Κυρτή Ανάλυση
Εαρινό Εξάμηνο 2019–2020
Φυλλάδιο 2

1. Έστω K ανοικτό, κυρτό και μη κενό υποσύνολο του $\mathbb{R}^d$, και $f\colon K \to \mathbb{R}$ κυρτή συνάρτηση. Για $i \in \{1, \dots, d\}$ και $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d \in \mathbb{R}$ έστω

$$K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d} := \{x \in \mathbb{R}\colon (x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_d) \in K\}$$

και $f^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}\colon K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d} \to \mathbb{R}$ η συνάρτηση

$$f^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}(x) := f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_d).$$

Δείξτε ότι το $K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$ είναι ανοικτό διάστημα και η $f^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$ κυρτή συνάρτηση.

Λύση. Έστω $i \in \{1, \dots, d\}$ και $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d \in \mathbb{R}$. Για να δείξουμε ότι το σύνολο $K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$ είναι ανοικτό διάστημα αρκεί να δείξουμε ότι είναι ανοικτό και κυρτό, αφού ένα υποσύνολο του $\mathbb{R}$ είναι κυρτό ανν είναι διάστημα. Για την κυρτότητα πρώτα, έστω $y, z \in K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$ και $\lambda \in [0, 1]$. Τότε από τον ορισμό του $K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$ και το γεγονός ότι $y, z \in K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$, έχει κανείς ότι

$$(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_d) \in K \quad \text{και} \quad (x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_d) \in K.$$

Έπεται άμεσα από την κυρτότητα του K ότι

$$\begin{aligned}
&(x_1, \dots, x_{i-1}, \lambda y + (1-\lambda)z, x_{i+1}, \dots, x_d)\\
&\quad = (\lambda x_1 + (1-\lambda)x_1, \dots, \lambda x_{i-1} + (1-\lambda)x_{i-1}, \lambda y + (1-\lambda)z, \lambda x_{i+1} + (1-\lambda)x_{i+1}, \dots, \lambda x_d + (1-\lambda)x_d)\\
&\quad = \lambda(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_d) + (1-\lambda)(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_d) \in K
\end{aligned}$$

και από τον ορισμό του $K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$ πάλι έπεται ότι $\lambda y + (1-\lambda)z \in K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$. Αυτό δείχνει ότι το $K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$ είναι κυρτό υποσύνολο του $\mathbb{R}$, άρα διάστημα.

Έστω τώρα $x \in K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$. Τότε $(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_d) \in K$, και επειδή το K είναι ανοικτό, έπεται ότι υπάρχει $\varepsilon > 0$ με την ιδιότητα ότι

$$U_2((x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_d), \varepsilon) \subseteq K,$$

όπου ως συνήθως $U_2(a, r)$ συμβολίζει την ανοικτή μπάλα με κέντρο a και ακτίνα r ($r > 0$) στον $\mathbb{R}^d$, ως προς την ευκλείδεια μετρική $\|\cdot\|_2$. Τότε όμως, για κάθε $y \in (x - \varepsilon, x + \varepsilon)$ έχει κανείς ότι

$$\begin{aligned}
&\left\|(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_d) - (x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_d)\right\|_2\\
&\qquad = \sqrt{\sum_{j \in \{1,\dots,d\} \setminus \{i\}} \left|x_j - x_j\right|^2 + |x - y|^2} = |x - y| < \varepsilon,
\end{aligned}$$

και επομένως

$$(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_d) \in U_2((x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_d), \varepsilon) \subseteq K,$$

και αυτό σημαίνει ότι $y \in K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$. Αυτό δείχνει ότι $(x-\varepsilon, x+\varepsilon) \in K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$, και αφού το x ήταν τυχόν σημείο του $K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$, το τελευταίο είναι ανοικτό σύνολο.

Η κυρτότητα της συνάρτησης $f^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$ αποδεικνύεται όπως και η κυρτότητα του συνόλου $K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$. Αν $y, z \in K^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}$ και $\lambda \in [0,1]$, τότε

$$(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_d) \in K \quad \text{και} \quad (x_1,\dots,x_{i-1},z,x_{i+1},\dots,x_d) \in K,$$

και άρα

$$\begin{aligned}
f^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}(\lambda x + (1-\lambda)y) &= f(x_1,\dots,x_{i-1},\lambda y + (1-\lambda)z, x_{i+1},\dots,x_d)\\
&= f(\lambda(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_d) + (1-\lambda)(x_1,\dots,x_{i-1},z,x_{i+1},\dots,x_d))\\
&\leqslant \lambda f(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_d) + (1-\lambda) f(x_1,\dots,x_{i-1},z,x_{i+1},\dots,x_d)\\
&= \lambda f^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}(x) + (1-\lambda) f^{(i)}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_d}(y),
\end{aligned}$$

η ανισότητα στο προτελευταίο βήμα από την κυρτότητα της συνάρτησης f. □

2. Έστω K κυρτό και μη κενό υποσύνολο του $\mathbb{R}^d$, και $f\colon K \to \mathbb{R}$ κυρτή συνάρτηση. Δείξτε ότι το σύνολο στάθμης $F_r := \{x \in K\colon f(x) \leqslant r\}$ είναι, για κάθε $r \in \mathbb{R}$, κυρτό (πιθανώς κενό). Δείξτε επίσης με παράδειγμα ότι το αντίστροφο δεν ισχύει· υπάρχει δηλαδή συνάρτηση f για την οποία το F_r είναι κυρτό για κάθε $r \in \mathbb{R}$, αλλά η f δεν είναι κυρτή.

Λύση. Έστω $r \in \mathbb{R}$ και $x, y \in F_r$ (αν F_r μη κενό)· τότε $f(x) \leqslant r$ και $f(y) \leqslant r$. Έστω επίσης $\lambda \in [0,1]$. Τότε από την κυρτότητα του K έπεται ότι $\lambda x + (1-\lambda)y \in K$, και από την κυρτότητα της f ότι

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

επειδή και $\lambda, 1-\lambda \geqslant 0$· επομένως $\lambda x + (1-\lambda)y \in F_r$. Άρα το F_r είναι κυρτό σύνολο.

