Optimization-based control

Objectives of the course

The objectives of the course are:
- To introduce optimization and convex optimization
- To introduce some of the main ideas in optimal control and more in general to the use of optimization techniques in control
- To detail some techniques and important results in optimal control
- To provide the students with some capability of defining an optimal control problem depending on the specific needs

Contents

The subjects will be touched in this courses are

- Some (not too rigorous) facts about optimization
    - Optimization problem
    - Some classical results (Lagrangian multipliers and Hamiltonian)
    - Convex optimization problems

- Introduction to optimal control

- "Classical" optimal control:
    - Quadratic Optimal Control
    - Some hint on Kalman Filter and LQG

- Linear Matrix Inequalities (LMIs) in control
    - Multiobjective Optimal Control

- Extensions
    - Robust Polytopic
    - LPV Polytopic
    - Systems with disturbance

- How about state/input constraints ?
    - Constraints through LMIs
    - MPC
        - Linear
        - Polytopic Uncertain/LPV
    - Reference and Command Governor

Evaluation
- your involvement in the class
- an oral: discussion of all the exercises, some indepth discussion

Experience 1 – An introduction

In previous courses some basic concept of control have been detailed.

Before starting, let us try to answer some questions

Question 1: Given two controllers both reaching a given objective (e.g. the asymptotically stabilize the output to zero), how to decide which one to use ?

Question 2: Given the controllers

$$u(t) = k_1(x(t)) = \begin{bmatrix} 0.2 & 0 \\ 0 & 0 \end{bmatrix} x(t)$$

$$u(t) = k_2(x(t)) = \begin{bmatrix} 0 & 0 \\ 0 & 0.1 \end{bmatrix} x(t)$$

$$u(t) = k_3(x(t)) = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix} x(t)$$

For the system

$$x(t+1) = \begin{bmatrix} 0.95 & 0.05 \\ 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} u(t)$$

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t)$$

With initial state $x(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$;

Which of the three is the best controller ?

An introduction to optimal control

We have seen that the best possible answer to "what is the best" controller is "**it depends**".

In particular it depends on the **criterion** (**performance index**) we use to evaluate it.

In **optimal control** the idea is to find a way to **generate commands** to our system (possibly subject to certain **constraints**) so as to **optimize some performance index** that typically gives a measure of the deviation of the system behavior from an ideal one (which is very often not attainable).

REMARK: Clearly the choice of an **appropriate performance index** is fundamental because it will determine the nature of the behavior of the optimally controlled system !

CONCLUSION: The control engineer has then the task to formulate **this index on the basis of the requirements** (the specification of the control).

WHAT IS THE PROBLEM ? The problem is that even if an index is define, to find the solution can be hard!

Typically:
1) Solutions are known in **analytical form** for very few classes of (simple) objective functions and systems
1) **Numerical solutions are viable** for some classes of (quite simple) objective functions and systems
3) In many cases the problem to solve, also numerically can be **prohibitive**

Solution: Tradeoff between reality (in terms of system/performance index) and… what we are able to do.

The first step is to try to learn together some of the very basic of optimization we need for this course.

Experience 2 - Some very basics theory of (convex) optimization

1.1 Optimization problems

This is meant to be a course where we are USERS of optimization a not interested in the optimization *per se*. So we will not enter in the precise technicalities. However some introduction is in order. The first question is **what is an optimization problem**

Usually an optimization problem is defined on the basis of
- A vector of **decision variables** x
- An **objective function**
- A set of **constraints** on the decision variables

$$\begin{array}{l} \min_x f(x) \\ \textit{subject to} \\ x \in X \end{array}$$

In general there are two possible approaches to solve this kind of problem
1) (quite rare) try to find an analytical parameterized solution… if it exists
2) solve it numerically. This is in line of principle always possible, however for many classes of problems can be "practically" infeasible.

Very often for certain problems only local minima can be found in a reasonable time.

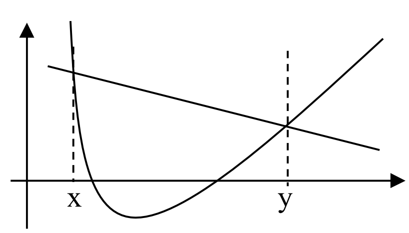
1.2 Convex Optimization problems

A very important class of problems that are **"easy" to solve in a numerical way** is the class convex optimization problems.

Convex optimization problems are problem in the form

$$\begin{array}{l} \min_x f(x) \\ \textit{subject to} \\ x \in X \end{array}$$

where
- $f(x)$ is a **convex function**, i.e. for any x,y, it is true that
$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y). \forall x, y, \ \ \forall \lambda \in [0,1]$



- X is a **convex set**, i.e. $\forall x \in X, y \in X$ all their convex combinations belong to X
$\lambda x + (1-\lambda) y \in X, \ \lambda \in [0 \quad 1]$

Remark: More in general a the convex combination of a set of vectors $x^i, i = 1,...,l$ is a point $\sum_{i=1}^{l} \lambda_i x^i$ where $\lambda_i, i = 1,...,l$ is such that $\sum_{i=1}^{l} \lambda_i = 1, \lambda_i \geq 0$

Why convex optimization is important ?

- In convex optimization a local minimum is also a global minimum. This is due to the fact that $f(x)$ is convex in x if and only if $\frac{\partial^2 f}{\partial x^2} \geq 0, \forall x$ and that the admissible region is convex
- There exists very efficient algorithms to solve numerical convex optimization problems

Question 3: Can a discontinuous set be convex ?

Question 4:

Let us assume X, Y are convex regions:

- Is $X \cap Y$ convex ?
- Is $X \cup Y$ convex ?
- Is the subtraction $X \setminus Y = \{z \mid z = x, x \in X, x \notin Y\}$ convex ?
- Is the Minkowski sum $X \oplus Y = \{z \mid z = x + y, x \in X, y \in Y\}$ convex ?
- Is $-X = \{z \mid z = -x, x \in X\}$ convex ?

Question 5:

Usually optimization problems are written in the form

$$\min_x f(x)$$
$$\begin{cases} c_i(x) \leq 0, & i \in I \\ c_j(x) = 0, & j \in E \end{cases}$$

Assuming $f(x)$ is convex, prove that

1) $c_i(x) \leq 0$ is a convex region if $c_i(x)$ is a convex function
2) $c_j(x) = 0$ is a convex region if $c_j(x)$ is an affine function
3) If $c_i(x)$ are all convex function for all $i \in I$ and $c_j(x)$ are affine functions for all $j \in E$ then the above optimization problem is convex
4) Is condition 3) also necessary ?

Question 6:

- is the sum of two convex functions convex ?
- is a costant function $f(x) = c_0$ convex ?
- is an affine function $f(x) = c^T x + c_0$ convex ?
- under which conditions the quadratic function $f(x) = x^T P x + c^T x + c_0$ is convex ?
- is the norm function $f(x) = \| x \|$ convex ?
- if $f(x)$ is convex, is $f(Tx + c)$ convex ?

Question 7: Consider the system

$$x(t+1) = \begin{bmatrix} 0.95 & 0.05 \\ 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} u(t)$$

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t)$$

with initial state $x(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$;

- $x_1(t)$ represents the percentage of people unable to use a computer in the Public Administration of a country in the year t

- $x_2(t)$ represents the percentage of young people unable to use a computer in the whole country in the year t

- $u_1(t)$ is the action of making courses to teach the use of the computer to a certain percentage of people of the Public Administration

- $u_2(t)$ is the action of making courses to teach the use of the computer to a certain percentage of young people in the country

It is assumed (it is a fictitious cost, just to simplify our life) that :

1 - the cost for the country of having employees not able to use the computer is $10x_1^2$
2 - the cost to making courses to the employees is $c_1 u_1^2$
3 - the cost to making courses to young citizen is $c_2 u_2^2$

If you where the ruler of the country for the next 10 years, which kind of policy would you like to use?

[Since we will use over and over again this example, let us denote this problem as the "computer education problem"]

Question 8: Consider the optimal control problem stated in the "Computer education problem". Is it possible to formulate it as a convex optimization problem ?

<u>Experience 3 - basic practice of (convex) optimization</u>

As "users of optimization", it is not in the aim of this course to enter in the details of the algorithm to solve optimization problems. We are more interested in how to formulate an optimization problem and how to give it to a solver.

There are many solvers available (open, free, commercial) more or less complex and effective and based on different framework (stand-alone, in Matlab, in C++, etc…) .

For the sake of this course we will mostly use some "embedded Matlab" function (not working too well, but a good first attempt) and an open-source Matlab library (The Multi-Parametric Toolbox).

For what regards Matlab we will use 2 Matlab functions

- *fminsearch* – for unconstrained optimization
- *fmincon* - for constrained optimization

<u>Question 9</u> Given the scalar function $f(x) = \left[(0.5 + \sin(x))/(0.1 + x^2)\right]^2$
- Try to find the minimum of this function using *fminsearch* using very different initial conditions
- Try to plot this function and to explain the results you obtained

<u>Question 10</u> Given the multivariable objective function
$f(x_1, x_2) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2$
- Try to find the minimum of this function using *fminsearch*
- Is this optimization problem convex ?

<u>Question 11</u> Consider the problem:

$$min\ 2\left\|x - \begin{bmatrix}0\\0\end{bmatrix}\right\| + \left\|x - \begin{bmatrix}10\\0\end{bmatrix}\right\|$$
$$subjectto$$
$$\left\|x - \begin{bmatrix}6\\7\end{bmatrix}\right\| \leq 1$$

- Try to find the minimum of this function using *fmincon*
- Is this optimization problem convex ?

Another tool we will use is the **MultiParametric Toolbox**.

The main functions we must learn to use are:

sdpvar(n,m,'full')   it generate a n by m matrix of varables
F=[expression]   it is needed to generate constraints
optimize(constraints, costfunction, options) to solve the optimization problem
double(variables)   to get the obtained optimal value

an example: solve the convex problem

$$\min\ 2\left\|x-\begin{bmatrix}0\\0\end{bmatrix}\right\|+\left\|x-\begin{bmatrix}10\\0\end{bmatrix}\right\|$$

subject to

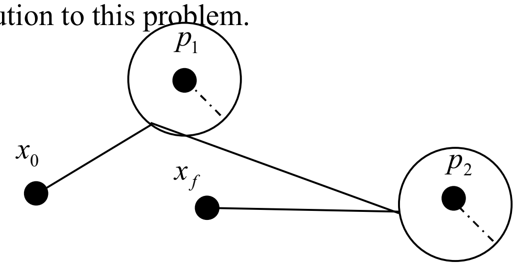
$$\left\|x-\begin{bmatrix}6\\7\end{bmatrix}\right\|\le 1$$

$$x_1 \ge 0,\ x_2 \ge 0$$

```
x= sdpvar(2,1, 'full ');
F= [norm(x-[6;7])<=1]
F=[F x(1) >= 0 x(2) >= 0]
optimize(F, 2*norm(x-[0;0])+norm(x-[10;0]))
optimalValue=double(x)
```

Question 12
- Try to find an optimal solution to the "Computer Education problem" using the MPT toolbox
- Try to find an optimal solution to the "Computer Education problem" using the MPT toolbox when you pass from 5 to 20 and to 40 years

Question 13: Given two targets points in $\Re^2$, p_1, p_2 a starting point x_0 and an end point x_g, try to find the minimal length trajectory composed of 3 segments such that it start in x_0, arrives at distance at least r from p_1, p_2 and then go to x_f. Define the associated optimization problem and build (in MPT) a function that computes the optimal solution to this problem.



Question 14: Generalize the function of Question 15 to the case of n points to be visited

Experience 4 – Unconstrained optimization and optimality conditions

Let us consider a scalar C[2] function $f(x), x \in \Re^n$. If we want to minimize

$$\min_x f(x)$$

we know that the **differential variation** around a vector $\bar{x}$ is

$$dL = \nabla f_{\bar{x}}^T dx + \frac{1}{2} dx^T \nabla^2 f_{\bar{x}} dx + H.O.T.$$

where:

1) $\nabla f_{\bar{x}} = \left.\begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \ldots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}\right|_{x=\bar{x}}$ is the **gradient** of f with respect to x evaluated in $\bar{x}$

2) $\nabla^2 f_{\bar{x}} = \nabla(\nabla f)^T\Big|_{x=\bar{x}} = \left.\begin{bmatrix} \frac{\partial f}{\partial x_1 \partial x_1} & \ldots & \frac{\partial f}{\partial x_1 \partial x_n} \\ \ldots & \ldots & \ldots \\ \frac{\partial f}{\partial x_n \partial x_1} & \ldots & \frac{\partial f}{\partial x_n \partial x_n} \end{bmatrix}\right|_{x=\bar{x}}$ is the **Hessian** of f with respect to x evaluated in

3) dx is the vector of displacement $dx = x - \bar{x}$

The following interesting facts can be proved:

1) If $\bar{x}$ is a **local minimum** then
   a. The gradient is null $\nabla f_{\bar{x}} = 0_n$
   b. The hessian is positive semidefinite $\nabla^2 f_{\bar{x}} \geq 0$

2) If the gradient is null, i.e. $\nabla f_{\bar{x}} = 0_n$ and the hessian is positive definite $\nabla f_{\bar{x}} > 0$
   then $\bar{x}$ is a **isolated local minimum**

Remark: These conditions translate the problem of finding the solution of a set of equality constraints into the one of finding the stationary points and then check the positivity of the Hessian. Pay attention ! Very often, this problem is not easier than the original one !

Question 15: Consider the optimization problem

$$\min_x \frac{1}{2}(Qx-b)^T(Qx-b)$$

where $Q \in \Re^{m\times n}, m > n, rank(Q) = n$

1) Characterize the local optimal solutions using the conditions on the gradient and on the Hessian
2) Have you ever seen this kind of problem ?
3) Is this problem convex ?
4) is this the global minimum ?

Question 16: Consider the optimization problem

$$\min_x \frac{\sin(x_1)+\cos(x_1)}{\sin^2(x_2)+\cos(x_2)}$$

- using the function Jacobian (in matlab) find the stationary points of this function and evaluate the hessian

- is it a convex function ?

Question 17 What is it possible to say on the unconstrained QP problem

$$\min_x \frac{1}{2}x^T Qx + bx + c$$

where $Q > 0$?

Answer to Question 15: This is a quite well-known problem. In fact, it is the well known Mean Least Square minimization :

$$\min_x \frac{1}{2}\varepsilon^T \varepsilon$$
$$Qx = b + \varepsilon$$

The characterization of the stationary points of the objective function is quite easy, in fact

$$\nabla f = \frac{\partial}{\partial x}\frac{1}{2}(Qx-b)^T(Qx-b) = Q^T(Qx-b) = 0$$
$$Q^T Qx = Q^T b$$
$$x = \left(Q^T Q\right)^{-1} Q^T b = Q^+ b$$

where Q^+ is the ***Moore-Penrose pseudo-inverse***.

This point is a minimum in fact

$$\nabla^2 \frac{1}{2}(Qx-b)^T(Qx-b) = Q^T Q > 0$$

Remark 1: This thing may be extended also to the case Q is not full column-rank. In such case one has to use the general definition of the Moore-Penrose pseudo-inverse computed using the SVD decomposition (*pinv* in matlab)

Remark 2: Note that very often this kind of analysis gives not results so nice as in this case and do not simplify the problem itself. However this kind of analysis can always be used as a "stop condition" to understand if we are in a local minimum or not.

Answer to Question 17

The Hessian is clearly positive definite $\nabla^2 f = Q > 0$

For what regard the gradient it is

$$\nabla f = Qx + b^T = 0$$

and the minimum is

$$\bar{x} = -Q^{-1} b^T$$

Experience 5 – Constrained optimization and Lagrange multipliers (equality constraints)

Let us consider a scalar objective function $f(x), x \in \Re^n$ and let the vector-evaluated function $c(x) \in \Re^m$. Assume these are smooth functions of class C^2. Assume we want to solve the following **constrained optimization problem**

$$\min_x f(x)$$
$$s.t.\ c(x) = 0_m$$

Interestingly, the characteristic of the extreme points (local maxima and minima) of this optimization problem are the same of the **Lagrangian function**

$$L(x, \lambda) = f(x) - \lambda^T c(x)$$

$L(x, \lambda)$ is sometimes also known as the Hamiltonian function (and denoted by H) and $\lambda \in \Re^m$ are the so-called **Lagrange multipliers**.

