A game theoretic approach to moving horizon control

Sanjay Lall and Keith Glover

Abstract

A control law is constructed for a linear time varying system by solving a two player zero sum differential game on a moving horizon, the game being that which is used to construct an $\mathcal{H}_\infty$ controller on a finite horizon. Conditions are given under which this controller results in a stable system and satisfies an infinite horizon $\mathcal{H}_\infty$ norm bound. A risk sensitive formulation is used to provide a state estimator in the observation feedback case.

1 Introduction

The moving horizon control technique was developed in the 1970's, with one of the first papers being that by Kleinman (Kleinman 1970) in 1970. The techniques in his paper were later reformulated and generalized by various authors, see for example (Chen & Shaw 1982, Kwon & Pearson 1977, Kwon, Bruckstein & Kailath 1983). Kwon and Pearson (Kwon & Pearson 1977) proved stability for linear time varying systems using a controller which optimised a quadratic cost function integrated over a time interval from the current time t to a fixed distance ahead $t+T$. The solution to this problem was given in terms of a Riccati differential equation, integrated backwards from time $t+T$ at each time t.

For time varying systems, this gives a practical method of stabilising the system. Kwon et. al. (Kwon et al. 1983) formulated a general procedure for the recursive update of such Riccati equations, which generalise easily to the indefinite Riccati equations in this paper.

Further, this method allows stabilization of systems which are known only for the short term future. In dealing with long term variation of systems, there are at present two extreme options. The first is to assume a time invariant system and design a controller robust against time varying perturbations, which usually leads to very conservative controller design. The other is to design fully for time varying systems. In the case of both linear quadratic and $\mathcal{H}_\infty$ controllers, this requires the backwards integration from infinity of a Riccati equation (see for example (Tadmor 1993, Ravi, Nagpal & Khargonekar 1991)), and somewhat optimistically assumes knowledge of the system throughout future time. The moving horizon method can be viewed as a compromise between these two methods.

The aim of this present paper is to extend the work of Tadmor (Tadmor 1992), in which a receding horizon controller is formulated with each finite horizon optimisation based upon an

$\mathcal{H}_\infty$ optimisation. It is hoped that the practical advantages of receding horizon control might be combined with the robustness advantages of $\mathcal{H}_\infty$ control. Tadmor proved that that this controller was stable and satisfied an infinite horizon norm bound. Although derived using different approaches to the $\mathcal{H}_\infty$ problem, the controller described in this paper is essentially the same in the state feedback case, differing only in the terminal constraint used.

In the observation feedback case, however, it is difficult to produce a natural formulation of the $\mathcal{H}_\infty$ receding horizon control problem. For the finite horizon problem, typically the assumptions made are that the initial state at the beginning of each optimization interval is completely known, or is completely unknown. In the latter case it is treated as part of the disturbance and subject to a quadratic weighting (Uchida & Fujita 1992) (Khargonekar, Nagpal & Poolla 1991). In the receding horizon problem, though, we have observations from *before* the optimization interval. We attempt to make use of prior observations using the theory of risk sensitive control, as developed by Whittle (Whittle 1990), which has been shown to be equivalent to $\mathcal{H}_\infty$ control in many situations.

2 Preliminaries

We consider the system

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B_1(t)w(t) + B_2(t)u(t) \\ z(t) &= C_1(t)x(t) + D_{12}(t)u(t) \\ y(t) &= C_2(t)x(t) + D_{21}(t)w(t) \end{aligned} \tag{1}$$

where the coefficient matrices are bounded matrix valued functions of time. We use the following assumptions

$$D'_{12}C_1 = 0 \quad D'_{12}D_{12} = I \quad D_{21}B'_1 = 0 \quad D_{21}D'_{21} = I \tag{2}$$

which can be removed by suitable changes in variables, see for example (Doyle, Glover, Khargonekar & Francis 1989, Ravi et al. 1991, Tadmor 1993). Further, we shall assume there exists $\varepsilon > 0$ such that

$$\begin{aligned} B_1(t)B_1(t)' &\geq \varepsilon I \text{ for all } t > 0 \\ C_1(t)'C_1(t) &\geq \varepsilon I \text{ for all } t > 0 \end{aligned} \tag{3}$$

These assumptions can be removed by assuming uniform complete controllability and observability of the system, and slightly modifying the following arguments.

Let $\mathcal{L}_2^m[t_i, t_f]$ be the Hilbert space of square integrable functions on $[t_i, t_f] \subset \mathcal{R}$ taking values in $\mathcal{R}^m$, and write $\mathcal{R}^+ = [0, \infty)$. Denote the usual norm on $\mathcal{L}_2$ by $\|\cdot\|_2$, and if $F\colon \mathcal{L}_2 \to \mathcal{L}_2$ then denote by $\|\cdot\|_\infty$ the norm on F induced by $\|\cdot\|_2$ Then $x\colon \mathcal{R}^+ \to \mathcal{R}^n$, $u \in \mathcal{L}_2^{m_1}[t_i, t_f]$,$w \in \mathcal{L}_2^{m_2}[t_i, t_f]$, $y\colon \mathcal{R}^+ \to \mathcal{R}^p$. The signal w represents all external inputs including disturbances, sensor noise, and commands, z represents the error signal, y is the measured variable and u is the control input. We shall omit the space dimensions of signals in the sequel.