Έστω τώρα η συνάρτηση

$$f(x) := \begin{cases} x & \text{για} \quad x < 0\\ 0 & \text{για} \quad 0 \leqslant x \leqslant 1\\ x-1 & \text{για} \quad x > 1\cdot \end{cases}$$

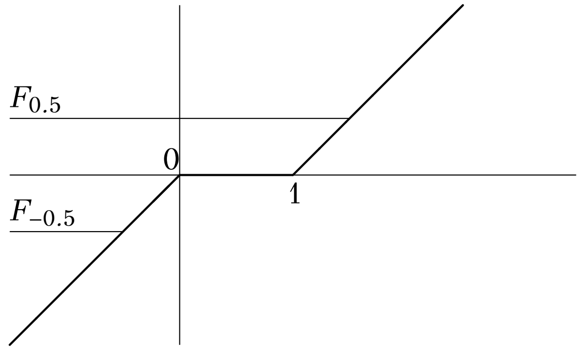
η f προφανώς δεν είναι κυρτή συνάρτηση: για παράδειγμα,

$$f(0) = 0 > \tfrac{1}{2} f(-1) + \tfrac{1}{2} f(1) = -\tfrac{1}{2}.$$

Από την άλλη, τα σύνολα στάθμης F_r της f είναι τα

$$F_r = \begin{cases} (-\infty, r] & \text{αν} \quad r < 0\\ (-\infty, 1] & \text{αν} \quad 0 \leqslant r \leqslant 1\\ (-\infty, r+1] & \text{αν} \quad r > 0, \end{cases}$$

που είναι φυσικά κυρτά σύνολα για όλα τα $r \in \mathbb{R}$.



$\square$

3. Έστω $K \subseteq \mathbb{R}^d$ κυρτό και μη κενό και $n \in \mathbb{N}$.

(α) Δείξτε ότι αν $f: K \to [0, \infty)$ κυρτή, τότε $x \mapsto [f(x)]^n$ είναι κυρτή.

(β) Δείξτε με (αντιπαράδειγμα) ότι το (α) δεν ισχύει εν γένει αν παραληφθεί η υπόθεση $f(x) \geqslant 0$ για κάθε $x \in K$.

Λύση. (α) **Λήμμα**: *Αν $f: K \to \mathbb{R}$ κυρτή, όπου $K \subseteq \mathbb{R}^d$ κυρτό και μη κενό, και $g: I \to \mathbb{R}$ κυρτή και αύξουσα (μη φθίνουσα), όπου I κυρτό υποσύνολο του $\mathbb{R}$ που περιέχει το $f(K)$, τότε $g \circ f$ είναι κυρτή.*
Απόδειξη: Έστω $x, y \in K$ και $\lambda \in [0, 1]$.

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

από την κυρτότητα της f, και τότε

$$g(f(\lambda x + (1-\lambda)y)) \leqslant g(\lambda f(x) + (1-\lambda)f(y)) \leqslant \lambda g(f(x)) + (1-\lambda)g(f(y)),$$

η πρώτη ανισότητα από την μονοτονία της g και η δεύτερη από την κυρτότητα της g. $\square$
Για την Άσκηση τώρα παίρνει κανείς $g(x) := x^n$, $x \in [0, \infty)$, η οποία είναι φυσικά αύξουσα και κυρτή όταν $n \geqslant 1$: $f'(x) = nx^{n-1} > 0$ και $f''(x) = n(n-1)x^{n-2} \geqslant 0$, για $x > 0$.
(β) η $f(x) := -\ln x$, $x > 0$, είναι κυρτή: $f''(x) = x^{-2} > 0$· όμως η $g(x) := (-\ln x)^2$, $x > 0$, έχει $g''(x) = 2(1 - \ln x)/x^2$, η οποία αλλάζει πρόσημο στο $x = \mathrm{e}$ και άρα η g δεν είναι κυρτή. $\square$

4. Ένα κυρτό υποσύνολο K του $\mathbb{R}^d$ με τουλάχιστον δύο σημεία λέγεται κυρτός κώνος με κορυφή την αρχή των αξόνων αν $x \in K \Rightarrow \lambda x \in K \ \forall \lambda \in [0, \infty)$. Έστω $n \in \mathbb{N}$ φυσικός αριθμός· μία συνάρτηση $f: K \to \mathbb{R}$, ορισμένη σε έναν κυρτό κώνο K, λέγεται ομογενής βαθμού n αν $f(\lambda x) = \lambda^n f(x)$ για κάθε $x \in K$ και $\lambda \geqslant 0$. Δείξτε ότι αν η f είναι κυρτή και ομογενής βαθμού n, και $f(x) > 0$ για κάθε $x \neq 0$, τότε η συνάρτηση $x \mapsto [f(x)]^{1/n}$ είναι κυρτή.

Λύση. Έστω $x, y \in [0,1]$ και $\lambda \in [0,1]$. Θέλουμε να δείξουμε ότι

$$
\begin{aligned}
& [f(\lambda x + (1-\lambda)y)]^{1/n} \leqslant \lambda [f(x)]^{1/n} + (1-\lambda)[f(y)]^{1/n} \\
\Leftrightarrow \quad & f(\lambda x + (1-\lambda)y) \leqslant \left(\lambda [f(x)]^{1/n} + (1-\lambda)[f(y)]^{1/n}\right)^n . \\
\Leftrightarrow \quad & \frac{f(\lambda x + (1-\lambda)y)}{\left(\lambda [f(x)]^{1/n} + (1-\lambda)[f(y)]^{1/n}\right)^n} \leqslant 1 \\
\Leftrightarrow \quad & f\left(\frac{\lambda x + (1-\lambda)y}{\lambda [f(x)]^{1/n} + (1-\lambda)[f(y)]^{1/n}}\right) \leqslant 1,
\end{aligned}
$$

