TOWARDS JOINT STATE ESTIMATION AND CONTROL IN MINIMAX MPC

Johan Löfberg

Department of Electrical Engineering
Linköpings universitet, SE-581 83 Linköping, Sweden
WWW: http://www.control.isy.liu.se
Email: johanl@isy.liu.se

Abstract: A new approach to minimax MPC for systems with bounded external system disturbances and measurement errors is introduced. It is shown that joint deterministic state estimation and minimax MPC can be written as an optimization problem with linear and quadratic matrix inequalities. By linearizing the quadratic matrix inequality, a semidefinite program is obtained. A simulation study indicates that solving the joint problem can improve performance.

Keywords: Predictive control, Robust control, Robust estimation, Minimax, Convex optimization

1. INTRODUCTION

In this paper, we introduce an approach to design MPC controllers in the case of estimated states and unknown but bounded disturbances acting on the system and the output measurements. The main contribution is an extension of the framework introduced in (Löfberg, 2001a). It is shown that joint state estimation and minimax MPC can be cast as an optimization problem involving a (unfortunately) quadratic matrix inequality. It is shown how this can be conservatively approximated as a linear matrix inequality (LMI) and thus enable us to approximately solve the joint problem using semidefinite programming.

Minimax MPC for systems with bounded disturbances has been studied before. The case with full state information is dealt with in, e.g., (Bemporad, 1998) and (Scokaert and Mayne, 1998). An approach for minimax MPC with both estimation error and disturbances is studied in (Bemporad and Garulli, 1997; Bemporad and Garulli, 2000). The problem with joint state estimation and control does however not seem to have been studied before.

2. UNCERTAINTY MODEL

The class of systems we address are linear time-invariant discrete-time systems with external system and measurement disturbances

\[ x(k + 1) = Ax(k) + Bu(k) + Fw(k) \]  
\[ y(k) = Cx(k) + E\eta(k) \]

The disturbances are assumed to be unknown but bounded

\[ \eta(k) \in \{ \eta : \eta^T \eta \leq 1 \} \]  
\[ w(k) \in \{ w : w^T w \leq 1 \} \]

Since we only measure a disturbed output, we have to use a state estimator. Regardless of how this is done, we can write

\[ x(k) = \hat{x}(k) + e(k) \]

The estimator used in this paper gives a state estimate with a guaranteed ellipsoidal error bound

\[ e^T(k)P(k)e(k) \leq 1 \]

The confidence matrix \( P(k) \) is an output from the state estimation procedure.
3. MINIMAX MPC

In the standard case, we typically use a quadratic finite horizon performance measure ($Q$ and $R$ for simplicity assumed positive definite)

$$J = \sum_{j=0}^{N-1} \|x(k+j+1)\|^2_Q + \|u(k+j)\|^2_R \tag{5}$$

There are typically constraints on inputs and outputs, but to keep the notation simple, we will not write this explicitly. We will however return to the constraints later on.

Since $x(k)$ is uncertain, this should be addressed in some way. The standard approach to robustify nominal MPC is to employ a minimax strategy, i.e. optimize worst-case behavior (Kothare et al., 1994; Bemporad and Garulli, 1997). In (Löfberg, 2001a), it was shown that, given an ellipsoidal estimation error bound

$$e^T(k)P(k)e(k) \leq 1 \tag{6}$$

and the previously introduced external system disturbances, a minimax strategy

$$\min_{u(k|k)} \max_{e(k)} J \tag{7}$$

can be turned into a problem that can be addressed with semidefinite programming. However, the estimation part was performed without any consideration on how the estimate would influence the control performance. The work here extends those results and the goal is to connect the estimation part with the minimax controller.