3 State feedback

3.1 Finite horizon $\mathcal{H}_\infty$ control

Define the cost function

$$J_\gamma(u,w) = \int_{t_i}^{t_f} \left\{ z(t)'z(t) - \gamma^2 w(t)'w(t) \right\} dt$$

We can regard J_γ as a function of either $\mathcal{L}_2$ signals or *feedback strategies*. Let $\mathcal{M} = \{\, \mu\colon [t_i,t_f] \times \mathcal{R}^n \to \mathcal{R}^{m1} \,\}$ and $\mathcal{N} = \{\, \nu\colon [t_i,t_f] \times \mathcal{R}^n \to \mathcal{R}^{m2} \,\}$. These spaces are the *strategy spaces*, and we shall write strategies as μ, ν to distinguish them from signals u, w. With the control $u(t) = \mu(t, x(t))$ in place, the operator $\mathcal{T}_{zw}$ maps w to z. If $x(t_i) = 0$, then we can define

$$\|\mathcal{T}_{zw}\|_\infty = \sup_{w \in \mathcal{L}_2[t_i,t_f]} \|\mathcal{T}_{zw} w\|_2 / \|w\|_2$$

The finite interval $\mathcal{H}_\infty$ problem is to find a linear causal control μ such that $\|\mathcal{T}_{zw}\|_\infty < \gamma$ for a given $\gamma > 0$. $\|\mathcal{T}_{zw}\|_\infty < \gamma$ if and only if there exists $\varepsilon > 0$ such that

$$J_\gamma(u,w) \le -\varepsilon \|w\|_2^2 \quad \forall w \in \mathcal{L}_2[t_i,t_f] \tag{4}$$

We formulate the finite time differential game

$$\inf_{\mu \in \mathcal{M}} \sup_{\nu \in \mathcal{N}} J_\gamma(\mu, \nu) \tag{5}$$

which is a *zero sum* game, where u is the minimizing player and w is the maximizing player. The designer chooses $u(t) = \mu(t, x(t))$ such that even if nature is malicious and chooses the worst case w, then equation (4) is satisfied.

If the extremizing operators in equation (5) are interchangeable, then we call the optimal u and worst case w *saddle point strategies*. Theorem 1 gives sufficient conditions for the existence of a saddle point solution. A saddle point solution $u(t) = \mu^*(t, x(t))$, $w(t) = \nu^*(t, x(t))$ will satisfy

$$J_\gamma(\mu^*, w) \le J_\gamma(\mu^*, \nu^*) \le J_\gamma(u, \nu^*) \quad \forall u, w \in \mathcal{L}_2[t_i, t_f].$$

Theorem 1 (Basar & Olsder 1982, Limebeer, Anderson, Khargonekar & Green 1992)
Let

$$J_\gamma(u,w) = \int_{t_i}^{t_f} \left\{ z(t)'z(t) - \gamma^2 w(t)'w(t) \right\} dt + x(t_f)'Fx(t_f)$$

where $F > 0$ is some weighting matrix. If

$$-\dot{P} = PA + A'P + C_1'C_1 - P(B_2B_2' - \gamma^{-2}B_1B_1')P, \quad P(t_f) = F \tag{6}$$

has a unique symmetric bounded solution on $t \in [t_i, t_f]$, then

$$\begin{aligned} J_\gamma(u,w) = x(t_i)'P(t_i)x(t_i) \quad &+ \quad \int_{t_i}^{t_f} |u(\tau) + B_2(\tau)'P(\tau)x(\tau)|^2 d\tau \\ &- \quad \int_{t_i}^{t_f} \gamma^2 |w(\tau) - \gamma^{-2}B_1(\tau)'P(\tau)x(\tau)|^2 d\tau \end{aligned}$$

Further, with full state information available to $u(t) = \mu^*(t, x(t))$ and $w(t) = \nu^*(t, x(t))$, there exists a unique feedback saddle point solution given by

$$\begin{aligned} \mu^*(t, x(t)) &= -B_2(t)'P(t)x(t) \\ \nu^*(t, x(t)) &= \gamma^{-2}B_1(t)'P(t)x(t) \end{aligned}$$

and the saddle point value of the game is given by

$$J_\gamma(\mu^*, \nu^*) = x(t_i)'P(t_i)x(t_i).$$

□

3.2 The moving horizon differential game

We now try to find $u(t) = \mu^*(t, x(t))$ at time t, where μ^* is the saddle point strategy for player u for

$$\inf_{\mu \in \mathcal{M}_t} \sup_{\nu \in \mathcal{N}_t} \left\{ \int_t^{t+T} \left\{ z(s)'z(s) - \gamma^2 w(s)'w(s) \right\} ds + x(t+T)'F(t+T)x(t+T) \right\}$$

where $\mathcal{M}_t = \{\, \mu \colon [t, t+T] \times \mathcal{R}^n \to \mathcal{R}^{m1} \,\}$, $\mathcal{N}_t = \{\, \nu \colon [t, t+T] \times \mathcal{R}^n \to \mathcal{R}^{m2} \,\}$, and $w(t) = \nu^*(t, x(t))$.

The idea behind this is to try to gain both the robustness benefits of the $\mathcal{H}_\infty$ formulation and the practical control of time varying systems associated with receding horizon control. At each time t we try to solve the finite horizon $\mathcal{H}_\infty$ problem on the optimisation interval $[t, t+T]$ to find a feedback strategy $u(t) = \mu(t, x(t))$.

In practice, this would be implemented for a finite time $\delta > 0$, after which a new calculation on the interval $[t+\delta, t+\delta+T]$ would be performed. In this paper we idealise this situation and assume that the controller is updated continuously.

For each separate $\mathcal{H}_\infty$ optimisation, the initial state may not be zero. Therfore, an induced norm interpretation on each finite horizon is not possible, since zero input u gives a nonzero output z on the interval $[t_i, t_f]$, and thus the induced norm is undefined.

