EL2320

APPLIED ESTIMATION

LABORATORY REPORT

LAB 2: PARTICLE FILTER

Jiang, Sifan
sifanj@kth.se

December 2, 2018

1 Introduction

Particle Filters:

1. What are the particles of the particle filter?

   Particles are also called samples. They are used to represent the posterior distribution of some stochastic process given noisy.

2. What are importance weights, target distribution, and proposal distribution and what is the relation between them?

   - Importance wights $w_t^{[m]}$ is the probability of a measurement z_t regarding a particle $x_t^{[m]}$. So that $w_t^{[m]} = p(z_t|x_t^{[m]})$.
   - Target distribution represents the belief $bel(x_t)$.
   - Proposal distribution represents the $\overline{bel}(x_t)$ based on the prior belief over state x_{t-1} and the control u_t.

   Target distribution can be sampled from the proposal distribution. Each sample could be drawn with a probability equals to importance weight.

3. What is the cause of particle deprivation and what is the danger?

   The cause of particle deprivation is the lack of particles which could be caused by re-sampling. The danger of particle deprivation is that there would be no particles around the true state of the system.

4. Why do we re-sample instead of simply maintaining a weight for each particle always.

   Without re-sampling most particles would represent states with a low likelihood after some time and the filter would loose track of the "good" hypotheses.

5. Give some examples of the situations which the average of the particle set is not a good representation of the particle set.

   When the posterior is multi-modal, the average of the particle set would be not good for representing the multiple regions of true interest.

6. How can we make inferences about states that lie between particles.

   Inferences about states that lie between particles can be made by histograms, where the bin is set by adjacent particles, or Gaussian kernels where each particle is used as the center of a Gaussian kernel.

7. How can sample variance cause problems and what are two remedies?

   High sample variance can cause incorrect estimation about the true state.

   Remedies:

   - Reducing the re-sampling frequency.
   - Using a sequential stochastic process instead.

8. For robot localization for a given quality of posterior approximation, how are the pose uncertainty (spread of the true posteriori) and number of particles we chose to use related.

   Higher pose uncertainty would result in larger spread of the posteriori, so, larger number of particles are needed.

2 Warm up problem with the Particle Filter

2.1 Prediction

- **Question 1:** What are the advantages/drawbacks of using (6) compared to (8)? Motivate.

  In section 2D State Space, (6) uses the same value of angle θ_0 for all time steps. This equation will give precise prediction when target moving on a line if there is no noise or diffusion. Also, same value for angle could improve computational time. However, since there is noise in the system, the estimate of the motion will result in a wave-like pattern. In section 3D State Space, (8) uses the angle obtained from the previous time stamp to replace the fixed angle, the estimate of the motion will result in a smoother line compared to what (6) gives.

- **Question 2:** What types of circular motions can we model using (9)? What are the limitations (what do we need to know/fix in advance)?

  We need to know the initial velocity v_0 and ω_0 to determine what exactly the type of circular motion would (9) modeled. Equation (9) could model any circular motion with angular velocity based on the previous time stamp.

2.2 Sensor Model

- **Question 3:** What is the purpose of keeping the constant part in the denominator of (10)?

  The purpose of keeping the constant part in the denominator of Equation (10) is to make the integral over the whole distribution to be 1.0 making the likelihood function for the observation a probability distribution.

2.3 Re-Sampling

- **Question 4:** How many random numbers do you need to generate for the Multinomial re-sampling method? How many do you need for the Systematic re-sampling method?

  Based on the algorithms, we need M random number for the Multinomial re-sampling method where M is the number of . And we need only 1 random number for the Systematic re-sampling.

- **Question 5:** With what probability does a particle with weight $\omega = \frac{1}{M} + \epsilon$ survive the re-sampling step in each type of re-sampling (vanilla and systematic)? What is this probability for a particle with $0 \leq \omega < \frac{1}{M}$? What does this tell you? (Hint: it is easier to reason about the probability of not surviving, that is M failed binary selections for vanilla, and then subtract that amount from 1.0 to find the probability of surviving.

  In the Multinomial re-sampling, for each particle $p^{[k]}$, there is a probability of $\omega^{[k]}$ to be selected in one iteration. The probability of being chosen is equal to its importance weight. So, the probability of not being selected for particle $p^{[k]}$ would be $1 - \omega^{[k]}$. So, the probability for a certain particle of not being selected for N iterations would be $(1-\omega^{[k]})^N$, thus, the probability of surviving the re-sampling would be $1 - (1 - \omega^{[k]})^N$.

