CHAPTER 6

Input-output behavior

6.1. State observation

Consider a MIMO system

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx. \end{aligned} \tag{6.1}$$

and the problem of reconstructing the state x from the output y.

As we have learned from a basic control course, one obvious way to estimate the state is to use an error correction on the system equation:

$$\dot{z} = Az + Bu + L(Cz - y), \tag{6.2}$$

where L is designable. The fact that the same A and C matrices are used is due to that if the observation error $Cz - y$ is identically zero, then z should x.

Let $e := z - x$ be the difference between the estimated and actual state trajectory, then

$$\dot{e} = (A + LC)e.$$

Naturally the difference will tend to zero if and only if all eigenvalues of $(A + LC)$ have negative real parts. Obviously we can find an L such that $(A + LC)$ is stable if and only if the pair (C, A) is detectable, namely, all eigenvalues of A restricted to the unobservable subspace must have negative real part.

It is well known that for a linear system, one can separate the designs for a stabilizing controller and for a state observer. This principle can be stated as follows.

PROPOSITION 6.1 (Separation principle). *Suppose* (6.1) *is stabilizable and detectable, then for any F and L such that $A + BF$ and $A + LC$ are stable, the closed-loop system resulting from using $u = Fz$ and the observer*

$$\begin{aligned} \dot{x} &= Ax + BFz \\ \dot{z} &= (A + BF)z + LC(z - x) \end{aligned} \tag{6.3}$$

is also stable.

PROOF
Denote

$$J := \begin{bmatrix} A & BF \\ -LC & A + BF + LC \end{bmatrix},$$

then by a similarity transformation we have

$$\begin{bmatrix} I & 0 \\ -I & I \end{bmatrix} J \begin{bmatrix} I & 0 \\ I & I \end{bmatrix} = \begin{bmatrix} A+BF & BF \\ 0 & A+LC \end{bmatrix}.$$

Thus, J is a stable matrix. ∎

6.2. Output tracking input

In this section we review some classical results on asymptotic input tracking. Consider a stable, controllable and observable SISO linear system:

$$\begin{aligned} \dot{x} &= Ax + bu \\ y &= cx \end{aligned} \tag{6.4}$$

where $x \in R^n$ and $\sigma(A) \in C^-$.

We will consider the case when the input u is generated by the following exogenous system:

$$\begin{aligned} \dot{w} &= \Gamma w \\ u &= qw \end{aligned} \tag{6.5}$$

where $w \in R^m$ and $\sigma(\Gamma) \in \bar{C}^+$. This exo-system can generally have a block diagonal Jordan realization

$$\begin{aligned} q &= \begin{pmatrix} q_1 & q_2 & \dots & q_M \end{pmatrix} \\ \Gamma &= \operatorname{diag}(\Gamma_1, \Gamma_2, \dots, \Gamma_M) \end{aligned} \tag{6.6}$$

where each $q_i = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}$ is a first unit vector of length $\dim(\Gamma_i)$ and the Jordan blocks correspond to polynomial, exponential, and sinusoidal functions. The output of the exo-system becomes

$$u(t) = \sum_{i=1}^{M} q_i e^{\Gamma_i t} w_{0_i} = q e^{\Gamma t} w_0.$$

Such exo-systems can generate, for example, step functions, ramp functions, polynomials, exponentials, sinusoidals, and combinations of such functions.

Proposition 6.2. *Suppose A is a stable matrix, then all trajectories of $(x(t), w(t))$ tend asymptotically to the invariant subspace $S := \{(x,w) : \ x = \Pi w\}$, where Π is the solution of*

$$A\Pi - \Pi\Gamma = -bq.$$

On the invariant subspace, we have

$$y(t) = c\Pi w(t).$$

Proof

We will establish that there exists a matrix Π so that the set

$$S = \{(x, w) : \ x = \Pi w\}$$

is invariant under the action of the linear system in the following sense. Let

$$u = qw$$

then in order to show S is invariant we only need to show $\dot{x} = \Pi\dot{w}$. Namely,

$$\Pi\Gamma w = A\Pi w + bqw \tag{6.7}$$

which gives us the Lyapunov equation

$$\Pi\Gamma - A\Pi = bq. \tag{6.8}$$

Now the eigenvalues of Γ are in the right half plane and the eigenvalues of $-A$ are in the strict right half plane and hence no sum of eigenvalues is zero. Thus there exists a unique solution Π to the equation.

