Machine Learning Working Group:

Inference in Dynamic Models
The Kalman Filter

25th May 2005

Christopher Tay Meng Keat

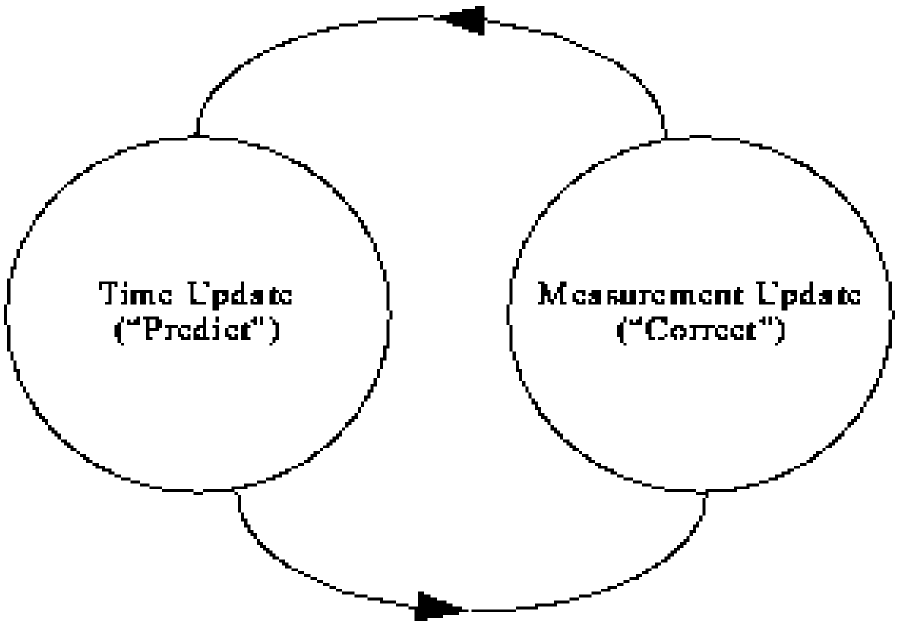
Presentation Outline

- What is a Kalman Filter
- Intuitive Notion (Least Squares example)
- Bayes Filter
- From Bayes filter to KF and EKF
- Alternatives
- Conclusion

What is a Kalman Filter ?

Goal: Infer hidden states of dynamical models

- Inference/Estimation in dynamic models
- States normally hidden
- Observations are noisy



Kalman Filter

Common Applications

Tracking, control, data fusion...

Advantages

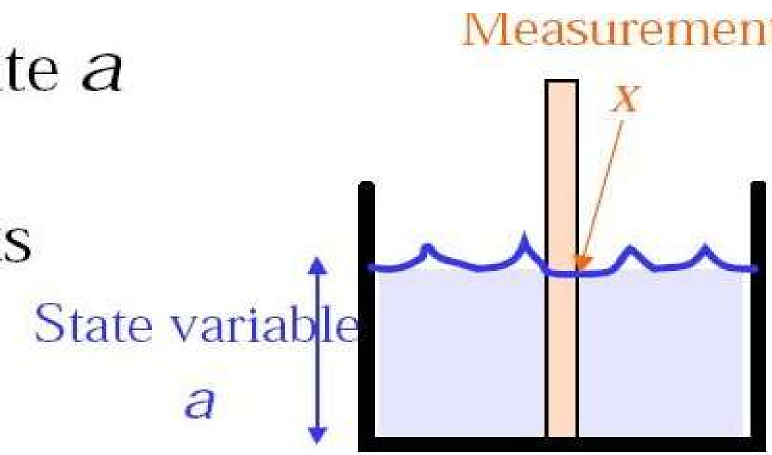
- Tractable –> On Line (Thanks to Gaussian)
- Recursive

Disadvantages

- Gaussian Assumption (non applicable at all times)
- Linear dynamics
- Unimodal distributions

Intuitive Notion (From Least Squares)

