Adaptive Dynamic Modelling for Kinematic Positioning

M. Moore, J. Wang
Satellite Navigation & Positioning Group
School of Surveying & Spatial Information Systems
The University of New South Wales, Sydney, NSW 2052, AUSTRALIA

Abstract. In order to optimize the dynamic model, careful consideration must be undertaken of the dynamics likely to be experienced by the system, while accounting for the measurement rate. The main problems faced by a dynamic model in a Kalman filter occur when the system experiences unexpected dynamic conditions, a change in data acquisition rate, or when in fact the dynamics of the system are non-linear. To minimize the errors produced from dynamic modelling in unusual conditions, a number of strategies can be employed. This paper investigates a range of these strategies that either try to compensate for the error in the dynamic model, or adapt the dynamic model to better suit the new conditions. The results from this study show that a constant acceleration model can be adapted to perform in the same manner as a constant velocity model. Also shown in this paper is that an adaptive variable dimensional dynamic model technique is superior to the others investigated, as it ensures the optimal performance of the Kalman filter and generation of reliable positioning results.

Keywords: Kalman filtering, Adaptive Dynamic Modelling, Kinematic Positioning.

1 Introduction

1.1 Dynamic Modelling

In Kalman filtering ultimately the accuracy of the state parameters will be dependent upon the accuracy of the measurements and the adequacy of the model assumptions. The accuracy of the measurements is most critical when moving from one kinematic representation to another. For instance accelerating from a constant velocity to a constant acceleration model involves numerical differentiation, which is an unstable process and therefore the noise level will be amplified. Conversely, decelerating from a constant acceleration to a constant velocity model involves numerical integration, which is a stable process but is very sensitive to errors in the initial values.

An adequate dynamic model for the problem at hand will be able to make the most use of the discrete noisy measurements of position from different sensors, and then combine them in such a way that time and position can be related unambiguously. This will enable interpolation between epochs for extra measurements.

The positions at time t_1 and t_{i+1} are related by:

$$p_{i+1} = p_i + \dot{p}_i \Delta t + \ddot{p}_i \frac{\Delta t^2}{2} + \dddot{p}_i \frac{\Delta t^3}{6} + \ldots \quad (1)$$

where the number of dots above the vector indicates the order of differentiation.

The above equation shows that for a proper position interpolation all of the derivatives of the position vector p are needed. In reality these are not available, and the problem is usually approximated by a truncated series, which either uses a constant velocity model:

$$p_{i+1} = p_i + \dot{p}_i \Delta t \quad (2)$$

or a constant acceleration model:

$$p_{i+1} = p_i + \dot{p}_i \Delta t + \ddot{p}_i \frac{\Delta t^2}{2} \quad (3)$$

where p and $\dot{p}$ are considered constant for the time interval Δt.

The accuracy of these assumptions will of course depend upon the length of Δt and the actual linearity in dynamics of the object. The dynamic model will be defined in terms of its state vector x, transition matrix **Φ,** and the process noise covariance matrix Q, while τ_k is the random error vector. These matrices are all required in order to implement the Kalman filter algorithm, which may be subdivided into two sections, the prediction and the filtering phase.

Prediction phase:

$$x_k = \Phi_{k,k-1} x_{k-1} + \tau_k \quad (4)$$

$$Q_{\bar{x}_k} = \Phi_{k,k-1} Q_{\hat{x}_{k-1}} \Phi^T_{k,k-1} + Q_k \quad (5)$$

Filtering Phase:

$$\hat{x}_k = (A_k^T P_k A_k)^{-1} A_k^T P_k l_k = \bar{x}_k + G_k d_k, \quad (6)$$

$$Q_{\hat{x}_k} = \Phi_{k,k-1} Q_{\hat{x}_{k-1}} \Phi_{k,k-1}^T + Q_k \quad (7)$$

where G_k is the *gain matrix*; d_k is the *innovation vector* and Q_{dk} is its covariance matrix, which are described as:

$$G_k = Q_{\bar{x}_k} A_k^T Q_{d_k}^{-1}, \quad (8)$$

$$d_k = z_k - A_k \bar{x}_k, \quad (9)$$

$$Q_{d_k} = R_k + A_k Q_{\bar{x}_k} A_k^T \quad (10)$$

where $\bar{x}_k$ is the predicted state, and $\hat{x}_k$ is the optimal estimator of the state parameters from the previous epoch (k-1), and z_k is the observation vector at time k.