  In the Systematic re-sampling, particles $p^{[k]}$ with importance weight $\omega^{[k]} = \frac{1}{M} + \epsilon$ where $\epsilon \in (0,1]$, would always be selected since the selection step size $\frac{1}{M} < \frac{1}{M} + \epsilon$. If the importance weight $\omega^{[k]} \in \left[0, \frac{1}{M}\right)$, the probability of surviving would be $M\omega^{[k]}$. The random number r_0 is uniformly distributed over $\left[0, \frac{1}{M}\right]$, also the probability of selection of a particle is proportional to its weight, the probability of surviving the re-sampling would be $M\omega^{[k]}$.

2.4 Experiments

- **Question 6:** Which variables model the measurement noise/process noise models?

  Variable `Sigma_Q` model the measurement noise model.

  Variable `Sigma_R` model the process noise model.

- **Question 7:** What happens when you do not perform the diffusion step? (You can set the process noise to 0)

  If the process noise is set to 0, according to equation (3), the predicted state would be equal to applying motion. Also, for a fixed target, $\bar{u}_t = 0$ for all time stamps, thus $\bar{x}_t^m = x_{t-1}^m$. So, it's predictable that the result would be M copies of the same particle.

- **Question 8:** What happens when you do not re-sample? (set RESAMPLE_MODE=0)

  If not re-sampling the particles, then all the particles would be remained no matter how much the importance weight. The motion of each particle simulates the estimation of the true state from a different initial state and no convergence would be obtained. As a result, the particles would spread over the whole region and with large estimation error.

- **Question 9:** What happens when you increase/decrease the standard deviations (diagonal elements of the covariance matrix) of the observation noise model? (Try values between 0.0001 and 10000)

  When increasing the standard deviations of the observation noise model from 0.0001 to 10000, the expectation to the accuracy of the estimation of the measurements is decreasing. If the standard deviations is small, then the estimates would not likely to be close to the true measurements, thus being regarded as outliers. So the estimate would hard to converge to the true state.

  When the standard deviations is big, then the estimates are easy to follow the true measurements closely, thus the estimate would easy to converge to the true state but with big variance and uncertainty.

- **Question 10:** What happens when you increase/decrease the standard deviations (diagonal elements of the covariance matrix) of the process noise model? (Try values between 0.0001 and 10000)

  When increasing the standard deviations of the process noise model, the diversity of the particle population is also increasing. When the deviation is small, the cluster of particles around the true state is also small and condensed in a certain region but with no guarantee to convergence. When the deviation is big, the particles spread around the true state and have a relatively big radius of the cluster.

- **Question 11:** How does the choice of the motion model affect a reasonable choice of process noise model?

  If the motion model is accurate to describe the true state, then the process noise can be small. If the motion model is known to be not accurate to model the true motion, the process noise should be large enough for particles to be able to estimate the true motion.

- **Question 12:** How does the choice of the motion model affect the precision/accuracy of the results? How does it change the number of particles you need?

  If a motion model model the actual motion accurately, then the results would also be precise and accurate. As a result, small number of particles would be needed because of the high accuracy of the particles. While larger number of particles would be needed for less accurate model so that the system could converge to true state.

- **Question 13:** What do you think you can do to detect the outliers in third type of measurements? Hint: what happens to the likelihoods of the observation when it is far away from what the filter has predicted?

  We can introduce a threshold value for the likelihoods of the observation. When the measurement is far away from what the filter has predicted, the likelihood of this observation would be very small and could be smaller than the threshold. So, in this way, outliers could be detected and removed in third type of measurements.

- **Question 14:** Using 1000 particles, what is the best precision you get for the second type of measurements of the object moving on the circle when modeling a fixed, a linear or a circular motion(using the best parameter setting)? How sensitive is the filter to the correct choice of the parameters for each type of motion?

  Change the function: `pf_track(Z, X, VERBOSE, Dimension, Motion, Sigma_Q, Sigma_R)`, so that we can easily change the parameters.

  - `Z`, the measurement of the object, is `circ_meas_2`.
  - `X`, the ground true information, is `circ_true`.
  - `Dimension` is 3.
  - `Sigma_Q`, the measurement noise covariance matrix, is chosen from 5 values: 100, 250, 500, 750, and 1000.
  - `Sigma_R`, the prediction noise covariance matrix, is chose from 5 values: 1, 10, 25, 50, 75, and 100.

  In all I would try 25 pairs of noises and from which I could obtain the best parameter setting so that having the minimized error.

