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 $\quad$ 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 $x^* = \mu + \frac{b - a^T\mu}{a^T\Sigma a}\Sigma a$ 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 $x^* = \mu + \frac{b - a^T\mu}{a^T\Sigma a}\Sigma a$ 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 $\lambda = 1/2$:

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

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