We also have a penalty on the terminal state, $F(t) > 0$. This is often incorporated into finite horizon problems to allow for compromises between the norm of z and the size of $x(t_f)$. In the moving horizon case, we shall show that a sufficiently large F will cause the closed loop to be stable.

In order to find $u(t)$ for the moving horizon problem we must integrate equation (6) backwards from the boundary condition at $t+T$ to time t. We can rewrite this as a partial differential equation as follows

$$\begin{aligned} -\frac{\partial P(\tau,\sigma)}{\partial \tau} &= P(\tau,\sigma)A(\tau) + A(\tau)'P(\tau,\sigma) - \\ &\quad P(\tau,\sigma)(B_2(\tau)B_2(\tau)' - \gamma^{-2}B_1(\tau)B_1(\tau)')P(\tau,\sigma) + C_1(\tau)'C_1(\tau) \\ P(t,t) &= F(t) \text{ for all } t \end{aligned}$$

The moving horizon saddle point controller is given by

$$u(t) = -B_2(t)'P(t, t+T)x(t)$$

In the case when all system matrices A, B_1, B_2, C_1, D_{12} are time invariant, then $P(t, t+T)$ is independent of t also.

3.3 Stability

Tadmor (Tadmor 1992) proves stability for the above controller when $F = \infty$. In this case, we can consider the Riccati equation for P^{-1} and the interpretation on the finite horizon is that the controller is subject to the constraint $x(t+T) = 0$.

The linear quadratic problem with a similar terminal weight was considered by Kwon, Bruckstein and Kailath (Kwon et al. 1983). We use a modified form of their methods to handle the terminal weight.

Before the stability proof, we need the following lemmas, in which we will write $P(\tau, \sigma, F)$ for the solution of the Riccati equation with boundary condition $P(\tau, \sigma, F) = F$ for $\tau = \sigma$, with σ fixed. The following results are derived in a slightly different way in (Kwon et al. 1983), and various similar results can be found in (Anderson & Moore 1989, Reid 1970, Bucy 1967, Ran & Vreugdenhil 1988, Basar & Bernhard 1991).

Lemma 2 Let P satisfy the following Riccati equation on $[t_i, t_f]$, with $F \geq 0$ and $Q \geq 0$.

$$-\dot{P} = PA + A'P - P(B_2B_2' - \gamma^{-2}B_1B_1')P + Q \quad P(t_f) = F$$

then if P exists, and either $F > 0$ or $Q > 0$, then $P(t) > 0 \ \forall t \in [t_i, t_f]$.

Lemma 3 Suppose

$$\dot{F} + A'F + FA - F(B_2B_2' - \gamma^{-2}B_1B_1')F + C_1'C_1 \leq 0, \quad F(t) > 0, t > 0 \tag{7}$$

and P satisfies

$$-\dot{P} = A'P + PA - P(B_2B_2' - \gamma^{-2}B_1B_1')P + C_1'C_1$$

then, if both sides exist,

$$P(\tau, \sigma_1, F(\sigma_1)) \geq P(\tau, \sigma_2, F(\sigma_2)), \quad \tau \leq \sigma_1 \leq \sigma_2$$

Proof We know (see (Basar & Bernhard 1991))

$$x(t_i)'P(t_i, t_f, F(t_f))x(t_i) = \tag{8}$$
$$\min_u \max_w \left\{ \int_{t_i}^{t_f} (x(s)'C_1(t)'C_1(t)x(s) + u(s)'u(s) - \gamma^2 w(s)'w(s))\, ds + x(t_f)'F(t_f)x(t_f) \right\}$$

then

(i) If $\tilde{C}_1(t)'\tilde{C}_1(t) \geq C_1(t)'C_1(t) \ \forall t \in [t_i, t_f]$, then replacing C_1 by $\tilde{C}_1$ results in a new solution, $\tilde{P}$ satisfying $\tilde{P}(t) \geq P(t) \ \forall t \in [t_i, t_f]$.

(ii) Similarly, if $\tilde{F}(t) \geq F(t) \ \forall t \in [t_i, t_f]$, then $\tilde{P}(t) \geq P(t) \ \forall t \in [t_i, t_f]$.

With either of these changes, equation(8) becomes

$$x(t_i)'\tilde{P}(t_i,t_f,F(t_f))x(t_i) = \\ \min_u \max_w \left\{ \int_{t_i}^{t_f} (x(s)'C_1(s)'C_1(s)x(s) + u(s)'u(s) - \gamma^2 w(s)'w(s))\, ds + x(t_f)'F(t_f)x(t_f) \\ + h(u,w) \right\}$$

and $h(u,w) > 0 \; \forall u, w \in \mathcal{L}_2$.

