EL2320 Applied Estimation - Lab 2: PF

Wenjie Yin

yinw@kth.se

PART I – Preparatory Questions

Particle Filters

1. **What are the particles of the particle filter?**

   The particles of the particle filter are a collection of samples drawn from a posterior distribution to represent continuous distribution.

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

   The *importance weights* $\omega_t^{[m]}$ is the probability of drawing particle of the measurement z_t, which is defined as $\omega_t^{[m]} = p(z_t|x_t^{[m]})$. And the *target distribution* is the probability distribution where we obtain samples, defined as $p(x_t|\{z_t\},\{u_t\})$. The *proposal distribution* is the distribution that generate particles, defined as $p(x_t|\{u_t\},\{z_{t-1}\})$.

   We cannot sample from the target distribution directly, but we can sample from the proposal distribution indirectly. We can generate the particles from the proposal distribution, then the appropriately weighted particles converge to the target distribution.

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

   Particle deprivation means there are no particles in the places where the state has some probability of being. Mostly, this is caused by the number of particles is too small to cover all relevant regions with high likelihood. And particle deprivation also is the result of the variance in random sampling. An unlucky series of random numbers can wipe out all particles near the true state.

4. **Why do we resample instead of simply maintaining a weight for each particle always?**

   We do resample to change the distribution of the particles from the proposal distribution to the target distribution. In this way the particles tend to distribute in high posterior probability. On the contrary, if we simply maintain a weight for each particle always, many particles will stay in low posterior probability.

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

   If the distribution is multimodal, (two or more peaks), then it is often better to do a separate resample of particles of each peak. If we use the average of the particle set, the result will be a bad representation.

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

   We can fit a Gaussian to the mean and variance of the particle set or create bins and count how many particles are in each bin to form a histogram. We also can place a Gaussian kernel around each particle.

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

   Sampling variance is the difference between the particle distribution and the target distribution. The sampling variance is amplified through repetitive resampling. The first remedy is reducing the frequency at which resampling takes place. The second remedy is using low variance sampling.

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.**

   There are some relations between the pose uncertainty and the number of particles. If the uncertainty of the posterior distribution of the robot localization is high, the region of the posterior distribution is wide, thus we need more number of particles to present the distribution.

PART II – Matlab Exercises

Warm up problem with Particle Filter

Question 1:

In the 2D state space, the equation (6) regards the angle θ as a constant value θ_0. In this way, the model is simpler, if there are no noise, this equation can tract the line accurately. However, the drawback is if there exists noise, the trace will be zigzag. On the contrary, in the 3D state space, the angle is not a constant value, the value of angle depends on the previous timestep value, which is affected by noise, thus the value exists variation. In this way the model can represent a more complex movement and the trace will be smoother.

Question 2:

The equation (9) can model a specific circular motion, and the trace only influenced by the noise. The limitation is we need to know the constant linear velocity v_0 and the constant angular velocity ω_0 in advance.

Question 3:

The constant part in the denominator is the integration of all distribution. The purpose of keeping the constant part in the denominator of (10) is to normalize the likelihood function.

Question 4:

In Multinomial re-sampling method, the numbers of random numbers we need to generate is *M*, which is the number of particles. And in the Systematic re-sampling method, we only need one random number.

Question 5:

For vanilla re-sampling method, the probability of a particle been selected in each selection equals to the weight ω, thus, the probability of a particle not been selected equals to $1 - \omega$. And in Vanilla re-sampling the selection repeats for *M* times. Thus, the probability of a particle not been selected is $(1 - \omega)^M$. Hence, the probability of a particle survive in vanilla re-sampling is $1 - (1 - \omega)^M$.

For systematic re-sampling method, if the weight is $\omega = 1/M + \epsilon$, the particle will always be selected, because the cumulative distribution function of this particle must can satisfy $i = \min j \colon \mathrm{CDF}(j) \geq r_0 + (m-1)/M$. So, this particle can survive without doubt. If the weight of a particle is $0 \leq \omega \leq 1/M$, then the probability of a particle can survive is proportional to the weight of this particle, thus, the probability of this particle can survive is $\omega \cdot M$.

Question 6:

The variable Sigma_Q models the measurement noise, and the variable Sigma_R models the process the process noise.

Question 7:

If we do not perform the diffusion step (set the process noise to 0), all particles survived from previous time step will have the same movement. After re-sampling for many times, all particles will converge to the same state space which has the highest weight at first.

