Least Squares and SLAM
Least Squares

Giorgio Grisetti

Part of the material of this course is taken from the Robotics 2 lectures given by G.Grisetti, W.Burgard, C.Stachniss, K.Arras, D. Tipaldi and M.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
- We acquire n noisy measurements $z_{1:n}$ about the state x

We want to estimate the state x which bests explains the measurements $z_{1:n}$

Graphical Explanation

$$\mathbf{x}$$

$$f_1(\mathbf{x}) = \mathbf{z}'_1 \qquad \mathbf{z}_1$$

$$f_2(\mathbf{x}) = \mathbf{z}'_1 \qquad \mathbf{z}_1$$

$$\cdots$$

$$f_n(\mathbf{x}) = \mathbf{z}'_n \qquad \mathbf{z}_n$$

state (unknown)

predicted measurements

real measurements

Example:

- $\boldsymbol{x}$= position of a set of 3d features
- $\boldsymbol{z_i}$=coordinates of the 3d features projected on an image plane w.r.t. the $\boldsymbol{i}$**th** observation point
- Estimate the most likely 3d position of the features in the scene given the images $\boldsymbol{z}$

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 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 of all measurements

$$
\begin{aligned}
\mathbf{x}^* &= \operatorname*{argmin}_{\mathbf{x}} F(\mathbf{x}) && \text{Global Error (scalar)} \\
&= \operatorname*{argmin}_{\mathbf{x}} \sum_i e_i(\mathbf{x}) && \text{Squared Error Terms (scalar)} \\
& && \text{Error Terms (vectors)} \\
&= \operatorname*{argmin}_{\mathbf{x}} \sum_i \mathbf{e}_i^T(\mathbf{x}) \boldsymbol{\Omega}_i \mathbf{e}_i(\mathbf{x})
\end{aligned}
$$

- 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

$\Rightarrow$ Numerical approaches

Approximations

- If
  - A good initial guess is available and
  - the measurement functions are "smooth" in the neighborhood of the (hopefully global) minima
- We can solve the problem by iterative local linearizations
  - Linearize the problem around the current initial guess
  - Solve a linear system
  - Determine a set of increments which 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

$$
\begin{aligned}
\mathbf{e}_i(\mathbf{x} + \boldsymbol{\Delta}\mathbf{x}) &\simeq \mathbf{e}_i(\mathbf{x}) + \frac{\partial \mathbf{e}_i}{\partial \mathbf{x}} \boldsymbol{\Delta}\mathbf{x} \\
&= \mathbf{e}_i + \mathbf{J}_i \boldsymbol{\Delta}\mathbf{x}
\end{aligned}
$$

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:

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

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:

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

Squared Error (cont)

- All summands are scalar so the transposition has no effect
- We can group the similar terms.

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

Global Error

- The global error is the sum of the squared errors of the 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***

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

Global Error (cont)

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

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

- The global error turns into a quadratic form in ***Δ***x and can now be minimized

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

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

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

- The optimum ***Δx**** is

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

Iterative Solution

- The minimization algorithm proceeds by repeatedly performing the following steps:
  - Linearizing the system around the current guess ***x*** and computing the following quantities for each measurement

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

  - Computing the terms for the linear system

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

  - Solving the system to get a new optimal 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 $\boldsymbol{u^*_i}$ which can, for example, be approximated by scan-matching or a SLAM procedure

Example: Odometry Calibration

- There is a function $\boldsymbol{f_i(x)}$ which, given some bias parameters $\boldsymbol{x}$, returns a an unbiased odometry for the reading $\boldsymbol{u_i'}$ as follows

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

- The goal is to find the parameters $\boldsymbol{x}$

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? $\Rightarrow$ ***e*** is linear, no need to iterate!

Some 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?***