UNIVERSIDADE FEDERAL DO CEARÁ - UFC
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA

LINEAR SYSTEM THEORY

ITALO AGUIAR DO NASCIMENTO PAULINO

Fortaleza - CE
2018

Contents

1 Introduction

2 Mathematical Descriptions of Systems

2.1 Causality and Lumpedness

In the studie of linear systems it's necessary to have some main notion of it's caracteristcs:

A system is called a causal or nonanticipatory system if it's current output depends on past and current inputs, but not on future input. If a system is not causal, then it's current output will depend on future input, so it's called noncausal or anticipatory system, cause it can predict what will be applied int he future.

A system is said to be lumped if it's number of state variables is finite or it's state is a finite vector. A system is called a distributed system if it's state has infinitely many state variables.

2.2 Linear Systems

A system is called a linear system if for every t_0 and any two state-input-output pairs:

$$\left.\begin{array}{l}\boldsymbol{x_i}(t_0)\\ \boldsymbol{u_i}(t), t \geq t_0\end{array}\right\} \longrightarrow \boldsymbol{y_i}(t), t \geq t_0 \tag{1}$$

for $i = 1, 2, 3, ..., n$, where: $x_i(t_0)$ is considered the initial state of the system analyzed and $u_i(t)$, $y_i i(t)$ are the input and output of the system at any time after t_0 , we can demonstrate the following two properties:

1. Additivity:

$$\left.\begin{array}{l}\boldsymbol{x_1}(t_0) + \boldsymbol{x_2}(t_0) + ... + \boldsymbol{x_n}(t_0)\\ \boldsymbol{u_1}(t) + \boldsymbol{u_2}(t) + ... + \boldsymbol{u_n}(t), t \geq t_0\end{array}\right\} \longrightarrow \boldsymbol{y_1}(t) + \boldsymbol{y_2}(t) + ... + \boldsymbol{y_n}(t), t \geq t_0 \tag{2}$$

2. Homogeneity:

$$\left.\begin{array}{l}\boldsymbol{\alpha x_i}(t_0)\\ \boldsymbol{\alpha u_i}(t), t \geq t_0\end{array}\right\} \longrightarrow \boldsymbol{\alpha} y_i(t), t \geq t_0 \tag{3}$$

for $i = 1, 2, 3, ..., n$.

Those two properties can be combined resulting in the superposition ppropertу:

$$\left.\begin{array}{l}\boldsymbol{\alpha_1 x_1}(t_0) + \boldsymbol{\alpha_2 x_2}(t_0) + ... + \boldsymbol{\alpha_n x_n}(t_0)\\ \boldsymbol{\alpha_1 u_1}(t) + \boldsymbol{\alpha_2 u_2}(t) + ... + \boldsymbol{\alpha_n u_n}(t), t \geq t_0\end{array}\right\} \longrightarrow \boldsymbol{\alpha_1 y_1}(t) + \boldsymbol{\alpha_2 y_2}(t) + ... + \boldsymbol{\alpha_n y_n}(t), t \geq t_0 \tag{4}$$

The Zero-input response of a system is the name given to the output excited exclusively by the initial state $x(t_0)$, it is the input u(t) is zero for $t \geq t_0$:

$$\left.\begin{array}{l}\boldsymbol{x}(t_0)\\ \boldsymbol{u}(t) \equiv 0, t \geq t_0\end{array}\right\} \longrightarrow \boldsymbol{y_{zi}}(t) \tag{5}$$

The Zero-state response of a system is the name given to the output excited exclusively by the input $u(t)$ for $t \geq t_0$, it is the input $x_(t_0)$ is zero:

$$\left.\begin{array}{l}\boldsymbol{x}(t_0) = 0 \\ \boldsymbol{u}(t), t \geq t_0\end{array}\right\} \longrightarrow \boldsymbol{y_{zs}}(t) \tag{6}$$

The response of every linear system can be decomposed into the zero-state response and the zero-input response by the superposition principle and that's how we usually analyze it;

Input-output description: Considering a system which initial state is zero, considering it's output being excited by an input u(t), we described mathematically as:

$$y(t) \approx \sum_i g_\Delta(t, t_i)u(t_i)$$

Where $u(t) \approx \sum_i u(t_i)\delta_\Delta(t - t_i)\Delta$ and $g_\Delta(t, t_i)$ is the output at time t excited by the pulse $u(t) = \delta_Delta(t - t_i)$ applied at time t_i.

If Δ aproaches zero, the pulse $\delta_Delta(t-t_i)$ becomes an impulse at t_i. denoted by $\delta(t-t_i)$, and the corresponding outuput will be denoted by $g(t, t_i)$. Thus, the approximation seen in the summation becomes an equality turning into an integration.

$$y(t) = \int_{-\infty}^{\infty} g(t, \tau)u(\tau)d\tau$$

Since $g(t, \tau)$ is the response excited by an impulse, it is called the impulse response.

A system is said to be relaxed at t_0 if it's initial sate at t_0 is 0. If the system is causal, then $g(t, \tau) = 0$ for $t < \tau$. Every linear system that is causal and relaxed at t_0 can be described by:

$$y(t) = \int_{t_0}^{t} g(t, \tau)u(\tau)d\tau$$

Thus, if a linear system has p input terminals and q output terminals, then we can extend 2.2 to:

$$y(t) = \int_{t_0}^{t} G(t, \tau)u(\tau)d\tau$$

where:

$$G(t, \tau) = \begin{bmatrix} \hat{g_{11}}(s) & \hat{g_{12}}(s) & \hat{g_{13}}(s) & . & . & . & \hat{g_{1p}}(s) \\ \hat{g_{21}}(s) & \hat{g_{22}}(s) & \hat{g_{23}}(s) & . & . & . & \hat{g_{2p}}(s) \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ \hat{g_{q1}}(s) & \hat{g_{q2}}(s) & \hat{g_{q3}}(s) & . & . & . & \hat{g_{qp}}(s) \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ . \\ . \\ . \\ \hat{u}_p \end{bmatrix}$$

2.3 Linear Time-Invariant (LTI) Systems

A system is said to be time invariant if for every state-input-output pair:

$$\left.\begin{array}{l}\boldsymbol{x}(t_0)\\ \boldsymbol{u}(t), t \geq t_0\end{array}\right\} \longrightarrow \boldsymbol{y(t)}, t \geq t_0 \tag{7}$$

and any T, we have:

$$\left.\begin{array}{l}\boldsymbol{x}(t_0+T)\\ \boldsymbol{u}(t-T), t \geq t_0+T\end{array}\right\} \longrightarrow \boldsymbol{y(t-T)}, t \geq t_0+T \tag{8}$$

Input-output description: The zero-state response of a linear system can be described as $y(t) = \int_{t_0}^{t} g(t,\tau)u(\tau)d\tau$. If the system is time invariant as well, then we have:

$$g(t,\tau) = g(t+T, \tau+T) = g(t-\tau, 0) = g(t-\tau)$$

for any T. Then, we can describe a system as:

$$y(t) = \int_0^t g(t-\tau)u(\tau)d\tau = \int_0^t g(\tau)u(t-\tau)d\tau$$

Transfer-function matrix: In Laplace domain, we know that the convolution of two function may be transformed in the multiplication of these two functions in Laplace domain:

$$\hat{y}(s) = \hat{g}(s)\hat{u}(s)$$

where:

$$\hat{g}(s) = \int_0^\infty g(t)e^{-st}dt$$

is called the transfer function of the system.

For p-input and q-output system, it can be extended as:

$$\begin{bmatrix}\hat{y_1}(s)\\ \hat{y_2}(s)\\ .\\ .\\ .\\ \hat{y_q}(s)\end{bmatrix} = \begin{bmatrix}\hat{g_{11}}(s) & \hat{g_{12}}(s) & \hat{g_{13}}(s) & . & . & . & \hat{g_{1p}}(s)\\ \hat{g_{21}}(s) & \hat{g_{22}}(s) & \hat{g_{23}}(s) & . & . & . & \hat{g_{2p}}(s)\\ . & . & . & . & . & . & .\\ . & . & . & . & . & . & .\\ . & . & . & . & . & . & .\\ \hat{g_{q1}}(s) & \hat{g_{q2}}(s) & \hat{g_{q3}}(s) & . & . & . & \hat{g_{qp}}(s)\end{bmatrix} \begin{bmatrix}\hat{u}_1\\ \hat{u}_2\\ .\\ .\\ .\\ \hat{u}_p\end{bmatrix}$$

$$\hat{y}(s = \hat{G}(s)\hat{u}(s)$$

where $\hat{g}_{ij}(s)$ is the transfer function from the jth input to the ith output. $\hat{G}(s)$ is called the transfer-function matrix or, simply, transfer matrix of the system.

State-space equation: Every linear time-invariant lumped system can be described by a set of the form:

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t)$$

For a system with p inputs,q outputs and n state variables, A, B, C, and D are, respectively, nxn, nxp, qxn and qxp constant matrices. Applying the Laplace transform to it, we will get:

$$\begin{aligned} s\hat{x}(s) - x(0) &= A\hat{x}(s) + B\hat{u}(s) \\ \hat{y}(s) &= C\hat{x}(s) + D\hat{u}(s) \end{aligned} \tag{9}$$

wich implies

$$\begin{aligned} \hat{x}(s) &= (sI - A)^{-1} + (sI - A)^{-1}B\hat{u}(s) \\ \hat{y}(s) &= C(sI - A)^{-1}x(0) + C(sI - A)^{-1}B\hat{u}(s) + D\hat{u}(s) \end{aligned} \tag{10}$$

Considering the initial state x(0) as zero, then:

$$\hat{y}(s) = [C(sI - A)^{-1} + D]\hat{u}(s)$$

where:

$$\hat{G}(s) = C(sI - A)^{-1}B + D$$

This relates input-output and state-space descriptions.

