SHOULD DELAYED MEASUREMENTS ALWAYS BE INCORPORATED IN FILTERING?

[1]Dimitris K. Tasoulis, [2]Niall M. Adams, and [1,2]David J. Hand

[1]Institute for Mathematical Sciences, [2] Department of Mathematics
Imperial College London, South Kensington Campus
London SW7 2AZ, United Kingdom

ABSTRACT

We consider situations in which the measurements an agent expects to receive are prone to delay. A number of procedures have been proposed to handle such delays. These procedures provide extended frameworks that incorporate delayed measurement once available. In a simple context, we explore the performance of these algorithms, and find that we can reduce the computational demands of these methods while retaining good performance.

Index Terms— Delayed measurements, out-of-sequence measurements, discrete time Kalman filter

1. INTRODUCTION

Most work on tracking and filtering is built on the standard assumption that measurements are available immediately to an agent. However, it is not difficult to conceive situations in which measurements are subject to non-negligible delays, such that the lag between measurement and receipt is of sufficient magnitude to impact on estimation.

This paper is concerned with methods for handling delayed measurements in the context of the simple discrete-time Kalman filter. The *delay mechanics*, the manner in which measurements are delayed, provides a fundamental distinction. The problem of *constant delay* involves every measurement being delayed by the same, constant lag. In this way measurements are never observed out of sequence, they are simply and consistently late. Such behaviour could be induced, for example, by a constant bandwidth restriction on a sensor network. In contrast, *random delays* provides for a number of possibilities, including that measurements are delayed with a constant probability but fixed lag, or constant probability and random lag. Such problems could arise as a result of intermittent bandwidth restrictions on a sensor network. All modes of random delay have the potential to cause *out-of-sequence* measurements (for example, see [1]). Both types of delay mechanics are illustrated in figure 1.

This research was undertaken as part of the ALADDIN (Autonomous Learning Agents for Decentralised Data and Information Systems) project and is jointly funded by a BAE Systems and EPSRC (Engineering and Physical Research Council) strategic partnership, under EPSRC grant EP/C548051/1.

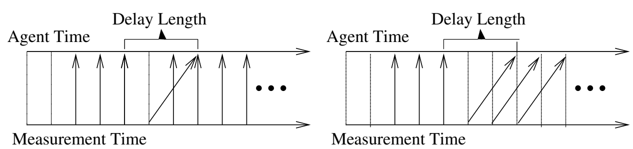

Fig. 1. Delay mechanics. Left: random delay. Right: Constant delay

In § 2 we give a more precise description of the problem. Much of the literature devoted to delay problems considers specific examples of delay mechanics and derives optimal (typically in the sense of minimum variance) estimators. Three of these methods are briefly described in § 3. We restrict attention to the case of the whole measurement vector being delayed. Some approaches are more general, treating the measurement vector as a collection of subvectors subject to independent delays. The more challenging problem of elementwise delay is explored in [2].

However, as fas as we are aware no comparative review or study has ever been conducted about the delay problem, in any context. Our extensive simulations suggest that there is often little to choose between these methods. A particular feature is that it is sometimes the case that incorporating delayed measurements is not useful. In this paper, we describe methods for determining whether to incorporate delayed measurements in these algorithms. In §4 we describe results of these methods in simulation. Finally, in § 5 we make some concluding remarks.

2. FRAMEWORK

The standard Gaussian formulation for the discrete-time dynamic linear model, relates the measurement vector y_t to the unobserved system vector x_t with the equation

$$y_t = C_t x_t + \epsilon_t \tag{1}$$

for $t = 0, 1, \ldots$, where the random vector innovation $\epsilon_t \sim N(0, R_t)$. The evolution of the system vector follows a random walk

$$x_t = \phi_t x_{t-1} + \nu_t \tag{2}$$

where the random vectors $\nu_t \sim N(0, Q_t)$, and both ϵ_t and ν_t are serially uncorrelated. In this work we regard the measurement and system covariances, R_t and Q_t respectively, as known. A useful concept with such a system is the Signal-to-noise ratio (SNR) defined as the ratio of a given signal to the background noise of the transmission medium. In the formulation above we define the SNR as $\frac{\|Q_t\|}{\|R_t\|}$.

