EMARO
European Master on Advanced Robotics
COMPUTER VISION

Some notes about Kalman filter

Manuela Chessa, Fabio Solari

Department of Biophysical and Electronic Engineering
University of Genoa, I-16145 Genova, ITALY
manuela.chessa@unige.it
fabio.solari@unige.it

Contents

---

1 Introduction

The Kalman filter (KF) [Kalman, 1960] is an optimal estimator: it evinces quantities of interest from indirect, inaccurate and uncertain observations. The KF approach is a statistically optimum way to combine sensor data: in particular, each sensor could provide measurement of different quantities and with different levels of precision. The KF is widely used (e.g.: tracking targets, global positioning system (GPS), computer graphics, robot localization and map building) because it is characterized by: good results in practice, convenient form for on-line real time processing and it is not necessary to invert the measurement equation.

In the following we introduce some properties of the KF:

- Estimator: it performs a stochastic estimation of the system state from noisy sensor measurements.
- Model based: it is based on (it makes use of) a system model described by a state equation and an output (measurement) equation.
- Linear: in the standard form all the equations are linear and the noises are Gaussian with zero mean.
- Least square: the KF minimizes the mean square error of the estimated parameters. In this respect, the KF is an optimal estimator.
- Recursive: it can process new measurements as they arrive. Thus the KF eliminates the need for storing the entire past observed data and is computationally more efficient.
- Stochastic: the confidence about Kalman quantities are expressed in terms of probability distributions.

Whether some of the KF hypothesis fails (e.g. nonlinear system), we can consider different forms of the KF to take into account these failures (e.g. Extended Kalman Filter).

2 The Kalman filter

Through the KF [Haykin, 1991] [Welch and Bishop, 2001] [Gelb, 1974] [Brown and Hwang, 1997] we can face the problem of estimating the state $\boldsymbol{x} \in \mathcal{R}^n$ of a (discrete-time) system described by the linear stochastic difference equation (**process equation**)

$$\boldsymbol{x}[k] = \boldsymbol{\Phi}[k, k-1]\ \boldsymbol{x}[k-1] + \boldsymbol{S}[k-1]\ \boldsymbol{s}[k-1] + \boldsymbol{n}_1[k-1] \tag{1}$$

and with a measurement $\boldsymbol{y} \in \mathcal{R}^m$ that is (**measurement equation**)

$$\boldsymbol{y}[k] = \boldsymbol{C}[k]\ \boldsymbol{x}[k] + \boldsymbol{n}_2[k] \tag{2}$$

The matrix $\boldsymbol{\Phi}[k, k-1]$ is a known state transition matrix that relates the state at the previous time step $k-1$ to the state at the current step k. The matrix $\boldsymbol{S}[k]$ takes into the account an optional control input to the state. The matrix $\boldsymbol{C}[k]$ is a known measurement matrix. the process and measurement uncertainty are represented by $\boldsymbol{n}_1[k] = N(0, \boldsymbol{\Lambda}_1[k])$ and $\boldsymbol{n}_2[k] = N(0, \boldsymbol{\Lambda}_2[k])$ The space spanned by the observations $\boldsymbol{y}[1], \boldsymbol{y}[2], \cdots, \boldsymbol{y}[k-1]$ is denoted by $\boldsymbol{\mathcal{Y}}_{k-1}$

All the matrices might change with each time step, but in a first approximation we can assume they are constant. Moreover, we assume the state dimension n is equal to the measurement dimension m.

To understand how the KF works, we define the following quantities: $\hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_{k-1}]$ is the *a priori* state estimate at step k given knowledge of the process at step $k-1$ and $\hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_k]$ denotes the *a posteriori* state estimate at step k given the measurement at the step k. The *a priori* estimate error covariance is

$$\boldsymbol{K}[k, k-1] = E\left\{(\boldsymbol{x}[k] - \hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_{k-1}])(\boldsymbol{x}[k] - \hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_{k-1}])^T\right\} \tag{3}$$

and the *a posteriori* estimate error covariance is

$$\boldsymbol{K}[k] = E\left\{(\boldsymbol{x}[k] - \hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_k])(\boldsymbol{x}[k] - \hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_k])^T\right\} \tag{4}$$

where $E\{\cdot\}$ means expected value and $(\cdot)^T$ denotes the matrix transpose.