η τελευταία σειρά από την ομογένεια της f. Τώρα για να χρησιμοποιήσουμε και την κυρτότητα της f, χρειαζόμαστε κυρτό συνδυασμό των x και y· γράφουμε λοιπόν

$$
\begin{aligned}
& \frac{\lambda x + (1-\lambda)y}{\lambda [f(x)]^{1/n} + (1-\lambda)[f(y)]^{1/n}} \\
& = \frac{\lambda [f(x)]^{1/n}}{\lambda [f(x)]^{1/n} + (1-\lambda)[f(y)]^{1/n}} \frac{x}{[f(x)]^{1/n}} + \frac{(1-\lambda)[f(y)]^{1/n}}{\lambda [f(x)]^{1/n} + (1-\lambda)[f(y)]^{1/n}} \frac{y}{[f(y)]^{1/n}} \\
& = \mu x' + (1-\mu)y',
\end{aligned}
$$

όπου

$$
\mu := \frac{\lambda [f(x)]^{1/n}}{\lambda [f(x)]^{1/n} + (1-\lambda)[f(y)]^{1/n}}, \quad x' := \frac{x}{[f(x)]^{1/n}} \quad \text{και} \quad y' := \frac{y}{[f(y)]^{1/n}}.
$$

Από την κυρτότητα της f παίρνουμε ότι

$$
f\left(\frac{\lambda x + (1-\lambda)y}{\lambda [f(x)]^{1/n} + (1-\lambda)[f(y)]^{1/n}}\right) = f(\mu x' + (1-\mu)y') \leqslant \mu f(x') + (1-\mu) f(y')\cdot
$$

από την ομογένεια της f και πάλι

$$
f(x') = f\left(\frac{x}{[f(x)]^{1/n}}\right) = 1 \quad \text{και} \quad f(y') = f\left(\frac{y}{[f(y)]^{1/n}}\right) = 1,
$$

οπότε τελικά πράγματι έχουμε ότι

$$
f\left(\frac{\lambda x + (1-\lambda)y}{\lambda [f(x)]^{1/n} + (1-\lambda)[f(y)]^{1/n}}\right) \leqslant \mu + (1-\mu) = 1.
$$

□

5. Δείξτε ότι μία συνεχής συνάρτηση $f\colon K \to \mathbb{R}$, ορισμένη σε ένα μη κενό κυρτό $K \subseteq \mathbb{R}^d$, είναι κυρτή ανν $f(\frac{1}{2}x + \frac{1}{2}y) \leqslant \frac{1}{2}f(x) + \frac{1}{2}f(y)$ για κάθε $x, y \in K$.
(Σημείωση: αυτό δεν ισχύει χωρίς την υπόθεση της συνέχειας της f.)

Λύση. Προφανώς όταν η f είναι κυρτή, η ανισότητα

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

αφού ισχύει για κάθε $\lambda \in [0,1]$ ισχύει και για $\lambda = \frac{1}{2}$, για κάθε $x, y \in K$.

Αντίστροφα, έστω ότι

$$f(\tfrac{1}{2}x+\tfrac{1}{2}y) \leqslant \tfrac{1}{2}f(x)+\tfrac{1}{2}f(y) \qquad \forall\, x, y \in K\cdot \tag{1}$$

ο ισχυρισμός τότε είναι ότι η ανισότητα

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

ισχύει και για κάθε $\lambda = k2^{-n}$, $n \in \mathbb{N}$, $k \in \{0, 1, \ldots, 2^n\}$. Πράγματι, αυτό μπορεί να δειχθεί με επαγωγή στο n. Η υπόθεση της Άσκησης μας δίνει ότι ισχύει για $n = 1$, για κάθε $x, y \in K$. Έστω τώρα ότι ισχύει για κάποιο $n \in \mathbb{N}$, και όλα τα $k \in \{0, 1, \ldots, 2^n\}$, και όλα τα $x, y \in K$. Τότε, για $k \in \left\{0, 1, \ldots, 2^{n+1}\right\}$, έχει κανείς ότι

$$\frac{k}{2^{n+1}}x + \left(1 - \frac{k}{2^{n+1}}\right)y = \tfrac{1}{2}\left[\frac{k}{2^n}x + \left(1 - \frac{k}{2^n}\right)y\right] + \tfrac{1}{2}y, \tag{2α$'$}$$

και επίσης

$$\frac{k}{2^{n+1}}x + \left(1 - \frac{k}{2^{n+1}}\right)y = \tfrac{1}{2}\left[\frac{k-2^n}{2^n}x + \left(1 - \frac{k-2^n}{2^n}\right)y\right] + \tfrac{1}{2}x\cdot \tag{2β$'$}$$

αν $k \leqslant 2^n$, τότε έχει κανείς από την πρώτη από αυτές τις ισότητες ότι

$$\begin{aligned}
f\left(\frac{k}{2^{n+1}}x + \left(1 - \frac{k}{2^{n+1}}\right)y\right) &\leqslant \tfrac{1}{2}f\left(\frac{k}{2^n}x + \left(1 - \frac{k}{2^n}\right)y\right) + \tfrac{1}{2}f(y) \\
&\leqslant \tfrac{1}{2}\left[\frac{k}{2^n}f(x) + \left(1 - \frac{k}{2^n}\right)f(y)\right] + \tfrac{1}{2}f(y) \\
&= \frac{k}{2^{n+1}}f(x) + \left(1 - \frac{k}{2^{n+1}}\right)f(y),
\end{aligned}$$

η πρώτη ανισότητα από την (1) και η δεύτερη από την επαγωγική υπόθεση· αν πάλι $k > 2^n$, τότε $k - 2^n \in \{0, 1, \ldots, 2^n\}$, και η (2β$'$), μαζί με την (1) και την επαγωγική υπόθεση πάλι, δίνουν όμοια ότι

$$\begin{aligned}
f\left(\frac{k}{2^{n+1}}x + \left(1 - \frac{k}{2^{n+1}}\right)y\right) &\leqslant \tfrac{1}{2}f\left(\frac{k-2^n}{2^n}x + \left(1 - \frac{k-2^n}{2^n}\right)y\right) + \tfrac{1}{2}f(x) \\
&\leqslant \tfrac{1}{2}\left[\frac{k-2^n}{2^n}f(x) + \left(1 - \frac{k-2^n}{2^n}\right)f(y)\right] + \tfrac{1}{2}f(x) \\
&= \frac{k}{2^{n+1}}f(x) + \left(1 - \frac{k}{2^{n+1}}\right)f(y).
\end{aligned}$$

Αυτό αποδεικνύει τον ισχυρισμό.

