ME 132, UC Berkeley Fall 2013

Nonlinear systems and linearization

1 State-space models and equilibrium points

2 Linearization

3 Example

15 Linearization

We have been learning about linear systems and controls for a while. In practice, many systems are nonlinear. Nonetheless, there are ways to use linear control techniques to handle nonlinear systems. This section shows one of such approaches.

15.1 State-space Representation of General Nonlinear Systems

Recall from Section 2.9, that a magnetically suspended ball can be modeled as

$$m\ddot{y} = mg - \frac{cu^2}{y^2} \tag{97}$$

Define the states: $x_1 = y, x_2 = \dot{y}$. We can have the state-space model:

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2, u) = x_2 \\ \dot{x}_2 &= f_2(x_1, x_2, u) = g - \frac{cu^2}{mx_1^2} = 10 - \left(\frac{u}{3.87x_1}\right)^2 \\ y &= h(x_1, x_2, u) = x_1 \end{aligned}$$

Similarly, for the pendulum model of

$$I\ddot{\theta}(t) = T_c - mglsin(\theta(t)),\ I = ml^2$$

we can model in the state space with $u = T_c$, $x_1(t) = \theta(t)$, and $x_2(t) = \dot{\theta}(t)$:

$$\begin{aligned} &\dot{x_1} = x_2 \\ &\dot{x_2} = -\frac{g}{l}sin(x_1) + \frac{1}{ml^2}u \\ &y = x_1 \end{aligned}$$

Both the above state-space models can be expressed as

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)) \\ y(t) &= h(x(t), u(t)) \end{aligned}$$

where $f(x, u)$ and $h(x, u)$ are some nonlinear functions of the state vector x and input vector u. Let us call this general nonlinear system $\mathcal{N}$.

15.2 Equilibrium Point and Linearization around an Equilibrium Point

Focus first on the state equation. Notice that if there exists some $\bar{x} \in \mathbf{R}^n, \bar{u} \in \mathbf{R}^m$ such that

$$f(\bar{x}, \bar{u}) = 0_n$$

then $\dot{\bar{x}} = 0$, namely, if we initialize the system with

$$\begin{aligned} x(0) &= \bar{x} \\ u(t) &= \bar{u}, \quad \forall t \geq 0 \end{aligned}$$

then the state will not move, such that

$$x(t) = \bar{x} \quad \forall t \geq 0, \; y(t) = \bar{y} \quad \forall t \geq 0$$

where

$$\bar{y} := h(\bar{x}, \bar{u}) \in \mathbf{R}^q$$

Such a pair of $\bar{x} \in \mathbf{R}^n, \bar{u} \in \mathbf{R}^m$ defines an *equilibrium point* of the nonlinear system, and the corresponding $\bar{y} := h(\bar{x}, \bar{u}) \in \mathbf{R}^q$ is the equilibrium output.

The remaining notes will explain that the behavior of a nonlinear system near an equilibrium point can be approximated by a linear system.

Behavior near an equilibrium point: Express the system dynamics in some new variables

$$\begin{aligned} \eta(t) &:= x(t) - \bar{x} \\ v(t) &:= u(t) - \bar{u} \\ w(t) &:= y(t) - \bar{y} \end{aligned}$$

which represent the offset from the equilibrium point. The word "near" can now be quantitatively explained as "(η, v) being small".

The dynamics of the states can now be expressed as

$$\begin{aligned} \dot{\eta}(t) &= \dot{x}(t) \\ &= f(\bar{x} + \eta(t), \bar{u} + v(t)) \end{aligned}$$

If $(\eta(t), v(t))$ are small, then the right hand side of the above equation is approximately (via first-order Taylor expansion):

$$f(\bar{x} + \eta(t), \bar{u} + v(t)) \approx f(\bar{x}, \bar{u}) + \left.\frac{\partial f}{\partial x}\right|_{\substack{x = \bar{x} \\ u = \bar{u}}} \eta(t) \quad + \quad \left.\frac{\partial f}{\partial u}\right|_{\substack{x = \bar{x} \\ u = \bar{u}}} v(t) \tag{98}$$

