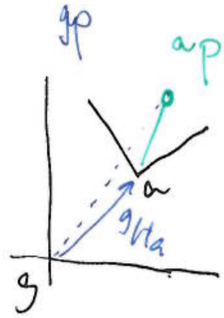
L07: Extended Kalman Filter, Unscented KF and Information Filter

* Summary of L06:

$${}^g p = {}^g H_a \, {}^a p \quad \text{(XYT parametrization)}$$



$${}^g p = {}^g H_b \, {}^b H_a \, {}^a p$$

SSG notation $\quad X_{ga} = X_{gb} \oplus X_{ab}$ (RBT + covariance projection)

${}^b H_a \in SE(2)$ Special Euclidean group.

Sensor models: Lidar (but there are others ...)

Landmarks: $m_i = [m_{i,x}, m_{i,y}]^T$ (locations)

$$f(z) = h(x, m_i) = \begin{bmatrix} r \\ \phi \\ s \end{bmatrix}$$

features from range, bearing, (appearance)

$$\Rightarrow f \sim \mathcal{N}(f(z);\ h(\mu_x, m), \Sigma_s)$$

- Kalman filter: linear systems + Gaussians priors

Joint Gaussians

$$\text{I} \quad \bar{\mu}_t = A_t \mu_t + B u_t$$
$$\text{II} \quad \bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t$$

prediction. (marginalize)

$$\text{III} \quad K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q)^{-1}$$
$$\text{IV} \quad \mu_t = \bar{\mu}_t + K_t (z_t - C_t \bar{\mu}_t)$$
$$\text{V} \quad \Sigma_t = (I - K_t C_t) \bar{\Sigma}_t$$

correction (conditioning)

- Motion model: first order Taylor expansion

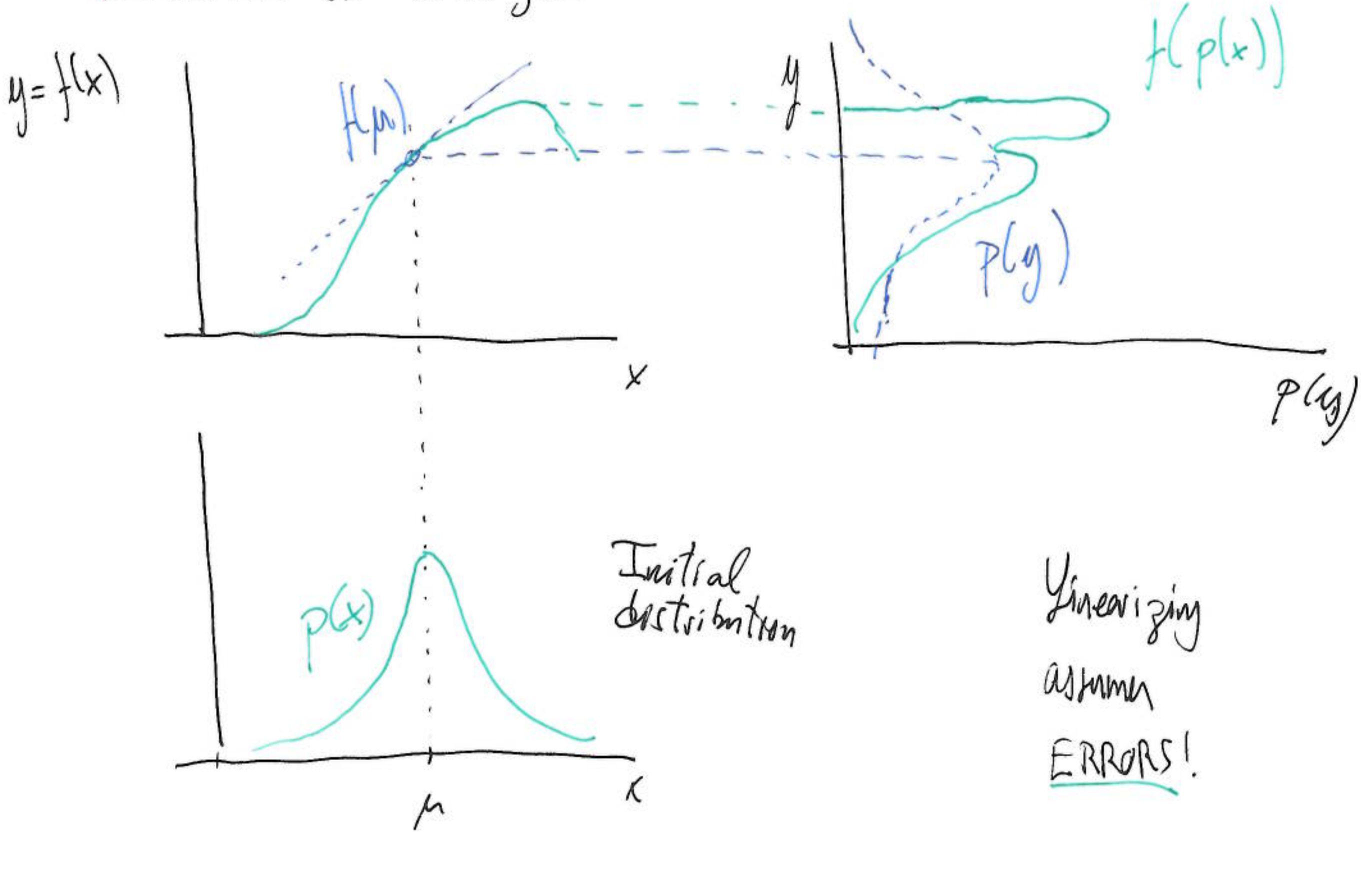
$$x_t = g(x_{t-1}, u_t, \varepsilon_t) \simeq g(\mu_{t-1}, u_t) + \underbrace{\left.\frac{\partial g}{\partial x_{t-1}}\right|_{\mu_{t-1}}}_{G_t} (x_{t-1} - \mu_{t-1}) + \varepsilon_t$$

