Kalman Filters

Note Title 3/29/2006

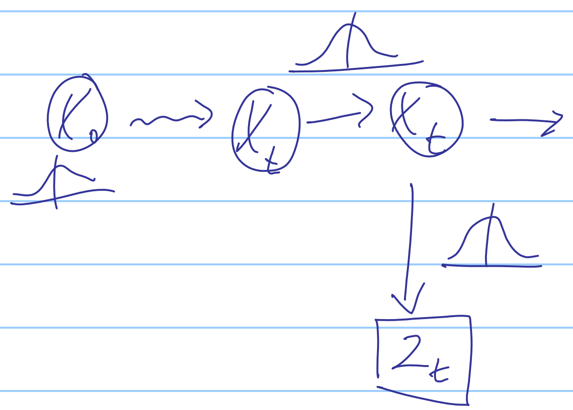
- Prelim Review: Monday 4/3 8pm Upson 109
- Final Exam: 5/15/06, 7-9:30 pm, HO 401

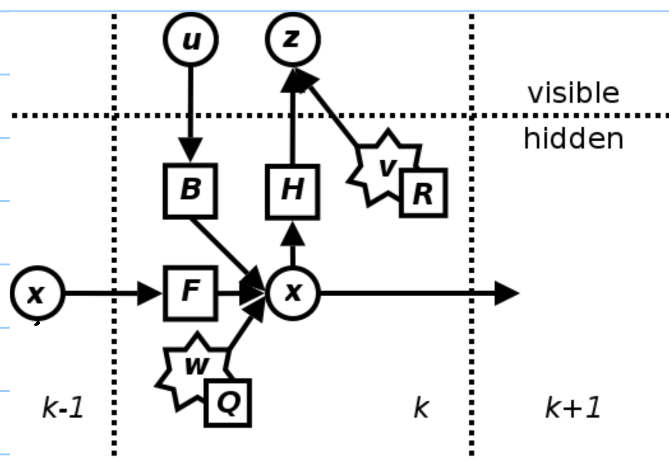
OUTLINE:
- intro
- motivating example & derivation
- full discrete KF algorithm
- Matlab demo

INTRODUCTION

- popular model for Stochastic Estimation:
  - estimate state of a system from noisy observations
- System:
  - i) initial state distribution
  - ii) transition model
  - iii) sensor model

  all based on Normal distribution



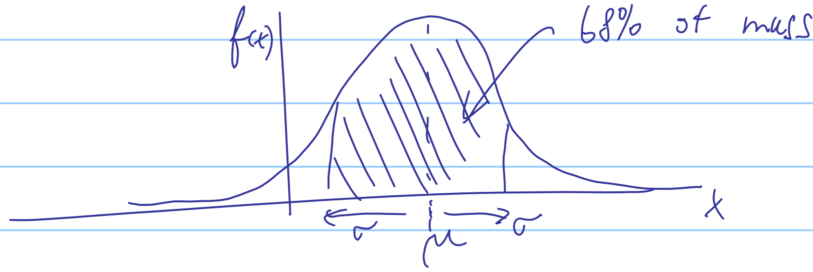


Normal Distribution (Gaussian)

- continuous distribution over $-\infty, +\infty$
- parameters: Mean (μ) $\in (-\infty, \infty)$; Variance (σ^2) $\in (0, +\infty)$ — $N(\mu, \sigma^2)$

- distribution function (pdf):

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$



- Aditivity of independent variables:

$$N(\mu_1, \sigma_1^2) + N(\mu_2, \sigma_2^2) = N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

- Central Limit Theorem:

$\{X_i\}$ iid random variables, with $E(X) = \mu$, $Var(X_i) = \sigma^2$

(ANY distribution with $\mu < \infty$, $\sigma^2 < \infty$)

$S_n = \sum_1^n X_i$, then $\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow[n\to\infty]{} \sim N(0,1)$

works even for small $n \sim 6$

- Multivariate Normal Distribution: $N(\vec{\mu}, \Sigma)$

mean covariance matrix

- higher-dimensional generalization of Normal
- random vector $\vec{X} = (X^{(1)}, \dots, X^{(k)})$