Under the hypothesis that the **local minimum is regular**, i.e. the lines of the Jacobian of $c(x)$ evaluated in $\bar{x}$, i.e. $Jc_{\bar{x}} = [(\nabla c_1(x)) \ \ldots \ (\nabla c_n(x))]^T$ are linearly independent, it is possible to prove that:

1) If $\bar{x}$ is a local minimum then it exists $\bar{\lambda}$ such that :
   a. $\nabla L_{\bar{x},\bar{\lambda}} = 0$
   b. $s^T \nabla_x^2 L_{\bar{x},\bar{\lambda}} s \geq 0 \qquad \forall s \in \{s \mid Jc_{\bar{x}}\, s = 0\}$, where ∇_x^2 is the Hessian w.r.t. only the x variable
2) Some sufficient conditions exists (e.g. if $\nabla L_{\bar{x},\bar{\lambda}} = 0$ and $\nabla_x^2 L_{\bar{x},\bar{\lambda}} > 0$)

Question 18

Make explicit first order necessary conditions for a point $(\bar{x}, \bar{\lambda})$. *[hint: compute the gradient of the Laplacian along x and λ separately]*

Question 19 Consider an optimization problem in the form

$$\min x^T R x$$
$$s.t.\ Ax = b$$

where $R > 0$. Moroever let $rank\{A\} = m, A \in \Re^{m \times n}$. Analyze this general problem. Using the Lagrange analysis, can we say something on the optimal solution ?

Question 20 Using the above conditions try to find the stationary point of the following optimization problem

$$\min x + y$$
$$s.t.\ x^2 + y^2 = 1$$

*[**Hint**: to solve the resulting set of nonlinear equation, try to use fminsearch]*

Answer to Question 18 and some comment

The first order conditions are:

$$\nabla_{\lambda} L_{\bar{\lambda},\bar{x}} = c(\bar{x}) = 0$$

$$\nabla_{x} L_{\bar{\lambda},\bar{x}} = \nabla f_{\bar{x}} + (Jc_{\bar{x}})^T \bar{\lambda} = 0$$

The first condition $c(\bar{x}) = 0$ implies the fact that **at the optimum, constraints must be satisfied**.

To **give an idea** about the second condition, try to think to the idea of **directions.** We know that, given a direction $s \in \Re^n$, the derivative of a function $f(x), \Re^n \to \Re$ is $s^T \nabla f$.

If $\nabla_x L_{\bar{\lambda},\bar{x}} = 0$ and $c(\bar{x}) = 0$ we know that

- $s^T \nabla f_{\bar{x}} = 0$ are direction along with the derivative of the cost function is 0
- $s^T (Jc_{\bar{x}})^T = 0$ for all the admissible directions

IMPORTANT OBSERVATION: **A point is a potential minimum/maximum** if the derivative of the cost function is 0 along all the admissible directions.

Let us consider now:

$$(\nabla_x L_{\bar{\lambda},\bar{x}}) = 0,$$

If the above equation is true is also true that

$$s^T (\nabla_x L_{\bar{\lambda},\bar{x}}) = s^T (\nabla f_{\bar{x}} + (Jc_{\bar{x}})^T \bar{\lambda}) = s^T \left(\nabla f_{\bar{x}} + \sum_{i=1}^{n} \lambda_i \nabla_{\bar{x}_i} c_i \right) = 0, \quad \forall s \in \Re^n$$

then for a given direction s' we can have the following cases:

1) if s' is an **admissible direction**, then $s'^T (Jc_{\bar{x}})^T = 0$ which implies that $s'^T \nabla f_{\bar{x}} = 0$ (the condition we wanted)

2) if $s'^T \nabla f_{\bar{x}} \neq 0$ and $\left(s'^T \nabla f_{\bar{x}} + s'^T \sum_{i=1}^{n} \lambda_i \nabla_{\bar{x}_i} c_i \right) = 0$, it must exist at least one constraints i such that $s'^T \nabla_{\bar{x}_i} c_i \neq 0$ and then s' direction cannot be an **admissible direction**,

Answer to Question 19

$$\min x^T R x$$

$$s.t. \; Ax = b$$

This is the so called Quadratic Programming problem. Using the Lagrangian approach we have the following conditions:

$$L(x, \lambda) = x^T R x - \lambda^T (Ax - b)$$

$$\Delta_{\lambda} L = (Ax - b) = 0$$

$$\Delta_{x} L = Rx - A^T \lambda = 0$$

And then since $\nabla_x^2 L(x, \lambda) = R > 0$ to find a solution for this problem is equivalent to evaluate the solutions of the following linear system

$$Ax = b$$
$$Rx - A^T\lambda = 0$$

Which if $R > 0$ can be solved in closed form !!!!!

Remark: There is another way to approach this problem where we rewrite the problem as follows

$$\min x^T Rx$$
$$s.t.\ \ Ax = b$$

Using the QR factorization we can write

$$A^T = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$

Where $R_1 \in \Re^{m\times m}$ is an upper triangular matrix, $Q_1 \in \Re^{n\times m}$, $Q_2 \in \Re^{n\times(n-m)}$ and $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$ is a unitary matrix, i.e. $Q^T Q = I$

At this point let us work on the constraint $Ax = b$ that becomes

$$Ax = \begin{bmatrix} R_1 & 0 \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} x = b$$

Let us introduce the change of variable

$$Q^T x = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} x = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

If we substitute we obtain

$$\begin{bmatrix} R_1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = b$$

From ths equation we can obtain that:

- $R_1 v_1 = b$ which implies $v_1 = R_1^{-1} b$ and then v_1 is fixed by the constraints
- v_2 is unconstrainted

At this point

$$\min x^T Rx$$
$$s.t.\ \ Ax = b$$

is equivalent to

$$\min_v v^T Q^T RQv$$
$$s.t.\ \ v_1 = R_1^{-1} b$$

and finally to the quadratic unconstrained quadratic programming problem

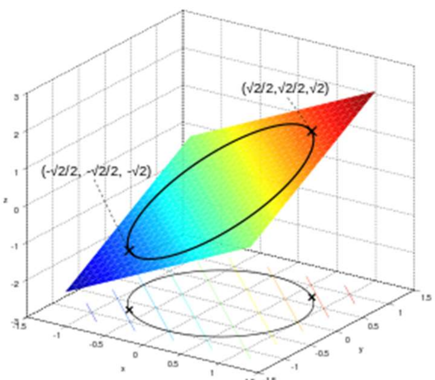
$$\min_{v_2} \begin{bmatrix} (R_1^{-1} b)^T & v_2 \end{bmatrix} Q^T RQ \begin{bmatrix} R_1^{-1} b \\ v_2 \end{bmatrix}$$

Answer to Question 20

$$\min x + y$$
$$s.t.$$
$$x^2 + y^2 = 1$$

The Lagrangian is
$$L(x, y, \lambda) = (x + y) + \lambda_1 (x^2 + y^2 - 1)$$

The stationary points are the points such that
$$x^2 + y^2 = 1$$
And
$$\nabla_{x,y} L = \begin{bmatrix} 1 + 2\lambda_1 x \\ 1 + 2\lambda_1 y \end{bmatrix} = 0$$
So we have to solve the following system of 3 equations in 3 unknowns
$$\begin{cases} x^2 + y^2 = 1 \\ 1 + 2\lambda_1 x = 0 \\ 1 + 2\lambda_1 y = 0 \end{cases}$$



The solution of which can be found with fminsearch rewriting everything as the following nonlinear least mean square problem:
$$\min \varepsilon^T \varepsilon$$
$$\begin{cases} x^2 + y^2 - 1 = \varepsilon_1 \\ 1 + 2\lambda_1 x = \varepsilon_2 \\ 1 + 2\lambda_1 y = \varepsilon_3 \end{cases}$$

Experience 7 – Constrained optimization and Lagrange multipliers (inequality constraints)

Ww will not enter into the details but in the case we also have equality constraints

$$\min_x f(x)$$
$$c_i(x) = 0, \quad i \in E$$
$$c_i(x) \geq 0, \quad i \in I$$

The Lagrangian is

$$L(x,\lambda) = f(x) - \sum_{i \in E} \lambda_i c_i(x) - \sum_{i \in I} \lambda_i c_i(x)$$

Let x^* a local minimum and let regularity conditions be true in x^*, Then it exists $\lambda^* \in R^{|E|+|I|}$ such that x^* e λ^* satisfy

$$\begin{cases} \nabla L(x^*, \lambda^*) = 0 \\ c_i(x^*) = 0 \;\; i \in E \\ c_i(x^*) \geq 0 \;\; i \in I \\ \lambda_i^* \geq 0 \quad i \in I \\ \lambda_i^* c_i(x^*) = 0 \qquad \forall i \in E \cup I \end{cases}$$

Experience 8 – Optimal Control for Discrete Time Systems – LQ - LTI

Let us assume that the system is

$$x_{k+1} = Ax_k + Bu_k$$

And the index:

$$J = \frac{1}{2}\left(x_N^T S x_N + \sum_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k \right)$$

Where $Q = Q^T \geq 0, R > 0, S \geq 0$

Question 21: Is this problem convex ? In your opinion, does there exists an analytic solution to this problem ?

Question 22: Find the stationary conditions of the Lagrangian for this optimization problem

Answer to Question 22:

The **Lagrangian is**

$$L = \frac{1}{2} x_N^T S x_N + \sum_{k=0}^{N-1} \frac{1}{2} x_k^T Q x_k + \frac{1}{2} u_k^T R u_k + \lambda_{k+1}^T \left(A x_k + B u_k - x_{k+1} \right)$$

The **gradient of the Lagrangian is**

1) $\nabla_{u_k} L = R u_k + B^T \lambda_{k+1}, \qquad k = 0,\dots, N-1$

2) $\nabla_{x_k} L = Q x_k + A^T \lambda_{k+1} - \lambda_k \qquad k = 0,\dots, N-1$

3) $\nabla_{\lambda_{k+1}} L = A x_k + B u_k - x_{k+1} \qquad k = 0,..,N-1$

And the further condition

4) $\nabla_{x_N} L = S x_N - \lambda_N$

5) x_0 is assigned

The **stationary conditions for extremes** are then

1b) $B^T \lambda_{k+1} + R u_k = 0, \qquad k = 0,\dots, N-1$

2b) $\lambda_k = A^T \lambda_{k+1} + Q x_k, \qquad k = 0,..,N-1$

3b) $x_{k+1} = A x_k + B u_k, \qquad k = 0,\dots,N-1$

4b) $\nabla_{x_N} L = S x_N - \lambda_N$

5b) x_0 is assigned

Now we have that the optimal command is a **feedback of the "co-state"**

$$u_k = -R^{-1} B^T \lambda_{k+1}, \qquad k = 0,\dots,N-1$$

We can substitute it into the state equation and obtain

$$x_{k+1} = A x_k - B R^{-1} B^T \lambda_{k+1}, \qquad k = 0,\dots, N-1$$

Let us consider now the "***co-state backward evolution***"

$$\lambda_k = A^T \lambda_{k+1} + Q x_k, \qquad k = 0,..,N-1$$

Moreover we know that x_0 is assigned and that $\lambda_N = \frac{\partial}{\partial x_N} x_N^T S x_N = S x_N$

We have then a "kind of strange system"

$$\begin{bmatrix} x_{k+1} \\ \lambda_k \end{bmatrix} = \begin{bmatrix} A & -BR^{-1}B^T \\ Q & A^T \end{bmatrix} \begin{bmatrix} x_k \\ \lambda_{k+1} \end{bmatrix}$$

Since λ_k, x_k have the same dimension, at each time instant we can imagine there exists a mapping matrix mapping P_k

$\lambda_k = P_k x_k$

At this point substituting we obtain:

1) **For the state**
   $x_{k+1} = Ax_k - BR^{-1}B^T P_{k+1} x_{k+1}$
   That is
   $x_{k+1} = \left(I + BR^{-1}B^T P_{k+1}\right)^{-1} Ax_k$

2) **For the co-state**
   $\lambda_k = P_k x_k = Qx_k + A^T P_{k+1} x_{k+1}$
   And then substituting the state evolution
   $P_k x_k = Qx_k + A^T P_{k+1}\left(I + BR^{-1}B^T P_{k+1}\right)^{-1} Ax_k$
   this equation is satisfied by
   $P_k = Q + A^T P_{k+1}\left(I + BR^{-1}B^T P_{k+1}\right)^{-1} A$

At this point by using the matrix inversion lemma:
$\left(\tilde{A} - \tilde{B}\tilde{D}^{-1}\tilde{C}\right)^{-1} = \tilde{A}^{-1} + \tilde{A}^{-1}\tilde{B}\left(\tilde{D} - \tilde{C}\tilde{A}^{-1}\tilde{B}\right)^{-1}\tilde{C}\tilde{A}^{-1}$, where
$\tilde{A} = I, \tilde{B} = -BR^{-1}\tilde{D} = I, \tilde{C} = BP_{k+1}$
we obtain
$\left(I + B^T R^{-1} BP_{k+1}\right)^{-1} = I - BR^{-1}\left(I + B^T P_{k+1} BR^{-1}\right)^{-1} B^T P_{k+1} = I - B\left(R + B^T P_{k+1} B\right)^{-1} B^T P_{k+1}$
And then we obtain the so called **Riccati iteration**

$$P_k = Q + A^T P_{k+1} A - A^T P_{k+1} B\left(R + B^T P_{k+1} B\right)^{-1} B^T P_{k+1} A$$

3) **For the final condition** $\lambda_N = Sx_N$, since $\lambda_k = P_k x_k$, we can say that $P_N = S$

At this point we can use the following procedure:

a) **Set** $P_N = S$

b) Using the **Riccati** equation $P_k = Q + A^T P_{k+1} A - A^T P_{k+1} B\left(R + B^T P_{k+1} B\right)^{-1} B^T P_{k+1} A$
compute all $P_k, N-1,\ldots,0$

c) The optimal command becomes a **feedback of the one step future state**
$u_k = -R^{-1}B^T P_{k+1} x_{k+1}$
Since $x_{k+1} = Ax_k + Bu_k$ we finally obtain:
$u_k = -R^{-1}B^T P_{k+1} Ax_k - R^{-1}B^T P_{k+1} Bu_k$
And finally
$u_k = -\left(I + R^{-1}B^T P_{k+1} B\right)^{-1} R^{-1}B^T P_{k+1} Ax_k = -\left(R + B^T P_{k+1} B\right)^{-1} B^T P_{k+1} Ax_k = -F_k x_k$

Where the sequence of $F_k = \left(R + B^T P_{k+1} B\right)^{-1} B^T P_{k+1} A$ is known as the **Kalman gain sequence**

Remark: Please note that we can rewrite the Riccati equation as follows
$P_k = Q + A^T P_{k+1}(A - BF_k)$

or in another very interesting form that is

$$\begin{aligned}
P_k &= Q + A^T P_{k+1}(A - BF_k) = \\
&= Q + A^T P_{k+1} A - A^T P_{k+1} BF_k = \\
&= Q + A^T P_{k+1} A - A^T P_{k+1} BF_k - F_k^T B^T P_{k+1} A + F_k^T B^T P_{k+1} A = \\
&= Q + A^T P_{k+1} A - A^T P_{k+1} BF_k - F_k^T B^T P_{k+1} A + F_k^T B^T P_{k+1} A \\
&= Q + A^T P_{k+1} A - A^T P_{k+1} BF_k - F_k^T B^T P_{k+1} A + A^T P_{k+1} B\left(R + B^T P_{k+1} B\right)^{-1} B^T P_{k+1} A = \\
&= Q + A^T P_{k+1} A - A^T P_{k+1} BF_k - F_k^T B^T P_{k+1} A + \\
&+ A^T P_{k+1} B\left(R + B^T P_{k+1} B\right)^{-1}\left(R + B^T P_{k+1} B\right)\left(R + B^T P_{k+1} B\right)^{-1} B^T P_{k+1} A = \\
&= Q + A^T P_{k+1} A - A^T P_{k+1} BF_k - F_k^T B^T P_{k+1} A + F_k^T\left(R + B^T P_{k+1} B\right) F_k = \\
&= Q + F_k^T RF_k + (A - BF_k)^T P_{k+1}(A - BF_k)
\end{aligned}$$

Please note that the Riccati equation **does not depend** on the initial conditions but only on S !