1.2 Constant Velocity Model

For the purposes of this paper we will set the state vector to contain two positional states {E, N} and two velocity states $\{v_E, v_N\}$ giving:

$$x_4 = \{E, N, v_e, v_N\}$$

where v_E and v_N represent vehicle velocities in the east and north directions respectively.

The corresponding transition matrix is of the form (Schwarz et al., 1989):

$$\Phi_4 = \begin{pmatrix} I & \vdots & SD \\ \cdots & \cdots & \cdots \\ 0 & \vdots & T \end{pmatrix} \quad (11)$$

where all the submatrices are diagonal and of dimension (2x2). Their diagonal elements are of the form:

$$S = diag\{s_i\} = diag\{\alpha^{-1}(1 - t_i)\} \quad (12)$$

$$T = diag\{t_i\} = diag\{e^{-\alpha_i \Delta t}\} \quad (13)$$

$$D = diag\{d_i\} = diag\{1;1;1;1\} \quad (14)$$

S and T take different values of α. The specific form of T suggests that the corresponding elements of the state vector can be modelled as 1st order Gauss-Markov processes satisfying a differential equation of the type:

$$\dot{x} = -\alpha x + w \quad (15)$$

where α is the inverse of the correlation length and w is a white noise process. To have a short correlation length, which will allow a large change in x between epochs, the alpha value should be set to a large value, and vice versa. If a strong correlation value is set then the transition matrix can be approximated by:

$$\Phi_4 = \begin{pmatrix} I & \vdots & D\Delta t \\ \cdots & \vdots & \cdots \\ 0 & \vdots & I \end{pmatrix} \quad (16)$$

1.3 Constant Acceleration Model

For a constant acceleration model an additional two acceleration states are required, an acceleration in easting, and northing (a_E, a_N) giving:

$$x_6 = \{E, N, v_E, v_N, a_E, a_N\}$$

Using the Gauss-Markov approach again, the transition matrix becomes(Schwarz et al., 1989):

$$\Phi_6 = \begin{pmatrix} I & SD & UD \\ 0 & I & S \\ 0 & 0 & T \end{pmatrix} \quad (17)$$

where U is a diagonal matrix with elements:

$$U = diag\{u_i\} = \frac{\{e^{-\alpha_i \Delta t} + \alpha_i \Delta t - 1\}}{\alpha_i^2} \quad (18)$$

As will be explained later, modifications to this model can be made to give an adaptive quality to the dynamic model. The effectiveness of this adaptive dynamic model will also be assessed.

2 Divergence

The prediction of the state is calculated from assumed models of the real world and for this reason they can never be exact. The real time operation of a Kalman filter will always be degraded from the theoretical projection of the state vector. This discrepancy between the theoretical and real-world situation can causes a large divergence in the Kalman filter which can be classified into two types (Brown & Hwang, 1992). The first of these, *bounded divergence*, is where the divergence of the filter is bounded within a finite range, resulting in the filter producing estimation errors much larger than those theoretically expected. The modelling errors in this instance cause the filter to perform sub-optimally. The latter, *true divergence*, is where the estimation errors become infinite, resulting in loss of stability in the filter algorithm.

In this paper the authors are mainly concerned with bounded divergence. There are a number of options available to help minimize bounded divergence caused by estimation errors. The unknown state can be modelled as a random process (i.e., as a white noise or auto-correlated noise), or the unknown state is assumed to be non-random (to be estimated, detected and compensated for in real time).

