Robust Kalman Filtering for Path Retrieval

Vinícius Miranda*
College of Computational Sciences, Minerva Schools at KGI
CS164: Optimization Methods

May 8, 2019

*The implementation is available at https://nbviewer.jupyter.org/gist/viniciusmss/9d2579c3f65c881af60c863e761da844

1 Problem Description

This project considers the problem of retrieving the true trajectory of a point-mass vehicle from a sequence of noisy location measurements. The vehicle behaves as a discrete-time linear time-invariant dynamic system (discrete LTI). I consider the use of Kalman filters to retrieve the correct trajectory under noise regimes where outliers are present. The remainder of this section provides a brief technical exposition of vehicle tracking as a discrete LTI and of Kalman filters. Section 2 presents the full specification of the optimization problem and section 3 concludes with an analysis of results.

1.1 Linear Dynamical Systems

A linear dynamical system is one whose state-space description varies dynamically as a function of time and its state at a given point in time (Dahleh, Dahleh, & Verghese, 2004; Schouwenaars, De Moor, Feron, & How, 2001). Generally, its trajectory follows the form

$$\dot{\mathbf{s}}(t) = \mathbf{A}_c(t)\mathbf{s}(t) + \mathbf{B}_c(t)\mathbf{u}(t) \tag{1}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{s}(t) \tag{2}$$

where $\mathbf{s}(t) \in \mathbb{R}^n$ is the state vector, $\mathbf{u}(t) \in \mathbb{R}^m$ the control vector, $\mathbf{y}(t) \in \mathbb{R}^p$ the output vector (e.g., the position), and $\mathbf{A}_c(t) \in \mathbb{R}^{n\times n}$, $\mathbf{B}_c(t) \in \mathbb{R}^{n\times m}$, and $\mathbf{C} \in \mathbb{R}^{p\times n}$ are context-specific matrices which describe the system dynamics (Calafiore & Ghaoui, 2014, p. 569; Schouwenaars et al., 2001). Note that this system is continuous and time-variant. Before proceeding, we will apply two simplifications to render the problem more tractable.

1.1.1 Linear Time-Invariant Systems

So-called LTI systems are those in which the laws governing the dynamics of the system do not change as a function of time. The motivation for such simplification will become clear once we define the application domain in Section 2. In LTIs, $\mathbf{A}_c$ and $\mathbf{B}_c$ are now constant matrices and (1) simplifies to

$$\dot{\mathbf{s}}(t) = \mathbf{A}_c\mathbf{s}(t) + \mathbf{B}_c\mathbf{u}(t) \tag{3}$$

1.1.2 Discrete-Time LTI systems

Our second and final simplification will be to discretize time. We shall now consider the system dynamics over N discrete time steps. We let $\mathbf{s}(k)$, $\mathbf{u}(k)$, and $\mathbf{y}(k)$ denote the state, control, and output vectors at time k, $k = 0, \dots, N-1$. (2) and (3) now become

$$\mathbf{s}(k+1) = \mathbf{A}\mathbf{s}(k) + \mathbf{B}\mathbf{u}(k) \tag{4}$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{s}(k) \tag{5}$$

where $\mathbf{A}$ and $\mathbf{B}$ are discretizations of $\mathbf{A}_c$ and $\mathbf{B}_c$ (Calafiore & Ghaoui, 2014, p. 569; Schouwenaars et al., 2001).

2 Solution Specification

2.1 Kalman Filtering

The Kalman filter is an algorithm which works on noisy time-series data to retrieve estimates of latent variables. We introduce uncertainty to the LTI system described above by rendering $\mathbf{u}$(k) a normally-distributed unknown variable and adding a noise component to the output or measurement vector $\mathbf{y}$(k). The final expression of the system becomes

$$\mathbf{s}(k+1) = \mathbf{A}\mathbf{s}(k) + \mathbf{B}\mathbf{u}(k) \qquad \mathbf{u}(k) \sim \mathcal{N}(0, \Sigma^s) \tag{6}$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{s}(k) + \boldsymbol{\eta}(k) \qquad \boldsymbol{\eta}(k) \sim \mathcal{N}(0, \Sigma^y) \tag{7}$$

