Course 07: Extended Information Filter (EIF)

Saturday, November 9, 2019 9:41 PM

slam07-eif

Robot Mapping

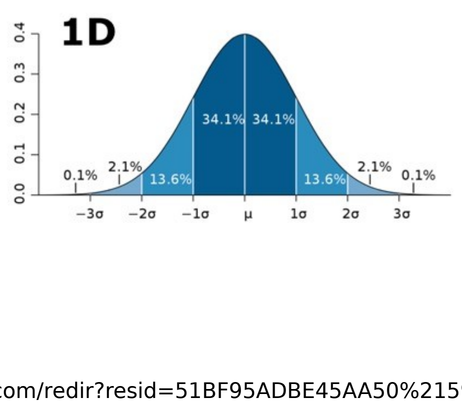
Extended Information Filter

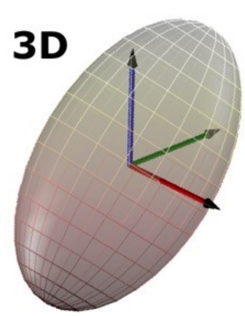
Cyrill Stachniss

Gaussians

- Gaussian described by **moments** μ, Σ

$$p(x) = \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$





Canonical Parameterization

- Alternative representation for Gaussians
- Described by **information matrix** Ω and **information vector** ξ

3

Canonical Parameterization

- Alternative representation for Gaussians
- Described by **information matrix** Ω

$$\Omega = \Sigma^{-1}$$

- and **information vector** ξ

$$\xi = \Sigma^{-1}\mu$$

Note: some things that are easy in regular space could be difficult in information matrix space; Vice Versa.

4

Complete Parameterizations

moments	canonical
$\Sigma = \Omega^{-1}$	$\Omega = \Sigma^{-1}$

$$\Sigma = \Omega^{-1} \qquad \Omega = \Sigma^{-1}$$

$$\mu = \Omega^{-1}\xi \qquad \xi = \Sigma^{-1}\mu$$

* cost of performing such process:

inverse the matrix, $O(n^{2.4})$ → we can do it, but typically we don't want to do it. as cost inefficient.

5

Towards the Information Form

$$
\begin{aligned}
& p(x) \\
= \quad & \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big)
\end{aligned}
$$

6

Towards the Information Form

$$
\begin{aligned}
& p(x) \\
= \quad & \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(\underline{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}\Big) \\
= \quad & \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu - \frac{1}{2}\mu^T\Sigma^{-1}\mu\Big)
\end{aligned}
$$

simply split it.

7

Towards the Information Form

$$
\begin{aligned}
&p(x)\\
&= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big)\\
&= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(\underline{-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu} \underline{- \frac{1}{2}\mu^T\Sigma^{-1}\mu}\Big)\\
&= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}\mu^T\Sigma^{-1}\mu\Big)\\
&\quad \exp\Big(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu\Big)
\end{aligned}
$$

special ???

↳ It's independent w/ x, i.e. nothing to do with our variable.

It's constant.

8

Towards the Information Form

$$
\begin{aligned}
&p(x)\\
&= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big)\\
&= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu - \frac{1}{2}\mu^T\Sigma^{-1}\mu\Big)\\
&= \underline{\det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}\mu^T\Sigma^{-1}\mu\Big)}\\
&\quad \exp\Big(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu\Big)\\
&= \eta \exp\Big(-\frac{1}{2}x^T\underbrace{\Sigma^{-1}}_{\Omega}x + x^T\underbrace{\Sigma^{-1}\mu}_{\xi}\Big)
\end{aligned}
$$

9

Towards the Information Form

$$
\begin{aligned}
p(x) &= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big) \\
&= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu - \frac{1}{2}\mu^T\Sigma^{-1}\mu\Big) \\
&= \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}\mu^T\Sigma^{-1}\mu\Big) \\
&\qquad \exp\Big(-\frac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu\Big) \\
&= \eta \exp\Big(-\frac{1}{2}x^T\underline{\Sigma^{-1}}x + x^T\underline{\Sigma^{-1}\mu}\Big) \\
&= \eta \exp\Big(-\frac{1}{2}x^T\Omega x + x^T\xi\Big)
\end{aligned}
$$