2.4 Lineariztion

Dealing with a system described as:

$$\begin{aligned} \dot{x}(t) &= h(x(t), u(t), t) \\ y(t) &= f(x(t), u(t), t) \end{aligned} \tag{11}$$

where h and f are nonlinear functions. Some nonlinear equations, however, can be approximated by linear equations under certain conditions. Suppose for input function $u_0(t)$ and some initial state, $x_0(t)$ is the solution of 11, $\dot{x}_0(t) = h(x_0(t), u_0(t), t)$. If the system's input is perturbed slightly to become $u_0(t) + \tilde{u}(t)$ and the initial state is perturbed too. In this case, we get:

$$\dot{x}_0(t) + \dot{\bar{x}}(t) = h(x_0(t) + \bar{x}(t), u_0(t) + \tilde{u}(t), t)$$
$$\dot{x}_0(t) + \dot{\bar{x}}(t) = h(x_0(t), u_0(t), t) + \frac{\partial h}{\partial x}\bar{x} + \frac{\partial h}{\partial u}\bar{u} + ...$$

where $h = [h_1 h_2 h_3]'$, $x = [x_1 x_2 x_3]'$ and $u = [u_1 u_2]'$:

$$A(t) = \frac{\partial h}{\partial x} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \frac{\partial h_1}{\partial x_3} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \frac{\partial h_2}{\partial x_3} \\ \frac{\partial h_3}{\partial x_1} & \frac{\partial h_3}{\partial x_2} & \frac{\partial h_3}{\partial x_3} \end{bmatrix}$$

$$B(t) = \frac{\partial h}{\partial u} = \begin{bmatrix} \frac{\partial h_1}{\partial u_1} & \frac{\partial h_1}{\partial u_2} \\ \frac{\partial h_2}{\partial u_1} & \frac{\partial h_2}{\partial u_2} \end{bmatrix}$$

Using it we can find C and D matrices too and, then, we are able to desbribe the system in the state-space equation form.

2.5 Example

For examplify all the subject of this section and add some informations, let us model a Three tank system:

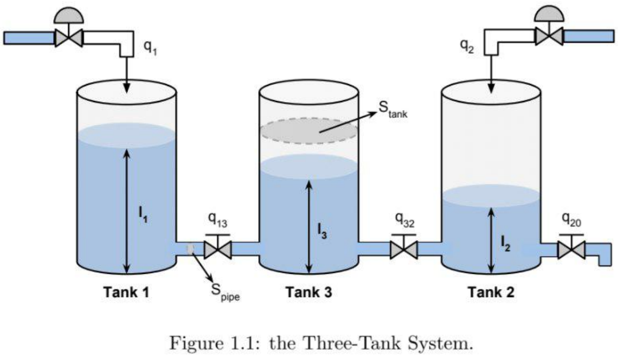


Figure 1: Three Tank System

It's known that thee flow rate is given by:

$$\begin{aligned} q_{13} &= \mu_{13} * s * \sqrt{2 * g * |I_1 - I_3|} * sig(|I_1 - I_3|) \\ q_{32} &= \mu_{32} * s * \sqrt{2 * g * |I_3 - I_2|} * sig(|I_3 - I_2|) \\ q_{20} &= \mu_0 * s * \sqrt{2 * g * |I_2|} \end{aligned}$$

Where:

- μ: Outflow coefficient;
- s: Pipe cross-sectional area;
- g: Gravity value;
- I: Liquid level;
- sig(): Signal function;

System Variable	Symbol	Value	Unit
Tanks cross-sectional area	S_{tank}	0.154	m^2
Pipes cross-sectional area	S_{pipe}	1×10^{-4}	m^2
Outflow coefficient	$\mu = \mu_{13} = \mu_{32}$	0.65	–
Nominal outflow coefficient	μ_0	0.825	–
Maximum supply flow-rate	$q_i^{\max}$ $(i = 1, 2)$	2.4×10^{-4}	$m^3 s^{-1}$
Maximum liquid level	$l_i^{\max}$ $(i = 1, 2, 3)$	0.6	m

Table 1.1: system constant variables table.

Figure 2: Data

Thus, it's possible to describe physically the relation between of liquid flow of the three tanks:

$$S\frac{dI_1}{dt} = q_1 - q_{13} = q_1 - \mu_{13} * s * \sqrt{2 * g * |I_1 - I_3|} * sig(|I_1 - I_3|)$$

$$S\frac{dI_2}{dt} = q_2 - q_{20} + q_{23} = q_2 - \mu_0 * s * \sqrt{2 * g * |I_2|} + \mu_{32} * s * \sqrt{2 * g * |I_3 - I_2|} * sig(|I_3 - I_2|)$$

$$S\frac{dI_3}{dt} = q_{13} - q_{32} = \mu_{13} * s * \sqrt{2 * g * |I_1 - I_3|} * sig(|I_1 - I_3|) - \mu_{32} * s * \sqrt{2 * g * |I_3 - I_2|} * sig(|I_3 - I_2|)$$

where:

- S: Tank cross-sectional area;
- q_1: Suplly flow-rate on tank 1 (Input);
- q_2: Suplly flow-rate on tank 2 (Input);
- q_20 Outflow-rate (output);

Now, it's possible to see that we are not dealing with a linear system. Thus, we need first to find a stationary state. It's possible to do it by calculating the state-values in the point where the variation is zero.

For $\frac{dI}{dt} = 0$:

$$\frac{dI_1}{dt} = \frac{q_1}{S} - \frac{\mu_{13} * s * \sqrt{2 * g * |I_1 - I_3|} * sig(|I_1 - I_3|)}{S} = 0$$

$$q_1 = \mu_{13} * s * \sqrt{2 * g * |I_1 - I_3|} * sig(|I_1 - I_3|) \rightarrow |I_1 - I_3| = \frac{{q_1}^2}{{\mu_{13}}^2 * s^2 * 2 * g}$$

$$\frac{dI_3}{dt} = \frac{\mu_{13} * s * \sqrt{2 * g * |I_1 - I_3|} * sig(|I_1 - I_3|)}{S} - \frac{\mu_{32} * s * \sqrt{2 * g * |I_3 - I_2|} * sig(|I_3 - I_2|)}{S} = 0$$

$\mu_{13}{}^2 * s^2 * 2 * g * |I_1 - I_3| = \mu_{32}{}^2 * s^2 * 2 * g * |I_3 - I_2|$, and it's know that the outflow coefficient are all equal, then:

$|I_1 - I_3| = |I_3 - I_2|$

$$\frac{dI_2}{dt} = \frac{q_2}{S} - \frac{\mu_0 * s * \sqrt{2 * g * |I_2|}}{S} + \frac{\mu_{32} * s * \sqrt{2 * g * |I_3 - I_2|} * sig(|I_3 - I_2|)}{S}$$

$\mu_0 * s * \sqrt{2 * g * I_2} = q_2 - \mu_{32} * s * \sqrt{2 * g * |I_3 - I_2|} * sig(|I_3 - I_2|)$

$\mu_0{}^2 * s^2 * 2 * g * I_2 = q_2{}^2 - 2 * q_2 * \mu_{32} * s * \sqrt{2 * g * |I_3 - I_2|} * sig(I_3 - I_2) + \mu_{32} * s^2 * 2 * g * |I_3 - I_2|$

Since $\mu_{32} * s^2 * 2 * g * |I_3 - I_2| = \mu_{13} * s^2 * 2 * g * |I_1 - I_3|$, we can substitute it and get:

$\mu_0{}^2 * s^2 * 2 * g * I_2 = q_2{}^2 - 2 * q_2 * \mu_{32} * s * \sqrt{2 * g * |I_3 - I_2|} + q_1{}^2$

$\mu_0{}^2 * s^2 * 2 * g * I_2 = (q_2 - q_1 * sig(I_3 - I_2))^2$, leading to:

$I_2 = \dfrac{(q_2 - q_1 * sig(I_3 - I_2))^2}{2 * g * \mu_0{}^2 * s^2}$

From these and be using the values on 3 (assuming the supply flow-rate equals to $q_1 = 0.40 * 10^{-4}$ and $q_2 = 0.30 * 10^{-4}$), we may get to: (Adopting $I_1 > I_3 > I_2$)

(i) $I_2 = \dfrac{(q_2 - q_1 * sig(I_3 - I_2))^2}{2 * g * \mu_0{}^2 * s^2} = 7.4884 * 10^{-4}$

(ii) $|I_1 - I_3| = \dfrac{q_1{}^2}{\mu_{13}{}^2 * s^2 * 2 * g} \rightarrow |I_1 - I_3| = 1.9302 * 10^{-2}$.

(iii) $|I_1 - I_3| = |I_3 - I_2| = 1.9302 * 10^{-2}$, applying (i), $I_3 = 2.0051 * 10^{-2}$.

(iv) For $|I_1 - I_3| = 1.9302 * 10^{-2}$, then $I_1 = 3.9353 * 10^{-2}$.