Now let $e = x - \Pi w$, one can easily show that $e \to 0$ as $t \to \infty$ by the assumption that A is a stable matrix. ∎

Using the matrix Π we have that the output of the linear system in the steady-state can be represented as

$$y = c\Pi w.$$

PROPOSITION 6.3. *Let the system*

$$\dot{w} = \Gamma w, \quad u = qw$$

be observable and no eigenvalue of Γ *is a transmission zero of (6.4). Then the system on the invariant subspace*

$$\dot{w} = \Gamma w, \quad y = c\Pi w$$

is also observable.

PROOF

We first need to establish that under the hypotheses, the composite system

$$\begin{aligned} \begin{pmatrix} \dot{x} \\ \dot{w} \end{pmatrix} &= \begin{pmatrix} A & bq \\ 0 & \Gamma \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} \\ y &= cx \end{aligned} \tag{6.9}$$

is observable. Methods for proving similar results can be found, for example, in [**4**].

Define

$$H(s) := \begin{pmatrix} sI - A & -bq \\ 0 & sI - \Gamma \\ c & 0 \end{pmatrix}.$$

By Hautus test we know that the system is observable if and only if

$$\text{rank}(H(s)) = n + m \ \forall s.$$

If s is not an eigenvalue of Γ, it is easy to see that $rank(H(s)) = n+m$ since (c, A) is observable. Now suppose s is an eigenvalue of Γ,

$$H(s) = \begin{pmatrix} sI - A & b & 0 \\ 0 & 0 & I_m \\ c & 0 & 0 \end{pmatrix} \begin{pmatrix} I_n & 0 \\ 0 & -q \\ 0 & sI - \Gamma \end{pmatrix}.$$

If s is not a transmission zero of (6.4), then and only then the first matrix on the right-hand side has rank $n+1+m$. The second has rank $n+m$ since (q, Γ) is observable. By Sylvester's inequality, we have

$$\mathrm{rank}(H(s)) \geq n+1+m+n+m-(n+m+1) = n+m.$$

Therefore $\mathrm{rank}(H(s)) = n+m$.

Now we do a coordinate change $\bar{x} := x - \Pi w$. Then (6.9) becomes

$$\begin{aligned} \begin{pmatrix} \dot{\bar{x}} \\ \dot{w} \end{pmatrix} &= \begin{pmatrix} A & 0 \\ 0 & \Gamma \end{pmatrix} \begin{pmatrix} \bar{x} \\ w \end{pmatrix} \\ y &= c\bar{x} + c\Pi w \end{aligned}$$

It is straight forward to see that

$$\left((c, c\Pi), \begin{pmatrix} A & 0 \\ 0 & \Gamma \end{pmatrix} \right)$$

is observable implies $(c\Pi, \Gamma)$ is so too. ∎

We will now discuss how Propositions 6.2 and 6.3 can be used to determine an appropriate output in order to track the input exactly in stationarity. It follows that the output tracks the input if the vector c is chosen such that

$$(c\Pi - q)e^{\Gamma t} w_0 = 0 \tag{6.10}$$

where w_0 is the initial state of (6.5) that generates the input. This is clearly the case if $c\Pi = q$. If Π, for example, has full column rank, then it is possible to design an output c for perfect input tracking in stationarity. We will show below that this is the case if (A, b) is controllable, (q, Γ) is observable and $\dim(A) \geq \dim(\Gamma)$.

In some way one may view this problem as a dual one to the output regulation problem discussed in [**7**]. In this section, we discuss some necessary and sufficient conditions.

