Outlier Robust State Estimation Through Smoothing on a Sliding Window

Daniela De Palma* and Giovanni Indiveri**

* Department of Innovation Engineering, University of Salento (ISME node), Italy. (e-mail: daniela.depalma@unisalento.it).
** DIBRIS, University of Genova (ISME node), Italy. (e-mail: giovanni.indiveri@unige.it).

Abstract: Measurement outliers can severely impact on the performance of conventional state estimators. The design of state estimators exhibiting enhanced robustness to measurement outliers is of interest in many areas of systems and control engineering. In marine robotics applications the issue is particularly relevant for navigation and model identification tasks exploiting acoustic based positioning and velocity sensors that are subject to relatively high rates of outliers. A sliding window state estimator is designed by minimizing the Least Median of Squares cost function evaluated by running a Rauch-Tung-Striebel smoother on the current window. The resulting estimator is tested on Doppler Velocity Log navigation data acquired on an underwater robot. Although these are only preliminary results, they confirm that the approach can be successfully used online.

Keywords: Kalman filtering techniques in marine systems control, Marine system navigation, guidance and control, Marine system identification and modelling, Filtering and smoothing, Estimation and filtering, Recursive identification.

1. INTRODUCTION

State estimation in the presence of measurement outliers is a relevant and challenging issue in many applications. Particularly so in marine robotics and oceanic engineering scenarios where acoustic measuring devices are most common. Multi-path and multi-reflection phenomena in underwater acoustic propagation may generate outliers in range and velocity measurements used for marine system navigation and localization applications. Possible approaches to cope with outliers in Ultra-Short Base Line and Long Base Line acoustic positioning systems are addressed by many authors including, by example, Vaganay et al. (1996), Bingham and Seering (2006), A. Alcocer (2006), Vasiljevic et al. (2012), Morgado et al. (2015), and Leonard and Bahr (2016). Similar problems arise when processing Doppler Velocity Logger (DVL) data for measuring velocity in marine vehicles in navigation or identification scenarios. Specific examples in this area include results by Martin and Whitcomb (2014), van de Ven et al. (2007), and Lekkas et al. (2015). Outliers may significantly impact also on cooperative robotics marine systems as accounted by Bahr et al. (2009) and Soares et al. (2013) as well as on underwater vehicle control (Caccia et al. (2003)) or diver tracking applications (Mišković et al. (2015)). It should be noticed that measurement outliers in underwater applications are not limited to acoustic transducers. Indeed, vision related processing is also potentially affected by outliers as discussed by Horgan and Toal (2006), Leone et al. (2006), and Distante and Indiveri (2011) just to name a few examples.

Fitting data to models in the presence of outliers as well as outlier identification are central topics in statistics. The related literature is extremely wide and interesting: significant textbooks addressing outlier related issues for statistical models include the ones by Hawkins (1980), Rousseeuw and Leroy (2003), and Huber and Ronchetti (2009). Although many results and ideas derived in the statistics literature are of paramount relevance in engineering applications, most of these results refer to static models and batch data processing approaches: indeed recursive algorithms as applied to dynamic (state space) models are basically absent in the statistics literature. An exception is the paper by Ruckdeschel et al. (2014) suggesting to saturate the correction term of the Kalman filter in order to prevent outliers from arbitrarily affecting the Kalman state estimate. The idea of limiting (i.e. saturating) the correction term in prediction - correction state estimation filters to limit the impact of potential measurement outliers is a quite common and simple. In marine applications a similar approach was used by Vike and Jouffroy (2005) designing a nonlinear diffusion based prediction - correction filter having an intrinsically limited correction term. The problem with such approaches as applied to linear state space estimation (like for the Kalman or Luenserger filters) is that a saturated correction term can jeopardize the filter stability, in particular when the state transition matrix is not stable. This problem was recently studied and solved by Alessandri and Zaccarian (2018) in the linear time-invariant system case, both in continuous and discrete time. In particular they designed an estimator with...
a correction term with an adaptive saturation threshold granting global exponential stability to the origin for the error dynamics.