L05 discussed on how to model $g(\cdot)$ for different systems and how to obtain the probabilistic Model.

- Sensor model: we observe features of landmarks (L06)

$$z_t = h(x_t, \eta_t) \simeq h(\mu_t) + \underbrace{\left.\frac{\partial h}{\partial x_t}\right|_{\mu_t}}_{H_t} (x_t - \mu_t) + \eta_t$$

- Intuition on linearization



Linearizing assumes ERRORS!

* Extended Kalman Filter

Inputs: $\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$

1: $\bar{\mu}_t = g(\mu_{t-1}, u_t)$

2: $\bar{\Sigma}_t = G_t \Sigma_{t-1} G_t^T + R_t$

3: $K_t = \bar{\Sigma}_t H_t^T \left(H_t \bar{\Sigma}_t H_t^T + Q_t \right)^{-1}$

4: $\mu_t = \bar{\mu}_t + K_t (\underbrace{z_t - h(\bar{\mu}_t)}_{\Delta z})$ $\qquad$ Δz (Innovation vector)

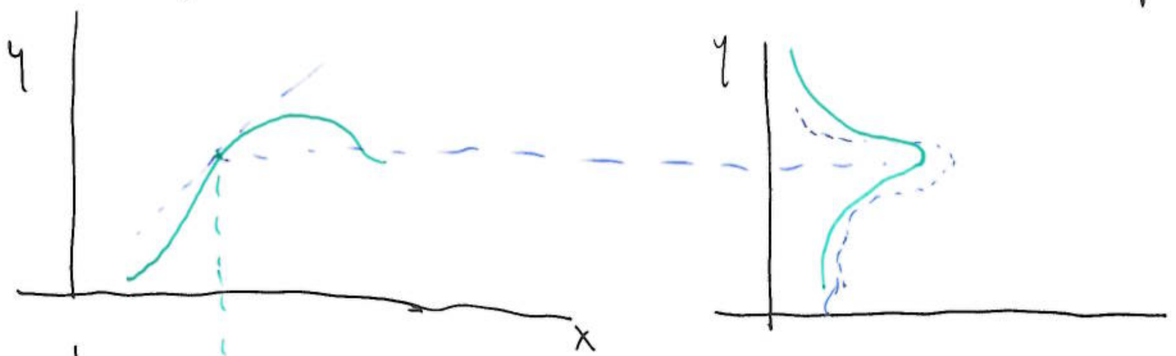
5: $\Sigma_t = (I - K_t H_t) \bar{\Sigma}_t$