10

Dual Representation

$$
p(x) = \frac{\exp(-\frac{1}{2}\mu^T\xi)}{\det(2\pi\Omega^{-1})^{\frac{1}{2}}} \exp\Big(-\frac{1}{2}\underbrace{x^T\Omega x}_{\text{quadratic form}} + x^T\xi\Big)
$$

canonical parameterization

$$
p(x) = \det(2\pi\Sigma)^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big)
$$

moments parameterization

11

Marginalization and Conditioning

① $\Sigma^{-1} \to \Omega$ ② $\Sigma^{-1}\mu \to \xi$

$$
p(\boldsymbol{\alpha},\boldsymbol{\beta}) = \mathcal{N}\Big(\begin{bmatrix}\boldsymbol{\mu}_\alpha \\ \boldsymbol{\mu}_\beta\end{bmatrix}, \begin{bmatrix}\Sigma_{\alpha\alpha} & \Sigma_{\alpha\beta} \\ \Sigma_{\beta\alpha} & \Sigma_{\beta\beta}\end{bmatrix}\Big) = \mathcal{N}^{-1}\Big(\begin{bmatrix}\boldsymbol{\eta}_\alpha \\ \boldsymbol{\eta}_\beta\end{bmatrix}, \begin{bmatrix}\Lambda_{\alpha\alpha} & \Lambda_{\alpha\beta} \\ \Lambda_{\beta\alpha} & \Lambda_{\beta\beta}\end{bmatrix}\Big)
$$

	MARGINALIZATION $p(\boldsymbol{\alpha}) = \int p(\boldsymbol{\alpha},\boldsymbol{\beta})\,d\boldsymbol{\beta}$	CONDITIONING $p(\boldsymbol{\alpha} \mid \boldsymbol{\beta}) = p(\boldsymbol{\alpha},\boldsymbol{\beta})/p(\boldsymbol{\beta})$
COV. FORM	$\boldsymbol{\mu} = \boldsymbol{\mu}_\alpha$ $\Sigma = \Sigma_{\alpha\alpha}$	$\boldsymbol{\mu}' = \boldsymbol{\mu}_\alpha + \Sigma_{\alpha\beta}\Sigma_{\beta\beta}^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu}_\beta)$ $\Sigma' = \Sigma_{\alpha\alpha} - \Sigma_{\alpha\beta}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta\alpha}$
INFO	$\boldsymbol{\eta} = \boldsymbol{\eta}_\alpha - \Lambda_{\alpha\beta}\Lambda_{\beta\beta}^{-1}\boldsymbol{\eta}_\beta$	$\boldsymbol{\eta}' = \boldsymbol{\eta}_\alpha - \Lambda_{\alpha\beta}\boldsymbol{\beta}$

INFO. FORM	$\Lambda = \Lambda_{\alpha\alpha} - \Lambda_{\alpha\beta}\Lambda_{\beta\beta}^{-1}\Lambda_{\beta\alpha}$	$\Lambda' = \Lambda_{\alpha\alpha}$

trivial

expensive

Note: depends what we want to do, one or the other representation would be advantageous.

Courtesy: R. Eustice 12

From the Kalman Filter to the Information Filter

① $\Sigma^{-1} \to \Omega$ ② $\Sigma^{-1}\mu = \xi$

- Two parameterization for Gaussian
- Same expressiveness
- Marginalization and conditioning have different complexities
- We learned about Gaussian filtering with the Kalman filter in Chapter 4
- Kalman filtering in information from is called information filtering

13

Kalman Filter Algorithm

1: **Kalman_filter**$(\mu_{t-1}, \Sigma_{t-1}, u_t, z_t)$**:**

2: $\boxed{\bar{\mu}_t = A_t\, \mu_{t-1} + B_t\, u_t}$
3: $\boxed{\bar{\Sigma}_t = A_t\, \Sigma_{t-1}\, A_t^T + R_t}$

4: $K_t = \bar{\Sigma}_t\, C_t^T (C_t\, \bar{\Sigma}_t\, C_t^T + Q_t)^{-1}$
5: $\mu_t = \bar{\mu}_t + K_t(z_t - C_t\, \bar{\mu}_t)$
6: $\Sigma_t = (I - K_t\, C_t)\, \bar{\Sigma}_t$
7: return μ_t, Σ_t

