

Massachusetts Institute of Technology

Department of Electrical Engineering and Computer Science

6.243j (Fall 2003): DYNAMICS OF NONLINEAR SYSTEMS

by A. Megretski

Lecture 13: Feedback Linearization¹

Using control authority to transform nonlinear models into linear ones is one of the most commonly used ideas of practical nonlinear control design. Generally, the trick helps one to recognize “simple” nonlinear feedback design tasks.

13.1 Motivation and objectives

In this section, we give a motivating example and state technical objectives of theory of feedback linearization.

13.1.1 Example: fully actuated mechanical systems

Equations of rather general mechanical systems can be written in the form

$$M(q(t))\ddot{q}(t) + F(q(t), \dot{q}(t)) = u(t), \quad (13.1)$$

where $q(t) \in \mathbf{R}^k$ is the position vector, $u(t)$ is the vector of actuation forces and torques, $F : \mathbf{R}^k \times \mathbf{R}^k \mapsto \mathbf{R}^k$ is a given vector-valued function, and $M : \mathbf{R}^k \mapsto \mathbf{R}^{k \times k}$ is a given function taking positive definite symmetric matrix values (the inertia matrix). When $u = u(t)$ is fixed (for example, when $u(t) = u_0 \cos(t)$ is a harmonic excitation), analysis of (13.1) is usually an extremely difficult task. However, when $u(t)$ is an unrestricted control effort to be chosen, a simple change of control variable

$$u(t) = M(q(t))(v(t) + F(q(t), \dot{q}(t))) \quad (13.2)$$

transforms (13.1) into a linear double integrator model

$$\ddot{q}(t) = v(t). \quad (13.3)$$

¹Version of October 29, 2003

The transformation from (13.1) to (13.3) is a typical example of *feedback linearization*, which uses a *strong control authority* to simplify system equations. For example, when (13.1) is an *underactuated* model, i.e. when $u(t)$ is restricted to a given subspace in \mathbf{R}^k , the transformation in (13.2) is not valid. Similarly, if $u(t)$ must satisfy an a-priori bound, conversion from v to u according to (13.2) is not always possible.

In addition, feedback linearization relies on access to *accurate information*, in the current example – precise knowledge of functions M, F and precise measurement of coordinates $q(t)$ and velocities $\dot{q}(t)$. While in some cases (including the setup of (13.1)) one can extend the benefits of feedback linearization to approximately known and imperfectly observed models, information flow constraints remain a serious obstacle when applying feedback linearization.

13.1.2 Output feedback linearization

Output feedback linearization can be viewed as a way of simplifying a nonlinear ODE control system model of the form

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t), \quad (13.4)$$

$$y(t) = h(x(t)), \quad (13.5)$$

where $x(t) \in U$ is the state vector ranging over a given open subset X_0 of \mathbf{R}^n , $u(t) \in \mathbf{R}^m$ is the control vector, $y(t) \in \mathbf{R}^m$ is the output vector, $f : X_0 \mapsto \mathbf{R}^n$, $h : X_0 \mapsto \mathbf{R}^m$, and $g : X_0 \mapsto \mathbf{R}^{n \times m}$ are given smooth functions. Note that in this setup $y(t)$ has same dimension as $u(t)$.

The simplification is to be achieved by finding a feedback transformation

$$v(t) = \beta(x(t)) + \alpha(x(t))u(t), \quad (13.6)$$

and a state transformation

$$z(t) = [z_l(t); z_0(t)] = \psi(x(t)), \quad (13.7)$$

where $\psi : X_0 \mapsto \mathbf{R}^n$, $\beta : X_0 \mapsto \mathbf{R}^m$, $\alpha : X_0 \mapsto \mathbf{R}^{m \times m}$ are continuously differentiable functions, such that the Jacobian of ψ is not singular on X_0 , and the relation between $v(t), y(t)$ and $z(t)$ subject to (13.6), (13.7) has the form

$$\dot{z}_l(t) = Az_l(t) + Bv(t), \quad y(t) = Cz_l(t), \quad (13.8)$$

$$\dot{z}_0(t) = a_0(z_l(t), z_0(t)), \quad (13.9)$$

where A, B, C are constant matrices of dimensions k -by- k , k -by- m , and m -by- k respectively, such that the pair (A, B) is controllable and the pair (C, A) is observable, and $a_0 : \mathbf{R}^k \times \mathbf{R}^{n-k} \mapsto \mathbf{R}^{n-k}$ is a continuously differentiable function.