The aim of the KF is to compute an *a posteriori* state estimate using an *a priori* state estimate and a weighted difference between the current measurement and the predicted measurement: thus, we compare what we know about the nature of the system (the model) and what the measurement (the data) points out. This relationship has the following formulation:

$$\hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_k] = \hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_{k-1}] + \boldsymbol{G}[k]\ \left(\boldsymbol{y}[k] - \boldsymbol{C}[k]\ \hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_{k-1}]\right) \tag{5}$$

The difference term in Eq. 5 is called **innovation** (or residual) $\boldsymbol{\alpha}[k]$:

$$\boldsymbol{\alpha}[k] = \boldsymbol{y}[k] - \boldsymbol{C}[k]\ \hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_{k-1}] \tag{6}$$

It takes into account the discrepancy between the actual measurement $\boldsymbol{y}[k]$ and the predicted measurement $\boldsymbol{C}[k]\ \hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_{k-1}]$. The matrix $\boldsymbol{G}[k]$ is the gain (called Kalman gain) that minimizes the *a posteriori* error covariance (Eq. 4):

$$\boldsymbol{G}[k] = \boldsymbol{K}[k, k-1]\ \boldsymbol{C}^T[k]\ \left(\boldsymbol{C}[k]\ \boldsymbol{K}[k, k-1]\ \boldsymbol{C}^T[k] + \boldsymbol{\Lambda}_2[k]\right)^{-1} \tag{7}$$

$(\cdot)^{-1}$ denotes the matrix inverse.

We can think the Eqs. 5 and 4 as mean and covariance expressions of the KF output.

2.1 The Kalman filter recursive algorithm

The KF estimates the system state at a some time and then obtains a feedback from the actual measurement (noisy data). We can consider the KF equations belonging to two groups: the time update equations and the measurement update equations. The first ones project forward (in time) the current state and the error covariance estimates to have the *a priori* estimates for the next time step. The measurement update equations (feedback) embody a new measurement into *a priori* estimate to gain an improved *a posteriori* estimate.

In the Table 1 we show a complete description of the steps of the KF algorithm, based on the filtered estimate of the state vector.

3 Bayesian background

In this section we provide some hints about the KF derivation from a general state estimation result [Gelb, 1974] [Barker *et al.*, 1995] [Brown and Hwang, 1997].

In general, we can consider system models of the class:

$$\boldsymbol{x}[k] = f(\boldsymbol{x}[k-1], \boldsymbol{\pi}_1[k-1], k, k-1) \tag{8}$$

$$\boldsymbol{y}[k] = g(\boldsymbol{x}[k], \boldsymbol{\pi}_2[k], k) \tag{9}$$

The Eq. 8 is the process equation and the Eq. 9 is the measurement equation. We assume that the probability density functions of $\boldsymbol{x}[0]$, $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ are known a priori and mutually statistically independent. We also assume that the densities conditioned $p(\bar{\boldsymbol{x}}[k]|\bar{\boldsymbol{x}}[k-1])$ and $p(\bar{\boldsymbol{y}}[k]|\bar{\boldsymbol{x}}[k])$ can be computed. The bar symbol above a variable denotes a realization of the variable: e.g. $\bar{\boldsymbol{x}}[k]$ denotes a realization of $\boldsymbol{x}[k]$.

It is worthy to note that the $\boldsymbol{x}[k]$ forms a discrete time Markov process.