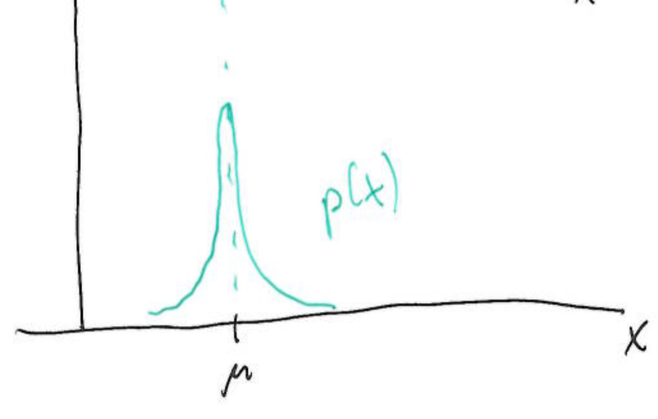
return μ_t, Σ_t $\qquad$ $(\mathcal{N}(\mu_t, \Sigma_t))$

Properties

- EKF is very efficient $O(k^{2.4} + n^2)$ (as KF)
- Not optimal, but in practice works well

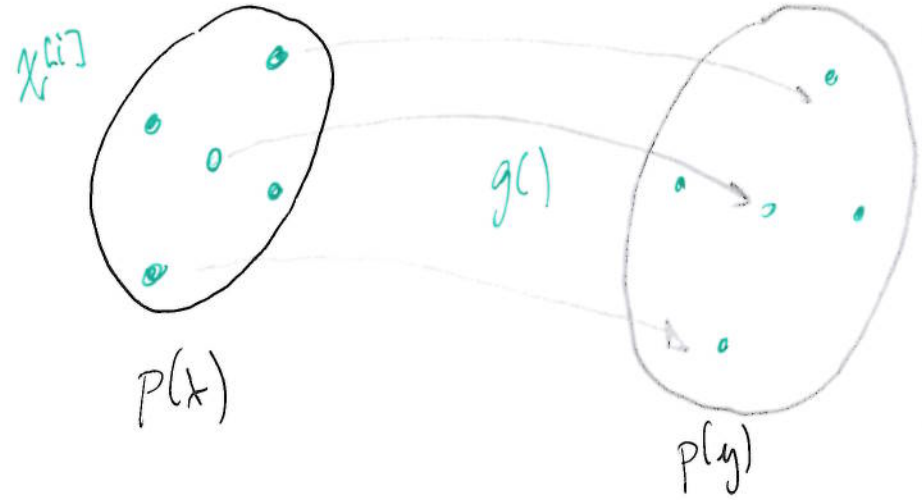
  Depends on the non-linearities (some are more problematic)





Compact initial distribution reduces the error because we are "near" the linearization point ($O(\|\Delta x\|)$!

* Unscented transformation



$$\mathcal{Y}^{[i]} = g(\chi^{[i]}, u_t)$$

We transform a set of sigma points instead of a 1st Taylor exp.

∘ Choosing the sigma points

$$\chi^{[0]} = \mu$$

$$\chi^{[i]} = \mu + \left(\sqrt{n+\lambda}\ \sqrt{\Sigma_x}\right)_i \quad , \quad i = 1, \dots, n$$

$$\chi^{[i]} = \mu - \left(\sqrt{n+\lambda} \cdot \sqrt{\Sigma_x}\right)_i \quad \quad i = n+1, \dots, 2n$$

ith column vector

dim(x)

recall $\Sigma = L \cdot L^T$ (Cholesky) $\Rightarrow \sqrt{\Sigma} = L$

$$\gamma = \sqrt{n+\lambda} = K \quad \text{(radius)}$$

∘ Sigma weights

mean $\quad w_m^{[0]} = \dfrac{\lambda}{n+\lambda}$

covariance $\quad w_c^{[0]} = \dfrac{\lambda}{n+\lambda} + (1 - \alpha^2 + \beta)$

$$w_m^{[i]} = w_c^{[i]} = \frac{1}{2(n+\lambda)} \quad \forall i = 1, \dots, 2n$$

Parameter (for reference)

$\beta = 2$
$\alpha \in [0, 1]$
$\lambda = \alpha^2 (n + K) - n$
$K \geq 0$

→ Unscented Kalman Filter

Inputs $\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$ (ProbRob 70)

I

1: $\mathcal{X}_{t-1} = \{ \mu_{t-1}, \mu_{t-1} + \gamma \sqrt{\Sigma_{t-1}}, \dots \}$ (Build 2n+1 Sigma p.)