We can write equation(7) as

$$-\dot{F} = A'F + FA - F(B_2B_2' - B_1B_1')F + (C_1'C_1 + \Delta),$$

where $\Delta \geq 0$. By definition,

$$P(\sigma_2, \sigma_2, F(\sigma_2)) = F(\sigma_2)$$

so, since the Riccati equation for P and F are equal at $t = \sigma_2$, and that for F has a bigger constant term,

$$F(t) \geq P(t, \sigma_2, F(\sigma_2)) \text{ for all } t \leq \sigma_2$$

We can rewrite this as

$$P(\sigma_1, \sigma_2, F(\sigma_2)) \leq P(\sigma_1, \sigma_1, F(\sigma_1))$$

so, when $t = \sigma_1$, the LHS is smaller than the RHS. We can take this point as a boundary condition, and, using (ii) above,

$$P(t, \sigma_2, F(\sigma_2)) \leq P(t, \sigma_1, F(\sigma_1)) \text{ for all } t \leq \sigma_1 \leq \sigma_2$$

□

Theorem 4 For the system described by equations (1-3), if at each time t the saddle point solution to the differential game on $[t, t+T]$ exists, then if

$$\dot{F} + A'F + FA - F(B_2B_2' - \gamma^{-2}B_1B_1')F + C_1'C_1 \leq 0, \quad F(t) > 0, t > 0$$

and if there exists constants $\alpha_1 > 0$ and $\alpha_2 > 0$ such that

$$\alpha_1 I \leq P(t, t+T, F(t+T)) \leq \alpha_2 I$$

then the control $u(t) = -B_2(t)'P(t, t+T, F(t+T))x(t)$ is stabilising.

Proof Let $\Lambda(t) = A(t) - B_2(t)B_2(t)'P(t, t+T, F)$. Then, with the controller in place, $\dot{x} = \Lambda(t)x(t) + B_1(t)w(t)$. Let $V(t) = x(t)'P(t, t+T, F(t+T))x(t)$. Then

$$\dot{V} = x' \left(-P'B_2B_2'P - \gamma^{-2}PB_1B_1'P - C_1'C_1 + \frac{\partial}{\partial\sigma}P(t, \sigma, F(\sigma))\bigg|_{\sigma=t+T} \right) x$$

Using Lemma 3,

$$\frac{\partial}{\partial\sigma}P(t, \sigma, F(\sigma))\bigg|_{\sigma=t+T} \leq 0$$

Since $C_1'C_1 > \varepsilon I$, $\dot{V}(t) < -\varepsilon x(t)'x(t)$. Then, since by assumption P is bounded, it is a Lyapunov function for the closed loop system, hence the system is globally exponentially stable. □

If we can prove boundedness of $P(t, t+T, F(t+T))$, then we shall have a sufficient condition for the controller to be stable. Tadmor (Tadmor 1992) proves that $P(t, t+T, \infty)$ is bounded above and below. Slight modification of his proof shows:

Theorem 5 *Let $G_i(t_0, t_1) = \int_{t_0}^{t_1} \Phi(t_0, t)B_i(t)B_i(t)'\Phi(t_0, t)'\,dt$. If (A, B_2) is uniformly controllable and $G_2(t_0, t_1)^{-1}G_1(t_0, t_1)$ is uniformly bounded for all $t_1 - t_0 \leq T$, then there exists $\gamma > 0$ such that $P(t, t+T, F(t+T))$ is bounded above for all $t > 0$.*

3.4 Infinite horizon norm bounds

Following (Tadmor 1992), with the controller $u(t) = -B_2(t)'P(t, t+T, F(t+T))x(t)$ and $w \in \mathcal{L}_2[t_0, \infty]$ we know

$$\int_{t_0}^{\infty} \frac{d}{dt}\left\{x(t)'P(t, t+T, F(t+T))x(t)\right\}\,dt =$$
$$\int_{t_0}^{\infty}\left\{-z'z + 2w'B_1'Px - \gamma^{-2}x'PB_1B_1'Px + x'\left(\frac{\partial}{\partial\sigma}P(t, \sigma, F(\sigma))\bigg|_{\sigma=t+T}\right)x\right\}\,dt$$

If the closed loop is stable, then

$$\int_{t_0}^{\infty} \frac{d}{dt}\left\{x(t)'P(t, t+T, F(t+T))x(t)\right\}\,dt = -x(t_0)'P(t_0, t_0+T, F(t_0+T))x(t_0)$$

Let $\hat{w}(t) = w(t) - \gamma^{-2}B_2'(t)P(t, t+T, F(t+T))x(t)$. Then, if $x(t_0) = 0$,

$$\|z\|_2^2 - \gamma^2\|w\|_2^2 = -\gamma^2\|\hat{w}\|_2^2 + \int_{t_0}^{\infty} x'\left(\frac{\partial}{\partial\sigma}P(t, \sigma, F(\sigma))\bigg|_{\sigma=t+T}\right)x\,dt$$

Using Lemma 3,

$$\|z\|_2^2 - \gamma^2\|w\|_2^2 \leq 0$$

Hence $\|\mathcal{T}_{zw}\|_\infty \leq \gamma$.

4 Calculation of $P(\tau, \sigma, F(\sigma))$

In order to implement the moving horizon controller, we need to calculate the value of $P(t, t+T, F(t+T))$ for all $t > 0$. This requires the solution of a Riccati differential equation over the interval $[t, t+T]$ for each time t, given boundary conditions $P(t, t+T, F(t+T)) = F(t+T)$ for all $t > 0$. Kwon, Bruckstein and Kailath (Kwon et al. 1983) applied results from scattering theory in their solution to the quadratic problem, to give a forwards differential equation for $P(t, t+T, F(t+T))$. Here we state the same solution for the indefinite Riccati equation. See (Verghese, Friedlander & Kailath 1980) for details on the derivation of these formulae. In this section, $P(\tau, \sigma)$ is used to mean $P(\tau, \sigma, 0)$. Define for convenience

$$N(\tau) = B_2(\tau)B_2(\tau)' - \gamma^{-2}B_1(\tau)B_1(\tau)'$$

Let

$$S(\tau,\sigma) = \begin{pmatrix} \Phi(\tau,\sigma) & L(\tau,\sigma) \\ P(\tau,\sigma) & \Psi(\tau,\sigma) \end{pmatrix}$$

