Kalman Filter

Linear Dynamical Model with Gaussian Noise

$$x_t = A x_{t-1} + \vec{n}_d \quad ; \quad \vec{n}_d \sim N(0, C_d)$$

Linear Observation Model with Gaussian observation noise

$$\cancel{x_t = A x_{t-1}}$$

$$z_t = M x_t + \vec{n}_m \quad , \quad \vec{n}_m \sim N(0, C_m)$$

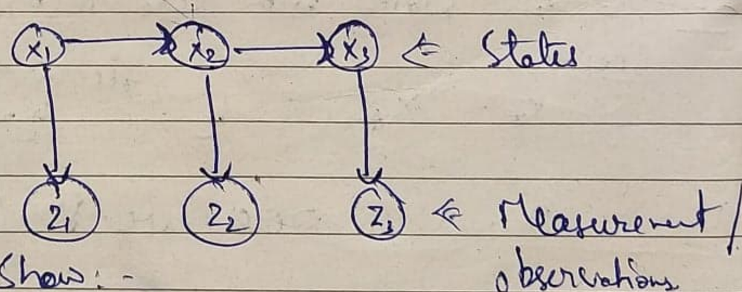
$\therefore$ $P(x_t | x_{t-1}) = G(x_t ; A x_{t-1}, C_d)$

Transition density

& $p(z_t | x_t) = G(z_t ; M x_t, C_m)$

Observation Density

Graph:



What does it show:-

→ Given the state, in which the measurement is taken, measurement is independent of other measurements.
i.e. given x_i, z_i is independent of z_j or x_j.

→ Also ~~x_t only dependent on x_{t-1}~~

$$P(x_t | x_{1:t}) = P(x_t | x_{t-1})$$

Date
Page

Step 3:- find new filtering distribution.

given samples x_t^j, compute weights $w_t^j = c\,p(z_t | x_t^j)$

→ $p(z_t | x_t^j)$ is data likelihood we know how to evaluate.

→ &

$$1/c = p(z_t | z_{1:t-1}) = \int p(z_t | x_t)\, p(x_t | z_{1:t-1})\, dx_t$$

$$\approx \sum_j p(z_t | x_t^j)$$

$$p(x_{t-1} | z_{1:t-1}) \xrightarrow{\text{Sample}} p(x_t | x_{t-1}) \xrightarrow{\text{sample}} p(z_t | x_t) \xrightarrow{\text{Normalise}} p(x_t | z_{1:t})$$

filter Distribution is approximated we weighted sample set.

temporal Dynamics

liklihood evaluati

filtering distribution

Recursion is important:-

1. express filtering distribution at time t in terms of the filter distribution at $t-1$ & evidence at time t.

2. All ~~infor~~ information from the past is summarised in prediction distribution.

↳ Without recursion one may have to store all previous image to compute the filtering distribution at time t.

x—x—x

Temporal Dynamics

→ Assume combination of Deterministic & Stochastic dynamics.

1. Random walk with zero vel0 & IID Gaussian noise.

$$x_t = x_{t-1} + \eta \xrightarrow{} \text{stochastic} \qquad \eta \sim N(0, c).$$

1. " with zero acc & Gaussian process noise.

$$\begin{pmatrix} x_t \\ \vec{v}_t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ v_{t-1} \end{pmatrix} + \begin{pmatrix} \vec{n}_t \\ \vec{\epsilon}_t \end{pmatrix}$$

Components of Tracking:-

1. Prediction:- Seeing $z_1, \ldots, z_{t-1}$, what can be predicted by x_t given these measurements. We obtain $p(x_t|z_1, \ldots z_{t-1})$

2. Filtering / Correction:- Now you have z_t as well, what can you say about x_t, we model $p(x_t|z_1, \ldots z_t)$.

$\Rightarrow$ So Now any further will depend upon then two terms which are there in prediction & update step. those two are:

$P(x_i|x_{i-1}) \Rightarrow$ Depends upon the dynamics models of the system.

Δ, Transition liklihood.

$P(y_i|x_i) \Rightarrow$ Depends upon the measurement model of the system.

Δ, ~~liklihood~~ observation liklihood.

For Ex:

In Kalman Filter, we have

$$x_i = Ax_{i-1} + n_d \quad \text{where } n_d \sim \mathcal{N}(0, \sigma_d^2)$$

Hence

$$P(x_i|x_{i-1}) = \mathcal{N}(x_i; Ax_{i-1}, \sigma_d^2)$$

and

$$y_i = Mx_i + n_m \qquad n_m \sim \mathcal{N}(0, \sigma_m^2)$$

$$P(y_i|x_i) = \mathcal{N}(y_i; Mx_i, \sigma_m^2)$$

(3) Update / Correction Step :-

In this we need to find : $P(x_i | y_0, \dots y_i)$

Given $P(x_i | y_0, \dots y_{i-1})$ ⇒ (we found this in prediction step.)

i.e when we get or measure y_i. we update our belief about x_i.