3 Motion Detection

In order for an adaptive dynamic model to be effectively implemented in a Kalman filter, there must be some process to detect a fault in the dynamic model. In any failure detection algorithm two essential tasks must be performed (Chow & Willsky, 1984):

(a) Generation of residuals or a signal that will reflect direct changes in the measurements. Ideally the residuals will be close to zero when no change occurs, and vary significantly when a failure is present.

(b) The design and development of algorithms through the use of statistical tools to decide when to declare a failure in the system.

The detector must be able to promptly react to any change in the system, and preferably be able to estimate the exact time of the failure and it's magnitude.

In this study it is assumed that a fault in the dynamic model can be detected and properly identified through the innovation sequence. In reality it is very hard to determine if an individual failure alarm is due to a measurement error or modelling error.

The innovation sequence has a property, that if the Kalman gain matrix is optimal, then:

$$E[v(t_1)v(t_2)^T] = 0, t_1 \neq t_2 \qquad (19)$$

In other words, the innovation sequence v is a white process noise (Gelb, 1974). Consequently this useful property allows a number of different change detection algorithms to be applied to the innovation sequence, such as control charts, CUSUM (Ogaja, 2001), FFT and wavelets.

The CUSUM algorithm has the advantage of being able to effectively summarize recent data, while offering a greater control on the number of false alarms. The algorithm detects small, but sustained, changes in a zero-mean random noise environment, and is able to characterize the origin of such changes. It is this algorithm which was used to detect any change in dynamics in the simulations described late in this paper. For more background detail on the CUSUM algorithm see Ogaja (2001).

4 Adaptive Dynamic Modelling

4.1 White noise model approach

If divergence is detected, one way to bound the divergence is to add a fictitious process noise to the dynamic model, which is then scaled up until the estimation errors are below the designated 'in error' threshold. An advantage of this approach is that it can be carried out on all of the states, or on a selective set of states, that are suspected to be causing the divergence.

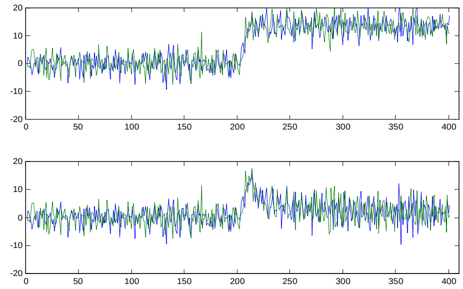

Figure 1 innovation sequence plots for a constant velocity model(top) without added white noise, and with white noise added (bottom)

Figure 1 show the response of the two filters, to the innovation sequence, for a simulated data set of an object moving at constant velocity (55km/h) at a bearing of 45°, until epoch 200 when the object is then simulated to be constantly accelerating (5km/h/s). The constant velocity filter without the added white noise (top) is clearly running sub-optimally with a large jump at epoch at 200, and a sustained bias for the remainder of the time. The filter with white noise added (bottom) experiences an initial jump at epoch 200 in the sequence, due to the delay in detecting the change of dynamics, but then begins to re-stabilize with only a small departure from the expected zero-mean for the remainder of the innovation sequence. Also of note is an increase in the variance of the innovation sequence resulting from the increase in added noise to the system.

The white noise approach offers a way of reducing the *bounded divergence* in a Kalman filter due to mismodelling, but still leaves the filter running in a sub-optimal state.

4.2 Auto-correlated noise model approach

A more realistic process noise can be added to the state(s) suspected of causing the divergence. This approach assumes that the change in dynamics is caused by a maneuver correlated in time (Singer, 1970). Typically the maneuver is corrected by assuming a zero-mean random process with an exponential auto-correlation function:

$$R(\tau) = E[a(t)a(t+\tau)] = \sigma_m^2 e^{-\alpha|\tau|} \qquad (20)$$

where σ^2_m is the variance of the noisy state, and $1/\alpha$ is the time constant of its auto-correlation

Note that as α increases, the process of the noisy state de-correlates faster, converging to the original dynamic model as α→∞. As α approaches 0 the integral of white noise is obtained. So an alpha value should be used that reflects the duration of the 'maneuvering period' (i.e. α ≅ 1/60 for a lazy turn, α ≅ 1/10 for fast turn).

