Optimization and Algorithms
February 6, 2023

Write your name: ______________________________

Write your student number: ______________________________

Write your answers (A, B, C, D, E, or F) to problems 1 to 3 in this box
Your answer to problem 1: ________
Your answer to problem 2: ________
Your answer to problem 3: ________

Exam

1. *Deconflicted trajectories.* A trajectory $\mathcal{T}$ of duration T in $\mathbf{R}^d$ is a sequence of T points in $\mathbf{R}^d$, denoted as $\mathcal{T} = \{x(1), x(2), \ldots, x(T)\}$, with $x(t) \in \mathbf{R}^d$ for $1 \leq t \leq T$. Note that t denotes discrete-time; thus t is an integer (such as $t = 0, 1, 2, 3. \ldots$).

Let $\mathcal{T}_1 = \{x_1(1), x_1(2), \ldots, x_1(T)\}$ and $\mathcal{T}_2 = \{x_2(1), x_2(2), \ldots, x_2(T)\}$ be two trajectories of duration T in $\mathbf{R}^d$. We say that $\mathcal{T}_1$ and $\mathcal{T}_2$ are space-deconflicted if $\|x_1(t) - x_2(s)\|_2 > \epsilon$ for $1 \leq t, s \leq T$, where ϵ is a given positive number. We say that $\mathcal{T}_1$ and $\mathcal{T}_2$ are time-deconflicted if $\|x_1(t) - x_2(t)\|_2 > \epsilon$ for $1 \leq t \leq T$.

Consider the following two controlled dynamic linear systems. The state of system 1 at time t is denoted by $x_1(t) \in \mathbf{R}^d$, for $1 \leq t \leq T$ and obeys the recursion

$$x_1(t+1) = A_1 x_1(t) + B_1 u_1(t), \quad 0 \leq t \leq T-1,$$

where $A_1 \in \mathbf{R}^{d \times d}$ and $B_1 \in \mathbf{R}^{d \times p}$ are given matrices, $x_1(0) \in \mathbf{R}^d$ is a given initial state and $u_1(t) \in \mathbf{R}^p$ is the control input of system 1 at time t, for $0 \leq t \leq T-1$. Note that the trajectory $\mathcal{T}_1$ depends on the inputs $u_1(t)$, $0 \leq t \leq T-1$.

Similarly, for system 2 we have

$$x_2(t+1) = A_2 x_2(t) + B_2 u_2(t), \quad 0 \leq t \leq T-1.$$

Note that the trajectory $\mathcal{T}_2$ depends on the inputs $u_2(t)$, $0 \leq t \leq T-1$.

Finally, let $\mathcal{T}_{\text{ref}} = \{r(1), r(2), \ldots, r(T)\}$ be a given, fixed reference trajectory of duration T in $\mathbf{R}^d$.

We want to design the control inputs $u_1(t)$ $(0 \leq t \leq T-1)$ and $u_2(t)$ $(0 \leq t \leq T-1)$ so that:

- the final state $x_1(T)$ of system 1 is as close as possible to a given, desired state $p_1 \in \mathbf{R}^d$;

- the final state $x_2(T)$ of system 2 is as close as possible to a given, desired state $p_2 \in \mathbf{R}^d$;
- the trajectories $\mathcal{T}_1$ and $\mathcal{T}_2$ are time-deconflicted;
- the trajectories $\mathcal{T}_1$ and $\mathcal{T}_{\text{ref}}$ are space-deconflicted;
- the trajectories $\mathcal{T}_2$ and $\mathcal{T}_{\text{ref}}$ are space-deconflicted.

One of the following problem formulations is suitable for the given context.