The Kalman filter estimates the process x, by using a form of feedback control: the filter estimates the process state at some time and then obtains feedback in the form of the (noisy) measurements y. We can therefore describe the Kalman filter operation with two steps; the *prediction* step and the *correction* step [3]. The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the a priori estimates for the next time step t, using all the previously received measurements $(t-1)$:

$$\textbf{Prediction} \quad \textbf{Step}$$

$$\hat{x}_{t|t-1} = \Phi_t \hat{x}_{t-1|t-1} \tag{3}$$

$$P_{t|t-1} = \Phi_t P_{t-1|t-1} \Phi_t^{\top} + Q_t \tag{4}$$

where $P_{t|t-1}$ denotes the a priori estimate for the error covariance at time t, give all the information available up to time $t-1$. The measurement update equations are responsible for fusing the new measurement y_t into the a priori estimates to obtain improved a posteriori estimates:

$$\textbf{Correction} \quad \textbf{Step}$$

$$K_t = P_{t|t-1} C_t^{\top} (C_t P_{t|t-1} C_t^{\top} + R_t)^{-1} \tag{5}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t (y_t - C_t \hat{x}_{t|t-1}) \tag{6}$$

$$P_{t|t} = (I - K_t C_t) P_{t|t-1} \tag{7}$$

The difference $(y_t - C_t \hat{x}_{t|t-1})$ in Eq.6, is called the measurement innovation, or the residual, and reflects the discrepancy between the predicted $(C_t \hat{x}_{t|t-1})$ and the actual measurement (y_t). The K matrix is called the *Kalman gain* or *blending factor*, and it is chosen as such to minimize the a posteriori error covariance.

In the context of this framework delay can be characterised as follows. First, consider constant delay. The measurement vector at time t, y_t, is not available to the agent until y_{t+d} (that is, it has *lag* d), for $d \in \{0, 1, 2, \ldots\}$ and all t. Note that this includes both the non-delayed situation and the problem of missing values as special cases. Next, consider random delays. We still have y_{t+d}, but d is a realisation of some lag random variable D. Different modes of random delay arise from different probability distributions for D. For example, consider the Bernoulli random variable D, with range $\{0, d\}$, for constant d, and

$$P(D = 0) = \theta \qquad P(D = d) = 1 - \theta$$

This is a random probability of delay, with a constant lag. Alternatively, consider $D \in \{0, 1, 2, \ldots\}$, with some probability mass function f providing $P(D = d) = f(f; \theta)$. This is an example of random delay, with random lag. We restrict attention to homogeneous delay mechanics, although non-homogeneous cases are also of interest. Note that random delays have the potential to cause no measurement be observed at some time point, and in consequence, more than one measurement to be observed at a later time point. For convenience, we ignore the potential operational problems this could imply.

3. METHODS

All methods proposed for the delay problem, including [4, 5, 6, 7, 8], have in common that delayed observations are always ultimately incorporated into the filtering process. We have extensive simulation evidence that suggests these methods provide similar empirical performance under varying characteristics of delay mechanics. Their performance however, is heavily depended on the SNR. Our simulations suggest that when the SNR is high, there is little value in incorporating delayed measurements. This should be unsurprising, since it is simply a consequence of data changing so rapidly that delayed data is no longer relevant.

Perhaps the crudest approach is simply to use k-step ahead forecast equations to fill in delayed values, and never to attempt to subsequently incorporate delayed measurements, when they become known. If a measurement is delayed, the whole correction step can be skipped, and the operation of the filter continues normally to the next prediction step. For the general case of subsequent k delayed measurements, a k-step ahead forecast is performed, up until new measurements start to be received. We refer to this approach as MKF and use it is as part of our performance assessment in the next section, since we certainly need to out-perform at least this approach.

In the context of known measurement and system covariance matrices, we can construct thresholding procedures, such that delayed measurements that do not alter the estimates significantly, when they are eventually observed, are disregarded. Such a decision can be made even before a delayed measurement is received by examining the variance of the innovations for the delayed measurements.

We consider three methods. The first, the measurement extrapolation Kalman filter (MEKF) is described in [9]. The MEKF filter essentially computes a correction term that is added to the state estimation when the delayed measurement arrives.

The second algorithm we consider, due to [1], operates in a similar fashion to MEKF. This approach, denoted ZKF, is designed to have minimum storage requirements, in response to the large storage requirements of other methods.

The final method, ASKF uses state augmentation, considered the classical approach for delay problems [10]. State augmentation proceeds by increasing the state space representation to accommodate delayed measurements. In detail the state of the system is now augmented as

$$x_t^\alpha = [x_t^\top, x_{t-1}^\top, \ldots, x_{t-\Delta}^\top],$$

where Δ is an upper bound for the delay. If we suppose that δ_t is the delay of the tth measurement, the measurement vector at each time t becomes $Y_t = (y_{s\in S(t)})$, where $S(t) = \{s : \delta_s + s = t\}$. This way constructing augmented forms for the matrices C, ϕ, R, Q as described [10], leads to an algorithm similar to a standard Kalman filter.