This approach best compensates for small intermittent changes in the assumed dynamic trajectory, retaining the possibility of being able to quickly resume to the original optimal solution, with only a small loss in performance from the filter.

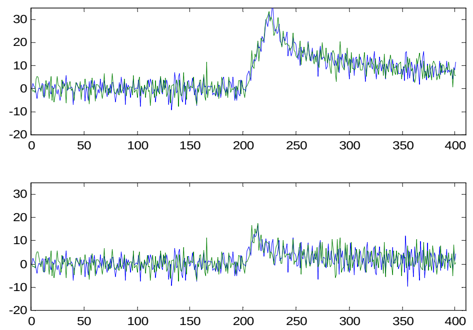

Figure 2 Innovation sequence plots for the added coloured noise approach, (top) beta = 3 (i.e. little correlation), bottom beta = 0.003 (i.e. highly correlated)

Figure 2 shows the response of the filters with the same simulated data set as used previously, to test the coloured noise approach. The filter with the low α value (bottom) applied to the coloured noisy state, produces a nearly identical result to the white noise approach. The filter with the higher α (top) has a higher initial jump in the innovation sequence at epoch 200 compared to the white noise approach and constant velocity approach (without white noise). The level of bias using the filter with high α values has a slow response in compensating for the bias in the filter, but has the advantage of not increasing the variance of the innovation sequence.

4.3 Input Estimation Approach

In this approach an extra state is added to the state vector to account for the unmodelled dynamics. The estimation of this additional state term is accounted for through the innovation sequence. A statistical test is then carried out to see if it is statistically significant, and then it is used to correct the state equations (Chan et al., 1979).

Using this approach, the prediction of the state vector can be expressed as:

$$x_{k+1} = Fx(k) + Gu(k) + v(k) \qquad (21)$$

where u is an unknown input that models the target maneuvers (u = 0 when there is no maneuver). v is the process noise.

The state is calculated from the model without the additional input parameter included. Using the innovation sequence from this filter the extra parameter u is detected, estimated, and then used to correct the state estimate. When the additional input is required, then the innovation becomes a white noise sequence plus a term related to the unaccounted-for parameter:

$$v^*_{k+1} = v_{k+1} + A \sum_{j=k}^{i} [\prod_{m=k}^{j-1} \Phi(m)] Gu(j) \qquad (22)$$

Assuming that the input is constant over the time interval [k,...,k+s], then:

$$v^*_{i+1} = \Psi_{i+1} u + v_{i+1}, \quad i = k,...,k+s-1 \qquad (23)$$

where

$$\Psi_{k+1} \cong A \sum_{j=k}^{i} \left[\prod_{m=k}^{j-1} \Phi(m) \right] G$$

From the above equation it can be seen that the innovation sequence of the non-maneuvering filter is a 'linear measurement' of the input (maneuver) u in the presence of the added 'white noise' v. Therefore the input parameter **u** can be estimated using least squares:

$$y = \Psi u + \varepsilon \qquad (24)$$

where

$$y \cong \begin{bmatrix} v^*_{k+1} \\ \vdots \\ v^*_{k+s} \end{bmatrix}, \quad \Psi \cong \begin{bmatrix} \Psi_{k+1} \\ \vdots \\ \Psi_{k+s} \end{bmatrix}, \quad \varepsilon \cong \begin{bmatrix} v_{k+1} \\ \vdots \\ v_{k+s} \end{bmatrix}$$

is zero-mean with a block diagonal covariance matrix:

$$S \cong diag\ [S(i)] \qquad (25)$$

Hence the estimation can be done in batch form:

$$\hat{u} = \left(\Psi^T S^{-1} \Psi \right)^{-1} \Psi^T S^{-1} y \qquad (26)$$

with the resulting covariance matrix:

$$L = \left(\Psi^T S^{-1} \Psi \right)^{-1} \qquad (27)$$

If a maneuver is detected, then the state will be corrected, by using the input term estimated from the least squares calculation as follows:

$$\hat{x}^u_{(k+s+1|k+s)} = \hat{x}^*_{(k+s+1|k+s)} + M\hat{u} \qquad (28)$$

where

$$M \cong \sum_{j=k}^{i} \left[\prod_{m=k}^{j-1} \Phi(m) \right] G$$

The maneuver is considered complete when the input estimate, based on the measurements from the sliding window of length s**,** become insignificant. The value of s is a design parameter. In cases where the duration of a maneuver is short relative to the sampling interval, values of 1 or 2 are appropriate. However, it is usually wise to consider data over a longer period in order to produce a reliable estimate.

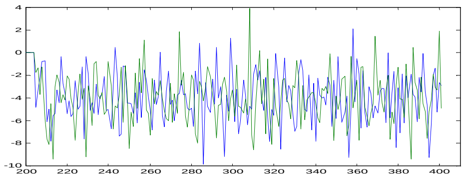

Figure 3 The corrected innovation sequence plot for the Input Estimation technique

Figure 3 shows the innovation sequence for the input estimation approach, with the same simulated data set used previously. As can be seen, the method has failed to remove the bias in the innovation sequence completely, but continues to operate at an acceptable level for the filter.

4.4 Variable Dimension Approach

In this approach, a dynamic model is selected as being the default mode for the Kalman filter (in this case a constant velocity model). Once a modelling error has been identified, a higher dimensional state dynamic model is applied to the filter. This is the most effective of all the techniques described in this paper, as it will ensures that the right order of dynamic model is used for the conditions being experienced by the filter. The rational for choosing a higher or lower order model is to ensure optimal filter performance in all dynamic situations (Bar-Shalom & Birmiwal, 1982).

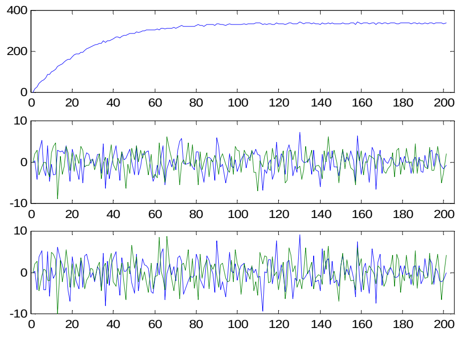

Figure 4 Innovation sequence plot of different degrees of dynamic model fitting an object moving at a constant velocity. Constant Position (top), Constant Velocity (middle), Constant Acceleration (bottom)

Figure 4 shows the response of different degree dynamic models to a simulated object moving at a constant velocity of 55km/h, with a random positional error of 2.5m. In the constant position (top) case it can be clearly seen that the innovation sequence has a large departure from the expected zero-mean. This indicates that the model used by the filter is erroneous. In the constant velocity case (middle), one can see that the innovation sequence is well centred around the expected zero-mean. For the constant acceleration model (bottom), one can see that the results are very similar to the constant velocity model, except that there is an increase in the variance of the innovation sequence. Therefore the constant acceleration is causing the filter to perform sub-optimally.

The effect of changing from one dynamic model to another can cause a large jump in the innovation sequence. This may be due to not correcting the states at the time when the fault occurred, or can be induced through a sudden drop in a higher order state (i.e. a sudden drop in acceleration), and should not be used until the acceleration states are required.

To reduce these jumps in the innovation sequence, a high dimensional state model must be switched to a lower dimensional state model when it is no longer seen as being a statistically significant improvement. If the jump in the innovation sequence is caused through the delay in correcting the bias, then the states will need to be re-computed at the time of failure with the newly adapted dynamic model. Once the states have be reprocessed then the Kalman filter will process the measurements at the time of arrival.

It is therefore important that the states are recomputed with the 'new' dynamic model, otherwise the Kalman filter may take a long time to recover from the initial divergence. If not corrected the results obtained from any interpolation from the dynamic model may be biased. This increases the probability of an out of control alarm being re-triggered, which an appropriately adapted filter would have avoided. An effective implementation strategy may be to down weight the dynamic noise matrix for a fixed period of time, or until a statistical test has confirmed that the filter has re-stabilized.

