned}
$$

Global Error

- Thus, the global error is the sum of the squared errors terms corresponding to the individual measurements
- We can use the new terms for the squared error to a new expression which approximates the global error in the neighborhood of the current solution ***x***

$$
\begin{aligned}
F(\mathbf{x}+\boldsymbol{\Delta}\mathbf{x}) &\simeq \sum_i \Big(c_i + \mathbf{b}_i^T \boldsymbol{\Delta}\mathbf{x} + \boldsymbol{\Delta}\mathbf{x}^T \mathbf{H}_i \boldsymbol{\Delta}\mathbf{x}\Big) \\
&= \sum_i c_i + 2(\sum_i \mathbf{b}_i^T)\boldsymbol{\Delta}\mathbf{x} + \boldsymbol{\Delta}\mathbf{x}^T(\sum_i \mathbf{H}_i)\boldsymbol{\Delta}\mathbf{x}
\end{aligned}
$$

Global Error (cont.)

$$
\begin{aligned}
F(\mathbf{x}+\Delta\mathbf{x}) &\simeq \sum_i \left(c_i + \mathbf{b}_i^T\Delta\mathbf{x} + \Delta\mathbf{x}^T\mathbf{H}_i\Delta\mathbf{x}\right) \\
&= \underbrace{\sum_i c_i}_{c} + 2\,\underbrace{(\sum_i \mathbf{b}_i^T)}_{\mathbf{b}^T}\,\Delta\mathbf{x} + \Delta\mathbf{x}^T\,\underbrace{(\sum_i \mathbf{H}_i)}_{\mathbf{H}}\,\Delta\mathbf{x} \\
&= c + 2\mathbf{b}^T\Delta\mathbf{x} + \Delta\mathbf{x}^T\mathbf{H}\Delta\mathbf{x}
\end{aligned}
$$

with

$$
\begin{aligned}
\mathbf{b}^T &= \sum_i \mathbf{e}_i^T\mathbf{\Omega}_i\mathbf{J}_i \\
\mathbf{H} &= \sum_i \mathbf{J}_i^T\mathbf{\Omega}\mathbf{J}_i
\end{aligned}
$$

Quadratic Form

- We can write the global error terms into a quadratic form in ***Δ***x

$$F(\mathbf{x}+\boldsymbol{\Delta}\mathbf{x}) \;\simeq\; c + 2\mathbf{b}^T\boldsymbol{\Delta}\mathbf{x} + \boldsymbol{\Delta}\mathbf{x}^T\mathbf{H}\boldsymbol{\Delta}\mathbf{x}$$

- The approximated derivative of ***F(x+Δx)*** with respect to ***Δx*** in the neighborhood of the current solution ***x*** is:

$$\frac{\partial F(\mathbf{x}+\boldsymbol{\Delta}\mathbf{x})}{\partial\boldsymbol{\Delta}\mathbf{x}} \;\simeq\; 2\mathbf{b} + 2\mathbf{H}\boldsymbol{\Delta}\mathbf{x}$$

Linear System

- The approximated derivative of ***F(x+Δx)*** is:

$$\frac{\partial F(\mathbf{x}+\boldsymbol{\Delta}\mathbf{x})}{\partial \boldsymbol{\Delta}\mathbf{x}} \simeq 2\mathbf{b} + 2\mathbf{H}\boldsymbol{\Delta}\mathbf{x}$$

- Setting it to zero leads to

$$0 = 2\mathbf{b} + 2\mathbf{H}\boldsymbol{\Delta}\mathbf{x}$$

- Which is the linear system

$$\mathbf{H}\boldsymbol{\Delta}\mathbf{x} = -\mathbf{b}$$

- Thus, the optimum ***Δx**** is

$$\boldsymbol{\Delta}\mathbf{x}^* = -\mathbf{H}^{-1}\mathbf{b}$$

Iterative Solution: Gauss-Newton

Repeatedly perform the following steps:

- Linearize the system around the current guess ***x*** and compute for each measurement

$$\mathbf{e}_i(\mathbf{x} + \boldsymbol{\Delta}\mathbf{x}) \simeq \mathbf{e}_i(\mathbf{x}) + \mathbf{J}_i \boldsymbol{\Delta}\mathbf{x}$$

- Compute the terms for the linear system

$$\mathbf{b}^T = \sum_i \mathbf{e}_i^T \boldsymbol{\Omega}_i \mathbf{J}_i \qquad \mathbf{H} = \sum_i \mathbf{J}_i^T \boldsymbol{\Omega}_i \mathbf{J}_i$$

- Solve the system to get a new increment

$$\boldsymbol{\Delta}\mathbf{x}^* \;=\; -\mathbf{H}^{-1}\mathbf{b}$$

- Updating the previous estimate

$$\mathbf{x} \leftarrow \mathbf{x} + \boldsymbol{\Delta}\mathbf{x}^*$$

Example: Odometry Calibration

- We have a robot which moves in an environment, gathering the odometry measurements $\boldsymbol{u}_i$
- The odometry is affected by a systematic error which we want to eliminate through calibration
- For each $\boldsymbol{u}_i$, we have a ground truth (estimate) $\boldsymbol{u}^*_i$, which can, for example, be approximated by scan-matching or a SLAM system

Example: Odometry Calibration

- There is a function ***f_i(x)*** which, given some bias parameters ***x***, returns a an unbiased (corrected) odometry for the reading ***u_i'*** as follows

$$\mathbf{u}'_i = f_i(\mathbf{x}) = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} \mathbf{u}_i$$

- To obtain this function ***f(x)***, we need to find the parameters ***x***

Odometry Calibration (cont.)

- The state vector is

$$\mathbf{x} = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{21} & x_{22} & x_{23} & x_{31} & x_{32} & x_{33} \end{pmatrix}^T$$

- The error function is

$$\mathbf{e}_i(\mathbf{x}) = \mathbf{u}_i^* - \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} \mathbf{u}_i$$

- Its derivative is:

$$\mathbf{J}_i = \frac{\partial \mathbf{e}_i(\mathbf{x})}{\partial \mathbf{x}} = -\begin{pmatrix} u_{i,x} & u_{i,y} & u_{i,\theta} & & & & & & \\ & & & u_{i,x} & u_{i,y} & u_{i,\theta} & & & \\ & & & & & & u_{i,x} & u_{i,y} & u_{i,\theta} \end{pmatrix}$$

Does not depend on ***x***, why? What are the consequences?

e is linear, no need to iterate!

Questions

- How do the parameters look like if the odometry is perfect?
- How many measurements (at least) are needed to find a solution for the calibration problem?
- ***H*** is symmetric. Why?
- How does the structure of the measurement function affects the structure of ***H?***

How to Solve the Linear System?

- Linear system $\mathbf{H}\Delta\mathbf{x} = -\mathbf{b}$
- Can be solved by matrix inversion (in theory)
- In practice:
  - Cholesky factorization
  - QR decomposition
  - Iterative methods such as conjugate gradients (for large systems)

Summary

- Technique to minimize a squared error function
- Presented approach is known as the Gauss-Newton method
- Start with an initial guess
- Approximate the error terms by linear functions
- This leads to a quadratic form
- One obtains a linear system by settings its derivative to zero
- Solving the linear systems leads to the next state
- Iterate this procedure