Optimization and Algorithms
February 22, 2022

Write your name: ____________________

Write your student number: ____________________

Exam

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: ________

1. *Simple convex function.* (3 points) One of the following six functions $\mathbf{R} \to \mathbf{R}$ is convex:

(A) $(1-(x-1)_+)_+$

(B) $|(x-1)_+ - 1|$

(C) -$(1-(x-1)_+)_+$

(D) $((x-1)_+ - 1)_+$

(E) -$((x-1)_+ - 1)_+$

(F) -$|(x-1)_+ - 1|$

Which one?

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

2. *Least-squares.* (2 points) Consider the following six optimization problems:

(A)

$$\underset{x \in \mathbf{R}^n}{\text{minimize}} \quad \|Ax - b\|_1^2 + \rho \|x\|_2^2$$

(B)

$$\underset{x \in \mathbf{R}^n}{\text{minimize}} \quad \|Ax - b\|_2^2 + \rho \|x\|_2^2$$

(C)

$$\underset{x \in \mathbf{R}^n}{\text{minimize}} \quad \|Ax - b\|_2^2 + \rho \|x\|_1^2$$

(D)

$$\underset{x \in \mathbf{R}^n}{\text{minimize}} \quad \|Ax - b\|_2 + \rho \|x\|_2^2$$

(E)

$$\underset{x \in \mathbf{R}^n}{\text{minimize}} \quad \|Ax - b\|_1^2 + \rho \|x\|_1^2$$

(F)

$$\underset{x \in \mathbf{R}^n}{\text{minimize}} \quad \|Ax - b\|_1 + \rho \|x\|_2$$

In each of the six problems above, the variable to optimize is $x \in \mathbf{R}^n$. The matrix A and the vector b are given. The scalar $\rho > 0$ is also given.

One of the optimization problems above is a least-squares problem.

Which one?

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

3. *Convex function.* (3 points) Let $f\colon \mathbf{R}^n \to \mathbf{R}$ be a convex function. One of the following functions is guaranteed to be convex:

(A) $|f(x)|$
(B) $f(x) + (f(x))^2$
(C) $(f(x))^2$
(D) $f(x)(f(x))^2$
(E) $|f(x)| + (f(x))^2$
(F) $f(x) + |f(x)|$

Which one?

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

4. *Robust portfolio selection.* (4 points) A problem that often occurs in finance has the following form

$$\begin{array}{ll} \underset{x \in \mathbf{R}^n}{\text{minimize}} & \max\{x^T V_1^T D V_1 x, x^T V_2^T D V_2 x\} \\ \text{subject to} & \min\{\mu_1^T x, \mu_2^T x\} \geq \alpha \\ & \mathbf{1}^T x = 1, \end{array} \tag{1}$$

where the variable to optimize is $x \in \mathbf{R}^n$.

The matrices $V_1 \in \mathbf{R}^{p \times n}$, $V_2 \in \mathbf{R}^{p \times n}$, and $D \in \mathbf{R}^{p \times p}$ are given, the matrix D being diagonal with positive entries in the diagonal:

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

with $d_i > 0$ for $i = 1, \ldots, p$.

The vectors $\mu_1 \in \mathbf{R}^n$, $\mu_2 \in \mathbf{R}^n$ and the scalar $\alpha \in \mathbf{R}$ are given. Finally, recall that the symbol $\mathbf{1}$ stands for the vector of dimension n with all components equal to one:

$$\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Show that the optimization problem (1) is convex.

5. *Mahalanobis projection.* (4 points) Consider the optimization problem

$$\begin{array}{ll} \underset{x}{\text{minimize}} & (x-\mu)^T \Sigma^{-1} (x-\mu) \\ \text{subject to} & a^T x = b, \end{array} \tag{2}$$

where the variable to optimize is $x \in \mathbf{R}^n$. The vector $\mu \in \mathbf{R}^n$ and the matrix $\Sigma \in \mathbf{R}^{n \times n}$ are given, with Σ being symmetric and positive definite.

Show that the optimal value of problem (2) is

$$\frac{(a^T \mu - b)^2}{a^T \Sigma a}.$$

6. *Strictly convex functions.* (4 points) Suppose that the functions $f_1 \colon \mathbf{R}^n \to \mathbf{R}$ and $f_2 \colon \mathbf{R}^n \to \mathbf{R}$ are both convex, and let $f \colon \mathbf{R}^n \to \mathbf{R}$ be defined as $f(x) = \max\{f_1(x), f_2(x)\}$ for each $x \in \mathbf{R}^n$. Is the function f strictly convex? If you think the answer is 'yes', then prove it; if you think the answer is 'no', then give a counter-example.

João Xavier

OPTIMIZATION AND ALGORITHMS

(Instituto Superior Técnico)

– Solution of the exam 2 –

Problem 1 D

Problem 2 B

Problem 3 F

Problem 4 The problem, rewritten in canonical form, is

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \underbrace{\max\{x^T V_1^T D V_1 x,\ x^T V_2^T D V_2 x\}}_{f(x)} \\ \text{subject to} & \underbrace{\max\{-\mu_1^T x, -\mu_2^T x\} + \alpha}_{g(x)} \leq 0 \\ & \underbrace{1^T x - 1}_{h(x)} = 0 \end{array}$$

This problem is convex only if f is convex, g is convex, and h is affine.