We can make the result a bit more compact by noticing that

$$f(\bar{x}, \bar{u}) = 0_n$$

and introducing

$$\left.\frac{\partial f}{\partial x}\right|_{\substack{x = \bar{x} \\ u = \bar{u}}} := A \in \mathbf{R}^{n \times n}, \quad \left.\frac{\partial f}{\partial u}\right|_{\substack{x = \bar{x} \\ u = \bar{u}}} := B \in \mathbf{R}^{n \times m}$$

This gives us the message that, while the deviations from equilibrium, namely, $(u(t) - \bar{u}, x(t) - \bar{x})$, both remain small, the deviations from equilibrium are approximately governed by the *linear* ODE

$$\dot{\eta}(t) = A\eta(t) + Bv(t)$$

The same can be done for the output equation

$$y(t) = h(\bar{x} + \eta(t), \bar{u} + v(t))$$

and we have

$$h(\bar{x} + \eta(t), \bar{u} + v(t)) \approx h(\bar{x}, \bar{u}) + \left.\frac{\partial h}{\partial x}\right|_{\substack{x = \bar{x} \\ u = \bar{u}}} \eta(t) \quad + \quad \left.\frac{\partial h}{\partial u}\right|_{\substack{x = \bar{x} \\ u = \bar{u}}} v(t) \tag{99}$$

Noticing $\bar{y} := h(\bar{x}, \bar{u})$ and letting

$$\left.\frac{\partial h}{\partial x}\right|_{\substack{x=\bar{x}\\ u=\bar{u}}} := C \in \mathbf{R}^{q\times n}, \quad \left.\frac{\partial h}{\partial u}\right|_{\substack{x=\bar{x}\\ u=\bar{u}}} := D \in \mathbf{R}^{q\times m}$$

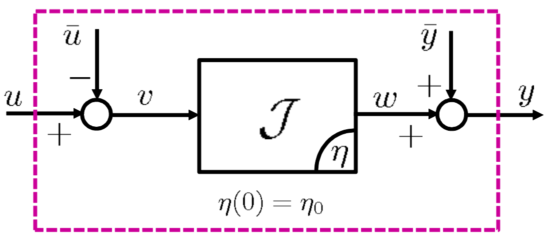
we get

$$y(t) - \bar{y} =: \quad w(t) = C\eta(t) + Dv(t)$$

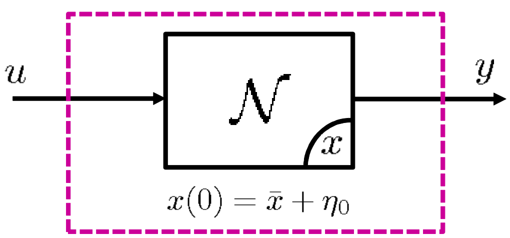
Summarizing the above, we have now derived a linear system $\mathcal{J}$:

$$\begin{aligned}\dot{\eta}(t) &= A\eta(t) + Bv(t)\\ w(t) &= C\eta(t) + Dv(t)\end{aligned}$$

with A, B, C, D as defined above, to approximate the nonlinear system around the equilibrium $\bar{x}$. In block diagrams, this means that the shifted linear system (input u, state η, output y)



behaves “approximately” like the appropriately initialized nonlinear system



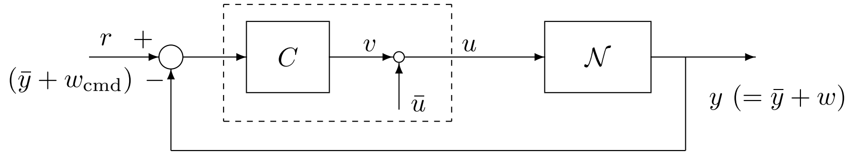
as long as the variables $(u(t) - \bar{u}, x(t) - \bar{x})$ remain small.

The linearized model above is called Jacobian linearization of the original nonlinear system.

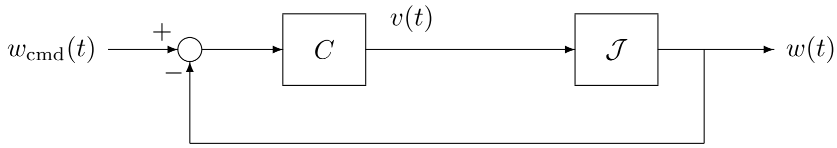
With such results, we can control the nonlinear system as follows:

1, find the equilibrium input and output $\bar{u}$ and $\bar{y}$ that define the equilibrium state $\bar{x}$ in the interested operation range of the system

2, decompose the reference signal as $y_{\text{desired}}(t) = \bar{y} + w_{\text{cmd}}(t)$ and the actual output as $y(t) = \bar{y} + w$. The control of the nonlinear system thus looks like



3, design the controller C as if we are to regulate the linearized system



4, implement the controller as shown in step 2

15.3 Multivariariate Partial Derivative

We learn in this section some details about obtaining the A, B, C, D matrices. This is just a generalization of the single-variable differentiation.

