João Xavier

OPTIMIZATION AND ALGORITHMS

(Instituto Superior Técnico)

– Solution of the Mock Exam –

Problem 1 B

Problem 2 E

Problem 3 B

Problem 4 The problem

$$\min_{c,R} \underbrace{\sum_{m=1}^{M} \omega_m \left(\left(\|c - x_m\|_2 - R \right)_+ \right)^2 + \rho R^2}_{f(c,R)}$$

subject to $R \geq 0$

has at least one global minimizer if the function $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ given above is continuous and coercive.

The function f is clearly continuous. We now show that f is coercive. To show that f is coercive we need to show that $f(c_k, R_k) \to +\infty$ (as $k \to +\infty$) whenever $\|(c_k, R_k)\|_2 \to +\infty$ (as $k \to +\infty$).

So, suppose $\|(c_k, R_k)\|_2 \to +\infty$, which means $\|c_k\|_2^2 + R_k^2 \to +\infty$. We are going to show that

$$\underbrace{\omega_1 \left(\left(\|c_k - x_1\|_2 - R_k \right)_+ \right)^2 + \rho R_k^2}_{\phi(c_k, R_k)} \to +\infty.$$

(Because $f(c_k, R_k) = \phi(c_k, R_k) + \sum_{m=2}^{M} \omega_m \left(\left(\|c_k - x_m\|_2 - R_k \right)_+ \right)^2 \geq \phi(c_k, R_k)$, it will follow that $f(c_k, R_k) \to +\infty$.)

Our first step is to create a lower bound for $\phi(c,R) = \omega_1 \left(\left(\|c - x_1\|_2 - R \right)_+ \right)^2 + \rho R^2$:

case 1 if $\|c-x_1\|_2 - R \geq 0$, then

$$\phi(c,R) = \omega_1 \left(\|c-x_1\|_2 - R\right)^2 + \rho R^2$$

$$= \omega_1 \begin{bmatrix} \|c-x_1\|_2 & R \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \|c-x_1\|_2 \\ R \end{bmatrix} + \rho \begin{bmatrix} \|c-x_1\|_2 & R \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \|c-x_1\|_2 \\ R \end{bmatrix}$$

$$= \begin{bmatrix} \|c-x_1\|_2 & R \end{bmatrix} \underbrace{\begin{bmatrix} \omega_1 & -\omega_1 \\ -\omega_1 & \omega_1+\rho \end{bmatrix}}_{A} \begin{bmatrix} \|c-x_1\|_2 \\ R \end{bmatrix}$$

$$\geq \lambda \left(\|c-x_1\|_2^2 + R^2\right),$$

where λ is the minimum eigenvalue of A (we used the inequality $v^T M v \geq \lambda_{min}(M) \|v\|_2^2$, valid for any $d \times d$ symmetric matrix M and any vector $v \in \mathbb{R}^d$). Note that $\lambda > 0$ because A is positive definite.

case 2 if $\|c-x_1\|_2 - R < 0$, then

$$\phi(c,R) = \rho R^2 = \frac{\rho}{2} R^2 + \frac{\rho}{2} R^2$$

$$\geq \frac{\rho}{2} \|c-x_1\|^2 + \frac{\rho}{2} R^2$$

$$= \frac{\rho}{2} \left(\|c-x_1\|^2 + R^2\right),$$

where the inequality is due to $\|c-x_1\|_2 < R$.

From case 1 and case 2, we see that

$$\phi(c,R) \geq \alpha \left(\|c-x_1\|^2 + R^2\right),$$

where $\alpha = \min\{\lambda, \rho/2\} > 0$.

Now, if $\|(c_k, R_k)\|_2 \to +\infty$, then $\|c_k - x_1\|^2 + R_k^2 \to +\infty$ (indeed, $\|(c_k, R_k)\|_2 \to +\infty$ means that the distance from (c_k, R_k) to the origin $(0,0)$ grows to infinity; thus, the distance from (c_k, R_k) to the point $(x_1, 0)$ also grows to infinity; finally, note that $\|c_k - x_1\|^2 + R_k^2$ is just the squared-distance from (c_k, R_k) to $(x_1, 0)$).

Therefore, $\alpha \left(\|c_k - x_1\|^2 + R_k^2\right) \to +\infty$, which implies $\phi(c_k, R_k) \to +\infty$.

Problem 5 For the problem

$$\min_{u_1,\dots,u_K} \quad \underbrace{\sum_{k=1}^{K} \max\{|a_k^T u_k - b_k|, |c_k^T u_k - d_k|\}}_{f(u_1,\dots,u_K)}$$

$$\text{s.t.} \quad \underbrace{\|u_2 - u_1\|_2 - U}_{g_1(u_1,\dots,u_K)} \le 0$$

