This exam contains 8 pages (including this cover page) and 5 exercises.

1. **Minimum energy control**

   Consider the single-input multiple-output discrete-time linear dynamical system

   $$\boldsymbol{x}(k+1) = \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{B}u(k)$$
   $$\boldsymbol{y}(k) = \boldsymbol{C}\boldsymbol{x}(k), \tag{1}$$

   with $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, $\boldsymbol{B} \in \mathbb{R}^{n \times 1}$ and $\boldsymbol{C} \in \mathbb{R}^{q \times n}$. Starting from a random initial condition $\boldsymbol{x}(0)$, your goal is to devise a minimum energy control scheme such that, at $k = t$, we have $\boldsymbol{y}(t) = \boldsymbol{y}_{\text{des}}$ where $\boldsymbol{y}_{\text{des}}$ is the desired output of the system. From a mathematical point of view, you thus need to the sequence of inputs $u(0), u(1), \cdots, u(t-1)$ that minimize the cost function

   $$\mathcal{J} = \sum_{k=0}^{t-1} u(k)^2$$

   under the constraints that the dynamics of the system are given by Eq. (1) and that $\boldsymbol{y}(t) = \boldsymbol{y}_{\text{des}}$. You can assume that the system is both controlable and observable.

   (a) (1 point) Given all the problem data (i.e. $\boldsymbol{x}(0)$, $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{y}_{\text{des}}$ and t), detail how you can reformulate this problem as a least-norm problem. If you need to make any assumption regarding the properties of the matrices, say so.

   > **Solution:** Starting from the arbitrary initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_0$, we can write
   >
   > $\boldsymbol{y}(0) = \boldsymbol{C}\boldsymbol{x}_0$
   > $\boldsymbol{y}(1) = \boldsymbol{C}\boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{C}\boldsymbol{B}u_0$
   > $\boldsymbol{y}(2) = \boldsymbol{C}\boldsymbol{A}^2\boldsymbol{x}_0 + \boldsymbol{C}\boldsymbol{A}\boldsymbol{B}u_0 + \boldsymbol{C}\boldsymbol{B}u_1$
   > $\boldsymbol{y}(3) = \boldsymbol{C}\boldsymbol{A}^3\boldsymbol{x}_0 + \boldsymbol{C}\boldsymbol{A}^2\boldsymbol{B}u_0 + \boldsymbol{C}\boldsymbol{A}\boldsymbol{B}u_1 + \boldsymbol{C}\boldsymbol{B}_2$
   > $\quad\vdots$
   > $\boldsymbol{y}(k) = \boldsymbol{C}\boldsymbol{A}^k\boldsymbol{x}_0 + \boldsymbol{C}\boldsymbol{A}^{k-1}\boldsymbol{B}u_0 + \boldsymbol{C}\boldsymbol{A}^{k-2}\boldsymbol{B}u_1 + \boldsymbol{C}\boldsymbol{A}^{k-3}\boldsymbol{B}u_2 + \cdots + \boldsymbol{C}\boldsymbol{B}u_{k-1}.$
   >
   > We can rewrite this equation in matrix form as follows
   >
   > $$\boldsymbol{y}(k) - \boldsymbol{C}\boldsymbol{A}^k\boldsymbol{x}_0 = \begin{bmatrix} \boldsymbol{C}\boldsymbol{A}^{k-1}\boldsymbol{B} & \boldsymbol{C}\boldsymbol{A}^{k-2}\boldsymbol{B} & \boldsymbol{C}\boldsymbol{A}^{k-3}\boldsymbol{B} & \cdots & \boldsymbol{C}\boldsymbol{B} \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_{k-1} \end{bmatrix}.$$

Finding the minimum energy input sequence $\boldsymbol{u} = \{u_k\}_{k=0,t-1}$ such that $\boldsymbol{y}(t) = \boldsymbol{y}_{\text{des}}$ then amounts to solve the following least-norm problem

$$\underset{\boldsymbol{u}}{\text{minimize}} \ \|\boldsymbol{u}\|_2^2$$

$$\text{subject to } \boldsymbol{H}\boldsymbol{u} = \boldsymbol{y}_{\text{des}} - \boldsymbol{C}\boldsymbol{A}^t\boldsymbol{x}_0$$

with

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{C}\boldsymbol{A}^{t-1}\boldsymbol{B} & \boldsymbol{C}\boldsymbol{A}^{t-2}\boldsymbol{B} & \boldsymbol{C}\boldsymbol{A}^{t-3}\boldsymbol{B} & \cdots & \boldsymbol{C}\boldsymbol{B} \end{bmatrix}.$$