Question 23: Remembering that

$$P_k = Q + F_k^T RF_k + (A - BF_k)^T P_{k+1}(A - BF_k)$$
$$u_k = -F_k x_k$$

Find the optimal cost of $J = \left(x_N^T P_N x_N + \sum_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k \right)$

[hint: Start from

$$J = \left(x_N^T P_N x_N + x_{N-1}^T Q x_{N-1} + u_{N-1}^T R u_{N-1} + \sum_{k=0}^{N-2} x_k^T Q x_k + u_k^T R u_k \right)$$

]

Answer to Question 23

$P_k = Q + F_k^T R F_k + (A - BF_k)^T P_{k+1}(A - BF_k)$

$u_k = -F_k x_k$

the optimal cost becomes

$$J = \left(x_{N-1}^T (A + BF_N)^T P_N (A + BF_N) x_{N-1} + x_{N-1}^T Q x_{N-1} + x_{N-1}^T F_{N-1}^T R F_{N-1} x_{N-1} + \sum_{k=0}^{N-2} x_k^T Q x_k + u_k^T R u_k \right) =$$

$$= \left(x_{N-1}^T \left[(A + BF_N)^T P_N (A + BF_N) + Q + F_{N-1}^T R F_{N-1} \right] x_{N-1} + \sum_{k=0}^{N-2} x_k^T Q x_k + u_k^T R u_k \right) =$$

$$x_{N-1}^T P_{N-1} x_{N-1} + \sum_{k=0}^{N-2} x_k^T Q x_k + u_k^T R u_k$$

By repeating the process we will arrive to

$$J = \left(x_{N-1}^T (A + BF_N)^T P_N (A + BF_N) x_{N-1} + x_{N-1}^T Q x_{N-1} + x_{N-1}^T F_{N-1}^T R F_{N-1} x_{N-1} + \sum_{k=0}^{N-2} x_k^T Q x_k + u_k^T R u_k \right) =$$

$$= \left(x_{N-1}^T \left[(A + BF_N)^T P_N (A + BF_N) + Q + F_{N-1}^T R F_{N-1} \right] x_{N-1} + \sum_{k=0}^{N-2} x_k^T Q x_k + u_k^T R u_k \right) =$$

$$x_{N-1}^T P_{N-1} x_{N-1} + \sum_{k=0}^{N-2} x_k^T Q x_k + u_k^T R u_k$$

Remark: This result it is very important because gives a very easy way to evaluate the optimal cost simply using $J = x_0^T P_0 x_0$

Question 24: Try to understand if the same technicalities to find the optimal control strategy and the optimal cost be applied to time varying systems:

$x_{k+1} = A_k x_k + B_k u_k$

Question 25: Using the seen approach, solve the "computer education problem" for an horizon of 20 years using matlab to compute the P_k and F_k matrices

Question 26: Assume the optimal solution for 20 years is known ($P_0, P_1, ..., P_{20}, F_0, ..., F_{19}$). Can you compute "pencil and paper" the optimal solution for 21 years ? [try to exploit the fact you know $J = x^T(0) P_0 x(0)$ from the previous solution]

<u>Experience 10 – Optimal Control for LTI DT Systems – Dynamic Programming</u>

The same results derived with the Lagrangian approach, can be derived with the idea of **dynamic programming**.

The main idea of dynamic programming is the **Bellman's Principle of Optimality**

<u>Principle of Optimality (Bellman):</u> *An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.*

The idea behind is quite straightforward, in practice if we have if we have a certain cost function in the form

$$J(x_i, i) = \sum_{k=i}^{N} J_k(x_k, u_k)$$

$u^*_{[0,N]}$ is the optimal sequence for the cost $J(x_0, 0)$ if and only if $u^*_{[i,N]}$ is the optimal sequence for each subproblem $J(x_i, 0), i = 0, \ldots, N-1$.

<u>What is the idea then ?</u> Let us consider again the quadratic regulator problem

$$J(x_i, i) = \frac{1}{2}\left(\sum_{k=i}^{N-1} x_k^T Q x_k + u_k^T R u_k\right) + \frac{1}{2} x_N^T S x_N$$

$$x_{k+1} = Ax_k + Bu_k$$

we know that $J(x_N, N) = x_N^T S x_N$

At this point the idea is that we can start "from the end" and see first the strategy with horizon 1, then 2, and up to the horizon N, but "starting from the end".

Let $V(i, x) = \min_{u(i,N)} J(x, u_{[i,N]})$ we know that for the Bellman principle

$$V(i, x) = \min_{u(i,N)} \frac{1}{2}\left(\sum_{k=i}^{j} x_k^T Q x_k + u_k^T R u_k\right) + V(j, x_j)$$

we have to find a strategy that optimizes

$$V(i, x) = \min_{u_i} x_i^T Q x_i + u_i^T R u_i + V(i+1, Ax_i + Bu_i)$$

where moreover

$$V(N, x) = x^T S x$$

<u>Exercise 27:</u> Determine the optimal u_{N-1} in $V(N-1, x)$ and determine the optimal cost of $V(N-1, x)$. Can you write this cost as a quadratic form ? [hint: use the gradient and set $P_N = S$]

<u>Exercise 28:</u> Determine the optimal u_i in the generic $V(i, x)$ [hint: try to use matrices P_i]

Solution of Exercise 27:

$$V(N-1,x)=\min_{u_{N-1}}\frac{1}{2}\left[x_{N-1}^TQx_{N-1}+u_{N-1}^TRu_{N-1}+\left(Ax_{N-1}+Bu_{N-1}\right)^TP_N\left(Ax_{N-1}+Bu_{N-1}\right)\right]$$

Doing the gradient we obtain:

$$\left(R+B^TP_NB\right)u_{N-1}+B^TP_NAx_{N-1}=0$$

And then the minimum is

$$\left(R+B^TP_NB\right)u_{N-1}+B^TP_NAx_{N-1}=0$$

And then

$$u_{N-1}=-\left(R+B^TP_NB\right)^{-1}B^TP_NAx_{N-1}=-F_{N-1}x_{N-1}$$
$$F_{N-1}=\left(R+B^TP_NB\right)^{-1}B^TP_NA$$

The optimal cost is

$$V(N-1,x)=\frac{1}{2}\left[x_{N-1}^TQx_{N-1}+x_{N-1}^TF_{N-1}^TRF_{N-1}x_{N-1}+x_{N-1}^T\left(A+BF_{N-1}\right)^TP_N\left(A+BF_{N-1}\right)x_{N-1}\right]=$$
$$\frac{1}{2}x_{N-1}^T\left(Q+F_{N-1}^TRF_{N-1}+\left(A+BF_{N-1}\right)^TP_N\left(A+BF_{N-1}\right)\right)x_{N-1}=$$
$$\frac{1}{2}x_{N-1}^T\left(Q+F_{N-1}^TRF_{N-1}+\left(A+BF_{N-1}\right)^TP_N\left(A+BF_{N-1}\right)\right)x_{N-1}$$

Please note that this is nothing but the **Riccati equation**

$$P_{N-1}=\left(Q+F_{N-1}^TRF_{N-1}+\left(A+BF_{N-1}\right)^TP_N\left(A+BF_{N-1}\right)\right)$$

Solution of exercise 28:

Similarly to before we have

$$P_i=\left(Q+F_i^TRF_i+\left(A+BF_i\right)^TP_{i+1}\left(A+BF_i\right)\right)$$
$$F_i=\left(R+B^TP_{i+1}B\right)^{-1}B^TP_iA$$
$$P_N=S$$

Moreover it comes for free in this case that the optimal cost is

$$V(N,x_0)=\frac{1}{2}x_0P_0x_0$$

Exercise 34: Solve (in matlab) the "computer education problem" for a horizon of 100,500 and 1000 years. Give a look to the first F_k,P_k you obtain and try to guess something about what happens if the horizon goes to infinity.

Experience 11 – Optimal Control for LTI DT Systems – Infinite Horizon

The question is: what happens as N goes to infinity ?

We consider as a problem the one without terminal cost

$$\lim_{N\to\infty} J_N = \frac{1}{2}\left(\sum_{k=0}^{N} x_k^T Q x_k + u_k^T R u_k\right)$$

where $R = R' > 0,\ Q = Q' \geq 0$

which is equivalent to write everything as:

$$\lim_{N\to\infty} J_N = \frac{1}{2}\left(\sum_{k=0}^{N} x_k^T Q x_k + u_k^T R u_k\right) + \frac{1}{2} x_N^T S x_N$$

with $S = 0$. This is without loss of generality for any $S \in span(Q^{1/2})$.

We know that the solution for the finite horizon would be:

1) set $P_N = S$
2) Iterate N times the Riccati iteration from N to 0
$$P_k = Q + A^T P_{k+1} A - A^T P_{k+1} B\left(R + B^T P_{k+1} B\right)^{-1} B^T P_{k+1} A$$
3) Apply $u_k = -F_k x_k$ where $F_k = \left(R + B^T P_{k+1} B\right)^{-1} B^T P_{k+1} A$

In our case we should start to infinity and iterate infinite time to arrive to the first action. P_0, K_0. Note that (for the classical Cantor paradox), if P_0, K_0 then $P_0 = P_i = P, K_0 = K_i = K$ for all i>0.

To obtain P_0, K_0 let us **"inverse" the sense of the Riccati iteration** and let

$$P^0 = S$$

$$P^{j+1} = Q + A^T P^j A - A^T P^j B\left(R + B^T P^j B\right)^{-1} B^T P^j A$$

Clearly the problem solution is obtained by iterating continuously

$$\lim_{j\to\infty} P^{j+1} = Q + A^T P^j A - A^T P^j B\left(R + B^T P^j B\right)^{-1} B^T P^j A$$

<u>**Exercise 29:**</u> Try to understand (numerically) under which conditions this limit admits a solution and when it does not *[hint: try for instance to see what happens if you don't have stabilizability]*

We will use the following Lemma

Prop. Consider a sequence of matrices $\{P_j\}_{j=0}^{\infty}$ such that:

1. Each matrix P_j is symmetric and positive semidefinite $P_j = P_j^T \geq 0$
2. $\{P_j\}_{j=0}^{\infty}$ is monotonically non-decreasing $i \leq j \Rightarrow P_i \leq P_j$
3. $\{P_j\}_{j=0}^{\infty}$ admits an upper bound, i.e. a matrix Π exists such that $P_j \leq \Pi, \forall j$

Then a limit matrix exists and $\lim_{j\to\infty} P_j = P$

This result allows one to state the following result:

Theorem Consider the matrix sequence P^j obtained by iterating

$$P^0 = 0_{n\times n}$$

$$P^{j+1} = Q + A^T P^j A - A^T P^j B\left(R + B^T P^j B\right)^{-1} B^T P^j A$$

Then, if (A,B) is a stabilizable pair, there exists the limit of

$$\lim_{j\to\infty} P^j = P$$

which is symmetric nonnegative definite and satisfies the Discrete Algebraic Riccati Equation (DARE)

$$P = Q + A^T PA - A^T PB\left(R + B^T PB\right)^{-1} B^T PA$$

Under the above circumstances, the infinite–horizon or steady–state Linear Quadratic Regulator (LQR) for which

$$\lim_{N\to\infty} \min J_N = x^T(0)Px(0)$$

And the optimal control is

$$u_k = -Fx_k$$

where $F = \left(R + B^T PB\right)^{-1} B^T PA$

proof.

It is enough to note that:

1)if $P^j \geq 0$

Then

$$P^{j+1} = Q + F_k^T RF_k + \left(A - BF_k\right)^T P^j \left(A - BF_k\right) \geq 0$$

2)

We know that $J_{N+1}(x(0)) = x^T(0)P^{j+1}x(0) \geq x^T(0)P^j x(0) = J_N(x(0)), \forall x(0)$ which implies

$$P^{j+1} \geq P^j$$

3)

We know that at each time the optimal cost is

$$J_N(x(0)) = x^T(0)P^N x(0)$$

is the minimal cost one can obtain. Then this is for sure minor than the cost obtained with a given control law $u_k = Kx_k$. If (A,B) is stabilizable, let us choose a generic control law such that $A + BK$ has asymptotically stable eigenvalues.

The cost would be:

$$J_N^* = x(0)P^N x(0) \le J_{N,A+BK} = \left(\sum_{k=0}^{N} x_k^T Q x_k + u_k^T R u_k\right) \le \lim_{N\to\infty}\left(\sum_{k=0}^{N} x_k^T Q x_k + u_k^T R u_k\right)$$

$$\lim_{N\to\infty}\left(\sum_{k=0}^{N} x_k^T Q x_k + u_k^T R u_k\right) = \lim_{N\to\infty}\left(\sum_{k=0}^{N} x_k^T \left(Q + F^T R F\right) x_k\right)$$

In general $x_k = (A+BF)^k x_0$ then:

$$\lim_{N\to\infty}\left(\sum_{k=0}^{N} x_k^T \left(Q + F^T R F\right) x_k\right) = \lim_{N\to\infty} x_0^T\left(\sum_{k=0}^{N}\left[(A+BF)^k\right]^T \left(Q+F^T RF\right)(A+BF)^k\right)x_0$$

Let $\widetilde{A} = A + BF, \Phi = \left(Q + F^T RF\right)$

$$\sum_{k=0}^{\infty}\left(\widetilde{A}^k\right)^T \Phi \widetilde{A}^k = \Phi + \widetilde{A}^T\left(\sum_{k=0}^{\infty}\left(\widetilde{A}^k\right)^T \Phi \widetilde{A}^k\right)\widetilde{A}$$

Let

$$\Pi = \left(\sum_{k=0}^{\infty}\left(\widetilde{A}^k\right)^T R \widetilde{A}^k\right)$$

Then

$$\Pi = \Phi + \widetilde{A}^T \Pi \widetilde{A}$$

That is

$$\widetilde{A}^T \Pi \widetilde{A} - \Pi = -\Phi$$

which for sure being a Lyapunov equation admits only one solution which is moreover positive definite solution if $\widetilde{A}$ is asymptotically stable. At this points the three conditions of the lemma are satisfied and then we can say that the series of the P^j converge.

Remark: It is worth to note that the above theorem:

1) Is only a **sufficient condition**
2) **does not ensure the stability** of $x_{k+1} = (A - BF)x_k$

In order to complete the analysis please note that it is possible to rewrite the LQR (Linear Quadratic Regulator) problem as a **Linear Quadratic Output Regulator**, i.e.

$$\min\ y_k^T y_k + u_k^T R u_k$$
$$x_{k+1} = Ax_k + Bu_k$$
$$y_k = Cx_k$$

That is equivalent to the previous one where $C^T C = Q$.

The following results can be proved:

Lemma: Consider the time-invariant LQOR problem and the corresponding matrix sequence P^j generated by the Riccati iterations with $P^0 = 0$. Let $\Sigma_0 = \left(A_0, B_0, C_0\right)$ the completely observable subsystem obtained via canonical observability decomposition of the plant $\Sigma = (A, B, C)$. By denoting with $A_{o\bar{r}}$ the unreachable subsystem of the observable subsystem, **then the limit** $P = \lim_{j\to\infty} P^j$ **exists if and only if the eigenvalues of** $A_{o\bar{r}}$ **are asymptotically stable**.