Now, we are able to linearize using the stationay values:

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 3.9353 * 10^{-2} \\ 7.4884 * 10^{-4} \\ 2.0051 * 10^{-2} \end{bmatrix}$$

Therefore, we can rewrite our equations without module and signal functions:

$$S\frac{dI_1}{dt} = q_1 - \mu_{13} * s * \sqrt{2 * g * (I_1 - I_3)}$$

$$S\frac{dI_2}{dt} = q_2 - \mu_0 * s * \sqrt{2 * g * I_2} + \mu_{32} * s * \sqrt{2 * g * (I_3 - I_2)}$$

$$S\frac{dI_3}{dt} = \mu_{13} * s * \sqrt{2 * g * (I_1 - I_3)} - \mu_{32} * s * \sqrt{2 * g * (I_3 - I_2)}$$

Now, we can make our matrix of partial derivative:

$$\text{S}\frac{\partial I_1/dt}{\partial I_1} = -\mu_{13}*s*\frac{1}{2}*\sqrt{2*g}*\frac{1}{\sqrt{(I_1-I_3)}} = -1.0362*10^{-3} \rightarrow \frac{\partial I_1/dt}{\partial I_1} = -6.7284*10^{-3}$$

$$\text{S}\frac{\partial I_1/dt}{\partial I_2} = 0$$

$$\text{S}\frac{\partial I_1/dt}{\partial I_3} = \mu_{13}*s*\frac{1}{2}*\sqrt{2*g}*\frac{1}{\sqrt{(I_1-I_3)}} = 1.0362*10^{-3} \rightarrow \frac{\partial I_1/dt}{\partial I_3} = 6.7284*10^{-3}$$

$$\text{S}\frac{\partial I_2/dt}{\partial I_1} = 0$$

$$\text{S}\frac{\partial I_2/dt}{\partial I_2} = -\mu_0*s*\sqrt{2*g}*\frac{1}{2}*\frac{1}{\sqrt{I_2}} - \mu_{32}*s*\sqrt{2*g}*\frac{1}{\sqrt{I_3-I_2}} = -2.0723*10^{-3} \rightarrow$$
$$\frac{\partial I_2/dt}{\partial I_2} = -2.99*10^{-2}$$

$$\text{S}\frac{\partial I_2/dt}{\partial I_3} = \mu_{32}*s*\sqrt{2*g}*\frac{1}{\sqrt{I_3-I_2}} = 2.0733*10^{-3} \rightarrow \frac{\partial I_2/dt}{\partial I_3} = 1.3458*10^{-2}$$

$$\text{S}\frac{\partial I_3/dt}{\partial I_1} = \mu_{13}*s*\sqrt{2*g}*\frac{1}{2}*\frac{1}{\sqrt{I_1-I_3}} = 1.0362*10^{-3} \rightarrow \frac{\partial I_3/dt}{\partial I_1} = 6.7284*10^{-3}$$

$$\text{S}\frac{\partial I_3/dt}{\partial I_2} = \mu_{32}*s*\sqrt{2*g}*\frac{1}{2}*\frac{1}{\sqrt{I_3-I_2}} = 1.0362*10^{-3} \rightarrow \frac{\partial I_3/dt}{\partial I_2} = 6.7284*10^{-3}$$

$$\text{S}\frac{\partial I_3/dt}{\partial I_3} = -\mu_{13}*s*\sqrt{2*g}*\frac{1}{2}*\frac{1}{\sqrt{I_1-I_3}} - \mu_{32}*s*\sqrt{2*g}*\frac{1}{2}*\frac{1}{\sqrt{I_3-I_2}} = -2.0724*10^{-3} \rightarrow$$
$$\frac{\partial I_3/dt}{\partial I_3} = -1.3457*10^{-2}$$

Therefore, we have a linear description of our system in the state-space model:

$$A(t) = \frac{\partial h}{\partial x} = \begin{bmatrix} -6.7284*10^{-3} & 0 & 6.7284*10^{-3} \\ 0 & -5.0085*10^{-2} & 1.3458*10^{-2} \\ 6.7284*10^{-3} & 6.7284*10^{-3} & -1.3457*10^{-2} \end{bmatrix}$$

$$B(t) = \frac{\partial h}{\partial u} = \begin{bmatrix} 6.4935 & 0 \\ 0 & 6.4935 \\ 0 & 0 \end{bmatrix}$$

Thus, we may describe our system as:

$$\begin{bmatrix} \frac{dI_1}{dt} \\ \frac{dI_2}{dt} \\ \frac{dI_3}{dt} \end{bmatrix} = \begin{bmatrix} -6.7284 * 10^{-3} & 0 & 6.7284 * 10^{-3} \\ 0 & -3.663 * 10^{-2} & -6.7284 * 10^{-3} \\ 6.7284 * 10^{-3} & 6.7284 * 10^{-3} & -1.3457 * 10^{-2} \end{bmatrix} \begin{bmatrix} \Delta I_1 \\ \Delta I_2 \\ \Delta I_3 \end{bmatrix} + \begin{bmatrix} 6.4935 & 0 \\ 0 & 6.4935 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta q_1 \\ \Delta q_2 \end{bmatrix}$$

On Matlab, we can simulate the system using the following script:

```
A=[-6.7284*10^(-3) 0 6.7284*10^(-3); 0 -2.99 *10^(-2) 1.3458*10^(-2) ;6.7284*10^(-3) 6.7284*10^(-3) -1.3457*10^(-2) ]
B=[6.4935 0;0 6.4935;0 0 ]
C=[1 0 0 ; 0 1 0; 0 0 1]re
D=[0 0 ; 0 0 ; 0 0]
u = [ones(1,10000) * 1.2e-4; ones(1,10000) * 0.6e-4];
ue=[0.4*10^(-4); 0.3*10^(-4)]
xe=[3.9353*10^(-2) 7.4884*10^(-4) 2.0051*10^(-2)]
tank=ss(A,B,C,D)
t=linspace(0,5000,10000);
y = lsim(tank, u - ue, t,xe);
plot(t, y+xe)
```

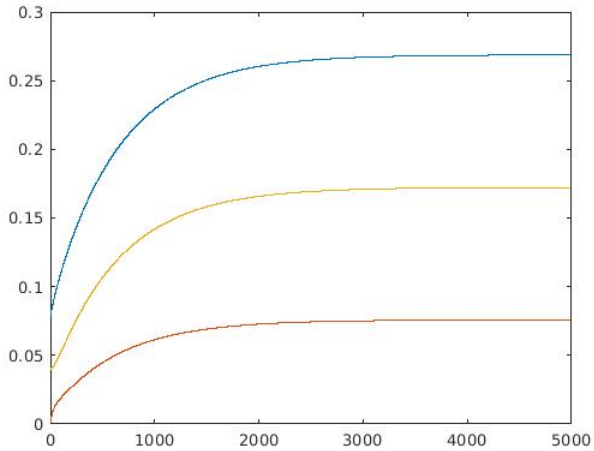


Figure 3: Data

It's also possible to calculate the transfer function of our system by

$$\hat{G}(s) = C(sI - A)^{-1}B + D$$

, where the inverse matrix is given by: $A^{-1} = \frac{adj(A)}{det(A)}$

3 State-Space Solutions and Realizations

3.1 Solution of LTI State Equations

Consider the linear time-invariant (LTI) state-space equation:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \tag{12}$$

If we want to find a solution for the system given an initial state x(0) and the input u(t). Using the following propertie for matrices:

$$\frac{d}{dt}e^{At} = Ae^{At} = e^{At}A$$

See that we can solve it as a first order differencial equation:

Multiplying both sides by e^{At}:

$$e^{-At}\dot{x}(t) - e^{-At}Ax(t) = e^{-At}Bu(t)$$

Then, we get:

$$\frac{d}{dt}(e^{-At}x(t)) = e^{-At}Bu(t)$$

Integrating from 0 to t yields:

$$e^{-At}x(t) - e^{0}x(0) = \int_0^t e^{-A\tau}Bu(\tau)d\tau$$

Thus we have:

$$x(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau$$

Using the same process, we will get:

$$y(t) = Ce^{At}x(0) + C\int_0^t e^{A(t-\tau)}Bu(\tau)d\tau + Du(t)$$

It's also possible to compute this solutions using the Laplace Transform:

$$\hat{x}(s) = (sI - A)^{-1}[x(0) + B\hat{u}(s)]$$

$$\hat{y}(s) = C(sI - A)^{-1}[x(0) + B\hat{u}(s)] + D\hat{u}(s)$$

Calculating the exponential of a matrix: Since $e^{At} = \mathcal{L}^{-1}(sI - A)^{-1}$, we can simple calculate it by doing some mathematical manipulation.

3.1.1 Example

3.2 Equivalent State Equations

Consider the n-dimensional state equation:

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t)$$

where $A \in \mathbb{R}^{nxn}$ and $x \in \mathbb{R}^n$. This equation may be considered associated with the following orthnormal basis:

$$i_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ . \\ . \\ . \\ 0 \\ 0 \end{bmatrix} \quad i_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ . \\ . \\ . \\ 0 \\ 0 \end{bmatrix} \quad i_{n-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ . \\ . \\ . \\ 1 \\ 0 \end{bmatrix} \quad i_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ . \\ . \\ . \\ 0 \\ 1 \end{bmatrix}$$

Thus, it's possible to study the effect of describing this system in a different basis.

Definition: Let P be an nxn real nonsingular matrix and let $\bar{x} = Px$. Then the state equation,

$$\dot{\bar{x}}(t) = \bar{A}\bar{x}(t) + \bar{B}u(t)$$
$$y(t) = \bar{C}\bar{x}(t) + \bar{D}u(t)$$

where, $\bar{A} = PAP^{-1}$, $\bar{B} = PB$, $\bar{C} = CP^{-1}$, $\bar{D} = D$, is said to be (algebraically) equivalent to the initial description and $\bar{x} = Px$ is called an equivalence transformation.

It can be obatined substituing $x(t) = P^{-1}\bar{x}(t)$ and $\dot{x}(t) = P^{-1}\dot{\bar{x}}$. In this way, we have changed the basis vectors of the state space from the orthonormal basis to the columns of P^{-1}.

It's important to know that the two descriptions have the same set of eigenvalues and the samee transfer matrix. Thus, equivalent state equations have the same characteristic polynomial and, consequently, the same set of eigenvalues and same transfer matrix. In fact, all properties are preserved or invariant under any equivalence transformation.

Theorem: Two linear time-invariant state equations [A,B,C,D] and $[\bar{A}, \bar{B}, \bar{C}, \bar{D}]$ are zero-state equivalent or have the same transfer matrix if and only if $D = \bar{D}$ and $CA^mB = \bar{C}\bar{A}^m\bar{B}$, for m=0, 1, 2,...,n .

3.3 Realizations

Every linear time-invariant (LTI) system can be described by the input-output description:

$$\hat{y}(s) = \hat{G}(s)\hat{u}(s)$$

and, if the system is lumped by the state-space equation:

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$

If the state equation is known, the transfer matrix can be computed as $\hat{G}(s) = C(sI - A)^{-1}B + D$. Now, we want to study the reverse process, it is, if we have the trasnfer matrix, then how could we get the state-space equation, that is what is called *realization.*

A transfer matrix $\hat{G}(s)$ is said to be realizable if there exists a finite-dimensional state equation such that

$$\hat{G}(s) = C(sI - A)^{-1}B + D$$

and [A,B,C,D] is called a realization of $\hat{G}(s)$.If $\hat{G}(s)$ is realizable, then it has infinitely many realizations, not necessarily of the same dimension.