Consider the system of equations

$$\begin{aligned}
\frac{\partial}{\partial\tau}\Phi(\tau,\sigma) &= \Phi(\tau,\sigma)[N(\tau)P(\tau,\sigma) - A(\tau)] && (9)\\
\frac{\partial}{\partial\tau}\Psi(\tau,\sigma) &= [P(\tau,\sigma)N(\tau) - A'(\tau)]\Psi(\tau,\sigma) && (10)\\
\frac{\partial}{\partial\tau}L(\tau,\sigma) &= \Phi(\tau,\sigma)N(\tau)\Psi(\tau,\sigma) && (11)\\
\frac{\partial}{\partial\tau}P(\tau,\sigma) &= A'(\tau)P(\tau,\sigma) + P(\tau,\sigma)A(\tau) - P(\tau,\sigma)N(\tau)P(\tau,\sigma) + C_1(\tau)'C_1(\tau) && (12)
\end{aligned}$$

with boundary conditions

$$S(\sigma,\sigma) = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

The following equations

$$\begin{aligned}
\frac{\partial}{\partial\sigma}\Phi(\tau,\sigma) &= [A(\sigma) + L(\tau,\sigma)C_1(\sigma)'C_1(\sigma)]\Phi(\tau,\sigma) && (13)\\
\frac{\partial}{\partial\sigma}\Psi(\tau,\sigma) &= \Psi(\tau,\sigma)[A'(\sigma) + C_1(\sigma)'C_1(\sigma)L(\tau,\sigma)] && (14)\\
\frac{\partial}{\partial\sigma}L(\tau,\sigma) &= A(\sigma)L(\tau,\sigma) + L(\tau,\sigma)A'(\sigma) + L(\tau,\sigma)C_1(\sigma)'C_1(\sigma)L(\tau,\sigma) - N(\sigma) && (15)\\
\frac{\partial}{\partial\sigma}P(\tau,\sigma) &= \Psi(\tau,\sigma)C_1(\sigma)'C_1(\sigma)\Phi(\tau,\sigma) && (16)
\end{aligned}$$

give the partial derivatives with respect to σ.

It is straightforward to verify this by taking both sets of second partial derivatives, and showing that they are equal, and verifying that along the boundary the total derivative

$$\frac{d}{dt}S(t,t) = \left.\frac{\partial S(\tau,t)}{\partial\tau}\right|_{\tau=t} + \left.\frac{\partial S(t,\sigma)}{\partial\sigma}\right|_{\sigma=t}$$

is zero, since Ψ, Φ, L, P are constant along the boundary. We can find a differential equation for $P(t,t+T)$ in terms of t since

$$\frac{d}{dt}S(t,t+T) = \left.\frac{\partial S(\tau,t+T)}{\partial\tau}\right|_{\tau=t} + \left.\frac{\partial S(t,\sigma)}{\partial\sigma}\right|_{\sigma=t+T}$$

Hence

$$\begin{aligned}
\frac{d\Phi}{dt} &= \Phi[N(t)P - A(t)] + [A(t+T) + LC_1(t+T)'C_1(t+T)]\Phi && (17)\\
\frac{d\Psi}{dt} &= [PN(t) - A'(t)]\Psi + \Psi[A'(t+T) + C_1(t+T)'C_1(t+T)L] && (18)\\
\frac{dL}{dt} &= \Phi N(t)\Psi + A(t+T)L + LA'(t+T) + LC_1(t+T)'C_1(t+T)L - N(t+T) && (19)\\
\frac{dP}{dt} &= A'(t)P + PA(t) - PN(t)P + C_1(t)'C_1(t) + \Psi C_1(t+T)'C_1(t+T)\Phi && (20)
\end{aligned}$$

where $\Phi = \Phi(t, t+T)$ etc. This gives a differential equation for $P(t, t+t, 0)$. The boundary conditions for equations (17–20) are given by solving equations (9–12) backwards in time from $\tau = \sigma = T$ to $\tau = 0, \sigma = T$. This backwards solution has to be performed only once, when $t = 0$.

We now make use of the following identity.

$$P(\tau, \sigma, F(\sigma)) = P(\tau, \sigma) + \Psi(\tau, \sigma)F(\sigma)[I - L(\tau, \sigma)F(\sigma)]^{-1}\Phi(\tau, \sigma)$$

This can be verified in a straightforward manner. Then

$$P(t, t+T, F(t+T)) = P + \Psi F(t+T)[I - LF(t+T)]^{-1}\Phi$$

These equations are simpler than they appear, since $\Phi = \Psi'$. They give us a quadratic differential equation in matrices of size $2n$, which is integrated forwards to give the controller.

5 State estimation and output feedback

In this section we consider the problem of optimally estimating the state. Let

$$J(t_i, t_f) = \int_{t_i}^{t_f} \left\{ z(t)'z(t) - \gamma^2 w(t)'w(t) \right\} dt$$

Let $u_{[0,t_i]} = \{\, u(t); 0 \le t \le t_i \,\}$. Let $t \in [t_i, t_f]$ be the current time. We then wish to find the feedback saddle point strategy for the u player given information $u_{[0,t]}$ and $y_{[0,t]}$. For the moving horizon controller we only need to know this strategy at time $t = t_i$.

In recent years $\mathcal{H}_\infty$ theory has been furnished with a certainty equivalence principle in several different formulations (Doyle et al. 1989, Whittle 1990, Basar & Bernhard 1991).