$\Rightarrow \vec{\mu} = (E(X^{(1)}), \dots, E(X^{(k)}))$

$\Sigma = cov(\vec{X}, \vec{X}) = \left(E(X^{(i)} - \vec{\mu}^{(i)})(X^{(j)} - \vec{\mu}^{(j)})\right)_{ij}$

Kalman Filter

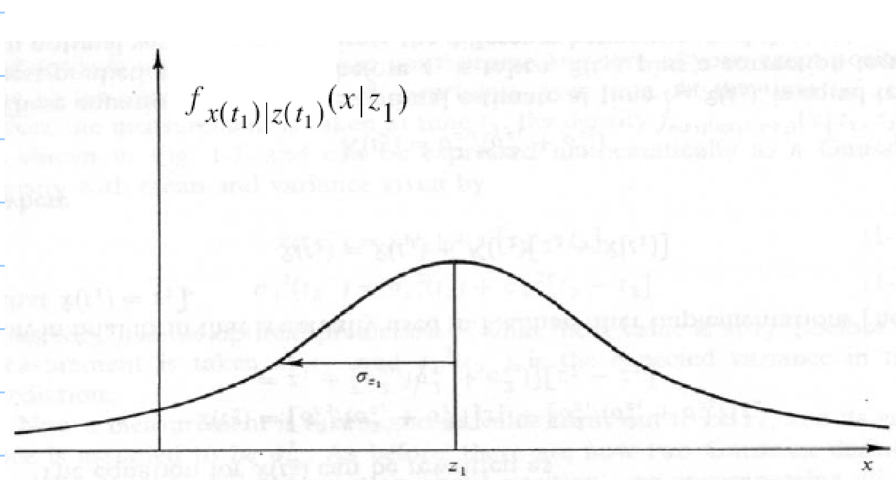
i) initial distribution: $P(x_0) \sim N(\mu_0, \Sigma_0)$
ii) transition model: $P(x_{t+1} | x_t) \sim N(\cdot, \Sigma_x)$
iii) sensor model: $P(z_t | x_t) \sim N(\cdot, \Sigma_z)$

- posterior probability $P(x_t)$ stays $N(\mu_t, \Sigma_t)$ for all t
- continuous state & evidence, discrete time (discrete KF)

EXAMPLE & MODEL DERIVATION

- lost at an unknown location $x(t)$ on a boat
- 2 ways to estimate location: (assume normal error: $Z = z + N(0, \sigma^2)$)
  - you (amateur) z_1 σ_1^2
  - friend (skilled) z_2 σ_2^2 $< \sigma_1^2$

