Homework 2 (due: 10/11/2023)

1. (a) For all t, we have the filtering formula

$$\mathbf{P}(R_t|u_{1:t}) = \alpha \mathbf{P}(u_t|R_t) \sum_{R_{t-1}} \mathbf{P}(R_t|R_{t-1}) P(R_{t-1}|u_{1:t-1})$$

At the fixed point, we additionally expect that $\mathbf{P}(R_t|u_{1:t}) = \mathbf{P}(R_{t-1}|u_{1:t-1})$. Let the fixed-point probabilities be $\langle \rho, 1-\rho \rangle$. This provides us with a system of equations:

$$\begin{aligned}
\langle \rho, 1-\rho \rangle &= \alpha \langle 0.9, 0.2 \rangle (\langle 0.7, 0.3 \rangle \rho + \langle 0.3, 0.7 \rangle (1-\rho)) \\
&= \alpha \langle 0.9, 0.2 \rangle (\langle 0.4\rho, -0.4\rho \rangle + \langle 0.3, 0.7 \rangle) \\
&= \frac{1}{0.9(0.4\rho + 0.3) + 0.2(-0.4\rho + 0.7)} \langle 0.9, 0.2 \rangle (\langle 0.4\rho, -0.4\rho \rangle + \langle 0.3, 0.7 \rangle)
\end{aligned}$$

Solving this system, we find that $\rho \approx 0.8967$. Let's suppose $p_t = \mathbf{P}(R_t = true|u_{1:t})$. Then we can compute $p_t - p_{t-1}$ as:

$$p_t - p_{t-1} = \frac{0.27 + 0.36 p_{t-1}}{0.41 + 0.28 p_{t-1}} - p_{t-1} = \alpha(0.27 - 0.5 p_{t-1} - 0.28 p_{t-1}^2),$$

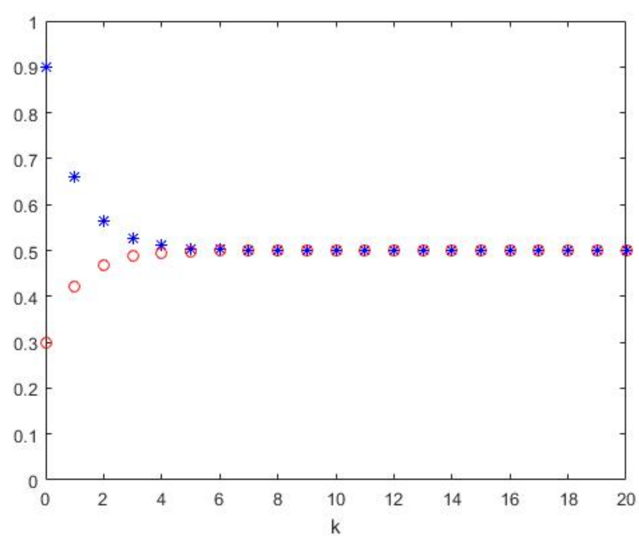
which is larger than 0 until $p_t \to 0.8967$. So, p_t increases monotonically.

(b) The probability converges to $\langle 0.5, 0.5 \rangle$. This convergence makes sense if we consider a fixed-point equation for $\mathbf{P}(R_{2+k}|U_1, U_2)$:

$$\mathbf{P}(R_{2+k}|U_1, U_2) = \langle 0.7, 0.3 \rangle P(r_{2+k-1}|U_1, U_2) + \langle 0.3, 0.7 \rangle P(\neg r_{2+k-1}|U_1, U_2)$$

$$\begin{aligned}
\mathbf{P}(r_{2+k}|U_1, U_2) &= 0.7 P(r_{2+k-1}|U_1, U_2) + 0.3(1 - P(r_{2+k-1}|U_1, U_2)) \\
&= 0.4 P(r_{2+k-1}|U_1, U_2) + 0.3
\end{aligned}$$

That is, $P(r_{2+k}|U_1, U_2) = 0.5$. Notice that the fixed point does not depend on the initial evidence. Here is the plot when $\mathbf{P}(r_2|U_1, U_2) = 0.9$, and 0.3.