For the finite horizon time varying $\mathcal{H}_\infty$ problem Basar and Bernhard (Basar & Bernhard 1991, p. 119) give a derivation of the optimal estimator for output feedback. Stated in simple terms, their result indicates that one should look, at each instant of time t, for the worst possible disturbance $w_{[t_i,t]}$ compatible with information available at that time about $u_{[t_i,t]}$ and $y_{[t_i,t]}$. This generates a corresponding worst possible state trajectory, which should be used by the feedback law as if it were the actual state.

The worst case approach to our problem is to extend the maximisation problem into the past, and maximise the cost function with respect to unobservables before t_i, back to time $t = 0$. This would require us to assume that, for times $0 < t < t_i$, w was trying to maximize the cost function on $[t_i, t_f]$. However, if we do this then w can make J_γ infinite, since there is no disincentive on the size of w outside $[t_i, t_f]$. This means that we cannot use this principle to estimate the state without further assumptions.

In fact, given the values of past observables, we can only draw conclusions about the current value of the state if we know some information about w (since we have the non-degeneracy assumption that $D_{21}D_{21}' = I$). In the finite horizon problem, we assume that in the past player w was trying to maximize the cost function over the entire finite interval.

One possible alternative approach is to assume that w is trying to maximize $J(0, t_f)$, while u is trying to minimize $J(t_i, t_f)$. For u, in fact this is equivalent to playing a moving horizon game in which only the endpoint moves.

Alternatively, it is possible to consider a stochastic formulation. In this case we should like to assume that outside the optimisation interval, w is accurately modelled by white noise. In order to do this, we need a stochastic model for $\mathcal{H}_\infty$ control.

5.1 Risk sensitive control

There is a strong connection between the formulation of risk sensitive control developed by Whittle (Whittle 1990) and $\mathcal{H}_\infty$ control. This is specified in (Glover & Doyle 1988). We consider here the system described by equations (1 – 3). However, we now change our assumptions so that w is white noise with covariance function $\delta(t)I$, the identity matrix multiplied by a delta function. Define the cost function

$$\mathcal{C} = \int_{t_i}^{t_f} z(t)'z(t)\,dt + x(t_f)'F(t_f)x(t_f)$$

where $x(t_i)$ is normally distributed with mean $\hat{x}_0$ and covariance matrix V, and $F(t) > 0$. Let

$$L_\gamma = 2\gamma^2 \log(\mathcal{E}_\mu(\exp(\gamma^2\mathcal{C}/2)))$$

where in this case $\mu(t, y(t))$ is a *policy*, and $\mathcal{E}_\mu$ indicates the expectation when the input is $u(t) = \mu(t, y(t))$. The risk sensitive control problem is

$$\min_{\mu\in\mathcal{M}} L_\gamma$$

where $\mathcal{M}$ is the space of all policy functions μ. Glover and Doyle (Glover & Doyle 1988) show that this is equivalent to an $\mathcal{H}_\infty$ problem. However, in this case, a sufficient statistic for the initial conditions are the mean and variance of the initial state. If the system has been running previous to the implementation of a risk sensitive controller, these can be provided by a Kalman filter.

The solution to the risk sensitive control problem is given by Whittle (Whittle 1990):

Theorem 6 *If there exists an $X_\infty \geq 0$ satisfying*

$$-\dot{X}_\infty = A'X_\infty + X_\infty A - X_\infty(B_2'B_2 - \gamma^{-2}B_1B_1')X_\infty + C_1'C_1 \quad X_\infty(t_f) = F(t_f)$$

and Y_∞ satisfying

$$\dot{Y}_\infty = AY_\infty + Y_\infty A' - Y_\infty(C_2'C_2 - \gamma^{-2}C_1'C_1)Y_\infty + B_1B_1' \quad Y_\infty(t_i) = V$$

such that $Y_\infty(t)^{-1} - \gamma^{-2}X_\infty(t) > 0$, then let

$$\dot{\hat{x}} = A\hat{x} + B_2u + Y_\infty C_2'(y - C_2\hat{x}) + \gamma^{-2}Y_\infty C_1'C_1\hat{x} \quad \hat{x}(t_i) = \hat{x}_0$$

and

$$\check{x} = (I - \gamma^{-2}Y_\infty X_\infty)^{-1}\hat{x}$$

Then the optimal risk sensitive feedback control law is given by

$$u = -B_2'X_\infty\check{x}$$

This has exactly the form of an $\mathcal{H}_\infty$ controller (see, for example, Basar and Bernhard (Basar & Bernhard 1991, pp. 124-126)) with unknown weighted initial state. However, it is important to note that the assumptions are very different, and that w is *white noise.*

In the moving horizon case, we only need $u(t)$ at $t = t_i$. At this time, $\hat{x}(t_i)$ and $Y_\infty(t_i)$ are the mean and variance of the state. Note that the moving horizon objective at time t_i is the expectation of L_γ over the current state variability and future disturbances. The information required, that is the mean and variance of the current state, is precisely that provided by a Kalman filter, as follows.

Theorem 7 *Suppose $x(0)$ has mean $\hat{x}_0$ and covariance matrix V. Let $Y_2 > 0$ be the solution to*

$$\dot{Y}_2 = AY_2 + Y_2A' - Y_2(C_2'C_2)Y_2 + B_1B_1' \quad Y_2(t_i) = V$$

and let $\hat{x}$ satisfy

$$\dot{\hat{x}} = A\hat{x} + B_2u + Y_2C_2'(y - C_2\hat{x}) \quad \hat{x}(t_i) = \hat{x}_0$$

then $\hat{x}(t)$ is the mean of x and $Y_2(t)$ is the variance of x at time t.