1) you estimate at time t_1: $z_1 = z_1(t_1)$

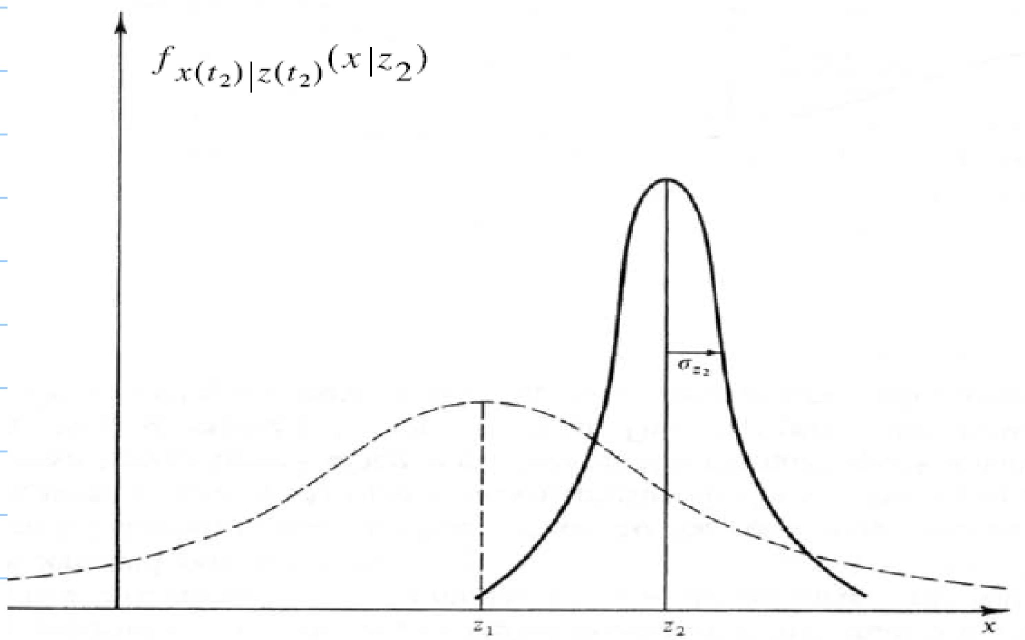


⇒ best estimate of the position:

mode
median
mean

$\hat{x}(t_1) = z_1$
$\hat{\sigma}^2(t_1) = \sigma_1^2$

2) friend estimates at the same time: $z_2 = z_2(t_1)$



$\Rightarrow$ now what is the best estimate of $\hat{x}(t_1)$?
how is the new information incorporated?

Model derivation (static)

- linearly combine the observations:

$$\hat{x} = w z_1 + (1-w) z_2$$

(w is unknown weight to be calculated)

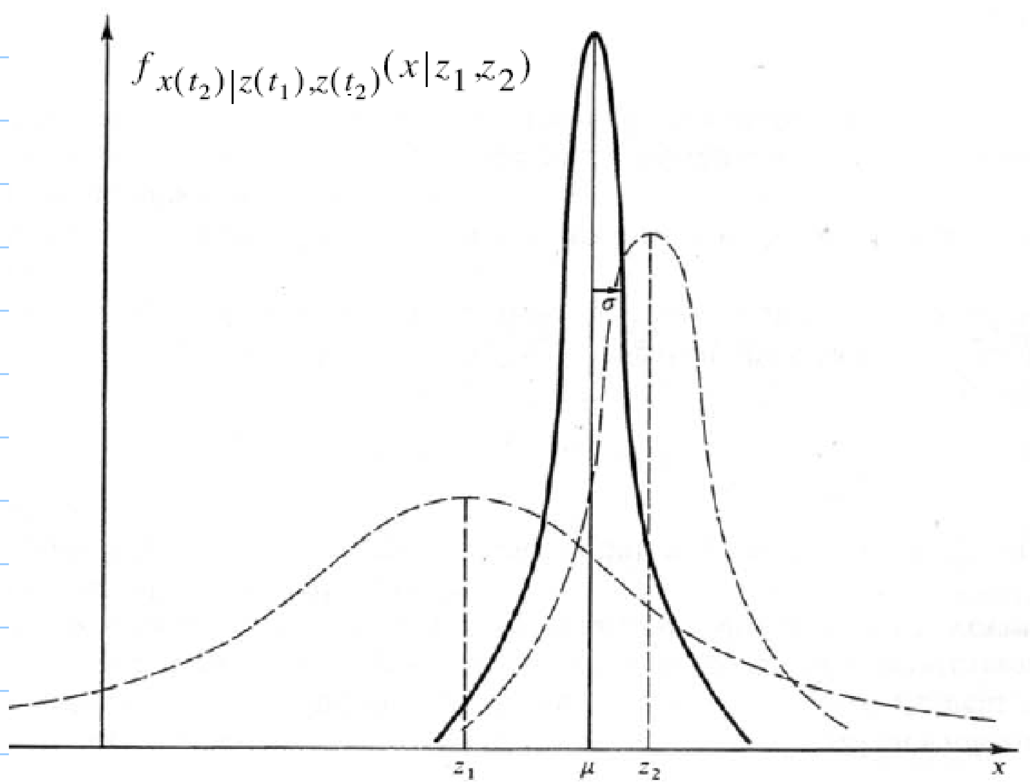
$$\Rightarrow \quad \hat{\sigma}^2 = w^2 \sigma_1^2 + (1-w)^2 \sigma_2^2$$

