EL2320

Applied Estimation

Laboratory Report

Lab 1: Extended Kalman Filter

Jiang, Sifan
sifanj@kth.se

December 11, 2018

1 Preparatory Questions

Linear Kalman Filter:

1. What is the difference between a 'control' $\mathbf{u_t}$, a 'measurement' $\mathbf{z_t}$ and the state $\mathbf{x_t}$? Give examples of each?

   - $\mathbf{u_t}$ contains information about the change of the state in the environment. Control is part of dynamic model between $\mathbf{x_{t-1}}$ and $\mathbf{x_t}$. For example: velocity of a robot.
   - $\mathbf{z_t}$ measures data through sensor measurements. The data is about the state of its environment. Measurement is modeled as depending on the state at time t with function $\mathbf{h}(\mathbf{x_t})$. For example, range scans.
   - $\mathbf{x_t}$ is used to describe environments which in the case of robot is the collection of all aspects of the robot and its environment that can impact the future. $\mathbf{x_t}$ is modeled as depending on both $\mathbf{u_t}$ and $\mathbf{x_{t-1}}$ as $\mathbf{g}(\mathbf{u_t}, \mathbf{x_{t-1}})$. For example: the robot pose.

   $\mathbf{u_t}$ is used in predict while $\mathbf{z_t}$ is used in update.

2. Can the uncertainty in the belief increase during an update? Why (or not)?

   The uncertainty in the belief during a Kalman filter update step is given by:

   $$\Sigma_t = (I - K_t C_t)\bar{\Sigma}_t = \bar{\Sigma}_t - K_t C_t \bar{\Sigma}_t \tag{1}$$

   where the Kalman gain K_t is:

   $$K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1} \tag{2}$$

   If plugging Equation 2 into Equation 1 we have:

   $$\begin{aligned} \Sigma_t &= \bar{\Sigma}_t - \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1} C_t \bar{\Sigma}_t && (3) \\ &= \bar{\Sigma}_t - M M^T && (4) \end{aligned}$$

   where $\mathbf{x}^T M M^T \mathbf{x} = \mathbf{y}^T \mathbf{y} \geq 0$ for any $\mathbf{x}$. Alternatively, applying Sherman-Morrison formula:

   $$\Sigma_t = (\bar{\Sigma}_t^{-1} + C_t^T Q_t^{-1} C_t)^{-1} \tag{5}$$

   where $C_t^T Q_t^{-1} C_t$ is positive semi-definite. Hence, the uncertainty in the belief during an update cannot increase.

3. During update what is it that decides the weighing between measurements and belief?

   $$\begin{aligned} \mu_t &= \bar{\mu}_t + K_t(z_t - C_t \bar{\mu}_t) = (I - K_t C_t)\bar{\mu}_t + K_t z_t && (6) \\ &= (I - W)\bar{\mu}_t + W \mu_z && (7) \end{aligned}$$

   where $W = \Sigma_t C_t^T Q_t^{-1} C_t$ and $C_t \mu_z = (z_t - \bar{z}_t + C_t \bar{\mu}_t)$. Hence, the relation between Σ_t (uncertainty) and Q_t (measurement error) after the update is done decides the weighting between the measurements and the belief.

4. What would be the result of using a too large a covariance (Q matrix) for the measurement model?

   If Q is large, the Kalman Gain will be small. Estimates would be pessimistic (conservative) and convergence would be slower.

5. What would give the measurements an increased effect on the updated state estimate?

   To have an increased effect for the measurements on the updated state estimate, the Kalman gain should be large, and since $\mu_t = \bar{\mu}_t + K_t(z_t - C_t\bar{\mu}_t)$ the uncertainty Q regarding the measurements should be small.

6. What happens to the belief uncertainty during prediction? How can you show that?

   The belief uncertainty during prediction is given by equation 8:

$$\bar{\Sigma}_t = R_t + A_t\Sigma_{t-1}A_t^T \tag{8}$$

   If $A_t \geq I$ then $A_t\Sigma_{t-1}A_t^T \geq \Sigma_{t-1}$. Also noise R_t is added. Hence, in general, uncertainty $\bar{\Sigma}_t$ increases in most real systems but it could still decreases depending on A_t and R_t.

7. How can we say that the Kalman filter is the optimal and minimum least square error estimator in the case of independent Gaussian noise and Gaussian priori distribution? (Just describe the reasoning not a formal proof.)

   The Kalman filter gives the true posterior distribution for a linear Gaussian system. So, suppose μ_{better} for the state has lower expected square error than the Gaussian mean μ.

$$\begin{aligned}
& \int_{-\infty}^{\infty} (x - \mu_{better})^2 G(x, \mu, \Sigma)dx && (9)\\
\Leftrightarrow \quad & \int_{-\infty}^{\infty} (x^2 - 2\mu_{better}x + \mu_{better}^2) G(x, \mu, \Sigma)dx && (10)\\
\Leftrightarrow \quad & \Sigma + \mu^T\mu - \mu^T\mu_{better} - \mu_{better}^T\mu + \mu_{better}^T\mu_{better} && (11)
\end{aligned}$$

   Minimizing the equation by differentiating and we have:

$$\begin{aligned}
-2\mu + 2\mu_{better} &= 0 && (12)\\
\mu_{better} &= \mu && (13)
\end{aligned}$$

   So, the Kalman filter is the optimal and minimum least square error estimator.

8. In the case of Gaussian white noise and Gaussian priori distribution, is the Kalman Filter a MLE and/or MAP estimator?

   If there is no a priori knowledge about the distribution and update the state, then the mean of the Kalman Filter can be an MLE of the Gaussian posterior. If there is a priori knowledge about the distribution, then a Kalman Filter can be an MAP if the distribution is Gaussian.

Extended Kalman Filter:

9. How does the extended Kalman filter relate to the Kalman filter?

   The extended Kalman filter applies the Kalman filter to the linearized nonlinear system. Furthermore, $A_t x_{t-1} + B_t u_t$ is replaced by $g(u_t, x_{t-1})$, A_t by G_t and C_t by H_t.

10. Is the EKF guaranteed to converge to a consistent solution?

    No. Consistent solution depends on the the local nonlinearity of the function g.

11. If our filter seems to diverge often can we change any parameter to try and reduce this?

    Yes. If divergence occurs on update, we can change modeled uncertainties Q and R and increase the relative size of the measurement covariance Q. If the divergence is due to bad data association, we can change the matching threshold.

Localization:

12. If a robot is completely unsure of its location and measures the range r to a know landmark with Gaussian noise what does its posterior belief of its location $p(x, y, \theta|r)$ look like? So a formula is not needed but describe it at least.

    Uniform distribution would be given over the heading θ between $-\pi$ and π. The position would be a sort of concentric rings. So a Gaussian on ρ in radial coordinate with uniform distribution on the angle ϕ. So, $x = \rho\cos\phi$, $y = \rho\sin\phi$, and $e^{-\frac{(\rho-r)^2}{2\sigma_r^2}}$.

13. If the above measurement also included a bearing how would the posterior look?

    The heading and angle around the concentric rings change to Gaussian distribution with a completely correlated covariance, $e^{-\left[\frac{(\rho-r)^2}{2\sigma_r^2}+\frac{(b-\phi+\theta)^2}{2\phi_b^2}\right]}$

14. If the robot moves with relatively good motion estimation (prediction error is small) but a large initial uncertainty in heading θ how will the posterior look after traveling a long distance without seeing any features?

    It will look like an arc and the heading θ will be correlated with position along the arc.

15. If the above robot then sees a point feature and measures range and bearing to it how might the EKF update go wrong?

    The Jacobian will produce an update direction that is a straight line which can not move the estimate along the arc. The Gaussian will not be able to represent the arc shape and thus the update might diverge.

2 MATLAB Exercises

2.1 Warm up problem with Standard Kalman Filter

Consider a car example. The acceleration of the car is controlled by the throttle/break. The system is described by Equation 14 $\sim$ 20.

$$x_k = \begin{bmatrix} p_k \\ v_k \end{bmatrix} \tag{14}$$

$$u_k = a_0 \tag{15}$$

$$A = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \tag{16}$$

$$B = \begin{bmatrix} 0 \\ \Delta t \end{bmatrix} \tag{17}$$

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{18}$$

$$x_{k+1} = Ax_k + Bu_k + \varepsilon_k \tag{19}$$

$$z_k = Cx_k + \delta_k \tag{20}$$

where p_k, v_k and are the position and velocity of the car in the $k^{\mathbf{th}}$ time step, a_0 is the constant acceleration of the car and ε_k and δ_k are white Gaussian noises.

- **Question 1:** What are the dimensions of ε_k and δ_k? What parameters do you need to define in order to uniquely characterize a white Gaussian?