Hence we propose a moving horizon controller of the form

$$\dot{Y}_2 = AY_2 + Y_2A' - Y_2(C_2'C_2)Y_2 + B_1B_1' \quad Y_2(t_i) = V$$

$$\dot{\hat{x}} = A\hat{x} + B_2u + Y_2C_2'(y - C_2\hat{x}) \quad \hat{x}(t_i) = \hat{x}_0$$

$$\begin{aligned} -\frac{\partial X_\infty(\tau,\sigma)}{\partial \tau} &= X_\infty(\tau,\sigma)A(\tau) + A(\tau)'X_\infty(\tau,\sigma) - \\ &\quad X_\infty(\tau,\sigma)(B_2(\tau)B_2(\tau)' - \gamma^{-2}B_1(\tau)B_1(\tau)')X_\infty(\tau,\sigma) + C(\tau)'C(\tau) \\ X_\infty(t,t) &= F(t) \text{ for all } t \end{aligned}$$

Then the controller is given by

$$u(t) = -B_2(t)'X_\infty(t,t+T)(I - \gamma^{-2}Y_2(t)X_\infty(t,t+T))^{-1}\hat{x}(t)$$

Let $Z_{2\infty}(t) = (I - \gamma^{-2}Y_2(t)X_\infty(t,t+T))^{-1}Y_2(t)$ Then we can rewrite equations (5.1–5.1) as

$$\dot{\check{x}} = (A - (B_2B_2' - \gamma^{-2}B_1B_1')X_\infty)\check{x} - Z_{2\infty}(C_2'C_2 + \gamma^{-2}R)\check{x} + Z_{2\infty}C_2'y$$

$$u(t) = -B_2'X_\infty(t,t+T)\check{x}(t)$$

where

$$R(t) = C_1(t)'C_1(t) - \left.\frac{\partial X_\infty(t,\sigma)}{\partial \sigma}\right|_{\sigma=t+T}$$

Theorem 8 *If X_∞ and $Z_{2\infty}$ satisfy*

$$\alpha_1 I \le X_\infty(t,t+T) \le \alpha_2 I$$

$$\alpha_1 I \le Z_{2\infty}(t) \le \alpha_2 I$$

for fixed α_1, α_2 for all $t > 0$, then the above controller is globally exponentially stable.

Proof Let $e = x - \check{x}$. Then

$$\frac{d}{dt}\begin{pmatrix} x \\ e \end{pmatrix} = F \begin{pmatrix} x \\ e \end{pmatrix}$$

where

$$F = \left[\begin{array}{cc} A - B_2 B_2' X_\infty & B_2 B_2' X_\infty \\ -\gamma^{-2} B_1 B_1' X_\infty + \gamma^{-2} Z_{2\infty} R & A - Z_{2\infty}(C_2' C_2 + \gamma^{-2} R) + \gamma^{-2} B_1 B_1' X_\infty \end{array} \right]$$

Let

$$\Omega = \left[\begin{array}{cc} X_\infty & 0 \\ 0 & \gamma^2 Z_{2\infty}^{-1} \end{array} \right]$$

Let $G = F'\Omega + \Omega F + \dot{\Omega}$. Then

$$\frac{d}{dt}\left(\begin{pmatrix} x \\ e \end{pmatrix}' \Omega \begin{pmatrix} x \\ e \end{pmatrix} \right) = \begin{pmatrix} x \\ e \end{pmatrix}' G \begin{pmatrix} x \\ e \end{pmatrix}$$

Using

$$\frac{dZ_{2\infty}^{-1}}{dt} = -Z_{2\infty}^{-1} A - A' Z_{2\infty}^{-1} + \gamma^{-2} R + C_2' C_2 - Y_2^{-1} B_1 B_1' Y_2^{-1} - \gamma^{-2} X_\infty (B_2 B_2' - \gamma^{-2} B_1 B_1') X_\infty$$

some algebra reveals

$$G = \left[\begin{array}{cc} -X_\infty B_2 B_2' X_\infty - \gamma^{-2} X_\infty B_1 B_1' X_\infty - R & X_\infty B_2 B_2' X_\infty - X_\infty B_1 B_1' Z_{2\infty}^{-1} + R \\ X_\infty B_2 B_2' X_\infty - Z_{2\infty}^{-1} B_1 B_1' X_\infty + R & \begin{array}{c} -X_\infty B_2 B_2' X_\infty - \gamma^2 Z_{2\infty}^{-1} B_1 B_1' Z_{2\infty}^{-1} \\ -\gamma^{-2}(C_2' C_2 + \gamma^{-2} R) \end{array} \end{array} \right]$$

hence

$$\frac{d}{dt}\left[\begin{pmatrix} x \\ e \end{pmatrix}' \Omega \begin{pmatrix} x \\ e \end{pmatrix} \right] = -\begin{pmatrix} x \\ e \end{pmatrix}' \hat{G} \begin{pmatrix} x \\ e \end{pmatrix} - |B_2' X_\infty \check{x}|^2 - \gamma^2 |C_2 e|^2 - \check{x}' R \check{x}$$

where

$$\hat{G} = -\gamma^{-2} \left[\begin{array}{c} X_\infty \\ \gamma^2 Z_{2\infty}^{-1} \end{array} \right] B_1 B_1' \left[\begin{array}{c} X_\infty \\ \gamma^2 Z_{2\infty}^{-1} \end{array} \right]'$$

We know that $\check{x} = 0$ if and only if $x = e$, and in this case

$$-\begin{pmatrix} x \\ e \end{pmatrix}' \hat{G} \begin{pmatrix} x \\ e \end{pmatrix} = -\gamma^2 |B_1' Y_2^{-1} x|^2 < \delta |x|^2 \text{ for some } \delta > 0$$

since by assumption $B_1 B_1' > \varepsilon I$ for some $\varepsilon > 0$. Using lemma 3, if F satisfies equation (7), then $R > 0$, since by assumption $C_1' C_1 > 0$. □

If we have a bound for X_∞, then in order to bound $Z_{2\infty}$ we need only bound Y_2. Bounds for Y_2 are given by Kalman (Kalman 1960). Therefore stability is again conditional on the existence of a bound for X_∞, of which we are assured in the case when $F = \infty$.