In this paper a sliding windows approach is discussed: similarly to the solution we previously derived in (De Palma and Indiveri (2017)), the estimate is computed by minimizing an outlier robust cost function on the moving window. The novelty here is twofold: the considered cost function is the Least Median of Squares (LMS); moreover, such cost function is evaluated using a Rauch-Tung-Striebel (RTS) smoother (Rauch et al. (1965)) on the window rather than the pure prediction of the state estimate. Numerical simulation results suggest that this significantly improves the performances of the filter as compared to an implementation computing the cost function on the prediction only. Notice that the proposed filter structure could be also implemented using alternative outlier robust cost functions on the sliding window as, for example, the Least Trimmed Squares (LTS) one. Indeed the idea of exploiting the structure of the LTS parameter identification approach for state estimation is also exploited by Alessandri and Awaddeh (2016) where smoothing is not performed on the moving window.

2. BACKGROUND AND SETTING

Consider the dynamic linear model given by

\[ x_k = A_k x_{k-1} + B_k u_{k-1} + w_{k-1} \]  
\[ y_k = C_k x_k + e_k \]

where \( x_k \in \mathbb{R}^n \) is the state vector, \( u_k \in \mathbb{R}^q \) is the known input vector, \( y_k \in \mathbb{R}^p \) is the observation vector, \( w_k \in \mathbb{R}^n \) is the state noise, \( e_k \in \mathbb{R}^p \) is the observation noise, \( A_k \in \mathbb{R}^{n \times n} \) is the state transition matrix, \( B_k \in \mathbb{R}^{n \times q} \) is the input matrix and \( C_k \in \mathbb{R}^{p \times n} \) is the observation matrix at time step \( k \). Assume \( w_k \) and \( e_k \) to be zero-mean, Gaussian, white and uncorrelated noise, that is, \( w_k \sim \mathcal{N}(0,Q_k) \) and \( e_k \sim \mathcal{N}(0,R_k) \) where \( Q_k \) and \( R_k \) are the covariance matrices of the state and the observation noise, respectively. Suppose further the initial state \( x_0 \) to be a Gaussian random vector \( x_0 \sim \mathcal{N}(\bar{x}_0,P_0) \) independent of \( w_k \) and \( e_k \), where \( P_0 \) denotes its covariance matrix.

The state estimation problem consists in determining the state vector \( x_k \) given the knowledge of the model matrices (including the noise covariances), the input and output sequences up to time \( k \). In the stated hypothesis and under suitable observability conditions Simon (2006), Jazwinski (2007), Anderson and Moore (2012), the optimal (i.e. with least estimate covariance) solution is given by the Kalman filter (KF). Following Anderson and Moore (2012), the equations of the recursive KF are:

\[ \hat{x}_k^- = A_k \hat{x}_{k-1}^- + B_k u_{k-1} \]  
\[ P_k^- = A_k P_{k-1}^- A_k^T + Q_{k-1} \]  
\[ K_k = P_k^- C_k^T (C_k P_k^- C_k^T + R_k)^{-1} \]  
\[ \hat{x}_k^+ = \hat{x}_k^- + K_k (y_k - C_k \hat{x}_k^-) \]  
\[ P_k^+ = P_k^- - K_k C_k P_k^- \]

where \( \hat{x}_k^- \) denotes the estimate of \( x_k \) before processing the measurements at time \( k \), \( \hat{x}_k^+ \) denotes the estimate of \( x_k \) after processing the measurements at time \( k \), \( P_k^- \) denotes the covariance of the estimation error of \( \hat{x}_k^- \) and \( P_k^+ \) denotes the covariance of the estimation error of \( \hat{x}_k^+ \). It should be noticed that the standard Kalman filter estimates simultaneously both, the state and the error covariance matrix.