Question 8:

If we do not use re-sampling, the initial uniform distribution will not be changed. All particles changed through prediction depend on the model and process noise. The weight also updated but because we do not use re-sampling, so the distribution is still uniform.

Question 9:

If we increase the standard deviations of the observation noise by setting the Sigma_Q to a high value 10000, the measurement's uncertainty is high, then the particles can converge quickly but the uncertainty around the ground truth also is high.

If we decrease the standard deviations of the observation noise by setting the Sigma_Q to a low value 0.0001, the measurement is very accurate, then the particles cannot converge because many measurements are regarded as outliers.

Question 10:

If we increase the standard deviations of the process noise by setting the Sigma_R to a high value 10000, then the particles will converge to the true state quickly, but the variance is high, these particles will spread around the true state.

If we decrease the standard deviations of the process noise by setting the Sigma_R to a low value 0.0001, the diffusion is small, and the distribution of particle will be tight, then these particles converge to the true state slowly.

Question 11:

The choice of the process noise model depends on the choice of the motion model. If the motion model is accurate enough, then we can use low process noise. On the contrary, if the motion model is not precise, we need to increase the process noise to compensate the error between the true motion.

Question 12:

If the motion model is accurate enough, then the results will be accurate, and we can use fewer particles. On the contrary, if the motion is lake of precision, the results also will in low accuracy, and the process noise will be high, we need more particles to cover the reasonable states around the true state.

Question 13:

In the third type of measurements, there exists approximately 50% outliers, when there exists large bias between the measurement and prediction, the likelihood will be very small. Thus, we can set a threshold, when we detect the likelihood is smaller than the threshold, we can regard this measurement as an outlier.

Question 14:

Using 1000 particles and the second type of measurements of the object moving on the circle to model a fixed, a linear and a circular motion. The second type of measurements has white gaussian noise with a big deviation. The best parameters and results are shown below:

Fixed:

$$Q = \begin{bmatrix} 400 & 0 \\ 0 & 400 \end{bmatrix},\ R = \begin{bmatrix} 30 & 0 & 0 \\ 0 & 30 & 0 \\ 0 & 0 & 0.01 \end{bmatrix},\ Estimation\ Error = 11.0 \pm 5.5$$

Linear:

$$Q = \begin{bmatrix} 300 & 0 \\ 0 & 300 \end{bmatrix},\ R = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0.01 \end{bmatrix},\ Estimation\ Error = 7.4 \pm 3.8$$

Circular:

$$Q = \begin{bmatrix} 400 & 0 \\ 0 & 400 \end{bmatrix},\ R = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0.01 \end{bmatrix},\ Estimation\ Error = 7.0 \pm 3.5$$

According to the results, the linear model and circular model can get similar result and use low values of process noise. The fixed model is sensitive to the choice of the parameter, and we need to increase the process noise to compensate the error between the true motion. As for the parameters, these models are more sensitive to the value of R.

Main problem: Monte Carlo Localization

Question 15:

The parameters affect the outlier detection approach is the threshold on the average likelihood λ_Ψ and the measurement noise model Q. if we model a very weak measurement noise $|Q| \to 0$, the distribution of the likelihood will be narrow, then more measurements will be regarded as outliers. When set Q to 0, only if the measurements equal to previous true state can survive. Other measurements will be regarded as outliers.

Question 16:

If we do not detect outliers, then we use all measurements to calculate the weights. But some measurements are not reliable, which should be regarded as outliers and discarded. Since these unreliable measurements are taken into consideration, the weight will be unreliable, then the estimation of the true state will also be unreliable.

2.3 Data sets

Map_sym2 + so_sym2_nk

According to the notes and results, we know this map is a symmetric environment with 4 landmarks. According to the code below:

```
if ~isempty(start_pose)
    S = [repmat(start_pose,1,M); 1/M*ones(1,M)];
else
    S = [rand(1,M)*(bound(2) - bound(1)+2*part_bound) + bound(1)-part_bound;
         rand(1,M)*(bound(4) - bound(3)+2*part_bound) + bound(3)-part_bound;
         rand(1,M)*2*pi-pi;
         1/M*ones(1,M)];
end
```

We know the value of part_bound will influence the initial distribution of the particles. If the part_bound is higher, the variance of the initial distribution around the landmarks will be larger, then these particles are more possible to cover the area around landmarks. In order to track all valid hypotheses in this symmetric environment, we need to increase part_bound.

when we perform a tracking task, the results are good, because we know the initial position and pose in this symmetric environment, thus the particles can spread around the true state initially. Just as Figure 1, Figure 2 shown. As for the following tasks we need to compensate noises to make the filter work correctly. So we use stronger noises.

