Automatica 97 (2018) 167–171

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Technical communique

Metamorphic moving horizon estimation[☆]

He Kong [*], Salah Sukkarieh

Australian Centre for Field Robotics, The University of Sydney, NSW, 2006, Australia

ARTICLE INFO

Article history:
Received 15 March 2018
Received in revised form 22 March 2018
Accepted 5 July 2018

Keywords:
Least-squares estimation
State estimation
Constrained estimation
Recursive estimation

ABSTRACT

This technical communique considers a practical scenario where a classical estimation method might have already been implemented on a certain platform when one tries to apply more advanced techniques such as Moving horizon estimation (MHE). We are interested to utilize MHE to upgrade, rather than completely discard, the existing estimation technique. This immediately raises the question how one can improve the estimation performance gradually based on the pre-estimator. To this end, we propose a general methodology which incorporates the pre-estimator with a tuning parameter $\lambda \in [0, 1]$ into the quadratic cost functions that are usually adopted in MHE. We examine the above idea in two standard MHE frameworks that have been proposed in the existing literature. For both frameworks, when $\lambda = 0$, the proposed strategy exactly matches the existing classical estimator; when the value of λ is increased, the proposed strategy exhibits a more aggressive normalized forgetting effect towards the old data, thereby increasing the estimation performance gradually.

1. Introduction

MHE is an optimization-based framework to handle constraints in estimation (Chu, Chen, & Marquez, 2007; Chu, Keshavarz, Gorinevsky, & Boyd, 2012; Dissanayake, Sukkarieh, Nebot, & Durrant-Whyte, 2001; Ge & Kerrigan, 2016; Ko & Bitmead, 2007; Mahata & Söderström, 2004; Rawlings & Mayne, 2009). Various forms of MHE have been proposed in the literature. For example, in Kong and Sukkarieh (2018b) and Rao, Rawlings, and Lee (2001), the cost is optimized over the initial state and the process noise sequence. Some MHE frameworks estimate only the initial state (Alessandri, Baglietto, & Battistelli, 2003; Sui & Johansen, 2014; Sui, Johansen, & Feng, 2010). Other limited memory filtering methods such as the finite impulse response (FIR) filter have also been developed (Ahn, Shi, & Basin, 2016; Shmaliy, Zhao, & Ahn, 2017). Both MHE and FIR filters only use recent measurements within a time window. A major difference between them is that the information contained in the measurements outside the moving horizon is captured by the arrival cost in MHE (Rao et al., 2001), such information, however, is completely ignored in FIR filtering.

A situation that one often encounters when trying to apply MHE is that some traditional estimators might have already been implemented. For example, there would be some forms of Kalman filters embedded in today's most GPS devices. Replacing the existing estimation methods and related software and hardware by MHE is often time consuming and costly, if possible. A similar situation is faced by control engineers and this has motivated some works to combine the merits of predictive and linear methods (Hartley & Maciejowski, 2013; Kong, Goodwin, & Seron, 2012, 2013). Especially, in Kong et al. (2013), a general framework has been proposed to gradually improve the control performance using predictive control, incorporating an existing linear controller. The question considered in this paper is to propose a MHE framework to gradually improve the estimation performance based on a pre-estimator. As such, we borrow the concept that is proposed in Kong et al. (2013) for the control case, and consolidate the idea in two MHE frameworks (Alessandri et al., 2003; Rao et al., 2001; Sui et al., 2010). For both frameworks, we propose a methodology that can gradually improve the estimation performance with MHE, incorporating an existing estimator. The above result is achieved by the introduction of cost functions parameterized by $\lambda \in [0, 1]$. When λ changes, optimizing the cost functions renders a new estimator, we thus term the framework metamorphic[1] MHE (MMHE).

An advantage of the proposed framework is that it can upgrade an existing classical method using MHE, thereby obtaining the constraint handling capabilities of MHE and avoiding the trouble involved in a completely new design of the estimator. A disadvantage of the proposed framework, compared to classical estimation

☆ This work was supported in part by the Australian Centre for Field Robotics and in part by the Faculty of Engineering and Information Technologies, University of Sydney, Australia. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Tong Zhou under the direction of Editor André L. Tits.

