Automatica 47 (2011) 2016–2022

Contents lists available at SciVerse ScienceDirect

Automatica

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

Brief paper

A global piecewise smooth Newton method for fast large-scale model predictive control[☆]

Panagiotis Patrinos[a], Pantelis Sopasakis[b], Haralambos Sarimveis[b,*]

[a] Department of Mechanical and Structural Engineering, University of Trento, Via Mesiano 77, 38100 Trento, Italy

[b] School of Chemical Engineering, National Technical University of Athens, 9 Heroon Polytechneiou Street, 15780 Zografou Campus, Athens, Greece

ARTICLE INFO

Article history:
Received 29 March 2010
Received in revised form
2 February 2011
Accepted 2 March 2011
Available online 28 June 2011

Keywords:
Model predictive control
Online optimization
Large-scale systems

ABSTRACT

In this paper, the strictly convex quadratic program (QP) arising in model predictive control (MPC) for constrained linear systems is reformulated as a system of piecewise affine equations. A regularized piecewise smooth Newton method with exact line search on a convex, differentiable, piecewise-quadratic merit function is proposed for the solution of the reformulated problem. The algorithm has considerable merits when applied to MPC over standard active set or interior point algorithms. Its performance is tested and compared against state-of-the-art QP solvers on a series of benchmark problems. The proposed algorithm is orders of magnitudes faster, especially for large-scale problems and long horizons. For example, for the challenging crude distillation unit model of Pannocchia, Rawlings, and Wright (2007) with 252 states, 32 inputs, and 90 outputs, the average running time of the proposed approach is 1.57 ms.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Model predictive control (MPC) owes its popularity to the fact that it is one of the few control methodologies that can stabilize linear or nonlinear systems subject to hard input and state constraints. The popularity of the method is due to the strong theoretical background that has been developed over the past few years (Mayne, Rawlings, Rao, & Scokaert, 2000; Rawlings & Mayne, 2009), the development of efficient optimization algorithms and codes, and the substantial increase in computational power. For linear systems with linear constraints and convex quadratic stage cost, the resulting optimal control problem can be formulated as a quadratic program (QP), and off-the-shelf QP solvers allow the application of MPC for small-scale to medium-scale processes with a moderate prediction horizon and slow dynamics.

However, the repeated solution of the finite-horizon optimal control online remains the main bottleneck of the methodology. Specifically, for high sampling rates and large-scale systems the computational time needed to solve the QP problem becomes a limiting factor of the method.

There are two major categories of QP algorithms: active set and interior point methods. Active set methods try to identify the set of inequality constraints that are satisfied as equalities at the optimal solution, and they can be classified into primal and dual feasible algorithms. Interior point algorithms move from an interior point of the feasible set towards the optimal solution by following the so-called central path. Specifically, Rao, Wright, and Rawlings (1998) applied Mehrotra's predictor–corrector algorithm (Mehrotra, 1992) using a discrete-time Riccati recursion to solve the system of linear equations in MPC problems, thus considerably reducing the computational effort. Recently, Wang and Boyd (2010) proposed an infeasible start primal barrier method, in which the structured system of linear equations at each Newton iteration is solved efficiently using block elimination. Active set methods usually require a large number of computationally cheap iterations, while interior point methods need only a few but more expensive steps.

To overcome the limitations of the aforementioned techniques, some other alternatives have been proposed in the literature for solving MPC problems efficiently. Specifically, taking advantage of the simplicity of the constraints appearing in the dual of a strictly convex QP, Axehill and Hansson (2008) proposed a gradient projection method (Bertsekas, 1999; Nocedal & Wright, 2006) applied to the dual QP of the MPC problem. Unlike the dual active set algorithm, gradient projection permits large changes of the working set. In Richter, Jones, and Morari (2009), the optimal gradient method of Nesterov (1983) is employed to solve MPC problems with lower and upper bound constraints on the manipulated variables only. In Cannon, Liao, and Kouvaritakis (2008), Pontryagin's minimum principle is employed to replace

[☆] The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Lalo Magni under the direction of Editor Frank Allgöwer.

* Corresponding author. Tel.: +30 210 7723237; fax: +30 210 7723138.
E-mail addresses: patrinos.panagiotis@ing.unitn.it (P. Patrinos), chvng@mail.ntua.gr (P. Sopasakis), hsarimv@central.ntua.gr (H. Sarimveis).

0005-1098/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2011.05.024

matrix factorizations during standard active set algorithms by recursions of state and co-state variables, resulting in linear complexity per iteration with respect to the prediction horizon for input constrained linear MPC.