Theorem: Consider the time-invariant LQOR problem and the corresponding matrix sequence P^j generated by the Riccati iterations with $P^0 = 0$, then the limit $P = \lim_{j\to\infty} P^j$ exists such that the corresponding optimal control is

$u_k = -Fx_k$ where $F = \left(R + B^T PB\right)^{-1} B^T PA$

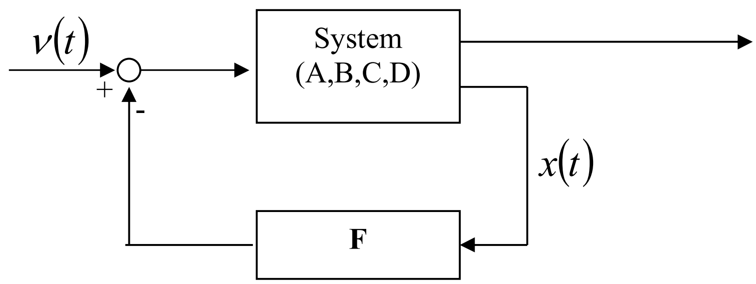
and stabilizes the system $x_{k+1} = (A - BF)x_k$ **if and only if the plant** $\Sigma = (A, B, C)$ **is stabilizable and detectable**. Moreover under the same conditions

1) Whatever is $P^0 = \left(P^0\right)^T \geq 0$ the limit $\lim_{j\to\infty} P^j$ has the same solution P
2) The ARE admits a unique symmetric non-negative solution coinciding with the above limit

Exercise 30 In Matlab using the function DARE or DLQR compute the infinite horizon optimal solution to the "computer education problem"

Experience 12 – Quadratic Optimal Control of a servo-system

In the last experiences we have introduced the classical LQR



This is a powerful control design method that also allows us to work with MIMO systems. However note that in the seen scheme the **reference signal does not appear**. In line of principle we could add a variable $\nu(t)$ to manage the reference.

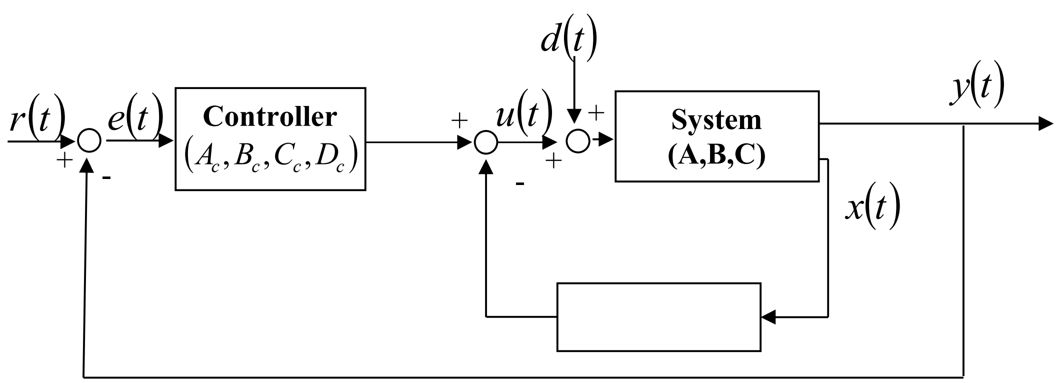
In general in control, besides **regulation to zero** we have other two more general classic control problems:

- **regulation to a set-point:** where we want to be able to stabilize the output to a constant set-point $r(t) = r$
- **trajectory tracking:** where we would like to make the output follow an arbitrary $r(t)$

It is clear that to face these more general problems we need to learn **how to manage** $\nu(t)$ to make the output $y(t)$ to follow a possible reference signal $r(t)$. Many techniques exist.

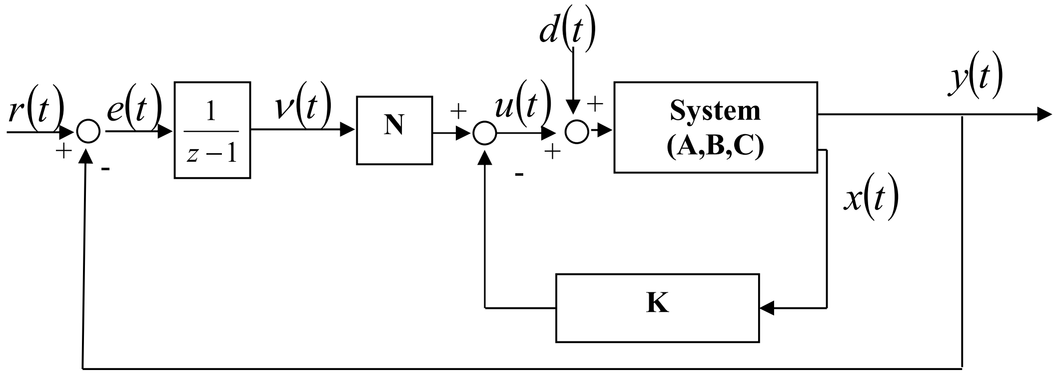
One of the most interesting is in the case we want to reach a steady state. A classical scheme to use the seen state feedback and to have some output tracking performance is to use a control schemes with two loops:

- an internal state loop for stabilization
- an external output loop for the regulation



This scheme gives us the flexibility to perform some trajectory tracking.
In particular we will detail a special case for the **regulation to a set-point** where it makes sense to use optimal LQR. In the classical input/output design, we know that the

zero error to a constant reference is obtained by using an integral action in the compensator. In a similar way it is possible to add an integrator in the control action by considering the following scheme



Note that in this scheme we have added a possible disturbance signal $d(t)$.

Let us analyze the signal after the integrator, being an integrator it can be written as
$v(t+1)=v(t)+e(t)=v(t)+r(t)-y(t)=v(t)+r(t)-Cx(t)$
The state evolution is
$x(t+1)=Ax(t)+Bu(t)+Bd(t)$
where $u(t)=Nv(t)-Kx(t)$ and then
$x(t+1)=Ax(t)+BNv(t)-BKx(t)+Bd(t)$

By composing the above two recurrent equations we obtain the following system
$$\begin{bmatrix} x(t+1) \\ v(t+1) \end{bmatrix} = \begin{bmatrix} A & 0 \\ -C & I \end{bmatrix} \begin{bmatrix} x(t) \\ v(t) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} \begin{bmatrix} -K & N \end{bmatrix} \begin{bmatrix} x(t) \\ v(t) \end{bmatrix} + \begin{bmatrix} Bd(t) \\ r(t) \end{bmatrix}$$
That we can rewrite as
$$x_{ext}(k+1)=(A_{ext}+B_{ext}F_{ext})x_{ext}(k)+\begin{bmatrix} Bd(k) \\ r(k) \end{bmatrix}$$
where
$$x_{ext}(t)=\begin{bmatrix} x(t) \\ v(t) \end{bmatrix},\ A_{ext}=\begin{bmatrix} A & 0 \\ -C & I \end{bmatrix},\ B_{ext}=\begin{bmatrix} B \\ 0 \end{bmatrix},\ F_{ext}=\begin{bmatrix} -K & N \end{bmatrix}$$

Note that we can **arbitrarily move the reachable eigenvalues** of the closed loop system by using the standard techniques.

In fact, once F_{ext} is determined so as to stabilize the closed loop system and to answer some transient response. Moreover since there is an integrator in the control (exactly like in the basic control courses) we may ensure that we are able **to follow the step reference** with zero error and **to reject constant disturbances**.

Remark 1: It is worth to remark that this scheme, while accomplishing the goal **introduces a delay**, in fact note that $v(t+1)$ depends on $e(t)$ and not on $e(t+1)$. This can be avoided but the control structure is more complicated

Remark 2: this method always works for **the case of single output systems with nonzero transfer function** but there are cases (depending on how the system is done) that the latter scheme does not work. An example is the case in which the number of output is higher than the dimension of the input.

Exercise 31: Consider a system

$x_1(t+1) = 0.9x_1(t) + x_2(t)$

$x_2(t+1) = 0.8x_2(t) + u(t)$

$y(t) = x_1(t)$

Design a controller using the above explained method

Experience 13 – Optimal filtering

Some general result

The dual of optimal control is the problem of "compute **the best possible estimation of the state**" from **noisy measurements**. The performance index is usually given in terms of the **covariance of the estimation**.

It is out of the scope of this course to enter in the detail on optimal estimation but some remarks are in order.

In particular let us start from the following fundamental theorem:

THEOREM: Given $x \in R^n, y \in R^m$ two random variables and let $g: R^m \to R^n$ a measurable function. Define $\hat{x}_g = g(y)$ an **estimator of** x **,** $\hat{x}_g$ the estimate and $e_g = x - g(y) = x - \hat{x}_g$ the estimation error.

The ***conditioned expectation*** of the variable x w.r.t. y, i.e. $\hat{x} = \hat{g}(y) = E[x \mid y]$ is the **minimum mean square error** estimator in fact

$$E\left[(x-\hat{x})(x-\hat{x})^T\right] \le E\left[(x-\hat{x}_g)(x-\hat{x}_g)^T\right], \forall g(\cdot)$$

and

$$E\left[\| x-\hat{x} \|_Q^2\right] \le E\left[\| x-\hat{x}_g \|_Q^2\right], \forall g(\cdot), \forall Q \ge 0$$

Moreover it is possible to prove that:

- The error of the optimal estimator $e = x - \hat{x}$ and $\hat{g}(y)$ are uncorrelated

$$E\left[e\hat{g}(y)^T\right] = 0$$

- The optimal estimator is unbiased $E[\hat{x}] = E[g(y) \mid y] = E[x]$

Remark: In the general case the optimal estimator is a nonlinear function of the measurements and then is not very useful in practice.

However in the case of **white Gaussian** processes and **for certain systems** (e.g. linear) this optimal estimator can be computed explicitly. The well-known Kalman filter is nothing but this kind of estimator shaped in a recursive form.

Linar case - The Kalman filter

In particular let us consider the following plant

$$x_{k+1} = Ax_k + Bu_k + v_k^x$$

$$y_k = Cx_k + v_k^y$$

where

- x_k is the state (to be estimated)
- y_k is the measured output
- v_k^x is a **disturbance** on the process and it is assumed to be **Gaussian white noise** with expected value $E(v_k^x) = 0_m$ and covariance matrix

$$E\left(\left[v_k^x - 0_n\right]\left[v_k^x - 0_n\right]^T\right) = Q = Q^T \ge 0$$

- v_k^y is **noise** on the measurements and it is assumed to be **Gaussian white noise** with expected value $E(v_k^y) = 0_p$ and covariance matrix $E\left([v_k^y - 0_p][v_k^y - 0_p]^T\right) = R = R^T > 0$
- The **initial state** x_0 is a **gaussian random variable with expected value** $\bar{x}_0$ **and covariance** $E\left([x_0 - \bar{x}_0][x_0 - \bar{x}_0]^T\right) = \hat{P}_0$

The processes v_k^x, v_k^y, x_0 are assumed independent.

Let us define:
1) $\hat{x}_{k|h} = E[x_k \mid y_0,\ldots,y_h,u_0,\ldots,u_h]$
2) $P_{k|h} = E\left[(x_k - \hat{x}_{k|h})(x_k - \hat{x}_{k|h})^T \mid y_0,\ldots,y_h,u_0,\ldots,u_h\right]$

The **optimal estimator** is the Kalman filter that can be defined as two steps:

1. Prediction

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + Bu_k$$

$$P_{k+1|k} = AP_{k|k}A^T + Q$$

2. Correction

$$K_{k+1|k+1} = P_{k+1|k}C^T\left(CP_{k+1|k}C^T + R\right)^{-1}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1|k+1}\left(y_{k+1} - C\hat{x}_{k+1|k}\right)$$

$$P_{k+1|k+1} = P_{k+1|k} - K_{k+1|k+1}CP_{k+1|k}$$

With **initial conditions:**

$$\hat{x}_{0|-1} = E[x_0]$$

$$P_{0|-1} = \hat{P}_0$$

Some remark:

1) The process can be Linear Time varying, it is enough to use $A_k, B_k, C_{k+1}, R_{k+1}, Q_k$
2) It is possible to extend to the case v_k^x, v_k^y are correlated
3) R can be also positive-semidefinite (the inversion becomes a pseudoinverse)
4) The fact the expected values of the noise is 0 is not fundamental and can be relaxed adding an offset in the predictions and in the estimation
5) For optimality, whiteness cannot be easily relaxed as well as the fact that noise and disturbances are Gaussian

Exercise 32: Using the latter equations try to write $P_{k+1|k}$ in function of $P_{k|k-1}$ and using previous try to infer something of what happens as time goes on.

Exercise 33: Consider a system you like with $n = 4, m = 1, p = 2$ and apply a Kalman filter to it

Experience 14 – LQG

The LQG control problem that can be formulated as follows

Consider the **plant**

$$x_{k+1} = Ax_k + Bu_k + v_k^x$$
$$y_k = Cx_k + v_k^y$$

where

- x_k is the state (to be estimated)
- y_k is the measured output
- v_k^x is a disturbance on the process and it is assumed to be Gaussian white noise with expected value $E(v_k^x) = 0_m$ and covariance matrix $E\left([v_k^x - 0_n][v_k^x - 0_n]^T\right) = Q = Q^T \geq 0$
- v_k^y is noise on the measurements and it is assumed to be Gaussian white noise with expected value $E(v_k^y) = 0_p$ and covariance matrix $E\left([v_k^y - 0_p][v_k^y - 0_p]^T\right) = R = R^T > 0$
- The initial state x_0 is a gaussian random variable with expected value $\bar{x}_0$ and covariance $E\left([x_0 - \bar{x}_0][x_0 - \bar{x}_0]^T\right) = \hat{P}_0$

The processes v_k^x, v_k^y, x_0 are assumed to be independent.

Determine the optimal control that minimizes the following cost:

$$\min_{u_k = f(I_k)} J_N(\bar{x}_0, P_0) = E\left[x_N^T W_N x_N + \sum_{k=0}^{N-1} x_k^T W x_k + u_k^T U u_k \mid x_0, P_0\right]$$

Where I_k it the information available at time k

Let us assume the state is estimated through an estimator and that $\hat{x}_k$ is the estimate and $P_{k|k}$ the covariance matrix. Remark that in general it is possible to prove that

$$E\left[x_k^T \Phi x_k\right] = trace\left(\Phi P_{k|k}\right) + \hat{x}_k^T \Phi \hat{x}_k$$

Let us use **Dynamic programming** :

$$V_N(x_N) = E\left[x_N^T W_N x_N \mid I_N\right]$$
$$V_k(x_k) = E\left[x_k^T W x_k + u_k^T U u_k + V_{k+1}(x_{k+1}) \mid I_N\right]$$

And let us assume that the "***cost to go***" is $V_{k+1}(x_{k+1}) = E\left[x_{k+1}^T S_{k+1} x_{k+1} \mid I_{k+1}\right] + c_{k+1}$

<u>*At time k*</u>

$$V_k(x_k) = E\left[x_k^T W x_k + u_k^T U u_k + E\left[x_{k+1}^T S_{k+1} x_{k+1} + c_{k+1} \mid I_{k+1}\right] \mid I_k\right] =$$
$$= E\left[x_k^T W x_k + u_k^T U u_k + x_{k+1}^T S_{k+1} x_{k+1} + c_{k+1} \mid I_k\right]$$

At this point it is enough to note that

$$E\left[x_{k+1}^T S_{k+1} x_{k+1} \mid k\right] = E\left[\left(Ax_k + Bu_k + \nu_k\right)^T S_{k+1}\left(Ax_k + Bu_k + \nu_k\right) \mid I_k\right] =$$
$$E\left[x_k^T A^T S_{k+1} Ax_k + u_k^T B^T S_{k+1} Bu_k + 2u_k^T B^T S_{k+1} Ax_k + \nu_k^T S_{k+1} \nu_k \mid I_k\right] =$$
$$= E\left[x_k^T A^T S_{k+1} Ax_k\right] + u_k^T B^T S_{k+1} Bu_k + 2u_k^T B^T S_{k+1} A\hat{x}_k + trace\left(S_{k+1} Q\right)$$
$$= \hat{x}_k^T A^T S_{k+1} A\hat{x}_k + u_k^T B^T S_{k+1} Bu_k +$$
$$+ trace\left(A^T S_{k+1} AP_{k|k}\right) + 2u_k^T B^T S_{k+1} A\hat{x}_k + trace\left(S_{k+1} Q\right)$$