4. DETERMINISTIC STATE ESTIMATION

What is an optimal state estimate in a minimax framework? Clearly, the optimal choice is to find the smallest set $\mathcal{X}(k)$ such that

$$x(k) \in \mathcal{X}(k)$$

can be guaranteed, given all measurement obtained since startup and perhaps some prior knowledge on the initial state $x(0) \in \mathcal{X}(0)$. Hence the problem is

$$\min \text{Vol}(\mathcal{X}(k)) \text{ given } y(1), y(2), \ldots, y(k), \mathcal{X}(0)$$

The crux is that this is not practically implementable, not even for our simple model. The problem is that the complexity of the set $\mathcal{X}(k)$ grows when more measurements are obtained. The standard way to overcome this problem is to restrict $\mathcal{X}(k)$ to have some special geometry, such as ellipsoidal (Schweppe, 1968; Schweppe, 1973; Ghaoui and Calafiore, 1999) or parallelotopic (Bemporad and Garulli, 1997). Furthermore, a recursive scheme is employed. Unfortunately, assuming that $\mathcal{X}(k-1)$ has some particular geometry does not imply that $\mathcal{X}(k)$ also will have this. Hence, if we force $\mathcal{X}(k)$ to be an ellipsoid, we have to settle with an approximation. When we resort to an approximation, there will be some degree of freedom, and this is the fact we will exploit in order to improve the performance of the minimax MPC controller.

4.1 Ellipsoidal state estimates

In this work, we use an ellipsoidal approximations of the set $\mathcal{X}(k)$. Given a guaranteed ellipsoidal bound of the prior estimation error

$$e^T(k-1)P(k-1)e(k-1) \leq 1 \tag{8}$$

and a new measurement $y(k)$, use the model (1) and the disturbance bounds (2) to find a new state estimate guaranteed to satisfy

$$e^T(k)P(k)e(k) \leq 1 \tag{9}$$

It can be shown that an LMI in the following form is obtained as a sufficient condition

$$\begin{bmatrix} \Gamma & S^T \\ S & P^{-1}(k) \end{bmatrix} \succeq 0 \tag{10}$$

The definition of $\Gamma$ and $S$ are given in the appendix, but for a more detailed discussion on the estimation procedure, the reader is referred to (Ghaoui and Calafiore, 1999) or (Löfberg, 2001a). The important thing to know is that the matrix $\Gamma$ is a linear function of four scalar optimization variables, and the matrix $S$ depends linearly on the state estimate $\hat{x}(k)$.

Having this sufficient condition is a first step in a state estimation procedure. The next step is to select a particular solution $\hat{x}(k)$ and $P^{-1}(k)$. To do this, some performance measure on $P^{-1}(k)$ is minimized under the constraint (10). A typical choice (Ghaoui and Calafiore, 1999) is the trace, $\text{tr}(P^{-1}(k))$. We call this problem $\mathcal{P}_1$

$$\mathcal{P}_1 : \min_{\Gamma, P^{-1}(k), \hat{x}(k)} \text{tr}(P^{-1}) \text{ subject to } (10)$$

However, when this problem is solved, there is no connection to the control problem in which the state estimate will be used. The main result in this paper is to show that the estimation, i.e. calculation of $\hat{x}(k)$ and $P(k)$, can be done simultaneously with the calculation of the control, thus leading to some sort of joint estimation and control.
5. THE JOINT PROBLEM

We first derive the LMI for the minimax MPC problem. The calculations are done in a vectorized form so we introduce the predicted future states, unknown disturbances and the control sequence

\[
X = \begin{bmatrix}
x(k + 1 | k) \\
x(k + 2 | k) \\
\vdots \\
x(k + N | k)
\end{bmatrix}, \quad U = \begin{bmatrix}
u(k | k) \\
u(k + 1 | k) \\
\vdots \\
u(k + N - 1 | k)
\end{bmatrix}
\]

\[
W = \begin{bmatrix}
w(k | k) \\
w(k + 1 | k) \\
\vdots \\
w(k + N - 1 | k)
\end{bmatrix}
\]

