INVESTIGATION OF SEMIDEFINITE RELAXATION IN MODEL PREDICTIVE CONTROL FORMULATION

P.E. Orukpe

Department of Electrical/Electronic Engineering, University of Benin, Nigeria.

Email: patience.orukpe01@imperial.ac.uk

ABSTRACT

Model Predictive Control (MPC) is a control technique that is widely used in the industries. In the problem formulation of MPC, the constrained optimization problem subject to inequality constraints is a standard optimization problem, know as quadratic programming (QP). In this paper, we present a method to exploit semidefinite relaxation in MPC problems. In particular, we use the solution of semidefinite relaxation to solve MPC problems instead of QP. We also use the solution value of the semidefinite relaxation as a bound for the objective function. We implemented our method using Matlab and provide some numerical results for a system using the MPC algorithm we have developed.

Keywords: Model predictive control, linear matrix inequality, quadratic programming, semidefinite relaxation, linear systems, discrete-time systems.

1.0 Introduction

Model Predictive Control (MPC) is a form of control in which the current control action is obtained by solving on-line, at each sampling instant, a finite horizon open-loop optimal control problem, using the current state of the plant as the initial state. The optimization yields an optimal control sequence and the first control in the sequence is applied to the plant. This is its main difference from conventional control which uses a pre-computed control law [1].

QP methods are widely used in applications of MPC and occur frequently in control theory, optimal filtering, operations research etc. Most MPC applications require a linear model to represent the process of interest over a moving time horizon with a quadratic objective or cost function to drive the controlled variables back to their set-points. In [2], QP was used for large-scale MPC. This paper focuses on optimization problems in MPC, which can be formulated as linear-quadratic problems. Such linear-quadratic optimization problems have been a recurring theme in semidefinite optimization for a very long time. A linear quadratic optimization problem is formulated in a semidefinite relaxation manner. Semidefinite programming (SDP) is one of the fastest developing branches of mathematical programming. The reason is twofold: efficient solution algorithms for SDP have come to light in the past few decades, and SDP finds applications in combinatorial optimization and engineering. One could easily be led to believe that the field of SDP originated in this decade. A glance at a bibliography of SDP papers indeed indicates an explosion of research effect, starting around 1991. A closer look reveals that interest in this class of problems is somewhat older, and dates back to the 1960's [3].

Semidefinite optimization is concerned with choosing a symmetric matrix to optimize a linear function subject to linear constraint and a further constraint that the matrix be positive semidefinite [4]. Semidefinite programs can be regarded as an extension of linear programming where the componentwise inequalities between vectors are replaced by matrix inequalities. SDP unifies several standard problems (e.g., linear and quadratic programming), and finds many applications in engineering and combinatorial optimization [5 - 6].

In this paper, we will exploit semidefinite relaxation in MPC framework and this work was motivated by a previous work in [7], where semidefinite relaxation was applied to an optimal production problem. In [7], semidefinite relaxation was applied to a linear program subject to inequality constraints, while in this paper we will consider quadratic program subject to inequality constraints in MPC setting. This paper is organized as follows: section 2 outlines the preliminaries needed in solving this kind of problem, section 3 presents the problem formulation of semidefinite relaxation, section 4 presents the simulation results and section 5 provides conclusion.

2.0 Preliminaries

In this section, we will look at the theory needed to carryout semidefinite relaxation.

2.1 Semidefinite Program

A wide variety of problems can be cast or recast as SDP problems, that is, problems of the form

$$\begin{aligned}&\min c^T x\\ &subject\ to\ F(x) \geq 0 \qquad (2.1)\end{aligned}$$

where

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i.$$

where $x \in \mathcal{R}^m$ is the variable. The problem data are the vector $c \in \mathcal{R}^m$ and $m+1$ symmetric matrices $F_0, \dots, F_m \in \mathcal{R}^{n \times n}$. The inequality sign in $F(x) \geq 0$ is positive semidefinite, i.e., $z^T F(x) z \geq 0$ for all $z \in \mathcal{R}^n$.

