Solution of the quiz

Problem 1 Answer: C

Details

(1) Because we want to maximize the distance between the position of the vehicle at time T (given by $Ex(T)$) and point b, we must minimize the negative of this quantity.

Problem 2 Answer: A

Details

(1) Note that all the six functions are continuous and coercive. Thus, they all have (at least) one global minimizer.

(2) Let's start by looking at the function in option A: $f(x) = e^x + 1/2\,|x|$.

The candidates for being global minimizers are the points $N \cup S$, where

- N = set of points at which f is not differentiable
- S = set of stationary points of f (points at which $\dot{f}(x) = 0$).

For the set N, we have $N = \{0\}$.

For the set S, we must solve $\dot{f}(x) = 0$ with $x \neq 0$:

- case $x > 0$: here, $f(x) = e^x + 1/2\,x$; thus, $\dot{f}(x) = e^x + 1/2$.

  This gives the equation $e^x = -1/2$, which does not have a solution for $x > 0$ (in fact, for no $x \in \mathbb{R}$);

• <u>case $x<0$</u> here, $f(x) = e^x - 1/2\, x$; thus, $\dot{f}(x) = e^x - 1/2$.

This gives the equation $e^x = 1/2$, which has the solution $x^* = -\log 2$ (which we accept because $x^* < 0$).

The set of candidates is $C = \{0, -\log 2\}$.

We now evaluate f on those points:

- $f(0) = e^0 + 1/2\,|0| = 1$
- $f(-\log 2) = e^{(-\log 2)} + 1/2\,|-\log 2| = 1/2 + 1/2\,\underbrace{\log 2}_{\simeq 0.7} < 1$

Because $f(-\log 2) < f(0)$, we conclude that 0 is not a global minimizer of f.

(We also found something extra: the global minimizer is $x^* = -\log 2$.)

③ For the functions in options B, C, D, E, and F, we always have $N = \{0\}$ and $S = \emptyset$.

Thus, for those functions, we just have one candidate, $x^* = 0$, which must be therefore the global minimizer.

Problem 3 Answer: C

<u>Details</u> ① Note that $C(u)\,v = C(v)\,u$ for any vectors $u, v \in \mathbb{R}^n$

(this is easily verified and highly suggested by looking at the available options)

② We have:

use ①

$$\|x - r\|_2^2 + \rho \sum_{k=1}^{K} \|C(x)a_k - b_k\|_2^2 = \|x - r\|_2^2 + \rho \sum_{k=1}^{K} \|C(a_k)x - b_k\|_2^2$$

for $\alpha \geq 0$: $\alpha\,\|u\|^2 = \|\sqrt{\alpha}\,u\|^2$

$$= \|x - r\|_2^2 + \sum_{k=1}^{K} \|\sqrt{\rho}\,C(a_k)x - \sqrt{\rho}\,b_k\|_2^2$$

for vectors $v_1, v_2, \ldots, v_K$:

$$\left\| \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_K \end{bmatrix} \right\|_2^2 = v_1^T v_1 + v_2^T v_2 + \cdots + v_K^T v_K = \|v_1\|_2^2 + \|v_2\|_2^2 + \cdots + \|v_K\|_2^2$$

$$= \left\| \begin{bmatrix} \sqrt{\rho}\,C(a_1)x - \sqrt{\rho}\,b_1 \\ \vdots \\ \sqrt{\rho}\,C(a_K)x - \sqrt{\rho}\,b_K \\ x - r \end{bmatrix} \right\|_2^2$$

For matrices $M_1, M_2, \ldots, M_K$, and vector u:

$$\begin{bmatrix} M_1 u \\ M_2 u \\ \vdots \\ M_K u \end{bmatrix} = \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_K \end{bmatrix} u$$

$$= \left\| \underbrace{\begin{bmatrix} \sqrt{\rho}\, C(a_1) \\ \vdots \\ \sqrt{\rho}\, C(a_K) \\ I \end{bmatrix}}_{\mathcal{A}} x - \underbrace{\begin{bmatrix} \sqrt{\rho}\, b_1 \\ \vdots \\ \sqrt{\rho}\, b_K \\ r \end{bmatrix}}_{\beta} \right\|_2^2$$

Problem D Answer: C

Details (1) Note that $D(u)v = D(v)u$ for any vectors $u, v \in \mathbb{R}^n$

(2) We have

$$f(x) = \| (A + R D(x))\theta + b \|_2^2$$

$$= \| R D(x)\theta + A\theta + b \|_2^2$$

use (1)

$$= \| \underbrace{R D(\theta)}_{\mathcal{A}} x + \underbrace{A\theta + b}_{\beta} \|_2^2$$

observe that the given optimization problem is a least-squares problem; thus, there is a global minimizer

$$= x^T \mathcal{A}^T \mathcal{A} x + 2\beta^T \mathcal{A} x + \beta^T \beta$$

(3) Because there is a global minimizer and f is differentiable everywhere, the global minimizer is to be found in $S = \{x : \nabla f(x) = 0\}$.

We now focus on solving $\nabla f(x) = 0$:

if $\phi(x) = x^T M x + p^T x$ (where $M = M^T$), then $\nabla \phi(x) = 2Mx + p$; in our case: $M = \mathcal{A}^T\mathcal{A}$ and $p = 2\mathcal{A}^T\beta$

$$\nabla f(x) = 0 \iff 2\mathcal{A}^T\mathcal{A}x + 2\mathcal{A}^T\beta = 0$$

$$\iff D(\theta)R^T R D(\theta)\, x = -D(\theta) R^T \beta$$

OK step because the matrix $D(\theta)R^T R D(\theta)$ is invertible: it's the product of three invertible matrices $D(\theta)$, $R^T R$, and $D(\theta)$. [$D(\theta)$ is invertible because $\theta_i \neq 0$ for all i; $R^T R$ is invertible because R is full column-rank]

$$\iff x = -\left(D(\theta) R^T R D(\theta)\right)^{-1} D(\theta) R^T \beta$$

$$= -D(\theta)^{-1} (R^T R)^{-1} R^T \beta$$

$$= -D(\theta)^{-1} (R^T R)^{-1} R^T (A\theta + b)$$