This constrained minimization problem can be transformed into an equivalent unconstrained problem by introducing a *Lagrange multiplier* $\boldsymbol{\lambda}$ such that it reads

$$\underset{\boldsymbol{u},\boldsymbol{\lambda}}{\text{minimize}} \ \boldsymbol{u}^T\boldsymbol{u} + \boldsymbol{\lambda}^T \left(\boldsymbol{H}\boldsymbol{u} - \boldsymbol{y}_{\text{des}} - \boldsymbol{C}\boldsymbol{A}^t\boldsymbol{x}_0 \right).$$

The components of the gradient of this augmented cost function $\mathcal{L}(\boldsymbol{u}, \boldsymbol{\lambda})$ are given by

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{u}} = 2\boldsymbol{u} + \boldsymbol{H}^T\boldsymbol{\lambda}$$

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda}} = \boldsymbol{H}\boldsymbol{u} - \left(\boldsymbol{y}_{\text{des}} - \boldsymbol{C}\boldsymbol{A}^t\boldsymbol{x}_0 \right).$$

Setting the first component to zero yields

$$\boldsymbol{u} = -\frac{1}{2}\boldsymbol{H}^T\boldsymbol{\lambda}.$$

Introducing this expression for $\boldsymbol{u}$ into the second equation and setting it to zero yields

$$\boldsymbol{\lambda} = -2 \left(\boldsymbol{H}\boldsymbol{H}^T \right)^{-1} \left(\boldsymbol{y}_{\text{des}} - \boldsymbol{C}\boldsymbol{A}^t\boldsymbol{x}_0 \right).$$

Combining these two expressions finally provides us with the minimum energy control sequence

$$\boldsymbol{u} = \boldsymbol{H}^T \left(\boldsymbol{H}\boldsymbol{H}^T \right)^{-1} \left(\boldsymbol{y}_{\text{des}} - \boldsymbol{C}\boldsymbol{A}^t\boldsymbol{x}_0 \right)$$

where $\boldsymbol{H}^T \left(\boldsymbol{H}\boldsymbol{H}^T \right)^{-1}$ is the Moore-Penrose pseudoinverse of $\boldsymbol{H}$. Note that in this process, we implicitly assume that $\boldsymbol{H}\boldsymbol{H}^T$ is invertible.

(b) (1 point) Using your computer, carry out your method on the particular problem instance with data

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0.1 \\ 0 & 0.9 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$\boldsymbol{x}(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{y}_{\text{des}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \text{and } t = 10.$$

Please plot the optimal sequence of inputs and return the value of the cost function $\mathcal{J}(\boldsymbol{u})$ for this particular sequence.

> **Solution:** Please refer to `ex1.jl` for the Julia implementation and `ex1.py` for the python one. A Matlab implementation can also be provided on demand.

2. **Estimating direction and amplitude of a light beam**

   A light beam with (non-negative) amplitude a comes from a direction $\boldsymbol{d} \in \mathbb{R}^3$, where $\|\boldsymbol{d}\|_2 = 1$ (this means that the beam travels in the direction $-\boldsymbol{d}$). The beam falls on $m \geq 3$ photodetectors, each of which generates a scalar signal that depends on the beam's amplitude and direction as well as on the direction in which the photodetector is pointed. Specifically, photodetector i generates an output signal p_i

   $$p_i = \alpha a \cos(\beta_i) + v_i$$

   where β_i is the angle between the beam direction $\boldsymbol{d}$ and the outward normal vector $\boldsymbol{q}_i$ on the surface of the i$^{\text{th}}$ photodetector. For the sake of simplicity, we'll assume that all photodetectors have the same sensitivity α. You can interpret $\boldsymbol{q}_i \in \mathbb{R}^3$ (with $\|\boldsymbol{q}_i\| = 1$) as the direction the i$^{\text{th}}$ photodetector is pointed. We'll assume furthermore that $|\beta_i| \leq 90°$, i.e. the beam illuminates the top of the photodetectors. The numbers v_i are small measurement errors. Assume you are given the photodetector direction vectors $\boldsymbol{q}_1, \cdots, \boldsymbol{q}_m \in \mathbb{R}^3$, their sensitivity α and the noisy output $p_1, \cdots, p_m \in \mathbb{R}$. Your job is to estimate the beam direction $\boldsymbol{d} \in \mathbb{R}^3$ and its amplitude $a \in \mathbb{R}^+$.

