TECHNICAL NOTE

ADAPTIVE ESTIMATION IN THE PRESENCE OF ORDER AND PARAMETER CHANGES

L. XIA, V. KRISHNAMURTHY AND J. B. MOORE

Department of Systems Engineering, The Australian National University, GPO Box 4, Canberra, ACT, Australia 2601

SUMMARY

This paper gives an approach to adaptive estimation/control when there are jump parameter changes which include order changes. Order changes can be viewed as the introduction of overparametrization, which in conventional algorithms causes ill-conditioning. Here, modified algorithms which involve the introduction of noise into the calculations are proposed and studied by theory and simulations.

KEY WORDS Kalman filter Overparametrization Jump order changes Identification

1. INTRODUCTION

Current adaptive estimation and control algorithms are designed to cope with plants of fixed order. However, many practical applications involve jump order changes for a plant. A typical example is the adaptive control of a robot arm when it grasps a flexible rod. There is a sudden 'jump' change in system order when the arm grasps the rod owing to the additional flexure modes of the rod. Another significant application is the adaptive control of a structure in space under construction or involved in a docking procedure. From our studies of the literature, the problem of dealing with jump order changes seems not to have been addressed seriously.

In the presence of jump order changes, pole–zero cancellations in the estimates of the plant is unavoidable, or equivalently there is overparametrization in the algorithm. As studied in References 1 and 2, some adaptive control schemes, such as standard pole assignment schemes, are ill-conditioned when there is overparametrization, unless pole–zero cancellations in the plant estimates occur at the origin. For indirect adaptive control schemes, based on least squares or gradient/stochastic approximation, the overparametrization in the estimates of the plant causes ill-conditioning in the estimation algorithm. More precisely, there is lack of sufficient excitation in the regression vectors and consequent parameter estimate drift. Corresponding ill-conditioning occurs in the direct adaptive control schemes, and for short-memory schemes there can be bursting phenomena.

In this paper, with a view to coping with time-varying plants with jump order changes, we develop novel algorithms, building on earlier works. These earlier studies propose least-squares-based adaptive estimation and control algorithms for unknown fixed plants, with only an upper bound on the plant order. The key contribution of References 1 and 2 is the...
introduction of noise signals into standard estimation algorithms according to ill-conditioning measures. This ensures that identification of any pole–zero cancellation takes place at the origin and any overparametrization does not lead to ill-conditioning.

In this paper, the approach of References 1 and 2 is modified for the case when the plant may have time-varying parameters by working with Kalman-filter-based estimation schemes or recursive least squares with forgetting factor schemes. Again noise is introduced to the calculations according to ill-conditioning measures. However, the new measures are non-trivial generalizations of earlier ones and a good deal simpler to implement. They are also able to detect and cope with jump order changes. For simplicity, the paper focuses on the special case of piecewise constant parameter values which permit jump order changes, but where parameter/order changes are relatively infrequent. Attention is also focused on the adaptive estimation case, this being the basis for adaptive control.

The algorithm is detailed in Section 2 and some simulation results are given in Section 3. In Section 4 we draw a few conclusions.

2. PERTURBED KALMAN FILTER DETECTION AND IDENTIFICATION SCHEME

Signal model

In the first instance we consider a scalar, time-varying, linear system with changing order described by the DARMA model as

\[ A_k(q^{-1})y_k = B_k(q^{-1})u_k \]  

(1)

where

\[ A_k(q^{-1}) = 1 + \sum_{i=1}^{n} a_{ki} q^{-i} \quad B_k(q^{-1}) = \sum_{i=1}^{m} b_{ki} q^{-i}, \]

and \( u_k \) and \( y_k \) are the system input and output respectively; \( n \) and \( m \) are the upper bounds on the degree of the polynomials \( A_k(q^{-1}) \) and \( B_k(q^{-1}) \) respectively. We say that the system is overparametrized when the polynomials \( q^n A_k(q^{-1}) \) and \( q^m B_k(q^{-1}) \) have common zeros. The degree of overparametrization is the number of common zeros. Thus if there are no common zeros, i.e. \( q^n A_k(q^{-1}) \) and \( q^m B_k(q^{-1}) \) are coprime, the degree of overparametrization is zero.

