# EL2320 - Applied Estimation - Lab 1

## Extended Kalman Filtering (EKF)

### PART I - Preparatory Questions

Alexandra Hotti

October 2019

# 1 Linear Kalman Filter

## 1. What is the difference between a "control" $u_t$, a "measurement" $z_t$ and the state $x_t$? Give examples of each!

**Definition:**

- **The state** $x_t$: A state is the robots internal belief of the world (Thrun et al. 2005, p 19). The state consists of all information about a robot and the robots environment that might impact the future. Relevant information about the robot is for instance its pose, velocity and whether its sensors are functioning (Thrun et al. 2005, p 20). Relevant information about the robot surroundings could be the location and velocity of objects in its environment (Thrun et al. 2005, p 21).
- **The measurement** $z_t$: When a robot wants to acquire momentary knowledge at time t about its environments state it uses sensors to gather measurements: $z_t$. This could for instance be images from a camera or the result of a range scanner (Thrun et al. 2005, p. 22).
- **The measurement** $u_t$: A robot impacts its environments state at time $t$ via control actions $u_t$. It does so by for instance moving around itself or by moving objects in its environment(Thrun et al. 2005, p. 22).

**Differences:**

- One important distinction between control $u_t$ and measurements $z_t$ is that the former reduces the robots certainty while the later increases the robots knowledge. Control $u_t$ actions result in loss of knowledge due to uncertainty in robot actuation. This occurs during the prediction step. Measurement data on the other hand increases the robots information about its surroundings and thus reduces its uncertainty later during the update (Thrun et al. 2005, p. 24).
- As a measurement at time step t $z_t$ is a function of the current statet $x_t$, it creates a way for us to infer information about the state $x_t$. This function is defined for the KF as $z_t = C_t x_t + \delta_t$. An example of this could be taking a sensor measurement of the temperature with a certain uncertainty.
- Our next state $x_t$ is a function of a control action $u_t$ at the same time step t and the former state $x_{t-1}$: $x_t = A_t x_{t-1} + B_t u_t + \epsilon_t$. Meaning that a control action allows us to transition between two states. An example of a control action could be to make the robot move to a new position.

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

The uncertainty during an update can either remain unchanged or decrease. Therefore, evidently it cannot increase. During the update step we acquire more information about the state via our measurement $z_t$ and thus the uncertainty will decrease. This is shown during the update when the covariance is updates in eq. (1).

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

In (1) the positive semi definite matrix $C_t$ is multiplied by $K_t$ which is in the range [0,1]. Thus when we compute $I - K_t C_t$ the elements of the resulting matrix will have values less then or equal to one. Accordingly when we multiply $\bar{\Sigma_t}$ by this difference, it is weighted such that: $\Sigma_t \leq \bar{\Sigma_t}$ and therefore the uncertainty cannot increase. It can remain unchanges or decrease.

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

The weighting between measurements and beliefs is decided by the Kalman gain. In the Kalman filter algorithm for linear Gaussian state transition it is incorporated in the mean and covariance updates:

$$\mu_t = \bar{\mu_t} + K_t(z_t - C_t\bar{\mu_t}) \tag{2}$$

Here the expected measurement at time t is given by: $C_t\bar{\mu_t}$ and the actual measurement by: $z_t$. Thus, the Kalman gain dictates how much importance we place on the actual compared to the predicted updated mean value (Thrun et al. 2005, p. 42-43).

$$\Sigma_t = (I - K_t C_t)\bar{\Sigma_t} \tag{3}$$

Here the change in $\bar{\Sigma_t}$ is weighted by the Kalman gain (Thrun et al. 2005, p. 42-43).

The Kalman gain is then defined as:

$$K_t = \bar{\Sigma_t} C_t^T (C_t \bar{\Sigma_t} C_t^T + Q_t)^{-1} \tag{4}$$

We see from the equation that $K_t$ depends on the co-variance ,i.e. the uncertainty, after the prediction: $\bar{\Sigma_t}$ and also the measurement noise $Q_t$. Since we multiply $\bar{\Sigma_t} C_t^T$ by the inverse of the sum of the measurement noise and the term dependent on $\bar{\Sigma_t}$, we get a weighting factor between these two terms. Such that if the measurement noise $Q_t$ is large compared to $C_t \bar{\Sigma_t} C_t^T$ then $K_t$ becomes small and if $Q_t$ is small compared to $C_t \bar{\Sigma_t} C_t^T$ then $K_t$ becomes large. Meaning that we are weighing the update based on the uncertainty $\Sigma$ and the errors of the measurements $Q_t$.