Then

$$V_k\left(x_k\right) = E\left[x_k^T W x_k + u_k^T U u_k + x_{k+1}^T S_{k+1} x_{k+1} + c_{k+1} \mid I_k\right] =$$
$$= \hat{x}_k^T\left[A^T S_{k+1} A + W\right]\hat{x}_k + u_k^T\left[B^T S_{k+1} B + U\right]u_k + 2u_k^T B^T S_{k+1} A\hat{x}_k +$$
$$+ trace\left(\left(A^T S_{k+1} A + W\right)P_{k|k}\right) + trace\left(S_{k+1} Q\right) + E\left[c_{k+1} \mid I_k\right]$$

The optimal control action then is

$$\frac{\partial V_k\left(x_k\right)}{\partial u_k} = 2\left[B^T S_{k+1} B + U\right]u_k + 2B^T S_{k+1} A\hat{x}_k = 0$$
$$u_k = -\left[B^T S_{k+1} B + U\right]^{-1} B^T S_{k+1} A\hat{x}_k$$

Substituting back we obtain

$$V_k\left(x_k\right) = \hat{x}_k^T\left[A^T S_{k+1} A + W\right]\hat{x}_k - x_k^T AS_{k+1} B\left[B^T S_{k+1} B + U\right]B^T S_{k+1} A\hat{x}_k +$$
$$+ trace\left(\left(A^T S_{k+1} A + W\right)P_{k|k}\right) + trace\left(S_{k+1} Q\right) + E\left[c_{k+1} \mid I_k\right] =$$
$$= x_k^T\left[A^T S_{k+1} A + W - AS_{k+1} B\left[B^T S_{k+1} B + U\right]B^T S_{k+1} A\right]\hat{x}_k +$$
$$+ trace\left(\left(A^T S_{k+1} A + W\right)P_{k|k}\right) + trace\left(S_{k+1} Q\right) + E\left[c_{k+1} \mid I_k\right]$$

And we have that:

$$S_k = \left[A^T S_{k+1} A + W - AS_{k+1} B\left[B^T S_{k+1} B + U\right]B^T S_{k+1} A\right]$$
$$c_k = trace\left(\left[A^T S_{k+1} A + W\right]P_{k|k}\right) + trace\left(S_{k+1} Q\right) + E\left[c_{k+1} \mid I_k\right]$$

Note that:

1) The **optimal control** is independent from the **estimation and depend only on $\hat{x}_{k|k}$ (separation principle). Note that the solution is the same of the LQR !**
2) the **optimal estimator** is the **Kalman filter**
3) Note that both the Kalman filter and the control have Riccati iterations. It is possible to prove that if the two Riccati converges then the cost $\lim_{N\to\infty} \frac{1}{N} J_N$ converges.

Experience 15 – Continuous Time

The results for what regard LQR, Kalman and LQG also holds true in continuous time.

Example: LQR

1)Plant and cost

$\dot{x} = Ax + Bu$

$J = \frac{1}{2}x^T(t_1)Q_f(t_1 x(t_1)) + \int_{t_0}^{t_1} x^T Qx + u^T Ru \ dt$

2)Optimal control

$u = -R^{-1}B^T P(t)$

3)Riccati Iteration

$A^T P(t) + P(t)A - P(t)BR^{-1}B^T P(t) + Q = -\dot{P}(t)$

3)Cost

$x^T(0)P(0)x(0)$

4)Infinite Horizon

$A^T P + PA - PBR^{-1}B^T P + Q = 0$

$u = -R^{-1}B^T P$

Similarly one can found the equation for the Kalman. LQG again satisfy the separation principle.

Experience 15 – Geometry and Optimization

Exercise 34 Consider the admissible region resulting from the convex hull of the points

$$p_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, p_2 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}, p_3 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, p_4 = \begin{bmatrix} -1 \\ -2 \end{bmatrix}, p_5 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

a) If the resulting optimization problem is convex, determine the minimal euclidean distance between this admissible region and the set

$$\left(x - \begin{bmatrix} 20 \\ 20 \end{bmatrix}\right)^T \begin{bmatrix} 10 & 4 \\ 4 & 10 \end{bmatrix} \left(x - \begin{bmatrix} 20 \\ 20 \end{bmatrix}\right) \le 3$$

b) If the resulting optimization problem is convex, determine the maximal euclidean distance between this admissible region and the set

$$\left(x - \begin{bmatrix} 20 \\ 20 \end{bmatrix}\right)^T \begin{bmatrix} 10 & 4 \\ 4 & 10 \end{bmatrix} \left(x - \begin{bmatrix} 20 \\ 20 \end{bmatrix}\right) \le 3$$

Exercise 35

Plot the sets

- $\left(x - \begin{bmatrix} 20 \\ 20 \end{bmatrix}\right)^T \begin{bmatrix} 10 & 4 \\ 4 & 10 \end{bmatrix} \left(x - \begin{bmatrix} 20 \\ 20 \end{bmatrix}\right) \le 3$
- the convex hull of $p_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, p_1 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}, p_3 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, p_4 = \begin{bmatrix} -1 \\ -2 \end{bmatrix}, p_5 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Exercise 36

a) If the resulting optimization problem is convex, Determine the minimal distance between the line $x_1 = x_2 + 100$and

$$\left(x - \begin{bmatrix} 20 \\ 20 \end{bmatrix}\right)^T \begin{bmatrix} 10 & 4 \\ 4 & 10 \end{bmatrix} \left(x - \begin{bmatrix} 20 \\ 20 \end{bmatrix}\right) \le 3$$

Exercise 37

a) If the resulting optimization problem is convex, determine the minimum radius circle that contains the convex hull of the points

$$p_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, p_1 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}, p_3 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, p_4 = \begin{bmatrix} -1 \\ -2 \end{bmatrix}, p_5 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

b) Is it possible to solve the same problem without solving a convex optimization problem ? How ?

Exercise 38

If the resulting optimization problem is convex, determine the maximum radius circle that is contained in the convex hull of the points

$$p_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, p_1 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}, p_3 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, p_4 = \begin{bmatrix} -1 \\ -2 \end{bmatrix}, p_5 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

[hint:** try to represent the polyhedron as a set of inequalities**]

Exercise 39

Consider the intersection of the following regions

$$x^T\begin{bmatrix}5 & 2\\ 2 & 10\end{bmatrix}x \le 1$$

$$x^T\begin{bmatrix}15 & -2\\ -2 & 10\end{bmatrix}x \le 1$$

If the resulting optimization problem is convex, determine the minimal and the maximal distance with respect to the set

$$99 \le x_1 \le 100$$

$$4 \le x_2 \le 4$$

Exercise 40

If the resulting optimization problem is convex, determine the maximal distance between two points belonging to the following set

$$x^T\left(\lambda_1\begin{bmatrix}5 & 2\\ 2 & 10\end{bmatrix} + (1-\lambda_1)\begin{bmatrix}15 & -2\\ -2 & 10\end{bmatrix}\right)x \le 1, \forall 0 \le \lambda_1 \le 1$$

Exercise 47

Consider the constraint

$$\left(x_1\begin{bmatrix}5 & 2\\ 2 & 10\end{bmatrix} + x_2\begin{bmatrix}-15 & -2\\ -2 & -10\end{bmatrix}\right) > 0$$

where > means "positive definite". Is this constraint convex ?

Under the further constraint $(x_1 + x_2) < 1$ find x_1, x_2 that maximize x_2

Exercise 41

Consider the ellipsoid

$$x^T\left(x_1\begin{bmatrix}5 & 2\\ 2 & 10\end{bmatrix} + x_2\begin{bmatrix}-15 & -2\\ -2 & -10\end{bmatrix}\right)x < 1$$

Obtained under the condition

$$\left(x_1\begin{bmatrix}5 & 2\\ 2 & 10\end{bmatrix} + x_2\begin{bmatrix}-15 & -2\\ -2 & -10\end{bmatrix}\right) > 0$$

Determine the solution that minimize the trace of $\left(x_1\begin{bmatrix}5 & 2\\ 2 & 10\end{bmatrix} + x_2\begin{bmatrix}-15 & -2\\ -2 & -10\end{bmatrix}\right) > 0$

And such that the points $x_1 = \begin{bmatrix}1\\ 2\end{bmatrix}, x_2 = \begin{bmatrix}5\\ 7\end{bmatrix}$ belong to that ellipsoid

Exercise 42

For all the exercises try to plot the admissible regions and to explain graphically the optimal solution

Experience 16 – Introduction to LMI

Definition An **affine combination of matrices** is a linear combination of matrices plus a constant matrix

$$M(x) = M_0 + \sum_{i=1}^{m} M_i x_i$$

With M_i matrices

Roughly speaking: An affine combination of matrices s a matrix $M(x)$ such that each entry of the matrix is an affine function of x.

Notation: It is usual to not show explicitly single variables but to introduce variables of matrices, for instance, let $P = P^T \in R^{n\times n}$ a symmetric variable matrix and $A \in R^{n\times n}$, then $A^T P + PA$ is an affine combination of matrices

Definition: A **Linear Matrix Inequality** is an affine combination of symmetric square matrices constrained to be positive definite. Let $F_i = F_i^T \in R^{n\times n}, i = 0,...,m$ be given and let $x \in R^m$ a vector of unknown, an LMI is

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i > 0$$

Remark 1: An LMI can be defined also w.r.t. to the negative definiteness, $F(x) < 0$ is also an LMI.

Remark 2: It is possible also to consider semidefinite conditions, $F(x) \geq 0, F(x) \leq 0$. There are formulas to pass from definite to semidefinite conditions.

Exercise 43 Prove that an LMI defines a convex region

Exercise 44 Prove that given a set of LMI

$$L_1(x) > 0$$

$$...$$

$$L_N(x) > 0$$

it can be rewritten as a single LMI condition

Exercise 45 Given a square full-rank matrix $Q \in R^{n\times n}$, prove that the admissible region of $L(x) > 0$ and the one of $Q^T L(x) Q > 0$ coincide

Experience 17 – Schur Complements

There are many tricks to manipulate LMIs. One of the most important are the Schur complements

LEMMA 1: The following Matrix Inequalities are equivalent

1) $\begin{bmatrix} Q(x) & S(x) \\ S(x)^T & R(x) \end{bmatrix} > 0$

2) $R(x) > 0, \quad Q(x) - S(x)R^{-1}(x)S^T(x) > 0$

3) zx\zzzzzzzzzzzzza`

LEMMA 2: The following Matrix Inequalities are equivalent

1) $\begin{bmatrix} Q(x) & S(x) \\ S(x)^T & R(x) \end{bmatrix} < 0$

2) $R(x) < 0, \quad Q(x) - S(x)R^{-1}(x)S^T(x) < 0$

3) $Q(x) < 0, \quad R(x) - S^T(x)Q^{-1}(x)S(x) < 0$

LEMMA 3: The following Matrix Inequalities are equivalent

1) $\begin{bmatrix} Q(x) & S(x) \\ S(x)^T & R(x) \end{bmatrix} \geq 0$

2) $R(x) \geq 0, \quad Q(x) - S(x)R^{+}(x)S^T(x) \geq 0, \; S(x)\big(I - R(x)R^{+}(x)\big) = 0$

3) $Q(x) \geq 0, \quad R(x) - S^T(x)Q^{+}(x)S(x) \geq 0, \big(I - Q(x)Q^{+}(x)\big)S(x) = 0$

LEMMA 4: The following Matrix Inequalities are equivalent

1) $\begin{bmatrix} Q(x) & S(x) \\ S(x)^T & R(x) \end{bmatrix} \leq 0$

2) $R(x) \leq 0, \quad Q(x) - S(x)R^{+}(x)S^T(x) \leq 0, \; S(x)\big(I - R^{+}(x)R^T(x)\big) = 0$

3) $Q(x) \leq 0, \quad R(x) - S^T(x)Q^{+}(x)S(x) \leq 0, \; \big(I - Q(x)Q^{+}(x)\big)S(x) = 0$

Exercise 46 Consider the following matrix inequality $Z(x)Z^T(x) < I$ where $Z(x)$ is an affine combination of matrices depending on x. Using the Schur complements try to make this an LMI

Exercise 47 Consider the following matrix inequality $A^T P + PA + PBR^{-1}B^T P + Q < 0$ where $P = P^T > 0$ is a matrix variable and $R > 0$. Using the Schur complements try to make this an LMI

Exercise 48 Consider the following matrix inequalities

$P - (A + BK)^T P(A + BK) > 0, \quad P > 0$

a) If $P = P^T$ is a matrix variable and K is a matrix variable. Using the Schur complements try to make this an LMI (*hint:* if you find at a certain point P^{-1} somewhere try to pre and post multiplying for something !). Have you ever seen this kind of conditions ?

Experience 18 - Problems solvable via LMI

1. LMI Problem (LMIP)

Given an LMI $F(x) > 0$, an **LMI Problem** o **LMIP** consists in finding a x_{feas} such tat $F(x_{feas}) > 0$.

This is an convex **admissibility problem**

Exercise 49 Given a polytopic uncertain system

$$x(t+1) = \left(\sum_{i=1}^{l} p_i(t) A_i\right) x(t)$$

Where

$$0 \le p_i(t) \le 1$$

$$\sum_{i=1}^{l} p_i(t) = 1$$

find a Lyapunov function $V(x) = x^T P x$ that proves the asymptotic stability of the system

Exercise 50 Given a polytopic uncertain system

$$x(t+1) = \left(\sum_{i=1}^{l} p_i(t) A_i\right) x(t) + \left(\sum_{i=1}^{l} p_i(t) B_i\right) u(t)$$

find a stabilizing control law $u(t) = Fx(t)$ and a Lyapunov function $V(x) = x^T P x$ that proves the asymptotic stability of the system

Exercise 51 Given a polytopic Linear Parameter Varying (LPV) system

$$x(t+1) = \left(\sum_{i=1}^{l} p_i(t) A_i\right) x(t) + \left(\sum_{i=1}^{l} p_i(t) B_i\right) u(t)$$

i.e. a polytopic uncertain system where the parameter $p(t)$ is measurable at each time instant, find a stabilizing control law $u(t) = \sum_{i=1}^{l} p_i(t) F_i x(t)$ and a Lyapunov function $V(x) = x^T P x$ that proves the asymptotic stability of the system

The EVP (EigenValue Problem)

The EVP consists in the minimization of the highest eiganvalue of a matrix that depends affinely on a variable and it is constrained through a LMI constraints:

$$\min \lambda$$

$$\begin{cases} \lambda I - A(x) > 0 \\ B(x) > 0 \end{cases}$$

Where A e B are affine combination w.r.t. x. This is a convex optimization problem which can be rewritten as

$$\min c^T x$$

$$F(x) > 0$$

Or equivalently

$$\min \lambda$$
$$A(x,\lambda) > 0$$

Where $A(x,\lambda)$ is affine w.r.t. (x,λ)

Example:

$$\min \gamma^2$$
$$\begin{bmatrix} -A^T P - PA - C^T C & PB \\ B^T P & \gamma^2 I \end{bmatrix} > 0$$
$$P > 0$$

The GEVP (Generalized EVP)

The GEVP is the problem of minimizing something in the form

$$\min \lambda$$
$$\begin{cases} \lambda B(x) - A(x) > 0 \\ B(x) > 0 \\ C(x) > 0 \end{cases}$$

Which can be expressed as

$$\min \lambda_{\max}\left(A(x), B(x)\right)$$
$$\begin{cases} B(x) > 0 \\ C(x) > 0 \end{cases}$$

Were $\lambda_{max}(X,Y)$ is the biggest eigenvalue of $\lambda Y - X > 0$ con $Y > 0$.

A further way to describe the GEVP is the notation:

$$\min \lambda$$
$$A(x,\lambda) > 0$$

Where A is affine w.r.t. x for a given λ and affine w.r.t. λ for a given .

The GEVP is a **quasi-convex problem**

An example of GEVP is the following

The CP

A last class of problem that can be proved to be convex is

$$\min \log\det A^{-1}(x)$$
$$A(x) > 0$$
$$B(x) > 0$$

An example is for instance:

$$\min \log\det P^{-1}$$
$$v_i^T P v_i \le 1 \quad i = 1,\ldots,N$$
$$P > 0$$

where v_i constant vectors and P a variable.

