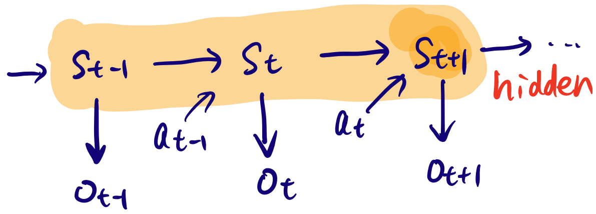
① "State Space Model"



Goal of World Model with SSM

$p(o_{t+1:t+H+1} \mid h_t, a_{t:t+H})$, h_t "history"

$\downarrow$

$p(o_{t+1} \mid h_t, a_t)$

$$p(o_{t+1} \mid h_t, a_t) = \int p(o_{t+1} \mid s_{t+1})\, p(s_{t+1} \mid h_t, a_t)\, ds_{t+1}$$

$$= \int \underbrace{p(o_{t+1} \mid s_{t+1})}_{\text{readout}} \int \underbrace{p(s_{t+1} \mid s_t, a_t)}_{\text{transition}} \cdot \underbrace{p(s_t \mid h_t)}_{\text{belif}}\, ds_t\, ds_{t+1}$$

② LGSSM = Linear Gaussian SSM

★ Both integrals are tractable under LGSSM, and the results are still Gaussian

$$\begin{cases} p(s_{t+1} \mid s_t, a_t) = N(A s_t + B a_t,\ Q) \\ p(o_{t+1} \mid s_{t+1}) = N(C s_{t+1},\ R) \end{cases}$$

③

In fact, if we assume we already know $p(s_t \mid h_t) = N(\mu_t, P_t)$

then

$$\begin{cases} p(s_{t+1} \mid h_t, a_t) = \int p(s_{t+1} \mid s_t, a_t)\, p(s_t \mid h_t)\, ds_t \\ \qquad = N(A\mu_t + Ba_t,\ A P_t A^T + Q) = N(\bar{\mu}, \bar{P}) \\ p(o_{t+1} \mid h_t, a_t) = \int p(o_{t+1} \mid s_{t+1})\, p(s_{t+1} \mid h_t, a_t)\, ds_{t+1} \\ \qquad = N(C\bar{\mu},\ \underbrace{C\bar{P}C^T + R}_{\to S}) \end{cases}$$

④ Then how can we compute $p(s_t \mid h_t)$?

|

The idea is to derive $p(s_t|h_t) \longrightarrow p(s_{t+1}|h_{t+1})$. If this is possible we can then propagate. Both have to be "Gaussian".

$$p(s_{t+1}|h_{t+1}) \propto p(o_{t+1}|s_{t+1})\cdot p(s_{t+1}|h_t,a_t)$$

★ This is Bayes' not the integral.

$$\Rightarrow N(\mu_{t+1},P_{t+1}) \propto \underline{N(o_{t+1}|Cs_{t+1},R)\,N(s_{t+1}|\bar{\mu},\bar{P})}$$

5

$\downarrow \propto$

$$\exp\left[-\tfrac{1}{2}(s-\mu_{t+1})^T P_{t+1}^{-1}(s-\mu_{t+1})\right]$$

$\downarrow \nabla_s \big|_{s=\mu_{t+1}} = 0$

$$P_{t+1}^{-1}(s-\mu_{t+1})$$

$\downarrow \propto$

$$\exp\left[-\tfrac{1}{2}(o_{t+1}-Cs)^T R^{-1}(o_{t+1}-Cs) - \tfrac{1}{2}(s-\bar{\mu})^T \bar{P}^{-1}(s-\bar{\mu})\right]$$

$\downarrow \nabla_s \big|_{s=\mu_{t+1}} = 0$

$$\mu_{t+1}-\bar{\mu} = \bar{P}C^T R^{-1}(o_{t+1}-C\mu_{t+1})$$

Let $\underline{v = o_{t+1} - C\bar{\mu}}$, $\underline{S = C\bar{P}C^T + R}$

$$\Rightarrow \mu_{t+1}-\bar{\mu} = \boxed{\bar{P}C^T S^{-1}}\, v = Kv$$

$K \to$ Kalman Gain

∇_s^2 ∇_s^2

$$\boxed{P_{t+1}^{-1} = C^T R^{-1} C + \bar{P}^{-1}}$$

$$\Rightarrow \boxed{P_{t+1} = \bar{P} - KSK^T}$$

In this case, we have derived the propagation of states

$$\underset{\boxed{p(s_t|h_t)}}{\mu_t, P_t} \xrightarrow[a_t]{} \underset{\boxed{p(s_{t+1}|h_t,a_t)}}{\bar{\mu}, \bar{P}} \xrightarrow[o_t]{} \begin{array}{l} \mu_{t+1} = \bar{\mu} + Kv \\ P_{t+1} = \bar{P} - KSK^T \end{array}$$

prediction from the model

new observation

$\Rightarrow \underline{v = o_t - C\bar{\mu} \text{ (error)}}$

Appendix. I

$$\mu_{t+1} - \bar{\mu} = \bar{P} C^T R^{-1} (o_{t+1} - C\mu_{t+1})$$

$$v = o_{t+1} - C\bar{\mu} = o_{t+1} - C\mu_{t+1} + C(\mu_{t+1} - \bar{\mu})$$

$$= o_{t+1} - C\mu_{t+1} + C\bar{P}C^T R^{-1}(o_{t+1} - C\mu_{t+1})$$

$$= [I + C\bar{P}C^T R^{-1}](o_{t+1} - C\mu_{t+1})$$

$$= [\underbrace{R + C\bar{P}C^T}] R^{-1} (o_{t+1} - C\mu_{t+1})$$

$$= S R^{-1} (o_{t+1} - C\mu_{t+1})$$

$$\boxed{\mu_{t+1} - \bar{\mu} = \underbrace{\bar{P} C^T S^{-1}}_{K \,:\, \text{Kalman Gain}} v}$$

Appendix II.

$$P_{t+1}^{-1} = C^T R^{-1} C + \bar{P}^{-1}$$

$$\Rightarrow P_{t+1}^{-1} K = [C^T R^{-1} C + \bar{P}^{-1}] \bar{P} C^T S^{-1}$$

$$= C^T R^{-1} C \bar{P} C^T S^{-1} + C^T S^{-1}$$

$$= C^T R^{-1} [\underbrace{C\bar{P}C^T + R}_{S}] S^{-1}$$

$$= C^T R^{-1}$$

$$\Rightarrow P_{t+1}^{-1} = P_{t+1}^{-1} K C + \bar{P}^{-1}$$

$$\Rightarrow P_{t+1} = (I - KC)\bar{P} = \bar{P} - KC\bar{P}, \quad K = \bar{P} C^T S^{-1}$$

$$\Rightarrow \boxed{P_{t+1} = \bar{P} - KSK^T}$$