A stochastic version of the signal model in (1) for estimation purposes can be rewritten as

\[ \theta_{k+1} = \theta_k + w_k \quad y_k = \theta_k^T \phi_k + v_k \]  

(2)

\[ \theta_k^T = [a_{k}^{(1)} \ldots a_{k}^{(n)} b_{k}^{(1)} \ldots b_{k}^{(m)}] \quad \phi_k = [-y_{k-1} \ldots -y_{k-n} u_{k-1} \ldots u_{k-m}] \]

Here \( \phi_k \) is termed a regression vector. The measurement noise term \( v_k \) is involved more for the purposes of algorithm design than to reflect any assumed persistently exciting measurement disturbance. The term \( w_k \) is the noise sequence which describes the changes in the parameter vector \( \theta_k \). For simplicity we assume that \( \theta_k \) is piecewise constant with infrequent changes, i.e. \( w_k = 0 \) for most \( k \). An order change is merely a change in parameters such that there is a change in the degree of overparametrization in plant (1).

Excitation

Assume that \( u_k \) is persistently exciting so that

\[ 0 < \alpha_2 I < \frac{1}{N} \sum_{i=k}^{k+N} \bar{u}_i \bar{u}_i^T < \alpha_1 I \quad \bar{u}_i^T = [u_i \ldots u_{i-n-m}] \]  

(3)

for some \( \alpha_1, \alpha_2, k_0, N \) and all \( k > k_0 \).
It is well known\(^6\) that for intervals when \(A_k(q^{-1})\) and \(B_k(q^{-1})\) are constant and coprime, there is no pole–zero cancellation in the system model and the input excitation assumption (3) translates to an excitation condition on the regression vector \(\phi_k\) as

\[
0 < \beta_z l < \frac{1}{N} \sum_{i=k}^{k+N} \phi_i \phi_i^T \beta_1 l
\]  

(4)

It is important that this excitation condition holds even in the absence of any actual persistently exciting measurement noise \(v_k\). Of course, if \(v_k\) is persistently exciting, then \(\phi_k\) is also persistently exciting and problems associated with overparametrization do not arise. To exclude this ‘trivial’ case, we assume that \(v_k\) may not be persistently exciting. If the degree of overparametrization is \(L < \min(n, m)\), it can be proved that \(\phi_k\) is persistently exciting only in the \((n + m - L)\)-dimensional subspace of the whole space \(\mathbb{R}^{n+m}\).

**Standard Kalman filter identification**

We consider first a standard Kalman filter form for estimating the system parameters.

\[
\hat{\theta}_{k+1} = \hat{\theta}_k + \frac{P_k \phi_k e_k}{R + \phi_k^T P_k \phi_k} e_k = y_k - \phi_k^T \hat{\theta}_k
\]  

(5a)

\[
P_{k+1} = P_k - \frac{P_k \phi_k \phi_k^T P_k}{R + \phi_k^T P_k \phi_k} + Q
\]  

(5b)

\[R > 0, \quad Q > 0\]  

are design parameters

(5c)

It has been proved\(^7\) that if \(v_k\) and \(w_k\) have bounded variance and \(\phi_k\) is suitably exciting, then the standard Kalman filter algorithm (5) gives an estimate \(\hat{\theta}_k\) of \(\theta_k\) with bounded tracking errors. Moreover,\(^8\) if \(w_k\) and \(v_k\) are zero mean and Gaussian with variances \(Q\) and \(R\) respectively, then (5) will yield the conditional minimum variance (conditional mean) estimate, with conditional error covariance \(P_k\). However, if \(\phi_k\) is not suitably exciting, the estimate of \(\hat{\theta}_k\) using algorithm (5) does not necessarily have a bounded tracking error.\(^7\) In fact, as will be shown later, \(P_k\) in (5) becomes unbounded when the plant is overparametrized, and consequently \(\hat{\theta}_k\) becomes unbounded. To avoid any possible ill-conditioning associated with the lack of suitable excitation in the regressor, the approach in References 1 and 2 of using a perturbed estimation algorithm will be used in the Kalman filter estimation. (Notice that the Kalman filter algorithm specializes to the least squares estimation scheme of Reference 1 and 2 when \(Q = 0\) and \(R = 1\).)