Let f be a function of x_1 and x_2. Pick a fixed vector $z = [z_1, z_2]^T$. For this two-variable function, the first-order Taylor expansion (around the point z) is

$$\begin{aligned} f(x) &\approx f(z) + \left.\frac{\partial f}{\partial x_1}\right|_{x=z} (x_1 - z_1) + \left.\frac{\partial f}{\partial x_2}\right|_{x=z} (x_2 - z_2) \quad \forall x \in \mathbf{R^2} \text{ close to } z \\ &= f(z) + \left[\begin{array}{cc} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{array}\right]_{x=z} \left[\begin{array}{c} x_1 - z_1 \\ x_2 - z_2 \end{array}\right] \\ &= f(z) + \nabla f(x)|_{x=z}^T (x - z) \end{aligned} \tag{100}$$

The term

$$\nabla f(x) = \left[\begin{array}{c} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{array}\right]$$

is called the gradient of $f(x)$, which is a generalization of $df(x)/dx$ in single-variable calculus. It is a 2 by 1 column vector if $f(x)$ is a mapping from $\mathbf{R^2}$ and $\mathbf{R}$. For instance, if $f(x_1, x_2) = x_1 + 2x_2$, then

$$\nabla f(x) = \left[\begin{array}{c} 1 \\ 2 \end{array}\right]$$

Note: by convention, the gradient of a $\mathbf{R^n} \to \mathbf{R}$ mapping $f(x_1, x_2, \ldots, x_n)$ is defined as a column vector:

$$\nabla f(x) = \left[\begin{array}{c} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{array}\right]$$

as it defines a direction in a vector space. A corresponding definition is the derivative

$$Df(x) = \left[\begin{array}{cccc} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \ldots & \frac{\partial f}{\partial n} \end{array}\right]$$

which is a row vector and

$$\nabla f(x) = [Df(x)]^T$$

We can generalize the above result. For instance, if $f_1(x,u) = f_1(x_1, x_2, u_1, u_2)$ then the Taylor approximation around the point $(\bar{x}, \bar{u})$ is

$$f_1(x,u) \approx f_1(\bar{x},\bar{u}) + \left[\frac{\partial f_1(x,u)}{\partial x_1}, \frac{\partial f_1(x,u)}{\partial x_2}\right]\Bigg|_{\substack{x=\bar{x}\\u=\bar{u}}} \begin{bmatrix} x_1-\bar{x}_1 \\ x_2-\bar{x}_2 \end{bmatrix} + \left[\frac{\partial f_1(x,u)}{\partial u_1}, \frac{\partial f_1(x,u)}{\partial u_2}\right]\Bigg|_{\substack{x=\bar{x}\\u=\bar{u}}} \begin{bmatrix} u_1-\bar{u}_1 \\ u_2-\bar{u}_2 \end{bmatrix}$$
$$= f_1(\bar{x},\bar{u}) + \nabla_x^T f_1(x,u)\big|_{\substack{x=\bar{x}\\u=\bar{u}}} (x-\bar{x}) + \nabla_u^T f_1(x,u)\big|_{\substack{x=\bar{x}\\u=\bar{u}}} (u-\bar{u})$$

If we have another similar function

$$f_2(x,u) \approx f_2(\bar{x},\bar{u}) + \left[\frac{\partial f_2(x,u)}{\partial x_1}, \frac{\partial f_2(x,u)}{\partial x_2}\right]\Bigg|_{\substack{x=\bar{x}\\u=\bar{u}}} \begin{bmatrix} x_1-\bar{x}_1 \\ x_2-\bar{x}_2 \end{bmatrix} + \left[\frac{\partial f_2(x,u)}{\partial u_1}, \frac{\partial f_2(x,u)}{\partial u_2}\right]\Bigg|_{\substack{x=\bar{x}\\u=\bar{u}}} \begin{bmatrix} u_1-\bar{u}_1 \\ u_2-\bar{u}_2 \end{bmatrix}$$
$$= f_2(\bar{x},\bar{u}) + \nabla_x^T f_2(x,u)\big|_{\substack{x=\bar{x}\\u=\bar{u}}} (x-\bar{x}) + \nabla_u^T f_2(x,u)\big|_{\substack{x=\bar{x}\\u=\bar{u}}} (u-\bar{u})$$