(A)

$$\begin{array}{ll} \underset{\{x_1(t),u_1(t),x_2(t),u_2(t)\}}{\text{minimize}} & \|x_1(T)-p_1\|_2^2+\|x_2(T)-p_2\|_2^2 \\ \text{subject to} & x_1(t+1)=A_1x_1(t)+B_1u_1(t), \quad 0\le t\le T-1 \\ & x_2(t+1)=A_2x_2(t)+B_2u_2(t), \quad 0\le t\le T-1 \\ & \min\{\|x_1(t)-x_2(s)\|_2 : 1,s\le t\le T\}<\epsilon \\ & \min\{\|x_1(t)-r(s)\|_2 : 1\le t,s\le T\}>\epsilon \\ & \min\{\|x_2(t)-r(s)\|_2 : 1\le t,s\le T\}>\epsilon \end{array} \tag{1}$$

(B)

$$\begin{array}{ll} \underset{\{x_1(t),u_1(t),x_2(t),u_2(t)\}}{\text{minimize}} & \|x_1(T)-p_1\|_2^2+\|x_2(T)-p_2\|_2^2 \\ \text{subject to} & x_1(t+1)=A_1x_1(t)+B_1u_1(t), \quad 0\le t\le T-1 \\ & x_2(t+1)=A_2x_2(t)+B_2u_2(t), \quad 0\le t\le T-1 \\ & \max\{\|x_1(t)-x_2(t)\|_2 : 1\le t\le T\}>\epsilon \\ & \max\{\|x_1(t)-r(s)\|_2 : 1\le t,s\le T\}>\epsilon \\ & \max\{\|x_2(t)-r(s)\|_2 : 1\le t,s\le T\}>\epsilon \end{array} \tag{2}$$

(C)

$$\begin{array}{ll} \underset{\{x_1(t),u_1(t),x_2(t),u_2(t)\}}{\text{minimize}} & \|x_1(T)-p_1\|_2^2+\|x_2(T)-p_2\|_2^2 \\ \text{subject to} & x_1(t+1)=A_1x_1(t)+B_1u_1(t), \quad 0\le t\le T-1 \\ & x_2(t+1)=A_2x_2(t)+B_2u_2(t), \quad 0\le t\le T-1 \\ & \min\{\|x_1(t)-x_2(t)\|_2 : 1\le t\le T\}<\epsilon \\ & \min\{\|x_1(t)-r(s)\|_2 : 1\le t,s\le T\}<\epsilon \\ & \min\{\|x_2(t)-r(s)\|_2 : 1\le t,s\le T\}<\epsilon \end{array} \tag{3}$$

(D)

$$\begin{array}{ll} \underset{\{x_1(t),u_1(t),x_2(t),u_2(t)\}}{\text{minimize}} & \|x_1(T)-p_1\|_2^2+\|x_2(T)-p_2\|_2^2 \\ \text{subject to} & x_1(t+1)=A_1x_1(t)+B_1u_1(t), \quad 0\le t\le T-1 \\ & x_2(t+1)=A_2x_2(t)+B_2u_2(t), \quad 0\le t\le T-1 \\ & \max\{\|x_1(t)-x_2(t)\|_2 : 1\le t\le T\}<\epsilon \\ & \max\{\|x_1(t)-r(s)\|_2 : 1\le t,s\le T\}<\epsilon \\ & \max\{\|x_2(t)-r(s)\|_2 : 1\le t,s\le T\}<\epsilon \end{array} \tag{4}$$

(E)

$$
\begin{array}{ll}
\underset{\{x_1(t),u_1(t),x_2(t),u_2(t)\}}{\text{minimize}} & \|x_1(T)-p_1\|_2^2 + \|x_2(T)-p_2\|_2^2 \\
\text{subject to} & x_1(t+1) = A_1x_1(t) + B_1u_1(t), \quad 0 \le t \le T-1 \\
& x_2(t+1) = A_2x_2(t) + B_2u_2(t), \quad 0 \le t \le T-1 \\
& \min\{\|x_1(t)-x_2(t)\|_2 : 1 \le t \le T\} > \epsilon \\
& \min\{\|x_1(t)-r(s)\|_2 : 1 \le t, s \le T\} > \epsilon \\
& \min\{\|x_2(t)-r(s)\|_2 : 1 \le t, s \le T\} > \epsilon
\end{array} \tag{5}
$$

(F)

$$
\begin{array}{ll}
\underset{\{x_1(t),u_1(t),x_2(t),u_2(t)\}}{\text{minimize}} & \|x_1(T)-p_1\|_2^2 + \|x_2(T)-p_2\|_2^2 \\
\text{subject to} & x_1(t+1) = A_1x_1(t) + B_1u_1(t), \quad 0 \le t \le T-1 \\
& x_2(t+1) = A_2x_2(t) + B_2u_2(t), \quad 0 \le t \le T-1 \\
& \min\{\|x_1(t)-x_2(s)\|_2 : 1 \le t, s \le T\} > \epsilon \\
& \min\{\|x_1(t)-r(s)\|_2 : 1 \le t, s \le T\} > \epsilon \\
& \min\{\|x_2(t)-r(s)\|_2 : 1 \le t, s \le T\} > \epsilon
\end{array} \tag{6}
$$

Which one?