4.5 Adaptive Dynamic Model

In this paper the constant acceleration model (17) is adapted by modifying the α values for T, UD and S. The model is set in the default mode of constant velocity by setting the correlation length α to be very large. This effectively sets the third column and row of (17) to 0s, giving a model identical a constant velocity model as set out in (11). When the filter detects an error in the dynamic model, through the innovation sequence, the α values are reduced to give the constant acceleration model.

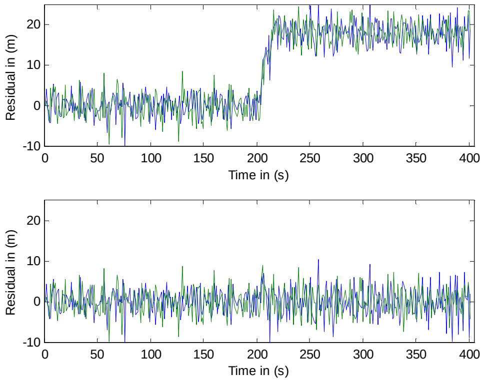

Figure 5 Innovation sequence for the Constant velocity model (top), and the Adaptive Variable Dimensional Dynamic model (bottom).

A simulation was run to see how effective the adaptive dynamic model would perform compared to a constant velocity model under different dynamic conditions. The conditions simulated was a period of constant velocity at 55km/h for 200 epochs, and then finally a period of constant acceleration at 5km/h/s for 200 epochs.

Figure 5 shows the results of the constant velocity model (top) and the adaptive dynamic model (bottom). As shown above both models perform identically during the constant velocity stage of the simulation. During the constant acceleration phase of the simulation, the constant velocity model begins to shift the mean of the innovation sequence, implying a bias in the filter. For the adaptive dynamic model, the shift in the mean of the innovation sequence can be picked up by the CUSUM test. The states are then recomputed for the appropriate number of epochs from the point of failure with the new dynamic model, thus ensuring consistent and optimal filter performance throughout all of the different dynamics experienced by the filter.

5 Conclusion

If a filter experiences a change in conditions not adequately accounted for in the dynamic model then *bounded divergence* will result. This will cause the filter to work sub-optimally producing estimates with a higher degree of uncertainty. This paper has investigated a number of techniques to limit or avoid *bounded divergence* occurring in a Kalman filter. The most effective technique was found to be the adaptive variable dimensional dynamic model, which ensures that the filter is operating in its most optimal state for the conditions being experienced at the time.

References

Bar-Shalom Y., & Birmiwal K.(1982), Variable Dimension Filter for Maneuvering Target Tracking, IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-18, No. 5, pp. 621-628.

Brown RG, Hwang PYC (1992), Introduction to random signals and applied Kalman Filtering (2 edn), Jown Wiley and Sons, New York.

Chan YT, Hu AGC, & Plant JB (1979), A Kalman filter based tracking scheme with input estimation, IEEE Tran. Aerospace & Electronic Systems, Vol. AES-15, 237-244.

Chow EY, Willsky (1984), Analytical redundancy and the detection of robust failure detection systems, IEEE Trans. Autom Control 29, pp603-613.

Gelb A (ed.) (1974), Applied Optimal Estimation, The M.I.T. Press, Cambridge, Massachusetts.

Schwarz KP, Cannon ME, & Wong RVC (1989), A comparison of GPS kinematic models for the determination of position and velocity along a trajectory, Manuscripta Geodaetica, 14, 345-353.

Singer R A.(1970), Estimating Optimal Tracking Filter Performance for Manned Maneuvering Targets, IEEE Transactions on Aerospace and Electronic Systems, VOL. AES-6, No. 4., pp. 473-483.

Ogaja C (2001), On-line GPS integrity monitoring and deformation analysis for structural monitoring applications, accepted for ION GPS-2001.