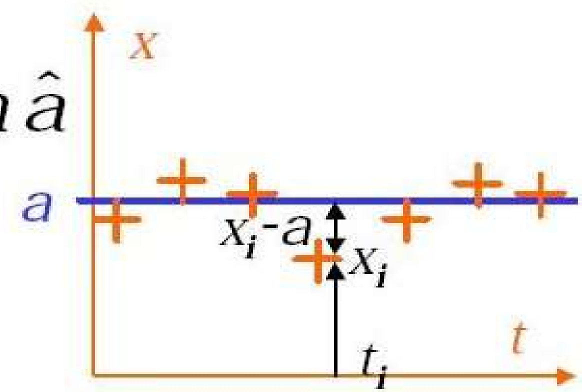
- Goal: Find estimate $\hat{a}$ of state a such that the least square error between measurements and the state is minimum



$$C = \frac{1}{2}\sum_{i=1}^{n}(x_i - a)^2$$

$$\frac{\partial C}{\partial a} = 0 = \sum_{i=1}^{n}(x_i - \hat{a}) = \sum_{i=1}^{n} x_i - n\,\hat{a}$$

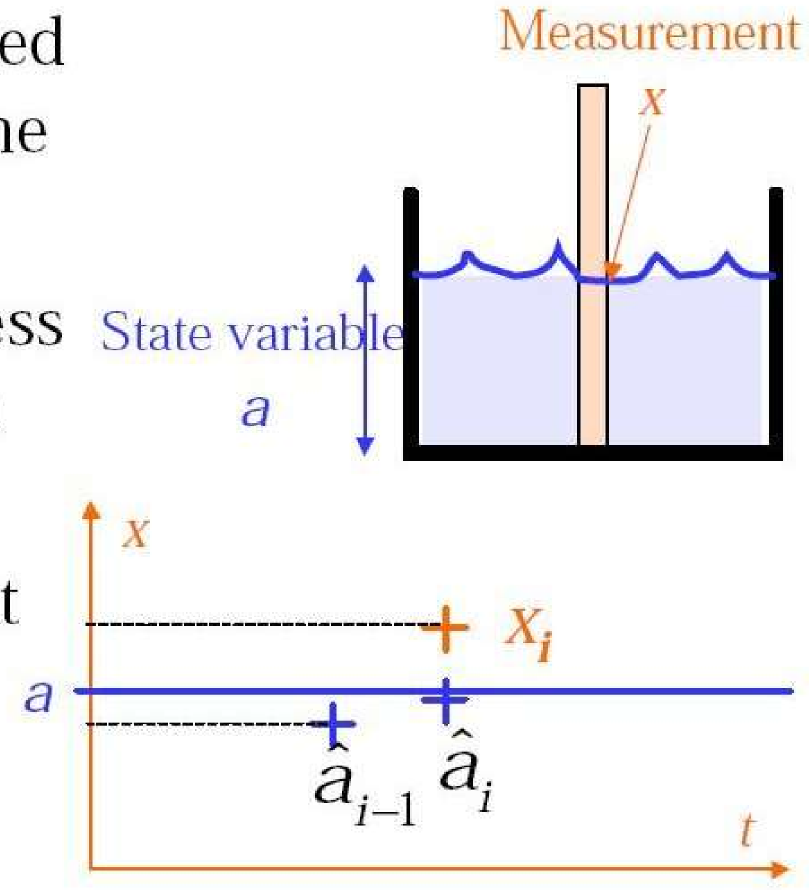
$$\hat{a} = \frac{1}{n}\sum_{i=1}^{n} x_i$$



Example from Cornelia Fermüller

Intuitive Notion (2)

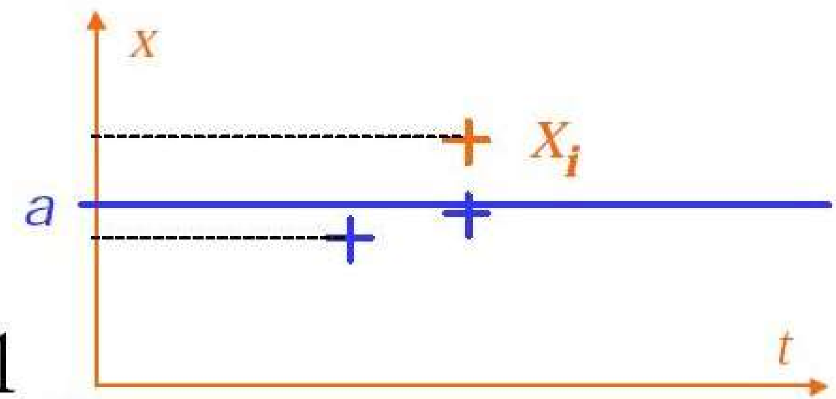
- We don't want to wait until all data have been collected to get an estimate $\hat{a}$ of the depth
- We don't want to reprocess old data when we make a new measurement
- Recursive method: data at step *i* are obtained from data at step *i*-1



