

A NOTE ON THE SMOOTHING FORMULATION OF MOVING HORIZON ESTIMATION

UDC 681.5.01 621.372.852.1

Morten Hovd

Engineering Cybernetics Department, Norwegian University of Science and Technology,
N-7491 Trondheim, Norway

Abstract. This note points out two discrepancies in the formulation of the arrival cost in the smoothing formulation of Moving Horizon Estimation (MHE). The first is an apparent misprint in the expression for the arrival cost in [1], the second is that we cannot assume that the control inputs and the state excitation noise sequences are uncorrelated over the estimation horizon.

Key words: MPC, moving horizon, estimation

1. INTRODUCTION

Model Predictive Control (MPC) has found wide application in the process industries [2]. The main reason for the popularity of MPC compared to other advanced control methods is commonly held to be its ability to handle constraints. Inspired by the success of MPC, Moving Horizon Estimation (MHE) - occasionally also termed Receding Horizon Estimation (RHE) - adopting similar constraint handling mechanism as MPC [3]. It may reasonably be argued that these constraint handling mechanisms are more systematic than the *a posteriori* 'clipping' of the estimated states that is commonly used in the (E)KF to enforce state constraints.

However, the focus in this note is not the constraint handling mechanism of MHE. Rather, we focus on the formulation of the 'arrival cost' for MHE, i.e., the term in the objective function that accounts for the information received prior to the start of the estimation horizon. Weighting this information too heavily may cause convergence or instability problems, whereas too little weigh on this information may cause poor performance [4].

This note is organized as follows; Section 2 introduces the MHE formulation and the difference between the *filtering* and *smoothing* formulation of MHE. In Section 3 the arrival cost for the smoothing formulation is presented. Section 4 addresses the fact that

Received November 22, 2012

Corresponding author: Morten Hovd

Engineering Cybernetics Department, Norwegian University of Science and Technology, N-7491 Trondheim, Norway • E-mail: morten.hovd@itk.ntnu.no

when considering the estimation problem over a horizon, the state excitation noise cannot be assumed to be independent from the control inputs - if the control inputs are determined using feedback from (estimated) states.

2. MHE FORMULATION

The state estimation problem is formulated as follows:

$$\min_{\tilde{x}, w, v} (\hat{x}_{k-N} - \tilde{x}_{k-N})^T S (\hat{x}_{k-N} - \tilde{x}_{k-N}) + \sum_{i=1}^N (v_{k-N+i}^T V^{-1} v_{k-N+i} + w_{k-N+i}^T W^{-1} w_{k-N+i}) \quad (1)$$

subject to constraints

$$\begin{aligned} \hat{x}_{k-N} & \text{ given} \\ y_{k-N+i} & \text{ given; } i = 1, \dots, N \\ u_{k-N+i} & \text{ given; } i = 0, \dots, N-1 \\ y_{k-N+i} & = C \tilde{x}_{k-N+i} + v_{k-N+i}; i = 1, \dots, N \\ \tilde{x}_{k-N+i+1} & = A \tilde{x}_{k-N+i} + B u_{k-N+i} + E w_{k-N+i}; i = 0, \dots, N-1 \end{aligned}$$

The noises w and v are assumed to be zero mean, independent and normally distributed, with covariances W and V , respectively. Here constraints other than the model equations are ignored - although in the absence of such constraints there would seem to be little reason to use an MHE formulation instead of the much simpler Kalman filter. The variables \tilde{x} , w and v are defined as:

$$\begin{aligned} \tilde{x} &= \begin{bmatrix} \tilde{x}_k^T & \cdots & \tilde{x}_{k-N}^T \end{bmatrix}^T \\ w &= \begin{bmatrix} w_{k-1}^T & \cdots & w_{k-N}^T \end{bmatrix}^T \\ v &= \begin{bmatrix} v_k^T & \cdots & v_{k-N+1}^T \end{bmatrix}^T \end{aligned}$$

Remark: The above formulation reflects the situation where the estimation is performed *before* the MPC calculations for each timestep, i.e., when the MPC at time k is supplied with the state estimate $x_{k|k}$. Thus, the MPC will be supplied with the *a posteriori* state estimate at time k . Likewise, the \hat{x}_{k-N} supplied to the state estimation at time k is (in the filtering formulation of the MHE) the *a posteriori* state estimate obtained at time $k-N$, and the weight S (also for the filtering formulation) is the inverse of the *a posteriori* Kalman filter covariance at time $k-N$. The information in the measurement equation for time $k-N$ is therefore accounted for by \hat{x}_{k-N} , and the measurement y_{k-N} is therefore not included in the MHE. Instead, the measurement y_k is included, as this is assumed to be available at time k . Other MHE formulations, such as the one used in [1], instead provide the *a priori* state estimate $x_{k|k-1}$, which leads to slight differences in the MHE formulation.