THEOREM 6.4. *Suppose (q, Γ) is observable and (A, b) controllable. Then a necessary and sufficient condition for the existence of c, such that $c\Pi = q$, is that the dimension of A is greater than or equal to that of Γ.*

PROOF

We can rewrite $\dot{x} = Ax + bu$ in the canonical form:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ &\vdots \\ \dot{x}_{n-1} &= x_n \\ \dot{x}_n &= -\sum_{i=1}^{n} a_i x_i + ku, \end{aligned} \tag{6.11}$$

where $k \neq 0$ and $\rho(s) := s^n + \sum_{i=1}^{n} a_i s^{i-1}$ is Hurwitz. In the steady state, by Proposition 6.2 we have

$$x_1 = \pi_1 w,$$

where π_1 is the first row of Π. Since $x_i = \pi_1 \Gamma^{i-1} w$, for $i = 1, \ldots, n$, we have

$$\Pi = \begin{pmatrix} \pi_1 \\ \pi_1 \Gamma \\ \vdots \\ \pi_1 \Gamma^{n-1} \end{pmatrix}. \tag{6.12}$$

Then from the last equation of (6.11) we have

$$\pi_1 \Gamma^n = -\sum_{i=1}^{n} a_i \pi_1 \Gamma^{i-1} + kq.$$

Since by assumption Γ does not have any eigenvalue in the open left-half plane, we have

$$\pi_1 = kq\rho(\Gamma)^{-1}. \tag{6.13}$$

If there exists a c, such that

$$q = c\Pi = \sum_{i=1}^{n} c_i \pi_1 \Gamma^{i-1},$$

then

$$q = kq\rho(\Gamma)^{-1} \sum_{i=1}^{n} c_i \Gamma^{i-1},$$

or equivalently,

$$q\left(I - k\rho(\Gamma)^{-1} \sum_{i=1}^{n} c_i \Gamma^{i-1}\right) = 0.$$

Denote $\Delta = I - k\rho(\Gamma)^{-1} \sum_{i=1}^{n} c_i \Gamma^{i-1}$. It is easy to show that

$$\Delta = \left(\Gamma^n + \sum_{i=1}^{n} (a_i - kc_i)\Gamma^{i-1}\right) \rho(\Gamma)^{-1}$$

thus $q\Delta = 0$ if and only if

$$q\Gamma^n + \sum_{i=1}^{n}(a_i - kc_i)q\Gamma^{i-1} = 0. \tag{6.14}$$

Since (q, Γ) is observable, (6.14) has a solution if and only if n is greater than or equal to the dimension of Γ. ∎

Corollary 6.5. *If dim$(A) \geq$ dim(Γ), then Π has full column rank, and thus there exists c, such that,*

$$c\Pi = q.$$

Moreover, if dim$(A) =$ dim(Γ), then such c is unique.

Proof
It follows from (6.12), (6.13), (6.14), and observability of the exo-system (6.5). Indeed,

$$\Pi = \begin{pmatrix} \pi_1 \\ \pi_1\Gamma \\ \vdots \\ \pi_1\Gamma^{n-1} \end{pmatrix} = k \begin{pmatrix} q \\ q\Gamma \\ \vdots \\ q\Gamma^{n-1} \end{pmatrix} \rho(\Gamma)^{-1}$$

which has full rank since (q, Γ) is observable. ∎

Corollary 6.6. *Suppose dim$(A) = n \geq$ dim$(\Gamma) = m$, then there exists c such that $c\Pi = q$ and the resulting system* (6.9) *is observable and* (6.4) *does not have any transmission zero that is also an eigenvalue of Γ.*