Έστω τώρα αυθαίρετο $\lambda \in (0,1)$. Για κάθε $n \in \mathbb{N}$, έστω k_n εκείνος ο ακέραιος στο $\{0, 1, \ldots, 2^n - 1\}$ για τον οποίο $k_n 2^{-n} \leqslant \lambda < (k_n + 1)2^{-n}\cdot$ τέτοιος ακέραιος υπάρχει και είναι μοναδικός γιατί τα διαστήματα $[k2^{-n}, (k+1)2^{-n})$, $k \in \{0, 1, \dots, 2^n - 1\}$ διαμερίζουν το $[0, 1)$. Τότε $k_n 2^{-n} \to \lambda$ καθώς $n \to \infty$ διότι

$$0 \leqslant \lambda - \frac{k_n}{2^n} \leqslant \frac{k_n + 1}{2^n} - \frac{k_n}{2^n} = 2^{-n} \xrightarrow{(n\to\infty)} 0.$$

Από την συνέχεια της f και τον ισχυρισμό έπεται τώρα ότι

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= \lim_{n\to\infty} f\left(\frac{k_n}{2^n}x + \left(1 - \frac{k_n}{2^n}\right)y\right) \\ &\leqslant \lim_{n\to\infty}\left[\frac{k_n}{2^n}f(x) + \left(1 - \frac{k_n}{2^n}\right)f(y)\right] = \lambda f(x) + (1-\lambda)f(y), \end{aligned}$$

για δοθέντα, αλλά αυθαίρετα $x, y \in K$. Άρα η f είναι κυρτή. □

6. (Μέθοδος ελαχίστων τετραγώνων.) Έστω $(x_1, y_1), \dots, (x_n, y_n)$ σημεία του $\mathbb{R}^2$, με τουλάχιστον δύο x_i διάφορα μεταξύ τους. Δείξτε ότι υπάρχει μοναδική ευθεία $y = ax+b$ «σε ελάχιστη απόσταση» από τα σημεία $(x_1, y_1), \dots, (x_n, y_n)$· δηλαδή δείξτε ότι η $\sum_{i=1}^n [y_i - (ax_i + b)]^2$ έχει μοναδικό ολικό ελάχιστο ως προς a, b.

Λύση. Έστω

$$Q(a, b) := \sum_{i=1}^n [y_i - (ax_i + b)]^2 \qquad a, b \in \mathbb{R}.$$

Αν η Q έχει ελάχιστο σε κάποιο σημείο (a, b), τότε πρέπει $\nabla Q(a, b) = 0$. Υπολογίζει κανείς

$$\nabla Q(a, b) = \begin{bmatrix} (\partial/\partial a)Q(a, b) \\ (\partial/\partial b)Q(a, b) \end{bmatrix} = -2\begin{bmatrix} \sum_{i=1}^n (y_i - ax_i - b)x_i \\ \sum_{i=1}^n (y_i - ax_i - b) \end{bmatrix},$$

και βλέπει εύκολα ότι $\nabla Q(a, b) = 0$ ανν

$$\begin{aligned} a &= \frac{n\sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right)}{n\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2} \\ b &= \frac{1}{n}\sum_{i=1}^n y_i - \frac{1}{n}\sum_{i=1}^n x_i \frac{n\sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i\right)\left(\sum_{i=1}^n y_i\right)}{n\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}. \end{aligned}$$

εισάγοντας τους συμβολισμούς $\overline{x} := n^{-1}\sum_{i=1}^n x_i$ και $\overline{y} := n^{-1}\sum_{i=1}^n y_i$, αυτά μπορούν να γραφούν και ως

$$\begin{aligned} a &= \frac{\sum_{i=1}^n x_i y_i - n\overline{xy}}{\sum_{i=1}^n x_i^2 - n\overline{x}^2} = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^n (x_i - \overline{x})^2} \\ b &= \overline{y} - \overline{x}\frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^n (x_i - \overline{x})^2}. \end{aligned}$$

Οι μερικές παράγωγοι δεύτερης τάξης είναι

$$\frac{\partial^2 Q}{\partial a^2} = 2\sum_{i=1}^{n} x_i^2 \qquad \frac{\partial^2 Q}{\partial b \partial a} = 2\sum_{i=1}^{n} x_i$$

$$\frac{\partial^2 Q}{\partial a \partial b} = 2\sum_{i=1}^{n} x_i \qquad \frac{\partial^2 Q}{\partial b^2} = 2n\cdot$$

δηλαδή η Εσσιανή της Q σε ένα σημείο (a,b) είναι

$$HQ(a,b) = 2\begin{bmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{bmatrix}.$$

Ο πίνακας αυτός είναι θετικά ορισμένος αφού το $(1,1)$ στοιχείο του $2\sum_{i=1}^{n} x_i^2$ είναι (γνήσια) θετικό, δηλ. $2\sum_{i=1}^{n} x_i^2 > 0$ (δεν είναι όλα τα x_i ίδια άρα δεν είναι όλα μηδέν), και το ίδιο και η ορίζουσά του:

$$\det(HQ(a,b)) = 4n\sum_{i=1}^{n} x_i^2 - 4\left(\sum_{i=1}^{n} x_i\right)^2 = 4n\sum_{i=1}^{n}(x_i - \overline{x})^2 > 0\cdot$$

το > 0 προκύπτει από το γεγονός ότι δεν είναι όλα τα x_i ίδια. Έπεται από αυτό κατ΄ αρχήν ότι η Q έχει τοπικό ελάχιστο στο σημείο (a,b) στο οποίο $\nabla Q(a,b) = 0$, επειδή η εσσιανή σε αυτό το σημείο είναι θετικά ορισμένη (Απειροστικός Λογισμός ΙΙΙ), και κατόπιν, επειδή η εσσιανή είναι θετικά ορισμένη παντού, ότι η Q είναι γνήσια κυρτή συνάρτηση και επομένως έχει μοναδικό ελάχιστο. □

7. Έστω K μη κενό ανοικτό και κυρτό υποσύνολο του $\mathbb{R}^d$ και $f\colon K \to \mathbb{R}$ κυρτή. Δείξτε ότι:

(α) Για κάθε $x \in K$ και $u \in \mathbb{R}^d$, το όριο

$$D_u^+ f(x) := \lim_{t\downarrow 0} \frac{f(x+tu) - f(x)}{t}$$

υπάρχει.