From Equation 14 to 17 and Equation 19, we have:

$$x_{k+1} = \begin{bmatrix} p_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_k \\ v_k \end{bmatrix} + \begin{bmatrix} 0 \\ \Delta t \end{bmatrix} a_0 + \varepsilon_k \tag{21}$$

$$= \begin{bmatrix} p_k + v_k \Delta t \\ v_k \end{bmatrix} + \begin{bmatrix} 0 \\ a_0 \Delta t \end{bmatrix} + \varepsilon_k \tag{22}$$

$$= \begin{bmatrix} p_k + v_k \Delta t \\ v_k + a_0 \Delta t \end{bmatrix} + \begin{bmatrix} \varepsilon_p^k \\ \varepsilon_v^k \end{bmatrix} \tag{23}$$

So, $\varepsilon_k = \begin{bmatrix} \varepsilon_p^k & \varepsilon_v^k \end{bmatrix}^{\mathbb{T}}$ is of dimension 2. Similarly, from Equation 18 and Equation 20, we have:

$$z_k = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} p_k \\ v_k \end{bmatrix} + \delta_k \tag{24}$$

$$= p_k + \delta_k \tag{25}$$

So, δ_k is of dimension 1.

A Gaussian function is characterized by a mean value μ and a variance σ^2. A white Gaussian has $\mu = 0$. So, δ_k can be uniquely characterized by a variance σ^2. Since ε is of dimension 2, it is characterized by a covariance Σ, which is a diagonal matrix.

- **Question 2:** Make a table showing the roles/usages of the variables(x, xhat, P, G, D, Q, R, wStdP, wStdV, vStd, u, PP). To do this one must go beyond simply reading the comments in the code to seeing how the variable is used. (Hint: some of these are our estimation model and some are for simulating the car motion).

Table 1: Usages of the Variables.

Variable	Usage
x	The true state of the system.
$\hat{x}$	The estimate of the true state of the system by the Kalman Filter.
P	Estimate error covariance matrix.
G	Identity matrix for dimensionality consistency.
D	Identity matrix for dimensionality consistency. Scalar here.
Q	Measurement noise variance.
R	Process noise covariance matrix.
$wStdP$	The simulated standard deviation of the noise in position.
$wStdV$	The simulated standard deviation of the noise in velocity.
$vStd$	The simulated standard deviation of the noise in position estimation.
u	Control signal which is the acceleration.
PP	Estimate error covariances over time.

In common definition, Q is used to represent the process noise covariance matrix and R is used to represent the measurement noise variance. However, in this example, the parameters are defined oppositely.

- **Question 3:** Please answer this question with one paragraph of text that summarizes broadly what you learn/deduce from changing the parameters in the code as described below. Chose two illustrative sets of plots to include as demonstration. What do you expect if you increase/decrease the covariance matrix of the modeled measurement noise 100 times (one change in the default parameters each time) for the same underlying system? Characterize your expectations. Confirm your expectations using the code (save the corresponding figures so you can analyze them in your report). Do the same analysis for the case of increasing/decreasing both parameters by the same factor at the same time. (Hint: It is the mean and covariance behavior over time that we are asking about.)

  The results with default parameters are illustrated in Figure 1, 2, and 3. The standard deviation of error in position (second half) is 0.075916 m, and the standard deviation of error in velocity (second half) is 0.263859 m/s.

  When increasing the covariance matrix of the modeled measurement noise Q, the error in estimating position and velocity would increase and the Kalman gain would decrease, thus the system would take more time to converge. The results of increasing 100 times are illustrated in Figure 4, 5, and 6. The standard deviation of error in position (second half) is 0.073729 m, and the standard deviation of error in velocity (second half) is 0.167867 m/s.

  When increasing the covariance matrix of the modeled process noise R, the measurement would have larger weight over determining the state. So, the Kalman gain would increase and so would the estimation error do. However, the convergence time would not change but the estimation of position and velocity would floating badly. The results of increasing 100 times are illustrated in Figure 7, 8, and 9. The standard deviation of error in position (second half) is 0.073415 m, and the standard deviation of error in velocity (second half) is 0.459833 m/s.

  When increasing the covariance matrix of the both noises Q and R, The modeled covariance would increase while the Kalman gain would decrease. Also, it would take the system more time to converge. The results of increasing 100 times are illustrated in Figure 10, 11, and 12. The standard deviation of error in position (second half) is 0.059333 m, and the standard deviation of error in velocity (second half) is 0.204902 m/s.

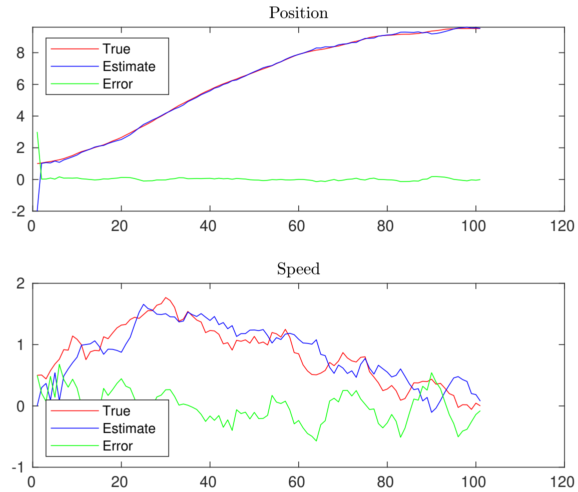

Figure 1: Default parameters: $wStdP = 0.01$, $wStdV = 0.1$, and $vStd = 0.1$.

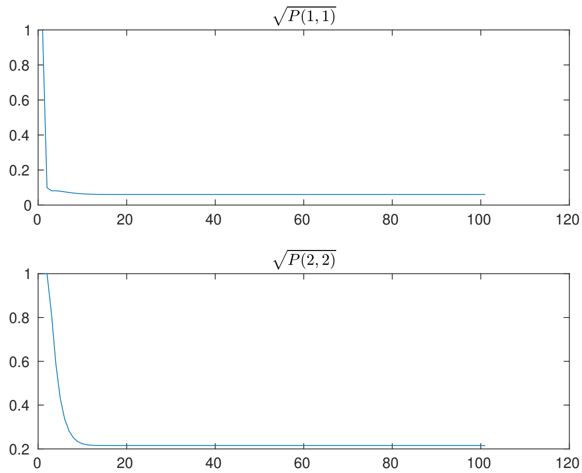

Figure 2: Default parameters: $wStdP = 0.01$, $wStdV = 0.1$, and $vStd = 0.1$.

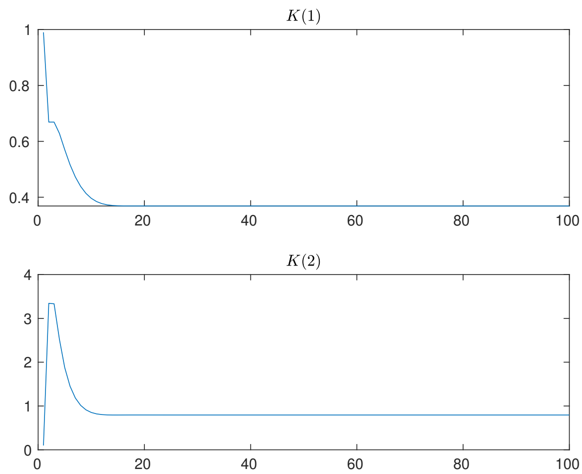

Figure 3: Default parameters: $wStdP = 0.01$, $wStdV = 0.1$, and $vStd = 0.1$.

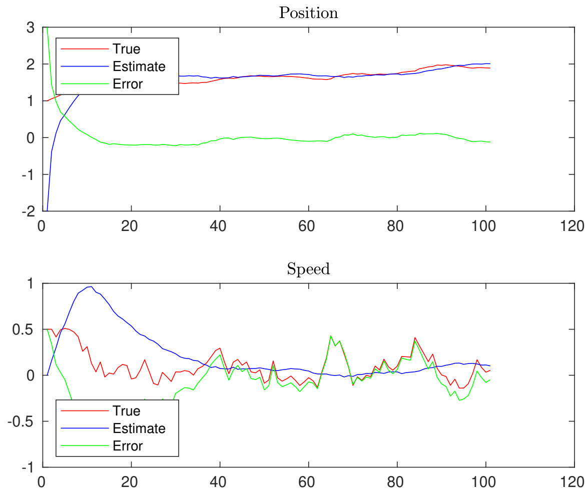

Figure 4: Increase measurement noise: $wStdP = 0.01$, $wStdV = 0.1$, and $vStd = 1$.

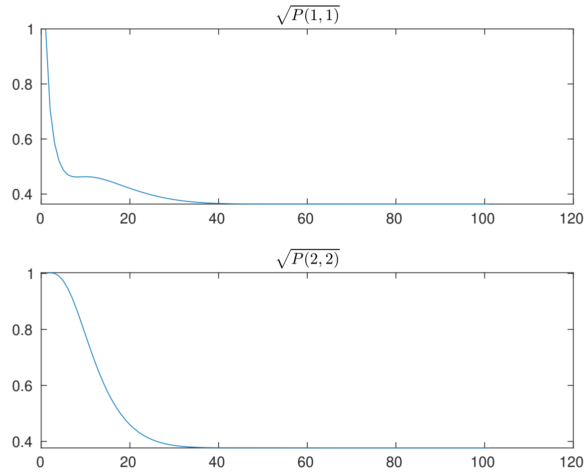

Figure 5: Increase measurement noise: $wStdP = 0.01$, $wStdV = 0.1$, and $vStd = 1$.