Experience 19 – Ellipsoid

Let's start from exercise on the maximum volume circle that can be defined as
$x:(x-c)^T(x-c)<\gamma^2$

Or equivalently
$\{x \mid (\gamma I)y + c, \| y \| \le 1\}$

If I have a set of constraints $a_i^T x \le b_i, i = 1,...,n$ to constraint the circle inside the constraints it must happen that:

$a_i^T x \le b \Leftrightarrow a_i^T(\gamma I)y + a_i^T c \le b, \forall \| y \| < 1 \Leftrightarrow$
$\sup_{\|y\|\le 1} a_i^T \gamma I \; y \le b - a_i^T c \Leftrightarrow$
$\gamma \| a_i^T I \| + c \le b$

And so we arrive to a convex formulation

In general **it is possible to define ellipsoids in two ways:**

1) $\varepsilon = \{x \mid \| Ax - b \| \le 1\}$ where
   a) the center is $A^{-1}b$
   b) $A = A^T$ and the volume is proportional to $\det A^{-1}$
2) $\varepsilon = \{x = By + d \mid \| y \| \le 1\}$ where
   a) the center is d
   b) $B = B^T$ and the volume is proportional to $\det B$

Exercice 52 Using the above definition and what you know on LMIs compute the minimal volume ellipsoid containing the convex hull of a set of point $x_i, i = 1,...,N$

Exercice 53 Using the above definition and what you know on LMIs compute

a) the maximal volume ellipsoid contained in the constraints $a_i^T x \le b$

Experience 20 – Some other things about ellipsoids

Ellipsoids representation 1)

We have seen that to force that an ellipsoid contains a convex combination of points $x_i, i = 1,...,N$ (a polyhedron) it is enough to use the notation $\varepsilon = \{x \mid \| Ax - b \| \le 1\}$ and force
$\| Ax_i - b \| \le 1, i = 1,...,N$
which is a convex set.

Please note that since the center is $x_c = A^{-1}b$ the notation $\varepsilon = \{x \mid \| Ax - b \| \le 1\}$ is equivalent to another representation we have used that is
$\varepsilon = \{x \mid \| Ax - Ax_c \| \le 1\} = \{x \mid (x - x_c)^T A^2 (x - x_c) \le 1\} = \{x \mid (x - x_c)^T P (x - x_c) \le 1\}$
where we can assume without loss of generality
$P = A^2$ and $P \ge 0$

Remark: Please note that, using Schur complements we can make become
The constraints $\| Ax_i - b \| \le 1, i = 1,...,N$ an LMI as follows

$$\begin{bmatrix} 1 & (x_i A - b)^T \\ (x_i A - b) & I \end{bmatrix} \ge 0, i = 1,..., N$$

or the more complicated $(x_i - x_c)^T P (x_i - x_c) \le 1, i = 1,..., N, P \ge 0$ that under the hypothesis $P > 0$ is equivalent to the following LMI

$$\begin{bmatrix} 1 & (x_i - x_c)^T \\ (x_i - x_c) & P^{-1} \end{bmatrix} \ge 0, i = 1,..., N$$

Ellipsoids representation 2)

We have seen that to force that an ellipsoids satisfies a set of linear constraints $\alpha_i^T x \le \beta_i, i = 1,...,N$ it is enough to use the representation
$\varepsilon = \{x = By + d \mid \| y \| \le 1\}$
And impose
$\|B\alpha_i\|_2 + \alpha_i^T d \le \beta_i$

The relationship between the two representation: Please note that, as already seen $d = x_c$. Moreover please note that

$$\varepsilon = \{x = By + d \mid \| y \| \le 1\} = \{x = By + x_c \mid \| y \| \le 1\} = \{x \ \textit{such that} \ x - x_c = By \mid \| y \| \le 1\} =$$
$$= \{x \ \textit{such that} \ B^{-1}(x - x_c) = y \mid \| y \| \le 1\} = \{x \mid \| B^{-1}(x - x_c) \| \le 1\}$$

And then (when it makes sense) $B = A^{-1} = P^{-1/2}$.

Remark Note that the constraint $\|B\alpha_i\|_2 + \alpha_i^T d \le \beta_i$ can be converted in an LMI as follows:

$$\sqrt{\alpha_i^T B^T B\alpha_i} + \alpha_i^T d \le \beta_i$$

$$\left(\beta_i - \alpha_i^T d\right)^2 - \alpha_i^T B^T B\alpha_i \ge 0$$

And finally

$$\begin{bmatrix} \left(\beta_i - \alpha_i^T d\right)^2 & \alpha_i^T B^T \\ B\alpha_1 & I \end{bmatrix} \ge 0$$

Please note that moreover the latter is equivalent to $\beta_i - \alpha_i^T d - \alpha_i^T P^{-1}\alpha_i \ge 0$, that under the hypothesis of $P > 0$ gives the LMI

$$\begin{bmatrix} \left(\beta_i - \alpha_i^T d\right)^2 & \alpha_i^T \\ \alpha_i & P \end{bmatrix} \ge 0$$

Exercice 54

If $A \ge 0$ or if $B \ge 0$ we have degenerate ellipsoids. Try to understand graphically which kind of degenerate ellipsoid correspond to the case $A \ge 0$ and which one to the case $B \ge 0$

Exercice 55

Determine, if exists and if the resulting optimization problem is convex, an ellipsoid centered in the origin such that

- Contains the polytope $\begin{matrix} -3 \le x_1 \le 3 \\ -5 \le x_2 \le 5 \end{matrix}$
- Is contained in the polyhedral set $-9 < x_1 + x_2 < 9$

Exercice 56 If possible generalize the constraints necessary to determine an ellipsoid centered in the origin contained in a polyhedron P_1 and that contains a polytope P_2

Experience 21 – Positive invariance and Constraned Control

Definition: Given an autonomous system $\left.\begin{array}{l}\dot{x}(t)\\ x(t+1)\end{array}\right\} = f(x(t))$, a set Ω is a **positive invariant set** if $x(t) \in \Omega$ implies $x(t+\tau) \in \Omega, \forall \tau \geq 0$

Exercise 57 Assuming a system is stable (even not asymptotically) and a Lyapunov function proving stability exists, does it exist a set $\Omega \subset \Re^n$ that is positive invariant ?

Exercise 58 Consider a system $x(t+1) = Ax(t) + Bu(t)$ and an initial point $x(0)$. Moreover assume the system is subject to state constraints in the form $\alpha_i x < \beta_i, i = 1,..., N$ and by input constraints in the form $\alpha_{u,i} u < \beta_{u,i}, i = 1,..., N_u$. Using the idea if invariant set formulate as an LMI a possible way to solve the problem of determining a stabilizing linear state feedback control law.

Exercise 59 Invent a 2x2 unstable system and apply the latter formulation to it (solve in matlab). Does the solution depends on the initial point ?

Experience 22 – LQ and LMIs

Let us consider the problem of finding a feedback $u(t)=Fx(t)$ that minimizes the quadratic cost $J(x(0),u(\cdot))=\sum_{i=0}^{\infty}x^T(i)R_x x(i)+u(i)R_u u(i)$
With $R_x=R_x^{\ T}>0, R_u=R_u^{\ T}>0$.

To do so follows the following steps:

it is easy to prove that given a static feedback F and let a generic $P>0$ such that

$$x^T(i+1)Px(i+1)-x^T(i)Px(i)\le -x^T(i)R_x x(i)-u^T(i)R_u u(i)$$

Then

$$\sum_{i=0}^{\infty}x^T(i)R_x x(i)-u^T(i)R_u u(i)\le x^T(0)Px(0)$$

Then a "non strict" (at least in principle) condition to minimize the cost is:

$$\begin{aligned}&\min \gamma\\ &(A+BF)^T P(A+BF)-P+R_x+F^T R_u F\le 0\\ &x^T(0)Px(0)\le\gamma\\ &P>0\end{aligned}$$

Which can be rewritten as

$$\begin{aligned}&\min_{\gamma,P,F}\gamma\\ &P-\begin{bmatrix}(A+BF)^T & R_x^{1/2} & F^T R_u^{1/2}\end{bmatrix}\begin{bmatrix}P&0&0\\0&I&0\\0&0&I\end{bmatrix}\begin{bmatrix}A+BF\\ R_x^{1/2}\\ R_u^{1/2}F\end{bmatrix}\ge 0\\ &\gamma-x^T(0)Px(0)\ge 0\\ &P>0\end{aligned}$$

Exercise 60: Using Schur complements and an appropriate change of basis try to convert the latter problem into an LMI based optimization problem.

Exercise 61: Apply the solution of Exercise 70 to the "*computer education problem*" and compare with the classical LQ results

Exercise 62: Consider a system $x(t+1)=Ax(t)+Bu(t)$ and an initial point $x(0)$. Moreover assume the system is subject to state constraints in the form $\alpha_i x<\beta_i, i=1,\ldots,N$ and by input constraints in the form $\alpha_{u,i}u<\beta_{u,i}, i=1,\ldots,N_u$. Using the ideas of exercise 68 and the LMIs of exercise 70 can we try to find a "LQ state feedback with constraints"

Exercise 63:
Consider the system

$$A = \begin{bmatrix} 1 & 0.4 \\ 0 & 2 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and input constraints $|u| \leq 10$ and implement the last exercise outcomes

- does the solution depend on the initial condition ?
- is this in your opinion the optimal LQ strategy under constraints ? [to answer this question formulate the minimization of the LQ cost using "free inputs" and a finite (but very long) horizon and try to compare solutions.]
- is this in your opinion the optimal linear state feedback ? [as above, call me and discuss with me if you need].

Experience 23 – Model Predictive Control

Exercise 64: make a MPC scheme recomputing at each time step the linear state feedback used in exercise 72.

Exercise 65: make a MPC scheme (following what seen in CSD course), using explicit prediction and a fixed terminal set (computed using the function in exercise 72). Make a code that allows changing the number of explicit moves N. Test it on a system with some "real meaning" and compare its behavior for different N.

Exercise 66: make a MPC scheme (following what seen in CSD course), using explicit prediction and a terminal set that is computed together with the explicit predictions. Make a code that allows changing the number of explicit moves N. Test it on a system with some "real meaning" and compare its behavior for different N.

Experience 24 – Reference Governor

Exercise 67: Compute $\tilde{O}_\infty$ for a system with constraint with some "real meaning" and test the reference governor and the command governor on it.

Experience 25 – Bonus exercises

Exercise 68: Build a function to compute $\tilde{O}_\infty$ in the case of linear systems subject to disturbance and with unmeasured state and a measured output z (use Luenberger observer and assume that you can start with ε_∞ as in the survey)

Exercise 69: Build all what is needed to use the Extended Command Governor

Exercise 70: Try to devise an MPC scheme for polytopic uncertain systems (suggestion: look at the Casavola Giannelli Mosca paper on the subject).

Experience 26 – Norms of signals

An interesting way to characterize the performances of a (controlled) system is to characterize the amplitude of some interesting signal. For instance, the effectiveness of a tracking system could be measured by means of the amplitude of the error signal

The easiest way to characterize a signal or a vector of signals is to make use of norms. Consider piece-wise continuous signals and (without loss of generality) assume the signal is zero for t<0.

The properties that a norm must satisfy are:

1) $||u|| \geq 0$
2) $||u||=0 \Leftrightarrow u(t)=0 \; \forall t$
3) Given $a \in \Re$, $\| au \| = | a | \| u \|$
4) $||u+v|| \leq ||u||+||v||$ triangular inequalities

The most used norms for vector of signals are

DEF. 1-Norm

$$\| u \|_1 = \int_{-\infty}^{\infty} \| u(t) \| \, dt$$

DEF. 2-Norm

$$\| u \|_2 = \left(\int_{-\infty}^{\infty} \| u(t) \|^2 \, dt \right)^{1/2}$$

Esempio: Assuming that u is the current of a resistance of 1 Ω, this norm is the instantaneous power. It's integral is the dissipated energy

DEF. n-th norm

$$\| u \|_n = \left(\int_{-\infty}^{\infty} \| u(t) \|^n \, dt \right)^{1/n}$$

DEF. ∞-norm

The ∞- norm of a vector of signals is the minimal upper-bound of its absolute value

$\| u \|_\infty = \max_i \sup_t | u_i(t) |$

DEF. the **power** of a signal u is the square root of the average power

$$pow(u) = \left(\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \| u(t) \|^2 \, dt \right)^{1/2}$$

Please note that **this is not a norm but a semi- norm**. In fact it can be zero even if the signal is not zero. All the other properties hold true

Remark Clearly in discrete time it is the same but we have sums instead of integrals

Experience 27 –Norms of systems – Continuous time

Let us consider continuous time LTI systems.

We know that in the time domain an input-output model for this kind of systems has the form of a convolution product

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

or equivalently in time

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

Where causality implies $g(t) = 0, \forall t < 0$ and $G(s) = C(sI - A)^{-1}B + D$

It is quite easy to define 2 norms for transfer function $\hat{G}(s)$:

DEF. Norm 2 (H2):

$$\left\|\hat{G}(j\omega)\right\|_2 = \left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} tr\{G^*(j\omega)G(j\omega)\}d\omega\right)^{1/2}$$

N.B. If $\hat{G}(s)$ is stable then for the Parseval theorem:

$$\left\|\hat{G}(s)\right\|_2 = \left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} |\hat{G}(j\omega)|^2 \, d\omega\right)^{1/2} = \left(\int_{-\infty}^{+\infty} |g(t)|^2 \, dt\right)^{1/2}$$

DEF. ∞-norm (H∞):

$$\left\|\hat{G}(j\omega)\right\|_\infty = \sup_{\omega\in R} \overline{\sigma}\left(\hat{G}(j\omega)\right)$$

In the SISO case the ∞-norm of $\hat{G}(s)$ is the maximal distance in the Nyquist diagram of $\hat{G}(j\omega)$ w.r.t. the origin, or equivalently as the peak value of the Bode diagram of $\hat{G}(j\omega)$.

DEF. 1-norm (L1):

$$\left\|\hat{G}(s)\right\|_1 = \int_{-\infty}^{+\infty} |g(t)| \, dt$$

We will not enter into the details but in practice these norms links the input and the output of signals. In particular if we stick to the SISO case we have the following link

	$u(t)=\delta(t)$	$u(t)=\sin(\omega t)$
$\|y\|_2$	$\left\|\hat{G}\right\|_2$	∞
$\|y\|_\infty$	$\|G\|_\infty$	$\left\|\hat{G}(j\omega)\right\|$
$pow(y)$	0	$\frac{\left\|\hat{G}(j\omega)\right\|}{\sqrt{2}}$

If we instead consider input that are unknown but bounded in some norm the relationship are the following

	$\|u(t)\|_2 \leq 1$	$\|u(t)\|_\infty \leq 1$	$pow(u(t)) \leq 1$
$\|y\|_2$	$\|\hat{G}\|_\infty$	∞	∞
$\|y\|_\infty$	$\|\hat{G}\|_2$ *	$\|\hat{G}\|_1$	∞
$pow(y)$	0	$\leq \|\hat{G}\|_\infty$	$\|\hat{G}\|_\infty$

We will not enter further into the details but it is possible to say that:

- The L1 norm links bounded input with bounded output (and in line of principle is the best thing to manage constraints and characterize results)
- The $H2$ norm is a norm that takes into account stochastic aspects (disturbances, measurement noise); a possible interpretation is that it gives the asymptotic covariance of the output when the system gets as an input white noise with unitary covariance. Moreover this norm is linked to the LQ cost if we define the output as

$$y = \begin{bmatrix} Q^{1/2} \\ 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ R^{1/2} \end{bmatrix} u$$

-The H∞ norm measures the highest gain across frequency in the singular value norm and maps the gain between finite energy or finite power of a system.

