Κεφάλαιο 4

Στοιχεία θεωρίας κυρτών συναρτήσεων

4.1 Κυρτές συναρτήσεις

Ορισμός 4.1.1. Μία συνάρτηση $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ καλείται **κυρτή** (convex) αν το σύνολο $\mathbf{dom} f$ είναι κυρτό και αν, για κάθε $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$ και $0 \leq \theta \leq 1$, ισχύει ότι

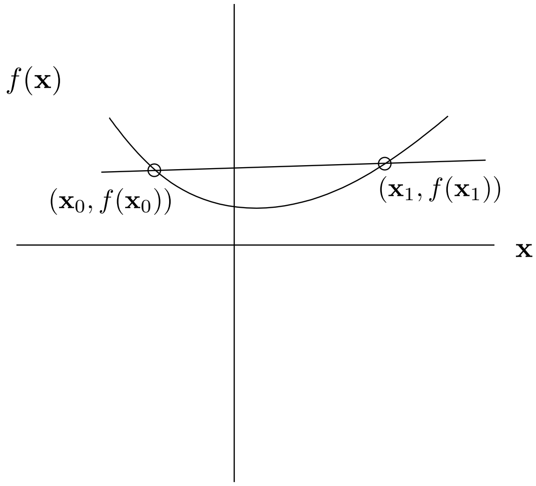
$$f(\theta \mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}). \tag{4.1}$$

Γεωμετρικά, η ανισότητα (4.1) σημαίνει ότι κανένα σημείο του ευθύγραμμου τμήματος το οποίο συνδέει τα σημεία $(\mathbf{x}, f(\mathbf{x}))$ και $(\mathbf{y}, f(\mathbf{y}))$ δεν ευρίσκεται κάτω από το γράφημα της f (δείτε το Σχήμα 4.1).

Ορισμός 4.1.2. Μία συνάρτηση $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ καλείται **αυστηρά κυρτή** (strictly convex) αν το σύνολο $\mathbf{dom} f$ είναι κυρτό και αν, για κάθε $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$, με $\mathbf{x} \neq \mathbf{y}$ και $0 < \theta < 1$, ισχύει ότι

$$f(\theta \mathbf{x} + (1-\theta)\mathbf{y}) < \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}). \tag{4.2}$$

Γεωμετρικά, η ανισότητα (4.2) σημαίνει ότι, με εξαίρεση των σημείων $(\mathbf{x}, f(\mathbf{x}))$ και $(\mathbf{y}, f(\mathbf{y}))$, το ευθύγραμμο τμήμα το οποίο συνδέει τα σημεία $(\mathbf{x}, f(\mathbf{x}))$ και $(\mathbf{y}, f(\mathbf{y}))$ ευρίσκεται πάνω από το γράφημα της f (δείτε το Σχήμα 4.1).



Σχήμα 4.1: Κυρτή συνάρτηση.

Η συνάρτηση f καλείται **κοίλη** (concave) αν η $-f$ είναι κυρτή.

Μία αφφινική συνάρτηση είναι, ταυτόχρονα, κυρτή και κοίλη.

Η σχέση (4.1) γενικεύεται ως εξής. Έστω $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ κυρτή συνάρτηση, $\mathbf{x}_i \in \mathbf{dom} f$, για $i = 1, \dots, n$, και $a_i \geq 0$, για $i = 1, \dots, n$, με $a_1 + \cdots + a_n = 1$. Τότε

$$f\left(\sum_{i=1}^{n} a_i \mathbf{x}_i\right) \leq \sum_{i=1}^{n} a_i f(\mathbf{x}_i). \tag{4.3}$$