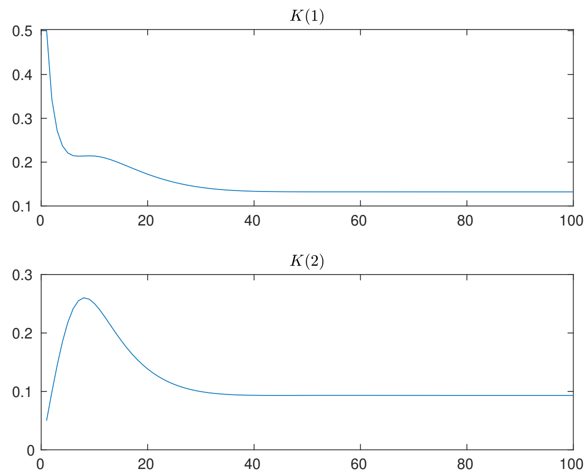

Figure 6: Increase measurement noise: $wStdP = 0.01$, $wStdV = 0.1$, and $vStd = 1$.

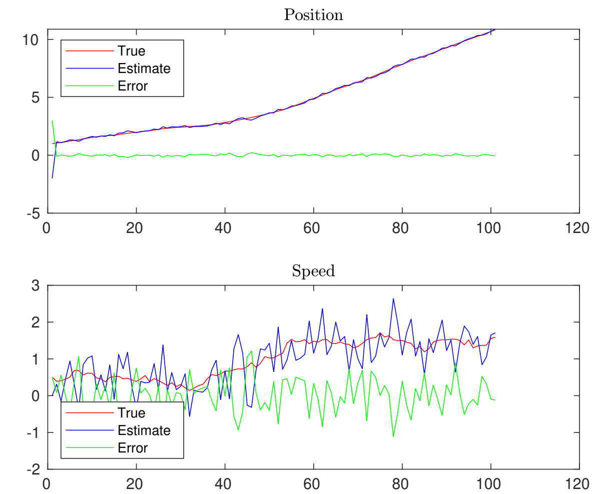

Figure 7: Increase process noise: $wStdP = 0.1$, $wStdV = 1$, and $vStd = 0.1$.

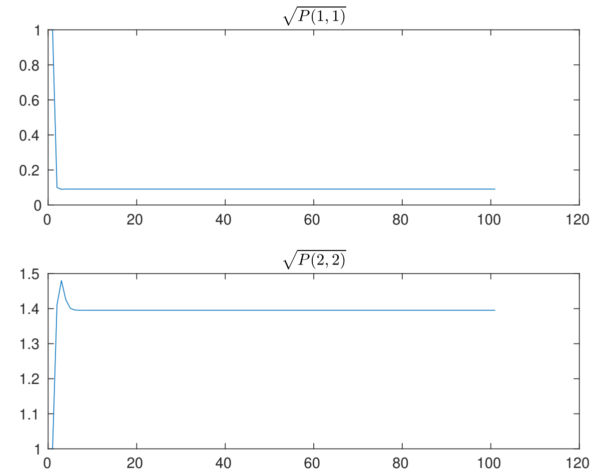

Figure 8: Increase process noise: $wStdP = 0.1$, $wStdV = 1$, and $vStd = 0.1$.

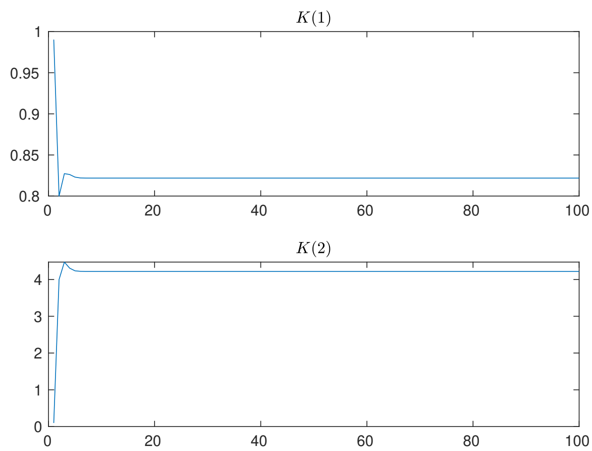

Figure 9: Increase process noise: $wStdP = 0.1$, $wStdV = 1$, and $vStd = 0.1$.

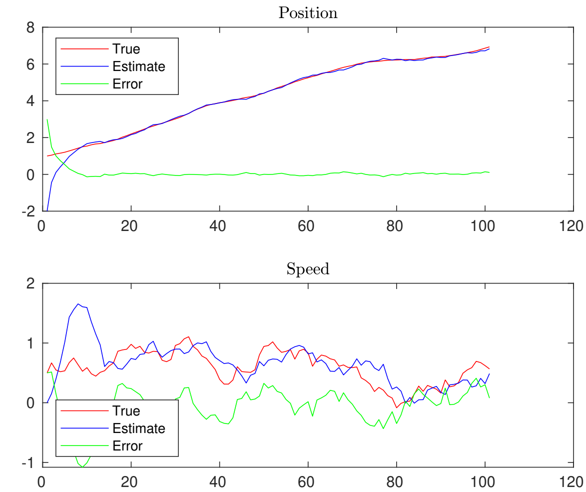

Figure 10: Increase both noises: $wStdP = 0.1$, $wStdV = 1$, and $vStd = 1$.

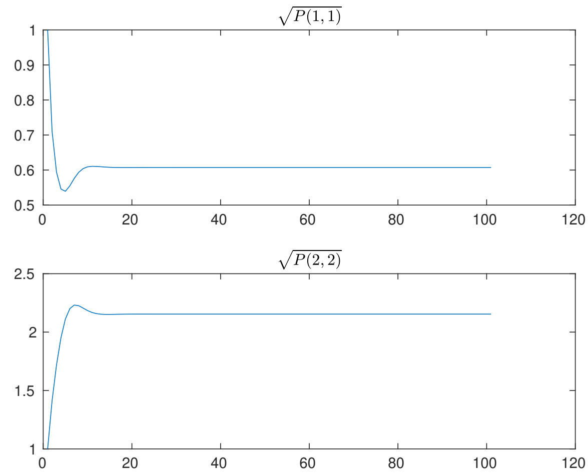

Figure 11: Increase both noises: $wStdP = 0.1$, $wStdV = 1$, and $vStd = 1$.

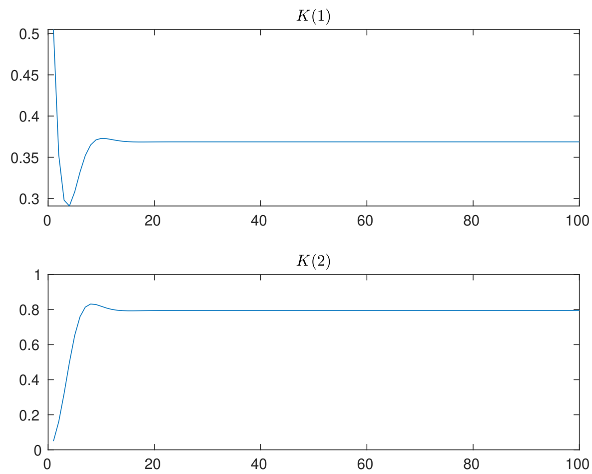

Figure 12: Increase both noises: $wStdP = 0.1$, $wStdV = 1$, and $vStd = 1$.

- **Question 4:** How do the initial values for P and $xhat$ affect the rate of convergence and the error of the estimates (try both much bigger and much smaller)?

  If the initial value of P is large, the uncertainty of the true state would be large and also the Kalman gain would be large. Meaning the measurement would have large weight. So the convergence time would not be affected largely. Also, the error of the estimates would not change much. The result is illustrated in Figure 13.

  If the initial value of P is small, the system would have more confidence in the prediction, so the Kalman gain would increase slower, thus the convergence time would be longer. Also, the error of the estimates would not be changed much. The result is illustrated in Figure 14.

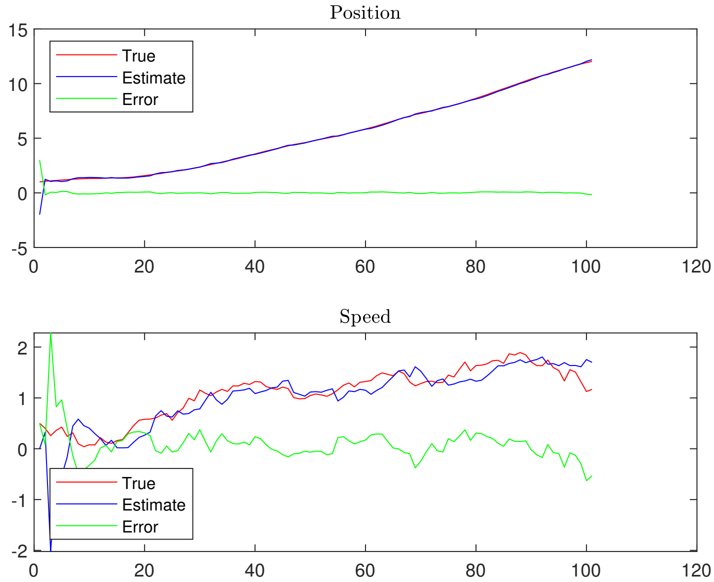

Figure 13: Increase P to the 100 times of the original values.