The Kalman filter is highly sensitive to outliers and has an asymptotic breakdown point of zero Rousseeuw and Leroy (2003). Indeed, if the assumption of Gaussian measurements noise is violated or the observations are contaminated by outliers, the optimality of the filter fails and its performance can deteriorate significantly. Even a single outlier in the observed data can compromise the result of the estimation. A similar situation may occur in regression and parameter identification scenarios when measurements include outliers. In such cases one may consider robust regression approaches as the Least Median of Squares (LMS) estimator described in Rousseeuw and Leroy (2003). In short, the LMS parameter estimate \( \theta_{LMS} \) is obtained by minimizing with respect to the unknown parameter \( \theta \) the median of the square residual \( r_k = y_k - y_k(\theta) \) norms being \( y_k(\theta) \) the model estimated measurement. In particular

\[ J_{LMS}(\theta) = \text{median} \{ ||r_1||^2, ||r_2||^2, \ldots , ||r_n||^2 \} \]
\[ \theta_{LMS} = \arg \min_{\theta} J_{LMS}(\theta). \]

The resulting estimator has the highest possible breakdown point, i.e. 50%. For an intuitive interpretation of the LMS, notice that in case of linear regression of data points on a line in the plane, the LMS line corresponds to the center line of the thinnest stripe of the plane containing half plus one of the data points.

3. LMS BASED ROBUST STATE ESTIMATOR

The proposed estimator consists in performing state estimation by minimizing a least median of square (LMS) cost function on a sliding window of data, i.e. according to a moving horizon strategy.

Let be \( N \) the size of the sliding window, the proposed approach is built on the derivation of a state estimate at the current time \( k \) by using the information given by the measurements and inputs within the window, i.e. \( y_{k-N+1}, \ldots , y_k; u_{k-N+1}, \ldots , u_k \) with the integer \( N \geq 1 \). The estimates of \( x_{k-N+1}, \ldots , x_k \) are based on such information and on an “estimation” \( \hat{x}_{k-N}^+ \) of the state \( x_{k-N} \) at the beginning of the moving window. In the following, the estimates of \( x_{k-N+1}, \ldots , x_k \) at time \( k \) are denoted by \( \hat{x}_{k-N+1|k}, \ldots , \hat{x}_{k,k-N+1} \). When \( k \) is an integer this is equivalent to a smoothing problem within the sliding window that can be solved with the standard Rauch-Tung-Striebel (RTS) smoother Rauch et al. (1965). The RTS smoother is implemented by first running the standard Kalman filter equations (3-7) forward in time to the current time \( k \), and then implementing the RTS smoother equations (10-15) backward in time to the initial time \( k - N + 1 \):
1. Initialize the RTS smoother as follows:
\[
\hat{x}_{k|k} = \hat{x}^+_{k|k} \quad (10)
\]
\[
P_{k|k} = P^+_{k|k} \quad (11)
\]
2. For \(i = k-1, \ldots, k-N+1 \) execute:
\[
I_{i+1} = (P^+_{i+1})^{-1} \quad (12)
\]
\[
K_i = P^+_{k|k} A_i^T I_{i+1} \quad (13)
\]
\[
P_{i|k} = P^+_{i|k} - K_i (P^+_{i+1} - P_{i+1|k}) K_i^T \quad (14)
\]
\[
\hat{x}_{i|k} = \hat{x}^+_{i|k} + K_i (\hat{x}_{i+1|k} - \hat{x}^+_{i+1} \quad (15)
\]

By seeking inspiration from the algorithms for the computation of the LMS estimator, the RTS estimation is proposed by repeatedly estimating subsamples of measurements in a sliding window. Let \( y_k = \{y_{k-N+1}, y_{k-N+2}, \ldots, y_k\} \). Subsamples are denoted as \( S_k \) where \( k \) is the current time index and \( j \) is an index running on all possible \( m \) subsamples obtained keeping only \( L \) data values out of the \( N \) ones in \( y_k \). In practice, \( m \) can be computed as the combinations without repetitions of \( L \) elements out of \( N \):
\[
m = C^L_N = \binom{N}{L} = \frac{N!}{L!(N-L)!}
\]

By example if \( N = 3 \) and \( L = 2 \) then \( m = 3 \) and at time, say, \( k = 8 \) the corresponding subsamples of \( y_k \) would be \( S^1_k = \{y_6, y_7, y_8\} \), \( S^2_k = \{y_6, y_7, y_8\} \), \( S^3_k = \{y_5, y_7, y_8\} \). For each subsample \( S^i_k \), one can determine the RTS estimates \( \hat{x}_{k-N+1|k} \), \( \hat{x}_{k-N+2|k} \), \( \ldots, \hat{x}^+_{k|k} \).