* Corresponding author.
E-mail addresses: h.kong@acfr.usyd.edu.au (H. Kong), salah.sukkarieh@sydney.edu.au (S. Sukkarieh).

[1] As noted in Kong et al. (2013), metamorphism is the recrystallization of pre-existing rocks due to physical/chemical changes.

https://doi.org/10.1016/j.automatica.2018.08.018
0005-1098/

techniques, is that one has to solve an optimization problem at each sampling instant. Due to limited space, proofs of all theorems or lemmas are not included and details can be found in Kong and Sukkarieh (2018a). The remainder of the paper is organized as follows. In Sections 2–3, we embellish the metamorphic concept with the MHE frameworks of Alessandri et al. (2003), Rao et al. (2001) and Sui et al. (2010), and present some analysis thereof, respectively. Section 4 uses simulation results to illustrate the proposed strategy. Conclusions are drawn in Section 4.

Notation: $[a_1, \dots, a_n]$ denotes $[a_1^{\mathrm{T}} \cdots a_n^{\mathrm{T}}]^{\mathrm{T}}$, where $a_1, \dots, a_n$ are scalars, vectors or matrices of proper dimensions; $\mathscr{I}_i^j$ denotes the set of integers between i and j; a set $\mathscr{U} \subset \mathbf{R}^n$ is a $\mathscr{C}$-set if it is a compact, convex set containing the origin in its (non-empty) interior; $diag(M_1, \dots, M_s)$ denotes a block diagonal matrix with $M_1, \dots, M_s$ as its block diagonal entries, and $diag_N(\cdot)$ denotes a block diagonal matrix with N blocks. $\mathbf{1}_n$ denotes a n-dimensional column vector with all its elements as 1.

2. Metamorphic MHE

2.1. Embellishing a pre-estimator into MHE

Consider the MHE framework in Rao et al. (2001) for the system

$$x_{k+1} = Ax_k + Gw_k,\ \ y_k = Cx_k + \nu_k \tag{1}$$

where $x_k \in \mathcal{X} \subset \mathbf{R}^n$, $w_k \in \mathcal{W} \subset \mathbf{R}^m$ and $\nu_k \in \mathcal{V} \subset \mathbf{R}^p$, respectively; the pair (A, C) is assumed to be observable; the set $\mathcal{X}$ is compact and convex; $\mathcal{W}$ and $\mathcal{V}$ are both $\mathscr{C}$-sets. Assume that for (1), we have the following Luenberger observer or stationary Kalman filter:

$$\widetilde{x}_{k+1} = A\widetilde{x}_k + L(y_k - \widetilde{y}_k),\ \ \widetilde{y}_k = C\widetilde{x}_k, \tag{2}$$

where L is chosen such that $A_L = A - LC$ is Schur stable. Define $e_{k+1} = x_{k+1} - \widetilde{x}_{k+1}$. Then it holds that

$$e_{k+1} = A_L e_k + \vartheta_k, \tag{3}$$

with $\vartheta_k = Gw_k - L\nu_k \in \mathcal{Q} = G\mathcal{W} \ominus L\mathcal{V}$. Note that $\mathcal{Q}$ is also a $\mathscr{C}$-set since both $\mathcal{W}$ and $\mathcal{V}$ are $\mathscr{C}$-sets. Given $\rho(A_L) < 1$, there exists a robust positively invariant $\mathscr{C}$-set $\mathscr{E}$ satisfying $A_L\mathscr{E} \oplus \mathcal{Q} \subseteq \mathscr{E}$ for system (3) (see Rawlings & Mayne, 2009, pp. 377). Define $x_k^e = [\widetilde{x}_k, e_k]$ and $\overline{w}_k = [w_k, \nu_k]$. From (1), (2), and (3), we have the augmented system

$$x_{k+1}^e = A_e x_k^e + G_e \overline{w}_k,\ \ y_k = C_e x_k^e + \nu_k, \tag{4}$$

where

$$A_e = \begin{bmatrix} A & LC \\ 0 & A_L \end{bmatrix},\ G_e = \begin{bmatrix} 0 & L \\ G & -L \end{bmatrix},\ C_e = \begin{bmatrix} C & C \end{bmatrix}.$$