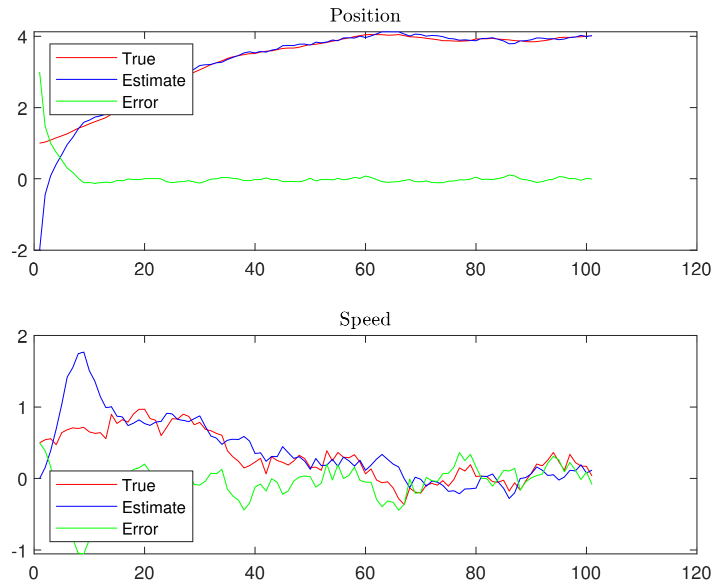

Figure 14: Decease P to the 0.01 times of the original values.

The initial values for $xhat$ would not affect the rate of convergence in a obvious way. Also, it would not obviously affect the error of the estimates. The results are illustrated in Figure 15 and 16.

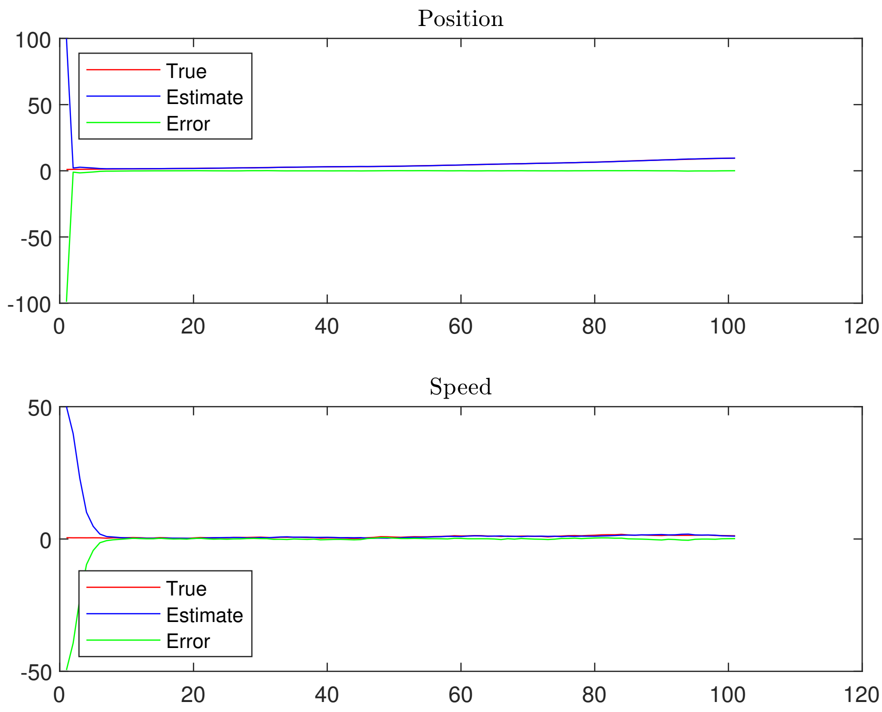

Figure 15: Increase $xhat$ to $[100; 50]$.

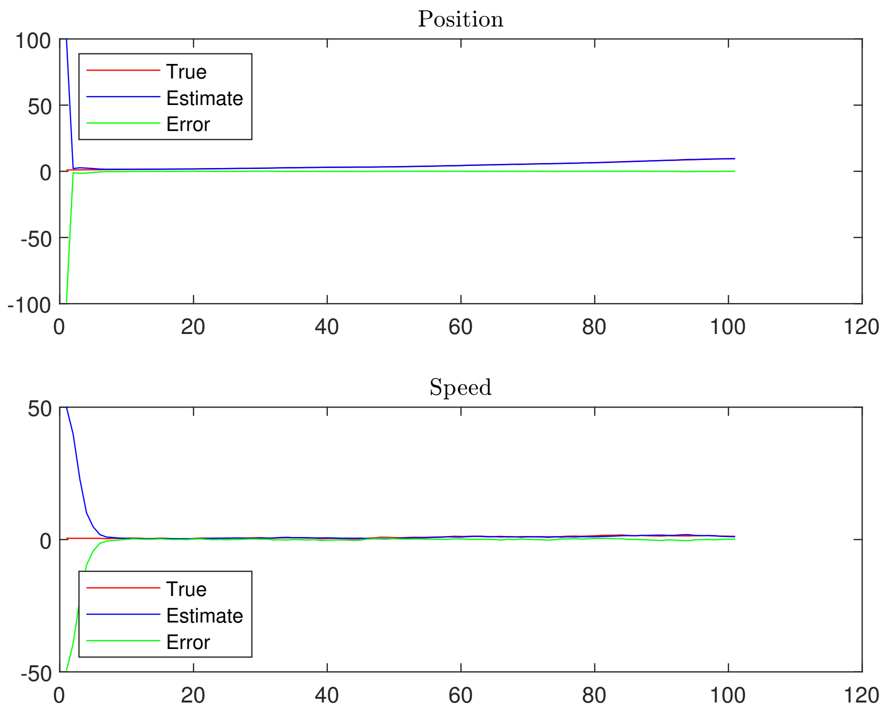

Figure 16: Decrease $xhat$ to $[1.5; 0.75]$ which is close to the true position and velocity.

2.2 Main problem: EKF Localization

Using a first order Markov assumption and Bayesian update:

$$\begin{cases} p(\mathbf{x_t}|\mathbf{u_{1:t}}, \mathbf{z_{1:t}}, \mathbf{\bar{x}_0}, M) & = \eta p(\mathbf{z_t}|\mathbf{x_t}, M) \int p(\mathbf{x_t}|\mathbf{u_t}, \mathbf{x_{t-1}}) p(\mathbf{x_{t-1}}|\mathbf{z_{1:t-1}}, \mathbf{u_{1:t-1}}, \mathbf{\bar{x}_0}, M) \mathbf{dx_{t-1}} \\ p(\mathbf{x_0}|\mathbf{\bar{x}_0}) & = \delta(\mathbf{x_0} - \mathbf{\bar{x}_0}) \end{cases} \tag{26}$$

or equivalently

$$\begin{cases} bel(\mathbf{x_t}) & = p(\mathbf{x_t}|\mathbf{u_{1:t}}, \mathbf{z_{1:t}}, \mathbf{\bar{x}_0}, M) = \eta p(\mathbf{z_t}|\mathbf{x_t}, M) \overline{bel}(\mathbf{x_t}) \\ \overline{bel}(\mathbf{x_t}) & = p(\mathbf{x_t}|\mathbf{u_{1:t}}, \mathbf{z_{1:t-1}}, \mathbf{\bar{x}_0}, M) = \int p(\mathbf{x_t}|\mathbf{u_t}, \mathbf{x_{t-1}}) bel(\mathbf{x_{t-1}}) \mathbf{dx_{t-1}} \\ bel(\mathbf{x_0}) & = p(\mathbf{x_0}|\mathbf{\bar{x}_0}) = \delta(\mathbf{x_0} - \mathbf{\bar{x}_0}) \end{cases} \tag{27}$$

- **Question 5:** Which parts of (26) and (27) are responsible for prediction and update steps?

  In Equation (27) as the example, the first line of the equation is responsible for the update step. The second line of the equation is responsible for the prediction steps. While in Equation (26), the first line of the equation is responsible for both of the prediction and update steps.

2.2.1 Maximum Likelihood Data Association

- **Question 6:** In the maximum likelihood data association, we assumed that the measurements are independent of each other. Is this a valid assumption? Explain why.

  It is valid to assume that the measurements are independent of each other. The reason is that, the measurement model for each landmark at each time stamp only depends on the landmark, and the state at this time stamp, thus the measurement noise is white Gaussian noise and the covariance matrix of the measurement noise is diagonal meaning all the measurements are uncorrelated with each other.