Then for the $\mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}^2$ function

$$f(x,u) = \begin{bmatrix} f_1(x,u) \\ f_2(x,u) \end{bmatrix}$$

we have

$$f(x,u) \approx \begin{bmatrix} f_1(\bar{x},\bar{u}) \\ f_2(\bar{x},\bar{u}) \end{bmatrix} + \underbrace{\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix}}_{\nabla_x^T f(x,u)}\Bigg|_{\substack{x=\bar{x}\\u=\bar{u}}} \begin{bmatrix} x_1-\bar{x}_1 \\ x_2-\bar{x}_2 \end{bmatrix} + \underbrace{\begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} \end{bmatrix}}_{\nabla_u^T f(x,u)}\Bigg|_{\substack{x=\bar{x}\\u=\bar{u}}} \begin{bmatrix} u_1-\bar{u}_1 \\ u_2-\bar{u}_2 \end{bmatrix}$$
$$= f(\bar{x},\bar{u}) + A(x-\bar{x}) + B(u-\bar{u})$$

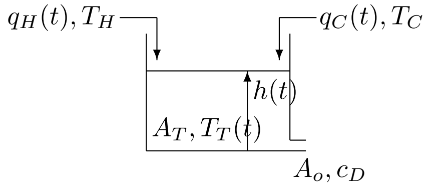
From here we learnt how to compute the derivative and gradient of a multi-input multi-output function:

$$D_x \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix}, \ \nabla_x \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_1} \\ \frac{\partial f_1}{\partial x_2} & \frac{\partial f_2}{\partial x_2} \end{bmatrix}$$

With the above results, we can writing down (98) and (99).

15.4 Example: Tank System

Consider the following water tank system:



There are hot and cold water supplies (at fixed temperatures T_C and T_H) going into the tank. We can control the hot and cold inflows q_C and q_H (in the unit of m^3/sec).

The orifice outflow is related to its area A_o and the discharge coefficient c_D. Torricelli's law states that the speed of a fluid through a sharp-edged hole under the force of gravity is the same as the speed that a body would acquire in falling freely from a height h, i.e. $v_{out}(t) = \sqrt{2gh(t)}$, where g is the acceleration due to gravity. Hence the outflow is

$$q_{out}(t) = c_D A_o \sqrt{2gh(t)}$$

Suppose the water supplies are instantaneously mixed in the tank, and the temperature of water in tank is T_T after mixing. By using the conservation law we can state that

$$\begin{aligned} \dot{h}(t) &= \tfrac{1}{A_T}\left(q_C(t) + q_H(t) - c_D A_o \sqrt{2gh(t)}\right) \\ \dot{T}_T(t) &= \tfrac{1}{h(t)A_T}\left(q_C(t)\left[T_C - T_T(t)\right] + q_H(t)\left[T_H - T_T(t)\right]\right) \end{aligned}$$

Nonlinear System and Linearization Define state and input vectors as

$$x(t) := \begin{bmatrix} h(t) \\ T_T(t) \end{bmatrix} \quad , \quad u(t) := \begin{bmatrix} q_C(t) \\ q_H(t) \end{bmatrix}$$

Then, with

$$f_1(x_1, x_2, u_1, u_2) = \frac{1}{A_T}\left(u_1 + u_2 - c_D A_o \sqrt{2gx_1}\right)$$

$$f_2(x_1, x_2, u_1, u_2) = \frac{1}{x_1 A_T}\left(u_1\left[T_C - x_2\right] + u_2\left[T_H - x_2\right]\right)$$

the dynamic equations of the system are of the form

$$\begin{aligned} \dot{x}_1(t) &= f_1(x(t), u(t)) \\ \dot{x}_2(t) &= f_2(x(t), u(t)) \end{aligned}$$