   (a) (1 point) Explain how you can recast this problem in the least-squares framework. If some matrix (or matrices) needs to be full rank for it to work, say so.

   > **Solution:** Let us begin by reformulating slightly the problem. Given that $\|\boldsymbol{d}\| = 1$ and $\|\boldsymbol{q}_i\| = 1$, the i$^{\text{th}}$ measurement can be written as
   >
   > $$p_i = \alpha \boldsymbol{q}_i^T \boldsymbol{d} a + v_i.$$
   >
   > Moreover, if we introduce $\boldsymbol{x} = a\boldsymbol{d}$, this equation becomes
   >
   > $$p_i = \alpha \boldsymbol{q}_i^T \boldsymbol{x} + v_i.$$
   >
   > Considering all of the measurements simulataneously leads to
   >
   > $$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_m \end{bmatrix} = \alpha \begin{bmatrix} \boldsymbol{q}_1^T \\ \boldsymbol{q}_2^T \\ \boldsymbol{q}_3^T \\ \vdots \\ \boldsymbol{q}_m^T \end{bmatrix} \boldsymbol{x} + \boldsymbol{v}$$
   >
   > $$\boldsymbol{p} = \alpha \boldsymbol{Q} \boldsymbol{x} + \boldsymbol{v}$$
   >
   > where $\boldsymbol{v}$ is the vector of measurement errors. The unknown vector $\boldsymbol{x}$ can then be found as the solution of the following least-squares problem
   >
   > $$\underset{\boldsymbol{x}}{\text{minimize}} \ \frac{1}{2} \|\boldsymbol{p} - \alpha \boldsymbol{Q} \boldsymbol{x}\|_2^2.$$

The gradient of this cost function $\mathcal{J}(\boldsymbol{x})$ is given by

$$\frac{\partial \mathcal{J}}{\partial \boldsymbol{x}} = \alpha^2 \boldsymbol{Q}^T \boldsymbol{Q} \boldsymbol{x} - \alpha \boldsymbol{Q}^T \boldsymbol{p}.$$

Setting this gradient to zero yields to the normal equation

$$\boldsymbol{Q}^T \boldsymbol{Q} \boldsymbol{x} = \frac{1}{\alpha} \boldsymbol{Q}^T \boldsymbol{p}.$$

Assuming $\boldsymbol{Q}^T \boldsymbol{Q}$ is full-rank, the least-squares solution is then

$$\boldsymbol{x} = \frac{1}{\alpha} \left(\boldsymbol{Q}^T \boldsymbol{Q} \right)^{-1} \boldsymbol{Q}^T \boldsymbol{p}.$$

Once $\boldsymbol{x}$ has been identified, the estimated amplitude of the beam is given by $a = \|\boldsymbol{x}\|$ while its direction is $\boldsymbol{d} = \boldsymbol{x}/a$.

(b) (1 point) To describe unit vectors $\boldsymbol{q}_1, \cdots, \boldsymbol{q}_m \in \mathbb{R}^3$, we will use azimuthal and elevation defined as follows

$$\boldsymbol{q}_i = \begin{bmatrix} \cos(\varphi)\cos(\vartheta) \\ \cos(\varphi)\sin(\vartheta) \\ \sin(\varphi) \end{bmatrix}$$

where φ is the elevation (between $0°$ and $90°$ since all photodetectors point upward). The azimuth angle $0° \le \vartheta \le 360°$ gives the direction in the plane spanned by the first and second coordinates. Carry out your method on the particular problem instance

$$\boldsymbol{p} = \begin{bmatrix} 1.58 \\ 1.50 \\ 2.47 \\ 1.1 \\ 0.001 \end{bmatrix}, \quad \boldsymbol{\varphi} = \begin{bmatrix} 88 \\ 34 \\ 30 \\ 20 \\ 50 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{\vartheta} = \begin{bmatrix} 3 \\ 10 \\ 80 \\ 150 \\ 275 \end{bmatrix}.$$

Note that you may have to express these angles in radians rather than degrees. Give your final estimate of the beam's amplitude a and direction $\boldsymbol{d}$ (with azimuth and elevation angles in degrees).