2.2.2 Outlier Detection

It is possible to define a threshold on the Mahalanobis distance between the measurement(z_t^i) and the most likely association which is given by:

$$D_M = (\bar{\boldsymbol{\nu}}_t^i)^T (\bar{S}_{t,i})^{-1} (\bar{\boldsymbol{\nu}}_t^i) \tag{28}$$

D_M follows the chi square (X_n^{-2}) cumulative distribution with n degree of freedom and therefore threshold be defined basing on a probability. ν has two degrees of freedom in our case, the threshold (λ_M) is given by:

$$\lambda_M = X_2^{-2}(\delta_M) \tag{29}$$

- **Question 7:** What are the bounds for δ_M in (29)? How does the choice of δ_M affect the outlier rejection process? What value do you suggest for λ_M when we have reliable measurements all arising from features in our map, that is all our measurements come from features on our map? What about a scenario with unreliable measurements with many arising from so called clutter or spurious measurements?

  The bounds for δ_M in (29) are $\delta_M \in [0, 1]$. Since the inverse chi square $\mathcal{X}_n^{-2}$ cumulative distribution is monotonically increasing, the increase in δ_M would result in the increase of the threshold λ_M. The larger the threshold λ_M, the more confidence would the system have on the measurement and the less measurements would be rejected as outliers. When we have reliable measurements come from features on our map, the covariance matrix for measurement noise would be small and few measurement should be viewed as outliers, so the threshold λ_M should be large. On the contrast, when the measurements are unreliable, the threshold λ_M should be small.

2.2.3 Update

The most simple type of update namely sequential update, performs one update for each observation z_t^i in the t^{th} time step (Algorithm 1).

Algorithm 1 Sequential update algorithm for the i^{th} observation

for all Observations i in z_t **do**

$$
\begin{aligned}
&... && \texttt{\{Compute the } i^{th} \texttt{ association using } \bar{\boldsymbol{\mu}}_t\texttt{\}} \\
K_{t,i} &= \bar{\Sigma}_t(\bar{H}_{t,i})^T(\bar{S}_{t,i})^{-1} \\
\bar{\boldsymbol{\mu}}_t &= \bar{\boldsymbol{\mu}}_t + K_{t,i}\bar{\boldsymbol{\nu}}_t^i \\
\bar{\Sigma}_t &= (I - K_{t,i}\bar{H}_{t,i})\bar{\Sigma}_t
\end{aligned}
$$

end for

$$
\begin{aligned}
\boldsymbol{\mu}_t &= \bar{\boldsymbol{\mu}}_t \\
\Sigma_t &= \bar{\Sigma}_t
\end{aligned}
$$

- **Question 8:** Can you think of some down-sides of the sequential update approach (Algorithm 1)? Hint: How does the first noisy measurements affect the intermediate results?

  Since the first measurements are noisy and would result in non-zero innovations in the data association step. The non-zero (or not close to zero) innovation would cause the shift of the estimated mean. Also the noisy measurements would cause the covariance matrix $\bar{\Sigma}$ being reduced much more, thus resulting in smaller $S_{t,j}$ and higher Kalman gain in future time stamps. Then the Mahalanobis distance would increase and more reasonable measurements would be rejected as outliers.

2.2.4 Batch Update

An alternative to the sequential update is the batch update algorithm. Algorithm 2 shows the complete EKF localization problem with the batch update.

- **Question 9:** How can you modify Algorithm 2 to avoid redundant re-computations?

  In the Algorithm 2, $\hat{\mathbf{z}}_{t,j}$, $H_{t,j}$, and $S_{t,j}$ do not depend on the observation z_t while still appear in the loop of z_t which is causing redundant re-computations and can be avoided by taking them out of the z_t loop so that these three parameters would be calculated only in the M loop.

- **Question 10:** What are the dimensions of $\bar{\boldsymbol{\nu}}_t$ and $\bar{H}_t$ in Algorithm 2? What were the corresponding dimensions in the sequential update algorithm? What does this tell you?

  The dimension of $\bar{\boldsymbol{\nu}}^{\boldsymbol{i}}{}_t$ is of 2×1, so the dimension of $\bar{\boldsymbol{\nu}}_t$ is of $2n \times 1$, where n is the number of inliers.

  The dimension of $\bar{H}_{t,i}$ is of 2×3, so the dimension of $\bar{H}_t$ is of $2n \times 3$, where n is the number of inliers.

Algorithm 2 EKF Localization with Batch update for the i^{th} time step

for all Observations i in z_t **do**

  for all Landmarks j in M **do**

$$\begin{aligned}
\hat{\mathbf{z}}_{t,j} &= \mathbf{h}(\bar{\boldsymbol{\mu}}_t, M, j) \\
H_{t,j} &= H(\bar{\boldsymbol{\mu}}_t, M, j, \hat{\mathbf{z}}_{t,j}) \\
S_{t,j} &= H_{t,j}\bar{\Sigma}_t(H_{t,j})^T + Q \\
\boldsymbol{\nu}_t^{i,j} &= \mathbf{z_{t,i}} - \hat{\mathbf{z}}_{\mathbf{t,j}} \\
D_t^{i,j} &= (\boldsymbol{\nu}_t^{i,j})^T(S_{t,j})^{-1}\boldsymbol{\nu}_t^{i,j} \\
\psi_t^{i,j} &= \det(2\pi S_{t,j})^{-\frac{1}{2}} exp[-\tfrac{1}{2}D_t^{i,j}]
\end{aligned}$$

  end for

$$\begin{aligned}
\hat{c}_t^i &= \arg\max_j \ \psi_t^{i,j} \\
\hat{o}_t^i &= D^{i,c_t^i} \geq \lambda_M \\
\bar{\boldsymbol{\nu}}_t^i &= \boldsymbol{\nu}_t^{i,\hat{c}_t^i} \\
\bar{H}_{t,i} &= H_{t,\hat{c}_t^i}
\end{aligned}$$

end for

{For inlier indices $1, ..., n$}

$$\begin{aligned}
\bar{\boldsymbol{\nu}}_t &= \left[\begin{array}{cccc} (\bar{\boldsymbol{\nu}}_t^1)^T & (\bar{\boldsymbol{\nu}}_t^2)^T & ... & (\bar{\boldsymbol{\nu}}_t^n)^T \end{array} \right]^T \\
\bar{H}_t &= \left[\begin{array}{cccc} (\bar{H}_{t,1})^T & (\bar{H}_{t,2})^T & ... & (\bar{H}_{t,n})^T \end{array} \right]^T \\
\bar{Q}_t &= \begin{bmatrix} Q & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & ... & 0 \\ 0 & 0 & 0 & Q \end{bmatrix}_{(2n\times 2n)} \\
K_t &= \bar{\Sigma}_t(\bar{H}_t)^T(\bar{H}_t\bar{\Sigma}_t(\bar{H}_t)^T + \bar{Q}_t)^{-1} \\
\boldsymbol{\mu}_t &= \bar{\boldsymbol{\mu}}_t + K_t\bar{\boldsymbol{\nu}}_t \\
\Sigma_t &= (I - K_t\bar{H}_t)\bar{\Sigma}_t
\end{aligned}$$

2.3 Data sets

2.3.1 map_o3.txt + so_o3_ie.txt

In this case, the covariance matrix of the process and measurement noises are:

$$R = \begin{bmatrix} 0.01^2 & 0 & 0 \\ 0 & 0.01^2 & 0 \\ 0 & 0 & (\frac{2\pi}{360})^2 \end{bmatrix}$$
$$Q = \begin{bmatrix} 0.01^2 & 0 \\ 0 & (\frac{2\pi}{360})^2 \end{bmatrix}$$

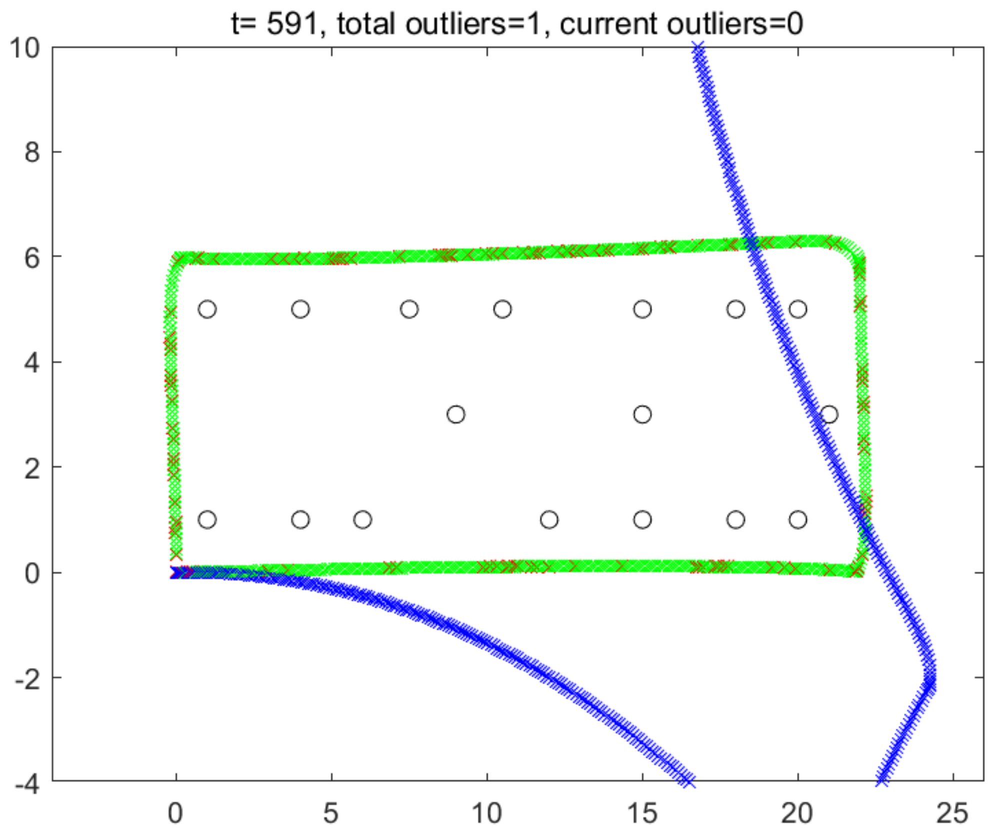

Figure 17: Path in case 1.