$\bar{\Sigma}_t = A_t \bar{\Sigma}_{t-1} A_t^T + R_t$

$\bar{\Omega}_t = \bar{\Sigma}_t^{-1} = (A_t \bar{\Sigma}_{t-1} A_t^T +$

$= (A_t \Omega_{t-1}^{-1}$

$\bar{\xi}_t = \bar{\Omega}_t \bar{\mu}_t$

$= \bar{\Omega}_t (A\mu_{t-1} + Bu_t)$

$= \bar{\Omega}_t (A\Omega_{t-1}^{-1}\xi_{t-1} + B$

14

Prediction Step (1)

- Transform $\bar{\Sigma}_t = A_t \, \Sigma_{t-1} \, A_t^T + R_t$
- Using $\Sigma_{t-1} = \Omega_{t-1}^{-1}$
- Leads to

$$\bar{\Omega}_t \quad = \quad (A_t \, \Omega_{t-1}^{-1} \, A_t^T + R_t)^{-1}$$

15

Prediction Step (2)

- Transform $\bar{\mu}_t = A_t \, \mu_{t-1} + B_t \, u_t$
- Using $\bar{\mu}_{t-1} = \Omega_{t-1}^{-1} \xi_{t-1}$
- Leads to

$$\begin{aligned} \bar{\xi}_t \quad &= \quad \bar{\Omega}_t (A_t \, \mu_{t-1} + B_t \, u_t) \\ &= \quad \bar{\Omega}_t (A_t \, \Omega_{t-1}^{-1} \xi_{t-1} + B_t \, u_t) \end{aligned}$$

16

Information Filter Algorithm

$O(n^3)$ costly !

1: **Information_filter**$(\xi_{t-1}, \Omega_{t-1}, u_t, z_t)$**:**

2: $\bar{\Omega}_t = (A_t \, \Omega_{t-1}^{-1} \, A_t^T + R_t)^{-1}$

prediction step becomes costly.

3:	$\xi_t = \Omega_t(A_t \, \Omega_{t-1}^{-1} \, \xi_{t-1} + B_t \, u_t)$
4:	
5:	
6:	

17

Correction Step

- Use the Bayes filter measurement update and replace the components

$$bel(x_t) = \eta \, p(z_t \mid x_t) \, \overline{bel}(x_t)$$

$$= \eta' \exp\left(-\frac{1}{2}(z_t - \underbrace{C_t x_t}_{\text{predicted observation}})^T Q_t^{-1} (z_t - C_t x_t)\right) \exp\left(-\frac{1}{2}(x_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1} (x_t - \bar{\mu}_t)\right)$$

18

Correction Step

- Use the Bayes filter measurement update and replace the components

$$bel(x_t) = \eta \, p(z_t \mid x_t) \, \overline{bel}(x_t)$$

$$= \eta' \exp\left(-\frac{1}{2}(z_t - C_t x_t)^T Q_t^{-1} (z_t - C_t x_t)\right) \exp\left(-\frac{1}{2}(x_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1} (x_t - \bar{\mu}_t)\right)$$

$$= \eta' \exp\left(-\frac{1}{2}(z_t - C_t x_t)^T Q_t^{-1} (z_t - C_t x_t) - \frac{1}{2}(x_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1} (x_t - \bar{\mu}_t)\right)$$

19

Correction Step

- Use the Bayes filter measurement update and replace the components

$$
\begin{aligned}
bel(x_t) &= \eta\, p(z_t \mid x_t)\, \overline{bel}(x_t) \\
&= \eta' \exp\left(-\frac{1}{2}(z_t - C_t x_t)^T Q_t^{-1}(z_t - C_t x_t)\right) \exp\left(-\frac{1}{2}(x_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1}(x_t - \bar{\mu}_t)\right) \\
&= \eta' \exp\left(-\frac{1}{2}(z_t - C_t x_t)^T Q_t^{-1}(z_t - C_t x_t) - \frac{1}{2}(x_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1}(x_t - \bar{\mu}_t)\right) \\
&= \eta'' \exp\left(-\frac{1}{2} x_t^T C_t^T Q_t^{-1} C_t x_t + x_t^T C_t^T Q_t^{-1} z_t - \frac{1}{2} x_t^T \bar{\Omega}_t x_t + x_t^T \bar{\xi}_t\right)
\end{aligned}
$$

constant term absorbs other info and changes itself.