Example from Cornelia Fermüller

Intuitive Notion (3)

- Recursive method: data at step *i* are obtained from data at step *i*-1



$$\hat{a}_i = \frac{1}{i}\sum_{k=1}^{i} x_k = \frac{1}{i}\sum_{k=1}^{i-1} x_k + \frac{1}{i} x_i$$

$$\hat{a}_{i-1} = \frac{1}{i-1}\sum_{k=1}^{i-1} x_k$$

$$\hat{a}_i = \frac{i-1}{i}\hat{a}_{i-1} + \frac{1}{i} x_i$$

$$\boxed{\hat{a}_i = \hat{a}_{i-1} + \frac{1}{i}\left(x_i - \hat{a}_{i-1}\right)}$$

Example from Cornelia Fermüller

Intuitive Notion

$$\underbrace{\hat{a}_i}_{\text{Estimate at step } i} = \hat{a}_{i-1} + \overbrace{\frac{1}{i}}^{\text{Gain}} \underbrace{(\overbrace{x_i}^{\text{Actual measure}} - \overbrace{\hat{a}_{i-1}}^{\text{Predicted measure}})}_{\text{Innovation}}$$

Gain specifies how much
do we pay attention
to the difference
between what we expected
and what we actually get

Example from Cornelia Fermüller

Starting From Bayes Filter

$$p(x_t|z_{1:t}) = \frac{p(z_t|x_t\, z_{1:t-1})\, p(x_t|z_{1:t-1})}{p(z_t|z_{1:t-1})}$$

$$= \eta\, p(z_t|x_t\, z_{1:t-1})\, p(x_t|z_{1:t-1})$$

$$= \eta\, p(z_t|x_t) \int p(x_t|x_{t-1}\, z_{1:t-1})\, p(x_{t-1}|z_{1:t-1})\, dx_{t-1}$$

$$= \eta\, p(z_t|x_t) \int p(x_t|x_{t-1})\, p(x_{t-1}|z_{1:t-1})\, dx_{t-1}$$

Correction
(Observation model)
$N(z_t; C_t x_t, Q_t)$

Prediction
(motion model)
$N(x_t; A x_{t-1}, Q_t)$

Previous State
$N(x_{t-1}; \mu_{t-1}, \Sigma_{t-1})$

From Bayes To Kalman

KF Prediction:

$$\bar{\mu}_t = A_t \mu_{t-1}$$

Motion Model

$$\bar{\Sigma}_t = \left(R_t + A_t \Sigma_{t-1} A_t^T \right)^{-1}$$

KF Correction:

Kalman Gain

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

$$\mu_t = \bar{\mu}_t + K_t \left(z_t - C_t \bar{\mu}_t \right)$$

Innovation

$$\Sigma_t = \left(I - K_t C_t \right) \bar{\Sigma}_t$$

When things go non linear - Extended Kalman Filter

Motion model: $x_t = g(x_{t-1}) + \epsilon_t$

Observation model: $z_t = h(x_t) + \delta_t$

<u>Taylor's Expansion:</u>

$$g(x_{t-1}) \approx g(\mu_{t-1}) + \frac{\partial g(x_{t-1})}{\partial x_{t-1}}(x_{t-1} - \mu_{t-1})$$

$$h(x_t) \approx h(\bar{u}_t) + \frac{\partial h(x_t)}{\partial x_t}(x_t - \bar{\mu}_t)$$

Extended Kalman Filter

$$G_t = \frac{\partial g(x_{t-1})}{\partial x_{t-1}} \qquad H_t = \frac{\partial h(x_t)}{\partial x_t}$$

KF Prediction:

$$\bar{\mu}_t = A_t \mu_{t-1}$$

$$\bar{\Sigma}_t = \left(R_t + A_t \Sigma_{t-1} A_t^T \right)^{-1}$$

KF Correction:

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

$$\mu_t = \bar{\mu}_t + K_t \left(z_t - C_t \bar{\mu}_t \right)$$

$$\Sigma_t = \left(I - K_t C_t \right) \bar{\Sigma}_t$$

EKF Prediction:

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

$$\bar{\Sigma}_t = \left(R_t + G_t \Sigma_{t-1} G_t^T \right)^{-1}$$

EKF Correction:

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

$$\mu_t = \bar{\mu}_t + K_t \left(z_t - h(\bar{\mu}_t) \right)$$

$$\Sigma_t = \left(I - K_t H_t \right) \bar{\Sigma}_t$$

Alternatives (1)