2.1. Filtering and smoothing formulations of the MHE

The Kalman filter efficiently represents the 'prior knowledge' using the *a priori* and *a posteriori* state estimate covariance matrices together with the corresponding state estimates. Unfortunately, such a simple and exact representation of the prior knowledge (of what has occurred prior to the start of the backward-looking estimation horizon) is not readily available for MHE when constraints (or nonlinearity) are present. However, to simple approximations are available [5]:

- The filtering formulation of MHE.
- The smoothing formulation of MHE.

Properties of these two formulations are discussed in [5, 1, 6], and stability analysis is provided in [4, 1].

2.1.1. The filtering formulation of MHE

In the filtering formulation, the state estimate \hat{x}_{k-N} in (1) is the *a posteriori* state estimate at time $k - N$, $\tilde{x}_{k-N|k-N}$, and the corresponding weight S is the inverse of the Kalman filter *a posteriori* covariance matrix at time $k - N$, *i.e.*, $S = \Pi_{k-N|k-N}^{-1}$. This covariance matrix is easily available by running a Kalman filter covariance updates in parallel with the MHE - as this would incur a modest computational cost compared to the MHE optimization solution. If some transient performance loss is acceptable, the steady state Kalman filter covariance matrix may also be used.

2.1.2. The smoothing formulation of MHE

In the smoothing formulation, the state estimate \hat{x}_{k-N} in (1) is the *a posteriori* state estimate of the state at time $k - N$ available at time $k - 1$, $\tilde{x}_{k-N|k-1}$. The expression for the arrival cost weight S in (1) is more complex when using the smoothing formulation. This will be addressed in more detail in the next section.

3. THE ARRIVAL COST IN THE SMOOTHING FORMULATION

The term $(\hat{x}_{k-N} - \tilde{x}_{k-N})^T S (\hat{x}_{k-N} - \tilde{x}_{k-N})$ in (1) is often termed the 'arrival cost', as it represents the cost related to information from times prior to the estimation horizon. It is shown in [4], that the arrival cost weight S in the smoothing formulation consists of two terms:

1. One term reflecting the covariance of the smoothed estimate of \hat{x}_{k-N} , and
2. One term preventing the measurement information inside the moving horizon to be used twice (both in the estimate \hat{x}_{k-N} supplied, and when performing the MHE calculations). The first term then corresponds to $(\hat{x}_{k-N} - \tilde{x}_{k-N})^T S_1 (\hat{x}_{k-N} - \tilde{x}_{k-N})$, with $S_1 = \Pi_{k-N|k-1}^{-1}$, where $\Pi_{k-N|k-1}$ is the covariance of the smoothed estimate of x_{k-N} accounting for all measurements up to and including time $k - 1$. The determination of $\Pi_{k-N|k-1}$ is described in, *e.g.* [7].

To calculate the second term, we need the covariance of the estimate of the measurement sequence $Y_{N-1} = [y_{k-N+1}^T \quad \cdots \quad y_{k-1}^T]^T$, given x_{k-N} .

Manipulating the model equations, we get

$$\begin{aligned}
 Y_{N-1} &= \begin{bmatrix} CA \\ CA^2 \\ \vdots \\ CA^{N-2} \\ CA^{N-1} \end{bmatrix} x_{k-N} + \begin{bmatrix} 0 & 0 & \cdots & 0 & CB \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & CB & \cdots & CA^{N-3}B \\ 0 & CB & \cdots & CA^{N-3}B & CA^{N-2}B \end{bmatrix} \begin{bmatrix} u_{k-1} \\ u_{k-2} \\ \vdots \\ u_{k-N+1} \\ u_{k-N} \end{bmatrix} + \\
 &+ \begin{bmatrix} 0 & 0 & \cdots & 0 & CE \\ 0 & 0 & \ddots & CE & CAE \\ \vdots & \ddots & \cdots & \ddots & \vdots \\ 0 & CE & \cdots & CA^{N-3}E & CA^{N-2}E \end{bmatrix} \begin{bmatrix} w_{k-1} \\ w_{k-2} \\ \vdots \\ w_{k-N+1} \\ w_{k-N} \end{bmatrix} + \begin{bmatrix} 0 & I_{(N-1) \cdot n_y \times (N-1) \cdot n_y} \end{bmatrix} \begin{bmatrix} v_k \\ v_{k-1} \\ \vdots \\ v_{k-N+1} \end{bmatrix}
 \end{aligned} \tag{2}$$