Theorem: A transfer matrix $\hat{G}(s)$ is realizable if and only if $\hat{G}(s)$ is a proper rational matrix.

First, we decompose $\hat{G}(s)$ as

$$\hat{G}(s) = \hat{G}(\infty) + \hat{G}_{sp}(s)$$

where $\hat{G}_{sp}(s)$ is the strictly proper part of $\hat{G}(s)$. Let

$$d(s) = s^r + \alpha_1 s^{r-1} + ... + \alpha_{r-1}s + \alpha_r$$

be the least common denominator of all entries of $\hat{G}_{sp}(s)$ can be expressed as:

$$\hat{G}_{sp}(s) = \frac{1}{d(s)}[N(s)] = \frac{1}{d(s)}[N_1 s^{r-1} + N_2 s^{r-2} + ... + N_{r-1}s + N_r]$$

where N_i are qxp constant matrices. Now, it's possible to see that the set of equations:

$$\dot{x} = \begin{bmatrix} -\alpha_1 I_p & -\alpha_2 I_p & ... & -\alpha_{r-1} I_p & -\alpha_r I_p \\ I_p & 0 & ... & 0 & 0 \\ 0 & I_p & ... & 0 & 0 \\ . & . & & . & . \\ . & . & & . & . \\ . & . & & . & . \\ 0 & 0 & ... & I_p & 0 \end{bmatrix} + \begin{bmatrix} I_p \\ 0 \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} N_1 & N_2 & ... & N_{r-1} & N_r \end{bmatrix} x + \hat{G}(\infty)u$$

is a realizarion of $\hat{G}(s)$, where the matrix I_p is the pxp unit matrix and every 0 is a pxp zero matrix.

Defining

$$Z = \begin{bmatrix} Z_1 \\ Z_2 \\ . \\ . \\ . \\ Z_r \end{bmatrix} = (sI - A)^{-1}B$$

where Z_i is pxp and Z is rpxp. The transfer matrix of the obtained state-space equation is

$$C(sI - A)^{-1}B + \hat{G}(\infty) = N_1Z_1 + N_2Z_2 + ... + N_rZ_r + \hat{G}(s)$$

We know that $Z = (sI - A)^{-1}B$, then $(sI - A)Z = B$ or $sZ = AZ + B$.

From this, we easily see that $sZ_2 = Z_1$, $sZ_3 = Z_2$, ... , $sZ_r = Z_{r-1}$, which implies in $Z_2 = \frac{1}{s}Z_1$, $Z_3 = \frac{1}{s^2}Z_1$, ... , $Z_r = \frac{1}{s^{r-1}}Z_1$.

Thus, we can apply on $sZ = AZ + B$:

$$sZ_1 = -\alpha_1Z_1 - \alpha_2Z_2 - ... - \alpha_rZ_r + I_p = -(\alpha_1 + \frac{\alpha_2}{s} + ... + \frac{\alpha_r}{s^{r-1}})Z_1 + I_p$$

Thus, we have:

$$(s + \alpha_1 + \frac{\alpha_2}{s} + ... + \frac{\alpha_r}{s^{r-1}})Z_1 = \frac{d(s)}{s^{r-1}}Z_1 = I_p$$

Substituting these, we have:

$$C(sI - A)^{-1}B + \hat{G}(\infty) = \frac{1}{d(s)}[N_1s^{r-1} + N_2s^{r-2} + ... + N_r] + \hat{G}(\infty)$$

what proves that is a realization of $\hat{G}(s)$.

4 Stability

4.1 Input-Output Stability of LTI Systems

Consider a SISO linear time-invariant (LTI) system described by:

$$y(t) = \int_0^t g(t - \tau)u(\tau)d\tau = \int_0^t g(\tau)u(t - \tau)$$

where g(t) is the impulse resonse or the output excited by an impulse input applied at t=0.

An input u(t) is said to be bounded if u(t) does not grow to positive or negative infinity or, equivalently, there exists a constant u_m such that:

$$|u(t)| \leq u_m < \infty \text{ for all } t \geq 0$$

.

A system is said to be BIBO stable (bounded-input bounded-output stable) if every bounded input excites a bounded output. This stability is defined for the zero-state response and is applicable only if the system is initially relaxed.

Theorem: A SISO system is said to be BIBO stable if and only if g(t) is absoouletly integrable in $[0, \infty)$ or

$$\int_0^\infty |g(t)|dt \leq M < \infty$$

for some constant M.

Theorem: A SISO system with proper rational transfer function $\hat{g}(s)$ is BIBO stable if and only if every pole of $\hat{g}(s)$ has a negative part or, equivalently, lies inside the left-half s-plane.

4.2 Internal Stability

Definition: The zero-input response of the equation $\dot{x} = Ax$ is marginally stable or stable in the sense of Lyapunov if every finite initial state x_0 excites a bounded response. It is asymptotically stable if every finite initial state excites a bounded response, which, in addition, approaches 0 as $t \to \infty$.

Theorem:

1. A system is said to be marginally stable if and only if all eigenvalues of A have zero or negative real parts and those with zero real parts are simple roots of the minimal polynomial of A.

2. A system is said to be asymptotically stable if and only if all eigenvalues of A have negative real parts.

Theorem: All eigenvalues of A have negative real parts if and only if for any given positive definite symmetric matrix N, the Lyapunov equation $A'M + MA = -N$ has unique symmetric solution M and M is positive definite.

4.3 MATLAB

To know if a system is stable or not, we can use the following function:

```
%Passing as input the state matrix (A) to get as output the eigenvalues
    and a binary variable 'flag' that is '1' if the system is stable and
    '0' if the system is unstable.
function [e,flag]=isStable(A)
    e=eig(A);
    [m,n]=size(e);
    flag=0;
    if (sum(e(:,1)>0)) > 0; %if there is any positive eigenvalue, then
        the system is unstable
            disp("unstable ")
            return
    end
    minpol=minpoly(A);
```

```
    raizes=roots(minpol);
    [m2,n2]=size(raizes);
    comparing=zeros(m2,n2);
    s=0;
    for i=1:m2
        if raizes(m2,1)==0
        s=s+1;
        end
    end
    s=sum(comparing<=raizes(:,1))
    if s>1 %zero isn't simple root of the minimal polynomial
            disp("unstable")
            return
    end
    if s==1 %zero is simple root of the minimal polynomial
            disp("marginally stable")
            flag=1;
            return
    end
    if s==0 %all the eigenvalues are negative
            disp("Assymptotically stable")
            flag=1;
            return
    end
end
```

Applying it on the three thank example, we get that it's assymptotically stable.

5 Controllability and Observability

5.1 Controllability

Definition: The state equation or pair (A,B) is said to be controllable if for any initial state x(0)=x_0 and any final state x_1, there exists an input that transfers x_0 to x_1 in a finite time. Otherwise, it's said to be uncontrollable.

Dealing with controllability, we can use many ways to proof if a system is controllable or not, provided by the following theorem:

Theorem: The following statements are equivalent.

1. The n-dimensional pair(A,B) is controllable.

2. The nxn matrix

$$W_c(t) = \int_0^t e^{A\tau} BB' e^{A'\tau} d\tau = \int_0^t e^{A(t-\tau)} BB' e^{A'(t-\tau)}$$

   is nonsingular for any t> 0.

3. The *nxnp* controllability matrix

$$C = [BABA^2B...A^{n-1}B]$$

   has rank n(full row rank).

4. The $nx(n+p)$ matrix [A-λI B] has full row rank at every eigenvalue λ of A.
5. If, in addition, all eigenvaleus of A have negative real parts, then the unique solution of
$$AW_c + W_c A' = -BB'$$
is positive definite. The solution is called the controllability Gramian and can be expressed as:
$$W_c = \int_0^\infty e^{A\tau} BB' e^{A'\tau} d\tau$$

Corollary: The n-dimensional pair(A,B) is controllable if and only if the matrix

$$C_{n-p+1} := [BAB...A^{n-p}B]$$

where $\rho(B) = p$, has rank n or the nxn matrix $C_{n-p+1}C'_{n-p+1}$ is nonsingular.

5.1.1 Controllability Indices

Corollary: The n-dimensional pair (A,B) is controllable if and only if the matrix: C_{n-p+1} =[B AB ... $A^{n-p}B$], where $\rho(B) = p$ has rank n or the nxn matrix $C_{n-p+1}C'{}_{n-p+1}$ is nonsingular.

5.1.2 MATLAB

Studying the controllability of a system on MATLAB, we can use the following function to see it's gramian, controllability matrix and see if it's controllable or not:

```
%Using matlab's function ctrb to get the controllability matrix and gram to get the controllability gramian to see if the system is controlable or not.
function [C]=isCtrl(A,B,flag)
    disp("Controllaybility Matrix")
    C=ctrb(A,B)
    display(C)
    [cm,n]=size(C')
    if n>rank(C')
        if flag==1
       disp("Controllability Gramian")
       Wc=gram(A,B)
        end
        disp("Uncontrollable")
    end
    if n==rank(C')
        if flag==1
       disp("Controllability Gramian")
       Wc=gram(A,B)
        end
        disp("Controllable")
    end
end
```

Applying it on the three thank example, we get that it's controllable, getting:

$$\text{C} = \begin{pmatrix} 6.4935 & 0 & -0.0437 & 0 & 0.0006 & 0.0003 \\ 0 & 6.4935 & 0 & -0.3252 & 0.0003 & 0.0166 \\ 0 & 0 & 0.0437 & 0.0437 & -0.0009 & -0.0009 \end{pmatrix}$$

$$W_c = 1.0 * 10^3 * \begin{pmatrix} 5.4983 & 0.3057 & 2.3649 \\ 0.3057 & 0.4501 & 0.2167 \\ 2.3649 & 0.2167 & 1.2908 \end{pmatrix}$$

5.2 Observability

Definition: The state equation is said to be observable if for any unknown initial state x(0), there exists a finite $t_1 > 0$ such that the knowledge of the input u and output y over $[0\ ,\ t_1]$ suffices to determine uniquely the initial state x(0). Otherwise, the equation is said to be unobservable.