Unscented Kalman Filter (Julier et. al 97)

- Easier to approx. Gaussian than linearizing
- deterministic sampling -> Propagate points non linearly
- Unscented Transform
- Relatively untested under experiments

Advantages

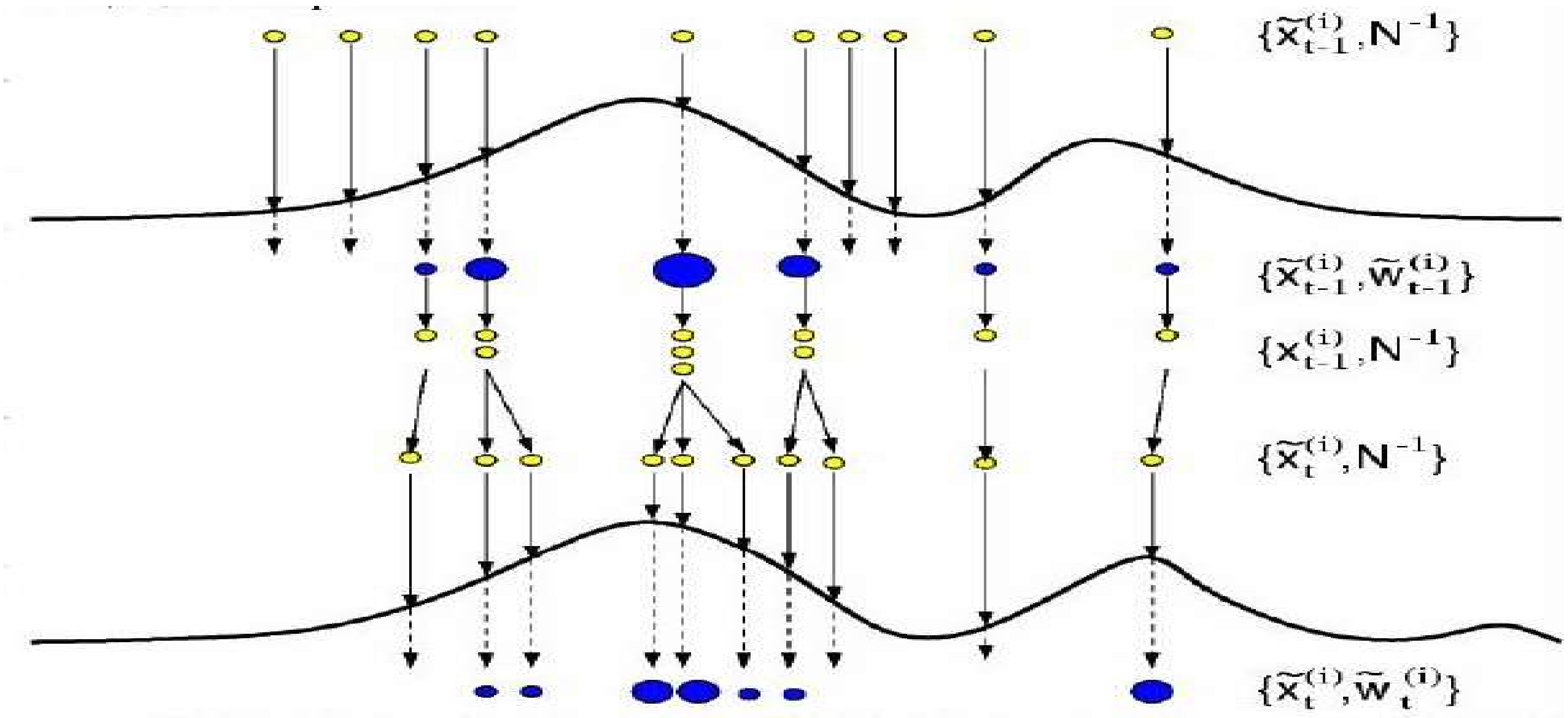
- Handle non linearities better
- Easier to implement

Disadvantages

- Gaussian assumption still holds
- Many parameters to tune (5)

Alternatives (2)

Particle Filters



- Nonlinear
- multimodal

- Degeneracy problems
- Number of particles?

Taken from Van der Merwe et al.

Conclusion

- Kalman filter similar to least squares
- Bayesian filter -> Kalman filter
- Advantages and Disadvantages
- Unscented KF
- Particle Filters
- Humans are not perfect, so are our mathematics