- find w that minimizes the uncertainty: $\frac{\partial \hat{\sigma}^2}{\partial w} = 0$

$$2w\sigma_1^2 - 2\sigma_2^2 + 2w\sigma_2^2 = 0$$

$$\Rightarrow \quad w = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

So $\hat{\sigma}^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$



OBSERVATIONS:

1) $\hat{x}$ vs. (σ_1^2, σ_2^2) nicely follows intuition

$$\sigma_2^2 = \sigma_1^2 \rightarrow \hat{x} = 1/2\,(z_1 + z_2)$$
$$\sigma_2^2 \gg \sigma_1^2 \rightarrow \hat{x} \sim z_1$$
$$\sigma_1^2 \gg \sigma_2^2 \rightarrow \hat{x} \sim z_2$$

2) $\hat{\sigma}^2$ is smaller than both σ_1^2 and σ_2^2
(follows from $\frac{1}{\hat{\sigma}^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$)
$\Rightarrow$ any information is used (even if very noisy)

3) such w also makes $\hat{x}$ minimize squared weighted distances from z_1 and z_2 to any x:

$$\hat{x} = \arg\min_x \sum_1^2 \left(\frac{z_i - x}{\sigma_i}\right)^2$$

$$\frac{\partial}{\partial x} = \sum_1^2 \frac{-2}{\sigma_i^2}(z_i - x) = 0$$

$$\sigma_2^2(z_1 - x) + \sigma_1^2(z_2 - x) = 0$$

$$\sigma_2^2 z_1 + \sigma_1^2 z_2 = x(\sigma_2^2 + \sigma_1^2)$$

$$w\, z_1 + (1-w)\, z_2 = x$$

RECURSIVE FORMULATION

$$\hat{x} = w\,\hat{x}_{prev} + (1-w)\,z_2 = \hat{x}_{prev} + \underbrace{(1-w)}_{\text{Update gain } K = \left(\frac{\hat{\sigma}^2_{prev}}{\hat{\sigma}^2_{prev} + \sigma_2^2}\right)}\underbrace{(z_2 - \hat{x}_{prev})}_{\text{innovation}}$$

$$\Rightarrow \hat{x} = \hat{x}_{prev} + K(z - \hat{x}_{prev}) \qquad \hat{\sigma}^2 = (1-K)\,\hat{\sigma}^2_{prev}$$

Model derivation (dynamic)

- similar situation as before, but the boat is moving with speed $v \sim N(\mu_v, \sigma_v^2)$ $\mu = \mu_u + v$
- another measurement is done at time $t_2 > t_1$

$$z_3 = z_3(t_2) \quad \text{with} \quad \sigma_3^2$$

- What is $\hat{x}(t_2)$?

Let t_2^- $(\approx t_2)$ be time just before z_3 is taken

PREDICTION:

$$\hat{x}(t_2^-) = \hat{x}(t_1) + \mu_v(t_2^- - t_1)$$
$$\hat{\sigma}^2(t_2^-) = \hat{\sigma}^2(t_1) + \sigma_v^2(t_2^- - t_1)^2$$

$\rightsquigarrow$ observation z_3: again, we need to combine 2 Gaussians (z_3, σ_3^2) and $(\hat{x}(t_2^-), \hat{\sigma}^2(t_2^-))$

CORRECTION:

$$\hat{x}(t_2) = \hat{x}(t_2^-) + K(z_3 - \hat{x}(t_2^-))$$
$$\hat{\sigma}^2(t_2) = (1-K)\,\hat{\sigma}^2(t_2^-)$$