For (4), we have $x_k^e \in \overline{\mathcal{X}}$, $\overline{w}_k \in \overline{\mathcal{W}}$, where $\overline{\mathcal{X}} = \mathcal{X} \times \mathscr{E}$, $\overline{\mathcal{W}} = \mathcal{W} \times \mathcal{V}$. The variables $(x_k^e, \overline{w}_k, y_k, \nu_k)$ in (4) represent the parameters of the *real* augmented process, and we denote $(\chi_k^e, \overline{\omega}_k, \eta_k, \upsilon_k)$ and $(\widehat{x}_k^e, \widehat{\overline{w}}_k, \widehat{y}_k, \widehat{\nu}_k)$ as the corresponding decision variables and the optimal solutions in the optimization, respectively. For system (4), consider the constrained estimation problem

$$\overline{\mathcal{M}}_T : \begin{cases} \min\limits_{\chi_{T-N}^e, \overline{\omega}_{T-N}^{T-1}} \overrightarrow{\phi}_T \text{ s.t. } \chi_k^e \in \overline{\mathcal{X}}, k \in \mathscr{I}_{T-N}^T \\ \overline{\omega}_k \in \overline{\mathcal{W}}, \upsilon_k \in \mathcal{V},\ k \in \mathscr{I}_{T-N}^{T-1} \end{cases}, \tag{5}$$

where $\chi_k^e = \chi^e(k; \chi_{T-N}^e, \overline{\omega}_{T-N}^{k-1})$, $\upsilon_k = y_k - C\chi_k^e$, $\lambda \in [0, 1]$, $\overline{\omega}_{T-N}^{T-1} = \{\overline{\omega}_i\}_{i=T-N}^{T-1}$,

$$\overrightarrow{\phi}_T = \lambda(\chi_{T-N}^e - \widehat{x}_{T-N}^{em})\Phi_{T-N}^{-1}(\chi_{T-N}^e - \widehat{x}_{T-N}^{em}) + \lambda\overrightarrow{\phi}_{T-N}^* + \sum_{k=T-N}^{T-1}\left[(1-\lambda)\overline{\omega}_k^{\mathrm{T}} M\overline{\omega}_k + \lambda(\upsilon_k^{\mathrm{T}} R^{-1}\upsilon_k + \omega_k^{\mathrm{T}} Q^{-1}\omega_k)\right],$$

in which $R, Q, M > 0$; Φ_{T-N} is a positive definite matrix that is to be discussed in the sequel; $\overrightarrow{\phi}_{T-N}^*$ is the optimal cost of (5) at time $T-N$, and thus is a constant parameter and can be safely ignored in the optimization; $\widehat{x}_{T-N}^{em}$ is the optimal moving horizon state prediction at time $T-N$, i.e., $\widehat{x}_{T-N}^{em} = \widehat{x}_{T-N|T-N-1}^{em}$. When $\lambda = 0$, it holds that $\overrightarrow{\phi}_T = \sum_{k=T-N}^{T-1}\overline{\omega}_k^{\mathrm{T}} M\overline{\omega}_k$. Given $0 \in \overline{\mathcal{W}}$, the optimal decision variables are $\widehat{\overline{w}}_i = 0$, for $i \in \mathscr{I}_{T-N}^{T-1}$. In this case, the optimal decision variables $(\widehat{x}_k^e, \widehat{\overline{w}}_k, \widehat{y}_k, \widehat{\nu}_k)$ satisfy $\widehat{x}_{k+1}^e = A_e\widehat{x}_k^e$, $\widehat{y}_k = C_e\widehat{x}_k^e$, i.e., the strategy reduces to a deterministic observer with the same gain as the pre-estimator (2). When $\lambda = 1$, one has