Write your answer (A, B, C, D, E, or F) in the box at the top of page 1

2. *Unconstrained optimization.* Consider the optimization problem

$$\underset{x\in\mathbf{R}}{\text{minimize}} \quad e^{x-a} + e^{-x} + x^2 - 2x + x_+. \tag{7}$$

The point $x^\star = 0$ is a global minimizer of (7) for one of the following choices of a:

(A) $a = -2$

(B) $a = -1$

(C) $a = 0$

(D) $a = 1$

(E) $a = 2$

(F) $a = 3$

Which one?

Write your answer (A, B, C, D, E, or F) in the box at the top of page 1

Hint: the numerical values $\log(2) \simeq 0.7$ and $\log(3) \simeq 1.1$ might be useful

3. *Gradient descent algorithm.* Consider the function $f\colon \mathbf{R}^2 \to \mathbf{R}$ given by $f(a,b) = \frac{1}{2}a^2 + (a-b)^2$. Suppose we do one iteration of the gradient descent algorithm (applied to f) starting from the point

$$x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

and using the stepsize 1.

Which of the following points is the next iteration x_1?

(A)

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

(B)

$$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

(C)

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

(D)

$$\begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

(E)

$$\begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

(F)

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Write your answer (A, B, C, D, E, or F) in the box at the top of page 1

4. *Signal-denoising as a least-squares problem.* Consider the function $f\colon \mathbf{R}^n \to \mathbf{R}$, $f(r) = r^T D r$, where D is a given $n \times n$ diagonal matrix with positive diagonal entries:

$$D = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix},$$

with $d_i > 0$ for $1 \le i \le n$.

Consider the following optimization problem

$$\begin{array}{ll} \underset{s,v}{\text{minimize}} & \|s - \overline{s}\|_2^2 + f(v) \\ \text{subject to} & y = As + v, \end{array} \tag{8}$$

where the variables to optimize are $s \in \mathbf{R}^p$ and $v \in \mathbf{R}^n$; the matrix $A \in \mathbf{R}^{n \times p}$ and the vectors $y \in \mathbf{R}^n$, and $\overline{s} \in \mathbf{R}^p$ are given. This problem can be interpreted as a signal-denoising problem: we observe y and want to decompose it as the sum of a signal of interest s and noise v; we know that s should be close to the nominal signal $\overline{s}$ and that v should be close to zero (the larger the d_i, the more confident we are that the component v_i should be close to zero).

Problem (8) can be reduced to a least-squares problem involving only the variable s, that is, it can be reduced to a problem of the form

$$\underset{s}{\text{minimize}} \quad \|\mathcal{A}s - \beta\|_2^2 \tag{9}$$

for some matrix $\mathcal{A}$ and vector β.

Give $\mathcal{A}$ and β in terms of the constants D, y, A, and $\overline{s}$.

5. *A simple optimization problem.* Consider the function $f\colon \mathbf{R}^2 \to \mathbf{R}$, $f(x) = \frac{1}{2}x^T M x$, where

$$M = \begin{bmatrix} a & b \\ b & a \end{bmatrix}.$$

The constants a and b satisfy $0 < a < b$.

Solve in closed-form the optimization problem

$$\begin{array}{ll} \underset{x}{\text{maximize}} & f(x) \\ \text{subject to} & \mathbf{1}^T x = 1, \end{array} \tag{10}$$

where $\mathbf{1}$ denotes the vector $\mathbf{1} = (1, 1)$.