Equilibrium Points Equilibrium points are characterized by $f(\bar{x}, \bar{u}) = 0$. In this case, with

$$f_1(x_1, x_2, u_1, u_2) = \frac{1}{A_T}\left(u_1 + u_2 - c_D A_o \sqrt{2gx_1}\right)$$

$$f_2(x_1, x_2, u_1, u_2) = \frac{1}{x_1 A_T}\left(u_1\left[T_C - x_2\right] + u_2\left[T_H - x_2\right]\right)$$

Writing with barred-quantities, and setting to 0 gives (assuming $\bar{x}_1 \neq 0$)

$$\begin{bmatrix} 1 & 1 \\ T_C - \bar{x}_2 & T_H - \bar{x}_2 \end{bmatrix} \begin{bmatrix} \bar{u}_1 \\ \bar{u}_2 \end{bmatrix} = \begin{bmatrix} c_D A_o \sqrt{2g\bar{x}_1} \\ 0 \end{bmatrix}$$

The matrix on the left hand side of the equation is invertible if and only if $\Leftrightarrow T_C \neq T_H$. Thus for any choice of $\bar{x}$, there is a unique equilibrium input $\bar{u}$, given by

$$\begin{bmatrix} \bar{u}_1 \\ \bar{u}_2 \end{bmatrix} = \frac{1}{T_H - T_C} \begin{bmatrix} T_H - \bar{x}_2 & -1 \\ \bar{x}_2 - T_C & 1 \end{bmatrix} \begin{bmatrix} c_D A_o \sqrt{2g\bar{x}_1} \\ 0 \end{bmatrix}$$

which gives

$$\bar{u}_1 = \frac{c_D A_o \sqrt{2g\bar{x}_1}\,(T_H - \bar{x}_2)}{T_H - T_C} \quad , \quad \bar{u}_2 = \frac{c_D A_o \sqrt{2g\bar{x}_1}\,(\bar{x}_2 - T_C)}{T_H - T_C}$$

Since u_1 and $u_2 x$ represent flow rates **into** the tank, physical considerations restrict them to be nonegative real numbers. This implies that $\bar{x}_1 \geq 0$ and $T_C \leq \bar{T}_T \leq T_H$. Looking at the differential equation for T_T, we see that its rate of change is inversely related to h. Hence, the differential equation model is valid while $h(t) > 0$, so we further restrict $x_1 > 0$. Under those restrictions, the state $\bar{x}$ is indeed an equilibrium point, and there is a unique equilibrium input given by the equations above.

Partial Derivatives

Obtain the "A" and "B" matrices, by first taking partial derivatives

$$D_x f(x,u) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} -\frac{g c_D A_o}{A_T \sqrt{2g x_1}} & 0 \\ -\frac{u_1[T_C - x_2] + u_2[T_H - x_2]}{x_1^2 A_T} & \frac{-(u_1+u_2)}{x_1 A_T} \end{bmatrix}$$

$$D_u f(x,u) = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{A_T} & \frac{1}{A_T} \\ \frac{T_C - x_2}{x_1 A_T} & \frac{T_H - x_2}{x_1 A_T} \end{bmatrix}$$

We can do some numerical examples by assuming

$$T_C = 10^\circ, T_H = 90^\circ, A_T = 3\mathtt{m}^2, A_o = 0.05\mathtt{m}, c_D = 0.7$$

We can compute, for instance, linearization at 4 different equilibrium points

$$(\bar{h} = 1\mathtt{m}, \bar{T}_T = 25^\circ),\ (\bar{h} = 3\mathtt{m}, \bar{T}_T = 25^\circ),\ (\bar{h} = 1\mathtt{m}, \bar{T}_T = 75^\circ),\ (\bar{h} = 3\mathtt{m}, \bar{T}_T = 75^\circ)$$

The results for $\left(\bar{h}, \bar{T}_T\right) = (1\mathtt{m}, 25^\circ)$ are as follows

$$\bar{u}_1 = \bar{q}_C = 0.126 \quad , \quad \bar{u}_2 = \bar{q}_H = 0.029$$

$$A = \begin{bmatrix} -0.0258 & 0 \\ 0 & -0.517 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 0.333 & 0.333 \\ -5.00 & 21.67 \end{bmatrix}$$