The Kalman filter recursive algorithm

- *Input vector process*
  Observations $= \{\boldsymbol{y}[1], \boldsymbol{y}[2], \cdots, \boldsymbol{y}[k]\}$
- *Known parameters*
  State transition matrix $\boldsymbol{\Phi}[k, k-1]$
  Measurement matrix $\boldsymbol{C}[k]$
  Process noise covariance $\boldsymbol{\Lambda}_1[k]$
  Measurement noise covariance $\boldsymbol{\Lambda}_2[k]$
- *Initial conditions*
  $\hat{\boldsymbol{x}}[0|\boldsymbol{\mathcal{Y}}_0]$
  $\boldsymbol{K}[0]$
- *Computation*: $k = 1, 2, \cdots$
  - **Time update ("prediction equations")**
    1. *A priori* state estimate (predicted estimate of the state)
       $\hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_{k-1}] = \boldsymbol{\Phi}[k, k-1]\ \hat{\boldsymbol{x}}[k-1|\boldsymbol{\mathcal{Y}}_{k-1}] + \boldsymbol{S}[k-1]\ \boldsymbol{s}[k-1]$
    2. *A priori* estimate error covariance
       $\boldsymbol{K}[k, k-1] = \boldsymbol{\Phi}[k, k-1]\ \boldsymbol{K}[k-1]\ \boldsymbol{\Phi}^T[k, k-1] + \boldsymbol{\Lambda}_1[k-1]$
  - **Measurement update ("correction equations")**
    1. Innovation
       $\boldsymbol{\alpha}[k] = \boldsymbol{y}[k] - \boldsymbol{C}[k]\ \hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_{k-1}]$
    2. Covariance of the innovation
       $\boldsymbol{\Sigma}[k] = \boldsymbol{C}[k]\ \boldsymbol{K}[k, k-1]\ \boldsymbol{C}^T[k] + \boldsymbol{\Lambda}_2[k]$
    3. Kalman gain
       $\boldsymbol{G}[k] = \boldsymbol{K}[k, k-1]\ \boldsymbol{C}^T[k]\ \boldsymbol{\Sigma}^{-1}[k]$
    4. *A posteriori* state estimate (filtered estimate of the state)
       $\hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_k] = \hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_{k-1}] + \boldsymbol{G}[k]\ \boldsymbol{\alpha}[k]$
    5. *A posteriori* estimate error covariance
       $\boldsymbol{K}[k] = \boldsymbol{K}[k, k-1] - \boldsymbol{G}[k]\ \boldsymbol{C}[k]\ \boldsymbol{K}[k, k-1]$

Table 1: A complete description of the operations of the KF, based on the filtered estimate of the state vector. The time update equations, projecting forward the current state, obtain a prediction. The measurement update equations correct this prediction through an actual measurement.

The aim is to determine the posterior density $p(\bar{\boldsymbol{x}}[k]|\boldsymbol{\mathcal{Y}}_k)$ of the state $\boldsymbol{x}[k]$ given the observed data $\boldsymbol{\mathcal{Y}}_k$.

Applying Bayes rule (and the Markovianity), we can get the posterior density as a pair of mutually recursive equations:

$$p(\bar{\boldsymbol{x}}[k] \mid \boldsymbol{\mathcal{Y}}_{k-1}) = \int d\bar{\boldsymbol{x}}[k-1] \;\; p(\bar{\boldsymbol{x}}[k] \mid \bar{\boldsymbol{x}}[k-1])\, p(\bar{\boldsymbol{x}}[k-1] \mid \boldsymbol{\mathcal{Y}}_{k-1}) \tag{10}$$

$$p(\bar{\boldsymbol{x}}[k] \mid \boldsymbol{\mathcal{Y}}_{k}) = p(\bar{\boldsymbol{y}}[k] \mid \bar{\boldsymbol{x}}[k])\, p(\bar{\boldsymbol{x}}[k] \mid \boldsymbol{\mathcal{Y}}_{k-1}) \tag{11}$$