Moreover it is very used to enforce robust stability by checking $\left\| \frac{1}{1+G(s)} \right\|_\infty$

<u>Remark</u> Similar results hold true also for discrete time systems

Experience 28 – Norms of signals and LMIs

H-infinity as a Constraint

Consider a system

$\dot{x} = Ax + Bw$

$z = Cx + Dw$

And the associated transfer function $T(s)$. The constraints

$\|T\|_\infty < \gamma$

If x(0)=0 in terms of input out this maps as the relationship between the 2 norms which is

$$\int_0^\infty z(t)^T z(t) - \gamma^2 w(t)^T w(t) \leq 0, \quad (1)$$

under the assumption $w(t)^T w(t)$ is bounded.

If we take a Lyapunov function V(x)=x^TPx that satisfies:

$$\frac{dV(x)}{dt} + \gamma^{-1} z(t)^T z(t) - \gamma w(t)^T w(t) < 0 \qquad (2)$$

Then integrating between t=0 and t=T, we obtain:

$$V(x(T)) - V(x(0)) + \gamma^{-1} \int_0^T z(t)^T z(t) - \gamma^2 w(t)^T w(t) dt < 0$$

Since x(0)=0 if we do the limit for $T \to \infty$

$$V(x(T)) + \gamma^{-1} \int_0^T z(t)^T z(t) - \gamma^2 w(t)^T w(t) dt < 0$$

Since V(x(t))≥0 ∀x(T) then the integral must be negative and then (1) is satisfied.
We can then say that condition (2) ensures $\|T\|_\infty < \gamma$.

This latter (2) can be rewritten as an LMI, in fact:

$V(x) = x^T Px$

$$\frac{dV(x)}{dt} = x^T P\dot{x} + \dot{x}^T Px = x^T P(Ax + Bw) + (Ax + Bw)^T Px$$

And then

$$x^T P(Ax + Bw) + (Ax + Bw)^T Px + \gamma^{-1}(Cx + Dw)^T (Cx + Dw) - \gamma w(t)^T w(t) < 0$$

That is :

$$\begin{bmatrix} x \\ w \end{bmatrix}^T \begin{bmatrix} A^T P + PA + \gamma^{-1} C^T C & PB + \gamma^{-1} C^T D \\ B^T P + \gamma^{-1} D^T C & -\gamma I + \gamma^{-1} D^T D \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} < 0 \; \forall \mathrm{x}, \forall \mathrm{w}$$

So we obtain that this is true for P=P’>0 such taht:

$$\begin{bmatrix} A^T P + PA + \gamma^{-1} C^T C & PB + \gamma^{-1} C^T D \\ B^T P + \gamma^{-1} D^T C & -\gamma I + \gamma^{-1} D^T D \end{bmatrix} < 0 \quad (3)$$

Which is an LMI in the unknown P since γ **is assumed given**.

Computing H-infinity

If we want to determine the minimal value of γ the problem becomes

$$\min_{P,\gamma} \gamma$$

$$\begin{bmatrix} A^T P + PA + \gamma^{-1} C^T C & PB + \gamma^{-1} C^T D \\ B^T P + \gamma^{-1} D^T C & -\gamma I + \gamma^{-1} D^T D \end{bmatrix} < 0$$

Which is not an LMI since contains γ and γ^{-1}.

This can be converted in an LMI as follows

Theorem. (Kalman-Yacubocich-Popov Lemma o r Bounded Real Lemma)
A is asymptotically stable and $\|T\|_\infty < \gamma$ if and only if it exists a solution to the following LMI

$$P_\infty > 0$$

$$\begin{bmatrix} A^T P_\infty + P_\infty A & P_\infty B & C^T \\ B^T P_\infty & -\gamma I & D^T \\ C & D & -\gamma I \end{bmatrix} < 0 \qquad (4)$$

Computing the H2 norm

1) $\begin{bmatrix} A^T P_2 + P_2 A & P_2 B \\ B^T P_2 & -I \end{bmatrix} < 0$

2) $\begin{bmatrix} P_2 & (C + DK)^T \\ C + DK & X \end{bmatrix} > 0$

3) $tr(X) < \nu$

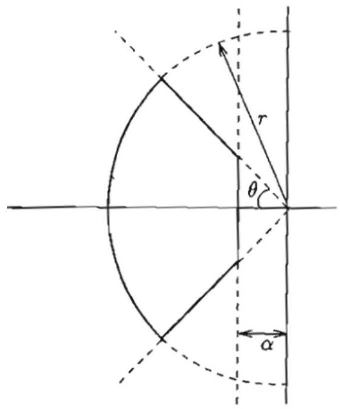
Computing an upper bound on the L-1
With further manipulation it is possible to prove the following results

Theorem $\|T\|_1^2 < \gamma$ if the following Matrix Inequalities in the P_1, μ, λ variables hold true that:

$$\begin{bmatrix} A^T P_1 + P_1 A + \lambda P_1 & P_1 B \\ B^T P_1 & -\mu I \end{bmatrix} < 0$$

$$\begin{bmatrix} \lambda P_1 & 0 & C^T \\ 0 & (\gamma - \mu) I & D^T \\ C & D & \gamma I \end{bmatrix} > 0$$

Experience 29 – Pole Location and LMIs



The eigenvalues of A are in the Region $S(\alpha, r, \theta)$ if the following LMIs old true

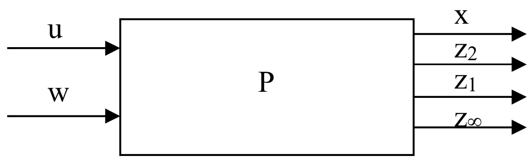
$$\begin{bmatrix} rP_r & A^T P_r \\ P_r A & rP_r \end{bmatrix} < 0$$
$$P_r > 0$$

$$\begin{bmatrix} \sin(\theta)(P_c A + A' P_c) & \cos(\theta)(A' P_c - P_c A) \\ \cos(\theta)(P_c A - A' P_c) & \sin(\theta)(P_c A + A' P_c) \end{bmatrix} < 0$$
$$P_c > 0$$

$$AP_\alpha + P_\alpha A^T - 2P_\alpha \alpha < 0$$
$$P_\alpha > 0$$

Experience 30 – From analysis to state feedback control

Let us consider a given plant P



We assume that the state is available and that moreover we have certain output

$T_\infty(s), z_\infty = T_\infty(s)w$

$T_1(s), z_1 = T_1(s)w$

$T_2(s), z_2 = T_2(s)w$

The plant is characterized by the following

$$\dot{x} = Ax + B_1 u + B_2 w$$
$$z_\infty = C_1 x + D_{11} w + D_{12} u$$
$$z_2 = C_2 x + D_{22} u$$
$$z_1 = C_3 x + D_{31} w + D_{32} u$$

If we use as a control a simple state feedback u=Kx we obtain

$$\dot{x} = (A + B_1 K)x + B_1 w$$
$$z_\infty = (C_1 + D_{12} K)x + D_{11} w$$
$$z_2 = (C_2 + D_{22} K)x$$
$$z_1 = (C_3 + D_{32} K)x + D_{31} w$$

At this point let us substitute in the different LMIs and let's see if we can come out with some design condition for each of the possible performance index and pole placement constraints

Norm H2

1) $\begin{bmatrix} (A+B_1K)^T P_2 + P_2(A+B_1K) & P_2B_2 \\ {B_2}^T P_2 & -I \end{bmatrix} < 0$

2) $\begin{bmatrix} P_2 & (C_2 + D_{22}K)^T \\ C_2 + D_{22}K & X \end{bmatrix} > 0$

3) $tr(X) < \nu$

First condition H2

We can pre and post-multiply for $\begin{bmatrix} P_2^{-1} & 0 \\ 0 & I \end{bmatrix}^T = \begin{bmatrix} P_2^{-1} & 0 \\ 0 & I \end{bmatrix}$:

$$\begin{bmatrix} P_2^{-1} & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} (A+B_1K)^T P_2 + P_2(A+B_1K) & P_2B_2 \\ {B_2}^T P_2 & -I \end{bmatrix}\begin{bmatrix} P_2^{-1} & 0 \\ 0 & I \end{bmatrix} < 0$$

which becomes:

$$\begin{bmatrix} P_2^{-1}(A+B_1K)^T + (A+B_1K)P_2^{-1} & B_2 \\ {B_2}^T & -I \end{bmatrix} < 0$$

And then

$$\begin{bmatrix} P_2^{-1}A^T + P_2^{-1}K^T {B_1}^T + AP_2^{-1} + B_1KP_2^{-1} & B_2 \\ {B_2}^T & -I \end{bmatrix} < 0$$

Finally we get:

$$\begin{bmatrix} Q_2A^T + {Q_2}^T {B_1}^T + AQ_2 + BY_2 & B_2 \\ {B_2}^T & -I \end{bmatrix} < 0$$

where

$$Q_2 = P_2^{-1}, Y_2 = KQ_2$$

Second condition H2

$$\begin{bmatrix} P_2 & (C_1 + D_{12}K)^T \\ C_1 + D_{12}K & X \end{bmatrix} > 0$$

With the same operations

$$\begin{bmatrix} P_2^{-1} & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} P_2 & (C_2 + D_{22}K)^T \\ C_2 + D_{22}K & X \end{bmatrix}\begin{bmatrix} P_2^{-1} & 0 \\ 0 & I \end{bmatrix} > 0$$

We get:

$$\begin{bmatrix} P_2^{-1} & P_2^{-1}{C_2}^T + P_2^{-1}K^T {D_{22}}^T \\ C_2P_2^{-1} + D_{22}KP_2^{-1} & X \end{bmatrix} > 0$$

And then:

$$\begin{bmatrix} Q_2 & Q_2{C_2}^T + {Y_2}^T {D_{22}}^T \\ C_2Q_2 + D_{22}Y_2 & X \end{bmatrix} > 0$$

Third condition H2

It remains the same

$$tr(X) < \nu$$

Norm H∞

$$\begin{bmatrix} (A+B_1K)^T P_\infty + P_\infty(A+B_1K) & P_\infty B_2 & (C_1+D_{12}K)^T \\ * & -\gamma_\infty I & D_{11}{}^T \\ * & * & -\gamma_\infty I \end{bmatrix} < 0$$

$$\begin{bmatrix} P_\infty{}^{-1} & & \\ & I & \\ & & I \end{bmatrix}\begin{bmatrix} (A+B_1K)^T P_\infty + P_\infty(A+B_1K) & P_\infty B_2 & (C_1+D_{12}K)^T \\ * & -\gamma_\infty I & D_{11}{}^T \\ * & * & -\gamma_\infty I \end{bmatrix}\begin{bmatrix} P_\infty{}^{-1} & & \\ & I & \\ & & I \end{bmatrix} < 0$$

$$\begin{bmatrix} P_\infty{}^{-1}(A+B_1K)^T + (A+B_1K)P_\infty{}^{-1} & B_2 & P_\infty{}^{-1}(C_1+D_{12}K)^T \\ * & -\gamma_\infty I & D_{11}{}^T \\ * & * & -\gamma_\infty I \end{bmatrix} < 0$$

And finally

$$\begin{bmatrix} Q_\infty A^T + Y_\infty{}^T B_1{}^T + AQ_\infty + B_1 Y_\infty & B_2 & Q_\infty C_1^T + Y_\infty{}^T D_{12}{}^T \\ * & -\gamma_\infty I & D_{11}{}^T \\ * & * & -\gamma_\infty I \end{bmatrix} < 0$$

where

$Q_\infty = P_\infty^{-1}, Y_\infty = KQ_\infty$

L1 - Norm

$$\begin{bmatrix} (A+B_1K)^T P_1 + P_1(A+B_1K) + \lambda P_1 & P_1 B_2 \\ B_2^T P_1 & -\mu I \end{bmatrix} < 0$$

$$\begin{bmatrix} \lambda P_1 & 0 & (C_3+D_{32}K)^T \\ 0 & (\gamma_1-\mu)I & D_{31}{}^T \\ (C_3+D_{32}K) & D_{31} & \gamma_1 I \end{bmatrix} > 0$$

Using the usual congruence transformations

$$\begin{bmatrix} P_1^{-1} & \\ & I \end{bmatrix}\begin{bmatrix} (A+B_1K)^T P_1 + P_1(A+B_1K) + \lambda P_1 & P_1 B_2 \\ B_2^T P_1 & -\mu I \end{bmatrix}\begin{bmatrix} P_1^{-1} & \\ & I \end{bmatrix} < 0$$

$$\begin{bmatrix} P_1^{-1} & & \\ & I & \\ & & I \end{bmatrix}\begin{bmatrix} \lambda P_1 & 0 & (C_3+D_{32}K)^T \\ 0 & (\gamma_1-\mu)I & D_{31}{}^T \\ (C_3+D_{32}K) & D_{31} & \gamma_1 I \end{bmatrix}\begin{bmatrix} P_1^{-1} & & \\ & I & \\ & & I \end{bmatrix} > 0$$

Finally we get

$$\begin{bmatrix} Q_1 A^T + Y_1{}^T B_1{}^T + AQ_1 + B_1 Y_1 + \lambda Q_1 & B_2 \\ B_2^T & -\mu I \end{bmatrix} < 0$$

$$\begin{bmatrix} \lambda Q_1 & 0 & Q_1 C_3{}^T + Y_1{}^T D_{32}{}^T \\ * & (\gamma_1-\mu)I & D_{31}{}^T \\ * & * & \gamma_1 I \end{bmatrix} > 0$$

where $Q_1 = P_1^{-1}, Y_1 = KX_1$

Pole Placement

α-stability

Let's start from

$$\left(A+BK\right)^{T}P_{\alpha}+P_{\alpha}\left(A+BK\right)+2\alpha P_{\alpha}<0$$

Pre and post-multiplying for P^{-1}:

$$P_{\alpha}^{-1}A^{T}+P_{\alpha}^{-1}K^{T}B^{T}+AP_{\alpha}^{-1}+BKP_{\alpha}^{-1}+2\alpha P_{\alpha}^{-1}<0$$

Using sobstitution $Q_{\alpha}=P_{\alpha}^{-1}, Y_{\alpha}=KQ_{\alpha}$ we obtain

$$Q_{\alpha}A^{T}+Y^{T}B^{T}+AQ_{\alpha}+BY_{\alpha}+2\alpha Q_{\alpha}<0$$

Conical sector

For notational compactness for θ∈(0,2π) let divide for cosθ and denote k=tgθ:

$$\begin{bmatrix} k\left(\left(A+BK\right)^{T}P_{c}+P_{c}\left(A+BK\right)\right) & \left(-\left(A+BK\right)^{T}P_{c}+P_{c}\left(A+BK\right)\right) \\ \left(-P_{c}\left(A+BK\right)+\left(A+BK\right)^{T}P_{c}\right) & k\left(\left(A+BK\right)^{T}P_{c}+P_{c}\left(A+BK\right)\right) \end{bmatrix}<0$$

Then we can do

$$\begin{bmatrix} P_{c}^{-1} & 0 \\ 0 & P_{c}^{-1} \end{bmatrix}\begin{bmatrix} k\left(\left(A+BK\right)^{T}P_{c}+P_{c}\left(A+BK\right)\right) & \left(-\left(A+BK\right)^{T}P_{c}+P_{c}\left(A+BK\right)\right) \\ \left(-P_{c}\left(A+BK\right)+\left(A+BK\right)^{T}P_{c}\right) & k\left(\left(A+BK\right)^{T}P_{c}+P_{c}\left(A+BK\right)\right) \end{bmatrix}\begin{bmatrix} P_{c}^{-1} & 0 \\ 0 & P_{c}^{-1} \end{bmatrix}<0$$

And get:

$$\begin{bmatrix} k\left(P_{c}^{-1}\left(A+BK\right)^{T}+\left(A+BK\right)P_{c}^{-1}\right) & \left(-P_{c}^{-1}\left(A+BK\right)^{T}+\left(A+BK\right)P_{c}^{-1}\right) \\ \left(-\left(A+BK\right)P_{c}^{-1}+P_{c}^{-1}\left(A+BK\right)^{T}\right) & k\left(P_{c}^{-1}\left(A+BK\right)^{T}+\left(A+BK\right)P_{c}^{-1}\right) \end{bmatrix}<0$$