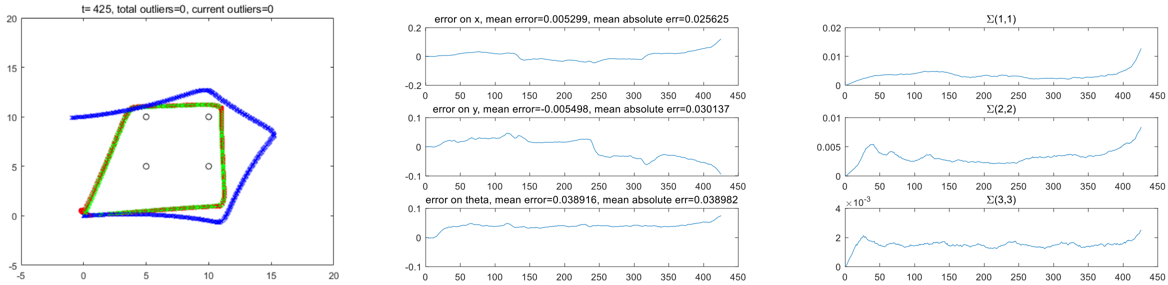


Figure 1: tracking task result, errors and covariance (1000 particles)

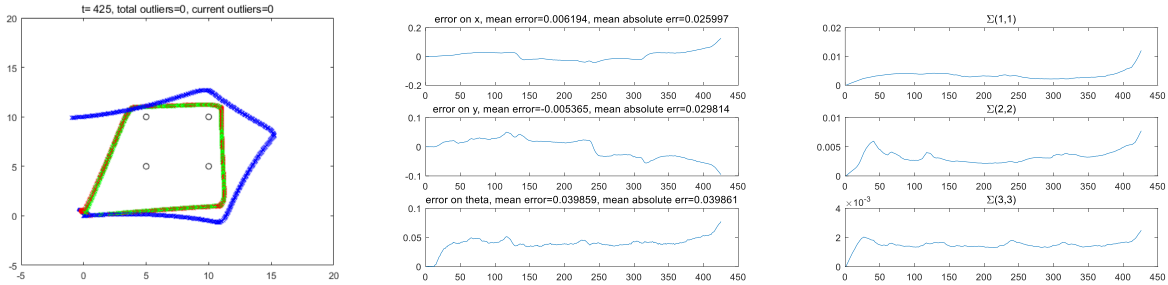


Figure 2: tracking task result, errors and covariance (10000 particles)

In the global localization problem, we can observe there are 4 valid hypotheses for these 4 landmarks. Starting with 1000 particles, we can observe some valid hypotheses will be discarded, because the small number of particles cannot cover all relevant regions with high likelihood, then will cause particle deprivation, which means there are no particles in the places where the state has some probability of being. When there are 10000 particles, the filter can keep all the hypotheses reliably, because more particles can cover the areas to prevent particle deprivation. Just as Figure 3 and Figure 4 shown.

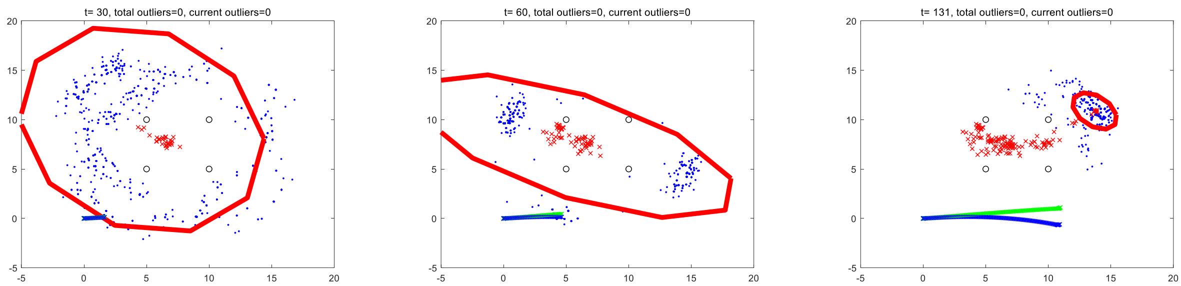


Figure 3: localization task of t = 30,60,131 (1000 particles)