**Perturbed Kalman filter**

We propose the following perturbed Kalman filter identification algorithm:

\[
\tilde{\theta}_{k+1} = \tilde{\theta}_k + \frac{\tilde{P}_k \psi_k e_k}{R + \psi_k^T \tilde{P}_k \psi_k} \quad e_k = y_k - \psi_k^T \tilde{\theta}_k
\]  

(6a)

\[
\tilde{P}_{k+1} = \tilde{P}_k - \frac{\tilde{P}_k \psi_k \psi_k^T \tilde{P}_k}{R + \psi_k^T \tilde{P}_k \psi_k} + Q
\]  

(6b)

\[\tilde{P}_0 > 0\]

\[
\psi_k = \phi_k + v_k
\]  

(6c)

where \(v_k\) is an injected noise which is white and zero mean to cope with the lack of persistent excitation in \(\phi_k\). This signal \(v_k\) is only injected into the Kalman filter algorithm and does not
directly influence the system or the system responses. With Kalman filter time-varying parameter identification, designing a suitable noise covariance is not a straightforward extension of that previously proposed in References 1 and 2 when parameters are assumed constant. However, in common with the approach of References 1 and 2, the selection is made on the basis of the appropriate Riccati equation solution, here $P_k$. A crucial property of $P_k$ is now studied.

**Kalman filter error covariance property**

**Theorem 1**

Consider the signal model as in (2) with constant parameters (namely $\theta_k = \theta$) and Kalman filter algorithm (5) with $Q = qI$. Consider also that the input $u_k$ is persistently exciting, i.e. (3) is satisfied.

1. If there is no overparametrization, then for some bounds $m_1, M_1$
   
   $0 < m_1 < \text{tr}[P_k] < M_1 < \infty$

2. If there is overparametrization by $L < \min(n, m)$, i.e. the degree of overparametrization is $L$, then
   
   $0 < m_1 + qL < \text{tr}[P_k] < M_1 + qL$

**Proof.** For part (1), from Reference 7 and (4) we see that $P_k$ is upper and lower bounded. Thus (7) is established. For part (2), since the system is overparametrized by $L$, $\phi_k$ in (2) is only exciting in the $(n + m - L)$-dimensional subspace of the $\mathbb{R}^{n \times m}$ space of $\phi_k$. Therefore there exists a unitary matrix $T$ such that

$$TT^\top = I \quad T\phi_k = [\phi_k^\top 0 ... 0]^\top = \Phi_k^* \quad \forall k$$

(9a,b)

After the transformation in (9b), the last $L$ rows of $T\phi_k$ are zero and $\phi_k^*$ is persistently exciting. Now recall Kalman filter algorithm (5), specialized here as

$$P_{k+1} = P_k - \frac{R + \phi_k^T P_k \phi_k}{P_k \phi_k \phi_k^T P_k} + qI$$

(10)

Premultiplying (10) by $T$ and postmultiplying by $T^\top$ we have

$$TP_{k+1}T^\top = TP_kT^\top - \frac{TP_k \phi_k \phi_k^T P_k T^\top}{R + \phi_k^T P_k \phi_k} + qTT^\top$$

(11)

Let $P_k^* = TP_kT^\top$. Then using (9) we can rewrite (11) as

$$P_{k+1}^* = P_k^* - \frac{P_k^* \Phi_k^* \Phi_k^{*\top} P_k^*}{R + \Phi_k^{*\top} P_k^* \Phi_k^*} + qI$$

(12)

By partitioning $P_k^*$ and $\Phi_k^*$ as

$$P_k^* = \begin{bmatrix} P_k(1) & P_k(2) \\ P_k(2)^\top & P_k(3) \end{bmatrix} \quad \Phi_k^* = \begin{bmatrix} \Phi_k^* \\ 0 \end{bmatrix}$$

(13)
(12) can be written as

\[ P_{k+1}(1) = P_k(1) - \frac{P_k(1)\phi_k^*\phi_k^T P_k(1)}{R + \phi_k^* P_k(1) \phi_k} + qI_1 \tag{14} \]

\[ P_{k+1}(2) = P_k(2) - \frac{P_k(1)\phi_k^*\phi_k^T P_k(2)}{R + \phi_k^* P_k(1) \phi_k} \tag{15} \]

\[ P_{k+1}(3) = P_k(3) - \frac{P_k(2)\phi_k^*\phi_k^T P_k(2)}{R + \phi_k^* P_k(1) \phi_k} + qI_2 \tag{16} \]