These estimates leave out the measurements not belonging to the subsample. In correspondence of such points, the standard RTS smoothing equation can still be used by assuming that the measurements have a diverging covariance \( R_i \), namely with \( R_i^{-1} \) tending to the zero matrix being \( i \) in the range \( [k - N + 1, k] \).

Having computed the RTS estimates, for each subsample \( S^i_k \), one can evaluate the median of the square residuals cost
\[
J^S_k(\hat{x}^+_{k-N}) = \min_{i=k-N+1,...,k} \left\| y_i - C_i S^i_k \right\|^2 (16)
\]
and select the subsample \( S^i_k \) with the least median of squares (LMS) cost, namely
\[
S^i_k = \arg \min_{j=1,\ldots,m} J^S_k(\hat{x}^+_{j-k-N}) (17)
\]
that allows to identify the current state estimate \( \hat{x}^+_k \) and its covariance \( P^+_{k|k} \) as \( \hat{x}^+_k := S^i_k \) and \( P^+_{k|k} := P^S_k \).

The LMS criterion to select the subsample \( S^i_k \) aims at excluding outliers from the measurement set used in the RTS smoothing phase. The computational effort of this approach is strongly dependent on the sliding window size \( N \) and on the number \( m \) of subsample sets where to compute the RTS estimator. If the inliers every \( N \) measurements are actually at least \( L \), the proposed approach has the potential (albeit with no formal guarantee) to exclude all the outliers. Figure 1 reports an example of the described algorithm with subsamples obtained considering \( N = 5 \) and \( L = 3 \) (hence \( m = 10 \)).

The pseudo-code for the proposed outlier robust state estimation is reported in algorithm 1.

As a side result the proposed approach allows to eventually rank the measurements according to the number of times they belong to an optimal subset minimizing the median of squares. Indeed when the sliding window of size \( N \) is shifted forward \( N - 1 \) measurements of the new window were also included in the previous window. As the window is shifted further, due to the overlap of successive windows, a certain number of measurements are re-evaluated multiple times. One can thus label measurements according to the number of times they are accepted in the (local) Least Median of Squares optimal subsets. Measurements that are more frequently accepted in the local optimal subsets according to the LMS criteria are candidate inliers. Eventually such data can be used to run an estimator over the whole batch of data or over longer windows than the \( N \) data points one used locally as previously described. Details of this extension are not included in this paper for the sake of brevity and will be subject to future work.

**Algorithm 1 LMS-RTS based estimation algorithm**

**Require:** \( L, N, \hat{x}^+_{k-N}, P^+_{k-N}, y_{k-N+1}, \ldots, y_k, u_{k-N+1}, \ldots, u_k \)

**Ensure:** \( \hat{x}^*_{k|k}, P^*_{k|k} \)

1. for \( j = 1 : m \)
2. compute the RTS smoothing estimate on mobile window using (10-15):
\[
\hat{x}^S_{k-N+1|k}, \hat{x}^S_{k-N+2|k}, \ldots, \hat{x}^S_{k|k}
\]
\[
P^S_{k-N+1|k}, P^S_{k-N+2|k}, \ldots, P^S_{k|k}
\]
3. compute the LMS cost \( J^S_k(\hat{x}^+_{k-N}) \) using (16)
4. end for
5. \( S^i_k \leftarrow \arg \min_{j=1,\ldots,m} J^S_k(\hat{x}^+_j) \)
6. \( \hat{x}^*_k, P^*_{k|k} \leftarrow \hat{x}^S_{k|k}, P^S_{k|k} \)
7. return \( \hat{x}^*_k, P^*_{k|k} \)