On a different approach, the observation that MPC for constrained linear systems can be formulated as a parametric QP (Bemporad, Morari, Dua, & Pistikopoulos, 2002) has enabled the offline solution of the optimal control problem and the explicit calculation of the MPC controller as a piecewise affine mapping of the measured state. However, the applicability of parametric programming is limited to small-scale to medium-scale systems and prediction horizons. To overcome the limitation of parametric MPC, Ferreau, Bock, and Diehl (2008) proposed an online active set strategy that exploits the solution information of the previous QP by moving along a homotopy path from the previous solution to the current solution while adding and dropping constraints like a usual active set algorithm. Based on the observation that in real time the system traverses only a small fraction of those critical regions, Pannocchia et al. (2007) proposed the partial enumeration method. As the name suggests, only a small number of active sets and expressions of the piecewise affine (PWA) control law are computed offline via simulations and stored in a look-up table.

In the present work, we show how MPC for linear systems with arbitrary polyhedral state and control constraints reduces to finding a zero of a system of piecewise affine (PWA) equations. One could then apply a piecewise smooth Newton method (Facchinei & Pang, 2003; Kojima & Shindo, 1986) in order to find the MPC control law. The algorithm converges in quadratic rate locally. In order to globalize the Newton method and obtain convergence from any starting point, whether it is feasible or not for the QP, one can follow the route used in most of the globalized versions of Newton method and perform a line search based on a merit function. For the MPC problem, one can easily obtain a convex, continuously differentiable, piecewise quadratic merit function, i.e. a convex quadratic spline, whose minimizers are the zeros of the system of the PWA equations. Therefore, MPC essentially reduces to the unconstrained minimization of a convex quadratic spline. In this work, we apply the technique developed by Li and Swetits (1997, 1999) to solve the unconstrained minimization problem, with some modifications that speed up convergence. It is worth noting that such kind of reformulations have been proposed for support vector machine classification with very encouraging results (Mangasarian, 2002). Furthermore, similar piecewise smooth Newton methods have been employed for the solution of Huber's M-estimation problems in linear regression (Chen & Pinar, 1998; Madsen & Nielsen, 1990) with very encouraging results.

The contribution of the paper lies in the application of reformulation techniques to solve QP problems arising in linear MPC. Furthermore, we offer new insights to the algorithm of Li and Swetits (1997, 1999) by establishing a connection with piecewise smooth Newton methods for nonsmooth equations (Section 4). Additionally, we propose some modifications with respect to the Newton approximation scheme and step-size selection (Section 6), that lead to significant convergence speed-up in MPC problems (see also Remark 1).

Exploiting the structure of the problem, it will be shown that the system of linear equations that needs to be solved at each iteration is positive semidefinite and of significantly smaller dimension than that of the dimensions of the original problem. The linear system of equations can be solved effectively by updating the Cholesky factor at each iteration. Furthermore, an exact line search can be performed very quickly, leading to very fast convergence rates and small computational times.

The proposed algorithm is compared against state-of-the-art solvers in random instances of MPC problems and various benchmark problems from the literature. The results are very encouraging, especially for large-scale systems and long prediction horizons.

2. Notation

The finite set of integers $\{1, \ldots, m\}$ is denoted by $\mathbb{N}_{[1,m]}$. For a set $\mathcal{I} \subseteq \mathbb{N}_{[1,m]}$, $\mathcal{I}^c$ denotes its complement in $\mathbb{N}_{[1,m]}$, i.e. $\mathcal{I}^c = \mathbb{N}_{[1,m]} \setminus \mathcal{I}$. If $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a matrix, $\mathbf{c} \in \mathbb{R}^m$ is a vector and $\mathcal{I} \subseteq \mathbb{N}_{[1,m]}$, $\mathcal{J} \subseteq \mathbb{N}_{[1,n]}$ are ordered subsets, then $\mathbf{A}_{\mathcal{I}\cdot}$ is the matrix formed by the rows of $\mathbf{A}$ whose indices are in $\mathcal{I}$, $\mathbf{A}_{\mathcal{I}\mathcal{J}}$ denotes the matrix formed by the rows and columns of $\mathbf{A}$ whose indices are in $\mathcal{I}$ and $\mathcal{J}$, respectively,and $c_{\mathcal{I}}$ denotes the vector formed by the elements of c whose indices are in $\mathcal{I}$. For a vector $\mathbf{y} \in \mathbb{R}^m$, $[\mathbf{y}]_+$ denotes a vector whose ith component is $\max\{\mathbf{y}_i, 0\}$. For a given pair of vectors $\boldsymbol{\ell} \in \mathbb{R}^m$, $\mathbf{u} \in \mathbb{R}^m$, with $\boldsymbol{\ell}_i \leqslant \mathbf{u}_i$ for $i \in [1, m]$, $\mathrm{mid}(\boldsymbol{\ell}, \mathbf{u}; \mathbf{y})$ denotes the vector whose ith component is $\max\{\min\{\mathbf{y}_i, \mathbf{u}_i\}, \boldsymbol{\ell}_i\}$ for any $\mathbf{y} \in \mathbb{R}^m$. This function is known as the mid function in the optimization community or the saturation function in the control community.