We call the inequality $F(x) \geq 0$ a linear matrix inequality and the problem (2.1) a semidefinite program. SDPs are convex optimization problems with a linear objective function and linear matrix inequality (LMI) constraints. Though the form of the SDP (2.1) appears very specialized, it turns out that it is widely encountered in systems and control theory, optimal production problem, merger of firms' problem and Huber penalty function [8].

2.2 Linear Matrix Inequality

The result in this paper is based on linear matrix inequalities LMIs. LMI techniques are now well-rooted as a unifying framework for formulating and solving problems in control theory with remarkable simplicity [9].

2.2.1 History: The most famous LMI in control is the Lyapunov inequality for the stability of LTI systems $A^T P + PA < 0$, which was originally considered over 100 years ago (i.e. 1890) and can be solved analytically by solving a set of linear equations. In the 1940's small LMIs were solved by hand, applying Lyapunov's methods to real control engineering problems (Lur'e, Postnikov etc.). Yakubovich was the first to make systematic use of LMIs along with the "*S-procedure*" for proving stability of nonlinear control systems. The works of Popov and Willems on optimal control outlined the relationship between the problem of absolute stability of automatic control, H_∞ theory and LMIs. Willems in particular, mentions LMIs as potentially powerful tools for system analysis:

The basic importance of the LMI seems to be largely unappreciated. It would be interesting to see whether or not it can be exploited in computational algorithms, for example [10].

For more detail, the reader is referred to [11].

2.2.2 Definition: A linear matrix inequality (LMI) has the form:

$$F(x) = F_0 + \sum_{i=1}^{l} x_i F_i > 0, \tag{2.2}$$

where $x \in \mathcal{R}^l$, $F_i = F_i^T \in \mathcal{R}^{n \times n}$ are given, and $F(x) > 0$ means that $F(x)$ is positive definite, i.e.

$$z^T F(x) z > 0, \forall z \neq 0, z \in \mathcal{R}^n. \tag{2.3}$$

The symmetric matrices F_i, $i = 1, \dots, l$ are fixed and x is the variable. Thus, $F(x)$ is an affine function of the elements of x. The set $\{x | F(x) > 0\}$ is *convex*, that is, the LMI (2.2) forms a convex constraint on x and need not have smooth boundary. This can be seen in the following [12]: let x and y be two vectors such that $F(x) > 0$ and $F(y) > 0$, and let $\lambda \in (0,1)$. Then

$$\begin{aligned} F(\lambda x + (1-\lambda) y) &= F_0 + \sum_{i=0}^{l} (\lambda x_i + (1-\lambda) y_i) F_i \\ &= F_0 + (1-\lambda) F_0 + \lambda \sum_{i=1}^{l} x_i F_i + (1-\lambda) \sum_{i=1}^{l} y_i F_i \end{aligned}$$

$$= \lambda F(x) + (1-\lambda)F(y)$$
$$> 0 \qquad (2.4)$$

Equation (2.4) is a *strict* LMI and is equivalent to a set of *n* polynomial inequalities in *x*, requiring only that *F*(*x*) be positive semidefinite and is referred to as *nonstrict* LMI. The strict LMI is feasible if the set $\{x|F(x) > 0\}$ is nonempty (a similar definition applies to nonstrict LMIs). Any feasible nonstrict LMI can be reduced to an equivalent strict LMI that is feasible by eliminating implicit equality constraints and then reducing the resulting LMI by removing any constant nullspace (see [5] page 19). Our focus is on nonstrict LMIs for convenience.