Dealing with controllability, we can use many ways to proof if a system is controllable or not provided by the following theorem:

Theorem: The state equation is observable if and only if the nxn matrix:

$$W_o(t) = \int_0^t e^{A'\tau} C' C e^{A\tau} d\tau$$

is nonsingular for any $t > 0$.

Theorem: The following statements are equivalent.

1. The n-dimensional pair(A,C) is observable.

2. The nxn matrix
$$W_o(t) = \int_0^t e^{A'\tau} C' C e^{A\tau} d\tau$$
is nonsingular for any t> 0.

3. The $nqxn$ observability matrix:
$$O = \begin{pmatrix} C \\ CA \\ . \\ . \\ . \\ CA^{n-1} \end{pmatrix}$$
has rank n(full column rank).

4. The $(n+q)xn$ matrix $\begin{pmatrix} A - \lambda I \\ C \end{pmatrix}$ has full column rank at every eigenvalue λ of A.

5. If, in addition, all eigenvaleus of A have negative real parts, then the unique solution of
$$A'W_o + W_oA = -C'C$$
is positive definite. The solution is called the Observability Gramian and can be expressed as
$$W_o = \int_0^{\infty} e^{A'\tau}C'Ce^{A\tau}d\tau$$

5.2.1 Observability Indices

Corollary: The n-dimensional pair (A,C) is observable if and only if the matrix:

$$O_{n-q+1} = \begin{bmatrix} C \\ CA \\ . \\ . \\ . \\ CA^{n-q} \end{bmatrix}$$

where $\rho(C) = q$, has rank n or the nxn matrix $O'_{n-q+1}O_{n-q+1}$ is nonsingular.

5.2.2 MATLAB

Studying the observability of a system on MatLab, we can use the following function to see it's gramian, observability matrix and see if it's observable or not:

```
%Using matlab's function obsv to get the observability matrix and gram to get the observability gramian to see if the system is observable or not.
function [O]=isObsv(A,C,flag)
    disp("Observability Matrix")
    O=obsv(A,C);
    display(O);
    [m,n]=size(O);
    if n>rank(O)
        if(flag==1)
            disp("Observability Gramian")
            Wo=gram(A',C');
            display(Wo);
        end
        disp("Unobservable")
    end
    if n==rank(O)
         if(flag==1)
            disp("Observability Gramian")
            Wo=gram(A',C');
            display(Wo);
        end
        disp("Observable")
    end
end
```

Applying it on the three thank example, we get that it's observable, getting:

$$O = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -0.0067 & 0 & 0.0067 \\ 0 & -0.0501 & 0.0067 \\ 0.0067 & 0.0067 & -0.0135 \\ 0.0001 & 0 & -0.0001 \\ 0 & 0.0026 & -0.0004 \\ -0.0001 & -0.0004 & 0.0003 \end{pmatrix}$$

$$W_o = \begin{pmatrix} 160.1531 & 11.5319 & 85.8412 \\ 11.5319 & 11.5322 & 11.5319 \\ 85.8412 & 11.5319 & 85.8412 \end{pmatrix}$$

6 Structural Controllability - Ching-Tai Lin

6.1 Strucutred System

Consider a linear control system described as $\dot{x} = Ax + bu$, where:

- $x \in \mathbb{R}^n$;
- $u \in \mathbb{R}$;
- $A \in \mathbb{R}^{nxn}$;
- $b \in \mathbb{R}^n$;

It's known that the set of all (completely) controlable pairs is open and dense (a subset $S \subset \mathbb{R}^d$ is said to be dense in $\mathbb{R}^d$ if, for each $r \in \mathbb{R}^d$ and every $\epsilon > 0$, there is an $s \in S$ such that the Euclidean distance $d(s, r) \geq \epsilon$) in the space of all pairs (A,b). If the pair (A_0, b_0) is not completely controllable, then for every $\epsilon > 0$, there exists a completely controllable pair (A,b with $||A - A_0|| < \epsilon$ and $||b - b_0|| < \epsilon$).

The following definitions will be the basis of our study:

Definition: (Structured Matrix) The elements of a structure matrix [A,B] are either fixed at zero or indeterminate values which are assumed to be independent of one another.

Definition: A numerically given matrix (A_0, B_0) is called an admissible numerical realization (with respect to [A,B]) if it can be obtained by fixing all indeterminate entries of [A,B] at some particular values.

Definition:(Equivalent Structured Matrices) The pair (A,b) has the same structure as another pair $(\bar{A}, \bar{b})$ if for every fixed (zero) entry of the matrix (A b), the corresponding entry of the matrix $(\bar{A}\bar{b})$ is fixed (zero) and, at the same time, for every fixed (zero) entry of $(\bar{A}\bar{B})$, the Given the following matrice $(A_1 b_1)$:

$$A_1 = \begin{pmatrix} 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} b_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \alpha \end{pmatrix}$$

corresponding entry of (A b) is also fixed (zero). For example, the following two matrices has the same structure:

$$A_1 = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 5 & 6 & 8 \end{pmatrix} A_2 = \begin{pmatrix} 4 & 3 & 0 \\ 3 & 5 & 0 \\ 1 & 8 & 7 \end{pmatrix}$$

Definition: A property holds structurally within a class of structurally equivalent systems if the property under investigation holds numerically for "almost all" admissible numerical realizations.

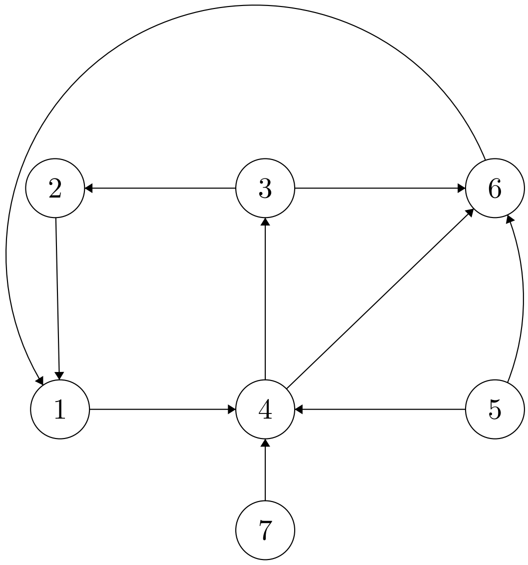
Definition (1): (Structural Controllability) *The pair* (A_0, b_0) *is said to be structurally controllable if and only if there exists a completely controllable pair (A,b) which has the same structure as* (A_0, b_0).;

Definition (2): (Structural Controllability) *The pair* (A_0, b_0) *is structurally controllable if and only if* $\forall \epsilon > 0$, *there exists a completely controllable pair* (A_1, b_1) *with* $||A_1 - A_0|| < \epsilon$ *and* $||b_1 - b_0|| < \epsilon$;

6.2 The Graph of a pair (A,b)

Given a pair (A,b), it's digraph G contains n+1 nodes, $v_1, v_2, ..., v_{n+1}$, and all of whose edges are defined for each nonfixed entry e_{ij} of the $n \times (n+1)$ matrix (A b), where there is a edge going from v_j to v_i and it's edge length is the numerical value on e_{ij}. By definition, we call the node v_{n+1} is called the "origin" of G. For example, the digraph of the following pair (A,b) is given by:

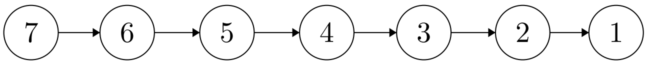
$$A = \begin{pmatrix} 0 & \alpha & 0 & 0 & 0 & \alpha \\ 0 & 0 & \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha & 0 & 0 \\ \alpha & 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha & \alpha & \alpha & 0 \end{pmatrix} b = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \alpha \\ 0 \\ 0 \end{pmatrix}$$



6.2.1 Main DiGraph Classes on the Subject

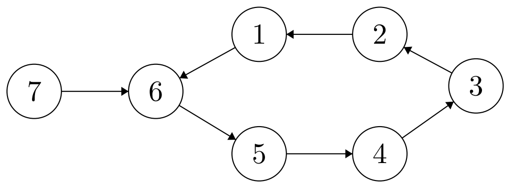
Definition: (*Stem*) A stem is a DiGraph in which it's possible to access all the nodes going throughout one only way leading to a node that doesn't go to any other node. Given the following pair $(A_1 b_1)$:

$$A_1 = \begin{pmatrix} 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} b_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \alpha \end{pmatrix}$$



Definition: (*Bud*) A bud is a DiGraph in which it's possible to access all nodes going throughout one only way leading to the initial node (which is accessed by the origin node). Given the following pair $(A_2 b_2)$:

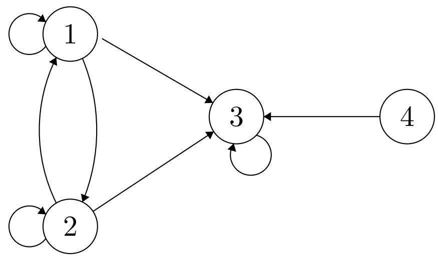
$$A_2 = \begin{pmatrix} 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha \\ \alpha & 0 & 0 & 0 & 0 & 0 \end{pmatrix} b_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \alpha \end{pmatrix}$$



Definition: (*Non-accessible node*) A node is said to be non-accessible if and only if there is no possibility of reaching it starting from the origin. For example, given

$$A = \begin{pmatrix} \alpha & \alpha & 0 \\ \alpha & \alpha & 0 \\ \alpha & \alpha & \alpha \end{pmatrix} b = \begin{pmatrix} 0 \\ 0 \\ \alpha \end{pmatrix}$$

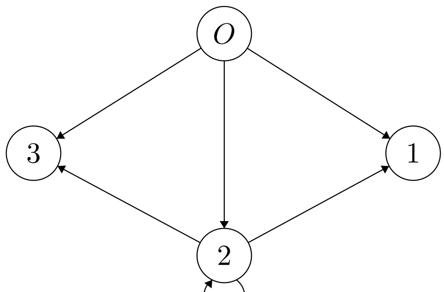
.