By introducing the matrices \(H, S\) and \(G\)

\[H = \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^N \end{bmatrix}, \quad S = \begin{bmatrix} B & 0 & \ldots & 0 \\ AB & B & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1} B & \ldots & AB & B \end{bmatrix}\]

\[G = \begin{bmatrix} F & 0 & \ldots & 0 \\ AF & F & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1} F & \ldots & AF & F \end{bmatrix} = \begin{bmatrix} G_0 & G_1 & \cdots & G_{N-1} \end{bmatrix}\]

we can write

\[X = Hx(k) + SU + GW \] (11)

The minimax problem can, after redefining \(Q := \text{diag}(Q, \ldots, Q)\) and \(R := \text{diag}(R, \ldots, R)\), be written as

\[
\min_{t, U} t \quad \text{subject to } \max_{c(k), W} X^T Q X + U^T R U \leq t
\]

The state estimate uncertainty

\[(x(k) - \hat{x}(k))^T P(k)(x(k) - \hat{x}(k)) \leq 1 \] (12)

can be written in a form more suitable for us

\[x(k) = \hat{x}(k) + P^{-\frac{1}{2}} z, \quad ||z|| \leq 1 \] (13)

This allows us to write

\[X = H\hat{x}(k) + SU + HP^{-\frac{1}{2}} z + GW \] (14)

From now on we skip the time-index on \(P\) in order to save space. For reasons that will be clear later, we also define

\[v_0 = GW \] (15)

The nominal part of the state predictions are gathered in \(\hat{X}\)

\[\hat{X} = H\hat{x}(k) + SU \] (16)

We use definition (16) and (15), and rewrite the constraint in the minimax optimization problem using a Schur complement

\[
\begin{bmatrix} t \\ X + HP^{-\frac{1}{2}} z + v_0 U^T \end{bmatrix} \begin{bmatrix} Q^{-1} & 0 \\ 0 & R^{-1} \end{bmatrix} \begin{bmatrix} t \\ U \end{bmatrix} \geq 0
\]

Extract the estimation error

\[
\begin{bmatrix} 0 \\ X + v_0 U^T \end{bmatrix} + \begin{bmatrix} 0 \\ H P^{-\frac{1}{2}} \end{bmatrix} z + \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} z^T \begin{bmatrix} 0 & P^{-\frac{1}{2}} H^T \end{bmatrix} \geq 0
\] (17)

The above matrix inequality should hold for all admissible normalized estimation errors \(z\). To proceed, we use the following theorem (Ghaoui and Lebret, 1997)

Theorem 1. (Robust LMI). Robust satisfaction of the uncertain matrix inequality

\[F + L \Delta R + R^T \Delta^T L^T \geq 0 \quad \forall ||\Delta|| \leq 1\]

is equivalent to the matrix inequality

\[
\begin{bmatrix} F & L \\ L^T & 0 \end{bmatrix} \succeq \begin{bmatrix} R & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \tau I & 0 \\ 0 & -\tau I \end{bmatrix} \begin{bmatrix} R & 0 \\ 0 & I \end{bmatrix} \tau \geq 0
\]

After introducing the multiplier \(\tau \geq 0\) and applying Theorem 1 to the uncertain LMI (17) we obtain

\[
\begin{bmatrix} t \\ \hat{X} + v_0 U^T \end{bmatrix} \begin{bmatrix} Q^{-1} & 0 \\ 0 & R^{-1} \end{bmatrix} \begin{bmatrix} t \\ U \end{bmatrix} \geq \begin{bmatrix} I \\ 0 \\ 0 \\ 0 \\ \tau_\alpha I \end{bmatrix} \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & -\tau_\alpha I & 0 & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & 0 & I \end{bmatrix}
\] (18)