The Eq. 10 is a prediction of a future state and the Eq. 11 is a correction of the prediction given a new observation. The problem is to evaluate these expressions: it depends on the form of the densities involved. However, if we consider a linear Gaussian case of the model equations 8 and 9, we can efficiently compute the posterior density $p(\bar{\boldsymbol{x}}[k]|\boldsymbol{\mathcal{Y}}_k)$ (by using the theorems related to Gaussian random vectors). Thence, the Eqs. 8 and 9 become the Eqs. 1 and 2. We can compute the expressions of the prediction and correction (Eqs. 10 and 11) by exploiting the following densities:

$$p(\bar{\boldsymbol{x}}[k] \mid \bar{\boldsymbol{x}}[k-1]) = N(\boldsymbol{\Phi}[k, k-1]\; \bar{\boldsymbol{x}}[k-1], \boldsymbol{\Lambda}_1[k-1]) \tag{12}$$

$$p(\bar{\boldsymbol{y}}[k] \mid \bar{\boldsymbol{x}}[k]) = N(\boldsymbol{C}[k]\; \bar{\boldsymbol{x}}[k], \boldsymbol{\Lambda}_2[k]) \tag{13}$$

$$p(\bar{\boldsymbol{x}}[0]) = N(\bar{\boldsymbol{x}}[0], \boldsymbol{K}[0]) \tag{14}$$

The result is known as the Kalman filter algorithm.

4 Convergence and consistency of the Kalman filter

In this section we address the validation problem of the KF output: we should perform a reliability analysis and statistical testing of the quantities characterizing the KF. To this purpose we can consider two different tools: the convergence monitoring and the consistency checking [De Schutter *et al.*, 1999b] [De Schutter *et al.*, 1999a]. The first one is also related to the reduction of the uncertainty of the state estimate. The second one is used to catch discrepancies between the model and the measurement.

4.1 Uncertainty of the state estimate

The *a posteriori* estimate error covariance matrix $\boldsymbol{K}[k]$ is a measure for the uncertainty of the state estimate $\hat{\boldsymbol{x}}[k|\boldsymbol{\mathcal{Y}}_k]$. In particular, the main diagonal of the matrix $\boldsymbol{K}[k]$ represents the mean squared error of the state estimate (see Eq. 4), thus we can use these values to monitor how the uncertainty of the state estimate changes as a function of the recursion time: the new measurements should reduce this uncertainty. It is worthy to note that we can monitor the behavior of each variable of the state.

We can, also, perform a convergence test for the KF using a singular value decomposition of the error covariance matrix $\boldsymbol{K}[k]$: in this case we can consider the singular directions as a new state space and the singular values as the semi-axes of an ellipsoid describing the uncertainty of the state estimate (see Fig. 1). Convergence of the estimate means that the ellipsoid contracts in all directions.

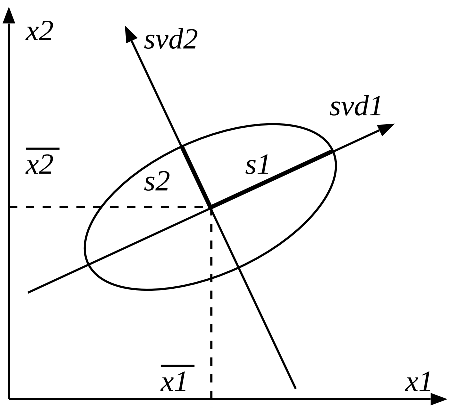


Figure 1: The uncertainty ellipsoid in a 2D state space. The semi-axes $s1$ and $s2$ are the square root of the singular values of the error covariance matrix $\boldsymbol{K}[k]$ and $svd1$ and $svd2$ the related singular directions. $x1$, $x2$ and $\overline{x1}$, $\overline{x2}$ represent the state space and the *a posteriori* state estimate, respectively

4.2 Detection of failures

The error covariance matrix $\boldsymbol{K}[k]$ provides us only information about the properties of convergence of the KF and not whether the KF converges to the correct values. It is due to the fact that the matrix $\boldsymbol{K}[k]$ is determined by the model and measurement equations, that is by the "*a priori* knowledge", and not by the actual measurement values, that is by the "*a posteriori* knowledge". We can verify the correctness of the KF output using a consistency check. In particular, we do not consider test that can be calculated only in simulation, but not in real condition where the state is unknown.

