Exercise: Kalman Filter

Consider the following stochastic linear dynamical system:

$$y_t = x_t^{(1)} + x_t^{(2)} + r_t$$

$$x_t^{(1)} = q_t^{(1)}$$

$$x_t^{(2)} = \gamma x_{t-1}^{(2)} + \alpha W_{t-1} x_{t-1}^{(1)} + q_t^{(2)}$$

$$W_{t-1} = \begin{cases} 0 & y_{t-1} \geq \theta \\ -1 & y_{t-1} < \theta \end{cases}$$

with parameters $\gamma < 1$ and $m_0 = 0$. $q_t^{(i)} \in \mathcal{N}\left(0, Q^{(i)}\right)$ and $r_t \in \mathcal{N}(0, R)$ are Gaussian distributed white noise (i.i.d.). As discussed in the lecture, this system describes the pitch trajectory of a bird during the reinforcement learning task where y_t is the mean-subtracted pitch of syllable rendition t, and where the hidden state $x_t^{(1)}$ determines the exploration and $x_t^{(2)}$ integrates the learned changes. The bird learns by correlating exploration and reward (whether or not it gets punished) with a learning rate α. W_{t-1} is zero if the bird escapes punishment (pitch y_{t-1} was above threshold θ) and negative if a white noise burst was delivered for punishment (pitch y_{t-1} was below threshold θ).

Analyze the system step by step:

1. Let's only look at $x_t^{(2)}$ and set $\alpha = 0$ ($x_t = \gamma x_{t-1} + q_t$). What is the variance of x_t after t timesteps? How does the variance behave for large t ($t \rightarrow \infty$) ?
2. Assume $\gamma = 1$: how does the variance of x_t grow as a function of t?

3. Now let's assume we have noisy measurements y_t of x_t (keep $\alpha = 0$ and ignore $x_t^{(1)}$, thus $y_t = x_t + r_t$). Using the Kalman filter, what is the variance of the estimation of x_t given all the measurements up to time t?
4. What happens now for $\gamma = 1$?

Matlab:

5. Simulate the dynamical system from question 3 and describe the behavior of y_t. Vary $0 \leq \gamma \leq 1$ and the proportion of state and measurement noise and discuss extreme cases. What is the influence of each of the two noise sources (R and Q) on the observed trajectories y_t?

Reinforcement learning of pitch:

6. Use the given Matlab program to simulate pitch reinforcement data using the equations above. Double α or divide it by two and describe the behavior of the trajectory.
7. Add the Kalman filter equations to the code and use them to estimate the hidden state $x_t^{(2)}$ from the generated measurements. How similar is the estimated state to the original one? Use 2α or $\frac{\alpha}{2}$ for the Kalman filter estimation (without changing the simulated data. How does the estimated state compare to the one where the correct α is used?