Noting that Y_{N-1} is independent of v_k , this may be reformulated as

$$Y_{N-1} - \mathcal{O}_{N-1} x_{k-N} - \tilde{B} u = \tilde{E} w + \tilde{I} v \tag{3}$$

We are starting from a given x_{k-N} , and u is the vector of (known) plant inputs. Y_{K1} is therefore the only uncertain term on the left hand side of ((3)). The variance of Y_{K-1} is therefore

$$S_2^{-1} = \tilde{E} \tilde{W} \tilde{E}^T + \tilde{I} \tilde{V} \tilde{I}^T \tag{4}$$

where $\tilde{W} = \text{diag}\{W\}$ and $\tilde{V} = \text{diag}\{V\}$. The arrival cost for the smoothing formulation of the MHE therefore becomes:

$$\begin{aligned}
 \Gamma(\tilde{x}_{k-N}) &= (\hat{x}_{k-N} - \tilde{x}_{k-N})^T S_1 (\hat{x}_{k-N} - \tilde{x}_{k-N}) - \\
 &- (Y_{N-1} - \tilde{B} u - \mathcal{O}_{N-1} \tilde{x}_{k-N})^T S_2 \times \\
 &\times (Y_{N-1} - \tilde{B} u - \mathcal{O}_{N-1} \tilde{x}_{k-N})
 \end{aligned} \tag{5}$$

This raises two points:

1. The expression for S_2 above yields a result that is slightly different from the one given in [1], there seems to be a slight misprint in that publication (the diagonal elements are the same, after accounting for changes in assumptions and nomenclature).

2. The above derivation assumes u to be independent of both x_{k-N} and w . This is not the case. In the linear, unconstrained case u is fully determined by x_{k-N} , w and v . For this situation, with a normal QP formulation of the MPC, the MPC will correspond to the LQ optimal controller. It therefore should be clarified how the arrival cost for the smoothing formulation should be modified to reflect this. This point is not clear from [1], since the problem statement therein does not include a manipulated variable u .

The latter point is the subject of the next section.

4. ACCOUNTING FOR CONTROL IN THE CALCULATION OF THE ARRIVAL COST

We here assume that a ('conventional') QP-based MPC formulation is used, for which the controller in the unconstrained case corresponds to the LQ-optimal state feedback controller, denoted K . Similarly, the MHE formulation used in the unconstrained case corresponds to a Kalman filter, with steady-state filter gain L . The plant model, together with the (unconstrained) control and estimation, then yields

$$x_{k+1|k+1} = Ax_{k|k} + Bu_k + L(y_k - Cx_{k|k}) + w_k = (A + BK)x_{k|k} + Lv_k + w_k \quad (6)$$

Starting from a given value of x_{k-N} , we then obtain

$$\begin{aligned} u_{k-N} &= Kx_{k-N} \\ u_{k-N+1} &= K(A + BK)x_{k-N} + KLv_{k-N} + Kw_{k-N} \\ u_{k-N+2} &= K(A + BK)^2 x_{k-N} + K(A + BK)Lv_{k-N} + \\ &\quad + KLv_{k-N+1} + K(A + BK)w_{k-N} + Kw_{k-N+1} \\ u_{k-N+i} &= K(A + BK)^i x_{k-N} + K[I \quad (A + BK) \quad \cdots \quad (A + BK)^{i-1}] \times \\ &\quad \times \begin{bmatrix} Lv_{k-N+i} + w_{k-N+i} \\ \vdots \\ Lv_{k-N} + w_{k-N} \end{bmatrix} \end{aligned} \quad (7)$$

Thus, we get

$$\begin{aligned} \begin{bmatrix} u_{k-1} \\ u_{k-2} \\ \vdots \\ u_{k-N} \end{bmatrix} &= \tilde{K} \begin{bmatrix} (A + BK)^{N-1} \\ (A + BK)^{N-2} \\ \vdots \\ (A + BK) \\ I \end{bmatrix} x_{k-N} + \\ &+ \tilde{K} \times \begin{bmatrix} I & (A + BK) & (A + BK)^{N-2} & (A + BK)^{N-1} \\ 0 & I & \cdots & (A + BK)^{N-2} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ddots & I & (A + BK) \\ 0 & \cdots & 0 & I \end{bmatrix} \times \end{aligned} \quad (8)$$

$$\begin{aligned} &\times \begin{bmatrix} Lv_{k-1} + w_{k-1} \\ \vdots \\ Lv_{k-N} + w_{k-N} \end{bmatrix} = \\ &= \tilde{K}A_K x_{k-N} + \tilde{K}B_K w + \tilde{K}B_K \tilde{L}v_{k-1} \end{aligned} \quad (9)$$

where $\tilde{K} = \text{diag}\{K\}$ and $\tilde{L} = \text{diag}\{L\}$. Substituting (9) into (3) we obtain

$$Y_{N-1} - (\mathcal{O}_{N-1} + \tilde{K}A_K)x_{k-N} = (\tilde{E} + \tilde{K}B_K)w + (\tilde{I} + \tilde{K}B_K\tilde{L})v_{k-1} \quad (10)$$