  When `Motion` is 0 which means modeling a fixed motion, we have following noises so that the error is 11.3 ± 5.4. It is very sensitive to the correct choice of the parameters and requires high level of process noise in order for the particles to be able to follow the true state.

  $$R = \begin{bmatrix} 25 & 0 & 0 \\ 0 & 25 & 0 \\ 0 & 0 & 0.01 \end{bmatrix}$$
  $$Q = \begin{bmatrix} 500 & 0 \\ 0 & 500 \end{bmatrix}$$

  When `Motion` is 1 which means modeling a linear motion, we have following noises so that the error is 7.5 ± 3.5. It is not very sensitive to the correct choice of the parameters.

  $$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.01 \end{bmatrix}$$
  $$Q = \begin{bmatrix} 250 & 0 \\ 0 & 250 \end{bmatrix}$$

  When `Motion` is 2 which means modeling a circular motion, we have following noises so that the error is 7.2 ± 3.3. It is not sensitive to the correct choice of the parameters and

allows low level noise to converge to true state.

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.01 \end{bmatrix}$$

$$Q = \begin{bmatrix} 250 & 0 \\ 0 & 250 \end{bmatrix}$$

3 Main Problem: Monte Carlo Localization

- **Question 15:** What parameters affect the mentioned outlier detection approach? What will be the result of the mentioned method if you model a very weak measurement noise $|Q| \to 0$?

  The threshold λ_Ψ affects the mentioned outlier detection approach. The higher the value of λ_Ψ is, the easier for a particle to be an outlier. Also, the measurement noise Q would also have effect on the detection. If the noise is very weak so that $|Q| \to 0$, the confidence in the measurement result would be very high and the tolerance between the difference between the prediction and true value decreases, thus making more measurements are regarded as outliers.

- **Question 16:** What happens to the weight of the particles if you do not detect outliers?

  If not detecting outliers, then all the measurements are considered to be valid so the measurements would be equally trusted. And because the wrong measurements would be taken into consideration, the weights of the particles would also be wrong.

3.1 Data Sets

3.1.1 map_sym2.txt + so_sym2_nk

`part_bound` would have influence on the initial position of each particles in the global localization. According to the code, we would have the initial x and y coordinates for the particles:

$$S(1,:) = rand(1, M) * [bound(2) - bound(1) + 2 * part\_bound] + bound(1) - part\_bound \quad (1)$$
$$S(2,:) = rand(1, M) * [bound(4) - bound(3) + 2 * part\_bound] + bound(3) - part\_bound \quad (2)$$

So, the higher the value of `part_bound` is, the more likely for the particles to spread out to a bigger region, thus the particles are easier to be distinguished from each and kept as reliable hypotheses.

In `map_sym3.txt`, there are 4 landmarks thus resulting in 4 hypotheses. Increase the number of particles M from 1000 to 10000 would making the hypotheses more reliable. When M is small and particles are spread over a large state space, the main drawback observed is particle deprivation which would discard particles near the correct state during the re-sampling step.

The multinomial sampling is not so effective when it comes to preserving multiple hypotheses since the multinomial sampling introduces random value for each particles.

When modeling stronger measurement noises, the particles would spread over a larger state space, and more hypotheses are preserved with lower accuracy.
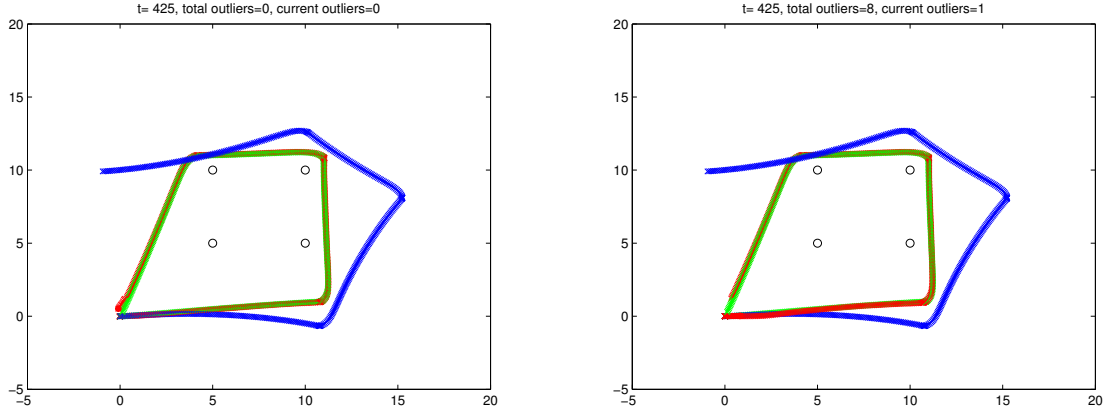
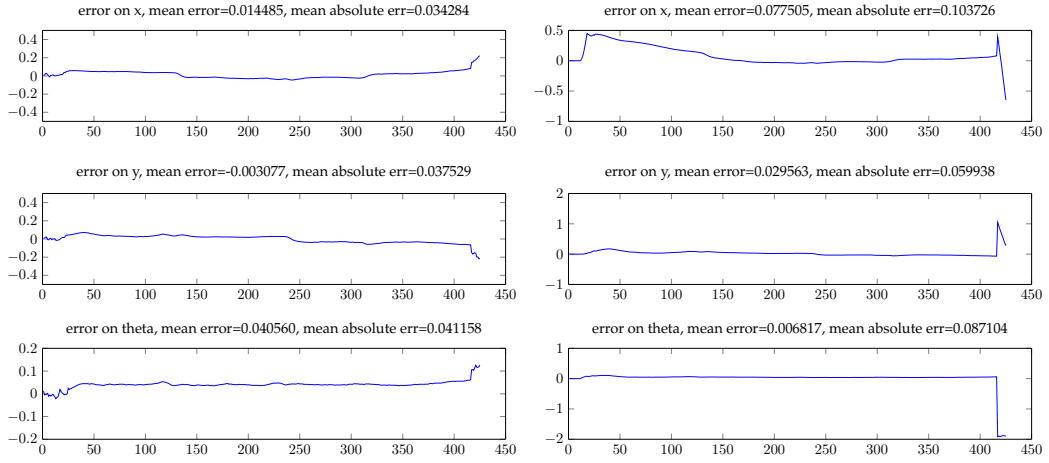Figure 1: Result when the number of particles $M = 1000$ (left) and $M = 10000$ (right).