More precisely, it is required that for every solution $x : [t_0, t_1] \mapsto X_0$, $u : [t_0, t_1] \mapsto \mathbf{R}^m$, $y : [t_0, t_1] \mapsto \mathbf{R}^m$ of (13.4), (13.5) equalities (13.8), (13.9) must be satisfied for $z(t), v(t)$ defined by (13.6) and (13.7).

As long as accurate measurements of the full state $x(t)$ of the original system are available, $X_0 = \mathbf{R}^n$, and the behavior of $y(t)$ and $u(t)$ is the only issue of interest, the output feedback linearization reduces the control problem to a linear one. However, in addition to sensor limitations, X_0 is rarely the whole \mathbf{R}^n , and the state $x(t)$ is typically required to remain bounded (or even to converge to a desired steady state value). Thus, it is frequently impossible to ignore equation (13.9), which is usually referred to as the *zero dynamics* of (13.4),(13.5). In the best scenario (the so-called “minimum phase systems”), the response of (13.9) to all expected initial conditions and reference signals $y(t)$ can be proven to be bounded and generating a response $x(t)$ confined to X_0 . In general, the area X_0 on which feedback linearization is possible does not cover of states of interest, the zero dynamics is not as stable as desired, and hence the benefits of output feedback linearization are limited.

13.1.3 Full state feedback linearization

Formally, full state feedback linearization applies to nonlinear ODE control system model of the form (13.4), without a need for a particular output $y(t)$ to be specified.

As in the previous subsection, the simplification is to be achieved by finding a feedback transformation (13.6) and a state transformation

$$z(t) = \psi(x(t)) \tag{13.10}$$

with a non-singular Jacobian. It is required that for every solution $x : [t_0, t_1] \mapsto X_0$, $u : [t_0, t_1] \mapsto \mathbf{R}^m$ of (13.4) equality

$$\dot{z}(t) = Az(t) + Bv(t) \tag{13.11}$$

must be satisfied for $z(t), v(t)$ defined by (13.6) and (13.10).

It appears that the benefits of having a full state linearization are substantially greater than those delivered by an output feedback linearization. Unfortunately, among systems of order higher than two, the full state feedback linearizable ones form a set of “zero measure”, in a certain sense. In other words, unlike in the case of output feedback linearization, which is possible, at least locally, “almost always”, full state feedback linearizability requires certain *equality constraints* to be satisfied for the original system data, and hence does not take place in a generic setup.

13.2 Feedback linearization with scalar control

This section contains basic results on feedback linearization of single-input systems (the case when $m = 1$ in (13.4)).

13.2.1 Relative degree and I/O feedback linearization

Assume that functions h, f, g in (13.4),(13.5) are at least $q + 1$ times continuously differentiable. We say that system (13.4),(13.5) has *relative degree* q on X_0 if

$$\nabla h_1(\bar{x})g(\bar{x}) = 0, \dots, \nabla h_{q-1}(\bar{x})g(\bar{x}) = 0, \quad \nabla h_q(\bar{x})g(\bar{x}) \neq 0 \quad \forall \bar{x} \in X_0,$$

where $h_i : X_0 \mapsto \mathbf{R}$ are defined by

$$h_1 = h, \quad h_{i+1} = (\nabla h_i)f \quad (i = 1, \dots, q).$$

By applying the definition to the LTI case $f(x) = Ax$, $g(x) = B$, $h(x) = Cx$ one can see that an LTI system with a non-zero transfer function always has a relative degree, which equals the difference between the degrees of numerator and denominator of its transfer function.

It turns out that systems with well defined relative degree are exactly those for which input/output feedback linearization is possible.