$$P(x_i | y_0, \dots y_i) = \frac{P(x_i, y_0 \dots y_i)}{P(y_0, \dots y_i)}$$

$$= \frac{P(y_i | x_i, y_0, \dots y_{i-1})\, P(x_i, y_0 \dots y_{i-1})}{P(y_0, \dots y_i)}$$

$$= \frac{P(y_i | x_i)\, P(x_i | y_0 \dots y_{i-1})\, P(y_0, \dots y_{i-1})}{P(y_0, \dots y_i)}$$

y_i independent of other $y_0, \dots y_{i-1}$ given x_i

$$= \frac{P(y_i|x_i)\, P(x_i | y_0, \dots y_{i-1})\, \cancel{P(y_0 \dots y_{i-1})}}{P(y_i | y_0 \dots y_{i-1})\, \cancel{P(y_0 \dots y_{i-1})}}$$

$$= \frac{P(y_i|x_i)\, P(x_i | y_0 \dots y_{i-1})}{\int \underbrace{P(y_i | x_i, y_0 \dots y_{i-1})}_{P(y_i|x_i)}\, P(x_i | y_0 \dots y_{i-1})\, dx_i}$$

Now, if $\omega(x) = P(x)/Q(x)$

then $E_Q[\omega(x) f(x)] = E_P[f(x)]$.

Sequential Monte Carlo :-

→ Sample set updated each time instant, possibly re-sampling the set of state samples x^i_t.

Idea :- Exploit the form of filtering distribution for importance sampling

$$p(x_t | z_{1:t}) = c \; \underbrace{p(z_t | x_t)}_{\text{easy to evaluate}} \; \underbrace{p(x_t | z_{1:t-1})}_{\text{can draw samples from this.}}$$

Simple Particle Filter :-

Sample from prediction Distribution $Q = P(x_t | z_{1:t-1})$. then the weights must be $\omega(x) = c\, p(z_t | x_t)$ with $c = 1/p(z_t | z_{1:t-1})$.

Step 1 :- Given $S_{t-1} = \{x^i_{t-1}, \omega^i_{t-1}\}$, sample the approximate filtering distribution $p(x_{t-1} | z_{1:t-1})$

Step 2 :- Give x^i_{t-1}, use ~~dynam~~ transition distribution

$$x^i_t \sim p(x_t | x^i_{t-1})$$

it is fair sample from the prediction distribution $P(x_t | z_{1:t-1})$.

Tracking Once Again

① Assume that we know $P(X_0)$.

Now,

$$P(X_0 \mid Y_0 = y_0) = \frac{P(y_0 \mid X_0)\, P(X_0)}{P(y_0)}$$

$$\propto P(y_0 \mid X_0)\, P(X_0)$$

② Prediction Step :-

Means, finding $P(X_i \mid y_0, \ldots y_{i-1})$
given $P(X_{i-1} \mid y_0, \ldots y_{i-1})$

i.e. find X_i using $y_0, \ldots y_{i-1}$
given we know
X_{i-1} using $y_0, \ldots y_{i-1}$

we will get this from update step at $i-1$.

Then $P(X_i \mid y_0, \ldots y_{i-1})$

$$= \int P(X_i, X_{i-1} \mid y_0, \ldots y_{i-1})\, dX_{i-1}$$

$$= \int_{X_{i-1}} P(X_i \mid X_{i-1}, y_0, \ldots y_{i-1})\, P(X_{i-1} \mid y_0, \ldots y_{i-1})$$

Because given X_{i-1}, X_i is independent of $y_0, \ldots y_{i-1}$

$$= \int P(X_i \mid X_{i-1})\, P(X_{i-1} \mid y_0, \ldots y_{i-1})\, dX_{i-1}$$

& $P(X_i \mid X_{i-1})$ is given by dynamic model.

will get from update Rule at $i-1$.

Liklihood and Prior:-
→ Writing Posterior in terms of liklihood & prior.

$$P(x_{1:t} \mid z_{1:t}) = \frac{\overbrace{P(z_{1:t} \mid x_{1:t})}^{\text{liklihood}}\ \overbrace{P(x_{1:t})}^{\text{prior}}}{P(z_{1:t})}$$

Model.

$p(x_{1:t}) \rightarrow$ prior
↳ represents beliefs about the state sequences (motions) are likely.

Ex:-
First order Markov model for temporal dependence

$$P(x_t \mid x_{1:t-1}) = P(x_t \mid x_{t-1})$$

using above

$$P(x_{1:t}) = P(x_1) \prod_{j=2}^{t} P(x_j \mid x_{j-1})$$

$P(z_{1:t} \mid x_{1:t}) \rightarrow$ liklihood function
→ represents the liklihood that the state generated the observed state.