20

Correction Step

- Use the Bayes filter measurement update and replace the components

$$
\begin{aligned}
bel(x_t) &= \eta\, p(z_t \mid x_t)\, \overline{bel}(x_t) \\
&= \eta' \exp\left(-\frac{1}{2}(z_t - C_t x_t)^T Q_t^{-1}(z_t - C_t x_t)\right) \exp\left(-\frac{1}{2}(x_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1}(x_t - \bar{\mu}_t)\right) \\
&= \eta' \exp\left(-\frac{1}{2}(z_t - C_t x_t)^T Q_t^{-1}(z_t - C_t x_t) - \frac{1}{2}(x_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1}(x_t - \bar{\mu}_t)\right) \\
&= \eta'' \exp\left(-\frac{1}{2} x_t^T C_t^T Q_t^{-1} C_t x_t + x_t^T C_t^T Q_t^{-1} z_t - \frac{1}{2} x_t^T \bar{\Omega}_t x_t + x_t^T \bar{\xi}_t\right) \\
&= \eta'' \exp\left(-\frac{1}{2} x_t^T \underbrace{[C_t^T Q_t^{-1} C_t + \bar{\Omega}_t]}_{\Omega_t} x_t + x_t^T \underbrace{[C_t^T Q_t^{-1} z_t + \bar{\xi}_t]}_{\xi_t}\right)
\end{aligned}
$$

21

Correction Step

- This results in a simple update rule

This results in a simple update rule

$$bel(x_t) \;=\; \eta \exp\left(-\frac{1}{2}\, x_t^T \underbrace{[C_t^T\, Q_t^{-1}\, C_t + \bar{\Omega}_t]}_{\Omega_t}\, x_t + x_t^T \underbrace{[C_t^T\, Q_t^{-1}\, z_t + \bar{\xi}_t]}_{\xi_t}\right)$$

come as our definition

$$\begin{aligned} \Omega_t &= C_t^T\, Q_t^{-1}\, C_t + \bar{\Omega}_t \\ \xi_t &= C_t^T\, Q_t^{-1}\, z_t + \bar{\xi}_t \end{aligned}$$

↳ dim of observation

↳ usually we can do it very efficiently.

Information Filter Algorithm

1: **Information_filter**$(\xi_{t-1}, \Omega_{t-1}, u_t, z_t)$**:**

2: $\bar{\Omega}_t = (A_t\, \Omega_{t-1}^{-1}\, A_t^T + R_t)^{-1}$

3: $\bar{\xi}_t = \bar{\Omega}_t(A_t\, \Omega_{t-1}^{-1}\, \xi_{t-1} + B_t\, u_t)$

4: $\Omega_t = C_t^T\, Q_t^{-1}\, C_t + \bar{\Omega}_t$

5: $\xi_t = C_t^T\, Q_t^{-1}\, z_t + \bar{\xi}_t$

6: *return* ξ_t, Ω_t

cost is lower than the pred step

Prediction and Correction

- Prediction

$$\begin{aligned} \bar{\Omega}_t &= (A_t\, \Omega_{t-1}^{-1}\, A_t^T + R_t)^{-1} \\ \bar{\xi}_t &= \bar{\Omega}_t(A_t\, \Omega_{t-1}^{-1}\, \xi_{t-1} + B_t\, u_t) \end{aligned}$$

- Correction

$$\Omega_t = C_t^T\, Q_t^{-1}\, C_t + \bar{\Omega}_t$$

$$\xi_t = C_t^T Q_t^{-1} z_t + \bar{\xi}_t$$

Discuss differences to the KF!

{ cheap in prediction step; expensive in correction step.

24

Exactly the opposite for Information Filte

Complexity

Note: "sparse" matrix v.s. "dense" matrix

majority

- Kalman filter
  - Efficient prediction step: $\mathcal{O}(n^2)^*$
  - Costly correction step: $\mathcal{O}(n^2 + k^{2.4})$
- Information filter
  - Costly prediction step: $\mathcal{O}(n^{2.4})$
  - Efficient correction step: $\mathcal{O}(n^2)^*$
- Transformation between both parameterizations is costly: $\mathcal{O}(n^{2.4})$

*Potentially faster, especially for SLAM; depending on type of controls and observations

In the n
to proce
to ma
it im
inversi

25

Extended Information Filter

- As the Kalman filter, the information filter suffers from the linear models
- The extended information filter (EIF) uses a similar trick as the EKF
- Linearization of the motion and observation function

Ethan: should be non-linear, right?