2: $\bar{\mathcal{X}}_t^* = g(u_t, \mathcal{X}_{t-1})$

3: $\bar{\mu}_t = \sum_{i=0}^{2n} w_m^{[i]} \bar{\mathcal{X}}_t^{*[i]}$

II

4: $\bar{\Sigma}_t = \sum_{i=0}^{2n} w_c^{[i]} (\bar{\mathcal{X}}_t^{*[i]} - \bar{\mu}_t)(\bar{\mathcal{X}}_t^{*[i]} - \bar{\mu}_t)^T + R_t$ (pi2 pi)

5: $\bar{\mathcal{X}}_t = \{$ create sigma points from $\mathcal{N}(\bar{\mu}_t, \bar{\Sigma}_t) \}$

6: $\bar{Z}_t = h(\bar{\mathcal{X}}_t)$

7: $\hat{z}_t = \sum_{i=0}^{2n} w_m^{[i]} \bar{Z}_t^{[i]}$ ($\hat{z} = H_t \cdot \bar{x}_t$) on EKF

8: $S_t = \sum_{i=0}^{2n} w_c^{[i]} (\bar{Z}_t^{[i]} - \hat{z}_t)(\bar{Z}_t^{[i]} - \hat{z}_t)^T + Q_t$ ($S = H\Sigma H^T + Q$)

9: $\bar{\Sigma}_t^{x,z} = \sum_{i=0}^{2n} w_c^{[i]} (\bar{\mathcal{X}}_t^{[i]} - \bar{\mu}_t)(\bar{Z}_t^{[i]} - \hat{z}_t)^T$ (Crosscovarian $\bar{\Sigma}_t \cdot H_t$)

III 10: $K_t = \bar{\Sigma}_t^{x,z} \cdot S_t^{-1}$

IV 11: $\mu_t = \bar{\mu}_t + K_t (z_t - \hat{z}_t)$

V 12: $\Sigma_t = \bar{\Sigma}_t - K_t S_t K_t^T$

$$\Sigma_t = \bar{\Sigma}_t - K_t S_t (S_t^{-1})^T (\Sigma_t^{x,z})^T$$
$$= \bar{\Sigma}_t - K_t (\bar{\Sigma}_t H_t)^T$$
$$= \bar{\Sigma}_t - K_t H_t \bar{\Sigma}_t$$

EKF

UKF summary:

Highly efficient, same complexity as EKF + extra constant.
Better linearization (Jacobian vs sigma points)
Derivative free
Still not optimal

* Gaussian Canonical parametrization

$$N(\mu, \Sigma) = \eta \cdot \exp\left\{ -\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu) \right\}$$

$$= \eta \cdot \exp\left\{ -\frac{1}{2} x^T \Sigma^{-1} x^T + x^T \Sigma \mu \underbrace{- \frac{1}{2} \mu^T \Sigma^{-1} \mu}_{\text{ctant.}} \right\}$$

$$= \eta' \cdot \exp\left\{ -\frac{1}{2} x^T \Sigma^{-1} x + x^T \Sigma^{-1} \mu \right\}$$

$\Lambda = \Sigma^{-1}$ information matrix (L03)

$\xi = \Sigma^{-1} \cdot \mu$ information vector

$$= \eta' \cdot \exp\left\{ -\frac{1}{2} x^T \Lambda x + x^T \xi \right\} = N^{-1}(\xi, \Lambda)$$

Canonical form: simplied and elegant form.
We can do the same derivations to obtain the KF equivalent (Information filter) for linear Gaussians and the Extended IF for linearized Gaussian systems.

* Information Filter

Inputs: $\xi_{t-1}, \Lambda_{t-1}, u_t, z_t$ (Prob Rob 73)

1: $\bar{\Lambda}_t = \left(A_t \Lambda_{t-1}^{-1} A_t^T + R_t \right)^{-1}$

2: $\bar{\xi}_t = \bar{\Lambda}_t \left(A_t \cdot \Lambda_{t-1}^{-1} \xi_{t-1} + B_t \cdot u_t \right)$

prediction (marginalization)

3: $\Lambda_t = C_t^T Q_t^{-1} C_t + \bar{\Lambda}_t$

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

correction (conditioning) (easier)

return ξ_t, Λ_t

There exists a duality between marginalizing a Gaussian and conditioning a canonical form, both are easy.

Next lecture: localization Prob Rob Ch 7