where $\boldsymbol{\eta} \in \mathbb{R}^p$ is the measurement noise and Σ^s and Σ^y are covariance matrices of the multivariate normal distribution $\mathcal{N}$ (Barber, 2012, p. 442). We can refer to (6) as the transition model and (7) as the emission model. Finally, we pose the problem of estimating the true trajectory $\mathbf{s}(k)$ from a series of noisy signals $\mathbf{y}(k)$ as an optimization problem of the form

$$\begin{aligned} \underset{\mathbf{s}_{0:N}, \mathbf{w}_{0:N-1}, \boldsymbol{\eta}_{0:N-1}}{\text{minimize}} \quad & \sum_{i=0}^{N-1} \|\mathbf{w}(k)\|_2^2 + \tau \sum_{i=0}^{N-1} \|\boldsymbol{\eta}(k)\|_2^2 \\ \text{subject to} \quad & \mathbf{s}(k+1) = \mathbf{A}\mathbf{s}(k) + \mathbf{B}\mathbf{u}(k), \\ & \mathbf{y}(k) = \mathbf{C}\mathbf{s}(k) + \boldsymbol{\eta}(k), \\ & k = 0, \dots, N-1. \end{aligned} \tag{8}$$

where τ is a trade-off parameter between the input and noise vectors (Diamond, Chu, Agrawal, & Boyd, 2018; Schon, Gustafsson, & Hansson, 2003).

2.2 Robust Kalman Filtering

Our final extension to the dynamical system is to render the noise component of the emission model non-Gaussian. Namely, $\boldsymbol{\eta}(k) \sim P(X)$ for some distribution $P(X)$. Specifically, we model the presence of a faulty instrument by introducing outliers to an otherwise Gaussian noise. Therefore,

$$P(X) = \begin{cases} \mathcal{N}(0, \Sigma^y), & \text{if } \rho > \rho_{out} \\ \mathcal{N}(0, \Sigma^y_{out}), & \text{otherwise} \end{cases} \tag{9}$$

for $\rho \sim \mathcal{U}(0, 1)$, where ρ_{out} is the probability of outliers and $\mathcal{N}(0, \Sigma^y_{out})$ their probability distribution. Note that the objective of (8) is sensitive to the presence of outliers due to the sum of squares of the l_2-norm. We can render the objective more robust by introducing the Huber loss function $\phi_\varrho(\mathbf{x})$, defined as

$$\phi_\varrho(\mathbf{x}) = \begin{cases} \|\mathbf{x}\|_2^2 & \text{if } \|\mathbf{x}\|_2 \leq \varrho \\ 2\varrho\|\mathbf{x}\|_2 - \varrho^2 & \text{otherwise} \end{cases} \tag{10}$$

where $\varrho \geq 0$ is the radius of a $\mathbb{R}^p$ ball (Diamond et al., 2018). $\phi_\varrho(\mathbf{x})$ grows quadratically inside the ball and linearly outside of it. We express the robust Kalman filter as an optimization problem of the form

$$\begin{aligned} \underset{\mathbf{s}_{0:N}, \mathbf{w}_{0:N-1}, \boldsymbol{\eta}_{0:N-1}}{\text{minimize}} \quad & \sum_{i=0}^{N-1} \|\mathbf{w}(k)\|_2^2 + \tau \sum_{i=0}^{N-1} \phi_\varrho(\boldsymbol{\eta}(k)) \\ \text{subject to} \quad & \mathbf{s}(k+1) = \mathbf{A}\mathbf{s}(k) + \mathbf{B}\mathbf{u}(k), \\ & \mathbf{y}(k) = \mathbf{C}\mathbf{s}(k) + \boldsymbol{\eta}(k), \\ & k = 0, \ldots, N-1, \end{aligned} \tag{11}$$

where ϱ and τ are tuning parameters. Note that there are no restrictions on the value of the decision variables, hence the feasible set is convex. Furthermore, (8) is a composition of l_2-norms, which are convex functions. Thus, (8) is a convex optimization problem since the objective and feasible set are convex. (11) includes the Huber loss function, which is a continuous concatenation of compositions of the l_2-norm. It is therefore a convex function, and the optimization problem described by (11) remains convex.

