UNIVERSITEIT·STELLENBOSCH·UNIVERSITY
jou kennisvennoot • your knowledge partner

Numerical integration for probabilisitc reasoning skripsie

by

Jacobus Martin Louw

Pectora roborant cultus recti

Bachelor of Engineering (Electrical and Electronic)

Final Report

Study leader: Dr. CE van Daalen

2018

Declaration

By submitting this report electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: ....

Abstract

Text in default language ...

1 Introduction

For modern mobile vehicles and robots it is important to have the ability to navigate their environment. It is usually critical for these devices to avoid collisions or dangerous environments. Some robots must move very precisely and therefore should have an accurate reading of their location.

Localisation is essential for robot navigation, but is used in various other applications such missile tracking. For applications like this, it is crucial to have accurate and instantaneous information of the missile's location in space. Measurements of any object's location will always have some noise, therefore the measured location is never 100% accurate. Therefore, one should rather approach the localisation problem in a probabilistic manner. A probability density region where the object is most likely can be calculated.

For systems with continuous random variables, most of the operations used in probabilistic reasoning use integration. These integrals can be solved analytically in the case of a problem with linear movement. Most systems are unfortunately nonlinear. The integrals in nonlinear systems cannot be computed analytically and one has to resort to numerical methods. Commonly-used techniques such as the extend or unscented Kalman filters use rudimentary numerical integration that are very inaccurate in some scenarios. There are several numerical techniques available that are more accurate.

The end goal of this project is to compare different numerical techniques to solve the nonlinear localisation problem. To reach the end goal, one should first have a good understanding of Gaussian random variables and traditional techniques such as the extended - and unscented kalman filters. Modeling the problem with Probabilistic Graphical Models has a lot of advantages and is therefore also investigated.

A relevant problem has been simulated in Python. Different techniques were implemented and compared in terms of accuracy and efficiency.

2 Gaussian Random Variables

The Gaussian or normal distribution is commonly used in probability theory. The Gaussian random variable (RV) is a very important concept in this paper as all probability distributions are approximated as Gaussian distributions. Key concepts and features of the Gaussian RV are discussed in this chapter. The canonical form and conditional distributions are very important concepts and will be used in following chapters.

2.1 Covariance Form

$$p(x) = \mathcal{N}(\mu; \sigma) = \eta \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] \tag{2.1}$$

$$\eta = \frac{1}{\sqrt{2\pi\sigma^2}} \tag{2.2}$$

$$p(x) = \mathcal{N}(\boldsymbol{\mu}; \Sigma) = \eta \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x}-\boldsymbol{\mu})\right] \tag{2.3}$$

$$\eta = \frac{1}{(2\pi^{n/2})|\Sigma|^{\frac{1}{2}}} \tag{2.4}$$

2.2 Canonical Form

Linear Gaussian CPDs (conditional probability densities) are conditional distributions and are generally not Gaussians. It will be handy to find general representation that accomodates both Gaussian distributions and linear Gaussian models. The canonical or information form is a viable option. It is also much easier to perform certain operations in the canonical form.

$$\mathcal{C}(\boldsymbol{X}; K, \boldsymbol{h}, g) = \exp\left(-\frac{1}{2}\boldsymbol{X}^T K \boldsymbol{X} + \boldsymbol{h}^T \boldsymbol{X} + g\right) \tag{2.5}$$

It is possible to represent every Gaussian as a canonical form. Equation 2.3 can be rewritten:

$$\eta \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]$$
$$= \exp\left(-\frac{1}{2}\boldsymbol{x}^T\Sigma^{-1}\boldsymbol{x} + \boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{x} - \frac{1}{2}\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu} + \ln\eta\right) \quad (2.6)$$

$\mathcal{N}(\boldsymbol{\mu};\Sigma) = \mathcal{C}(K, \boldsymbol{h}, g)$ by comparing 2.6 with 2.5:

$$K = \Sigma^{-1} \quad (2.7)$$
$$\boldsymbol{h} = \Sigma^{-1}\boldsymbol{\mu} \quad (2.8)$$
$$g = -\frac{1}{2}\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu} + \ln\eta \quad (2.9)$$

The covariance parameters can again be recovered.

$$\Sigma = K^{-1} \quad (2.10)$$
$$\boldsymbol{\mu} = \Sigma\boldsymbol{h} \quad (2.11)$$

The covariance form is not defined when K is not invertible, even though the distribution can still be represented in the canonical form. The canonical form is therefore more general than the covariance form. The canonical form is very useful to present linear Gaussian CPDs. From 2.9 it can be seen that only K and $\boldsymbol{h}$ are necessary to calculate g. Thus, g can be omitted when working in the canonical form.