26

Linearization of the EIF

- Taylor approximation analog to the EKF (see Chapter 3)

$$\begin{aligned} g(u_t, x_{t-1}) &\approx g(u_t, \mu_{t-1}) + G_t\,(x_{t-1} - \mu_{t-1}) \\ h(x_t) &\approx h(\bar{\mu}_t) + H_t\,(x_t - \bar{\mu}_t) \end{aligned}$$

- with the Jacobians G_t and H_t

27

Prediction: From EKF of EIF

- Substitution of the moments brings us from the EKF

$$\begin{aligned} \bar{\Sigma}_t &= G_t\,\Sigma_{t-1}\,G_t^T + R_t \\ \bar{\mu}_t &= g(u_t, \mu_{t-1}) \end{aligned}$$

- to the EIF

$$\begin{aligned} \bar{\Omega}_t &= (G_t\,\Omega_{t-1}^{-1}\,G_t^T + R_t)^{-1} \\ \bar{\xi}_t &= \bar{\Omega}_t\,g(u_t, \Omega_{t-1}^{-1}\,\xi_{t-1}) \end{aligned}$$

28

Prediction: From EKF of EIF

1: **Extended_Kalman_filter**$(\mu_{t-1}, \Sigma_{t-1}, u_t, z_t)$**:**

2: $\bar{\mu}_t = g(u_t, \mu_{t-1})$
3: $\bar{\Sigma}_t = G_t\,\Sigma_{t-1}\,G_t^T + R_t$

1: **Extended_information_filter**$(\xi_{t-1}, \Omega_{t-1}, u_t, z_t)$:

2: $\mu_{t-1} = \Omega_{t-1}^{-1}\, \xi_{t-1}$
3: $\bar{\Omega}_t = (G_t\, \Omega_{t-1}^{-1}\, G_t^T + R_t)^{-1}$
4: $\bar{\mu}_t = g(u_t, \mu_{t-1})$
5: $\bar{\xi}_t = \bar{\Omega}_t\, \bar{\mu}_t$

29

Correction Step of the EIF

- As from the KF to IF transition, use substitute the moments in the measurement update

$$bel(x_t) = \eta \exp\left(-\frac{1}{2}\,(z_t - h(\bar{\mu}_t) - H_t\,(x_t - \bar{\mu}_t))^T\, Q_t^{-1} \right.$$
$$\left.(z_t - h(\bar{\mu}_t) - H_t\,(x_t - \bar{\mu}_t)) - \frac{1}{2}(x_t - \bar{\mu}_t)^T \bar{\Sigma}_t^{-1}(x_t - \bar{\mu}_t)\right)$$

- This leads to

$$\Omega_t = \bar{\Omega}_t + H_t^T\, Q_t^{-1}\, H_t$$
$$\xi_t = \bar{\xi}_t + H_t^T\, Q_t^{-1}\,(z_t - h(\bar{\mu}_t) + H_t\, \bar{\mu}_t)$$

30

Extended Information Filter

1: **Extended_information_filter**$(\xi_{t-1}, \Omega_{t-1}, u_t, z_t)$:

2: $\mu_{t-1} = \Omega_{t-1}^{-1}\, \xi_{t-1}$
3: $\bar{\Omega}_t = (G_t\, \Omega_{t-1}^{-1}\, G_t^T + R_t)^{-1}$
4: $\bar{\mu}_t = g(u_t, \mu_{t-1})$
5: $\bar{\xi}_t = \bar{\Omega}_t\, \bar{\mu}_t$
6: $\Omega_t = \bar{\Omega}_t + H_t^T\, Q_t^{-1}\, H_t$
7: $\xi_t = \bar{\xi}_t + H_t^T\, Q_t^{-1}\,(z_t - h(\bar{\mu}_t) + H_t\, \bar{\mu}_t)$
8: return ξ_t, Ω_t

predicted mean. It's required for next steps.

31

EIF vs. EKF

- The EIF is the EKF in information form
- Complexities of the prediction and correction steps differ
- Same expressiveness than the EKF
- Unscented transform can also be used
- Reported to be numerically more stable than the EKF
- In practice, the EKF is more popular than the EIF

32

Summary

- Gaussians can also be represented using the canonical parameterization
- Allow for filtering in information form
- Information filter vs. Kalman filter
- KF: efficient prediction, slow correction
- IF: slow prediction, efficient correction
- The application determines which filter is the better choice!

33

Literature

Extended Information Filter

- Thrun et al.: “Probabilistic Robotics”, Chapter 3.5

34