Observers, invariance and autonomy

J. Trumpf

ANU, Canberra

based on joint work with C. Lageman and R. Mahony

July 2008

Outline

Outline

Left invariant systems

We consider systems of the form

$$\dot{X} = Xu$$

where the *state* X evolves on a finite dimensional, connected Lie group G and the (admissible) *input* is in the associated Lie algebra $\mathfrak{g}$.

Left invariant systems

We consider systems of the form

$$\dot{X} = Xu$$

where the *state* X evolves on a finite dimensional, connected Lie group G and the (admissible) *input* is in the associated Lie algebra $\mathfrak{g}$.

Xu is shorthand notation for T_eL_Xu where T_eL_X is the derivative of the left multiplication map

$$L_X: \ G \longrightarrow G, \ \ Y \mapsto XY$$

at the identity element $Y = e$ of G. This is then a map $\mathfrak{g} \simeq T_eG \longrightarrow T_XG$, so $X \mapsto Xu$ is a left invariant vector field.

Left invariant systems

This is a special case of a well studied class of systems on Lie groups (Brockett, Jurdjevic, Sussmann, ...) with no drift and a full basis set of control vector fields.

Left invariant systems

This is a special case of a well studied class of systems on Lie groups (Brockett, Jurdjevic, Sussmann, ...) with no drift and a full basis set of control vector fields.

For $G = SO(3)$ and $\mathfrak{g} = \mathfrak{so}(3)$ these are the kinematic equations

$$\dot{R} = R\Omega$$

describing the time evolution of the attitude (=orientation) of the center of mass of a rigid body in 3D space. Here, R is the rotation matrix relating the inertial coordinate frame to the body-fixed frame and Ω contains the angular velocities.

Similar for $G = SE(3) \simeq SO(3) \ltimes \mathbb{R}^3$ and $\mathfrak{g} = \mathfrak{se}(3)$.

Outputs

$$\dot{X} = Xu$$

We consider outputs of the form

$$y = h(X, y_0)$$

with a *right* action $h : G \times M \longrightarrow M$, where M is a smooth manifold (i.e. a homogeneous space of G).

Outputs

$$\dot{X} = Xu$$

We consider outputs of the form

$$y = h(X, y_0)$$

with a *right* action $h : G \times M \longrightarrow M$, where M is a smooth manifold (i.e. a homogeneous space of G).

For $G = SO(3)$ think of $M = S^2$ and

$$y = R^T y_0$$

which describes the direction a fixed landmark is seen in by, say, a camera mounted on board.

Note that this is *different* from the case

$$y = h(X, y_0)$$

where h is a *left* action.

For $G = SO(3)$ think of $M = S^2$ and

$$y = Ry_0$$

which describes the direction a fixed marking on the rigid body is seen in from the ground.

Note that this is *different* from the case

$$y = h(X, y_0)$$

where h is a *left* action.

For $G = SO(3)$ think of $M = S^2$ and

$$y = Ry_0$$

which describes the direction a fixed marking on the rigid body is seen in from the ground.

Only this latter case has been studied in the literature (in the context of control).

The problem

$$\dot{X} = Xu$$
$$y = h(X, y_0)$$

Suppose we have measurements of u and measurements of y. We want to construct observers that estimate X, i.e. systems with input (u, y) and state $\hat{X}$ where $\hat{X}$ is a reasonable estimate of X.

The problem

$$\dot{X} = Xu$$
$$y = h(X, y_0)$$

Suppose we have measurements of u and measurements of y. We want to construct observers that estimate X, i.e. systems with input (u, y) and state $\hat{X}$ where $\hat{X}$ is a reasonable estimate of X.

Think: noisy measurements ...

Outline

Systems with symmetry

A system that is globally given by

$$\dot{x} = f(x, u), \quad x \in N$$

can be regarded as a map $f : B \longrightarrow TN$ where B is a trivial bundle over N. (We allow general bundles here.)

Add an output $h : N \longrightarrow M$.

Systems with symmetry

A system that is globally given by

$$\dot{x} = f(x, u), \quad x \in N$$

can be regarded as a map $f : B \longrightarrow TN$ where B is a trivial bundle over N. (We allow general bundles here.)

Add an output $h : N \longrightarrow M$.

A Lie group H is called a *symmetry* of this system if there are left actions S^B and S^N and a right action S^M such that

$$f(S^B(X, v)) = TS^N_X f(v)$$

$$h(S^N(X, x)) = S^M(X, h(x))$$

(Cf. Grizzle/Marcus, Nijmeijer/van der Schaft, Tabuda/Pappas)

Projected systems

$$\begin{aligned} \dot{X} &= Xu \\ y &= h(X, y_0) \end{aligned} \tag{S}$$

Proposition: $\mathrm{stab}(y_0)$ is a symmetry of (S).

Theorem: (S) projects to the system on M

$$\dot{y} = T_X\pi(Xu), \quad X \in \pi^{-1}(y)$$

where $\pi : G \longrightarrow M$ is the canonical projection.

Projected systems

$$\begin{aligned} \dot{X} &= Xu \\ y &= h(X, y_0) \end{aligned} \qquad \text{(S)}$$

Proposition: stab(y_0) is a symmetry of (S).

Theorem: (S) projects to the system on M

$$\dot{y} = T_X\pi(Xu), \quad X \in \pi^{-1}(y)$$

where $\pi : G \longrightarrow M$ is the canonical projection.

Corollary: Two states $X, Y \in G$ are indistinguishable if and only if $XY^{-1} \in \text{stab}\, y_0$.