2.1.1 Operations using the canonical form

The main advantage of using the canonical form is that it is very easy to perform various operations on distributions. The variable g is included in the following operations for completeness sake, but can be omitted as stated above.

Multiplication of canonical forms:
When multiplying to canonical forms, it is important that their scopes are identical. Multiplying canonical forms with the same scope is simply:

$$\mathcal{C}(K_1, \boldsymbol{h_1}, g_1) \times \mathcal{C}(K_2, \boldsymbol{h_2}, g_2) = \mathcal{C}(K_1 + K_2, \boldsymbol{h_1} + \boldsymbol{h_2}, g_1 + g_2) \quad (2.12)$$

Division of canonical forms:
Again, it important that the scopes of the distributions are identical.

$$\frac{\mathcal{C}(K_1, \boldsymbol{h_1}, g_1)}{\mathcal{C}(K_2, \boldsymbol{h_2}, g_2)} = \mathcal{C}(K_1 - K_2, \boldsymbol{h_1} - \boldsymbol{h_2}, g_1 - g_2) \quad (2.13)$$

Marginalization of a canonical form:
A marginal distribution can be found by integrating over a subset of variables. For example the marginal distribution over $\boldsymbol{X}$ can be found by integrating over $\boldsymbol{Y}$.
Let $\mathcal{C}(\boldsymbol{X}, \boldsymbol{Y}; K, \boldsymbol{h}, g)$ be a canonical form with subsets $\boldsymbol{X}$ and $\boldsymbol{Y}$ where

$$K = \begin{bmatrix} K_{\boldsymbol{XX}} & K_{\boldsymbol{XY}} \\ K_{\boldsymbol{XY}} & K_{\boldsymbol{YY}} \end{bmatrix}; h = \begin{pmatrix} h_{\boldsymbol{X}} \\ h_{\boldsymbol{Y}} \end{pmatrix} \tag{2.14}$$

To obtain the marginal distribution over $\boldsymbol{X}$, we have to find the integral over $\boldsymbol{Y}$. Therefore

$$\int \mathcal{C}(\boldsymbol{X}, \boldsymbol{Y}; K, \boldsymbol{h}, g) d\boldsymbol{Y} = \mathcal{C}(\boldsymbol{X}; K', \boldsymbol{h}', g') \tag{2.15}$$

Where

$$K' = K_{\boldsymbol{XX}} - K_{\boldsymbol{XY}} K_{\boldsymbol{XY}}^{-1} K_{\boldsymbol{YX}} \tag{2.16}$$

$$h' = \boldsymbol{h}_{\boldsymbol{X}} - K_{\boldsymbol{XX}} K_{\boldsymbol{YY}}^{-1} \boldsymbol{h}_{\boldsymbol{Y}} \tag{2.17}$$

$$g' = g - \frac{1}{2} \left(\ln |2\pi K_{\boldsymbol{YY}}^{-1}| + \boldsymbol{h}_{\boldsymbol{Y}}^T K_{\boldsymbol{YY}}^{-1} \boldsymbol{h}_{\boldsymbol{Y}} \right) \tag{2.18}$$

It is important to that $K_{\boldsymbol{YY}}$ is positive definite for the result to be finite.

Reduction with evidence:
Let $\mathcal{C}(\boldsymbol{X}, \boldsymbol{Y}; K, \boldsymbol{h}, g)$ be a canonical form with subsets $\boldsymbol{X}$ and $\boldsymbol{Y}$. If subset $\boldsymbol{Y}$ is known ($\boldsymbol{Y} = \boldsymbol{y}$), then the canonical form can be reduced to $\mathcal{C}(\boldsymbol{X}; K', \boldsymbol{h}', g')$, where

$$K' = K_{\boldsymbol{XX}} \tag{2.19}$$

$$h' = \boldsymbol{h}_{\boldsymbol{X}} - K_{\boldsymbol{XY}} \boldsymbol{y} \tag{2.20}$$

$$g' = g + \boldsymbol{h}_{\boldsymbol{y}}^T \boldsymbol{y} - \frac{1}{2} \boldsymbol{y}^T K_{\boldsymbol{YY}} \boldsymbol{y} \tag{2.21}$$