$$\overrightarrow{\phi}_T = (\chi_{T-N}^e - \widehat{x}_{T-N}^{em})\Phi_{T-N}^{-1}(\chi_{T-N}^e - \widehat{x}_{T-N}^{em}) + \sum_{k=T-N}^{T-1}\left[\overline{\omega}_k^{\mathrm{T}}\overline{Q}\overline{\omega}_k + \upsilon_k^{\mathrm{T}} R^{-1}\upsilon_k\right],$$

with $\overline{Q} = diag(Q^{-1}, 0) \geq 0$. Note that this is not a well-posed problem since positive definiteness is required for the weight on $\overline{\omega}_k$. Therefore, we will only consider the cases of $\lambda \in (0, 1)$. Dividing $\overrightarrow{\phi}_T$ by λ gives us:

$$\overline{\phi}_T = \lambda^{-1}\overrightarrow{\phi}_T = \sum_{k=T-N}^{T-1}\left[\overline{\omega}_k^{\mathrm{T}} Q_e^{-1}\overline{\omega}_k + \upsilon_k^{\mathrm{T}} R^{-1}\upsilon_k\right] + (\chi_{T-N}^e - \widehat{x}_{T-N}^{em})\Phi_{T-N}^{-1}(\chi_{T-N}^e - \widehat{x}_{T-N}^{em}), \tag{6}$$

where

$$Q_e^{-1} = \frac{1-\lambda}{\lambda}M + diag(Q^{-1}, 0) > 0,$$

given $\lambda \in (0, 1)$ and $M > 0$. Moreover, one can consider a constrained estimation problem which replaces $\overrightarrow{\phi}_T$ in (5) with $\overline{\phi}_T$ (6). Note that doing so will not affect the optimal solution or stability.

2.2. Stability ingredients for metamorphic MHE

When one replaces $\overrightarrow{\phi}_T$ in (5) with $\overline{\phi}_T$(6), the associated ARE for system (4) is

$$\Phi_T = G_e Q_e G_e^{\mathrm{T}} + A_e\Phi_{T-1}A_e^{\mathrm{T}} - A_e R_e A_e^{\mathrm{T}} \tag{7}$$

with Φ_0 as the initial condition, $R_e = \Phi_{T-1}C_e^{\mathrm{T}}(R + C_e\Phi_{T-1}C_e^{\mathrm{T}})^{-1}C_e\Phi_{T-1}$, and Q_e being defined in (6). Without constraints, one obtains the metamorphic Kalman filter,

$$\widehat{x}_T^e = A_e\widehat{x}_{T-1}^e + L_e(y_T - C_eA_e\widehat{x}_{T-1}^e),$$

where $L_e = A_e\Phi_{T-1}C_e^{\mathrm{T}}(R + C_e\Phi_{T-1}C_e^{\mathrm{T}})^{-1}$. We have the following results regarding Φ_T (7).

Lemma 1. *Assume that Q, R, M, Φ_0 are positive definite, (A, C) is observable. For $\lambda \in (0, 1)$, we have $\Phi_k > 0$, for all $k \geq 0$, if either of the following two conditions is satisfied: (i) $(A, GQ^{-1/2})$ is controllable, and $\Phi_0 \geq \Phi_\infty$; (ii) G and L are both nonsingular.*

Theorem 1. *Assume that Φ_0 is chosen independently of λ, and Φ_T is updated according to the ARE (7). Suppose either of the two conditions in Lemma 1 is satisfied, i.e., $\Phi_k > 0$, for $k \geq 0$, then for $\lambda \in (0, 1)$, we have $\frac{d\Phi_k}{d\lambda} \geq 0$.*

When $\Phi_0 > \Phi_\infty$, the assumption that Φ_0 is chosen independently of λ can be satisfied by selecting a sufficiently large Φ_0. Therefore, the results in Theorem 1 can be applied for this case. When $\Phi_0 = \Phi_\infty$, $\Phi_k = \Phi_\infty$, for all $k \geq 0$, e.g., Φ_0 is dependent of λ, as Φ_∞ is. We have the following results complementary to Theorem 1.