

# Recursive Optimization

Andrius Buinovskij - 18-940-270

## 1 Kalman Filter

Is modelled after Bayesian tracking. The main difference is that the state is now a continuous random variable.

### 1.1 Model

Behold

$$x(k) = A(k-1)x(k-1) + u(k-1) + v(k-1) \quad (1)$$

$$z(k) = H(k)x(k) + w(k) \quad (2)$$

So everything is nice and linear, and the only change is that We've also added  $u(k-1)$  in in this time, it's the user input.

Well, the other real change is that the state  $x(k)$  is now continuous, so

$$p(x(k)|z(1:k-1)) = \int_{\lambda \in \mathcal{X}} p(x(k)|\lambda, z(1:k-1))p(\lambda|z(1:k-1))d\lambda \quad (3)$$

$$= \int_{\lambda \in \mathcal{X}} p(x(k)|\lambda)p(\lambda|z(1:k-1))d\lambda \quad (4)$$

So that's the prior update and then the measurement update is

$$p(x(k)|z(1:k)) = \frac{p(z(k)|x(k)) \cdot p(\bar{x}(k)|\bar{z}(1:k-1))}{\int_{\lambda \in \mathcal{X}} p(z(k)|\lambda) \cdot p(\lambda|\bar{z}(1:k-1))} \quad (5)$$

where  $b$  is the prior.

Cool.

But We need to simplify this or it'll be computationally infeasible.

### 1.2 Auxiliary variables

$$x_m(0) = x(0) \quad (6)$$

$$x_p(k) = A(k-1)x_m(k-1) + u(k-1) + v(k-1) \quad (7)$$

$$z_m(k) = H(k)x_p(k) + w(k) \quad (8)$$

Cool. We have one more piece That We need to define, and We'll do so through, well, a definition:

$$p_{x_m(k)}(\xi) = p_{x_p(k)|z_m(k)}(\xi|\bar{z}(k)) \quad (9)$$

Where  $x_m$  is supposed to represent the probability for state  $x(k)$  given  $z(1:k)$  of course,  $x_p$  is the variable representing prior update, and  $z_m$  is the measurement update.

Now We need to prove stuff about this parametrization. The claim is that

$$\text{Fact 1: } p_{x_p(k)}(\xi) = p_{x(k)|1(1:k-1)}(\xi|\bar{z}(1:k-1)) \quad \forall \xi \quad (10)$$

$$\text{Fact 2: } p_{x_m(k)}(\xi) = p_{x(k)|z(1:k)}(\xi|\bar{z}(1:k)) \quad \forall \xi. \quad (11)$$

Cool

*Proof.* The proof is by induction.

Okay so at step 0 the first condition does not make sense, since no  $k-1$ 'th measurement exists, but the second condition holds by initialization.

So then We assume the second fact for  $k-1$  and try to prove it for step  $k$ .

Fact 1:

By the total probability theorem:

$$p_{x_p(k)}(\xi) = \int p_{x_p(k)|x_m(k-1)}(\xi|\lambda) p_{x_m(k-1)}(\lambda) d\lambda. \quad (12)$$

So We're just conditioning on the prior state here using law of total probability, no big deal.

Cool. We want to show that  $p_{x_p(k)}$ , so, the prior at time  $k$ , is equal to  $p_{x(k)|z(1:k-1)}$ , which, recall, We worked out in eq.4.

$$p(x(k)|z(1:k-1)) = \int_{\lambda \in \mathcal{X}} p_{x(k)|x(k-1)}(\bar{x}(k)|\lambda) p_{x(k-1)|z(1:k-1)}(\lambda|\bar{z}(1:k-1)) d\lambda \quad (13)$$

Okay. So We have definitions for both our variables, now We just need to show that they are the same.

So, the first bit - by our inductive assumption We have that  $p_{x_m(k-1)}(\lambda) = p_{x(k-1)|z(1:k-1)}(\lambda|\bar{z}(1:k-1))$ , so that term checks out.

So now for that second equation. The idea is to express both terms, that is  $p_{x_p(k)|x_m(k-1)}(\xi|\lambda)$  and  $p_{x(k)|x(k-1)}(\bar{x}(k)|\lambda)$  in terms of the same variable, and then see what's up.

The variable both of those terms share is  $v(k-1)$ .

Recall that the formula for change of variables in the multivariate case is

$$f_Y = f_X \cdot \text{abs}\left(\det\left(\frac{\partial y}{\partial x}\right)\right)$$

Of course  $\partial x/\partial y$  also works.

So first, let's write down the change of variables I guess

$$x_p(k) = A(k-1)x_m(k-1) + u(k-1) + v(k-1) \quad (14)$$

$$x_p(k) - u(k-1) - A(k-1)x_m(k-1) = v(k-1) \quad (15)$$

$$\xi - u(k-1) - A(k-1)\lambda = v(k-1) \quad (16)$$

$$(17)$$

Aight. And if We take the derivative of that with respect to  $\xi$ , We just get one, and that's how We arrive at

$$p_{x_p(k)|x_m(k-1)}(\xi|\lambda) = p_{v(k-1)}(\xi - u(k-1) - A(k-1)\lambda) \quad (18)$$

Identical line of reasoning for  $p_{x(k)|x(k-1)}(\bar{x}(k)|\lambda)$  leads to an identical pdf, therefore We have inductively proven fact 1.

The reasoning is identical for the second fact (see course notes).

Up next is actually calculating  $x_p(k)$  and  $x_m(k)$ , since at the moment it is not defined how one would reach these values.

□