2.3 Application

The last step is to define the dynamics of the system. In the domain of vehicle tracking, we model a two-dimensional system according to the laws of kinematics. Let $\mathbf{s}(k) \in \mathbb{R}^4$ be such that the vector $[s_{1,k}, s_{2,k}]^T$ denote the position of the vehicle and $[s_{3,k}, s_{4,k}]^T$ its velocity at time k (Shekhar, n.d.). From the laws of kinematics, we know that

$$x(k+1) = x(k) + v(k)\Delta t + \frac{1}{2}a(k)\Delta t^2 \tag{12}$$

$$v(k+1) = v(k) + a(k)\Delta t \tag{13}$$

in the discrete and one-dimensional case for position, velocity, and acceleration scalars $x(k)$, $v(k)$, and $a(k)$ respectively. Now, we define the input vector $\mathbf{u}(k) \in \mathbb{R}^2$ and let the acceleration in a given axis $i = \{1, 2\}$ be determined by $a_i(k) = u_i(k) - \gamma v_i(k) = u_i(k) - \gamma s_{i+2}(k)$, since $v_i(k) = s_{i+2}(k)$. In other words, the acceleration at time k is given by the input vector minus a damping term with constant proportionality γ to the velocity (Diamond et al., 2018). Plugging the expressions of the acceleration and velocity in (12) and (13), we express the dynamics matrices as

$$A = \begin{bmatrix} 1 & 0 & (1-\frac{\gamma}{2}\Delta t)\Delta t & 0 \\ 0 & 1 & 0 & (1-\frac{\gamma}{2}\Delta t)\Delta t \\ 0 & 0 & 1-\gamma\Delta t & 0 \\ 0 & 0 & 0 & 1-\gamma\Delta t \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{2}\Delta t^2 & 0 \\ 0 & \frac{1}{2}\Delta t^2 \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

where Δt is the duration of one time step (Diamond et al., 2018).

3 Analysis

We simulate data across three experimental designs to test the effectiveness of Kalman filtering. Across all experiments, the data is generated over 1000 time steps for a total duration of 60 seconds. Table 1 summarizes the parameters of the data generation process.

The objective is that experiment 1 describe a relatively well-behaved regime, whereas in experiment 2 we increase the proportion of outliers substantially and in 3 we have both high variance on the inputs and standard measurements.

We test eight filters on each experiment. Table 2 provides the parameters of each filter.

Experiment	Σ^s	Σ^y	Σ^y_{out}	ϱ
1	I_2	I_2	$15I_2$	0.2
2	I_2	I_2	$15I_2$	0.4
3	$3I_2$	$3I_2$	$15I_2$	0.4

Table 1: Data generation parameters for the three experiments, showing the covariance matrices of inputs and noise vectors along with the probability of outliers. I_2 is the $\mathbb{R}^{2\times 2}$ identity matrix.

Parameter	1	2	3	4	5	6	7	8
Filter type	Kalman	Kalman	Kalman	Kalman	Robust	Robust	Robust	Robust
τ	0.01	0.1	1	10	1	1	10	10
ϱ	-	-	-	-	1	10	1	10

Table 2: Parameters of each of the eight filters tested. τ is the trade-off between the input and noise vectors in the objective and ϱ the radius of the $\mathbb{R}^2$ ball.

We focus on the analysis of the results of experiment 2, shown in Figure 1. The appendix contains the results across all experiments.

The row of standard Kalman filters shows that the high proportionality of outliers in the data substantially impairs the effectiveness of the filter. Furthermore, it becomes clear that τ works as a bias-variance trade-off tuner. For $\tau = 0.01$, the noise is virtually ignored and the trajectory is smoothed out. For $\tau = 10$ the opposite happens; since the noise dominates the objective, the estimated path becomes erratic. $\tau = 0.1$ best approximates the correct trajectory for the standard filter.