Proof
Consider the canonical form (6.11). Suppose the characteristic polynomial for Γ is $\rho_\Gamma(s) = s^m + \sum_{i=1}^{m} \gamma_i s^{i-1}$. It follows from (6.14) and Cayley-Hamilton that

$$c_i = \frac{1}{k}(a_i - \bar{\gamma}_i) \ i = 1, \ldots, n,$$

where $\bar{\gamma}_i = 0 \ \forall i < n - m + 1$ and $\bar{\gamma}_i = \gamma_{i-n+m}$ otherwise, is a solution such that $c\Pi = q$. It then follows from the fact that A and Γ do not share any eigenvalue, that no eigenvalue s_0 of A or Γ is a root of

$$\sum_{i=1}^{n} c_i s_0^{i-1} = \frac{1}{k}\sum_{i=1}^{n} a_i s_0^{i-1} - \frac{s_0^{n-m}}{k}\sum_{i=1}^{m} \gamma_i s_0^{i-1}.$$

Indeed, if s_0 is for example a root of the characteristic polynomial of A, the above expression reduces to

$$-\frac{s_0^{n-m}}{k}\rho_\Gamma(s_0),$$

which must be nonzero. Thus, no transmission zero of the corresponding (6.4) is an eigenvalue of Γ and the pair (c, A) is observable. From the proof of Proposition 6.3 we derive that (6.9) is observable. ∎

We have shown under the assumptions of Corollary 6.5 that the input u can be reconstructed simply as

$$\hat{u} = cx,$$

where, e.g. $c = q\Pi^{\dagger}$ ($\dagger$ denotes pseudo inverse). The tracking error satisfies (here $\bar{x} = x - \Pi w$)

$$\bar{u} := u - \hat{u} = cx - qw = c\bar{x} + c\Pi w - qw = c\bar{x}$$

Hence, the error dynamics in this case becomes

$$\begin{aligned} \dot{\bar{x}} &= A\bar{x} \\ \bar{u} &= c\bar{x} \end{aligned}$$

which has its rate of convergence limited by the eigenvalues of A.

EXAMPLE 6.1. *We consider a car-like base and a manipulator mounted on it. By using the homogeneous representation of rigid body motions, we can easily compute the position of the end-effector, relative to the base,* r_A^B*, and thus the kinematic model as*

$$\dot{x}_A = f(x_A, u).$$

For the car, we use the model we have studied in Chapter 1:

$$\begin{aligned} \dot{\alpha}_f &= a_{11}\alpha_f + r + \dot{\delta}_f \\ \dot{\psi} &= r \\ \dot{r} &= a_{21}\alpha_f + a_{22}r + b_{21}\delta_f + d(t) \\ y_1 &= \alpha_f \\ y_2 &= \psi \end{aligned}$$

The problem we want to study is that by measuring the orientation (ψ) and yaw rate (r), is it possible to recover $d(t)$ at least in some cases? Since this information shall be useful to know for the control of the manipulator. In other words, can we find a $c = (0\ c_2\ c_3)$ such that $y = d$ in stationarity?

We leave this as an exercise.

6.3. The partial stochastic realization problem

In this section, we will present the results in discrete time. Although one can easily translate these results into continuous time, it makes more sense to discuss them in discrete time, judging by their potential applications.

Given a partial covariance sequence

$$\{c_0, c_1, c_2, \ldots, c_n\}, \tag{6.15}$$

consider the problem to determine a stationary stochastic model

$$(\Sigma) \quad \begin{cases} x(t+1) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{cases} \tag{6.16}$$

such that

$$\mathrm{E}\{y(t+k)y(t)\} = c_k \quad k = 0,1,2,\dots,n.$$

Since clearly $c_0 > 0$, it is no restriction setting $c_0 = 1$. It just amounts to scaling the process.

This is an important problem in systems theory, appearing in many applications, among them speech processing and spectral estimation. A very popular solution of this problem is the so called *maximum entropy solution*, which we shall describe next.