Since, the Kalman filter strongly relies on the model of the system, wrong models result in estimate errors, so the KF output is not reliable. We have to check the consistency between the innovation and the model (between observed and predicted values) in statistical terms. A measure of the reliability of the KF output is the Normalized Innovation Squared (NIS), that has a χ^2 distribution with m degrees of freedom:

$$NIS_k = \boldsymbol{\alpha}^T[k]\; \boldsymbol{\Sigma}^{-1}[k]\; \boldsymbol{\alpha}[k] \tag{15}$$

where m is the number of statistically independent measurements (i.e. the dimension of $\boldsymbol{y}$). It is worthy to note that this quantity defines a local distance measure, namely the Mahalanobis distance.

Another measure is the Summed Normalized Innovation Squared ($SNIS$), that has a χ^2 distribution with $M \cdot m$ degrees of freedom. It is the sum of the M latest NIS values:

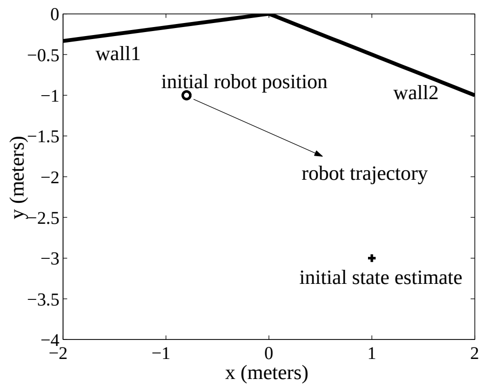
$$SNIS_k = \sum_{j=k-M+1}^{k} \boldsymbol{\alpha}^T[j]\; \boldsymbol{\Sigma}^{-1}[j]\; \boldsymbol{\alpha}[j] \tag{16}$$

The consistency test is performed by checking whether the $SNIS$ (or NIS) is within a given confidence interval (e.g. it can be chosen such that the 95% of the correct matches hit the interval). It is, also, possible to exploit the $SNIS$ (or NIS) properties to detect whether the actual observations are an instance of the model embedded in the KF [Bousquet *et al.*, 1998] [De Schutter *et al.*, 1999a].

5 An example: a mobile robot

Let's consider a robot moving on a planar surface. The robot task is to locate itself on an environment map (see Fig. 2): the internal sensors measure its speed (v) and orientation (θ); moreover it uses sensors that measure the distance $\boldsymbol{d} = (d_x, d_y)$ of the obstacles (e.g. walls).

Environment map



Model equations

State equations:
$x_1[k] = x_1[k-1] + v\cos(\theta)\Delta t$
$x_2[k] = x_2[k-1] + v\sin(\theta)\Delta t$
$v = 0.5\ m/s,\ \theta = -0.6\ rad$
and $\Delta t = 0.04\ s$

Measurement matrix:

$$\boldsymbol{C} = \begin{bmatrix} c_{1x} & c_{1y} \\ c_{2x} & c_{2y} \end{bmatrix}$$

Figure 2: On the left there is a sketch of the environment where the robot moves. On the right there are the equations describing the system (see text): Δt is the time step.