Definition: (*Dilation*) The graph pair (A,b) contains a "dilation" if and only if there is a set S of k nodes in the vertex set of the graph - not containing the origin v_{n+1} - such that there are no more than k-1 nodes v_j in T(S). For example, Given

$$A = \begin{pmatrix} 0 & \alpha & 0 \\ 0 & \alpha & 0 \\ 0 & \alpha & 0 \end{pmatrix} b = \begin{pmatrix} \alpha \\ \alpha \\ \alpha \end{pmatrix}$$

.



6.2.2 Cacti

Lemma 1: Suppose that G is a graph of a structurally controllable pair. Let B be a bud with origin e, and suppose e is the only node which belongs at the same time to the vertex set of G and to the vertex set of B. Then GUB is the graph of a structurally controllable pair.

The graph of a pair (A,b) is a cactus P if it can be obtained by starting from a stem S and by constructing a sequence of graphs $G_0 \subset G_1 \subset ... \subset G_k \subset ... \subset G_p$ as follows: The first graph G_0 in the sequence is the the stem S, and the last graph G_p in the sequence is the cactus P.

The graph of a pair (A,b) is a cactus if and only if one can write $P = SUB_1UB_2U...UB_p$ where S is a strem and B_i are buds and, for every i=1,2,...,p, the origin e_i of B_i is also the origin of an (oriented) edge of the graph $P = SUB_1UB_2U...UB_{i-1}$. Moreover e_i is the only node which belongs at the same time to the vertex set B_i and to the vertex set of $P = SUB_1UB_2U...UB_{i-1}$.

Proposition: If the graph of a pair (A,b) is a cactus, then the pair is structurally controllable.

Proposition: If the graph of a pair (A,b) is spanned by a cactus, then the pair (A,b) is structurally controllable.

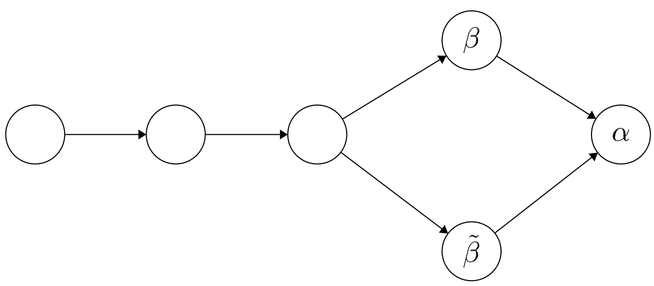
6.3 A Class of Graphs which are Cacti

Assume that G is the graph of a pair (A,b) and has the following properties:

1. There is no non-accessible node in the vertex set of G.
2. There is no dilation.
3. G is minimal. (after deleting any edge of the graph, one of the properties 1) and 2) is violated).

Lemma 2: Every node in G is accessible from the origin along one and only one simple path.

Proof: Suposse a node α can be reached along two distinct paths and let β and $\tilde{\beta}$ be the last nodes met before α on these paths. Since the two paths are distinct, then we may assume $\beta \neq \tilde{\beta}$. Deleting the edge (β, α) it's obtained a new graph G_1. By the third propertie of the graph, G_1 must be minimal, then we need to find a nonacessible node or a dilation, but it's impossible to get in G_1 a nonacessible node.



Thus, G_1 is suppose to have a dilation. Defining S=$\{\alpha_1, \alpha_2, \alpha_3, ..., \alpha_k\}$ and T_1(S)=$\{\beta_1, \beta_2, ..., \beta_p\}$, with $p \leq k-1$. Then, it's possible to see that $\tilde{\beta} \in T_1(S)$, $\alpha \in S$, $\beta \notin T_1(S)$ and T(S)=$T_1(S) \cup \beta$, knowing that $T_1(S) \cup \beta \neq T_1(S)$.

Similarly, it's possible to obtain a graph G_2 by deleting the edge $(\tilde{\beta}, \alpha)$. Then, there exists a set $\tilde{S}$ of $\tilde{k}$ nodes such that $T_2(\tilde{S})$ (with respect to G_2) contains no more that $\tilde{k}-1$ nodes. Defining $\tilde{S} = \{\tilde{\alpha}_1, \tilde{\alpha}_2, ..., \tilde{\alpha}_{\tilde{k}}\}$ and $T_2(\tilde{S}) = \{\tilde{\beta}_1, \tilde{\beta}_2, ..., \tilde{\beta}_{\tilde{p}}\}$ with $\tilde{p} < \tilde{k} - 1$.Then, it's also possible to see that $\beta \in T_2(\tilde{S})$, $\alpha \in \tilde{S}$, $\tilde{\beta} \notin T_2(\tilde{S})$ and T($\tilde{S}$)=$T_2(\tilde{S}) \cup \tilde{\beta}$, knowing that $T_2(\tilde{S}) \cup \tilde{\beta} \neq T_2(\tilde{S})$.

Now, let us define $\breve{S} = S \cup \tilde{S}$, it's possible to see that:

$$T(\breve{S}) = (T_1(S) \cup \beta) \cup (T_2(\tilde{S}) \cup \tilde{\beta})$$

$$T(\breve{S}) = T_1(S) \cup \beta \cup T_2(\tilde{S}) \cup \tilde{\beta}$$

$$T(\breve{S}) = T_1(S) \cup T_2(\tilde{S})$$

.

Using N(M) as the number of distinct elements in a set. Consider the first case in which S and $\tilde{S}$ have always in common the node $\alpha = \alpha_1 = \tilde{\alpha}_1$, then we may get $N(\breve{S}) = N(S) + N(\tilde{S}) - 1 = k + \tilde{k} - 1$. Since, $T(\breve{S}) = T_1(S) \cup T_2(\tilde{S})$, we get:

$$N(T(\breve{S})) = N(T_1(S) \cup T_2(\tilde{S})) \leq N(T_1(S)) + N(T_2(\tilde{S})) = p + \tilde{p} \leq (k-1) + (\tilde{k}-1) = k + \tilde{k} - 2$$

Thus, we found a set $\breve{S}$ of $k + \tilde{k} - 1$ nodes such that the set T($\breve{S}$) (in the original graph) contains no more than $k + \tilde{k} - 2$ nodes, what is a contradiction.

Now, consider the case in which S and $\tilde{S}$ have other nodes in common, besides α. Suppose that $\alpha_i = \tilde{\alpha}_i$, for i= 1, ...,j where $1 < j < min(k, \tilde{k})$, and that $\alpha_m \neq \tilde{\alpha}_l$, $\forall j < m \leq k$ and $\forall j < l \leq \tilde{k}$. Define $S_0 = \{\alpha_2, \alpha_3, ..., \alpha_j\}$ and consider the corresponding set $T(S_0)$ in the graph G. We claim that $\beta \notin T(S_0)$, $\tilde{\beta} \notin T(S_0)$, $T(S_0) \subset T_1(S) \cap T_2(\tilde{S})$ and $N(T(S_0)) \geq j - 1$.

Assume again $\breve{S} = S \cup \tilde{S}$, then:

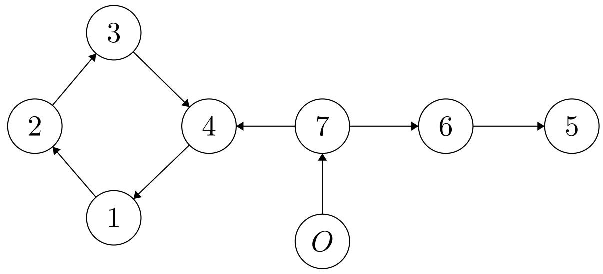
$$N(T(\breve{S})) = N(T_1(S) \cup T_2(\tilde{S})) = N(T_1(S)) + N(T_2(\tilde{S})) - N(T_1(S) \cap T_2(\tilde{S}))$$

$$\leq N(T_1(S)) + N(T_2(\tilde{S})) - N(T(S_0)) \leq p + \tilde{p} - (j-1) = k + \tilde{k} - j - 1$$

.

On the other hand, the number of distinct nodes in $\breve{S}$ is N($\breve{S}$) = $j + (\tilde{k} - j) + (k - j) = k + \tilde{k} - j$. Hence, G has a dilation. This contradiction proves Lemma 2.

Terminal bunch: If there exists a subset $S \subset V_i$, where V_i is the set of all the nodes which can be reached from the origin of G by passing through the edge n_i, which is an oriented edge in G with the property that their origins coincide with the origin of G, such that N(T(S))=N(S).



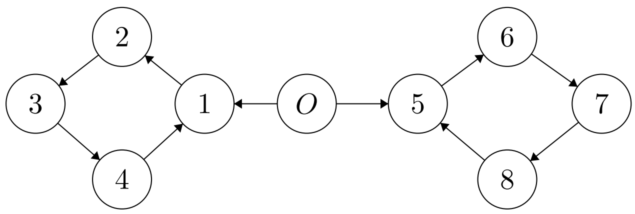
Lemma 3: If G_i is not a terminal bunch, then for every set $S \subset V_i$ such that T(S) contains the origin, one has N(T(S))-N(S) ≥ 1. (Indeed, the case N(T(S))<N(S) is contradictory since it implies the existence of a dilation in G.

Lemma 4: There exists at most one terminal bunch in G.