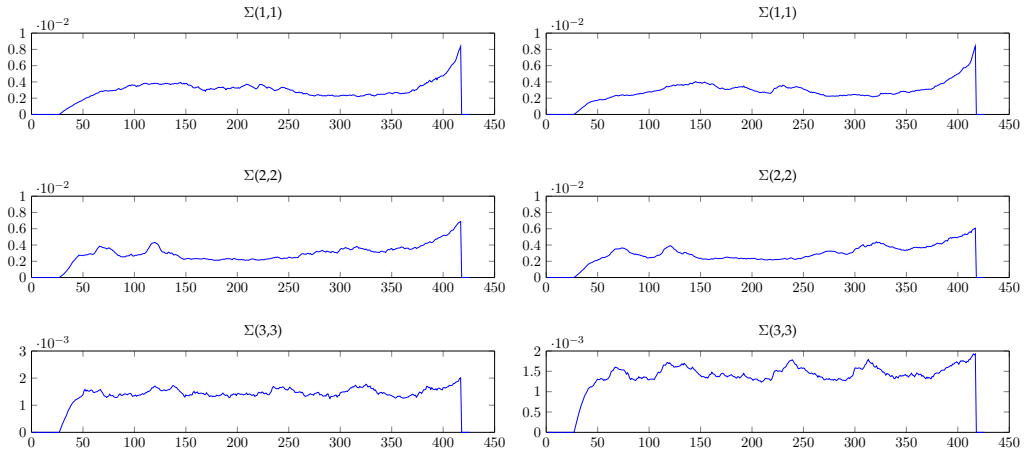Figure 2: Mean pose error when the number of particles $M = 1000$ (left) and $M = 10000$ (right).

Figure 3: Uncertainty of pose when the number of particles $M = 1000$ (left) and $M = 10000$ (right).

3.1.2   map_sym3.txt + so_sym3_nk
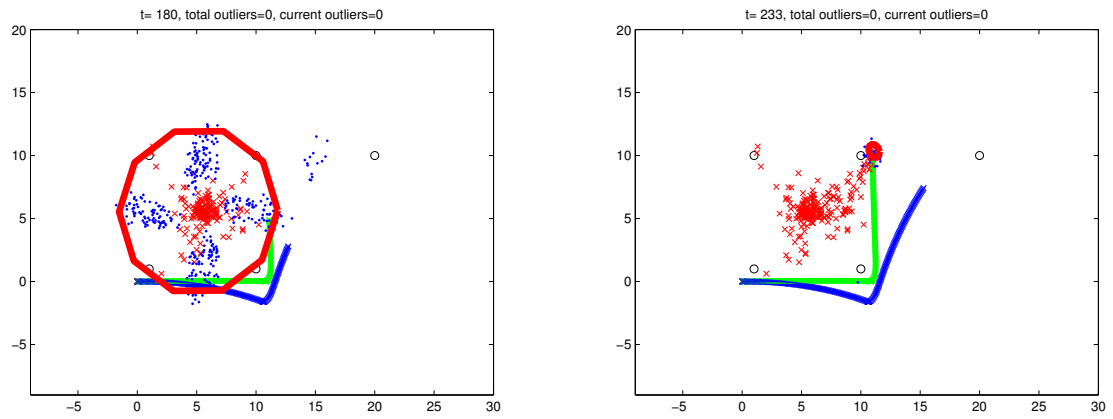
Figure 4 illustrate the images just before and just after the convergence.

Figure 4: Time stamps of convergence.