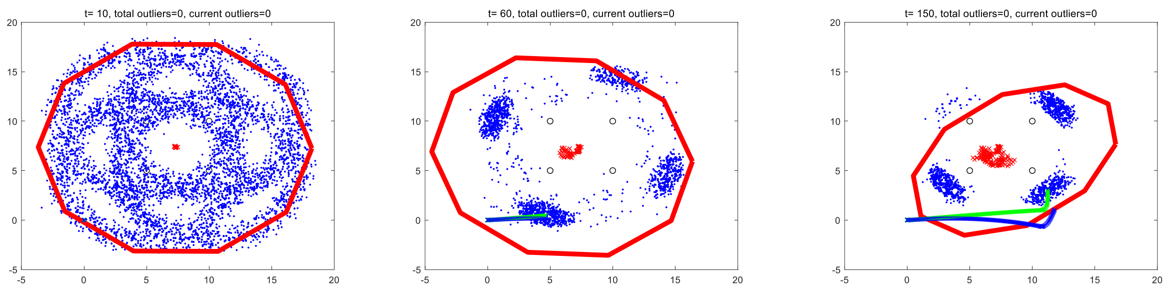
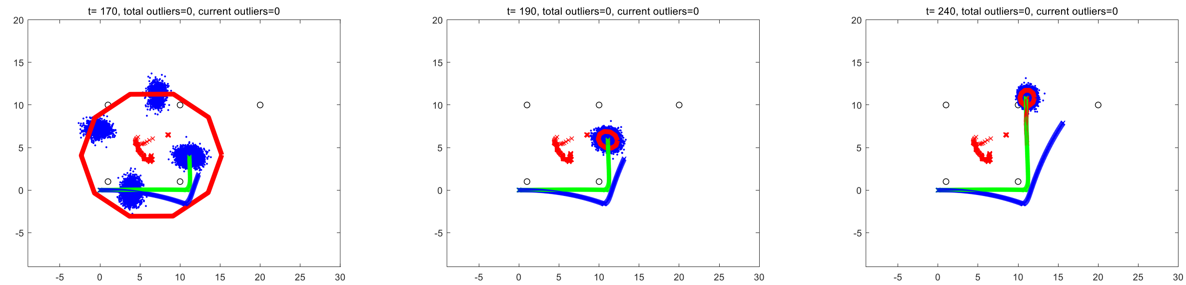


Figure 4: localization task of t = 10,60,150 (10000 particles)

When we try multinomial sampling, the particles will converge to one hypotheses faster than Systematic Re-sampling, because multinomial sampling choose particles with high weights. Thus, the ability to preserve multiple hypotheses of multinomial sampling is weak. If we use stronger measurement noises, the ability to preserve multiple hypotheses also improved.

Map_sym3 + so_sym3_nk

Now in this map, there are 5 nearly symmetric landmarks. Before the 180th time step, the robot cannot observe the nonsymmetric landmarks, so there are 4 hypotheses because the environment of these four landmarks is symmetric, just as Figure 5(a) shown. After 180 time-steps, the top right landmark has been observed, thus, there are only one hypothesis remains valid, just as Figure 5(b), 5(c) shown. And Figure 6 shows the complete localization result.



(a) (b) (c)

Figure 5: localization task of nonsymmetric map when t = 170, 190, 240 (10000 particles)

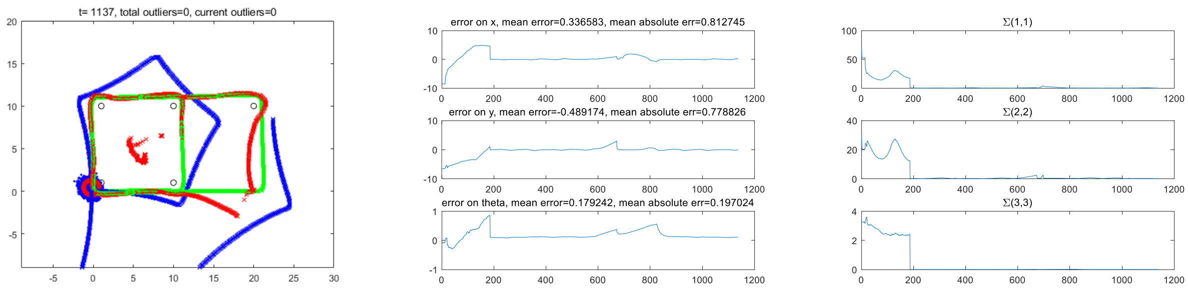


Figure 6: localization task of nonsymmetric map (10000 particles)