(β) Για $x \in K$ σταθερό, η συνάρτηση $G_x(u) := D_u^+ f(x)$, $u \in \mathbb{R}^d$, είναι κυρτή και θετικά ομογενής (δηλαδή $G_x(\lambda u) = \lambda G_x(u)$ για κάθε $u \in \mathbb{R}^d$ και $\lambda \geqslant 0$).

(γ) Για $x \in K$ σταθερό, υπάρχει $v \in \mathbb{R}^d$ τέτοιο ώστε $D_u^+ f(x) \geqslant \langle v, u\rangle$ $\forall\, u \in \mathbb{R}^d$.

(δ) Αν $\partial f(x)$ το υποδιαφορικό της f στο x, τότε

$$\partial f(x) = \{v \in \mathbb{R}^d : D_u^+ f(x) \geqslant \langle v, u\rangle \ \forall\, u \in \mathbb{R}^d\}.$$

(ε) Δείξτε επίσης ότι $D_u^+ f(x) \geqslant D_u^- f(x)$, όπου $D^- f(x) := -D^+ f_{-u}(x)$, για κάθε $x \in K$. (Η κατά κατεύθυνση παράγωγος $D_u f(x)$ της f στο x υπάρχει όταν $D_u^+ f(x) = D_u^- f(x)$.)

Λύση. (α) Έστω $g(t) := f(x+tu)$, $t \in \{s \in \mathbb{R} \colon x+su \in K\} =: I$· όπως έχει αποδειχθεί στο μάθημα, το I είναι κυρτό σύνολο (άρα διάστημα) και η g είναι κυρτή συνάρτηση. Άρα η δεξιά πλευρική παράγωγός της στο $0 \in I$ υπάρχει και αυτό σημαίνει ότι το

$$\lim_{t\downarrow 0} \frac{f(x+tu)-f(x)}{t} = \lim_{t\downarrow 0} \frac{g(t)-g(0)}{t} = g'_+(0)$$

υπάρχει, και μάλιστα

$$\frac{f(x+tu)-f(x)}{t} = \frac{g(t)-g(0)}{t} \geqslant g'_+(x) = D_u^+ f(x) \tag{3}$$

για κάθε $t > 0$ και τέτοιο ώστε $x+tu \in K$· το γεγονός ότι $0 \in I$ είναι συνέπεια του ότι $x \in K$ και το γεγονός ότι το I είναι μη τετριμμένο ανοικτό διάστημα έπεται από το γεγονός ότι το K είναι ανοικτό. Υπενθυμίζεται ότι η ύπαρξη του παραπάνω ορίου για την κυρτή συνάρτηση g (μίας μεταβλητής), καθώς επίσης και η (3), οφείλονται στο γεγονός ότι η συνάρτηση

$$t \mapsto \frac{g(t)-g(0)}{t} = \frac{g(t)-g(0)}{t-0}$$

είναι αύξουσα στο σύνολο $I \setminus \{0\}$, το οποίο με την σειρά του είναι συνέπεια του Λήμματος των τριών χορδών.

Σημειώνεται ακόμη ότι αν $u = 0$, τότε η συνάρτηση g είναι η σταθερή συνάρτηση $g(t) = f(x)$ $\forall\, t \in \mathbb{R}$ ($I = \mathbb{R}$ σ' αυτήν την περίπτωση)· επομένως $D_0^+ f(x) = g'_+(0) = 0$ για κάθε x.

(β) Από τον ορισμό της $G_x(u)$, έχει κανείς ότι

$$\begin{aligned} G_x(\lambda u) = D_{\lambda u}^+ f(x) = \lim_{t\downarrow 0} \frac{f(x+t\lambda u)-f(x)}{t} &= \lambda \lim_{t\downarrow 0} \frac{f(x+t\lambda u)-f(x)}{t\lambda} \\ &= \lambda \lim_{s\downarrow 0} \frac{f(x+su)-f(x)}{s} = \lambda D_u^+ f(x) = \lambda G_x(u), \end{aligned}$$

και εφόσον βέβαια $\lambda > 0$ (ώστε $s = \lambda t \downarrow 0$ όταν $t \downarrow 0$), και άρα η G_x είναι θετικά ομογενής. Επίσης, για $u, v \in \mathbb{R}$ και $\lambda \in [0,1]$,

$$\begin{aligned} G_x(\lambda u + (1-\lambda)v) &= D_{\lambda u+(1-\lambda)v}^+ f(x) \\ &= \lim_{t\downarrow 0} \frac{f(x+t[\lambda u+(1-\lambda)v])-f(x)}{t} \\ &= \lim_{t\downarrow 0} \frac{f(\lambda(x+tu)+(1-\lambda)(x+tv))-f(x)}{t} \\ &\leqslant \lim_{t\downarrow 0} \frac{\lambda f(x+tu)+(1-\lambda)f(x+tv)-f(x)}{t} \\ &= \lambda \lim_{t\downarrow 0} \frac{f(x+t\lambda u)-f(x)}{t} + (1-\lambda) \lim_{t\downarrow 0} \frac{f(x+t\lambda v)-f(x)}{t} \\ &= \lambda D_u^+ f(x) + (1-\lambda) D_v^+ f(x) \\ &= \lambda G_x(u) + (1-\lambda) G_x(v), \end{aligned}$$