6 Conclusions

With the terminal constraint $x(t+T)=0$, the moving horizon state feedback problem is known to be stable and to result in a global norm bound for the closed loop system, under the assumptions of Theorem 5. Several interesting problems remain. It is possible within the current formulation to use a time varying γ, which might allow both for attempts to improve performance if the system becomes easier to control with time, and to reduce performance demands if the system becomes more difficult to control.

It is also feasible to extend the above approach for the observation feedback case so that the horizon becomes $[t-T_1, t+T_2]$. In this case, we might expect a more conservative controller.

One motivation for this control methodology has been that it is possible to implement it knowing only details of the system up to time $t+T$ ahead. In this formulation that knowledge is discreet; we assume that up to time $t+T$ we know exactly the system matrices, after that we know nothing about them. This seems unnatural, since we might hope that an online identification process would give us a description of the future system with *gradually* increasing uncertainty in time. The linking of the present controller with a mathematical model for such uncertainty might produce a more realistic control system.

References

Anderson, B. D. O. & Moore, J. B. (1989), *Optimal Control - Linear Quadratic Methods*, Prentice Hall.

Basar, T. & Bernhard, P. (1991), $\mathcal{H}_\infty$ *Optimal Control and Related Minimax Design Problems. A Dynamic Game Approach*, Systems and Control: Foundations and Applications, Birkhauser.

Basar, T. & Olsder, G. J. (1982), *Dynamic Noncooperative Game Theory*, Mathematics in Science and Engineering, Academic Press.

Bucy, R. S. (1967), 'Global theory of the riccati equation', *Journal of Computer and System Sciences* **1**, 349–361.

Chen, C. C. & Shaw, L. (1982), 'On receding horizon feedback control', *Automatica* **18**(3), 349–352.

Doyle, J. C., Glover, K., Khargonekar, P. P. & Francis, B. A. (1989), 'State-space solutions to standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems', *IEEE Transactions on Automatic Control* **34**(8), 831–847.

Glover, K. & Doyle, J. C. (1988), 'State-space formulae for all stabilizing controllers that satisfy an $\mathcal{H}_\infty$-norm bound and relations to risk sensitivity', *Systems and Control Letters* **11**, 167–172.

Kalman, R. E. (1960), 'Contributions to the theory of optimal control', *Boletin de la Sociedad Matematica Mexicana* **5**(1), 102–119.

Khargonekar, P. P., Nagpal, K. M. & Poolla, K. R. (1991), '$\mathcal{H}_\infty$ control with transients', *SIAM Journal on Control and Optimization* **29**(6), 1373–1393.

Kleinman, D. L. (1970), 'An easy way to stabilize a linear constant system', *IEEE Transactions on Automatic Control* p. 692.

Kwon, W. H., Bruckstein, A. M. & Kailath, T. (1983), 'Stabilizing state-feedback design via the moving horizon method', *International Journal of Control* **37**(3), 631–643.

Kwon, W. H. & Pearson, A. E. (1977), 'A modified quadratic cost problem and feedback stabilization of a linear system', *IEEE Transactions on Automatic Control* **22**(5), 838–842.

Limebeer, D. J., Anderson, B. D. O., Khargonekar, P. P. & Green, M. (1992), 'A game theoretic approach to $\mathcal{H}_\infty$ control for time varying systems', *SIAM Journal on Control and Optimization* **30**(2), 262–283.

Mayne, D. Q. & Michalska, H. (1990), 'Receding horizon control of nonlinear systems', *IEEE Transactions on Automatic Control* **35**(7), 814–824.

Ran, A. C. M. & Vreugdenhil, R. (1988), 'Existence and comparison theorems for algebraic riccati equations for continuous and discrete time systems', *Linear Algebra and its applications* **99**, 63–83.

Ravi, R., Nagpal, K. M. & Khargonekar, P. P. (1991), '$\mathcal{H}_\infty$ control of linear time varying systems: A state space approach', *SIAM Journal on Control and Optimization* **29**(6), 1394–1413.

Reid, W. T. (1970), *Riccati Differential Equations*, Academic Press.

Tadmor, G. (1992), 'Receding horizon revisited: An easy way to robustly stabilize an ltv system', *Systems and Control Letters* **18**, 285–294.

Tadmor, G. (1993), 'The standard $\mathcal{H}_\infty$ problem and the maximum principle: the general linear case', *SIAM Journal on Control and Optimization* **31**(4), 813–846.

Uchida, K. & Fujita, M. (1992), 'Finite horizon $\mathcal{H}_\infty$ control problems with terminal penalties', *IEEE Transactions on Automatic Control* **37**(11), 1762–1767.

Verghese, G., Friedlander, B. & Kailath, T. (1980), 'Scattering theory and linear least squares estimation, part iii: The estimates', *IEEE Transactions on Automatic Control* **25**(4), 794–802.

Whittle, P. (1990), *Risk-sensitive optimal control*, Wiley Interscience series in Systems and Optimization, Wiley.