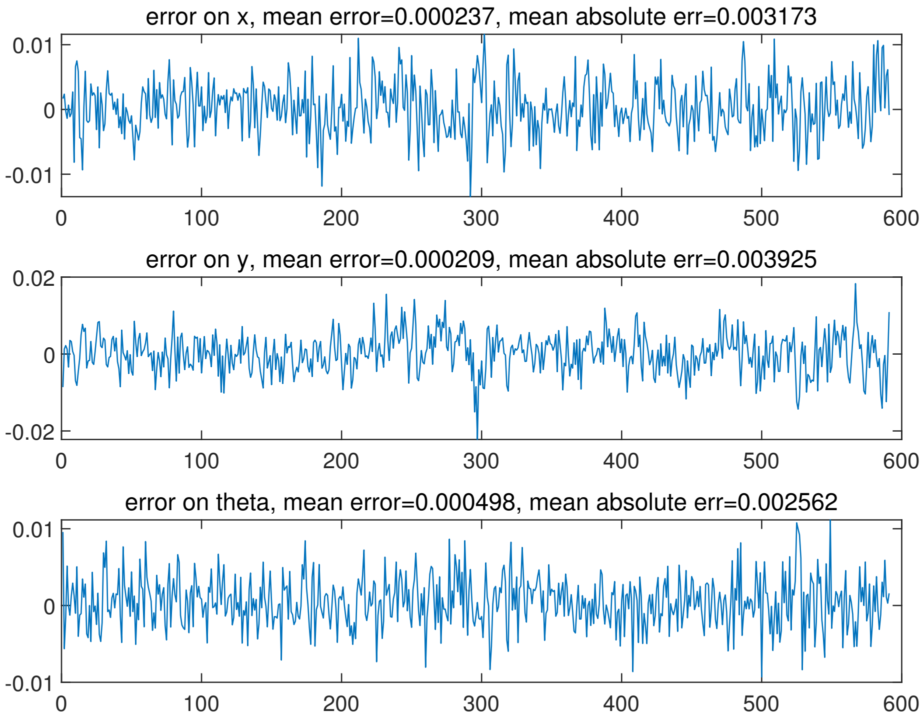

Figure 18: Error in case 1.

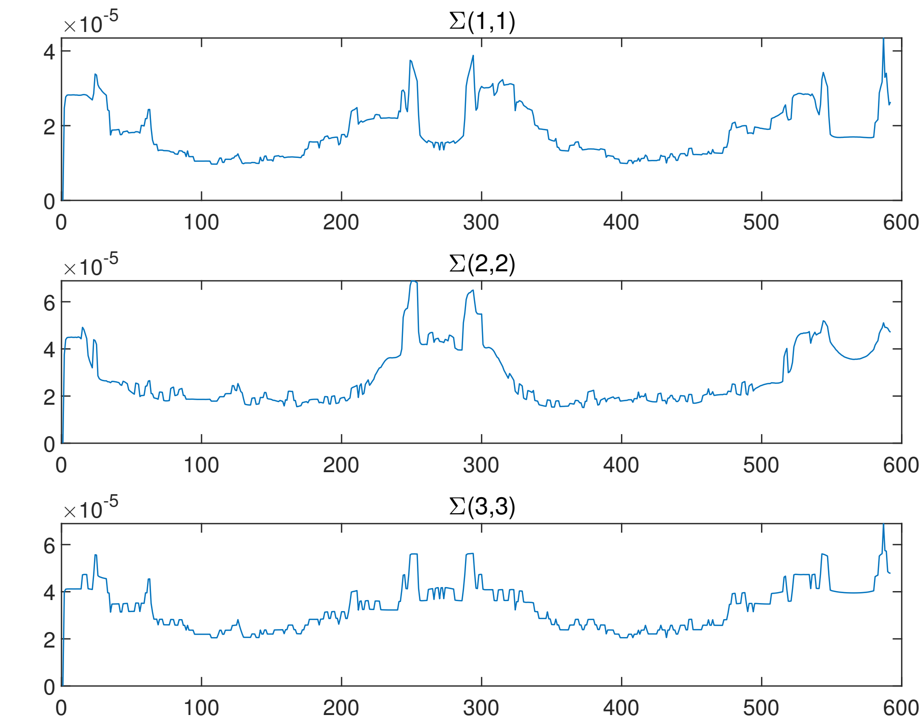

Figure 19: Covariance in case 1.

2.3.2 map_pent_big_10.txt + so_pb_10_outlier.txt

In this case, the covariance matrix of the process and measurement noises are:

$$R = \begin{bmatrix} 0.01^2 & 0 & 0 \\ 0 & 0.01^2 & 0 \\ 0 & 0 & (\frac{2\pi}{360})^2 \end{bmatrix}$$

$$Q = \begin{bmatrix} 0.2^2 & 0 \\ 0 & 0.2^2 \end{bmatrix}$$

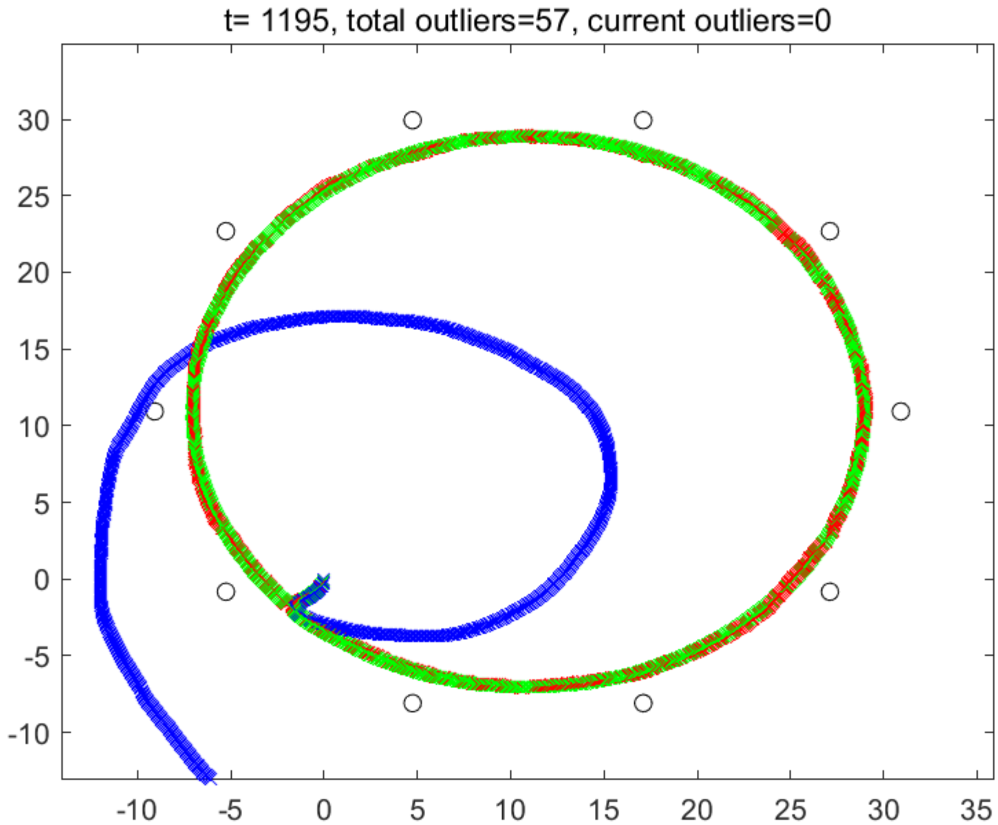

Figure 20: Path in case 2.