From (14) and also since \( \phi_k^* \) is persistently exciting, it can be concluded that \( P_k(1) \) is bounded from above and below. Also it can easily be proved\(^7\) that since \( \phi_k^* \) is persistently exciting, (15) implies that \( P_k(2) \) converges exponentially to zero. Thus when \( k \) is large in (16), the following approximation holds:

\[ P_k(3) \approx kqI_2 \tag{17} \]

Now since \( \text{tr}[P_k] = \text{tr}[P_k^*] \), we have

\[ \text{tr}[P_k] = \text{tr}[P_k(1)] + \text{tr}[P_k(3)] \tag{18} \]

From the properties of \( P_k(1) \) and \( P_k(3) \), the desired result (8) follows. \(\square\)

**Corollary 1**

Consider that the conditions for Theorem 1 hold. Consider the average on the trace of \( P_k \) over the past period \( l \) as

\[ P_m(k) = \frac{1}{l} \sum_{i=k-l}^{k} \text{tr}(P_i) \]

Then there is a constant \( N_l \) such that for all \( k > N_l \),

(1) if the system is not overparametrized, then the following inequality is true:

\[ \text{tr}[P_k] - P_m(k - N_l) < N_lq \tag{19} \]

(2) if the system is overparametrized by \( L_0 > 0 \), the following inequality is true:

\[ N_lqL_0 < \text{tr}[P_k] - P_m(k - N_l) < N_lq(L_0 + 1) \tag{20} \]

**Proof.** The proof follows from application of (7) and (8). \(\square\)

**Injected noise construction**

Now, on the basis of Theorem 1 and Corollary 1, we propose the following procedure for constructing the injected noise \( v_k \).

Step 1. Generate a zero-mean, unit-covariance white noise sequence \( \tilde{v}_k \).
Step 2. On the basis of the standard Kalman filter algorithm (5), update \( P_k \) and \( P_m(k) \).
Step 3. Check for some large constant \( N_l \) whether (19) holds, or if not, for what value of \( L_0 \) (20) holds. This latter constitutes a detection of the degree of the overparametrization \( L_0 \).
Step 4. If (19) holds then let \( v_k = 0 \); if (20) holds for \( L_0 \) then set

\[
v_k = D_k \hat{v}_k \quad D_k = \text{Diag}[0 \ldots 0 \ 1 \ldots 1 \ 0 \ldots 0 \ 1 \ldots 1]
\]

(21)

where the pattern for the diagonal \( D \) is \( n-L_0 \) zeros followed by \( L_0 \) ones and then \( m-L_0 \) zeros and \( L_0 \) ones again. (If we set \( L_0 = 0 \), for the case of no overparametrization as detected by (19) holding, then (21) is a comprehensive formula for \( v_k \) design.)

**Properties of perturbed Kalman filter**

When the system is not overparametrized, the perturbed Kalman filter specializes as the standard Kalman filter.

If the system is overparametrized, then the model parametrization is not unique, \( v_k \neq 0 \), and \( \hat{v}_k \text{ in (6c)} \) is persistently exciting by virtue of the presence of \( v_k \). To prove this, we first verify that \( \hat{v}_k = \phi_k + v_k \) is completely reachable from both \( v_k \) and \( u_k \) with any \( L_0 \) overparametrization. Then the persistency of excitation of \( \hat{v}_k \) follows from (3) and the stochastic nature of \( u_k \). Because of the excitation of \( \hat{v}_k \), \( \tilde{P}_k \) is bounded above and below. Moreover, if the detection of the degree of overparametrization is correct, the addition of \( v_k \) does not introduce any bias on the estimation. Furthermore, the injected noise \( v_k \) forces an estimation of the unique plant parameters which give rise to a pole–zero cancellation at the origin. The reason for this is as follows. From (6) and denoting \( \hat{v}_k = \theta - \hat{\theta}_k \), we have (with \( \theta^T v_k = 0 \))

\[
\phi_k^T \theta = \psi_k^T \theta \quad \hat{\theta}_{k+1} = \hat{\theta}_k - \frac{\tilde{P}_k \psi_k \psi_k^T \hat{\theta}_k}{R + \psi_k^T \tilde{P}_k \psi_k}
\]

(22)

where \( \theta \) is the unique plant parameter vector which forces a pole–zero cancellation at the origin for the interval in question. Then the argument as in Reference 7 can be used to conclude that \( \hat{\theta}_k \) converges to zero, or in other words, the algorithm tries to estimate the ‘true’ plant parameters.