2.2.3 Multiple LMIs can be expressed as a single LMI: One of the advantages of representing control problems with LMIs is the ability to consider multiple control requirements by appending additional LMIs. Consider a set defined by *n* LMIs:

$$F_1(x) > 0; F_2(x) > 0; \dots; F_n(x) > 0. \qquad (2.5)$$

Then an equivalent single LMI is given by

$$F(x) = F_0 + \sum_{i=0}^{m} x_i F_i = diag\{F_1(x), F_2(x), \dots, F_n(x) > 0\} \qquad (2.6)$$

where

$$F_i = diag\{F_{i1}(x), F_{i2}(x), \dots, F_{in}(x) > 0, \forall i = 0, \dots, m\}. \qquad (2.7)$$

The following Lemma will be used in the derivation of the main result.

Lemma 1: The Schur complement converts a class of convex quadratic nonlinear inequalities that appears regularly in control problems to an LMI. The basic idea is as follows: the LMI

$$\begin{bmatrix} Q(x) & S(x) \\ S(x)^T & R(x) \end{bmatrix} > 0, \qquad (2.8)$$

where $Q(x) = Q(x)^T, R(x) = R(x)^T$, and *S*(*x*) depend affinely on *x*, is equivalent to the matrix inequalities

$$R(x) > 0, Q(x) - S(x)R(x)^{-1}S(x)^T > 0 \qquad (2.9)$$

or equivalently,

$$Q(x) > 0, R(x) - S(x)^T Q(x)^{-1} S(x) > 0 \qquad (2.10)$$

Proof: Assume

$$\begin{bmatrix} Q(x) & S(x) \\ S(x)^T & R(x) \end{bmatrix} > 0,$$

and define

$$F(u,v) = \begin{bmatrix} u \\ v \end{bmatrix}^T \begin{bmatrix} Q(x) & S(x) \\ S(x)^T & R(x) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \qquad (2.11)$$

Then,

$F(u, v) > 0, \forall\ [u, v] \neq 0.$ (2.12)

First, consider $u = 0$. Then

$F(0, v) = v^T R(x) v, \forall\ v \neq 0 \Rightarrow R(x) > 0.$

Next consider

$v = -R(x)^{-1} S(x) u$, with $u \neq 0$.

Then,

$$F(u, v) = u^T (Q(x) - S(x) R(x)^{-1} S(x)^T) u > 0, \forall u \neq 0$$
$$\Rightarrow Q(x) - S(x) R(x)^{-1} S(x)^T > 0.$$

Now assume

$Q(x) - S(x) R(x)^{-1} S(x)^T > 0, R(x) > 0$ (2.13)

with $F(u, v)$ defined in (2.11).

We will fix u and optimize over v.

$\nabla_v F^T = 2Rv + 2S^T u = 0.$ (2.14)

Since $R > 0$, (2.14) gives a single extrema $v = -R^{-1} S^T u$. Substituting this into (2.11) gives$F(u) = u^T (Q - SR^{-1} S^T) u$. Since $(Q - SR^{-1} S^T) > 0$, the minimum of $F(u)$ occurs for $u = 0$, which also implies that $v = 0$. Thus the minimum of $F(u, v)$ is positive definite. QED

2.2.4 S-procedure: The S-procedure greatly extends the usefulness of LMIs by allowing non-LMI conditions that commonly arise in non-linear systems analysis to be represented as LMIs (although with some conservatism). The S-procedure can be applied to quadratic functions as well as quadratic forms as is discussed in this section. Let $F_0, \ldots, F_p$ be quadratic functions of $\rho \in \mathcal{R}^n$:

$F_i(\rho) = \rho^T T_i \rho + 2u^T \rho + v_i, i = 0, \ldots, p; T_i = T_i^T.$ (2.15)

The existence of $\tau_1 \geq 0, \ldots, \tau_p \geq 0$, such that

$F_0(\rho) - \sum_{i=1}^{p} \tau_i F_i(\rho) \geq 0,\ \forall \rho$ (2.16)

implies that $F_i(\rho) \geq 0, \forall \rho$ such that

$F_i(\rho) \geq 0, i = 1, \ldots, p.$ (2.17)

This is true, because if $\tau_1 \geq 0, \ldots, \tau_p \geq 0$ exist such that (2.16) holds for all $F_i(\rho) \geq 0, i = 1, \ldots, p$.

