Kalman Filter I

- We have illustrated the idea behind Filtering with a transition model of order 1 (This can be extended to any order). In practice, we need to be able to calculate or compute $p(s_j|o_{1:j})$. Kalman filtering is one elegant way of doing that.
- In 1960, R.E. Kalman published his famous paper describing a recursive solution to the discrete-data linear filtering problem.
- The Kalman filter uses hypotheses of linearity (in the state transition equation, and the observation equation) and Normal distributions.

Kalman Filter II

Exercise: Given

$$\begin{cases} p(s_1) = \mathcal{N}(s_1; \mu_1, \sigma_1^2) & \text{(initial state pdf)} \\ p(s_j|s_{j-1}) = \mathcal{N}(s_j; \alpha s_{j-1}, \sigma^2) & \text{(state transition pdf)} \\ p(o_j|s_j) = \mathcal{N}(o_j; s_j, \sigma_o^2) & \text{(observation pdf)} \end{cases}$$

with the notation $\mathcal{N}(x; \mu, \sigma^2)$ corresponding to the normal pdf for r.v. x with mean μ and variance σ^2. **Show by induction that** $p(s_j|o_{1:j})$ **is a normal distribution** $\mathcal{N}(s_j; \mu_j, \sigma_j^2)$ **and give the recursive formula to compute parameters** (μ_j, σ_j^2) **assuming** $p(s_{j-1}|o_{1:j-1}) = \mathcal{N}(s_{j-1}; \mu_{j-1}, \sigma_{j-1}^2)$**.**

Kalman Filter III

Answer.

1. at $j = 1$, we know the initial state s_1 is normally distributed (given in the hypothesis).
2. assuming at $j-1$ that $p(s_{j-1}|o_{1:j-1}) = \mathcal{N}(s_{j-1}; \mu_{j-1}, \sigma_{j-1}^2)$, then solving the prediction and update steps of filtering, we find that:

$$\begin{cases} \mu_j = \alpha\ (1 - K_j)\ \mu_{j-1} + K_j\ o_j \\ \sigma_j^2 = \sigma_o^2\ K_j \\ K_j = \frac{\sigma^2 + \alpha^2 \sigma_{j-1}^2}{\sigma^2 + \sigma_o^2 + \alpha^2 \sigma_{j-1}^2} \end{cases}$$

Kalman Filter IV

1. the prediction step corresponds to (using result $< \mathcal{N}_A | \mathcal{N}_B >$):

$$\begin{aligned} p(s_j|o_{1:j-1}) &= \int \mathcal{N}(s_j; \alpha s_{j-1}, \sigma^2)\ \mathcal{N}(s_{j-1}; \mu_{j-1}, \sigma_{j-1}^2)\ ds_{j-1} \\ &= \frac{1}{|\alpha|} \int \mathcal{N}\left(s_{j-1}; \frac{s_j}{\alpha}, \frac{\sigma^2}{\alpha^2}\right)\ \mathcal{N}(s_{j-1}; \mu_{j-1}, \sigma_{j-1}^2)\ ds_{j-1} \\ &= \frac{1}{|\alpha|}\ \mathcal{N}\left(0; \frac{s_j}{\alpha} - \mu_{j-1}, \frac{\sigma^2}{\alpha^2} + \sigma_{j-1}^2\right) \\ &= \mathcal{N}(s_j; \alpha\mu_{j-1}, \sigma^2 + \alpha^2 \sigma_{j-1}^2) \end{aligned}$$

2. update (rewriting is required to be able to identify (μ_j, σ_j^2)):

$$\begin{aligned} p(s_j|o_{1:j}) &\propto \mathcal{N}(o_j; s_j, \sigma_o^2)\ \mathcal{N}(s_j; \alpha\mu_{j-1}, \sigma^2 + \alpha^2 \sigma_{j-1}^2) \\ &= \mathcal{N}(s_j; \mu_j, \sigma_j^2) \end{aligned}$$

Kalman Filter V

- These two parameters (μ_j, σ_j^2) are enough to characterise fully the Normal distribution $p(s_j|o_{1:j})$.

- The observations available $\{o_j^{(1)}\}$ are plugged in the expression (μ_j, σ_j^2) so that they can be computed iteratively.