2. The propagation of the l message is identical to that for filtering:

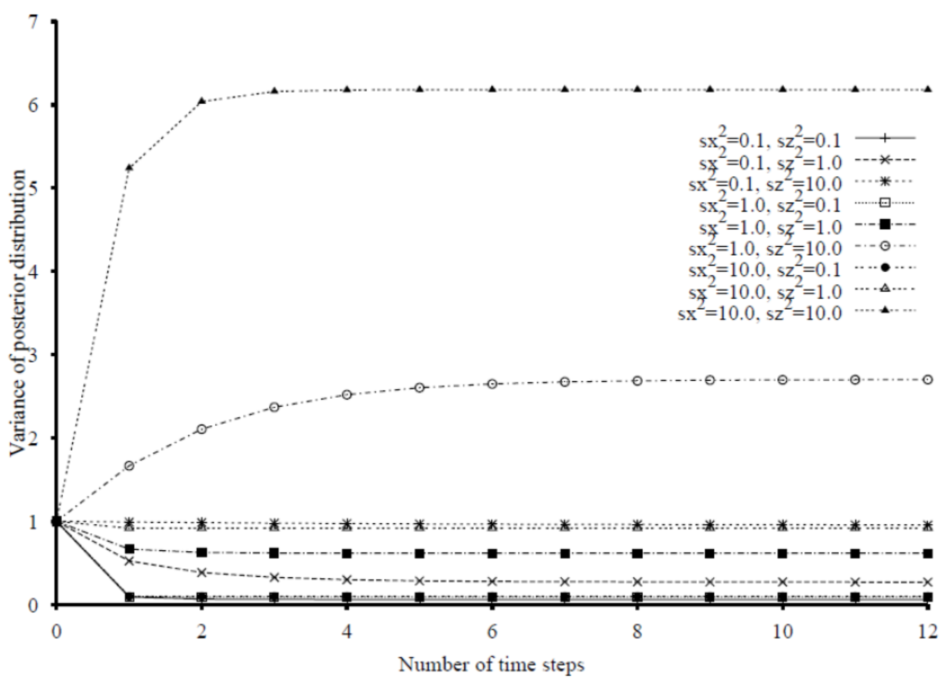
$$l_{1:t+1} = \alpha \mathbf{O}_{t+1} \mathbf{T}^T l_{1:t}$$

Since l is a column vector, each entry l_i of which gives $P(X_t = i, \mathbf{e}_{1:t})$, the likelihood is obtained simply by summing the entries:
$L_{1:t} = P(\mathbf{e}_{1:t}) = \sum_i l_i$

3.

$$\begin{aligned}
P(x_1|z_1) &= \alpha e^{-\frac{1}{2}\left(\frac{(z_1-x_1)^2}{\sigma_z^2}\right)} e^{-\frac{1}{2}\left(\frac{(x_1-\mu_0)^2}{\sigma_0^2+\sigma_x^2}\right)} \\
&= \alpha e^{-\frac{1}{2}\left(\frac{(\sigma_0^2+\sigma_x^2)(z_1-x_1)^2+\sigma_z^2(x_1-\mu_0)^2}{\sigma_z^2(\sigma_0^2+\sigma_x^2)}\right)} \\
&= \alpha e^{-\frac{1}{2}\left(\frac{(\sigma_0^2+\sigma_x^2)(z_1^2-2z_1x_1+x_1^2)+\sigma_z^2(x_1^2-2\mu_0x_1+\mu_0^2)}{\sigma_z^2(\sigma_0^2+\sigma_x^2)}\right)} \\
&= \alpha e^{-\frac{1}{2}\left(\frac{\sigma_0^2+\sigma_x^2+\sigma_z^2)x_1^2-2((\sigma_0^2+\sigma_x^2)z_1+\sigma_z^2\mu_0)x_1+c}{\sigma_z^2(\sigma_0^2+\sigma_x^2)}\right)} \\
&= \alpha' e^{-\frac{1}{2}\left(\frac{\left(x_1-\frac{(\sigma_0^2+\sigma_x^2)z_1+\sigma_z^2\mu_0}{\sigma_0^2+\sigma_x^2+\sigma_z^2}\right)^2}{(\sigma_0^2+\sigma_x^2)\sigma_z^2/(\sigma_0^2+\sigma_x^2+\sigma_z^2)}\right)}
\end{aligned}$$

4. (a) Here is the plot.



(b) We can find a fixed point by solving
$\sigma^2 = \frac{(\sigma^2+\sigma_x^2)\sigma_z^2}{\sigma^2+\sigma_x^2+\sigma_z^2}$
for σ^2. Using the quadratic formula and requiring $\sigma^2 \geq 0$, we obtain
$\sigma^2 = \frac{-\sigma_x^2+\sqrt{\sigma_x^4+4\sigma_x^2\sigma_z^2}}{2}$.
We omit the proof of convergence, which presumably, can be done by showing that the update is a contradiction (i.e., after updating, two different starting points for σ_t become closer).

(c) As $\sigma_x^2 \to 0$, we see that the fixed point $\sigma^2 \to 0$ also. This is because $\sigma_x^2 = 0$ implies a deterministic path for the object. Each observation supplies more information about this path, until its parameters are known completely.
As $\sigma_z^2 \to 0$, the variance update gives $\sigma_{t+1}^2 \to 0$ immediately. That is, if we have an exact observation of the object's state, then the posterior is a delta function about that observed value regardless of the transition variance.

5. (MATLAB)

6. (MATLAB)