3. Model predictive control for constrained linear systems

We consider the following discrete-time linear, time-invariant system:

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \tag{1}$$

where $\mathbf{x}(t) \in \mathbb{R}^{n_x}$ is the state and $\mathbf{u}(t) \in \mathbb{R}^{n_u}$ is the control input. We assume perfect state measurement and that $(\mathbf{A}, \mathbf{B})$ is stabilizable. The system's input and state should belong to the following polyhedral set:

$$Y = \{(\mathbf{u}, \mathbf{x}) \in \mathbb{R}^{n_u} \times \mathbb{R}^{n_x} | \mathbf{f}_{\min} \leqslant \mathbf{F_u}\mathbf{u} + \mathbf{F_x}\mathbf{x} \leqslant \mathbf{f}_{\max}\}, \tag{2}$$

where $\mathbf{F_u}$, $\mathbf{F_x}$, are $n_c \times n_u$ and $n_c \times n_x$ matrices, respectively, and $\mathbf{f}_{\min} \in \mathbb{R}^{n_c}$, $\mathbf{f}_{\max} \in \mathbb{R}^{n_c}$ with $-\infty \leqslant \mathbf{f}^i_{\min} < \mathbf{f}^i_{\max} \leqslant \infty$. Consider the following regulator problem:

$$\mathbb{P}_N(\mathbf{x}) \quad V_N^\star(\mathbf{x}) \triangleq \inf\{V_N(\mathbf{x}, \mathbf{u}) | \mathbf{u} \in \mathcal{U}_N(\mathbf{x})\}, \tag{3}$$

where the finite-horizon cost is

$$V_N(\mathbf{x}, \mathbf{u}) \triangleq \sum_{k=0}^{N-1} \ell(\mathbf{x}_k, \mathbf{u}_k) + V_f(\mathbf{x}_N) \tag{4}$$

and $\mathcal{U}_N : \mathbb{R}^{n_x} \rightrightarrows \mathbb{R}^{Nn_u}$ is

$$\mathcal{U}_N(\mathbf{x}) \triangleq \left\{ \mathrm{u} = (\mathbf{u}_0, \ldots, \mathbf{u}_{N-1}) \middle| \begin{array}{l} \mathbf{x}_0 = \mathbf{x} \\ \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \\ (\mathbf{x}_k, \mathbf{u}_k) \in Y, \; k \in \mathbb{N}_{[0,N-1]} \\ \mathbf{x}_N \in X_f \end{array} \right\}. \tag{5}$$

The stage cost is $\ell(\mathbf{x}, \mathbf{u}) \triangleq \frac{1}{2}(\mathbf{x}'\mathbf{Q}\mathbf{x} + \mathbf{u}'\mathbf{R}\mathbf{u})$, the terminal cost is $V_f(\mathbf{x}) \triangleq \frac{1}{2}\mathbf{x}'\mathbf{P}_f\mathbf{x}$, and $X_f \triangleq \{\mathbf{x} \in \mathbb{R}^{n_x} | \mathbf{k}_{\min} \leqslant \mathbf{H}_f\mathbf{x} \leqslant \mathbf{k}_{\max}\}$ is a polyhedral terminal set, where $\mathbf{H}_f$ is an $n_f \times n_x$ matrix and $\mathbf{k}_{\min} \in \mathbb{R}^{n_f}$, $\mathbf{k}_{\max} \in \mathbb{R}^{n_f}$ with $-\infty \leqslant \mathbf{k}^i_{\min} < \mathbf{k}^i_{\max} \leqslant \infty$. We assume that $\mathbf{R}$ is symmetric positive definite and that $\mathbf{Q}$, $\mathbf{P}_f$ are symmetric positive semidefinite matrices of compatible dimensions. Notice that we have used lower and upper bounds in the definition of Y and X_f. This assumption is not restrictive, since we allow upper or lower bounds to be equal to infinity. In fact, it is a common practice in MPC to have lower and upper bounds on the input and state variables. Furthermore, it will be shown next that this formulation significantly reduces the complexity of the proposed technique. Removing equality constraints, performing trivial algebraic manipulations, and omitting terms from the cost function that are independent of $\mathbf{u}$, the finite-horizon optimal control problem can be expressed as

$$\min \left\{ \frac{1}{2}\mathbf{u}'\mathbf{M}\mathbf{u} + \mathbf{c}(\mathbf{x})'\mathbf{u} | \mathbf{b}_{\min}(\mathbf{x}) \leqslant \mathbf{G}\mathbf{u} \leqslant \mathbf{b}_{\max}(\mathbf{x}) \right\}, \tag{6}$$

where $\mathbf{G} \in \mathbb{R}^{m\times Nn_u}$ with $m \triangleq Nn_c + n_f$. The matrices and vectors appearing in (6) can be found in Patrinos, Sopasakis, and Sarimveis (2010).

Download English Version:

https://daneshyari.com/en/article/10398830

Download Persian Version:

https://daneshyari.com/article/10398830

Daneshyari.com