Solution of Exam 1 – Part A

Problem 1

Answer: A

Details

① Note that $S \not\subset B(c)$ if and only if $V \not\subset B(c)$. That is, the ball $B(c)$ does not cover S if and only if there exists a vector $v \in V$ such that $v \notin B(c)$.

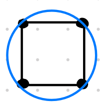
Thus:

$$S \not\subset B(c) \iff \exists_{v \in V} : v \notin B(c)$$

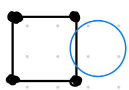
$$\iff \exists_{v \in V} : \|v - c\|_2 > r$$

$$\iff \underbrace{\max\{\|v-c\|_2 : v \in V\} > r.}_{\text{option A}}$$

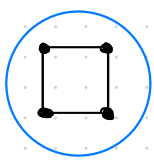
② Option B is incorrect because it allows



③ Option C is incorrect because it forbids



④ Option D is incorrect because it allows



⑤ Option E is incorrect because it allows the configuration shown in ②

⑥ Option F is incorrect because it allows the configuration shown in ②

Problem 2

Answer: E

Details

① The problem is convex, and the function

$$f(x) = (a^T x - b)^2 + (\|x\|_2 - c)_+^2 + \|x - d\|_2^2$$

is differentiable at $x^* = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$.

② So, x^* is a global minimizer for f if and only if

$$\nabla f(x^*) = 0 \iff 2(a^T x^* - b)a + 2(\|x^*\|_2 - c)_+ \frac{x^*}{\|x^*\|_2} + 2(x^* - d) = 0$$

$$\iff d = (a^T x^* - b)a + (\|x^*\|_2 - c)_+ \frac{x^*}{\|x^*\|_2} + x^*$$

$$\iff d = \underbrace{\left(\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ 4 \end{bmatrix} - 2 \right)}_{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \underbrace{\left(\left\| \begin{bmatrix} -3 \\ 4 \end{bmatrix} \right\|_2 - 10 \right)_+}_{0} \frac{1}{\left\| \begin{bmatrix} -3 \\ 4 \end{bmatrix} \right\|_2} \begin{bmatrix} -3 \\ 4 \end{bmatrix} + \begin{bmatrix} -3 \\ 4 \end{bmatrix}$$

$$\iff d = \begin{bmatrix} -1 \\ -1 \end{bmatrix} + \begin{bmatrix} -3 \\ 4 \end{bmatrix}$$

$$\iff d = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$$

Problem 3

Answer: B

Details

① Recall that an LS problem is an optimization problem of the form

$$\underset{x}{\text{min.}} \quad \| \mathcal{A} x - \beta \|_2^2 ,$$

for some matrix $\mathcal{A}$ and vector β

② In option A, the map $\text{rev}: \mathbb{R}^n \to \mathbb{R}^n$ is a linear map, that is, it can be written in the form $\text{rev}(x) = Mx$, for some matrix M. Indeed, taking the special case $n=4$, we have

$$\text{rev}\Bigg(\underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{x}\Bigg) = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{x} .$$

This means that

$$\underset{x}{\text{min.}} \quad \| A\,\text{rev}(x) - b \|_2^2 \iff \underset{x}{\text{min.}} \quad \| \underbrace{AM}_{\mathcal{A}} x - \underbrace{b}_{\beta} \|_2^2$$

that's a LS problem

③ The reasoning above applies also to options C, D, E, and F; only the matrix will change:

- option C

$$\text{circ}\Bigg(\underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{x}\Bigg) = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{x}$$

- option D

$$\text{cent}\Bigg(\underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{x}\Bigg) = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} - \begin{bmatrix} (x_1+x_2+x_3+x_4)/4 \\ (x_1+x_2+x_3+x_4)/4 \\ (x_1+x_2+x_3+x_4)/4 \\ (x_1+x_2+x_3+x_4)/4 \end{bmatrix} = \underbrace{\begin{bmatrix} 3/4 & -1/4 & -1/4 & -1/4 \\ -1/4 & 3/4 & -1/4 & -1/4 \\ -1/4 & -1/4 & 3/4 & -1/4 \\ -1/4 & -1/4 & -1/4 & 3/4 \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{x}$$

• option E

$$\text{trim}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{x}$$

$$\underbrace{}_{x}$$

• option F

$$\text{swap}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{x}$$

$$\underbrace{}_{x}$$

Problem 4

Answer: E

Details

① In the Newton algorithm $x_{k+1} = x_k + \alpha_k d_k$, the direction d_k is given by

$$d_k = - \nabla^2 f(x_k)^{-1} \nabla f(x_k)$$

② For $f(a,b) = (2a-b)^2 + b^2$, we have

$$\nabla f(a,b) = \begin{bmatrix} 8a - 4b \\ -4a + 4b \end{bmatrix} \qquad \nabla^2 f(a,b) = \begin{bmatrix} 8 & -4 \\ -4 & 4 \end{bmatrix}$$

