HOLON INSTITUTE OF TECHNOLOGY, DEPARTMENT OF APPLIED MATHEMATICS

# Sequential Models in Data Science
# Recursive Bayesian Filtering and Markov Models

Yirmeyahu Kaminski

## INTRODUCTION

**The exercise is to be done by pairs of students. Each pair must present a Jupyter notebook containing all theoretical answers as well as all Python functions. The notebook will be tested as this. No modification should be necessary for making it work.**
All over the questions, $x_k$ and $y_k$ denotes respectively the state vector and the measurements.

## 1 KALMAN FILTER WITH BIASED NOISE

Derive the Kalman filter equations for the following linear-Gaussian filtering model with non-zero mean noises:

$$\begin{aligned} x_k &= Ax_{k-1} + q_{k-1}, \\ y_k &= Hx_k + r_k, \end{aligned}$$

where $q_{k-1} \sim \mathcal{N}(m_q, Q)$ and $r_k \sim \mathcal{N}(m_r, R)$.
**It is recommended to use the two lemmas that were defined in task 2, section 3.**

## 2 EXTENDED KALMAN FILTER

Consider the following non-linear space model:

$$\begin{aligned} x_k &= x_{k-1} - 0.01\sin(x_{k-1}) + q_{k-1}, \\ y_k &= 0.5\sin(2x_k) + r_k, \end{aligned}$$

where $q_{k-1} \sim \mathcal{N}(0, 0.01^2)$ and $r_k \sim \mathcal{N}(0, 0.02)$.

1. Write the extended Kalman filter obtained by the **Taylor** linearization.
2. Write the extended Kalman filter obtained by the **statistical** linearization.
3. Implement in python these versions of the EKF. Draw on the same graph in three different colors (with a clear caption) the trajectory of the system and the estimated states by the two different methods of EKF.

## 3 Grid Based Filter

Consider a situation where the state space is discrete and finite: $\forall n, x_n \in S$ and $|S| = N < \infty$. Let $S = \{x^1, \cdots, x^N\}$.
Define the coefficients $w^i_{k-1|k-1}$ by the following relation:

$$p(x_{k-1}|y_{1:k-1}) = \sum_{i=1}^{N} w^i_{k-1|k-1}\delta(x_{k-1} - x^i).$$

The prediction step leads to the following expression:

$$p(x_k|y_{1:k-1}) = \sum_{i=1}^{N} w^i_{k|k-1}\delta(x_k - x^i).$$

Then the update step leads to:

$$p(x_k|y_{1:k}) = \sum_{i=1}^{N} w^i_{k|k}\delta(x_k - x^i).$$

Prove the following relations:

1. $\forall i, w^i_{k|k-1} = \sum_{j=1}^{N} w^j_{k-1|k-1} p(x^i|x^j)$, where $p(x^i|x^j)$ is the transition probability from state $x^j$ to state $x^i$,
2. $\forall i, w^i_{k|k} = \frac{w^i_{k|k-1} p(y_k|x^i)}{\sum_{j=1}^{N} w^j_{k|k-1} p(y_k|x^j)}$, where $p(y_k|x^l)$ is the probability distribution of the measurement $y_k$ given the system is in state $x^l$.
3. Consider the following situation. Let $x_k$ be a sequence of real numbers that satisfies:

$$\begin{aligned} x_k &= x_{k-1} + Lv_{k-1} \\ y_k &= (\tfrac{1}{2} + r_k)x_k, \end{aligned}$$

where $L$ is a positive real number, $v_{k-1}$ is a Rademacher random variable the distribution of which is given by:

$$\mathbb{P}(v_{k-1} = i) = \begin{cases} \frac{1}{2} \text{ if } i \in \{-1, 1\} \\ 0 \text{ otherwise} \end{cases}$$

and $r_k$ is a uniform random variable in $[0, 1/2]$. Let $x_0$ be a Gaussian random variable, such that $x_0 \sim \mathcal{N}(0, 1)$. Implement a grid based filter to track the state vector $x_k$ over 100 time steps, within the state space $S = \{x_0, x_0 \pm L, \ldots, x_0 \pm NL\}$, where $N = 100$. Take $L = 2$. Draw a plot with the actual state and the estimated one with two different colors and an implicit legend, over the tracking interval.

## 4 Rejection Sampling and Monte Carlo Integration

Consider a probability density distribution $f : \mathbb{R} \to \mathbb{R}_+$. Let $g$ be another density distribution such that there exists $c > 0$ such that $f(x) \leq cg(x)$ for all $x \in \mathbb{R}$. The purpose of rejection sampling is to make use of $g$, which is supposed to be simple to sample in order to produce samplings from $f$. The algorithm works as follows:

1. Draw a random variable $x$ following the distribution given by $g$,
2. Draw a random variable $u$ from the uniform distribution over $(0,1)$,
3. If $u \leq \frac{f(x)}{cg(x)}$, accept $x$ as a sample from the distribution given by $f$. Otherwise return to 1.

To prove that this algorithm indeed generates a sample from $f$, we shall prove that the cumulative distribution function of $x$ given by:

$$F_x(t) = \mathbb{P}\left(x \leq t \mid u \leq \frac{f(x)}{cg(x)}\right),$$

satisfies $F_x(t) = F(t)$, where $F$ is the cumulative distribution function of $f$. Answer the following questions, that will guide you.

1. Prove that $\mathbb{P}\left(x \leq t \mid u \leq \frac{f(x)}{cg(x)}\right) = \frac{\mathbb{P}\left(x \leq t, u \leq \frac{f(x)}{cg(x)}\right)}{\mathbb{P}\left(u \leq \frac{f(x)}{cg(x)}\right)}$,
2. Prove that $\mathbb{P}\left(u \leq \frac{f(x)}{cg(x)}\right) = \frac{1}{c}$,
3. Deduce that $F_x(t) = F(t)$.

Let $f(x) = 0.2\frac{1}{\sqrt{2\pi 0.1}}\exp\left(-\frac{(x-1)^2}{2\cdot 0.01}\right) + 0.8\frac{1}{\sqrt{2\pi 0.8}}\exp\left(-\frac{(x-2)^2}{2\cdot 0.64}\right)$ and $g(x) = \frac{1}{\sqrt{2\pi 10}}\exp\left(-\frac{x^2}{2\cdot 100}\right)$. Find the smallest $c$ such that $f \leq cg$. Apply the rejection algorithm to compute $n = 1000$ samples of $f$. Draw the graph of $f$ and the histogram of these samples on the same drawing.
Use this algorithm to compute the following integral with the Monte-Carlo method:

$$\int_{-\infty}^{+\infty} (x^2 - 3x + 5)\cos(x) f(x)\, dx.$$

# **May the force be with you!**