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

From equation (4) of the Kalman gain one can infer that if the measurement model error $Q_t$ is large then the resulting Kalman gain is small. Thus, in the update step the change in $\mu_t$ and $\Sigma_t$ becomes incremental and the algorithm becomes very slow.

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

For the measurement to have a high impact during the update step we must get a large Kalman gain value for that particular measurement. From (4) it is apparent that reducing the uncertainty in the measurement by decreasing the covariance of the measurement noise $Q_t$ will increase the impact of each update.

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

During the prediction step the uncertainty over the belief often increases due to a control action. We can see this in line 2 i the Kalman filter algorithm (Thrun et al. 2005, p. 42):

$$\bar{\Sigma_t} = A_t \Sigma_{t-1} A_t^T + R_t \tag{5}$$

If $A_t$ is larger then the identity matrix then the first term in (5), i.e: $A_t \Sigma_{t-1} A_t^T$ will be larger than the co-variance at the previous state: $\Sigma_{t-1}$. However, A does not have to be positive even though it usually is.

The next term in our expression $R_t$ is positive definite since it is the co-variance of a Gaussian. Meaning that we are adding a positive value to the updated $\bar{\Sigma_t}$.

Thus, the belief uncertainty often increases during prediction, but it does not necessarily have to increase since $A_t$ is not really constrained.

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

In this case we get a closed form computation and not an estimate for the posterior. This closed form computation of the posterior is another Gaussian distribution since "... any linear transformation of a Gaussian random variable results in another Gaussian random variable.." (Thrun et al. 2005, p. 54). This new Gaussian posterior is what the Kalman Filter locates and thus we cannot find a better solution than the exact solution.

If we would apply another type of estimator we would get a solution with larger variance.

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

A MAP estimator needs a prior distribution, while a MLE estimator does not. Thus, if we have access to a Normal state prior distribution the Kalman filter is a MAP estimator and if it is not included it becomes an MLE estimator.

# 2 Extended Kalman Filter

### 9. How does the Extended Kalman filter relate to the Kalman filter?

The Extended Kalman filter is a generalization of the Kalman filer. In EKF the linearity assumptions is relaxed such that it includes systems with nonlinear functions $x_t = g(u_t, x_{t-1}) + \varepsilon \approx g(u_t, \mu_{t-1}) + G_t(x_{t-1} - \mu_{t-1}) + \varepsilon$ and $z_t = h(x_t) + \delta_t \approx h(\bar{\mu_t}) + H_t(x_t - \bar{\mu_t}) + \delta_t$, instead of $x_t = A_t\mu_{t-1} + B_t u_t$ and $z_t = C_t x_t + \delta_t$. Instead the non-linear functions are linearized around the expectation $\mu_{t-1}$. (Thrun et al. 2005, p. 54-56).

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

No. As we do not have linearity we have to use approximate methods which means that we will not get a closed form solution. (Thrun et al. 2005, p. 54-56). Instead we acquire an estimate solution through local linearization. How consistent this solution is determined by the relation between the width of the Gaussian being passed through the linear approximation and the local linearity of the non-linear function. Meaning that we will get a better solution if the non-linear function is ”more” locally linear.

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

One possible way of avoiding divergence is to make sure that the filter is less certain about its predictions and updates by increasing the co-variance of the modeled measurement $Q_t$ and transition noise $R_t$.

# 3 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 the posterior belief of its location $p(x, y, \theta|r)$ look like? A formula is not needed but describe it at least.

The only information that the robot has is a measured range to this landmark. Therefore, the robot could be located anywhere at the edge of a circle around the landmark and thus it cannot tell what its theta angle is relative to the landmark. It could any angle is equally probable (uniformly distributed) between $[-\pi, \pi]$.
When it comes to the $(x, y)$ coordinates we can infer more information. At each point of the circle with radius r we get an error with a Gaussian spread centered at r. Such that the robot could be within a certain range inside or outside of the circle. All in all we get a circle with a Gaussian spread at every single point at its circumference.

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

It is still a circle and we are still unsure about r, thus we get a Gaussian spread again at every point in the circle centered at r. Also, we get a Gaussian uncertainty for the measured bearing.

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

Since we are quite certain about the distance the robot travel, but have a large uncertainty over the heading $\theta$ we get a posterior with a banana shape (depending on how we define large). Meaning that the large spread over $\theta$ corresponds to the bananas length and the small uncertainty in the range corresponds to the width of the banana.

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

One problem that arises is that since we get this half circular posterior, locally linearizing as an estimate to the non-linear functions might not be possible. Meaning that the Jacobians will not be good estimates in local regions on the circular shape.

# References

Thrun, S., Burgard, W. & Fox, D. (2005), *Probabilistic robotics*, MIT Press, Cambridge, Mass.