4. EXAMPLE ON MARINE ROBOTICS APPLICATIONS.

A specific example relative to the navigation of an underwater robot is reported. The experiment is performed employing an Underwater Vehicle Manipulator System (UVMS) developed within the ROBUST project ROBUST (2015-2020). The UVMS is composed by three Autonomous Underwater Vehicles (AUVs), each one equipped with 4 tunnel thrusters (2 vertical and 2 lateral at the bow and stern of the AUVs respectively), and 1 main rear thruster. During the experiment, the UVMS is controlled to obtain a motion along the surge axes at constant speed, and its velocity is acquired using a DVL sensor together with the thrusters commands. The velocity starts from zero and increases gradually until reaching a steady state value. The following dynamic equation for surge velocity is considered:
\[
m \ddot{v} + D_l v + D_q v |v| = \tau
\]
where \( v \) denotes the surge velocity, \( m \) the mass of the vehicle (including added mass), \( D_l \) and \( D_q \) the linear and quadratic surge drag coefficients, \( \tau \) the surge actuation term. The latter can be expressed as:
\[ Y_k = \{y_{k-4},y_{k-3},y_{k-2},y_{k-1},y_k\} \]

\[ S^1_k = \{y_{k-4},y_{k-3},y_{k-2}\} \rightarrow J^{S^1_k} = \text{median}_{i=k-4,\ldots,k} \| y_i - C_i \hat{x}_{i|k} \|^2 \]

\[ S^2_k = \{y_{k-4},y_{k-3}\} \rightarrow J^{S^2_k} = \vdots \]

\[ S^3_k = \{y_{k-4},y_{k-2}\} \rightarrow J^{S^3_k} = \vdots \]

\[ S^4_k = \{y_{k-4},y_k\} \rightarrow J^{S^4_k} = \vdots \]

\[ S^5_k = \{y_{k-4},y_{k-1}\} \rightarrow J^{S^5_k} = \vdots \]

\[ S^6_k = \{y_{k-3},y_{k-2}\} \rightarrow J^{S^6_k} = \vdots \]

\[ S^7_k = \{y_{k-3},y_k\} \rightarrow J^{S^7_k} = \vdots \]

\[ S^8_k = \{y_{k-3},y_{k-1}\} \rightarrow J^{S^8_k} = \vdots \]

\[ S^9_k = \{y_{k-2},y_{k-1}\} \rightarrow J^{S^9_k} = \text{median}_{i=k-4,\ldots,k} \| y_i - C_i \hat{x}_{i|k} \|^2 \]

\[ \min_{j=1,\ldots,10} J^{S^j_k} \rightarrow \hat{x}_{k|k} \]

Fig. 1. Example of subsamples with \( N = 5 \) and \( L = 3 \) (\( m = 10 \)).

\[ \tau = G \left[ \begin{array}{c}
\alpha_1 n_1 | n_1 | \\
\vdots \\
\alpha_n n_n | n_n |
\end{array} \right], \quad (19) \]

where \( n_i, i = 1,\ldots,15 \) represent the thruster revolutions per minute (RPM) values, while the matrix \( G \in \mathbb{R}^{1 \times 15} \) depends on the relative position of each thruster with respect to the body frame, and on the thruster characteristics. A quadratic relationship has been assumed between the thruster’s RPM values and its exerted force, neglecting other hydrodynamic terms in the thruster modelling depending, by example, on the water flow velocity through the propeller blades. With regards to the knowledge of the added mass, hydrodynamic parameters, and coefficients in (18, 19), we resort to the values estimated and experimentally identified exploiting the method in Ingrosso et al. (2019). Note that for the sake of simplicity, in order to approximate the non linear model (18) into a linear one, the hydrodynamic drag has been modelled with a linear term only leading to the following equation:

\[ m \ddot{v} + D \dot{v} = \tau \quad (20) \]

with \( D \) chosen as \( D = D_1 + D_2 \dot{v} \), being \( \dot{v} \) the mean value of the vehicle velocity during the experiment. Assuming to explicitly account for noise, and discretizing the system with a sampling time \( T_s \), the equation (20) can be reformulated in a linear time invariant (LTI) state space setting:

\[ x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}, \quad (21) \]

\[ y_k = Cx_k + \epsilon_k, \quad (22) \]

having defined the state as \( x = v \), the input as \( u = \tau \), and

\[ A = 1 - \frac{D}{m} T_s, \quad B = \frac{T_s}{m}, \quad C = 1. \quad (23) \]