Solution: Please refer to `ex2.jl` for the Julia implementation and `ex2.py` for the python one. A Matlab implementation can also be provided on demand.

3. **Convex optimization**

Consider the quadratic functional $f : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x_1, x_2) = 2x_1^2 + x_2^2 + x_1 x_2 + x_1 + x_2.$$

The aim of this exercise is to design an algorithm to find the solution (x_1, x_2) minimizing the function $f(x_1, x_2)$.

(a) (1 point) Show that the problem above can be cast as

$$\underset{\boldsymbol{x}}{\text{minimize}} \ \frac{1}{2}\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{q} + c$$

where $\boldsymbol{P} \in \mathbb{R}^{2\times 2}$ is a symmetric positive definite matrix, $\boldsymbol{q}$ is in $\mathbb{R}^2$ and c a constant.

> **Solution:** Straightforward algebraic manipulations show that
>
> $$\boldsymbol{P} = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}, \quad \boldsymbol{q} = - \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \text{and} \quad c = 0.$$
>
> It can moreover easily be shown that the eigenvalues of $\boldsymbol{P}$ are indeed positive.

(b) (1 point) Show that minimize of this problem is also solution of

$$\boldsymbol{P}\boldsymbol{x} = \boldsymbol{q}$$

and solve the proble analytically.

> **Solution:** The gradient of the function f with respect to $\boldsymbol{x}$ is given by
>
> $$\frac{\partial f}{\partial \boldsymbol{x}} = \boldsymbol{P}\boldsymbol{x} - \boldsymbol{q}.$$
>
> Setting this gradient to zero (optimilaty condition) yields
>
> $$\boldsymbol{P}\boldsymbol{x} = \boldsymbol{q}.$$
>
> The analytical resolution is left to the students.

(c) (1 point) Solve the problem numerically, return the optimal value of $\boldsymbol{x}$ as well as the corresponding value of the function $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{q} + c$.

4. **Specialized solver for tridiagonal matrices**

Tridiagonal matrices are of the form

$$\boldsymbol{A} = \begin{bmatrix} a_1 & b_1 & 0 & 0 & 0 & \cdots & 0 \\ c_1 & a_2 & b_2 & 0 & 0 & \cdots & 0 \\ 0 & c_2 & a_3 & b_3 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & c_{n-2} & a_{n-1} & b_{n-1} \\ 0 & \cdots & \cdots & \cdots & 0 & c_{n-1} & a_n \end{bmatrix}$$

They are ubiquituous in applied mathematics (see the discretization of differential operations). It is possible to leverage their particular structure to design efficient algorithms to solve linear problems involving such matrices. In this exercise, you'll have to design such an algorithm.

(a) (1 point) Design an algorithm solving $\boldsymbol{Ax} = \boldsymbol{b}$ that leverages the tridiagonal structure of this type of matrices. Your algorithm should work no matter the size n of the matrix. You can assume that all the entries a_i on the main diagonal are non-zero.

> **Solution:** The particular algorithm you had to come up with is known as *Thomas algorithm*. There are plenty of resources online about it so it won't be detailed hereafter.

(b) (1 point) Implement your algorithm and consider the following 10×10 matrix

$$\boldsymbol{A} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 \end{bmatrix}$$

also known as the *Strang matrix* and the right hand side vector

$$\boldsymbol{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T.$$

5. **It's the dreaded theory problem ! And it's mandatory !**

Consider the SISO discrete-time linear system

$$\boldsymbol{x}_{k+1} = \boldsymbol{Ax}_k + \boldsymbol{B}u_k$$
$$y_k = \boldsymbol{Cx}_k.$$

(a) (1 point) Under which condition is this system linearly unstable?

> **Solution:** In the absence of external forcing, the solution to this system is given by
> $$\boldsymbol{x}_k = \boldsymbol{A}^k \boldsymbol{x}_0.$$
> Introducing the eigendecomposition $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1}$, it can be rewritten as
> $$\boldsymbol{x}_k = \boldsymbol{V}\boldsymbol{\Lambda}^k\boldsymbol{V}^{-1}\boldsymbol{x}_0.$$
> Hence, if at least one eigenvalue of $\boldsymbol{A}$ is outside the unit circle, then
> $$\lim_{k \to \infty} \|\boldsymbol{\Lambda}^k\|_2^2 = \infty$$
> and the system is linearly unstable.