(Cf. Sussmann)

The idea

Construct an observer for the projected system and lift it up.

Outline

How to measure errors?

(Smooth) *error functions*

$$E: \; M \times M \longrightarrow N$$

Definition: Two systems

$$\dot{y} = f_y(u, t),$$
$$\dot{\hat{y}} = f_{\hat{y}}(u, t)$$

with common input are called *E-synchronous* if E is constant along corresponding trajectory pairs.

The canonical error function

There is a canonical error function

$$E_r(\hat{y}, y) = h(\hat{X}X^{-1}, y_0)$$

where $\pi(\hat{X}) = \hat{y}$ and $\pi(X) = y$.

The canonical error function

There is a canonical error function

$$E_r(\hat{y}, y) = h(\hat{X}X^{-1}, y_0)$$

where $\pi(\hat{X}) = \hat{y}$ and $\pi(X) = y$.

Theorem: Only $\dot{\hat{y}} = T_{\hat{X}}\pi(\hat{X}u)$ is E_r-synchronous to $\dot{y} = T_X\pi(Xu)$.

Theorem: If $\dot{\hat{y}} = T_{\hat{X}}\pi(\hat{X}u)$ is E-synchronous to $\dot{y} = T_X\pi(Xu)$ then $E = g \circ E_r$.

Outline

Internal models

Consider system pairs

$$\dot{y} = f_x(u, t), \tag{S}$$

$$\dot{\hat{y}} = f_{\hat{y}}(u, y, t). \tag{O}$$

Definition: (O) has an *internal model* of (S) if $\hat{y} = y$ for corresponding trajectories.

Definition: We call a map

$$\alpha : \; M \times M \times TM \times \mathbb{R} \longrightarrow TM$$

an *innovation term* if $\alpha(\hat{y}, y, u, t) \in T_{\hat{y}}M$ and α is zero along corresponding trajectories.

A nice little result

$$\dot{y} = T_X \pi(Xu), \tag{S}$$
$$\dot{\hat{y}} = f_{\hat{y}}(\hat{y}, y, u, t). \tag{O}$$

Theorem: (O) has an internal model of (S) if and only if it has the form

$$\dot{\hat{y}} = T_{\hat{X}} \pi(\hat{X}u) + \alpha(\hat{y}, y, u, t),$$

i.e. if and only if it splits into an E_r-synchronous term and an innovation term.

Outline

Good innovation terms

$$\dot{y} = T_X \pi(Xu), \tag{S}$$

$$\dot{\hat{y}} = T_{\hat{X}} \pi(\hat{X}u) + \alpha(\hat{y}, y, u, t). \tag{O}$$

Observer design in this framework amounts to finding a good choice for α.

Good innovation terms

$$\dot{y} = T_X\pi(Xu), \tag{S}$$

$$\dot{\hat{y}} = T_{\hat{X}}\pi(\hat{X}u) + \alpha(\hat{y}, y, u, t). \tag{O}$$

Observer design in this framework amounts to finding a good choice for α.

Theorem: α yields autonomous E_r dynamics if and only if it is equivariant, i.e.

$$T_{\hat{y}}h_S\alpha(\hat{y}, y, u, t) = \alpha(h_S(\hat{y}), h_S(y), u, t),$$

and independent of u and t. Then

$$\dot{E}_r = \alpha(E_r, y_0).$$

Gradient innovations

Introduce a smooth, non-negative *cost function*

$$f: \; M \times M \longrightarrow \mathbb{R}$$

and choose a Riemannian metric on M.

We propose observers of the form

$$\dot{\hat{y}} = T_{\hat{X}}\pi(\hat{X}u) - \text{grad}_1 f(\hat{y}, y)$$

Gradient innovations

Introduce a smooth, non-negative *cost function*

$$f : \; M \times M \longrightarrow \mathbb{R}$$

and choose a Riemannian metric on M.

We propose observers of the form

$$\dot{\hat{y}} = T_{\hat{X}}\pi(\hat{X}u) - \mathrm{grad}_1 f(\hat{y}, y)$$

Theorem: Assume now that M is reductive. If f is constant along the diagonal and both f and the Riemannian metric are invariant then

$$\dot{E}_r = -\mathrm{grad}_1 f(E_r, y_0).$$

A convergence result

$$\dot{y} = T_X\pi(Xu), \tag{S}$$

$$\dot{\hat{y}} = T_{\hat{X}}\pi(\hat{X}u) - \mathrm{grad}_1 f(\hat{y}, y) \tag{O}$$

where f and the Riemannian metric are both invariant.

Corollary: Let additionally $\hat{y} \mapsto f(\hat{y}, y_0)$ be a Morse-Bott function with a global minimum at y_0 and no other local minima. Then E_r converges to y_0 for generic initial conditions.

The lift

$$\dot{X} = Xu$$
$$y = h(X, y_0)$$

The lifted observer takes the form

$$\dot{\hat{X}} = \hat{X}u - \left(\text{grad}_1 f(\pi(\hat{X}), y)\right)^H$$
$$= \hat{X}u - \text{grad}_1 \tilde{f}(\hat{X}, X)$$

where $\tilde{f}(\hat{X}, X) = f(\pi(\hat{X}), \pi(X))$. We get

$$\hat{X}X^{-1} \to \text{stab}(y_0).$$

Outline

A first conclusion

Brockett is right:

> *Our main point is that this class of systems is in many ways not more difficult than linear systems of the usual type in $\mathbb{R}^n$.*

[System theory on group manifolds and coset spaces, SIAM J. Control, 10(2), 1972, p.265]