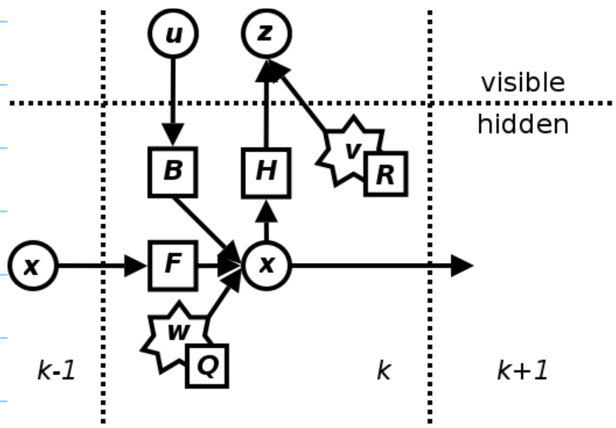
where
$$K = \frac{\hat{\sigma}^2(t_2^-)}{\hat{\sigma}^2(t_2^-) + \sigma_3^2}$$

OBSERVATIONS:

1) K and $\hat{\sigma}^2(t_2)$ does not depend on z_3, can be precomputed before observations are taken

2) the correction step again makes an optimal decision between how much to trust the new observation vs. the prediction

DISCRETE KALMAN FILTER



x ... system state
u ... (optional) control input
z ... observation (measurement)
F ... state transition matrix
B ... control input matrix
w ... transition noise $w \sim N(0, Q)$
H ... observation relation
v ... observation noise $v \sim N(0, R)$

MODELS:

transition model:

$$x(t) = F\,x(t-1) + B\,u(t) + w$$

sensor model:

$$z(t) = H\,x(t) + v$$

ASSUMPTIONS:

i) linear models (both transition & sensor)
ii) uncertainty Gaussian (Normally distributed)
iii) White (uncorrelated in time)

Algorithm:

$\hat{x}(t)$ .... estimate of $x(t)$
$P(t)$ .... covariance matrix of $\hat{x}(t)$ (uncertainty)

INPUT : $\hat{x}(t-1)$, $P(t-1)$, $u(t-1)$
OUTPUT: $\hat{x}(t)$, $P(t)$

prediction:
$$\hat{x}(t^-) = F\hat{x}(t-1) + Bu(t)$$
$$P(t^-) = FP(t-1)F^T + Q$$

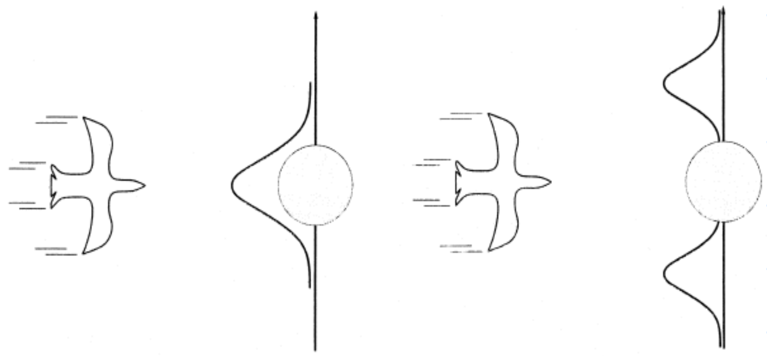
correction:
$$\hat{x}(t) = \hat{x}(t^-) + K\left(z(t) - H\hat{x}(t^-)\right)$$
$$P(t) = (I - KH)P(t^-)$$

$$\left(K = P(t^-)H^TS^{-1}\right.$$
$$\left.S = R + HP(t^-)H^T\right)$$

INITIAL ESTIMATE: $\hat{x}(0)$, $P(0)$

Extensions:

- Kalman Smoothing
- Extended KF (nonlinear transition & sensor models)
  - locally linearized using Hessian
- Switching KF


(a) (b)

REFERENCES

An Introduction to the Kalman Filter, SIGGRAPH 2001 Course, Greg Welch and Gary Bishop
Kalman filtering chapter from Stochastic Models, Estimation, by Peter Maybeck
http://en.wikipedia.org/wiki/Kalman_filter