Let us first return to the Schur parameters. It was shown by Schur at the beginning of the century that there is a one-one correspondence between infinite covariance sequences

$$\{c_0, c_1, c_2, c_3, \dots\}$$

and infinite sequences

$$\{\gamma_0, \gamma_1, \gamma_2, \dots\}$$

such that $|\gamma_t| < 1$ for all $t \geq 0$, under which the partial sequence $\{c_0, c_1, c_2, \dots, c_m\}$ is uniquely determined by the partial sequence $\{\gamma_0, \gamma_1, \dots, \gamma_m\}$ and vice versa, for each $m \geq 0$. In particular, (6.15) defines

$$\{\gamma_0, \gamma_1, \dots, \gamma_{n-1}\}$$

uniquely via the Levinson equation. Consequently, each extension

$$\{\gamma_n, \gamma_{n+1}, \gamma_{n+2}, \dots\}$$

of Schur parameters with the property $|\gamma_t| < 1$ corresponds to an extension of the covariance sequence for which the Toeplitz matrix

$$\begin{bmatrix} c_0 & c_1 & c_2 & c_3 & \cdots \\ c_1 & c_0 & c_1 & c_2 & \cdots \\ c_2 & c_1 & c_0 & c_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix} > 0.$$

However, this covariance sequence may not correspond to a rational spectral density and thus to a finite-dimensional stochastic model Σ.

One choice that does is

$$\gamma_n = \gamma_{n+1} = \gamma_{n+2} = \cdots = 0,$$

which yields precisely the maximum entropy solution. In fact, from the Levinson algorithm we have

$$\varphi_{n+k}(z) = z^k \varphi_n(z)$$

so that, for $t \geq n$,

$$\varphi_{tk} = \begin{cases} \varphi_{nk} & \text{for } k = 0,1,2,\dots,n \\ 0 & \text{for } k > 0 \end{cases}$$

Hence,

$$y(t) = -\varphi_{n1}y(t-1) - \varphi_{n2}y(t-2) - \cdots - \varphi_{nn}y(t-n) + \sqrt{r_n}u(t).$$

Introducing the state process

$$x(t) = \begin{bmatrix} y(t-n) \\ y(t-n+1) \\ \vdots \\ y(t-1) \end{bmatrix}$$

we obtain a state space representation Σ with

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \hdotsfor{5} \\ 0 & 0 & 0 & \cdots & 1 \\ -\varphi_{nn} & -\varphi_{n,n-1} & -\varphi_{n,n-2} & \cdots & -\varphi_{n1} \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \sqrt{r_n} \end{bmatrix}$$

$$C = \begin{bmatrix} -\varphi_{nn} & -\varphi_{n,n-1} & -\varphi_{n,n-2} & \cdots & -\varphi_{n1} \end{bmatrix} \quad D = \sqrt{r_n}$$

which is on reachable canonical form. Hence,

$$\begin{aligned} W(z) &= C(zI - A)^{-1}B + D \\ &= \frac{z^n - \varphi_n(z)}{\varphi_n(z)}\sqrt{r_n} + \sqrt{r_n} \\ &= \frac{\sqrt{r_n} z^n}{\varphi_n(z)} \end{aligned}$$

yielding the spectral density

$$\Phi(z) = W(z)W(1/z) = \frac{r_n}{\varphi_n(z)\varphi_n(1/z)},$$

i.e., the maximum entropy solution has no zeros.

In many applications, such as speech processing, we would like to have zeros. In fact, we may want to have a particular set of zeros. Can we achieve this? Schur-parameter extension will in general not do since we require the system Σ to be finite-dimensional of dimension at most n.

Very recently we have proved the following.

THEOREM 6.7. *To each stable polynomial*

$$\sigma(z) = z^n + \sigma_1 z^{n-1} + \cdots + \sigma_n$$

there is one and only one stable polynomial

$$a(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n \quad a_0 > 0$$

such that a system Σ with transfer function

$$W(z) = \frac{\sigma(z)}{a(z)}$$

satisfies

$$E\{y(t+k)y(t)\} = c_k \quad k = 0, 1, 2, \ldots .n.$$

This theorem gives a *complete parameterization* of the partial stochastic realization problem in terms of zeros. The zeros of

$$\Phi(z) = \frac{\sigma(z)\sigma(1/z)}{a(z)a(1/z)}$$

can be chosen arbitrarily (subject to the obvious limitations due to the fact that Φ is a spectral density). Moreover there is a *unique* partial stochastic realization to each admissible zero structure.