For the sake of simplicity, the robot moves on a straight line and we model the walls through the equations $a_1x + b_1y = 0$ and $a_2x + b_2y = 0$. Thus, the distances $\boldsymbol{d_1}$ and $\boldsymbol{d_2}$ are:

$$d_{1x} = c_{1x}x_1[k] \qquad\qquad d_{1y} = c_{1y}x_2[k]$$

and

$$d_{2x} = c_{2x}x_1[k] \qquad\qquad d_{2y} = c_{2y}x_2[k]$$

with $c_{1x} = a_1/\sqrt{a_1^2 + b_1^2}$, $c_{1y} = b_1/\sqrt{a_1^2 + b_1^2}$, $c_{2x} = a_2/\sqrt{a_2^2 + b_2^2}$, $c_{2y} = b_2/\sqrt{a_2^2 + b_2^2}$. The system state $\boldsymbol{x}[k] = (x_1[k], x_2[k])^T$ is the current position of the robot plus the motion vector, thus the transition matrix is: $\boldsymbol{\Phi}[k, k-1] = \boldsymbol{I}$. The model described is implemented with positions represented in two dimensional Cartesian coordinates (see Fig. 2).

Estimate of the robot position

In this first simulation we set optimal parameters to obtain good performances of the KF. The initial state has the value $[-0.8, -1]^T$ meters, the initial state estimate has the value $[1, -3]^T$ meters with the estimate error covariance matrix equals to $0.09\boldsymbol{I}$ squared meters. The standard deviation of the measurement noise and the process noise are, respectively, 0.03 meters and 0.001 meters. The walls are described by the equations $x - 6y = 0$ and $x + 2y = 0$ (Fig. 3).

The form factor of the ellipsoids in the Fig. 3 depends on the relative position of the two walls, that is, on the relationships between the measurements. We show the value of the error covariance matrix $\boldsymbol{K}[k]$ at time step 100 for two different placing of the walls.

$$\begin{cases} y = 0 \\ x = 0 \end{cases} \Rightarrow \boldsymbol{K}[100] = \begin{bmatrix} 2.95 & 0 \\ 0 & 2.95 \end{bmatrix} 10^{-5}\ m^2$$

$$\begin{cases} x - 6y = 0 \\ x + 2y = 0 \end{cases} \Rightarrow \boldsymbol{K}[100] = \begin{bmatrix} 7.48 & -0.79 \\ -0.79 & 2.30 \end{bmatrix} 10^{-5}\ m^2$$

It is worthy to note that the off-diagonal elements indicate correlation between the estimates.

About the sensibility of the KF to the noise, we can verify that if the standard deviation of the measurement noise increases, the behavior is similar, but the values of the uncertainty are higher. Consequently, the state estimate moves away from the correct value. If we increase the standard deviation of the process noise, the values of the uncertainty are similar, but do not decrease significantly.

Detection of a wrong model

To test the behavior of the KF in presence of a wrong model, we modify the actual trajectory of the robot. To this end, we change the equations describing the actual model of the trajectory in this way: $x_a[k] = x_a[k-1] + v\cos(\theta - \pi/3)\Delta t$ and $y_a[k] = y_a[k-1] + v\sin(\theta)\Delta t$.

In Fig. 4 we can see that the value of $SNIS$ is very high, it indicates that there is something of wrong in the model.

References

[Barker *et al.*, 1995] A.L. Barker, D.E. Brown, and W.N. Martin. Bayesian estimation and the kalman filter. *Computers and Mathematics with Applications*, 30(10):55–77, 1995.

[Bousquet *et al.*, 1998] O. Bousquet, K. Balakrishnan, and V. Honavar. Is the hippocampus a kalman filter? In *Proc. of the Pacific Symposium on Biocomputing*, pages 655–666, 1998.

[Brown and Hwang, 1997] R.G. Brown and P.Y.C. Hwang. *Introduction to Random Signals and Applied Kalman Filtering: with MATLAB Exercises and Solutions.* John Wiley & Sons, Inc., 1997.

[De Schutter *et al.*, 1999a] J. De Schutter, H. Bruyninckx, J. Dutré, S. De Geeter, J. Katupitiya, S. Demey, and T. Lefebvre. Estimating first order geometric parameters and monitoring contact transitions during force controlled compliant motion. *Int. J. of Robotics Research*, 18:1161–1184, 1999.