Simplification yields

\[
\begin{bmatrix} t - \tau_\alpha \\ \hat{X} + v_0 U^T \end{bmatrix} \begin{bmatrix} Q^{-1} & 0 \\ 0 & R^{-1} \end{bmatrix} \begin{bmatrix} t \\ U \end{bmatrix} \geq \begin{bmatrix} I \\ 0 \end{bmatrix} \begin{bmatrix} \tau_\alpha I & 0 & 0 & 0 \\ 0 & -\tau_\alpha I & 0 & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & 0 & I \end{bmatrix}
\] (19)

The condition still contains uncertain parts, more precisely the vector \(v_0\). To take care of these, we first define

\[v_i = \sum_{j=1}^{N-1} G_j w(k + j | k) \] (20)
and note that
\[ \dot{\theta}_0 = G_0 w(k|k) + \dot{\theta}_1 \]  
(21)
The uncertainty \( w(k|k) \) is now removed using Theorem 1, and the procedure is repeated. This will eventually give us the following LMI
\[
\begin{bmatrix}
-\tau_x - \text{tr}(\Omega) & \dot{X}^T & U^T & 0 & 0 \\
\dot{X} & Q^{-1} & 0 & H P^{-1} G & 0 \\
0 & 0 & R^{-1} & 0 & 0 \\
0 & Z^T H^T & 0 & \tau_x I & 0 \\
0 & G^T & 0 & 0 & \Omega
\end{bmatrix} \succeq 0 \tag{22}
\]
The matrix \( \Omega \) is a diagonal matrix containing the variables introduced when applying Theorem 1 on the future unknown disturbances.

Given a state estimate \( \hat{x}(k) \) and \( P(k) \), this is the LMI derived in (Lofterg, 2001a) for minimax MPC. We denote this problem \( \mathcal{P}_2 \).

\[
\mathcal{P}_2 : \min_{\tau_x, \Omega, U, t} \quad t \\
\text{subject to} \quad (22)
\]
We are now ready to proceed to the main idea in this paper. Recall the state estimation LMI and introduce \( Z = P^{-\frac{1}{2}} \hat{x}(k) \). The constraints for estimation and minimax MPC can be summarized as
\[
\begin{bmatrix}
-\tau_x - \text{tr}(\Omega) & \dot{X}^T & U^T & 0 & 0 \\
\dot{X} & Q^{-1} & 0 & H Z G & 0 \\
0 & 0 & R^{-1} & 0 & 0 \\
0 & Z^T H^T & 0 & \tau_x I & 0 \\
0 & G^T & 0 & 0 & \Omega
\end{bmatrix} \succeq 0 \tag{23}
\]
Since \( S \) is linear in \( \hat{x}(k) \), the equations are linear in \( \Gamma, \Omega, \tau_x, t, U \) and \( \hat{x}(k) \). Unfortunately it is quadratic in \( Z \). However, for future reference we define the problem as \( \mathcal{P}_3 \)
\[
\mathcal{P}_3 : \min_{\tau_x, \Omega, U, t, \Gamma, Z, \hat{x}(k)} \quad t \\
\text{subject to} \quad (23, 24)
\]
5.1 A tractable approximation
To obtain a tractable problem, we simply linearize the quadratic matrix inequality. From the trivial inequality
\[(Z - Z_0)^T(Z - Z_0) \geq 0 \tag{25}\]
we have
\[ Z^T Z \succeq Z^T Z_0 + Z_0^T Z - Z_0^T Z_0 \tag{26}\]
We use this and obtain an LMI that conservatively approximates the original quadratic matrix inequality
\[
\begin{bmatrix}
\Gamma & S^T \\
S & Z^T Z_0 + Z_0^T Z - Z_0^T Z_0
\end{bmatrix} \succeq 0 \tag{27}
\]
Clearly, the main problem now is to select the linearization point \( Z_0 \). The perhaps easiest solution is to solve the problem \( \mathcal{P}_1 \), and then use the solution to define \( Z_0 \). Of course, this can be repeated in order to find a local minimum of \( \mathcal{P}_3 \). We define the linearized and conservative approximation of \( \mathcal{P}_3 \)
\[
\mathcal{P}_4 : \min_{\tau_x, Z, \hat{x}(k), U, t} \quad t \\
\text{subject to} \quad (23, 27)
\]
5.2 State constraints
Typically there are state constraints in the MPC problem. Let us study the simple scalar case \( MX \leq 1 \). In other words, \( M \) is a row vector. Inserting the definition of \( X \) and the state estimate error yields the constraint
\[ MH \hat{x}(k) + MSU + MG W + MHP^{-\frac{1}{2}} z \leq 1 \]
It is easy to show (Lofterg, 2001a) that the constraint is satisfied for all possible estimation errors and future disturbances if
\[
MH \hat{x}(k) + MSU + \sqrt{MP^{-1}HT MT} + \sum_{j=0}^{N-1} \sqrt{MG_j G_j^T} M T \leq 1 \tag{28}
\]
To save space, we define
\[ \gamma = 1 - MH \hat{x}(k) - MSU - \sum_{j=0}^{N-1} \sqrt{MG_j G_j^T} M T \]
and the constraint can be written as
\[ \sqrt{MP^{-1}HT MT} \leq \gamma \tag{29} \]
We square the constraint and recall that \( P^{-1} = Z^T Z \). This allows us to perform a Schur complement and obtain an LMI
\[
\begin{bmatrix}
\gamma & M H Z \\
Z^T H^T M T & \gamma I
\end{bmatrix} \succeq 0 \tag{30}
\]
6. SIMULATION EXAMPLE
This example is adapted from (Bemporad and Garulli, 2000). Since the main result in this paper is the introduction of a joint estimation and control scheme,