η ανισότητα από την κυρτότητα της f και το γεγονός ότι $t > 0$. Επομένως η $u \mapsto G_x(u)$ είναι και κυρτή.
(γ) Αφού η $u \mapsto G_x(u)$ είναι κυρτή, έχει σε κάθε σημείο $u \in \mathbb{R}^d$, άρα και στο $u = 0$, υπερεπίπεδο στήριξης· υπάρχει δηλαδή $v \in \mathbb{R}^d \setminus \{0\}$, τέτοιο ώστε

$$G_x(u) \geqslant \langle v, u\rangle + G_x(0) = \langle v, u\rangle \qquad \forall\, u \in \mathbb{R}^d.$$

Άρα $D_u^+ f(x) = G_x(u) \geqslant \langle v, u\rangle$ $\forall\, u \in \mathbb{R}^d$.
(δ) Εξ΄ ορισμού,

$$\begin{aligned} \partial f(x) &= \{v \in \mathbb{R}^d : f(y) \geqslant \langle v, y-x\rangle + f(x)\ \forall\, y \in K\} \\ &= \{v \in \mathbb{R}^d : f(y) - f(x) \geqslant \langle v, y-x\rangle\ \forall\, y \in K\}\cdot \end{aligned}$$

έστω επίσης

$$F(x) := \{v \in \mathbb{R}^d :\ D_u^+ f(x) \geqslant \langle v, u\rangle\ \forall\, u \in \mathbb{R}^d\}.$$

Έστω τώρα $v \in \partial f(x)$ και $u \in \mathbb{R}^d$. Τότε

$$f(x+tu) - f(x) \geqslant \langle v, (x+tu) - x\rangle = t\langle v, u\rangle$$

για κάθε t τέτοιο ώστε $x + tu \in K$, και αν επιπλέον $t > 0$, έχει κανείς ότι

$$\frac{f(x+tu) - f(x)}{t} \geqslant \langle v, u\rangle$$

για κάθε $t > 0$ και τέτοιο ώστε $x + tu \in K$· αφήνοντας το $t \downarrow 0$ παίρνει κανείς ότι

$$D_u^+ f(x) = \lim_{t\downarrow 0} \frac{f(x+tu) - f(x)}{t} \geqslant \langle v, u\rangle.$$

Αφού το u ήταν αυθαίρετο, έπεται ότι $v \in F(x)$. Αυτό δείχνει ότι $\partial f(x) \subseteq F(x)$.

Αντίστροφα, έστω τώρα $v \in F(x)$ και $y \in K$. Τότε $D_u^+ f(x) \geqslant \langle v, u\rangle$ για κάθε $u \in \mathbb{R}^d$ και άρα και για $u := y - x$. Από την (3) έχει όμως κανείς ότι

$$\frac{f(x+tu) - f(x)}{t} \geqslant D_u^+ f(x)$$

για κάθε $t > 0$ και τέτοιο ώστε $x + tu \in K$, όποτε, για $t = 1$ παίρνει κανείς ότι

$$f(y) - f(x) = f(x+u) - f(x) \geqslant D_u^+ f(x) \geqslant \langle v, u\rangle = \langle v, y-x\rangle\cdot$$

επιτρέπεται να πάρει κανείς $t = 1$ γιατί $1 > 0$ και $x + 1u = y \in K$. Αφού η παραπάνω ανισότητα ισχύει για κάθε $y \in K$, έπεται ότι $v \in \partial f(x)$. Αυτό δείχνει τον αντίστροφο εγκλεισμό $F(x) \subseteq \partial f(x)$ και άρα πρέπει $F(x) = \partial f(x)$.
(ε) Έστω πάλι $g(t) := f(x+tu)$, $t \in \{s \in \mathbb{R} :\ x + su \in K\} =: I$· όπως παρατηρήθηκε στο (α), $0 \in I$ και το I είναι μη τετριμμένο ανοικτό διάστημα. Επίσης $D_u^+ f(x) = g'_+(0)$. Η g είναι κυρτή

συνάρτηση μίας μεταβλητής και άρα $g'_-(0) \leqslant g'_+(0)$. Όμως

$$\begin{aligned} D_u^- f(x) = -D_{-u}^+ f(x) &= -\lim_{t\downarrow 0} \frac{f(x+t(-u)) - f(x)}{t} \\ &= \lim_{t\downarrow 0} \frac{f(x+(-t)u) - f(x)}{(-t)} \\ &= \lim_{s\uparrow 0} \frac{f(x+su) - f(x)}{s} \\ &= \lim_{s\uparrow 0} \frac{g(s) - g(0)}{s} = g'_-(0) \leqslant g'_+(0) = D_u^+ f(x). \end{aligned}$$

□

8. Έστω $K \subseteq \mathbb{R}^d$ μη κενό, κυρτό και ανοικτό, $f\colon K \to \mathbb{R}$ κυρτή συνάρτηση, και $x \in K$. Δείξτε ότι το υποδιαφορικό $\partial f(x)$ της f στο x είναι μη κενό, κυρτό και συμπαγές υποσύνολο του $\mathbb{R}^d$.

Λύση. Έστω $x \in K$. Το ότι $\partial f(x) \neq \varnothing$ είναι Θεώρημα που έχει αποδειχθεί στο μάθημα. Το ότι το υποδιαφορικό είναι κυρτό έπεται άμεσα: αν $u, v \in \partial f(x)$ και $\lambda \in [0,1]$, τότε $f(y) \geqslant \langle u, y-x\rangle + f(x)$ για κάθε $y \in K$ και $f(y) \geqslant \langle v, y-x\rangle + f(x)$ για κάθε $y \in K$, και άρα

$$\begin{aligned} f(y) &= \lambda f(y) + (1-\lambda) f(y) \\ &\geqslant \lambda[\langle u, y-x\rangle + f(x)] + (1-\lambda)[\langle v, y-x\rangle + f(x)] = \langle \lambda u + (1-\lambda)v, y-x\rangle + f(x) \end{aligned}$$

για κάθε $y \in K$, επειδή $\lambda, 1-\lambda \geqslant 0$ και επειδή το εσωτερικό γινόμενο είναι γραμμική συνάρτηση ως προς κάθε μεταβλητή (εδώ φυσικά χρησιμοποιήθηκε μόνο η γραμμικότητα ως προς την πρώτη μεταβλητή). Έπεται ότι $\lambda u + (1-\lambda)v \in \partial f(x)$. Άρα το $\partial f(x)$ είναι κυρτό.

