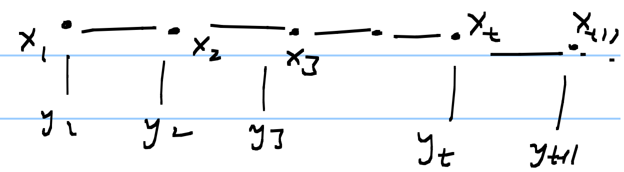
Lecture 14   Weighted Trajectories: Bayes-Kalman example.

Note Title   4/27/2006

Recall $Y_t = \{y_1, \ldots, y_t\}$,   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

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

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

$p(y_t \mid x_t)$

We will revisit the example in the previous lecture and derive an alternative sampling method which exploits the partial Gaussian structure of the problem.

Recall that there is a target with prior probability $p(x_{t+1} \mid x_t) = N(x_t + \mu, \sigma_p^2)$ ~ Gaussian model.

At time t there are m_t observations. If the target is visible then one observation corresponds to the target (but we do not know which). The other observations are distractors.

An indicator variable I_t specifies which observation corresponds to the target, and whether the target is visible.

$I_t = 0$,   if target is invisible at time t

$I_t = k$,   if target is the k^{th} observation at time t.

The observation model for the target is Gaussian if I_t is known.

$$P(y_{t,k} \mid x_t, I_k = k) = \frac{1}{\sqrt{2\pi}\, r} e^{-(y_{t,k} - x_t)^2 / 2r^2}$$

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

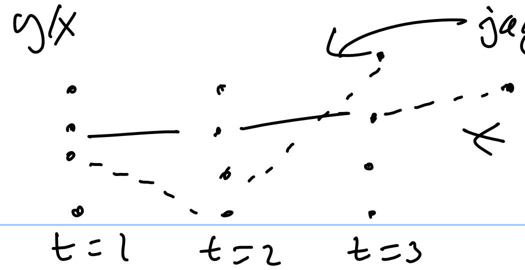
Details - recursive description. Suppose we have m trajectories $\{\Lambda^{(1)}_{t-1}, \ldots, \Lambda^{(m)}_{t-1}\}$ with weights $\omega^{(1)}_{t-1}, \ldots, \omega^{(m)}_{t-1}$.

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

Denote $KF^{(j)}_{t-1} = (\mu^{(j)}_{t-1}, \Sigma^{(j)}_{t-1})$

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

Intuition



Note: exponential number of trajectories
— possible assignments (I_t) of observations to trajectories.
Sample these as follows:

Select a trial distribution $g(I_t \mid \Lambda^{(j)}_{t-1}, KF^{(j)}_{t-1}, y_t)$ for assignment
— e.g. $(1-p_d)$ for $I_t = 0$; $\frac{p_d}{Z} e^{-(y_{t,k} - \mu^{(j)}_{t-1,k})^2}$ for $I_t = k$.
(Z normalization constant.)

(1) Generate $I_t^{(j)}$ from $g(I_t \mid \Lambda^{(j)}_{t-1}, KF^{(j)}_{t-1}, \underline{y}_t)$
— i.e. predict new assignment.

(2) Conditional on each $\{KF^{(j)}_{t-1}, y_t, I_t^{(j)}\}$ obtain $KF^{(j)}_t$ by one step of the Kalman filter (update mean & variance).

(3) Update the new weights by

$$w_t^{(j)} = w^{(j)}_{t-1} \, u_t^{(j)}$$

(Note: error in lecture here)

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

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

See p 107 Liu, for errors that can happen.

Page 4.

Need to compute $p(\Lambda_{t-1}^{(j)}, I_t^{(j)} \mid \underline{y}_t)$ and $p(\Lambda_{t-1}^{(j)} \mid \underline{y}_{t-1})$

standard Kalman.

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

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

we know $p(I_t)$

Gaussian

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

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

Similar manipulations to get.

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