[De Schutter *et al.*, 1999b] J. De Schutter, J. De Geeter, T. Lefebvre, and H. Bruyninckx. Kalman filters: A tutorial. *Journal A-Benelux Quarterly Journal on Automatic Control*, 40(4):538–546, 1999.

[Gelb, 1974] A. Gelb. *Applied Optimal Estimation.* Artech House, MIT Press, 1974.

[Haykin, 1991] S. Haykin. *Adaptive Filter Theory.* Prentice-Hall International Editions, 1991.

[Kalman, 1960] R.E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(D):35–45, 1960.

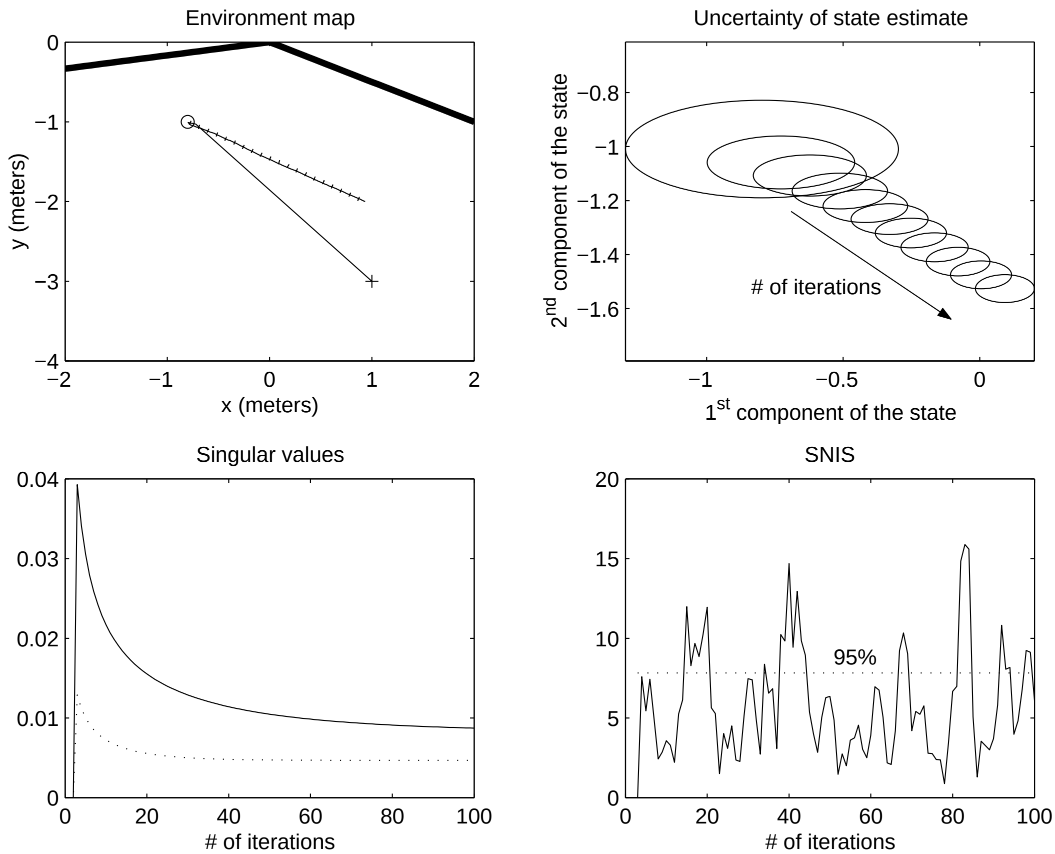


Figure 3: (top-left) We can see that the KF makes quickly a good estimate (solid line) of the robot trajectory (dotted line). It starts from the initial state estimate (cross symbol) to go in few steps near initial robot position (open circle). (top-right) This plot shows the uncertainties of the state estimate as ellipsoids in the state space: the semi-axes are related to the values of the main diagonal of the matrix $\boldsymbol{K}[k]$. These ellipsoids shrink (measurements reduce the uncertainty) as the number of iterations grows (arrow direction). The semi-axes are normalized to the initial value. (bottom-left) To perform a convergence test we can check whether the singular values of the error covariance matrix decrease or are constant. In this case the KF converges. (bottom-right) Since the $SNIS$ is below the threshold (dotted line), our model is right for the actual system, that is the KF output are reliable. We set $M = 3$ for the $SNIS$.