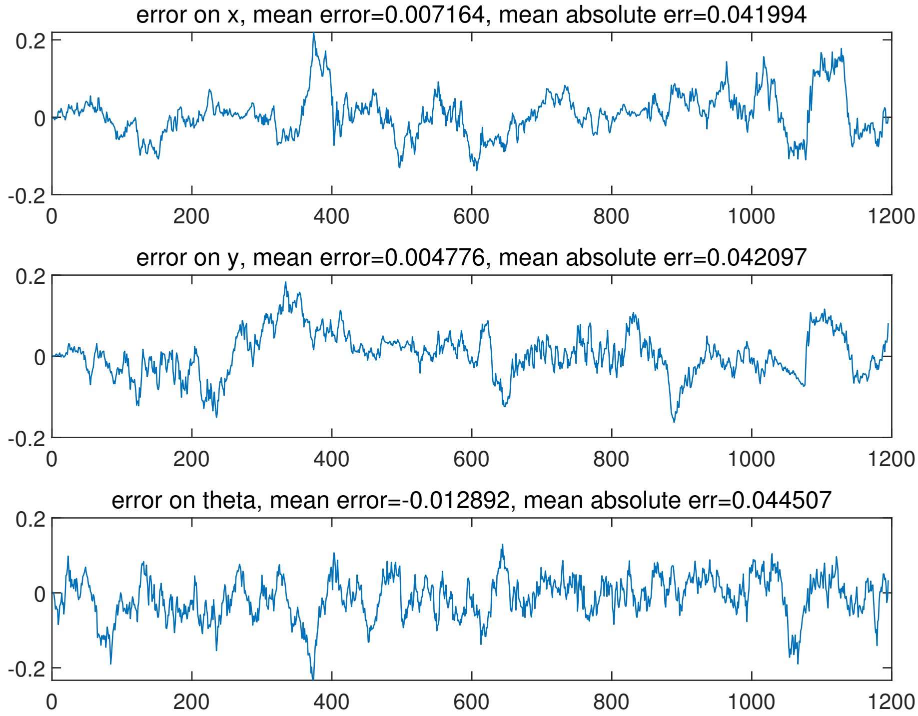

Figure 21: Error in case 2.

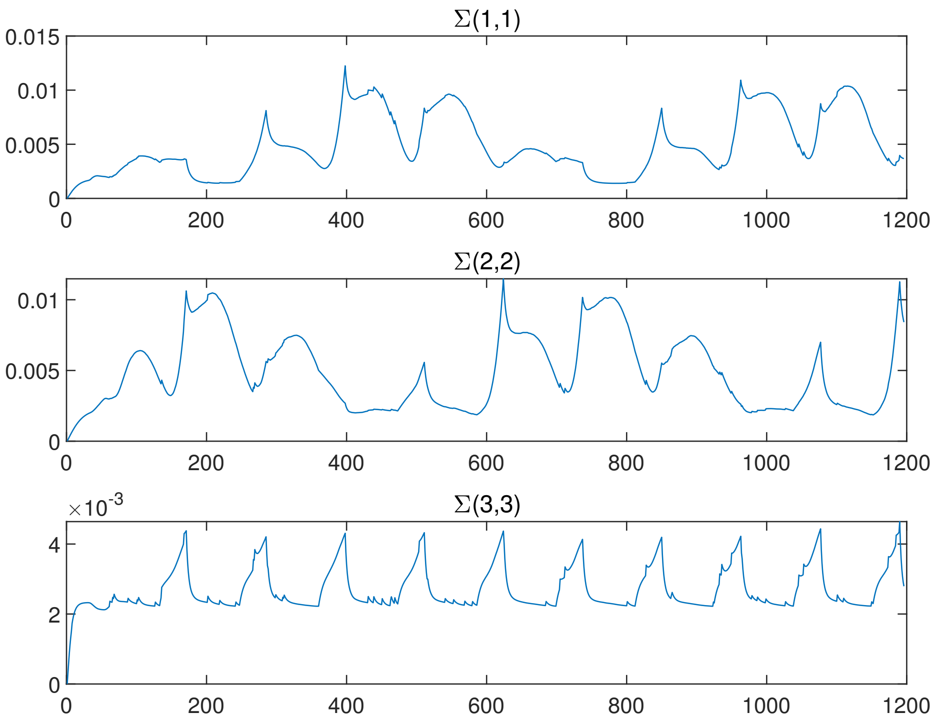

Figure 22: Covariance matrix in case 2.

2.3.3 map_pent_big_40.txt + so_pb_40_no.txt

In this case, the covariance matrix of the process and measurement noises are:

$$R = \begin{bmatrix} 1^2 & 0 & 0 \\ 0 & 1^2 & 0 \\ 0 & 0 & 1^2 \end{bmatrix}$$

$$Q = \begin{bmatrix} 0.1^2 & 0 \\ 0 & 0.1^2 \end{bmatrix}$$

also, the δ_M is set to 1.

- Sequential update

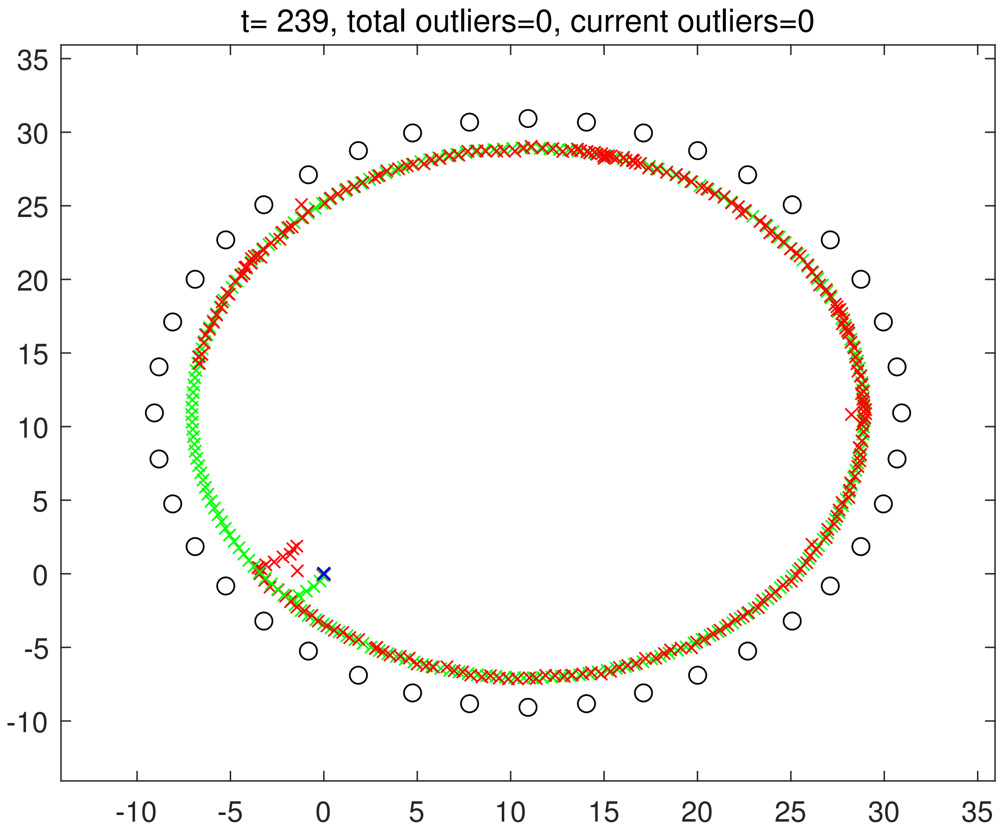

Figure 23: Path in case 3.

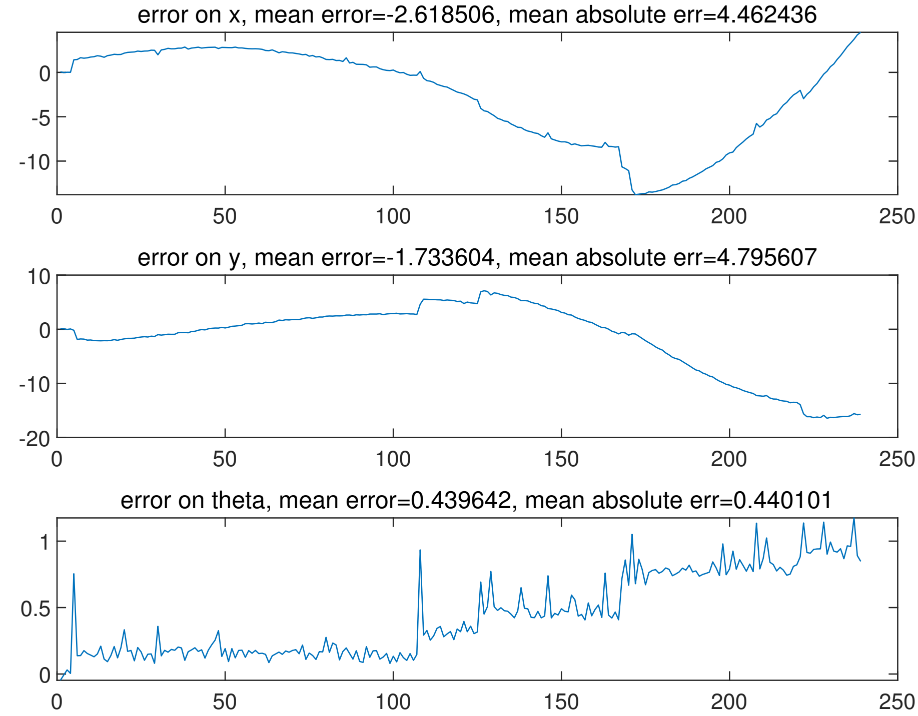

Figure 24: Error in case 3.