This corresponds to the arrival cost

$$\begin{aligned} \Gamma(\tilde{x}_{k-N}) = & (\hat{x}_{k-N} - \tilde{x}_{k-N})^T S_1 (\hat{x}_{k-N} - \tilde{x}_{k-N}) - (Y_{N-1} - (\mathcal{O}_{N-1} + \tilde{B}\tilde{K}A_K)\tilde{x}_{k-N})^T S_2 \\ & \times (Y_{N-1} - (\mathcal{O}_{N-1} + \tilde{B}\tilde{K}A_K)\tilde{x}_{k-N}) \end{aligned} \quad (11)$$

with

$$S_2^{-1} = (\tilde{E} + \tilde{B}\tilde{K}B_K)\tilde{W}(\tilde{E} + \tilde{B}\tilde{K}B_K)^T + (\tilde{I} + \tilde{B}\tilde{K}B_K\tilde{L})\tilde{V}(\tilde{I} + \tilde{B}\tilde{K}B_K\tilde{L})^T \quad (12)$$

4. DISCUSSION AND CONCLUSIONS

This note addresses two issues related to the smmothing formulation of MHE. The arrival cost formulation in (11 - 12) is based on the unconstrained solution to the MPC and MHE problems. This has the advantage of using values that can be pre-computed at the design stage. On the other hand, the presence of constraints is the main reason for introducing MPC and MHE. It is well known from the theory of explicit MPC [8] that when constraints are active, the optimal input can be calculated as

$$u_k = K_i x_k + k_i \quad (13)$$

where the index i reflects the set of constraints that are active at the optimum of the MPC problem for the state x_i . A similar expression can be derived for the optimal MHE solution when constraints are active. Whereas exploring the state space to find the explicit MPC controller (all the K_i 's and k_i 's) quickly becomes unmanageable even for modest size MPC problems, it would be possible to find the current K_i and k_i quite easily when the solution to the current MPC problem is available. One could consider using this information in real time to find an improved version of (12). Conceptually, this should not be too hard, although it would add to the required notational complexity. A more challenging problem would be how to handle cases when the state is at the border between two different sets of active constraints, as in such cases the optimal input will not be differentiable with respect to the state.

A more mundane - but nonetheless relevant - task for further work is to study the importance of the issues raised in this note, for realistic MPC/MHE problems.

REFERENCES

1. C. V. Rao, J. B. Rawlings, and J. H. Lee, "Constrained linear state estimation - a moving horizon approach," *Automatica*, vol. 37, pp. 1619–1629, 2001.
2. S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, pp. 733–764, 2003.
3. D. G. Robertson, J. H. Lee, and J. B. Rawlings, "A mov-ing horizon-based approach for least squares estimation," *AIChE Journal*, vol. 42, pp. 2209–2224, 1996.
4. C. V. Rao, "Moving horizon strategies for the constrained monitoring and control of nonlinear discrete-time systems," Ph.D. dissertation, Department of Chemical Engineering, University of Wisconsin - Madison, 2000.
5. P. K. Findiesen, "Moving horizon state estimation for discrete time systems," 1997.
6. F. Scibilia and M. Hovd, "Multirate moving horizon es-timation with erroneous infrequent measurements recovery," in *IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, Barcelona, Spain, 2009.
7. J. A. E. Bryson and Y.-C. Ho, *Applied Optimal Control*. Waltham, Massachusetts, USA: Blaisdell Publishing Company, 1969.
8. A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos, "The explicit linear quadratic regulator for con-strained systems," *Automatica*, vol. 38, pp. 3–20, 2002.

NAPOMENA O UBLAŽENOJ FORMULACIJI ESTIMACIJE POKRETNOG HORIZONTA

Morten Hovd

Ovaj rad ukazuje na dva neslaganja u definisanju kriterijumske funkcije u ublaženoj formulaciji estimacije pokretnog horizonta. Prvo neslaganje je očigledna greška u izrazu za kritierijumsku funkciju [1], dok je drugo u nepostojanju pretpostavke da su ulazni signali i sekvence šuma za pobuđivanje stanja nekorelaisani na horizontu estimacije.

Ključne reči: *MPC, pokretni horizont, estimacija*