Then,

$F_0(\rho) \geq \sum_{i=1}^{p} \tau_i F_i(\rho) \geq 0,\ \forall \rho.$ (2.18)

Remark 1: If the functions are affine, then (2.16) and (2.17) are equivalent; this is the *Farkas* [13] lemma.

Note that (2.16) can be written as

$$F(u,v) = \begin{bmatrix} T_0 & u_0 \\ u_0^T & v_0 \end{bmatrix} - \sum_{i=1}^{p} \tau_i \begin{bmatrix} T_i & u_i \\ u_i^T & v_i \end{bmatrix} > 0, \tag{2.19}$$

Hence the above S-procedure can be equivalently written in terms of quadratic forms and strict inequalities with similar proof to the one above. Let $T_0, \dots, T_p \in \mathcal{R}^{n\times n}$ be symmetric matrices. If there exists $\tau_i \geq 0, \dots, \tau_p \geq 0$ such that

$$T_0 - \sum_{i=1}^{p} \tau_i T_i > 0, \tag{2.20}$$

then, $\rho^T T_0 \rho > 0, \forall \rho \neq 0$ such that $\rho^T T_i \rho \geq 0, i = 1, \dots, p.$ (2.21)

3.0 Problem formulation of semidefinite relaxation

In this section, we formulate the MPC problem using the LMI approach explained in this paper, rather than QP method.

3.1 Basics of quadratic programming in MPC

In this section, we will remind the reader of the basic relations of MPC and fix notations.

Consider a nominal linear discrete-time system described by

$$\left.\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) \end{aligned}\right\} \tag{3.1}$$

where $x(k) \in \mathcal{R}^n$ and $u(k) \in \mathcal{R}^m$ are the state and control vectors respectively and $x(0)$ is assumed measured, $A \in \mathcal{R}^{n\times n}, B \in \mathcal{R}^{n\times m}$ and k belongs to the time set T of nonnegative integers: $T = [0,1,2,\dots]$. Since we are restricting our analysis to linear time invariant systems, we will take the initial time to be 0 for simplicity. When the horizon is updated, we assume that all indexing is changed so that the new initial time is still zero.

By iterating the model (3.1), we get a matrix which can be represented as

$$\tilde{x} = \tilde{A}x(0) + \tilde{B}\tilde{u}. \tag{3.2}$$

For input and state (or output) constraints of the form:

$$\left.\begin{aligned} Wu(k) &\leq w, \; k = 0,1,\dots \\ G_1 x(k) &\leq g, \; k = 1,2,\dots \end{aligned}\right\} \tag{3.3}$$

Using linear transformation, the state constraint is transformed into

$$G_b \tilde{u} \leq g_c - G_a x(k),$$

while the input constraint is transformed into

$$W_a \tilde{u} \leq w_c.$$

Augmented constraint for quadratic programming will be

$$\begin{bmatrix} G_b \\ W_a \end{bmatrix} \tilde{u} \leq \begin{bmatrix} g_c - G_a x(k) \\ w_c \end{bmatrix} \; or \; A_m \tilde{u} \leq b_m. \tag{3.4}$$

The standard quadratic cost function to be minimized is represented by

$$V(k) = \sum_{k=0}^{N-1} (x(k)^T Q x(k) + u(k)^T R u(k)), \tag{3.5}$$

in which Q and R are positive definite, symmetric weighting matrices. In model predictive control framework only the first control move $u(0)$ is injected into the plant. At the next sample time, (3.5) is solved with the new measured state as its initial condition. To construct the quadratic cost function, which can be solved on-line, we iteratively solve (3.5) and substitute (3.2) to obtain

$$V(k) \approx \tilde{u}^T H \tilde{u} + 2G\tilde{u}. \tag{3.6}$$

For further details see Chapter 2 of [14] and [15].