$$\underbrace{\|u_3 - u_2\|_2 - U}_{g_2(u_1,\dots,u_K)} \le 0$$

$$\vdots$$

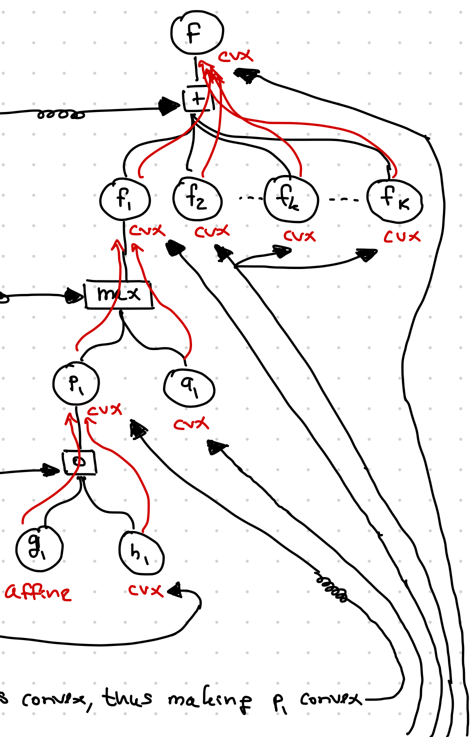
$$\underbrace{\|u_K - u_{K-1}\|_2 - U}_{g_K(u_1,\dots,u_K)} \le 0$$

to be convex, we need to show that

① f is convex

② $g_1, g_2, \dots, g_K$ are convex

① f is convex



- We start by decomposing $f = f_1 + \dots + f_K$, where

  $f_k(u_1,\dots,u_K) = \max\{|a_k^T u_k - b_k|, |c_k^T u_k - d_k|\}$

- Focusing now on f_1, we decompose it as

  $f_1(u_1,\dots,u_K) = \max\{p_1(u_1,\dots,u_K), q_1(u_1,\dots,u_K)\}$,

  where $p_1(u_1,\dots,u_K) = |a_1^T u_1 - b_1|$ and $q_1(u_1,\dots,u_K) = |c_1^T u_1 - d_1|$

- The function p_1 is convex because $p_1 = g_1 \circ h_1$,

  where $g_1(u_1,\dots,u_K) = a_1^T u_1 - b_1$ is affine and

  $h_1(z) = |z|$ is convex

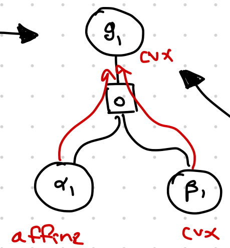
- An affine function followed by a convex function is convex, thus making p_1 convex

- The function q_1 is convex (same reasoning as for p_1)
- The maximum of convex functions is convex, thus making f_1 convex
- The functions $f_2, \dots, f_K$ are convex (same reasoning as for f_1)
- The sum of convex function is convex, thus making f convex

(2) g_1 is convex

- We decompose $g_1 = \alpha_1 \circ \beta_1$,

  where $\alpha_1(u_1, \dots, u_K) = u_2 - u_1$ and $\beta_1(z) = \|z\|_2 - v$



- The map α_1 is affine and β_1 is a convex function

  (β_1 is the sum of known convex functions)

- An affine map followed by a convex function is a convex function, making g_1 convex

The functions $g_2, \dots, g_K$ are convex, by the same reasoning

Problem 6

We are given the following data:

(a) x^* is the global minimizer of
$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & s^T x = r \end{aligned}$$

(b) x_k^* is the global minimizer of $\displaystyle\min_x \quad f(x) + \frac{1}{2} c_k (s^T x - r)^2$

(c) $c_k \uparrow +\infty$ and $x_k^* \to \bar{x}$.

We want to show $\bar{x} = x^*$.

One way to show this is to show that $\bar{x}$ is a global minimizer of $\displaystyle\min_x \ f(x)$ s.t. $s^T x = r$,

that is, to show that $\bar{x}$ satisfies the KKT system: $\begin{cases} \text{(i)} & \exists_\lambda : \nabla f(\bar{x}) = s\lambda \\ \text{(ii)} & s^T \bar{x} = r. \end{cases}$

We start by establishing (ii):

- From (b), we have

$$\nabla f(x_k^*) + c_k (s^T x_k^* - r) s = 0,$$

which implies

$$s^T \nabla f(x_k^*) + c_k \|s\|_2^2 (s^T x_k^* - r) = 0,$$

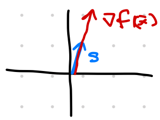
and, in turn,

$$s^T x_k^* - r = -\frac{s^T \nabla f(x_k^*)}{c_k \|s\|_2^2}.$$

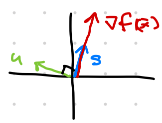
- Taking the limit $k \to +\infty$ on both sides (and recalling $c_k \uparrow +\infty$ and $x_k^* \to \bar{x}$) gives

$$s^T \bar{x} - r = 0.$$

We now establish (i):

- Note that (i) is equivalent to say that the vector $\nabla f(\bar{x}) \in \mathbb{R}^2$ is aligned with the vector $s \in \mathbb{R}^2$. (Example: 

)

- Let $u \in \mathbb{R}^2$ be a vector orthogonal to $s \in \mathbb{R}^2$. (Example: 

)

- $\nabla f(\bar{x})$ is aligned with s if $\nabla f(\bar{x})$ is orthogonal to u, that is, if

$$u^T \nabla f(\bar{x}) = 0,$$

which we now show to be the case

- From (b), we have

$$\nabla f(x_k^*) + c_k (s^T x_k^* - r) s = 0 \Rightarrow u^T \nabla f(x_k^*) + c_k (s^T x_k^* - r) \underbrace{u^T s}_{= 0} = 0$$

$$\Rightarrow u^T \nabla f(x_k^*) = 0$$

$$\Rightarrow \lim_{k \to +\infty} u^T \nabla f(x_k^*) = 0$$

$$\Rightarrow u^T \nabla f(\bar{x}) = 0$$