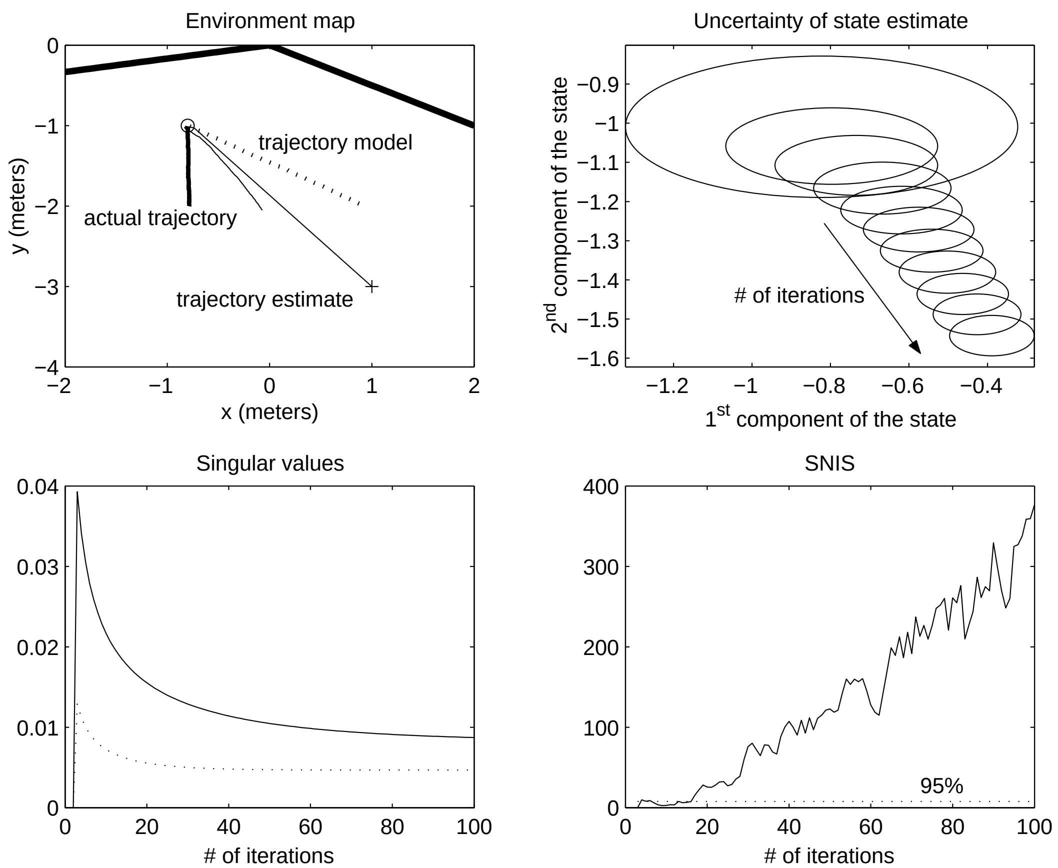


Figure 4: (top-left) In this case, we see the different trajectories due to wrong model: the trajectory of the model (dotted line) moves away from the actual trajectory (solid line), so the estimate (thin line) is not reliable. (top-right) and (bottom-left) The KF convergence is not influenced by the wrong model. Only the $SNIS$ (bottom-right) detects the unexpected characteristic in the nature of the system.

[Welch and Bishop, 2001] G. Welch and G. Bishop. An introduction to the kalman filter. In *SIGGRAPH 2001*, Los Angeles, USA, 2001.