---

1 Not entirely true. To obtain the LMI in (Lofterg, 2001a) some additional Schur complements are needed.
we want to study the impact of the estimation error. For that reason, the only uncertainty in the system is a measurement error, leading to an uncertain state estimate.

\[
x(k+1) = \begin{bmatrix} 1.64 & -0.79 \\ 1 & 0 \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(k)
\]

\[
y(k) = \begin{bmatrix} 0.14 & 0 \end{bmatrix} x(k) + 0.2\eta(k)
\]

There is a non-minimum phase output

\[
z(k) = \begin{bmatrix} -1.93 & 2.21 \end{bmatrix} x(k)
\]

with a hard constraint

\[-1 \leq z(k) \leq 3\]

In addition to the output constraint, the controller also has to satisfy \(|u(k)| \leq 2|\).

The goal is to have the (undisturbed) output \(y(k)\) follow a constant unit reference. In order to get good tracking, the following performance measure was chosen (\(N = 10\))

\[
\sum_{j=0}^{N-1} ||C x(k+j+1|k) - 1||^2 + 0.1||u(k+j|k) - 1||^2
\]

Since we have shifted the origin in the tracking formulation, some straightforward modifications of the algorithm are needed. For brevity, the details are omitted.

Three different controllers were implemented. In the first approach, the state estimation is performed by solving \(P_1\) and the estimate is then used in the minimax controller defined by \(P_2\). This is basically the controller proposed in (Löfberg, 2001a). We denote this controller \(C_1\). In a second controller \(C_2\), an initial state estimate is found by solving \(P_1\), and the matrix is then used to linearize the joint problem \(P_3\), yielding \(P_4\), which then is solved. In a third approach \(C_3\), the linearization procedure is repeated two times.

The three controllers were simulated 100 times with different initial conditions and disturbance realizations. The initial state estimate was \(\hat{x}(0) = 0\) and \(P(0) = I\), while the true initial state was uniformly distributed in the ellipsoid \(||x(0)|| \leq 1\). The measurement disturbances were uniformly distributed. Implementation and solution of the optimization problems were done using (Löfberg, 2001b) and (Vandenberge and Boyd, 1998).