6. *A convex optimization problem.* Consider the following optimization problem

$$\begin{array}{ll} \underset{x_1, x_2, \ldots, x_n}{\text{minimize}} & g(x_1 - c_1) + g(x_2 - c_2) + \cdots + g(x_n - c_n) \\ \text{subject to} & a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b, \end{array} \tag{11}$$

where the variables to optimize are $x_i \in \mathbf{R}$, for $1 \le i \le n$. The vectors $a_i \in \mathbf{R}^p$, $1 \le i \le n$ and $b \in \mathbf{R}^p$ are given. The constants c_i, $1 \le i \le n$, are also given. The function $g\colon \mathbf{R} \to \mathbf{R}$ is defined as follows:

$$g(x) = \begin{cases} x^2, & \text{if } x \ge 0 \\ -x, & \text{if } x < 0. \end{cases}$$

Show that (11) is a convex optimization problem.

7. *A convex function based on a worst-case representation.* Show that the function $f\colon \mathbf{R} \to \mathbf{R}$,

$$f(x) = \max\left\{ \|(a+u)x - b\|_2 : \ \|u\|_2 = r \right\} \tag{12}$$

is convex, where the vectors $a, b \in \mathbf{R}^n$ and the constant $r > 0$ are given.

In words: f takes as input a number x and returns as output the largest value of the expression

$$\|(a+u)x - b\|_2$$

as u ranges over the sphere centered at the origin and with radius r.

Solution of Exam 2

Problem 1 Answer: E

Problem 2 Answer: B

Details ① note that $x^\star = 0$ is a global minimizer of a function $\phi : \mathbb{R} \to \mathbb{R}$ of the form $\phi(x) = f(x) + x_+$ (where $f : \mathbb{R} \to \mathbb{R}$ is a convex, continuously differentiable function) only if

$$\dot\phi(0^-) = \lim_{x \to 0^-} \phi(x) \leq 0 \quad \text{and} \quad \dot\phi(0^+) = \lim_{x \to 0^+} \phi(x) \geq 0$$

To see why, suppose, for example, that $\dot\phi(0^-) > 0$; then ϕ would be increasing when approaching $x^\star = 0$ from the left — that invalidates $x^\star = 0$ being a minimum. Similarly, suppose $\dot\phi(0^+) < 0$, then ϕ would be decreasing when departing from $x^\star = 0$ to the right — that invalidates $x^\star = 0$ being a minimum.

② In our case, $\phi(x) = \underbrace{e^{x-a} + e^{-x} + x^2 - 2x}_{f(x)} + x_+$.

Thus, $x^\star = 0$ minimizes ϕ when

$$\begin{aligned}
\dot\phi(0^-) \leq 0 \text{ and } \dot\phi(0^+) \geq 0 &\iff \dot f(0) \leq 0 \text{ and } \dot f(0) + 1 \geq 0 \\
&\iff e^{-a} - 3 \leq 0 \text{ and } e^{-a} - 2 \geq 0 \\
&\iff 2 \leq e^{-a} \leq 3 \\
&\iff \underbrace{-\log 3}_{\substack{\simeq \\ -1.1}} \leq a \leq \underbrace{-\log 2}_{\substack{\simeq \\ -0.7}}
\end{aligned}$$

The only number a in $\{-2,-1,0,1,2,3\}$ that satisfies the inequalities above is $a=-1$

Problem 3 Answer: C

<u>Details</u> ① The gradient descent algorithm is $x_{k+1} = x_k + \alpha_k d_k$, where $\alpha_k > 0$ is the stepsize and $d_k = -\nabla f(x_k)$

② For $x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, we have $\nabla f(x_0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Thus, for the stepsize $\alpha_0 = 1$, we have $x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (1)\begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Problem 4

First, note that $f(r) = r^T D r = \| D^{1/2} r \|_2^2$, where $D^{1/2} = \begin{bmatrix} d_1^{1/2} & & & \\ & d_2^{1/2} & & \\ & & \ddots & \\ & & & d_n^{1/2} \end{bmatrix}$

Now:
$$\begin{array}{ll} \min\limits_{s,v} & \|s-\bar{s}\|_2^2 + f(v) \\ \text{s.t.} & y = As + v \end{array} \iff \min_s \|s-\bar{s}\|^2 + f(y-As)$$

use the constraint to eliminate the variable v: $v = y - As$

$$\iff \min_s \|s-\bar{s}\|^2 + \| D^{1/2}(As-y)\|_2^2$$

use the identity: $\|A_1 s - b_1\|_2^2 + \|A_2 s - b_2\|_2^2 = \left\| \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} s - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\|_2^2$

$$\iff \min_s \Big\| \underbrace{\begin{bmatrix} I \\ D^{1/2}A \end{bmatrix}}_{\mathcal{A}} s - \underbrace{\begin{bmatrix} \bar{s} \\ D^{1/2}y \end{bmatrix}}_{\mathcal{B}} \Big\|_2^2$$