The robust filters far dominate the standard ones. The parametrization of the Huber loss function stands out as particularly important. When the radius ϱ of the ball is large, variance increases as many outliers still influence the objective quadratically. This effect is especially prominent when τ is large since the noise impacts the optimization function more. When the noise is effectively filtered out by a small ball, the robust filter can retrieve the correct trajectory for both $\tau = 1$ and $\tau = 10$.

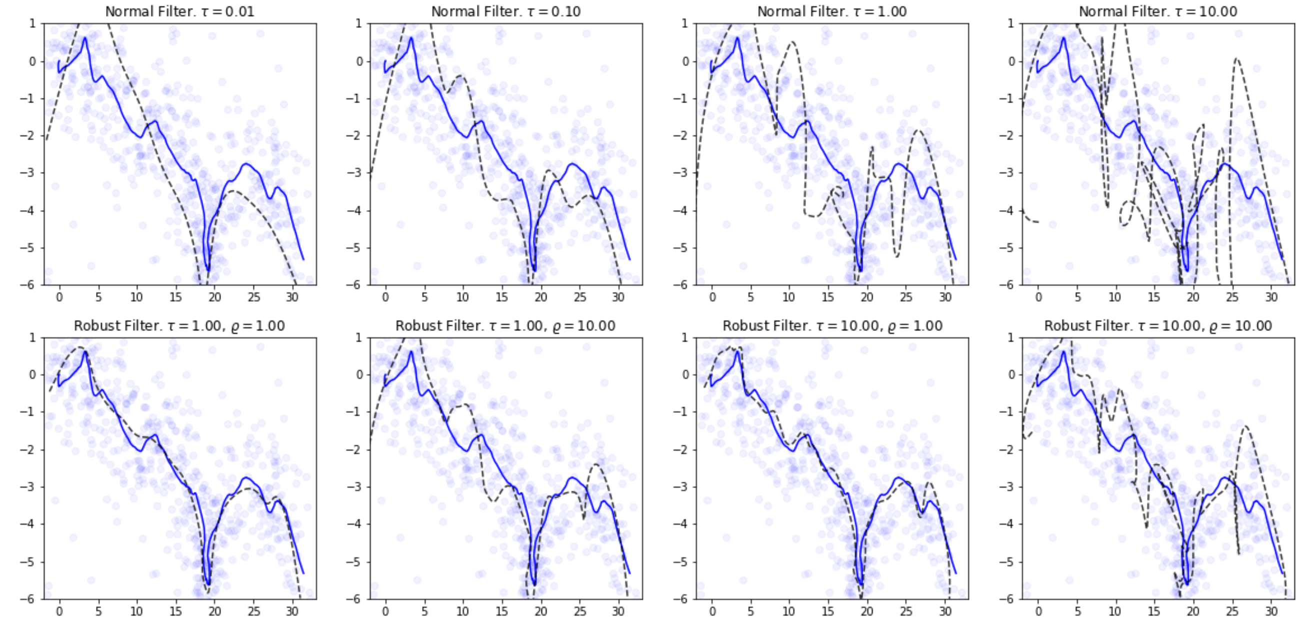

Figure 1: Eight filters applied on data from experiment 2. Solid line shows the true trajectory whereas the dashed line shows the estimated path.

References

Barber, D. (2012). Continuous-state Markov Models. In *Bayesian reasoning and machine learning*. Cambridge: Cambridge University Press.

Calafiore, G., & Ghaoui, L. E. (2014). *Optimization models*. Cambridge: Cambridge University Press.

Dahleh, M., Dahleh, M. A., & Verghese, G. (2004). Chapter 10: Discrete-Time Linear State-Space Models. In *Lectures on dynamic systems and control*.

Diamond, S., Chu, E., Agrawal, A., & Boyd, S. (2018). Robust Kalman filtering for vehicle tracking. Retrieved from `https://www.cvxpy.org/examples/applications/robust_kalman.html`.

Schon, T., Gustafsson, F., & Hansson, A. (2003). A note on state estimation as a convex optimization problem. In *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing*.

Schouwenaars, T., De Moor, B., Feron, E., & How, J. (2001). Mixed integer programming for multi-vehicle path planning. In *2001 European Control Conference (ECC)* (pp. 2603-2608).