3.1 Method

The problem to be minimized is

$$\gamma_{opt} = \underbrace{min}_{\tilde{u}} \tilde{u}^T H \tilde{u} + 2G\tilde{u}$$

$$sub.to\ A_m\tilde{u} \le b_m \Rightarrow b_m - A_m\tilde{u} \ge 0. \tag{3.7}$$

H is positive semidefinite, $H \in \mathcal{R}^{n\times n}, b_m \in \mathcal{R}^{m\times 1}, A_m \in \mathcal{R}^{m\times n}$ and $G \in \mathcal{R}^{n\times 1}$. We found the equivalent of γ_{opt} in terms of new variables $\mu \ge 0$ and $\mu_0 \in \mathcal{R}$ as follows. We can write

$$\tilde{u}^T H \tilde{u} + 2G\tilde{u} = \mu^T(b_m - A_m\tilde{u}) - \mu^T(b_m - A_m\tilde{u}) + \tilde{u}^T H \tilde{u} + 2G\tilde{u} + \mu_0 - \mu_0. \tag{3.8}$$

The right hand side of (3.8) is rewritten as

$$\mu^T(b_m - A_m\tilde{u}) + [\tilde{u}^T \quad 1]\begin{bmatrix} H & G + \frac{1}{2}A_m^T\mu \\ G^T + \frac{1}{2}\mu^T A_m & \mu_0 - \mu^T b_m \end{bmatrix}\begin{bmatrix}\tilde{u} \\ 1\end{bmatrix} - \mu_0. \tag{3.9}$$

If we impose the constraint that $\mu \ge 0$, then for all $\tilde{u}$ such that $b_m - A_m\tilde{u} \ge 0$, we have that $\mu^T(b_m - A_m\tilde{u}) \ge 0$. It follows (3.9) that $-\mu_0$ is a lower bound on $\tilde{u}^T H \tilde{u} + 2G\tilde{u}$ for all $\mu \ge 0$ such that

$$\begin{bmatrix} H & G + \frac{1}{2}A_m^T\mu \\ G^T + \frac{1}{2}\mu^T A_m & \mu_0 - \mu^T b_m \end{bmatrix} \ge 0.$$

It follows that

$$\gamma_{opt} \ge \max -\mu_0$$

$$sub.to \left\{ \begin{matrix} \mu \ge 0 \\ \begin{bmatrix} H & G + \frac{1}{2}A_m^T\mu \\ G^T + \frac{1}{2}\mu^T A_m & \mu_0 - \mu^T b_m \end{bmatrix} \\ \mu_0 \in \mathcal{R} \end{matrix} \right\}. \tag{3.10}$$

For ease of coding in matlab, (3.10) was solved using

$$\gamma_{opt} \ge -\min \mu_0$$

$$sub.to \left\{ \begin{matrix} \mu \ge 0 \\ \begin{bmatrix} H & G + \frac{1}{2}A_m^T\mu \\ G^T + \frac{1}{2}\mu^T A_m & \mu_0 - \mu^T b_m \end{bmatrix} \\ \mu_0 \in \mathcal{R} \end{matrix} \right\}. \tag{3.11}$$

Problems (3.10) and (3.11) are convex. Note that since the original problem in (3.7) is convex, it follows that (3.10) and (3.11) give the exact solution [16].

4.0 Simulation Results

In this section, we present an example that illustrates the implementation of this semidefinite relaxation approach for constrained MPC. The simulations were performed on a PC with Pentium IV processor and we use the software LMI control Toolbox in the MATLAB environment to compute the solution.