Theorem 13.1 *Assuming that h, f, g are continuously differentiable $n + 1$ times, the following conditions are equivalent:*

- (a) *system (13.4),(13.5) has relative degree q ;*
- (b) *system (13.4),(13.5) is input/output feedback linearizable.*

Moreover if conditions (a) is satisfied then

- (i) *the gradients $\nabla h_k(\bar{x})$ with $k = 1, \dots, q$ are linearly independent for every $\bar{x} \in X_0$ (which, in particular, implies that $q \leq n$);*

- (ii) *vectors $g_k(\bar{x})$ defined by*

$$g_1 = g, \quad g_{k+1} = [f, g_k] \quad (k = 1, \dots, q - 1)$$

satisfy

$$\nabla h_i(\bar{x})g_k(\bar{x}) = \nabla h_{i+j-1}(\bar{x})g(\bar{x}) \quad \forall \bar{x} \in X_0$$

for $i + j \leq q + 1$;

- (iii) *feedback linearization is possible with $k = q$,*

$$\alpha(\bar{x}) = \nabla h_q(\bar{x})g(\bar{x}), \quad \beta(\bar{x}) = \nabla h_q(\bar{x})f(\bar{x}),$$

$$\bar{z}_l = \psi_l(\bar{x}) = \begin{bmatrix} h_1(\bar{x}) \\ h_2(\bar{x}) \\ \vdots \\ h_q(\bar{x}) \end{bmatrix}.$$

Note that, unlike the Frobenius theorem, Theorem 13.1 is not local: it provides feedback linearization on every open set X_0 on which the relative degree is well defined. Also, in the case of *linear* models, where $f(x) = Ax$ and $g(x) = B$, it is always possible to get the zero dynamics depending on y only, i.e. to ensure that

$$a_0(z_l, z_0) = \bar{a}_0(Cz_l, z_0).$$

This, however, is not always possible in the nonlinear case. For example, for system

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ u \\ x_1 + x_2^2 \end{bmatrix}, \quad y = x_1$$

there exists no function $p : X_0 \mapsto \mathbf{R}$ defined on a non-empty open subset of \mathbf{R}^3 such that

$$\nabla p(x)f(x) = b(x_1, p(x)), \quad \nabla p(x)g(x) = 0, \quad \nabla p(x) \neq 0 \quad \forall x \in X_0.$$

Indeed, otherwise the system with new output $y_{new} = p(x)$ would have relative degree 3, which by Theorem 13.1 implies that $(\nabla p)g_1 = (\nabla p)g_2 = 0$, and hence by the Frobenius theorem the vector fields

$$g_1(x) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad g_2(x) = \begin{bmatrix} 1 \\ 0 \\ 2x_2 \end{bmatrix}$$

would define an involutive distribution, which they do not.

13.2.2 Involutivity and full state feedback linearization

It follows from Theorem 13.1 that system (13.4), (13.5) which has maximal possible relative degree n is full state feedback linearizable. The theorem also states that, given smooth functions f, g , existence of h defining a system with relative degree n implies linear independence of vectors $g_1(\bar{x}), \dots, \gamma_n(\bar{x})$ for all $\bar{x} \in X_0$, and involutivity of the regular distribution defined by vector fields g_1, \dots, g_{n-1} . The converse is also true, which allows one to state the following theorem.

Theorem 13.2 *Let $f : X_0 \mapsto \mathbf{R}^n$ and $g : X_0 \mapsto \mathbf{R}^n$ be $n + 1$ times continuously differentiable functions defined on an open subset X_0 of \mathbf{R}^n . Let g_k with $k = 1, \dots, n$ be defined as in Theorem 13.1.*

- (a) *If system (13.4) is full state feedback linearizable on X_0 then vectors $g_1(\bar{x}), \dots, \gamma_n(\bar{x})$ form a basis in \mathbf{R}^n for all $\bar{x} \in X_0$, and the distribution defined by vector fields g_1, \dots, g_{n-1} is involutive on X_0 .*
- (b) *If for some $\bar{x}_0 \in X_0$ vectors $g_1(\bar{x}), \dots, \gamma_n(\bar{x})$ form a basis in \mathbf{R}^n , and the distribution defined by vector fields g_1, \dots, g_{n-1} is involutive in a neighborhood of \bar{x}_0 , there exists an open subset \bar{X}_0 of X_0 such that $\bar{x}_0 \in \bar{X}_0$ and system (13.4) is full state feedback linearizable on \bar{X}_0 .*