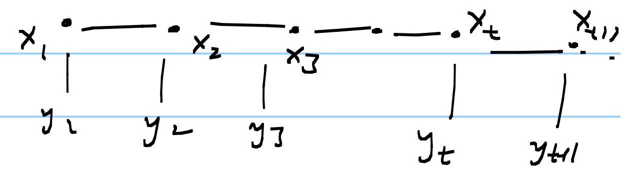
Particle Filtering (Continued)

Note Title — 4/27/2006

Recall. $Y_t = \{y_1, \dots, y_t\}$, state x_t



prediction

$$p(x_{t+1} \mid Y_t) = \sum_{x_t} p(x_{t+1} \mid x_t)\, p(x_t \mid Y_t)$$

correction

$$p(x_{t+1} \mid Y_{t+1}) = \frac{p(y_{t+1} \mid x_{t+1})\, p(x_{t+1} \mid Y_t)}{p(y_{t+1} \mid Y_t)}$$

$p(x_{t+1} \mid x_t)$

$p(y_t \mid x_t)$

Recall example from last lecture.

Prior model: $p(x_{t+1} \mid x_t) = N(x_t + \mu, \sigma_p^2)$ Gaussian

Likelihood function $p(\underline{y}_t \mid x_t)$

The observations occur in a window size Δ.

At time t, the number of observations in the window is a random variable m_t.

The observations are represented by a vector $\underline{y}_t = (y_{t,1}, \dots, y_{t,m_t})$. Each $y_{t,i}$ is an observation.

Model Assumptions.

The probability that target is visible is p_d.

The distribution of the number of distractors is Poisson, with parameter $\lambda\Delta$.

The position of a distractor is uniformly distributed in the window (i.e. density is $1/\Delta$)

(2)

Define an Indicator Variable $I_t = \begin{cases} 0, & \text{if target is invisible} \\ k, & \text{if } k^{th} \text{ object is the target} \end{cases}$

$$p(y_t \mid x_t, I_t = 0) = \Delta^{-m_t} \frac{(\lambda\Delta)^{m_t}}{m_t!} e^{-\lambda\Delta} = \frac{\lambda^{m_t}}{m_t!} e^{-\lambda\Delta}$$

$$p(y_t \mid x_k, I_t = k) = \frac{\lambda^{m_t-1}}{(m_t-1)!} e^{-\lambda\Delta} \frac{1}{\sqrt{2\pi}\, r} \exp\left\{-\frac{(y_{t,k} - x_t)^2}{2r^2}\right\} \quad (*)$$

By assumption

$$P(I_t = 0) = 1 - P_d, \quad P(I_t = k) = P_d / m_t.$$

So

$$p(y_t, I_t \mid x_t) \propto \begin{cases} (1-P_d)\lambda & \text{if } I_t = 0 \\ P_d \frac{1}{(2\pi)^{1/2} r} \exp\left\{-\frac{(y_{t,k}-x_t)^2}{2r^2}\right\} & \text{if } I_t = k \end{cases}$$

(we drop the common factor $\frac{\lambda^{m_t}}{(m_t-1)!} e^{-\lambda\Delta}$)

Now I_t is not known, so we have to sum it out to obtain

$$p(y_t \mid x_t) = \sum_{I_t} p(y_t, I_t \mid x_t)$$

$$\propto (1-P_d)\lambda + \sum_{k=1}^{m_t} P_d \frac{1}{(2\pi)^{1/2} r} \exp\left\{-\frac{(y_{t,k}-x_t)^2}{2r^2}\right\}$$

(3)

Mixture Kalman Filter

Recall Rao-Blackwellization:

- don't sample if you do things analytically
- do as much as possible analytically.

Let $\Lambda_t = (I_1, \ldots, I_t)$ be a trajectory to time (i.e. a selection of observation for each time step).

Key Ideas: If we know I_t, then we can solve by standard Kalman filter – i.e. update means and variances.

We don't know I_t, so let us have several sampled I_t's with weights.

The correct trajectory will have fit the data well (because there will usually be observations where it predicts them to be), so it will have a large weight.

Details – recursive description. Suppose we have m trajectories $\{\Lambda_{t-1}^{(1)}, \ldots, \Lambda_{t-1}^{(m)}\}$ with weights $\omega_{t-1}^{(1)}, \ldots, \omega_{t-1}^{(m)}$.

For each trajectory, use standard Kalman to estimate the mean $\mu_{t-1}^{(j)}$ and covariance $\Sigma_{t-1}^{(j)}$.

$$\text{Denote} \quad KF_{t-1}^{(j)} = (\mu_{t-1}^{(j)}, \Sigma_{t-1}^{(j)})$$

(Note: The trajectory is a hidden variable so we are representing it by a distribution of samples.)

(4) Select a trial distribution for assignment

$g(I_t \mid \Lambda_{t-1}^{(j)}, KF_{t-1}^{(j)}, y_t)$ — e.g. $(1-P_d)$ — for $I_t = 0$

normalization

$P_d \frac{e^{-\| y_{t,k} - \mu_{t-1}^{(j)} - \mu \|^2}}{Z}$, $I_t = k$

Do the 3 steps

(1) generate $I_t^{(j)}$ from $g(I_t \mid \Lambda_{t-1}^{(j)}, KF_{t-1}^{(j)}, \underline{y}_t)$
(e.g. predict assignment)

(2) conditional on each $\{KF_{t-1}^{(j)}, y_t, I_t^{(j)}\}$
obtain $KF_t^{(j)}$ by one step of the Kalman filter.

(3.) update the new weight as "times"

$$\omega_t^{(j)} = \omega_{t-1}^{(j)} \times u_t^{(j)}$$

where

$$u_t^{(j)} = \frac{p(\Lambda_{t-1}^{(j)} \mid \underline{y}_t)}{p(\Lambda_{t-1}^{(j)} \mid \underline{y}_{t-1})\, g(I_t \mid \Lambda_{t-1}^{(j)}, KF_{t-1}^{(j)}, \underline{y}_t)}$$

(4) If the coeff of variation of the w_t exceeds a threshold, we resample a new set of KF_t from $\{KF_t^{(1)}, \ldots, KF_t^{(m)}\}$ with probability proportional to the weights $\omega_t^{(j)}$.

See p 101 Liu, for examples of the errors of these methods.

(5) Need to compute $p(\Lambda^{(j)}_{t-1}, I^{(j)}_{t} \mid \underline{y}_t)$

and $p(\Lambda^{(j)}_{t-1} \mid \underline{y}_{t-1})$

standard Kalman.

We know $p(\underline{y}_t \mid x_t, I_t)$

we know $p(x_t \mid \Lambda^{(j)}_{t-1}) = \int dx_{t-1}\, p(x_t \mid x_{t-1})\, p(x_{t-1} \mid \Lambda^{(j)}_{t-1})$

we know $p(I_t)$

Gaussian

$$p(\underline{y}_t \mid \Lambda^{(j)}_{t-1}, I_t) = \int dx_t\; p(\underline{y}_t \mid x_t, I_t)\, p(x_t \mid \Lambda^{(j)}_{t-1}).$$

$$p(\Lambda^{(j)}_{t-1}, I_t \mid \underline{y}_t) = \frac{p(\underline{y}_t \mid \Lambda^{(j)}_{t-1}, I_t)\, p(I_t)\, p(\Lambda^{j}_{t-1})}{p(\underline{y}_t)}$$

Similar manipulations to get.

$p(\Lambda^{(j)}_{t-1} \mid \underline{y}_{t-1})$.