(b) (1 point) Give the expression of the impulse response filter $\boldsymbol{h}_k$ as a function of k, $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$.

> **Solution:** The impulse response is given by
>
> $$\boldsymbol{h}_k = \begin{cases} \boldsymbol{C}\boldsymbol{A}^k\boldsymbol{B}, & \text{for} k > 0, \\ \boldsymbol{0}, & \text{otherwise.} \end{cases}$$

(c) (1 point) Show that $\boldsymbol{h}_k$ is invariant with respect to an invertible change of coordinates.

> **Solution:** Consider the change of coordinates $\boldsymbol{z} = \boldsymbol{V}^{-1}\boldsymbol{x}$. The governing equations for $\boldsymbol{z}$ are given by
>
> $$\boldsymbol{z}_{k+1} = \boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{V}\boldsymbol{z}_k + \boldsymbol{V}^{-1}\boldsymbol{B}u_k$$
> $$\boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{V}\boldsymbol{z}_k.$$
>
> It can then easily be shown that the Markov parameters are the same.

(d) (1 point) Assume that $\boldsymbol{x}_0 = \boldsymbol{0}$. Consider the two vectors $\boldsymbol{u} = \begin{bmatrix} u_0 & u_1 & u_2 & \cdots & u_{N-1} \end{bmatrix}^T$ and $\boldsymbol{y} = \begin{bmatrix} y_1 & y_2 & y_3 & \cdots & y_N \end{bmatrix}^T$. Find the matrix $\boldsymbol{T} \in \mathbb{R}^{N \times N}$ that maps the input sequence $\boldsymbol{u}$ to the output sequence $\boldsymbol{y}$, i.e. $\boldsymbol{y} = \boldsymbol{T}\boldsymbol{u}$.

> **Solution:** Starting from the definition of the system and the fact that $\boldsymbol{x}_0 = \boldsymbol{0}$, we can write
>
> $y_0 = 0$
> $y_1 = \boldsymbol{C}\boldsymbol{B}u_0$
> $y_2 = \boldsymbol{C}\boldsymbol{A}\boldsymbol{B}u_0 + \boldsymbol{C}\boldsymbol{B}u_1$
> $y_3 = \boldsymbol{C}\boldsymbol{A}^2\boldsymbol{B}u_0 + \boldsymbol{C}\boldsymbol{A}\boldsymbol{B}u_1 + \boldsymbol{C}\boldsymbol{B}u_2$
> $\vdots$
> $y_N = \boldsymbol{C}\boldsymbol{A}^{N-1}\boldsymbol{B}u_0 + \boldsymbol{C}\boldsymbol{A}^{N-2}\boldsymbol{B}u_1 + \boldsymbol{C}\boldsymbol{A}^{N-3}\boldsymbol{B}u_2 + \cdots + \boldsymbol{C}\boldsymbol{B}u_{N-1}.$
>
> In matrix form, this set of equation can be rewritten as
>
> $$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}\boldsymbol{B} & 0 & 0 & \cdots & 0 \\ \boldsymbol{C}\boldsymbol{A}\boldsymbol{B} & \boldsymbol{C}\boldsymbol{B} & 0 & \cdots & 0 \\ \boldsymbol{C}\boldsymbol{A}^2\boldsymbol{B} & \boldsymbol{C}\boldsymbol{A}\boldsymbol{B} & \boldsymbol{C}\boldsymbol{B} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \boldsymbol{C}\boldsymbol{A}^{N-1}\boldsymbol{B} & \boldsymbol{C}\boldsymbol{A}^{N-2}\boldsymbol{B} & \boldsymbol{C}\boldsymbol{A}^{N-3}\boldsymbol{B} & \cdots & \boldsymbol{C}\boldsymbol{B} \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_{N-1} \end{bmatrix}.$$

(e) (1 point) What do the respective eigendirections of the observability and controllability Gramians represents?

> **Solution:** Assuming the eigenvalues of the observability (resp. controllability) Gramian have been sorted in decreasing order, the ith eigendirection corresponds to the ith most observable (resp. controllable) direction in phase space.