Shekhar, R. (n.d.). CS164 Lesson 13.2 - Preclass Work Exercises. Retrieved from `https://course-resources.minerva.kgi.edu/uploaded_files/mke/Y55AwY/preclasswork-13.2.pdf`.

Appendix

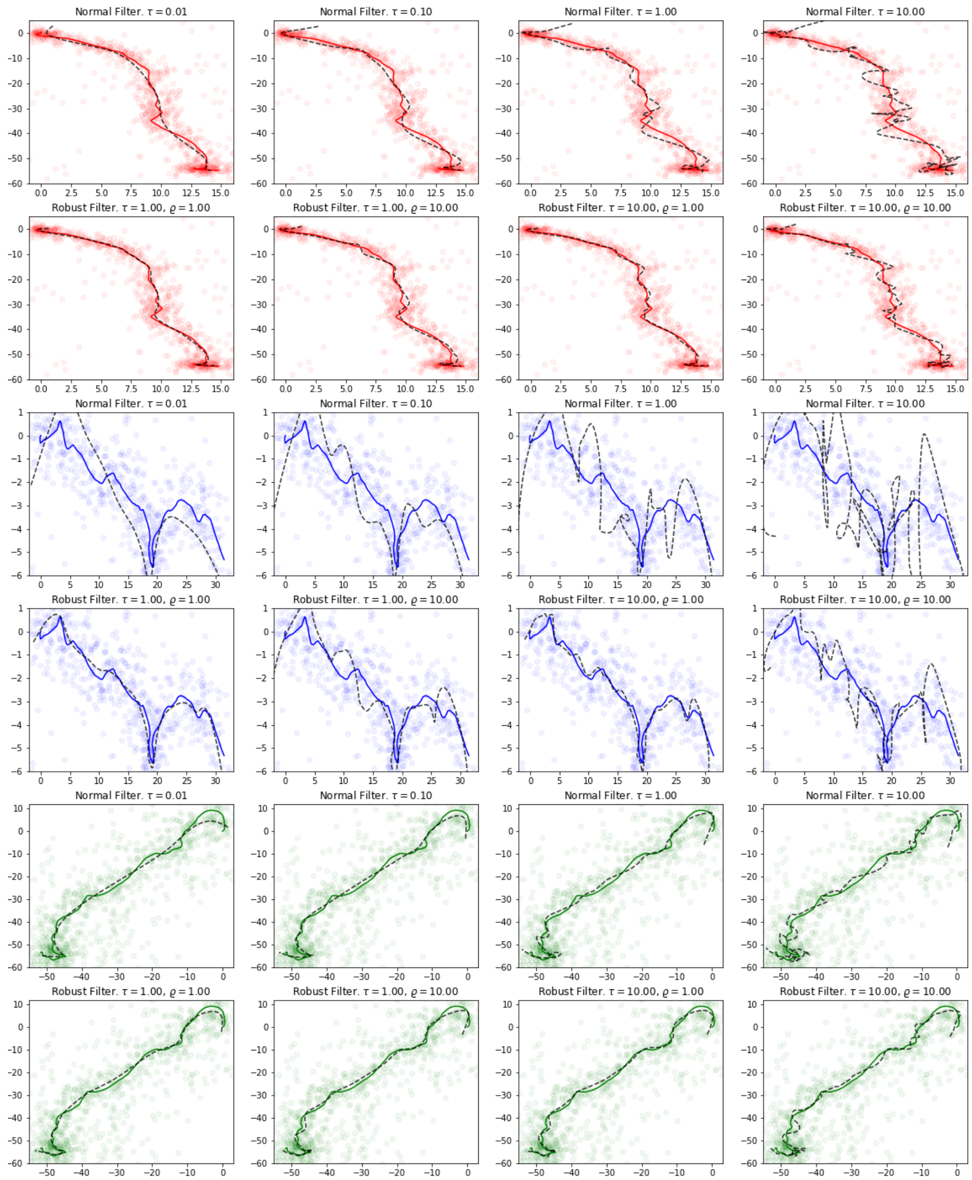

Figure 2: Results of all filters across the three experiments.