Consider the distillation column used in [17], which was modified for our purpose to obtain the following discrete-time parameters at a sampling time of $T = 2min$:

$$A = \begin{bmatrix} 1-\frac{T}{54} & 0 & 0 & 0 \\ 0 & 1-\frac{T}{78} & 0 & 0 \\ 0 & 0 & 1-\frac{T}{114} & 0 \\ 0 & 0 & 0 & 1-\frac{T}{42} \end{bmatrix}; \; B = \begin{bmatrix} \frac{34T}{54} \\ \frac{31.6T}{78} \\ 0 \\ 0 \end{bmatrix}; C = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

The input constraint was $|u(k+i|k)| \leq 0.05$, while the state constraint was $|x(k+i|k)| \leq 1$. We specify the design parameters $diag(1,1,1,1)$ and $R = 0.00002$. Given an initial state $x(0) = [0.05 \quad 0 \quad 0.05 \quad 0]^T$, Figures 1 and 2 illustrate the simulation results for the control and output of the system under consideration. Furthermore, in the presence of constraints, no violations occurred. Thus the existence of a feasible solution ensures constraint satisfaction.

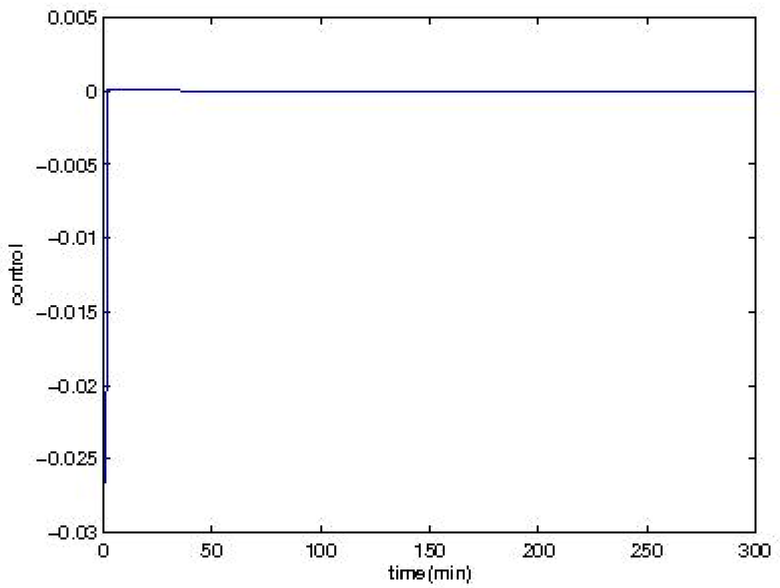

Figure 1: Control for the semidefinite relaxation method

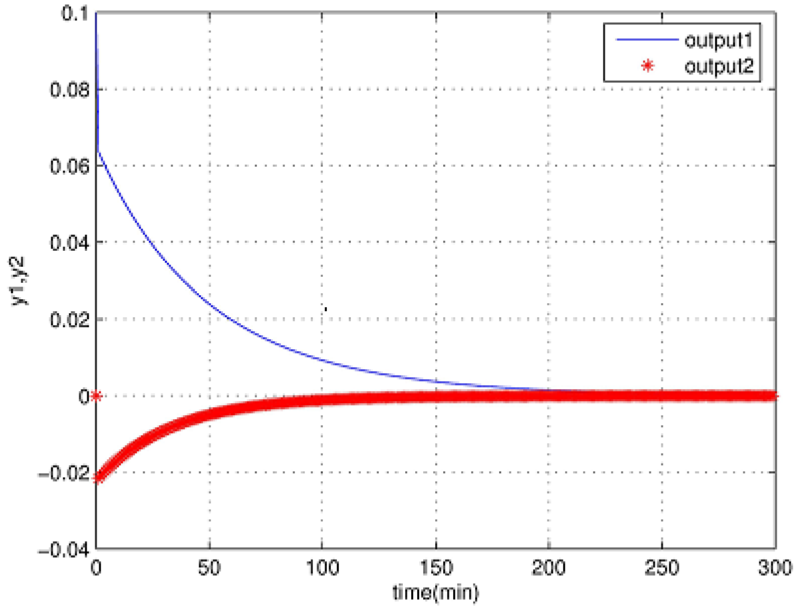

Figure 2: Output for the semidefinite relaxation method

6.0 Conclusion

In this paper, we developed a semidefinite relaxation algorithm for MPC problem subject to constraints on the input and states, which was solved in polynomial time. We have presented a method to use semidefinite relaxation within quadratic programming. The solution value of the relaxation is used as a bound for the corresponding quadratic programming cost function. The results were impressive with no constraint violation. In the example, the results were better in performance to the results obtained in [17], although a different technique was used, as our results converge to zero on time. The optimal cost obtained was the same for both the QP and semidefinite relaxation methods. The result guarantees a simpler approach that will be further investigated for other properties associated with MPC.

References

[1] Mayne D. Q., Rawlings J. B., Rao C.V. and Scokaert P. O. M. (2000) Constrained model predictive control: stability and optimality. Automatica, Vol. 36, pg 789 – 814.

[2] Bartlett R. A., Biegler L. T., Backstorm J. and Gopal V. (2002) Quadratic programming algorithms for large – scale model predictive control. Journal of Process Control, Vol. 12, pg 775 – 795.