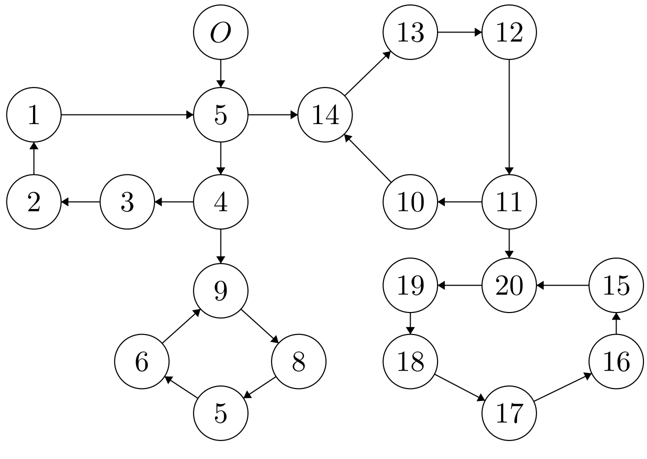
Proof: Suppose there are more than one terminal bunches. Let G_1 and G_2 be two terminal bunches. Then there exists $S_i \subset V_i$ such that $N(S_i) = N(T(S_i))$ and $T(S_i)$ contains the origin of G, with i $=$ 1,2. Define $S = S_1 \cup S_2$ and consider the correspondingg set T(S) in G. Since V_1 and V_2 are disjoint, S_1 and S_2 are also disjoint, and $T(S_1)$ and $T(S_2)$ have in common only the origin of G. Therefore, we get:

$$N(T(S)) = N(T(S_1)) + N(T(S_2)) - 1 = N(S_1) + N(S_2) - 1 = N(S) - 1$$

.

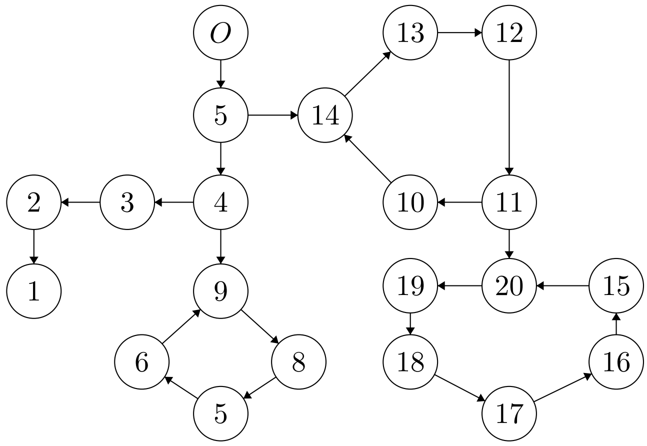


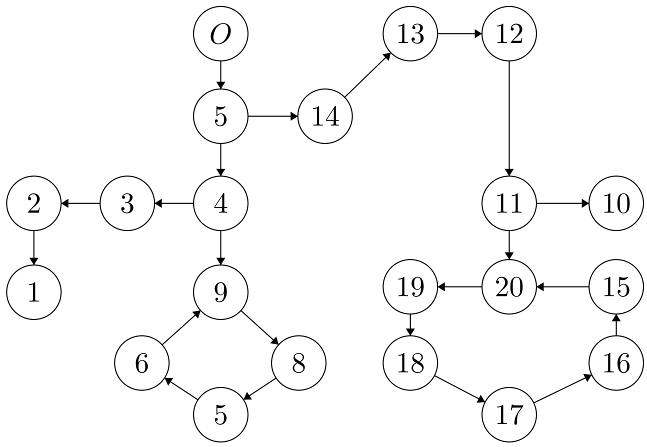
A graph H is a "precactus" if and only if one can write $H = B_1UB_2U...UB_p$, where B_i are the buds, such that for every i=1,2,3...,p, the origin e_i of B_i is also the origin of one oriented edge in the graph of $B_1UB_2U...B_{i-1}$



Lemma 5: Every precactus becomes a cactus after eliminating one or more suitable edges.

Proof: From the definition, a precactus can be written in the form $H = B_1 \cup B_2 \cup ... \cup B_p$. Let e_1 be the origin of B_1 and (e_1, β_1) be distinguished edge of B_1. Then besides the edge (e_1, β_1) the bud contains one and only one edge of the form (e_2, β_1) where e_2 is another node belonging to the vertex set of B_1. Remove the edge (e_2, β_1). Then we can write $B_1 = S_1 \cup (e_2, \beta_1)$, where S_1 is a stem whose vertex set is identical to the vertex set of B_1. If e_2 is not the origin of any other bud, $B_2, B_3, ...B_p$, then the proof is complete. If not, then e_2 is the origin of another bud (one can assume this bud is B_2), and the same procedure can be repeated obtaining a new node e_3 and a new relation $B_1 \cup B_2 = S_2 \cup (e_2, \beta_1) \cup (e_3, \beta_2)$ where S_2 is a stem whose vertex set is identical to the vertex set of $B_1 \cup B_2$. After applying the procedure the maximal number of times, it's possible to obtain a node $e_{n+1}(n \leq p)$ which is not the origin of any bud, $B_{n+1}, B_{n+2}, ..., B_p$. Therefore, one has $B_1 \cup B_2 \cup ... \cup B_n = S_n \cup (e_2, \beta_1) \cup (e_3, \beta_2) \cup ... \cup (e_{n+1}, \beta_n) = S_n \cup G_0$ where $G_0 = (e_2, \beta_1) \cup (e_3, \beta_2) \cup ... \cup (e_{n+1}, \beta_n)$ and S_n is a stem whose vertex set is identical to the vertex set of $B_1 \cup B_2 \cup ... \cup B_n$. Thus $B_1 \cup B_2 \cup ... \cup B_n \cup B_{n+1} \cup ... \cup B_p = S_n \cup G_0 \cup B_{n+1} \cup ... \cup B_p = P \cup G_0$ where $P = S_n \cup B_{n+1} \cup ... \cup B_p$.Thus, the graph P satisfies the conditions of a cactus.

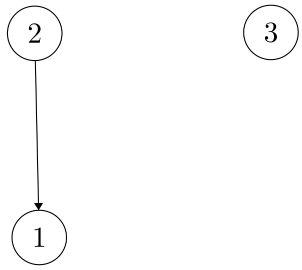


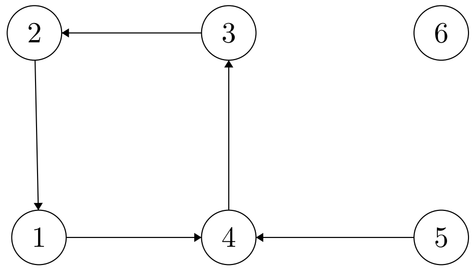


Lemma 6: Any nonterminal bunch becomes a precactus, possibly after eliminating some edges of thee bunch.

Proof: With n_i, V_i and G_i defined as before, let G_1 be a nonterminal buch and $\beta_1 \in V_1$ be the only successor of the origin along the oriented edge n_1 (issuing from the origin). Then thre must exist another edge entering β_1 from a node ξ_1. (Otherwise it's possible to define S=$\{\beta_1\}$) and $T(S) = e =$ the origin. This is contrary to Lemma 3.) One might have $\xi_1 = \beta_1$. Assume first that $\xi_1 \neq \beta_1$. There is one and only one simple path which connects ξ_1 to the origin (Lemma 2). This path goes through β_1 since G_1 is disjointed of the other bunches of G. Therefore, it's possible to find a loop whose set of nodes contains the node β_1. This loop together with the edge (e_1, β_1) entering β_1 forms a bud B_1 (where e_1 is the origin). If $\beta_1 = \xi_1$, it still has a bud.

If the vertex set of B_1 is the same as the vertex set of G_1, then G_1 IS "spanned" by B_1, which, by definition, is a precactus. If not, one takes a new node q. There is one and only one simple path π connecting the origin to q. Let e_2 be the last node in the vertex set of B_1 which is met along this path. Let β_2 be the first node which is met after e_2 on this path. Applying to the node β_2 the same arguments which were used before connecting the node β_1, we obatin another bud B_2. Moreover e_2 is the origin of B_2 and is the only node which belongs at the same time to the vertex set of B_2. The same procedure can be applied successively until all nodes of V_1 belong to the vertex set of some buds $B_1, B_2, ..., B_p$ (because V_1 contains only a finite number of nodes). Clearly, the origin e_i of the bud $B_i (1 < i \leq p)$ is the only node which belongs at the same time to the vertex set of B_i and to the vertex set of $B_1 \cup B_2 \cup ... \cup B_{i-1}$. Therefore, $H = B_1 \cup B_2 \cup ... \cup B_p$ is a precactus according to the previous definition.





Lemma 7: There always exists a terminal bunch in G.

Proof: Suppose there are no terminal bunches in G. It's possible to see that all these nonterminal bunches in G are disjoint from one another, but every nonterminal bunch is spanned by a precactus (Lemma 6). Thus G is spanned by a precactus. But then G does not satisfy assumption 3) at the beginning of this section.

Indeed, Lemma 5 shows that, after eliminating one or several situable edges from the precactus, it's possible to obtain a cactus P. By a past proposition, P is the graph of a structurally controllable pair (A,b). As already shown, this implies that P satisfies the assumptions 1) and 2) at the beginning of this section. Thus assumption 3) is contradicted and Lemma 7 is proved.

Lemma 8: Any terminal bunch G_1 is spanned by a cactus. (Proof similar to Lemma 6 proof)

Proposition 4: If the graph of a pair (A,b) satisfies the properties 1) through 3), then is a cactus.

6.4 Conclusuion

Theorem: The following properties are equivalent.

1. The pair (A,b) is structurally controllable.
2. There is no permutation of coordinates, bringing the pair (A,b) to one of the form: $A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} b = \begin{pmatrix} 0 \\ b_2 \end{pmatrix}$
3. The graph of (A,b) contains no non-accessible node and no dilation.
4. The graph of (A,b) is spanned by a cactus.

7 Multivariable Control - A Graph-Theoretic Approach

7.1 Structure Matrices and their associated digraphs

A linear multivariable control system described as:

$$\dot{x} = Ax + Bu$$

$$y = Cx$$

where:

1. $x(t) \in \mathbb{R}^n$ and $A \in \mathbb{R}^{nxn}$
2. $u(t) \in \mathbb{R}^m$ and $B \in \mathbb{R}^{nxm}$
3. $y(t) \in \mathbb{R}^r$ and $C \in \mathbb{R}^{rxn}$

Assume that every every state is available to be used in a scenery of state feedback control, then the feedback matrix F, that for this scenery will be denoted as E.