Θα αποδειχθεί τώρα ότι το $\partial f(x)$ είναι φραγμένο. Κατ' αρχήν υπάρχει $r > 0$ τέτοιο ώστε η κλειστή σφαίρα $B(x,r) = \{y \in \mathbb{R}^d : \|y-x\| \leqslant r\}$ να περιέχεται στο K. Η f είναι συνεχής και άρα φραγμένη στο συμπαγές $B(x,r)$· έστω M ένα φράγμα: $\left|f(y)\right| \leqslant M \ \forall\, y \in B(x,r)$. Έστω τώρα $v \in \partial f(x)$· χωρίς βλάβη της γενικότητας μπορούμε να υποθέσουμε ότι $v \neq 0$ (αφού αν το $\partial f(x) \setminus \{0\}$ είναι φραγμένο, τότε και το $\partial f(x)$ είναι επίσης φραγμένο). Το

$$I := \{t \in \mathbb{R} : x + tv \in K\}$$

είναι μη τετριμμένο ανοικτό διάστημα και πρέπει

$$f(x+tv) - f(x) \geqslant t\|v\|^2 \tag{4}$$

για κάθε $t \in I$, επειδή $v \in \partial f(x)$ και $y = x + tv \in K$. Το I περιέχει το $t = r/\|v\|$, επειδή $x + tv \in B(x,r) \subseteq K$ για αυτό το t, αφού

$$\left\|\left(x + \frac{r}{\|v\|}v\right) - x\right\| = r\cdot$$

επιπλέον, αφού $x \in B(x,r)$ και $x + tv \in B(x,r)$ για αυτό το t, έχει κανείς ότι $\left|f(x+tv)\right| \leqslant M$ και $\left|f(x)\right| \leqslant M$, και άρα

$$f(x+tv) - f(x) \leqslant \left|f(x+tv) - f(x)\right| \leqslant \left|f(x+tv)\right| + \left|f(x)\right| \leqslant 2M. \tag{5}$$

Από τις (4) και (5) έπεται τώρα ότι

$$r\|v\| = t\|v\|^2 \leqslant 2M.$$

Αυτό δείχνει ότι το $\partial f(x)$ φράσσεται από το $2M/r$.

Τέλος, θα δειχθεί ότι το υποδιαφορικό $\partial f(x)$ είναι και κλειστό, οπότε θα είναι συμπαγές (κλειστό και φραγμένο). Έστω $v_n \in \partial f(x)$, $n \in \mathbb{N}$, και έστω ότι $v_n \to v$ καθώς $n \to \infty$, για κάποιο $v \in \mathbb{R}^d$. Έστω $y \in K$. Τότε

$$f(y) - f(x) \geqslant \langle v_n, y - x\rangle \qquad \forall\, n \in \mathbb{N},$$

και άρα

$$f(y) - f(x) \geqslant \lim_{n\to\infty} \langle v_n, y - x\rangle = \langle v, y - x\rangle, \tag{6}$$

η ισότητα στο τέλος επειδή

$$\left|\langle v_n, y - x\rangle - \langle v, y - x\rangle\right| = \left|\langle v_n - v, y - x\rangle\right| \leqslant \|v_n - v\| \left\|y - x\right\| \xrightarrow{(n\to\infty)} 0,$$

από την ανισότητα Cauchy–Schwartz και την σύγκλιση $v_n \to v$. Αφού το $y \in K$ ήταν αυθαίρετο, η (6) δείχνει ότι $v \in \partial f(x)$. Αυτό τώρα δείχνει ότι το $\partial f(x)$ είναι κλειστό. □

9. (α) Βρείτε το υποδιαφορικό $\partial f(0)$ της συνάρτησης $f(x) := \|x\|_2$, $x \in \mathbb{R}^d$, στο 0.
(β) Ομοίως για την συνάρτηση $f(x) := \|x\|_1$, $x \in \mathbb{R}^d$.
[Υπενθύμιση: $\|x\|_p := \left(\sum_{i=1}^d |x_i|^p\right)^{1/p}$ για $x \in \mathbb{R}^d$ και $p > 0$, και $\|x\|_\infty := \max_{i\in\{1,\dots,d\}} |x_i|$.]

Λύση. (α) Από τον ορισμό έχει κανείς ότι

$$\begin{aligned}\partial f(0) &= \{u \in \mathbb{R}^d : f(x) \geqslant \langle u, x\rangle \ \forall\, x \in \mathbb{R}^d\}\\ &= \{u \in \mathbb{R}^d : \|x\|_2 \geqslant \langle u, x\rangle \ \forall\, x \in \mathbb{R}^d\}.\end{aligned}$$

Έστω $u \in B_2(0,1) = \{v \in \mathbb{R}^d : \|v\|_2 \leqslant 1\}$· τότε

$$\langle u, x\rangle \leqslant \|u\|_2 \|x\|_2 = \|x\|_2 \qquad \forall\, x \in \mathbb{R}^d,$$

από την ανισότητα Cauchy–Schwarz, και άρα $u \in \partial f(0)$. Αντίστροφα, αν $u \in \partial f(0)$, τότε πρέπει, για $x = u$,

$$\|u\|_2 \geqslant \langle u, u\rangle = \|u\|_2^2,$$

και άρα πρέπει $\|u\|_2 \leqslant 1$. Αυτό δείχνει ότι $\partial f(0) = B_2(0,1)$ τελικά.
(β) Από τον ορισμό πάλι έχει κανείς ότι

$$\begin{aligned}\partial f(0) &= \{u \in \mathbb{R}^d : f(x) \geqslant \langle u, x\rangle \ \forall\, x \in \mathbb{R}^d\}\\ &= \{u \in \mathbb{R}^d : \|x\|_1 \geqslant \langle u, x\rangle \ \forall\, x \in \mathbb{R}^d\}.\end{aligned}$$