[3] De Klerk E., Roos C. and Terlaky T. (1997) A short survey on semidefinite programming. Faculty of Technical Mathematics and Informatics, Delft University of Technology, The Netherlands, Tech. Report 28.

[4] Todd M. J. (2001) Semidefinite optimization. Acta Numerica, pg 515 – 560.

[5] Boyd S., Ghaoui L. E., Ferron E. and Balakrishnan V. (1994) Linear matrix inequalities in systems and control theory. Vol. 15 of Studies in Applied Mathematics, SIAM, Philiadelphia, PA.
[6] Vandenberge L. and Boyd S. (1996) Semidefinite programming. SAIM Review, Vol. 38, pg 49 – 95.
[7] Orukpe P. E. and Onohaebi S. O. (2006) Investigation of semidefinite relaxation in optimal production problem. Journal of Electrical and Electronic Engineering,Vol. 10, No. 1, pg 60 – 68.
[8] Joshi S. (2005) EE364: Convex optimization section 6. http://www.stanford.edu.
[9] Apkarian P. and Tuan H. D. (2000) Parameterized LMIs in control theory. SIAM Journal of Control Optimization, Vol. 38, No. 4, pg 1241 – 1264.
[10] Willems J. C. (1971) Least squares stationary optimal control and the algebraic riccati equation. IEEE Transcations on Automatic Control, Vol. 16, No. 6, pg 621 – 634.
[11] Boyd S., Ferron E., Balakrishnan V. and Ghaoui L. E. (1994) History of linear matrix inequalities in control theory. Proceedings of the American Control Conference, Baltimore, Maryland, pg 31 - 34.
[12] VanAntwerp J. G. and Braatz R. D. (2000) A tutorial on linear and bilinear matrix inequalities. Journal of Process Control, Vol. 10, pg 363 – 368.
[13] Rohn J. (1989) A farkas-type theorem for linear interval equations. Computing 43, pg 93 – 95.
[14] Orukpe P. E. (2009) Model predictive control for linear time invariant systems using linear matrix inequality techniques. PhD Thesis – Imperial College London.
[15] Orukpe P. E. (2012) Model predictive control fundamentals. Accepted for publication in Nigerian Journal of Technology, Nsukka, Nigeria.
[16] Ye. Y. and Zhang S. (2003) "New results on quadratic minimization". SIAM Journal of Optimization, Vol. 14, No. 1, pg 245 – 267.
[17] Wan Z. and Kothare M. V. (2002) Robust output feedback model predictive control using off-line linear matrix inequalities. Journal of Process Control, Vol. 12, pg 763 – 774.