The State Augmentation approach is the most general one, since it can handle all types of delay mechanics. Additionally, the whole algorithmic description is quite clear, a fact that allows easy implementation. However the computational complexity and storage requirement of State Augmentation increases linearly with with the lag. The ZKF algorithm, by construction, has a steady storage requirement, irrespective of the lag. However the storage as well as the computational requirement grow quadratically with the number of simultaneous delayed measurements. Measurement Extrapolation can also be regarded as efficient with respect to the extra storage it requires, as it is steady and unaffected by the delay length. However, the method is not generalised by [9] to the case simultaneous delayed measurements.

Now, we can consider taking each of these methods, and introducing a threshold such that the delayed measurement is only incorporated if the threshold is exceeded. In the case of the MEKF algorithm if we denote with T_{MEKF}, the covariance matrix of the delayed measurement to be incorporated at time step k, for a previously delayed measurement y_s, for d time ticks:

$$\begin{aligned} T_{\mathrm{MEKF}} &= Var(K * (y_s^{\mathrm{int}} - C_k\hat{x}_k)) \\ &= K * (C_k * P_k * C_k^\top + R_k) * K^\top, \end{aligned} \tag{8}$$

that is updated from the moment a measurement is delayed up until either the measurement arrives or the threshold condition is satisfied. The matrix K is computed as follows

$$K_k = M_* P_{s|s-1} C_s^\top [C_s P_{s|s-1} C_s^\top + R_k]^{-1} \tag{9}$$

where as analytically described in [9]:

$$M_* = \prod_{i=0}^{d-1} (I - K_{k-i}C_{k-i})\Phi_{k-i-1}. \tag{10}$$

The derivation of y_s^{int} is also described in [9].

The thresholding can be applied at each time the matrix T_{MEKF} is updated, using the respective components of the product in Eq. (10), to calculate K. For the ZKF algorithm the respective matrix is:

$$\begin{aligned} T_{\mathrm{ZKF}} &= Var(K_s(z_s - C_s\hat{x}_{s|k}) \\ &= K_s(C_s P_{s|k} C_s^\top + R_s) K_s^\top \end{aligned} \tag{11}$$

where the K_s, C_s and $P_{s|k}$ matrices are computed through the recursion proposed in [1].

For the ASKF algorithm, however, a different thresholding procedure can be used, by setting the size of the state augmentation such that only the delayed measurements with lag smaller than the augmentation size are incorporated. The delays with larger lag will be automatically disregarded as missing values.

We conducted an experiment to examine how the threshold value in the MEKF, ZKF case, and the augmentation size in the ASKF, affects the performance of the algorithms. We should note that in the case of MEKF, and ZKF, the "fusing or not" decision, is made by comparing the value of the norms $\|T_{\mathrm{MEKF}}\|$ and $\|T_{\mathrm{ZKF}}\|$ respectively with the threshold, while for the ASKF on the delay length.

4. EXPERIMENT

Since we are interested in the complete measurement vector being delayed, it is sufficient to restrict attention to the case of a univariate system equation, evolving as a random walk, and a univariate observation equation, both as per equations (2) and (1) with known error covariances. Our experiments have constant delay probability, θ, and lag uniformly distributed on the integers 1 to 10.

There are two aspects of performance here. First we need to measure the performance of the filtering method. In our simulations, we refer to the *error* of a delay algorithm as the mean absolute difference between the state estimates of the delay algorithm and the state estimates of the baseline Kalman filter, estimated on the true undelayed data. Our performance metric then computes the ratio of the error measure to the error of the MKF. The second aspect we have to account for is the impact of incorporating delayed measurements. We account for this by tracking the proportion of delayed measurements that are actually incorporated.

The plots in Figure 2 display the ratio of the performance measure on the vertical axis and the ratio of delayed measurements on the horizontal axis. The left-hand side plot refers to SNR=0.05, while the right-hand side refers to SNR=0.25. In general, we see that performance degrades when delayed measurements are not incorporated. However, the nature of this degradation appears to be related to the SNR. For the smaller SNR, all algorithms exhibit similar behaviour, and we see that 80% performance can be obtained incorporating as few as 60% of delayed measurements. For the greater SNR the degradation of performance is less marked although the ZKF generally degrades faster than the other methods. However, 80% performance can still be achieved incorporating as few as 50% of delayed measurements.