Thus, in order to investigate multivariable control systems we shall consider some matrices forms:

1. $Q_0 = \begin{pmatrix} 0 & C & 0 \\ 0 & A & B \\ 0 & 0 & 0 \end{pmatrix}$, which has information about how the states, inputs and observers are interconnected.
2. $Q_1 = \begin{pmatrix} A & B \\ E & 0 \end{pmatrix}$, which has information about how the states and inputs are interconnected, and indicates state feedback.

3. $Q_2 = \begin{pmatrix} 0 & C \\ E & A \end{pmatrix}$,which has information about how the states and observers are interconnected, and indicates state feedback.

4. $Q_3 = \begin{pmatrix} 0 & C & 0 \\ 0 & A & B \\ E & 0 & 0 \end{pmatrix}$, which has information about how the states, inputs and observers are interconnected, and indicates state feedback.

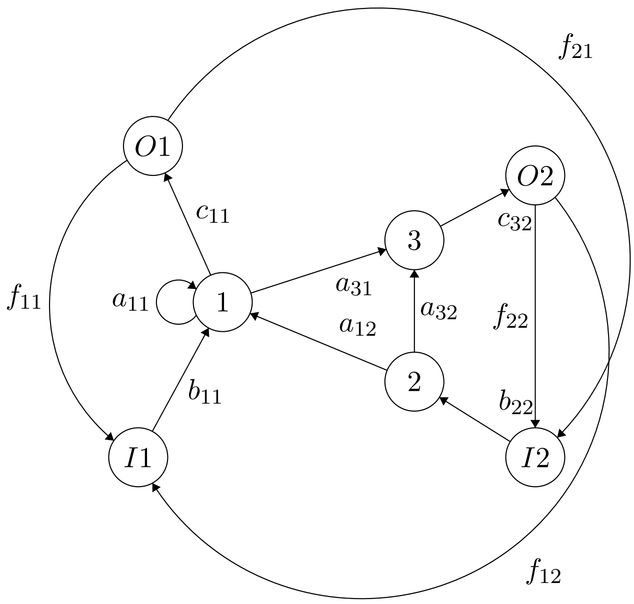
5. $Q_4 = \begin{pmatrix} 0 & C & 0 \\ 0 & A & B \\ F & 0 & 0 \end{pmatrix}$,which has information about how the states, inputs and observers are interconnected, and indicates state feedback too, but state feedback might not be freely assignable.

Definition: Let Q be a given square matrix of order q. Q may me represented by a digraph G(Q) with different $v_1, v_2, ..., v_q$. There exists an edge (v_i, v_j) from the vertex v_i to vertex v_j if and only if the entry q_ij of Q has some non-zero value. The edge weight is equal to the numerical value of q_{ji}.

For example, the digraph of the following system:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ 0 & 0 & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} b_{11} & 0 \\ 0 & b_{22} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} c_{11} & 0 & 0 \\ 0 & 0 & c_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$



7.1.1 Some properties of Irreducible structure matrices

Let [Q] be an irreducible square structure matrix and G([Q]) be the associated digraph. Since it's an irreducible structure matrix, it's possible to reach an arbitrary vertex i from every vertex j of G([Q]).

Theorem: Consider the lengths of all cycles within the digraph G([Q]) of an irreducible structure matrix [Q]. The greatest common divisor of those lengths is equal to the index d of periodicity.

Lemma: Let [Q] be an irreducible nxn structure matrix, [R] an $nx1$ non-zero strucutre matrix, and I the nxn identity matrix. Then for almost all admissible realizations $(Q, R) \in [Q, R]$ there is valid rank$(Q - \lambda I, R) = n$, for all scalars $\lambda \neq 0$.

7.2 Structural controllability, structural observability and structural completeness

A dynamic system is said to be "controllable" if it's state vector can be caused, by an appropriate manipulation of system inputs, to behave in desirable manner. We have already studied a numerical form to investigate the controllability of a dynamic system.

Definition: A class of systems given by it's structure matrix pair [A,B] is said to be structurally controllable if there exists at least one addmissible realization $(A, B) \in [A, B]$ being controllable in the usual numerical sense.

7.2.1 Input-Connectability and Structural Controllability

Definition: A class of systems is said to be input-connectable (or input-reachable) if in the digraph G([Q]) there is, for each state vertex, a path from at least one of the input vertices to the chosen state vertex.

Input-connectability, howerver, is not sufficient for s-controllability.

Lemma: If a class of systems characterized by the structure matrix pair [A,B] is input-connectable, then rank$(A-\lambda I, B) =$ n for all scalars $\lambda \neq 0$ for almost all admissible marices $(A, B) \in [A, B]$.

Algorithm searching for Input-Coonectability

Using the following algorithm it's possible to find the minimum number of vertices to reach all nodes of the system.

```
%minVertextoCtrl returns the minimum nodes to reach all nodes of the
    system.
function [B]=minVertextoCtrl(A)
[m,n]=size(A);
B=zeros(m,1);
C=A;
%check if there exists any node isolated node.
for i=1:n
    if and(sum(A(i,:))==0,sum(A(:,i))==0)
        B(i,1)=1;
    end
end
nonacessed=0;
```

```
%check if there exists any node that isn't reached, but may reach some other nodes.
for i=1:n
    if sum(A(i,:))==0
    B(i,1)=1;
        if nonacessed==0
            nonacessed=i;
        else
            nonacessed=[nonacessed,i];
        end
    end
end
%clean the nonacessed and nodes of the matrix A
if nonacessed~=0
for i=1:length(nonacessed)
    [A]=cleanEdges2(A,nonacessed(i),n);
end
end
%makes the program goes into a loop until all nodes being reached
while sum(sum(A))~=0
    maximum=0;
    col=0;
    %gets the node that reach the greatest number of nodes
    for j=1:n
        if maximum<sum(A(:,j))
            maximum=sum(A(:,j));
            col=j;
        end
    end
    %cleans the nodes that are reached by the node got previously
    [A]=cleanEdges2(A,col,n);
    %updates the node that was got previously if there is any node reaching him
    [A,col]=cleanBack(A,col,n);
    %fills the line of the node that must be used as an origin
    B(col,1)=1;
end
[B]=isItMinimal(C,B,n,nonacessed);
end

%cleanedges2 basically cleans all edges that has as origin the node referenced by the variable col and all nodes that it reaches (recursively, until all nodes of possible nodes are reached).
function [A]=cleanEdges2(B,col,n)
A=B;
pilha=col;
k=1;
  while k<=length(pilha)
    node=pilha(k);
    for i=1:n
        if A(i,node)==1
            A(i,:)=0;
            pilha=[pilha,i];
        end
```

```
    end
    k=k+1;
  end
end
%cleanBack does the same thing that cleanedges2 does, but it start in the referenced node to do the process backwards.
function [A,col]=cleanBack(A,col,n)
for i=1:n
        if sum(A(col,:))==0
            break;
        else
            if A(col,i)==1
                col=i;
                [A]=cleanEdges2(A,col,n);
                [A,col]=cleanBack(A,col,n);
            end
        end
    end
end

%isItMinimal guarantee that the returns are correct and it's possible to reach all nodes of the system.
function [B]=isItMinimal(A,B,n,nonacessed)
C=A;
for i=1:n
    if B(i)==1
    [C]=cleanEdges2(A,i,n);
    end
    for j=1:n
    if ~ismember(j,nonacessed)
    if and(and(B(j)==1,i~=j),sum(A(j,:))==0)
        B(j)=0;
    end
    end
C=A;
end
end
```

7.2.2 Criteria of Structural Controllability

Theorem: A class of systems characterized by the nx(n+m) structure matrix pair [A,B] is s-controllable if and only if: (a) it is input-connectable; (b)s-rank[A,B]=n.

Definition: A cycle is a closed path that reaches no other vertex if not the initial more than once.

Definition: A set of vertex disjoint cycles is said to be a cycle family.

Definition: Consider the digraph G(Q) associated with the square matrix Q. A given cycle family in G(Q) is said to be of width w if this cycle family touches exactly w state vertices.

Theorem: A class of systems characterized by the nx(n+m) structure matrix [A,B] is s-controllable if and only if the digraph G([Q_1]) meets both the following conditions:

- For each state vertex in G($[Q_1]$) there is at least one path from one of the m input vertices to the chosen state vertex.
- There is at least one cycle family of width n in G($[Q_1]$).

$S{-}rank[A, B] < n$ if and only if there is no cycle family of width n in G($[Q_1]$). If there are two or more cycle families of width n in G($[Q_1]$), then, for some admissible realizations $(A, B) \in [A, B]$ numerical cancellation can happen in such a way that $rank(A, B) < n$ despite $s - rank[A, B] = n$. However, if there exists exactly one cycle family in G($[Q_1]$) such a numerical cancellation is impossible.

Corollary: A class of systems characterized by the structure matrix pair [A,B] is strongly s-controllable if and ony if the digraph G($[Q_1]$) meets both the following conditions:

- Input-connectability
- There is exactly one cycle family of width n in G($[Q_1]$).

7.3 Algorithms to examine Digraphs

Connectability matrix (or reachability matrix): [R] is an nnnxn structure matrix defined by: r_{ij}:

- L if a path leads from vertex j to vertex i
- 0 else

```
function [R]=connectabilityMatrix(A)
[m,n]=size(A);
aux=zeros(n);
for i=1:n
    for j=1:n
    if A(i,j)==1
        for k=1:n
        if aux(i,k)==0
            aux(i,k)=j;
            break
        end
        end
    end
    end
end
k=1;
R=zeros(n);
while k<=n
    [R]=connectabilityMatrixRecursive(aux,R,k);
    k=k+1;
end
end
```

Strongly connected matrix: [S] can be obtained by using the connectability matrix: $[S] = [R] \wedge [R']$.

```
S=R & R';
```

Weakely connected matrix: [T] can also be obtained by using the connectability matrix: $[T] = [R] \wedge [\bar{R}']$

```
T=R & (~R)';
```

While $s_{ij} = L$ means that the vertices i and j belong to the same equivalence class, $t_{ij} = L$ means that there is a path from j to i but no path from i to j.