The covariances of the state and the observations are \( Q = (10^{-3} \text{m/s})^2 \) and \( R = (5 \cdot 10^{-3} \text{m/s})^2 \), respectively. In order to validate the performance of the proposed algorithm both in absence and in presence of outliers, we firstly considered an experimental data set free of outliers. A sliding window of size \( N = 9 \) with \( L = 5 \) and \( m = 126 \) is considered and the proposed LMS-RTS algorithm is applied. The resulting estimates are shown in figure 2. If compared with a standard KF, the LMS-RTS algorithm does not exhibit degraded performances. Then, a data set corrupted by outlier has been considered. Outliers can occur, for example, when the vehicle is very close to the sea bottom due to multi-path phenomena. Figure 3 shows an example of real data set of DVL measurements acquired from the UVMS, it should be noted that the measurement outliers occurred when the depth of the vehicle was maximum, i.e. when the vehicle was close to the sea bottom.

In order to compare the behaviour of the proposed filter with the KF estimate in the presence of measurement outliers, some artificial ones are added to the same data set.
set trying to replicate the typical frequency and entity of outliers in DVL measurements.

Fig. 3. Example of outliers in a real data set of DVL measurements acquired from the UVMS and corresponding depth of the UVMS.

The results of this experimental example are plotted in 4. The comparison with the KF estimate confirms the enhanced ability to cope with measurement outliers at the cost of an additional computational effort arising from the implementation of the smoother and the LMS minimization. From a computational point of view, Table 1 provides the amount of time required to compute the Kalman and RTS-LMS estimates (pseudo-code in Algorithm 1), respectively. The algorithms have been coded in MATLAB® (version R2017a) on an Apple Laptop with a 3.1 GHz Intel Core i7 processor, 16 GB RAM, running the MAC OS X Version 10.11.6 operating system. Also notice that our MATLAB implementations, for both filters, were not specifically optimized for execution time. Of course, by using dedicated hardware and a suitably optimized code, the execution time can be significantly less.

<table>
<thead>
<tr>
<th>Time to process</th>
<th>Average time per step</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kalman Filter</td>
<td>0.02 s</td>
</tr>
<tr>
<td>LMS-RTS Filter</td>
<td>8.6 s</td>
</tr>
</tbody>
</table>

Table 1. Computational cost of Kalman and LMS-RTS algorithm.

5. CONCLUSION

A sliding window state estimation filter is designed for a linear state space model aiming to achieve enhanced robustness to measurement outliers as compared to conventional Luenberger or Kalman filters. The novel contribution is relative to the use of a Least Median of Squares cost function on a moving horizon and its evaluation on the residuals arising from performing smoothing on the same window rather than estimation only. This allows to better highlight outlying data as due to smoothing on the moving window an outlier eventually entering at the current time step in the window will affect the past smoothed state estimates in the window and not only the future ones. Outlier robustness stems from the fact that not all measurements are processed on the moving window, but only the ones such that the median of squared residuals is minimized on the window. The identification of the specific subsample corresponding to the least median of squares cost on the window is performed through an exhaustive search. This can be quite demanding from a computational point of view as the number of subsamples to evaluate rapidly grows with the window size. Window size and the least number of candidate inliers in the window are tuning parameters. If the true number of inliers is actually larger or equal than the expected one, the proposed solution has the potential to reject all the outliers, otherwise not. The study is motivated by the need to cope with measurement outliers frequently arising in marine robotics navigation and localization scenarios. Numerical results exploiting experimental DVL measurements are illustrated.

ACKNOWLEDGEMENTS

The authors are very grateful to Dr. Enrico Simetti of the ISME node at the University of Genova (Italy) for having made available the DVL experimental data used to validate the proposed solution.

REFERENCES


Berlin, Germany, July 12-17, 2020

Preprints of the 21st IFAC World Congress (Virtual)