③ Evaluating ∇f and $\nabla^2 f$ at $x_k = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ gives

$$\nabla f(x_k) = \begin{bmatrix} 0 \\ 4 \end{bmatrix} \qquad \nabla^2 f(x_k) = \begin{bmatrix} 8 & -4 \\ -4 & 4 \end{bmatrix}$$

④ Thus,

$$d_k = - \begin{bmatrix} 8 & -4 \\ -4 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

Solution of Exam 1 – Part B

Problem 1

① The problem

$$\min_{x} \underbrace{\sum_{k=1}^{K} \phi(a_k^T x - b_k) + \lambda \Psi(\|x-c\|_2 - r)}_{f(x)}$$

$$\text{s.t.} \quad \underbrace{\|x\|_1 - r}_{g(x)} \leq 0$$

is convex if f and g are convex, which we now show

② g is convex

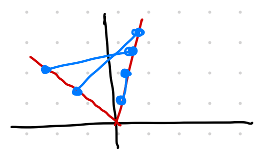
- g is the sum of a norm (a well-known convex function) with a constant (which does not affect convexity)
- so, g is convex

③ f is convex

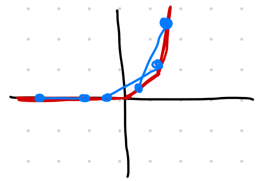
- write f as nonnegative linear combination of $f_1, \dots, f_K$, and h:

$$f = f_1 + f_2 + \dots + f_K + \lambda h,$$

where $f_k(x) = \phi(a_k^T x - b_k)$ and $h(x) = \Psi(\|x-c\|_2 - r)$

- if we show $f_1, \dots, f_K, h$ are convex, the function f will be convex
- f_k is convex: f_k is the composition $\phi \circ p_k$, where $p_k(x) = a_k^T x - b_k$ is affine and ϕ is convex (it suffices to look at the graph of ϕ: )

• <u>h is convex</u> h is the composition $\psi \circ q$, where ψ is convex and non-decreasing (as obvious from its graph



),

and q is convex: q is the composition of a convex function, $\|\cdot\|_2$, with the affine map $x \mapsto x - c$ (the additive constant $-r$ is irrelevant for the convexity of q)

Problem 2 Alice is right:

① Bob's problem is
$$\min_x \underbrace{\frac{1}{K}\sum_{k=1}^{K} \|x - p_k\|_2^2}_{f_{BOB}(x)}$$
$$\text{s.t. } s^T x = r,$$

whose KKT conditions are
$$\begin{cases} \nabla f_{BOB}(x) = s\lambda \\ s^T x = r \end{cases} \rightarrow \begin{cases} 2(x - \bar{p}) = s\lambda \\ s^T x = r, \end{cases}$$

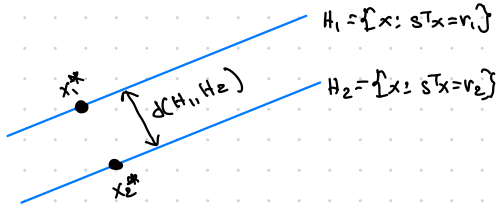
where $\bar{p} = \frac{1}{K}\sum_{k=1}^{K} p_k$

② Alice's problem is
$$\min_x \underbrace{\|x - \bar{p}\|_2^2}_{f_{ALICE}(x)}$$
$$\text{s.t. } s^T x = r,$$

whose KKT conditions are
$$\begin{cases} \nabla f_{ALICE}(x) = s\lambda \\ s^T x = r \end{cases} \rightarrow \begin{cases} 2(x - \bar{p}) = s\lambda \\ s^T x = r \end{cases}$$

③ The KKT conditions of Bob and Alice are the same; so, the minimizers are the same

Problem 3



① The goal is to find $d(H_1, H_2)$:

② The quantity $d(H_1, H_2)$ is equal to $\|x_1^* - x_2^*\|_2$ whenever x_1^*, x_2^* are solutions of the problem

$$\min_{x_1, x_2} \quad \underbrace{\tfrac{1}{2}\|x_1 - x_2\|^2}_{f(x_1, x_2)}$$

$$\text{s.t.} \quad \underbrace{s^T x_1 - r_1}_{h_1(x_1, x_2)} = 0$$

$$\underbrace{s^T x_2 - r_2}_{h_2(x_1, x_2)} = 0 ,$$

whose KKT conditions are

$$\begin{cases} \nabla f(x_1, x_2) = \nabla h_1(x_1, x_2)\lambda_1 + \nabla h_2(x_1, x_2)\lambda_2 \\ h_1(x_1, x_2) = 0 \\ h_2(x_1, x_2) = 0 \end{cases} \rightarrow \begin{cases} \begin{bmatrix} x_1 - x_2 \\ x_2 - x_1 \end{bmatrix} = \begin{bmatrix} s \\ 0 \end{bmatrix}\lambda_1 + \begin{bmatrix} 0 \\ s \end{bmatrix}\lambda_2 \\ s^T x_1 = r_1 \\ s^T x_2 = r_2 \end{cases} \rightarrow \begin{cases} x_1 - x_2 = s\lambda_1 & (i) \\ x_2 - x_1 = s\lambda_2 & (ii) \\ s^T x_1 = r_1 & (iii) \\ s^T x_2 = r_2 & (iv) \end{cases}$$

③ Subtract (iii) and (iv) to get

$$s^T(x_1 - x_2) = r_1 - r_2 ,$$

and plug $x_1 - x_2 = s\lambda_1$ (from (i)) to arrive at

$$s^T s \lambda_1 = r_1 - r_2 \quad \Rightarrow \quad \lambda_1 = \frac{r_1 - r_2}{\|s\|_2^2}$$

④ Using $\lambda_1 = \dfrac{r_1 - r_2}{\|s\|_2^2}$ in (i) gives $\quad x_1 - x_2 = \dfrac{(r_1 - r_2)}{\|s\|_2^2} s$

⑤ This means $\quad d(H_1, H_2) = \|x_1 - x_2\| = \dfrac{|r_1 - r_2|}{\|s\|_2}$

Problem 4

① The problem is
$$\min_x \underbrace{\left(\|x-c\|_2 - B\right)_+^2}_{f(x)}$$
$$\text{s.t.} \quad \underbrace{s^T x - r}_{h(x)} = 0,$$

whose KKT conditions are

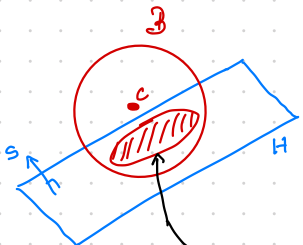
$$\begin{cases} \nabla f(x) = \nabla h(x)\lambda \\ h(x) = 0 \end{cases} \longrightarrow \begin{cases} 2\left(\|x-c\| - B\right)_+ \dfrac{x-c}{\|x-c\|_2} = s\lambda \\ s^T x = r \end{cases}$$

② Two cases can arise: $\left(\|x-c\|-B\right)_+ = 0$ or $\left(\|x-c\|-B\right)_+ > 0$

③ <u>case $\left(\|x-c\|-B\right)_+ = 0$</u>

- this means that we are investigating the existence of a minimizer x that satisfies $\|x-c\|_2 \le B$, that is, x is in the ball $\mathcal{B} = \{u : \|u-c\|_2 \le B\}$

- in this case, KKT system becomes
$$\begin{cases} 0 = s\lambda \ (\Rightarrow \lambda = 0 \text{ because } s \neq 0) \\ s^T x = r, \ (\Rightarrow x \in H) \end{cases}$$
with solutions being all pairs (x, λ) satisfying
$$x \in \underbrace{H \cap \mathcal{B}} \quad \text{and} \quad \lambda = 0$$



④ <u>case $\left(\|x-c\|_2-B\right)_+ > 0$</u>

- in this case, KKT system becomes
$$\begin{cases} 2\left(\|x-c\|_2 - B\right) \dfrac{x-c}{\|x-c\|_2} = s\lambda & \text{(i)} \\ s^T x = r & \text{(ii)} \end{cases}$$

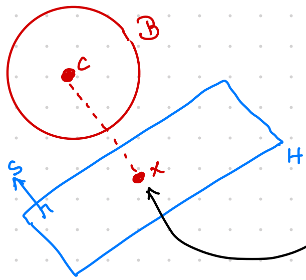
- (i) implies

$$x = c + \underbrace{\frac{\|x-c\|_2}{2(\|x-c\|_2 - B)}\lambda}_{\eta}\, s,$$

that is, x is of the form $x = c + \eta s$, for some $\eta \in \mathbb{R}$

- plug $x = c + \eta s$ into (ii) to get

$$s^T(c + \eta s) = r \Rightarrow \eta = \frac{r - s^T c}{\|s\|_2^2},$$



which means

$$x = c + \frac{r - s^T c}{\|s\|_2^2} s$$

- such x above is a solution in case it also satisfies $x \notin \mathcal{B}$