Conditional independence of observations:-

~~$P(z_{1:t} \mid x_{1:t}) = P(z_t \mid x_t)\, p($~~

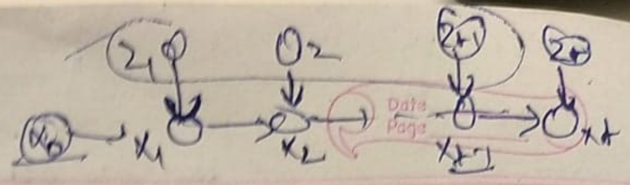
Given x_i, z_i is independent of all x_j, $j \neq i$.

$$P(z_{1:t} \mid x_{1:t}) = P(z_1, \ldots z_t \mid x_{1:t})$$

$$= P(z_1 \mid x_{1:t}) \cdots \cdots P(z_t \mid x_{1:t})$$

$$= P(z_1 \mid x_1) \cdots P(z_t \mid x_t)$$

$$= \prod_{i=1}^{t} P(z_i \mid x_i).$$



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

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

Given x_t, $z_{1:t-1}$ is independent of z_t

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

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

$$= c\; \underbrace{p(z_t \mid x_t)}_{\text{observation}}\; p(x_t \mid z_{1:t-1}).$$

Here $p(x_t \mid z_{1:t-1})$ is prediction Distribution

$$\underbrace{p(x_t \mid z_{1:t-1})}_{\text{prediction Distribution}} = \int_{x_{t-1}} p(x_t, x_{t-1} \mid z_{1:t-1})\, dx_{t-1}$$

$$= \int_{x_{t-1}} p(x_t \mid x_{t-1}, z_{1:t-1})\, p(x_{t-1} \mid z_{1:t-1})$$

$$= \int_{x_{t-1}} \underbrace{p(x_t \mid x_{t-1})}_{\text{Transition}}\, p(x_{t-1} \mid z_{1:t-1}).$$

filtering distribution at $t-1$.

Given x_{t-1}, x_t is independent of $z_{1:t-1}$.

Sequential Monte Carlo.

→ Approximate the filtering distribution, $p(x_t | z_{1:t})$, using a weighted sample set, $S = \{ x_t^j, w_t^j \}$ ~~with~~ where $w_t^j = w(x_t^j)$, $w(x)$ some weight function $w(x)$.

Monte Carlo :- Approximate, filtering distribution P with samples drawn from it, $S = \{x^{(j)}\}_{j=1}^{N}$. Then use sample statistics to approximate expectations under P; i.e. for function $f(x)$.

$$E_s[f(x)] = \frac{1}{N} \sum_{j=1}^{N} f(x^j) \xrightarrow{N \to \infty} \int f(x) P(x)\, dx \equiv E_P[f(x)].$$

But,
we don't know how to draw samples from P $p(x_t | z_{1:t})$

Particle Filter

Importance Sampling :

draw samples x^j from ~~prediction~~ ~~proposal~~ distribution, $Q(x)$.
with weights w^j. proposal

$$E_s[f(x)] = \sum_{j=1}^{N} \cancel{w(x^j)}\, w^j f(x^j) \xrightarrow{N \to \infty} E_Q[w(x) f(x)].$$

$z_1 \quad z_2 \quad \ldots \quad z_t$

$x_1 \to x_2 \to \ldots \to x_t$, $\quad p(x_t|x_{t-1})\, p(x_{1:t-1})$

Obser → estimate.

Can be said that:-

→ The observations at different times are independent when we know the ~~state~~ true underlying states.

~~Filtering Distribution~~

Posterior Distribution

$$p(x_{1:t}|z_{1:t}) \;\alpha\; p(z_{1:t}|x_{1:t})\, p(x_{1:t})$$

Posterior Distribution

$$= \prod_{\tau=1}^{t} p(z_\tau|x_\tau)\, p(x_1) \prod_{j=2}^{t} p(x_j|x_{j-1})$$

$$= p(z_t|x_t) \prod_{\tau=1}^{t-1} p(z_\tau|x_\tau)\, p(x_{1:t})$$

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

$$= p(z_t|x_t)\, p(x_t|x_{t-1})\; \underbrace{p(z_{1:t-1}|x_{1:t-1})\, p(x_{1:t-1})}_{\text{using Bayes Rule}}$$

observation ⇑ → transition

$$\underbrace{\alpha\; p(z_t|x_t)\, p(x_t|x_{t-1})\; p(x_{1:t-1}|z_{1:t-1})}$$

Recursive formulation of posterior distribution

Filtering Distribution:-

(Recursively).

$p(x_t|z$

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

~~$= \int\int \ldots \int p(x_{1:t}|z_{1:t})\, dx_{t-1} \ldots dx_1$~~

Filtering Distribution

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

Tracking:

State: $x_t \rightarrow$ state at time t. Past states: $x_{1:t}$
↳ Can be continuous (position, velocity, shape)
↳ Can be discrete (# of objects, gender, etc)

Observations: The data distribution (images) with which we constrain state estimate, based on observation equation at $z_t = f(x_t)$.
We also have observation history denoted by $z_{1:t} = (z_1, \ldots z_t)$.

Posterior: The conditional probability distribution over state which specifies all we can possibly know (according to the model) about the state history from the observations.

$$\rightarrow \quad p(x_{1:t} \mid z_{1:t})$$

Filtering Distribution:-
Marginal posterior distribution over the state at the current time given the observation history.

$$\underbrace{P(x_t \mid z_{1:t})}_{\text{filtering distribution}} = \int_{x_1} \int \int \cdots \int_{x_{t-1}} p(x_{1:t} \mid z_{1:t})\, dx_1 \ldots dx_{t-1}$$