Using sobstitution $Q_{c}=P_{c}^{-1}, Y_{c}=KQ_{c}$:

$$\begin{bmatrix} k\left(Q_{c}A^{T}+Y_{c}^{T}B^{T}+AQ_{c}+BY_{c}\right) & \left(-Q_{c}A^{T}-Y_{c}^{T}B^{T}+AQ_{c}+BY_{c}\right) \\ \left(-AQ_{c}-BY_{c}+XA_{c}^{T}+Y_{c}^{T}B^{T}\right) & k\left(Q_{c}A^{T}+Y_{c}^{T}B^{T}+AQ_{c}+BY_{c}\right) \end{bmatrix}<0$$

Circle

$$\begin{bmatrix} rP_{r} & \left(\left(A+BK\right)^{T}P_{r}\right) \\ \left(P_{r}\left(A+BK\right)\right) & rP_{r} \end{bmatrix}<0$$

$$\begin{bmatrix} rP_{r}^{-1} & \left(P_{r}^{-1}\left(A+BK\right)^{T}\right) \\ \left(\left(A+BK\right)P_{r}^{-1}\right) & rP_{r}^{-1} \end{bmatrix}<0$$

Using sobstitution $Q_{r}=P_{r}^{-1}, Y_{r}=KQ_{r}$ we obtain

$$\begin{bmatrix} rQ_{r} & \left(Q_{r}A^{T}+Y_{r}^{T}B^{T}\right) \\ \left(AQ_{r}+BY_{r}\right) & rQ_{r} \end{bmatrix}<0$$

Experience 31 – Multiobjective idea – State feedback

It is possible to combine altogether the seen synthesis method. In particular we want to find a single controller K such that

1) $\|T_1\|_1 < \overline{\gamma_1}$
2) $\|T_2\|_2 < \overline{\gamma_2}$
3) $\|T_\infty\|_\infty < \overline{\gamma_\infty}$
4) all the closed loop eigenvalue are in the area S(α,θ,r)

Moreover we may want to minimize the index
5) min $\alpha\|T_\infty\|_\infty + \beta\|T_2\|_2 + \phi\|T_1\|_1$

The idea is that we should find a single K solving all the single problems. For the seen reformulation we would have then

$X_i = P_i^{-1}$

$Y_i = KP_i^{-1}$

The main idea is introducing consertivism using a unique $P = P_1 = P_2 = P_\infty = P_\alpha = P_c = P_r$ (**Lyapunov Shaping**). In this case we would have a single problem with X and Y. To complete our synthesis we should just add the constraints on the norms

$\gamma_\infty < \overline{\gamma_\infty}$

$\gamma_2 < \overline{\gamma_2}$

$\gamma_1 < \overline{\gamma_1}$

and the minimization index

min $\alpha\gamma_\infty + \beta\gamma_2 + \phi\gamma_1$

If the problem has solution, the controller will be $K=YX^{-1}$

Remark Note that there are quasi LMIs parameterized in λ !

Experience 32 – Multiobjective idea – Output feedback

The same case of design can be done if instead of “reading” the state we have an output y that is sensed.

$$\begin{aligned}
\dot{x} &= Ax + B_1 u + B_2 w \\
z_\infty &= C_1 x + D_{11} w + D_{12} u \\
z_2 &= C_2 x + D_{22} u \\
z_1 &= C_3 x + D_{31} w + D_{32} u \\
y &= Cx + Dw
\end{aligned}$$

In this case the controller will be a dynamical compensator

$$\begin{aligned}
\dot{x}_K &= A_K x_K + B_K y \\
u &= C_K x_K + D_K y
\end{aligned}$$

And the closed loop system becomes

$$\begin{cases}
\begin{bmatrix} \dot{x} \\ \dot{x}_K \end{bmatrix} = \begin{bmatrix} A + B_1 D_k C & B_1 C_K \\ B_K C & A_K \end{bmatrix} \begin{bmatrix} x \\ x_K \end{bmatrix} + \begin{bmatrix} B_1 D_K D + B_2 \\ B_K D \end{bmatrix} w \\
z_\infty = \begin{bmatrix} C_1 + D_{12} D_k C & D_{12} C_K \end{bmatrix} \begin{bmatrix} x \\ x_K \end{bmatrix} + \left(D_{11} + D_{12} D_K D \right) w \\
z_2 = \begin{bmatrix} C_2 + D_{22} D_k C & D_{22} C_K \end{bmatrix} \begin{bmatrix} x \\ x_K \end{bmatrix} + \left(D_{21} + D_{22} D_K D \right) w \\
z_1 = \begin{bmatrix} C_3 + D_{32} D_k C & D_{32} C_K \end{bmatrix} \begin{bmatrix} x \\ x_K \end{bmatrix} + \left(D_{31} + D_{32} D_K D \right) w
\end{cases}$$

Or in compact form

$$\begin{cases}
\dot{x}_{cl} = A_{cl} x_{cl} + B_{cl} w \\
z_\infty = \overline{C}_1 x_{cl} + \overline{D}_1 w \\
z_2 = \overline{C}_2 x_{cl} + \overline{D}_2 w \\
z_1 = \overline{C}_3 x_{cl} + \overline{D}_3 w
\end{cases}$$

It remains to substitute in the conditions for the multi-objective. Also in this case we will use Lyapunov Shaping, and will consider a common P. Using A_K,B_K,C_K,D_K as unknown we do not obtain LMI

However, it is possible to obtain LMI **using some trick.** Let k be the dimension of the controller:

$$P = \begin{bmatrix} X & N \\ N^T & F \end{bmatrix} \in R^{(n+k)\times(n+k)}$$

$$P^{-1} = \begin{bmatrix} Y & M \\ M^T & Z \end{bmatrix} \in R^{(n+k)\times(n+k)}$$

with:

$$X = X' \in R^{n\times n},\ Y = Y' \in R^{n\times n},\ N \in R^{n\times k},\ M \in R^{n\times k}$$

Clearly PP^{-1}=I, but then we can write

$$P\begin{bmatrix} Y \\ M^T \end{bmatrix} = \begin{bmatrix} I_n \\ 0 \end{bmatrix}$$

If we introduce the following matrices

$$\pi_y = \begin{bmatrix} Y & I_n \\ M^T & 0 \end{bmatrix}$$

$$\pi_x = \begin{bmatrix} I_n & X \\ 0 & N^T \end{bmatrix}$$

We can write that

$$P\pi_y = \pi_x$$

Tese matrices $\pi_y \, e \, \pi_x$ are foundamental to obtain LMI conditions

Before using them please note that

$$R_1 = \pi_y{}^T P A_{cl} \pi_y = \pi_x{}^T A_{cl} \pi_y = \begin{bmatrix} AY + B_1\hat{C}_K & A + B_1\hat{D}_K C \\ \widehat{A}_k & XA + \hat{B}_K C \end{bmatrix}$$

$$R_2 = \pi_y{}^T P \pi_y = \pi_x{}^T \pi_y = \begin{bmatrix} Y & I \\ I & X \end{bmatrix}$$

$$R_3 = \pi_y{}^T P B_{cl} = \pi_x{}^T B_{cl} = \begin{bmatrix} B_2 + B\hat{D}_K D \\ XB_2 + \hat{B}_K D \end{bmatrix}$$

$$R_{4,i} = \overline{C_i}\pi_y = \begin{bmatrix} C_i Y + D_{i,2}\hat{C}_K & C_i + D_{i,2}\hat{D}_K C \end{bmatrix}$$

Where each one of these matrices is Affine in the unknowns and where the auxiliary variables are:

$$\hat{A}_K = NA_K M^T + NB_K CY + XBC_K M^T + X(A + BD_K C)Y$$

$$\hat{B}_K = NB_K + XBD_K$$

$$\hat{C}_K = C_K M^T + D_K CY$$

$$\hat{D}_K = D_K$$

Remark if M e N are full row-rank rango $N \in R^{nxk}$, $M \in R^{nxk}$ (k>n and n rows are linearly indipendent), given the auxiliary variables, there are infinite A_K, B_K, C_K, D_K solving these qualities. If N and M are square (k=n) and invertible, the solution is unique

H2 Norm

H2 Norm- first equation

$$\begin{bmatrix} A_{cl}{}^{T}P + PA_{cl} & PB_{cl} \\ B_{cl}{}^{T}P & -I \end{bmatrix} < 0$$

Doing the following transformation

$$\begin{bmatrix} \pi_y{}^{T} & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} A_{cl}{}^{T}P + PA_{cl} & PB_{cl} \\ B_{cl}{}^{T}P & -I \end{bmatrix}\begin{bmatrix} \pi_y & 0 \\ 0 & I \end{bmatrix} < 0$$

We obtain

$$\begin{bmatrix} \pi_y{}^{T}A_{cl}{}^{T}P\pi_y + \pi_y{}^{T}PA_{cl}\pi_y & \pi_y{}^{T}PB_{cl} \\ B_{cl}{}^{T}P\pi_y & -I \end{bmatrix} < 0$$

And finally

$$\begin{bmatrix} R_1{}^{T} + R_1 & R_3 \\ R_3{}^{T} & -I \end{bmatrix} < 0$$

H2 Norm, second equation

$$\begin{bmatrix} P & \overline{C}_2{}^{T} \\ \overline{C}_2 & Q \end{bmatrix} > 0$$

Using the same transformation

$$\begin{bmatrix} \pi_y{}^{T} & 0 \\ 0 & I \end{bmatrix}\begin{bmatrix} P & \overline{C}_2{}^{T} \\ \overline{C}_2 & Q \end{bmatrix}\begin{bmatrix} \pi_y & 0 \\ 0 & I \end{bmatrix} > 0$$

We have

$$\begin{bmatrix} \pi_y{}^{T}P\pi_y & \pi_y{}^{T}\overline{C}_2{}^{T} \\ \overline{C}_2\pi_y & Q \end{bmatrix} > 0$$

And then

$$\begin{bmatrix} R_2 & R_{4,2}{}^{T} \\ R_{4,2} & Q \end{bmatrix} > 0$$

Substitution H2 norm, third equation

$$tr(Q) < \gamma$$

Remains the same

Substitution D_2=0

$$\overline{D}_2 = \left(D_{21} + D_{22}\hat{D}_K D\right) = 0$$

Substitution P>0

$$\pi_y{}^{T}P\pi_y = R_2 > 0$$

Hinf norm

$$\begin{bmatrix} \pi_y{}^T & 0 & 0 \\ * & I & 0 \\ * & * & I \end{bmatrix}\begin{bmatrix} A_{cl}{}^T P_\infty + P_\infty A_{cl} & P_\infty B_{cl} & \overline{C}_1^T \\ * & -\gamma_\infty I & \overline{D}_1^T \\ * & * & -\gamma_\infty I \end{bmatrix}\begin{bmatrix} \pi_y & 0 & 0 \\ * & I & 0 \\ * & * & I \end{bmatrix} < 0$$

$$\begin{bmatrix} \pi_y A_{cl}{}^T P_\infty \pi_y + \pi_y{}^T P_\infty A_{cl} \pi_y & \pi_y{}^T P_\infty B_{cl} & \pi_y{}^T \overline{C}_1^T \\ * & -\gamma_\infty I & \overline{D}_1^T \\ * & * & -\gamma_\infty I \end{bmatrix} < 0$$

And then

$$\begin{bmatrix} R_1{}^T + R_1 & R_3 & R_{4,1} \\ * & -\gamma_\infty I & \overline{D}_1{}^T \\ * & * & -\gamma_\infty I \end{bmatrix} < 0$$

Norma L1

$$\begin{bmatrix} \pi_y{}^T & \\ & I \end{bmatrix}\begin{bmatrix} A_{cl}{}^T P_1 + P_1 A_{cl} + \lambda P_1 & P_1 B_{cl} \\ B_{cl}^T P_1 & -\mu I \end{bmatrix}\begin{bmatrix} \pi_y & \\ & I \end{bmatrix} < 0$$

$$\begin{bmatrix} \pi_y{}^T & & \\ & I & \\ & & I \end{bmatrix}\begin{bmatrix} \lambda P_1 & 0 & (C_3 + D_{32}K)^T \\ 0 & (\gamma_1 - \mu)I & D_{31}{}^T \\ (C_3 + D_{32}K) & D_{31} & \gamma_1 I \end{bmatrix}\begin{bmatrix} \pi_y & & \\ & I & \\ & & I \end{bmatrix} > 0$$

And finally

$$\begin{bmatrix} R_1{}^T + R_1 + \lambda R_2 & R_3 \\ * & -\mu I \end{bmatrix} < 0$$

$$\begin{bmatrix} \lambda R_2 & 0 & R_{4,3}{}^T \\ 0 & (\gamma_1 - \mu)I & \overline{D}_3{}^T \\ * & * & \gamma_1 I \end{bmatrix} > 0$$

Pole Placement

$$\left[\lambda_{k,l} P + \mu_{kl} P A_{cl} + \mu_{lk} A_{cl}{}^T P\right]_{1\le k,l\le m} < 0$$

$$\left[\pi_y{}^T \lambda_{k,l} P \pi_y + \mu_{kl} \pi_y{}^T P A_{cl} \pi_y + \mu_{lk} \pi_y{}^T A_{cl}{}^T P \pi_y\right]_{1\le k,l\le m} < 0$$

$$\left[\lambda_{k,l} R_1 + \mu_{kl} R_2 + \mu_{lk} R_2{}^T\right]_{1\le k,l\le m} < 0$$

How to get the controller

If we can find a solution, it remains to compute the control law.The first step is to compute N,M and P satisfying :

$P\pi_y = \pi_x$

That is

$$\begin{bmatrix} X & N \\ N^T & Z \end{bmatrix}\begin{bmatrix} Y & I \\ M^T & 0 \end{bmatrix} = \begin{bmatrix} I & X \\ 0 & N^T \end{bmatrix}$$

It is important to focus on the first element

$$\begin{bmatrix} XY + NM^T & \\ & \end{bmatrix} = \begin{bmatrix} I & \\ & \end{bmatrix}$$

That is

$NM^T = I - XY$

Please note that, since $\begin{bmatrix} Y & I \\ I & X \end{bmatrix} > 0$ then X>0,Y-X^{-1}>0, and then I-YX is nonsingular.

As a consequence we can factorize as

NM^T=I-XY

Once we get M and N it remain to solve the linear system

$$\hat{A}_K = NA_K M^T + NB_K CY + XBC_K M^T + X(A + BD_K C)Y$$
$$\hat{B}_K = NB_K + XBD_K$$
$$\hat{C}_K = C_K M^T + D_K CY$$
$$\hat{D}_K = D_K$$

If M and N are square and nonsingular we have

$$D_k = \hat{D}_K$$
$$C_k = \left(\hat{C}_K - D_K CY\right)\left(M^T\right)^{-1}$$
$$B_k = N^{-1}\left(\hat{B}_K - XBD_K\right)$$
$$A_k = N^{-1}\left(\hat{A}_K - NB_K CY - XBC_K M^T - X(A + BD_K C)Y\right)\left(M^T\right)^{-1}$$

Experience “bonus track” - The S-Procedure

Sometimes there is the need to use “**for all”** quantifier. A typical case is when one wants that **some quadratic function is negative for all the value belonging to a given set.**

It is possible to write sufficient condition guaranteeing the satisfaction of these constraints. The basic one is the following:

LEMMA: Let a problem in the form: $x^T T_0 x > 0 \quad \forall x : x^T T_i x \geq 0 i = 1,.., p$. A sufficient condition is given by the existence of scalars ∃ $\tau_1 \geq 0, \ldots, \tau_p \geq 0$: $T_0 - \sum_{i=1}^{p} \tau_i T_i > 0$. For p=1 the condition becomes also necessary if $\exists x_0 : x_0^T T_1 x_0 > 0$

Exercise 59:

$$\begin{bmatrix} x^T & \pi^T \end{bmatrix} \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \begin{bmatrix} x \\ \pi \end{bmatrix} < 0$$ has to be true

$\forall x \neq 0$ e $\forall \pi$ such that

$$\begin{bmatrix} x^T & \pi^T \end{bmatrix} \begin{bmatrix} C^T C & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} x \\ \pi \end{bmatrix} \geq 0$$

where P>0 is a variable