João Xavier

OPTIMIZATION AND ALGORITHMS

(Instituto Superior Técnico)

– Solution of the exam –

Problem 1

D

Problem 2

E

Problem 3

B

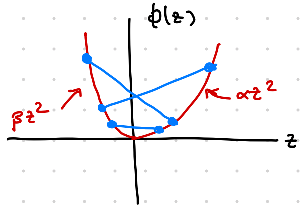
Problem 4

- Note that
$$f(s,r) = \sum_{k=1}^{K} \underbrace{\phi(s^T x_k + r - y_k)}_{f_k(s,r)} + \rho \underbrace{\|s\|_1}_{f_0(s,r)},$$

where $\phi : \mathbb{R} \to \mathbb{R}$, $\phi(z) = \alpha (z_-)^2 + \beta (z_+)^2$.

- The function ϕ is convex. This can be seen directly from its graph:

$$\phi(z) = \begin{cases} \alpha z^2 & ; \text{ if } z \geq 0 \\ \beta z^2 & ; \text{ if } z < 0 \end{cases} \longrightarrow$$



- The function $f_k : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, $f_k(s,r) = \phi(s^T x_k + r - y_k)$, can be decomposed as
$$f_k = \phi \circ g_k,$$
where $g_k : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, $g_k(s,r) = s^T x_k + r - y_k$.

Because g_k is affine and ϕ is convex, we conclude that f_k is convex (for $1 \leq k \leq K$)

- The function $f_0 : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, $f_0(s,r) = \|s\|_1$, can be decomposed as

$$f_0 = q \circ p,$$

where $p : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$, $p(s,r) = s$, and $q : \mathbb{R}^n \to \mathbb{R}$, $q(z) = \|z\|_1$.

Because p is an affine map and q is a (well-known) convex function, we conclude that f_0 is convex.

- Finally, $f = f_1 + \cdots + f_K + \rho f_0$, being a linear combination with nonnegative weights of convex functions, is itself convex.

Problem 5

The constraints $\begin{cases} x_1 = x_{\text{initial}} \\ x_{t+1} = x_t + D_t u_t, \quad \text{for } 1 \le t \le T-1 \end{cases}$ imply $x_T = x_{\text{initial}} + D_1 u_1 + \cdots + D_{T-1} u_{T-1}$.

Thus, our problem can be written as

$$\underset{u_1, \ldots, u_{T-1}}{\text{minimize}} \quad \tfrac{1}{2} \| x_{\text{initial}} + D_1 u_1 + \cdots + D_T u_T \|^2 + \frac{\rho}{2} \sum_{t=1}^{T-1} \|u_t\|^2,$$

or, in a more compact form, as

$$\underset{u}{\text{minimize}} \quad \underbrace{\tfrac{1}{2} \| x_{\text{initial}} + D u \|^2 + \frac{\rho}{2} \|u\|^2}_{f(u)},$$

where $D = [D_1 \; D_2 \; \cdots \; D_{T-1}]$ and $u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{T-1} \end{bmatrix}$.

The function f is strongly convex, which implies that it has a unique minimizer. Because f is differentiable, its global minimizer, say u^*, is the point satisfying

$$\nabla f(u^*) = 0 \quad \Rightarrow \quad D^T (x_{\text{initial}} + D u^*) + \rho u^* = 0$$

$$\Rightarrow \quad (D^T D + \rho I) u^* = -D^T x_{\text{initial}}$$

$$\Rightarrow \quad u^* = -(D^TD+\rho I)^{-1} D^T x_{initial}.$$

The last step is valid because $D^TD+\rho I$ is invertible; indeed, it is a positive-definite matrix: for $v \neq 0$, we have

$$v^T(D^TD+\rho I)v = \underbrace{\|Dv\|^2}_{\geq 0} + \underbrace{\rho}_{>0} \underbrace{\|v\|^2}_{>0} > 0.$$

Problem 6 Let x be fixed.

We have

$$e_{\lambda_1}[e_{\lambda_2}[f]](x) = \min\left\{ e_{\lambda_2}[f](u) + \frac{1}{2\lambda_1}(u-x)^2 : u \in \mathbb{R}\right\}$$

$$= \min\left\{ \min\left\{ f(v) + \frac{1}{2\lambda_2}(v-u)^2 : v \in \mathbb{R}\right\} + \frac{1}{2\lambda_1}(u-x)^2 : u \in \mathbb{R}\right\}$$

$$= \min\left\{ f(v) + \frac{1}{2\lambda_1}(u-x)^2 + \frac{1}{2\lambda_2}(v-u)^2 : u, v \in \mathbb{R}\right\}$$

$$= \min\left\{ f(v) + \boxed{\min\left\{\frac{1}{2\lambda_1}(u-x)^2 + \frac{1}{2\lambda_2}(v-u)^2 : u \in \mathbb{R}\right\}} : v \in \mathbb{R}\right\}$$

$$\overset{(*)}{=} \min\left\{ f(v) + \boxed{\frac{1}{2(\lambda_1+\lambda_2)}(v-x)^2} : v \in \mathbb{R}\right\}$$

$$= e_{\lambda_1+\lambda_2}[f](x)$$

(*) proof that $\min\left\{\frac{1}{2\lambda_1}(u-x)^2 + \frac{1}{2\lambda_2}(v-u)^2 : u \in \mathbb{R}\right\} = \frac{1}{2(\lambda_1+\lambda_2)}(v-x)^2$

- Let $\phi(u) = \frac{1}{2\lambda_1}(u-x)^2 + \frac{1}{2\lambda_2}(v-u)^2$, which is a strongly convex function

- Its global minimizer, say u^*, can be found by solving $\dot{\phi}(u^*)=0$:

$$\dot{\phi}(u^*)=0 \quad \Rightarrow \quad \frac{u^*-x}{\lambda_1} + \frac{u^*-v}{\lambda_2} = 0$$

$$\Rightarrow \quad u^* = \frac{\lambda_1^{-1}x + \lambda_2^{-1}v}{\lambda_1^{-1}+\lambda_2^{-1}}$$

• Finally,

$$\min\{\phi(u) : u \in \mathbb{R}\} = \phi(u^*)$$

$$= \tfrac{1}{2}\lambda_1^{-1}(u^* - x)^2 + \tfrac{1}{2}\lambda_2^{-1}(u^* - v)^2$$

$$= \tfrac{1}{2}\lambda_1^{-1}\left(\frac{\lambda_2^{-1}(v-x)}{\lambda_1^{-1}+\lambda_2^{-1}}\right)^2 + \tfrac{1}{2}\lambda_2^{-1}\left(\frac{\lambda_1^{-1}(x-v)}{\lambda_1^{-1}+\lambda_2^{-1}}\right)^2$$

$$= \tfrac{1}{2}\,\frac{\lambda_1^{-1}\lambda_2^{-2} + \lambda_2^{-1}\lambda_1^{-2}}{(\lambda_1^{-1}+\lambda_2^{-1})^2}(v-x)^2$$

$$= \tfrac{1}{2}\,\frac{\lambda_1^{-1}\lambda_2^{-1}\cancel{(\lambda_1^{-1}+\lambda_2^{-1})}}{(\lambda_1^{-1}+\lambda_2^{-1})^{\cancel{2}}}(v-x)^2$$

$$= \tfrac{1}{2}\,\frac{1}{\lambda_1+\lambda_2}(v-x)^2$$