The mean of the accumulated quadratic performance measure,

\[
\sum_{j=0}^{\infty} ||C x(j) - 1||^2 + 0.1||u(j) - 1||^2
\]

was calculated and was \(J_{C_1} = 17.4\), \(J_{C_2} = 10.4\) and \(J_{C_3} = 8.8\). The average improvement when looking at single realizations and comparing the controllers \(C_1\) and \(C_2\) was 21%, while \(C_3\) gave an additional 8% average improvement. Furthermore, the controller \(C_1\) became infeasible in 12 cases, while this never happened for \(C_2\) or \(C_3\). In Figure 1, we see a situation where the proposed approach has improved tracking performance substantially.

Fig. 1. Closed loop response.

The reason why the proposed approach gave such a substantial improvement in this example is the state constraint. The constrained output has a severe non-minimum phase behavior. If the uncertainty in the state estimate is too large, the uncertainty in the constrained output will force the controller to be very careful. Since the limiting factor is the constraint, it is important that the measurements are used in order to obtain an estimate that is certain along the constrained output directions. This will be done automatically in the joint approach, hence leading to improved performance.

7. CONCLUSION

We have shown that incorporation of the state estimation problem into minimax MPC yields a problem with a quadratic matrix inequality. By linearizing this inequality, a linear matrix inequality is obtained, and the joint estimation and control problem could be solved using semidefinite programming. A simple simulation study was carried out and showed that the approach indeed can improve performance in some cases.

Of course, the improved performance comes at a price, computational complexity. Various improvements can be done to reduce this. Currently, the initial guess on \(P(k)\) is found by solving problem \(P_1\). A cheaper way to find an initial guess could be to use approximative solutions based on ellipsoidal calculus (Schweppe, 1968; Schweppe, 1973; Kurzhanski and Vályi, 1997).
8. REFERENCES


Appendix A. DEFINITION OF $\Gamma$ AND $S$

The matrices $\Gamma$ and $S$, involved in the state estimation LMI (10), are derived as follows (Ghaoui and Calafiore, 1999; Löfberg, 2001a); Define

$$S = \begin{bmatrix} A & F & 0 & Bu(k-1) - \dot{x}(k) \end{bmatrix}$$

$$T_e = \begin{bmatrix} I & 0 & 0 & -\dot{x}(k-1) \end{bmatrix}$$

$$T_y = \begin{bmatrix} -CA & -CF & -E & y(k) - CBu(k-1) \end{bmatrix}$$

$$T_w = \begin{bmatrix} 0 & I & 0 \end{bmatrix}$$

$$T_\eta = \begin{bmatrix} 0 & 0 & I \end{bmatrix}$$

$$T_1 = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\Lambda = T_1^T T_1$$

and

$$\xi = \begin{bmatrix} x(k-1), w(k-1), \eta(k-1), 1 \end{bmatrix}$$

The constraint (9) on the new estimate can be written as

$$\xi^T S^T P(k) S \xi \leq \xi^T \Lambda \xi$$

when

$$\xi^T T_e^T P(k-1) T_e \xi \leq \xi^T \Lambda \xi$$

$$\xi^T T_w^T T_w \xi \leq \xi^T \Lambda \xi$$

$$\xi^T T_\eta^T T_\eta \xi \leq \xi^T \Lambda \xi$$

$$\xi^T T_y^T T_y \xi = 0$$

The implication is conservatively approximated using the S-procedure (Boyd et al., 1994), yielding the LMI (10). Without going into details, three non-negative scalars $\tau_e$, $\tau_w$, and $\tau_\eta$, and one indefinite scalar $\tau_y$ are introduced, and $\Gamma$ will be defined as

$$\Gamma = \Lambda - \tau_e S_e - \tau_w S_w - \tau_\eta S_\eta - \tau_y S_y$$

where

$$S_e = \Lambda - T_e^T P(k-1) T_e$$

$$S_w = \Lambda - T_w^T T_w$$

$$S_\eta = \Lambda - T_\eta^T T_\eta$$

$$S_y = T_y^T T_y$$