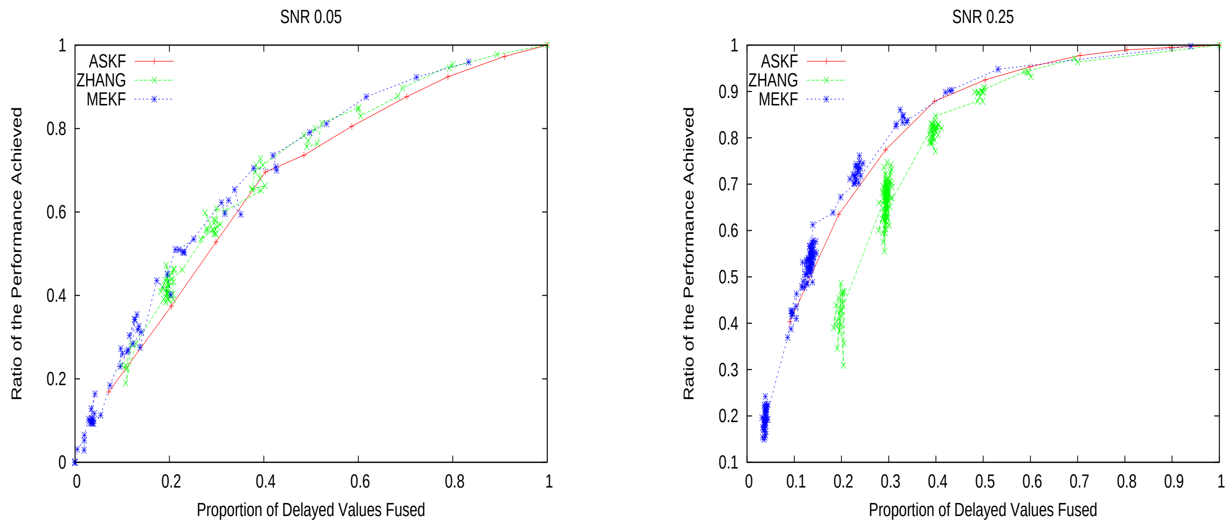

Fig. 2. The effect of ratio of the delays fused on the algorithms performance for a random delay length between 1 and 10.

5. CONCLUSIONS

We have outlined aspects of the delay problem and some solutions proposed to handle the problem. We note that the characteristic policy adopted by such solutions — always incorporate delayed data — has theoretical optimality, but induces both computational and storage overheads. In the simple context described above, we explore the merits of not always incorporating delayed measurements - only incorporating measurements that are in some sense informative. While performance necessarily degrades, we find that it degrades gracefully, and we can often obtain good performance for less work.

In the disaster recovery scenarios that are the concern of the ALADDIN project (`www.aladdinproject.org`), the capacity for an agent to retain a good model of the world while reducing the resource demands, is very desirable.

6. REFERENCES

[1] K. Zhang, X.R. Li, and Y. Zhu, "Optimal update with out-of-sequence measurements," *IEEE Transactions on Signal Processing*, vol. 53, no. 6, pp. 1992– 2004, 2005.

[2] Dimitris K. Tasoulis, Niall M. Adams, and David J. Hand, "Component-wise delay in discrete time filtering," Tech. Rep., Institute for Mathematical Sciences, Imperial College London, 2007.

[3] R.E. Kalman, "A new approach to linear filtering and prediction problems," *Transaction of the ASME–Journal of Basic Engineering*, pp. 35–45, 1960.

[4] K. Yamanaka, K. Uchida, and E. Shimemura, "Optimal control of systems with random delay," *Internat. J. Control*, vol. 29, no. 3, pp. 489–495, 1979.

[5] A. Ray, "Output feedback control under randomly varying distributed delays," *J. Guidance, Control Dynamics*, vol. 17, no. 4, pp. 701–711, 1994.

[6] Y. Bar-Shalom and T.E. Fortmann, *Tracking and Data Association*, Academic Press, New York, 1988.

[7] S. Challa, R.J. Evans, and X. Wang, "A Bayesian solution and its approximations to out-of-sequence measurement problems," *Information Fusion*, vol. 4, pp. 185–199, 2003.

[8] A.S. Matveev and A.V. Savkin, "The problem of state estimation via asynchronous communication channels with irregular transmission times," *IEEE Transactions on Automatic Control*, vol. 48, no. 4, pp. 670–676, 2003.

[9] T. Larsen, N. Poulsen, N. Andersen, and O. Ravn, "Incorporation of time delayed measurements in a discrete-time kalman filter," in *CDC '98, Tampa, Florida, 1998.*, 1998.

[10] X. Lu, H.S. Zhang, W. Wang, and K.L. Teo, "Kalman filtering for multiple time-delay systems," *Automatica*, vol. 41, no. 8, pp. 1455–1461, 2005.