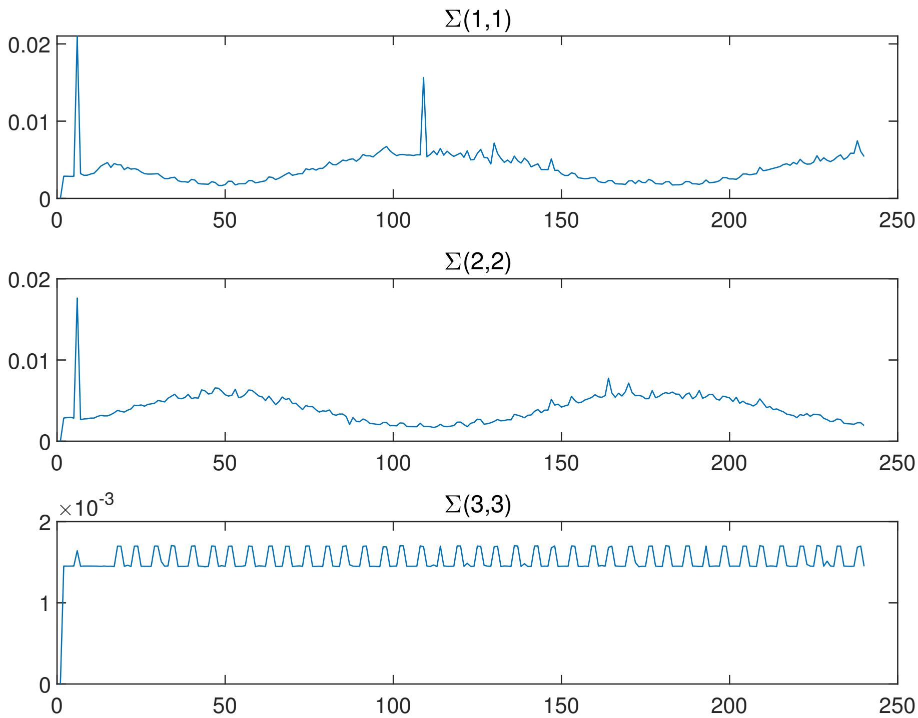

Figure 25: Covariance matrix in case 3.

- Batch update

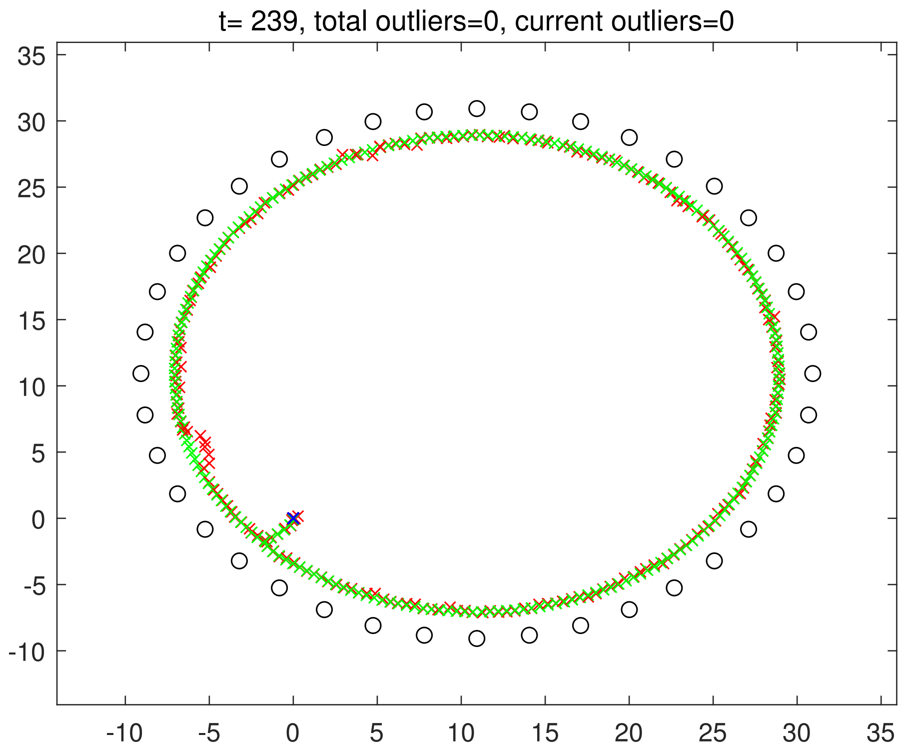

Figure 26: Path in case 3 with Batch update.

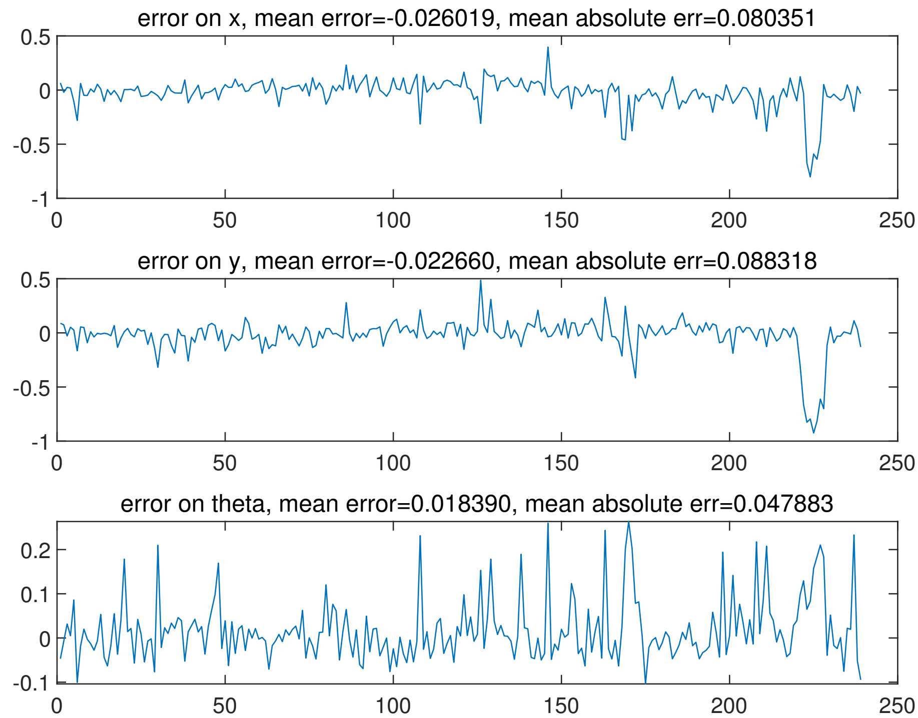

Figure 27: Error in case 3 with Batch update.

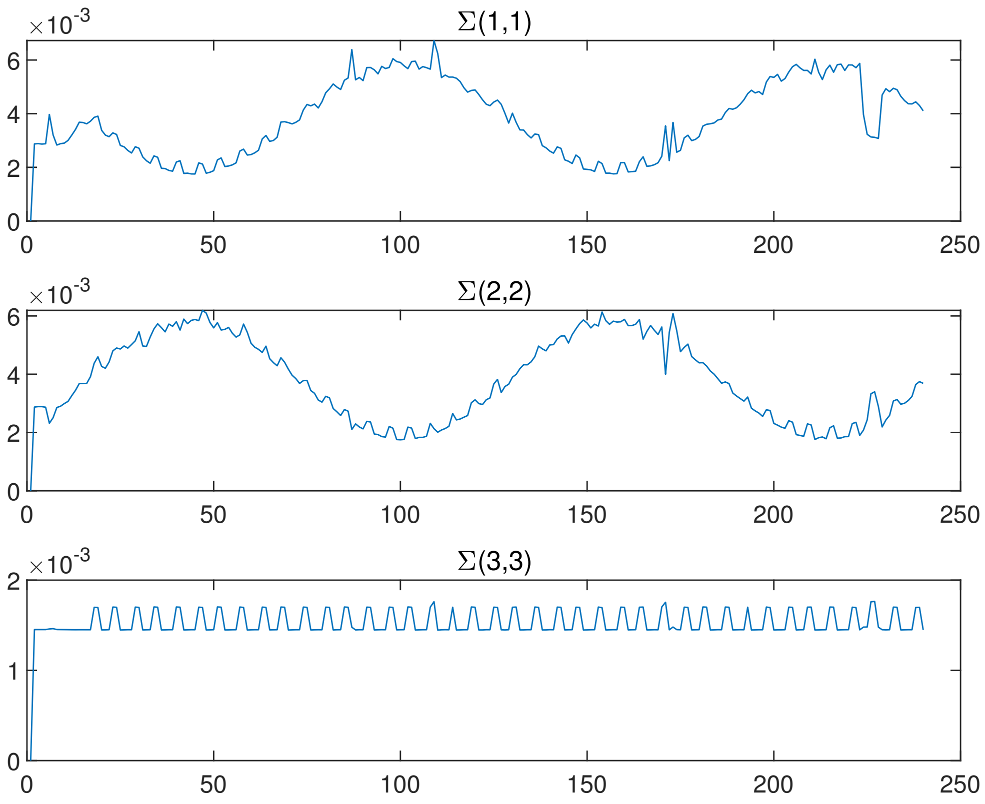

Figure 28: Covariance matrix in case 3 with Batch update.