Problem 5

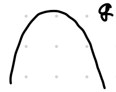
$$\max_{x_1, x_2} \quad f(x_1, x_2) \quad \text{s.t.} \quad x_1 + x_2 = 1 \quad \Longleftrightarrow \quad \max_{x_1} \quad f(x_1, 1 - x_1)$$

use the constraint to eliminate the variable x_2: $x_2 = 1 - x_1$

$$\Longleftrightarrow \quad \max_{x_1} \quad \underbrace{\tfrac{1}{2} a x_1^2 + b\, x_1 (1 - x_1) + \tfrac{1}{2} a (1 - x_1)^2}_{\shortparallel}$$

$$g(x_1) = (a - b)\, x_1^2 + (b - a)\, x_1 + \tfrac{1}{2} a$$

because $b > a$, this is a parabola that looks like this



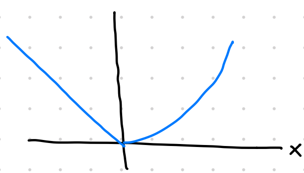
thus, the global maximizer occurs at the stationary point: $\dot{g}(x_1^*) = 0 \Leftrightarrow x_1^* = 1/2$

Because $x_2^* = 1 - x_1^* = 1/2$, the global maximizer of f (under the constraint $1^T x = 1$) is

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}.$$

Problem 6

① First, note that $g : \mathbb{R} \to \mathbb{R}$ is convex: this can be obtained directly from its graph:



② The given problem is

$$\min_{x} \quad \underbrace{\underbrace{g(x_1-c_1)}_{f_1(x)} + \cdots + \underbrace{g_n(x_n-c_n)}_{f_n(x)}}_{f(x)}$$

$$\text{s.t.} \quad Ax = b,$$

where

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

③ The constraints are linear. Thus, the problem is convex if f is convex

④ $f = f_1 + f_2 + \cdots + f_n$ where $f_i : \mathbb{R}^n \to \mathbb{R}$, $f_i(x) = g(x_i - c_i)$, for $1 \le i \le n$

<u>f_i is convex</u> because $f_i = g \circ h_i$, where $h_i(x) = [0 \cdots \underset{i}{1} \cdots 0]\,x - c_i$ is an affine map, and g is convex

<u>f is convex</u> because it is the sum of convex functions

Problem 7

① We start by obtaining a closed-form expression for $f(x)$, for fixed $x \in \mathbb{R}$:

$$f(x) = \max\{\|(a+u)x - b\|_2 : \|u\|_2 = r\}$$

$$= \max\{(\|ax-b+ux\|_2^2)^{1/2} : \|u\|_2 = r\}$$

$$= \max\{(\|ax-b\|_2^2 + \|u\|_2^2 x^2 + 2u^T(ax-b)x)^{1/2} : \|u\|_2 = r\}$$

$$= \max\{(\|ax-b\|_2^2 + r^2x^2 + 2u^T(ax-b)x)^{1/2} : \|u\|_2 = r\}$$

$$= \left(\|ax-b\|_2^2 + r^2x^2 + 2\max\{u^T(ax-b)x : \|u\|_2 = r\}\right)^{1/2}$$

$$= \left(\|ax-b\|_2^2 + r^2x^2 + 2r|x|\,\|ax-b\|_2\right)^{1/2}$$

$$= \left(\left(\|ax-b\|_2 + r|x|\right)^2\right)^{1/2}$$

$$= \underbrace{\|ax-b\|_2}_{f_1(x)} + r\underbrace{|x|}_{f_2(x)}$$

(2) <u>f_1 is convex</u> because $f_1 = g \circ h$, where $h(x) = ax-b$ is an affine map and $g = \|\cdot\|_2$ is convex

<u>f_2 is convex</u> because $f_2 = |\cdot|$

<u>f is convex</u> because $f = f_1 + r f_2$ is a nonnegative combination of convex functions