<u>f is convex</u> f can be expressed as $f = \max\{f_1, f_2\}$ where $f_i : \mathbb{R}^n \to \mathbb{R}$ is given by $f_i(x) = x^T V_i^T D V_i x$.

Function f_i is a quadratic function; so, f_i is convex if the matrix $V_i^T D V_i$ is semidefinite positive, that is, if

$$v^T \; V_i^T D V_i \; v \geq 0 \qquad \text{for each } v \in \mathbb{R}^n.$$

Now note that

$$v^T V_i^T D V_i \; v = v_i^T V_i^T D^{1/2} D^{1/2} V_i v = \|D^{1/2} V_i v\|_2^2 \geq 0.$$

$$D^{1/2} = \begin{bmatrix} \sqrt{d_1} & & \\ & \ddots & \\ & & \sqrt{d_p} \end{bmatrix}$$

Thus, each f_i is convex. Being the maximum of convex functions, the function f is convex.

<u>g is convex</u> $\quad g$ can be written as $g = g_1 + g_2$, where

$$g_1 : \mathbb{R}^n \to \mathbb{R} \qquad g_1(x) = \max\{-\mu_1^T x, -\mu_2^T x\}$$
$$g_2 : \mathbb{R}^n \to \mathbb{R} \qquad g_2(x) = \alpha.$$

The function g_1 is the pointwise maximum of two convex functions (in fact, of two affine functions). As such, g_1 is convex.

The function g_2 is convex (in fact, affine).

Being the sum of two convex functions, g is convex.

<u>h is affine</u> $\quad$ Obvious.

Problem 5 The problem, written in canonical form, is

$$\begin{aligned} \min_x \quad & \underbrace{(x-\mu)^T \Sigma^{-1} (x-\mu)}_{f(x)} \\ \text{s.t.} \quad & \underbrace{a^T x - b}_{h(x)} = 0 \end{aligned}$$

The associated KKT system is

$$\begin{cases} \nabla f(x) = \nabla h(x)\, \lambda \\ h(x) = 0 \end{cases} \to \begin{cases} 2\Sigma^{-1}(x-\mu) = a\lambda \\ a^T x = b \end{cases} \to \begin{cases} x = \mu + \frac{\lambda}{2}\Sigma^{-1} a \\ \text{———} \end{cases}$$

$$\begin{cases} \text{———} \\ a^T\left(\mu + \frac{\lambda}{2}\Sigma^{-1} a\right) = b \end{cases} \to \begin{cases} \text{———} \\ \lambda = \dfrac{2(b - a^T\mu)}{a^T \Sigma^{-1} a} \end{cases} \to \begin{cases} x = \mu + \dfrac{b - a^T\mu}{a^T\Sigma^{-1}a}\Sigma^{-1} a \\ \text{———} \end{cases}$$

note: OK because $a \neq 0$ and $\Sigma^{-1} \succ 0$

This computation shows $x^* = \mu + \dfrac{b - a^T\mu}{a^T\Sigma a}\Sigma a$ solves the KKT system.

However, by itself, it does not show x^* solves the optimization problem, which requires an extra argument. One such argument is through convexity: the problem is convex because f is convex and h is affine. Indeed:

<u>f is convex</u> The function f is a quadratic: $f(x) = x^T\Sigma^{-1}x$, with Hessian matrix $\nabla^2 f(x) = 2\Sigma^{-1}$. Because Σ is positive definite, so is Σ^{-1}, and therefore $2\Sigma^{-1}$. This implies f is convex.

<u>h is affine</u> Obvious.

Now that we know $x^* = \mu + \dfrac{b - a^T\mu}{a^T\Sigma a}\Sigma a$ solves the optimization problem, we can find the optimal value:

$$
\begin{aligned}
f(x^*) &= \left(\cancel{\mu} + \frac{b - a^T\mu}{a^T\Sigma a}\Sigma a - \cancel{\mu}\right)^T \Sigma^{-1}\left(\cancel{\mu} + \frac{b - a^T\mu}{a^T\Sigma a}\Sigma a - \cancel{\mu}\right) \\
&= \frac{(b - a^T\mu)^2}{(a^T\Sigma a)^2}\, a^T\Sigma\, \cancel{\Sigma^{-1}}\,\cancel{\Sigma}\, a \\
&= \frac{(a^T\mu - b)^2}{a^T\Sigma a}.
\end{aligned}
$$

Problem 6 A function $f: \mathbb{R}^n \to \mathbb{R}$ is strictly convex if

$$f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y)$$

for $\lambda \in]0,1[$ and $x \neq y$.

Consider the functions $f_1: \mathbb{R}^n \to \mathbb{R}$, $f_1(x) = 0$, and $f_2: \mathbb{R}^n \to \mathbb{R}$, $f_2(x) = 1$. These are convex functions (in fact, affine). For these functions, we have $f = \max\{f_1, f_2\}$ to be $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = 1$. This is not a strictly convex function because the inequality is not satisfied, say, by $x = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$, $y = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, and $\lambda = 1/2$:

$$f((1-\lambda)x + \lambda y) = 1$$

$$(1-\lambda) f(x) + \lambda f(y) = 1.$$