**Avoiding false alarms**

It is noticed that only when \( N_t \) is suitably large does (19) or (20) hold. However, in order to adapt to a rapidly changing environment, \( N_t \) must be small. In this case, there might exist intervals where \( \text{tr}[\tilde{P}_k] \) increases with a gradient more than predicted, given knowledge of the degree of overparametrization. This constitutes a false alarm, since in such an interval the degree of system overparametrization is incorrectly detected. To avoid false alarms, we propose to consider the previous \( j \) detections in the calculations as follows. Consider a modified noise selection as

\[
v_k = \left( \prod_{i=k-j}^k D_i \right) \hat{v}_k
\]

(23)

The choice of the constant \( j \) and \( N_t \) is a trade-off between the sensitivity of system order change and false alarm. Normally we choose \( N_t = 3(n + m) \) and \( j = (n + m)/2 \). Another ad hoc approach is to make \( j \) change on the basis of the information from the estimation errors \( e_k \) in (6). Details are omitted here.

### 3. SIMULATIONS

To give insights into the algorithm behaviour, a number of simulation studies have been made. Here we report a typical one as follows. The signal model is assumed as a third-order system
with \( n = 3 \) and \( m = 3 \) (recall (1) and (2)). The plant parameter sets are

\[
\theta_k = \begin{cases} 
[-0.8 0.8 -0.2 1.5 0.5 0.4] & 0 < k \leq 200 \\
[-0.6 0.15 0 1.25 -0.8 0] & 200 < k \leq 400 \\
[-0.8 0.8 -0.2 1.5 0.5 0.4] & 400 < k \leq 600 \\
[-0.6 0 0 1.25 0 0] & 600 < k \leq 800 
\end{cases}
\]

A white noise sequence is used as the input signal. The design parameters are chosen as \( R = 1 \) and \( Q = 0.1I \), the interval for calculation of \( P_m(k) \) is \( l = 18 \), and the interval for slope checking is \( N_l = 18 \). Figures 1(a) and 1(b) show the estimates of the plant parameters, and Figure 2
depicts the behaviour of the trace of $P_k$. To make an easy comparison, Figures 3(a) and 3(b) give the estimates of the plant parameters using the standard Kalman filter algorithm.

From Figures 1 and 3 it is clearly seen that for time periods 1–200 and 400–600 there is no overparametrization and the estimates of the plant parameters from both the perturbed Kalman filter (PKF) algorithm and the standard Kalman filter (SKF) algorithm are almost the same. For periods 200–400 and 600–800 there is overparametrization. The estimates from the PKF algorithm converge to the plant parameters which correspond to the pole–zero cancellation at the origin, or in other words, they converge to $\theta_k$. However, the estimates from the SKF algorithm do not converge to the plant parameters $\theta_k$; pole–zero cancellations may occur anywhere depending on the initial conditions (when $k_0 = 200$ or $k_0 = 600$).
From Figure 2 it is clearly seen that when the signal model is not overparametrized as in periods 1–200 and 400–600, and also when the signal model is overparametrized as in periods 200–400 ($L_0 = 1$) and 600–800 ($L_0 = 2$), the trace of $P_k$ has the gradient properties expected, being proportional to $L_0$. These and other simulations not reported here confirm that the PKF algorithm performs virtually as well as if jump order changes do not occur. The PKF algorithm certainly performs as well as SKF with order changes where appropriate, and copes better when there are false alarms.

4. CONCLUSIONS

An algorithm of adaptive estimation to cope with jump parameter changes and order jump changes is proposed. The key modification to the standard algorithm is to introduce injected noise into the algorithm to handle the ill-conditioning due to lack of persistence of excitation caused by overparametrization. Theoretical analysis and simulations confirm that the algorithm has attractive properties, in that the algorithm performs as if the jump parameter changes did not include order changes.

ACKNOWLEDGEMENT

This work was partially supported by DSTO, Australia.

REFERENCES