Έστω $u \in B_\infty(0,1) = \{v \in \mathbb{R}^d : \|v\|_\infty \leqslant 1\}$, όπου $\|u\|_\infty = \max_{i\in\{1,\dots,d\}} |u_i|$· τότε

$$\langle u, x\rangle = \sum_{i=1}^d u_i x_i \leqslant \left|\sum_{i=1}^d u_i x_i\right| \leqslant \sum_{i=1}^d |x_i||u_i| \leqslant \|u\|_\infty \|x\|_1 = \|x\|_1 \qquad \forall\, x \in \mathbb{R}^d$$

(αυτή βασικά είναι η ανισότητα Hölder για $p = 1$, $q = \infty$), και άρα $u \in \partial f(0)$. Αντίστροφα, αν $u \in \partial f(0)$, τότε πρέπει

$$\langle u, \mathrm{e}_i \rangle \leqslant \|\mathrm{e}_i\|_1 = 1 \qquad \forall\, i \in \{1, \ldots, d\}, \tag{7α′}$$

και

$$\langle u, -\mathrm{e}_i \rangle \leqslant \|-\mathrm{e}_i\|_1 = 1 \qquad \forall\, i \in \{1, \ldots, d\}, \tag{7β′}$$

όπου $\mathrm{e}_1, \ldots \mathrm{e}_d$ τα διανύσματα της συνήθους βάσης του $\mathbb{R}^d$. Όμως $\langle u, \mathrm{e}_i \rangle = u_i$, η i συντεταγμένη του u· άρα οι (7α′)–(7β′) δείχνουν ότι $|u_i| \leqslant 1$ για κάθε $i \in \{1, \ldots, d\}$ και άρα ότι $u \in B_\infty(0, 1)$. Επομένως $\partial f(0) = B_\infty(0, 1)$ τελικά. □

10. Έστω K μη κενό ανοικτό και κυρτό υποσύνολο του $\mathbb{R}^d$ και $f\colon K \to \mathbb{R}$ διαφορίσιμη συνάρτηση. Τότε η f είναι κυρτή ανν

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geqslant 0 \qquad \forall\, x, y \in K.$$

Λύση. Έστω πρώτα ότι η f είναι κυρτή και έστω $x, y \in K$. Αφού η f είναι διαφορίσιμη έχει ακριβώς ένα υπερεπίπεδο στήριξης σε κάθε σημείο του K, και αυτό ισχύει συγκεκριμένα για τα σημεία x και y του K· έχει συγκεκριμένα κανείς ότι

$$f(z) \geqslant \langle \nabla f(x), z - x \rangle + f(x) \qquad \forall\, z \in K$$

και

$$f(z) \geqslant \langle \nabla f(y), z - y \rangle + f(y) \qquad \forall\, z \in K.$$

Θεωρώντας αυτές τις ανισότητες για $z = y$ και $z = x$, αντίστοιχα, παίρνει κανείς ότι

$$f(y) \geqslant \langle \nabla f(x), y - x \rangle + f(x) \qquad \forall\, z \in K$$

και

$$f(x) \geqslant \langle \nabla f(y), x - y \rangle + f(y) \qquad \forall\, z \in K,$$

και προσθέτοντας κατά μέλη αυτές τις δύο ανισότητες προκύπτει η ζητούμενη $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geqslant 0$.

Αντίστροφα, έστω ότι ξέρει κανείς ότι,

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geqslant 0 \qquad \forall\, x, y \in K. \tag{8}$$

Έστω $x, y \in K$ και έστω πάλι η συνάρτηση μίας μεταβλητής

$$g(t) := f(x + t(y - x))$$

ορισμένη για $t \in \{s \in \mathbb{R} \colon x + s(y - x) \in K\} =: I$. Η συνάρτηση αυτή είναι κυρτή και έχει παράγωγο

$$g'(t) = \langle \nabla f(x + t(y - x)), y - x \rangle \cdot \tag{9}$$

αυτό είναι συνέπεια του κανόνα της αλυσίδας αλλά αποδείχθηκε και στο μάθημα. Υπενθυμίζεται η απόδειξη:

$$\begin{aligned}\frac{g(s+t)-g(t)}{s} &- \langle \nabla f(x+t(y-x)), y-x\rangle \\ &= \frac{f(x+(t+s)(y-x)) - f(x+t(y-x)) - \langle \nabla f(x+t(y-x)), s(y-x)\rangle}{s} \xrightarrow{(s\to 0)} 0,\end{aligned}$$

αφού

$$\frac{f(x+(t+s)(y-x)) - f(x+t(y-x)) - \langle \nabla f(x+t(y-x)), s(y-x)\rangle}{|s|\|y-x\|} \xrightarrow{(s\to 0)} 0,$$

από τον ορισμό της διαφορισιμότητας της f στο $x+t(y-x)$. Από τις (8) και (9) έπεται ότι αν $t_1 < t_2$,

$$\begin{aligned}g'(t_2) - g'(t_1) &= \langle \nabla f(x+t_2(y-x)) - \nabla f(x+t_1(y-x)), y-x\rangle \\ &= \frac{\langle \nabla f(x+t_2(y-x)) - \nabla f(x+t_1(y-x)), (t_2-t_1)(y-x)\rangle}{t_2-t_1} \geqslant 0,\end{aligned}$$

και άρα η g' είναι αύξουσα και επομένως η g κυρτή. Έπεται ότι, αν $t \in [0,1]$, τότε

$$g(t) = g((1-t)\cdot 0 + t\cdot 1) \leqslant (1-t)g(0) + tg(1),$$

που σημαίνει ότι

$$f((1-t)x+ty) = g(t) \leqslant (1-t)g(0) + tg(1) = (1-t)f(x) + tf(y).$$

Αφού τα $x, y \in K$ και $t \in [0,1]$ ήταν αυθαίρετα, έπεται ότι η f είναι κυρτή. □