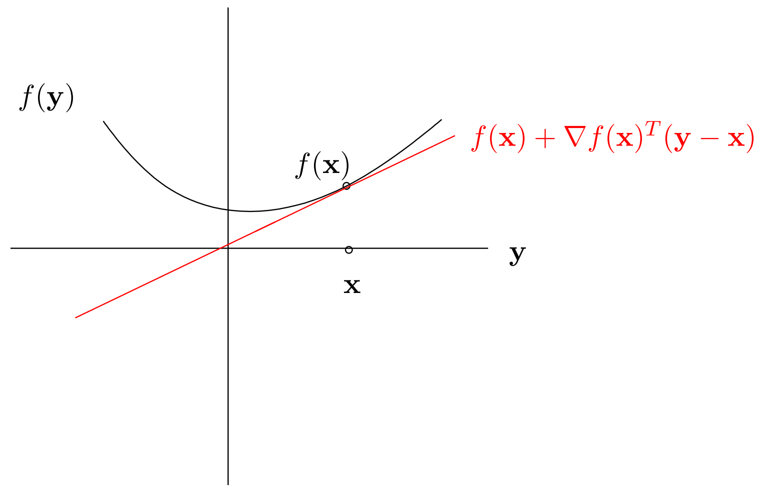
Μπορεί να αποδειχθεί ότι μία συνάρτηση f είναι κυρτή αν, και μόνο αν, ο περιορισμός της πάνω στην τομή οποιασδήποτε ευθείας γραμμής στον $\mathbb{R}^n$ και του πεδίου ορισμού της f είναι κυρτή συνάρτηση.

Θεώρημα 4.1.1. Η συνάρτηση $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ είναι κυρτή αν, και μόνο αν, για κάθε $\mathbf{x} \in \mathbf{dom} f$ και $\mathbf{v} \in \mathbb{R}^n$, η συνάρτηση $g(t) = f(\mathbf{x} + t\mathbf{v})$, με πεδίο ορισμού το σύνολο $\{t \in \mathbb{R} \,|\, \mathbf{x} + t\mathbf{v} \in \mathbf{dom} f\}$, είναι κυρτή.

Απόδειξη. Δείτε το Παράρτημα του Κεφαλαίου. □

Οι κυρτές συναρτήσεις έχουν κάποιες πολύ σημαντικές ιδιότητες.

Για παράδειγμα, είναι συνεχείς στο εσωτερικό του πεδίου ορισμού τους.



Σχήμα 4.2: Συνθήκη κυρτότητας πρώτης τάξης.

Ασυνέχειες μπορεί να υπάρχουν μόνο πάνω στο σύνορο του πεδίου ορισμού τους.

Η θεωρία των κυρτών συναρτήσεων είναι εξαιρετικά εκτενής. Στα πλαίσια του μαθήματος, θα καλύψουμε μόνο θέματα τα οποία είναι απολύτως απαραίτητα για τη συνέχεια.

4.2 Συνθήκες πρώτης τάξης

Θεώρημα 4.2.1. *Έστω ανοιχτό σύνολο* $\mathbf{dom} f \subseteq \mathbb{R}^n$ *και* $f : \mathbf{dom} f \to \mathbb{R}$ *διαφορίσιμη συνάρτηση. Τότε, η* f *είναι κυρτή αν, και μόνο αν, το* $\mathbf{dom} f$ *είναι κυρτό σύνολο και*

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}), \tag{4.4}$$

για κάθε $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$.

Η ανισότητα (4.4) είναι εξαιρετικά σημαντική. Δηλώνει ότι η προσέγγιση Taylor πρώτης τάξης σε οποιοδήποτε σημείο μίας κυρτής συνάρτησης είναι ολικός υποεκτιμητής της συνάρτησης.

Δηλαδή, από τοπική πληροφορία (τις τιμές της συνάρτησης και της παραγώγου της σε ένα σημείο), λαμβάνουμε ολική πληροφορία (έναν ολικό υποεκτιμητή της συνάρτησης).

Απόδειξη. (*Ευθύ*) Έστω ότι η f είναι κυρτή συνάρτηση και $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$.

Τότε, για $0 < \theta \leq 1$ έχουμε ότι $(1-\theta)\mathbf{x} + \theta\mathbf{y} = \mathbf{x} + \theta(\mathbf{y} - \mathbf{x}) \in \mathbf{dom} f$ και

$$f(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})) \leq (1-\theta) f(\mathbf{x}) + \theta f(\mathbf{y}).$$

Διαιρώντας και τα δύο μέλη της παραπάνω ανισότητας με θ, και μετά από απλές αλγεβρικές πράξεις, λαμβάνουμε

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \frac{f(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})) - f(\mathbf{x})}{\theta}.$$

Παίρνοντας όριο για $\theta \to 0$, λαμβάνουμε (γιατί;)

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}),$$

η οποία είναι η προς απόδειξη σχέση.

(*Αντίστροφο*) Ας υποθέσουμε ότι ισχύει η (4.4) για κάθε $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$.

Έστω $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{dom} f$, με $\mathbf{x}_1 \neq \mathbf{x}_2$, και $0 \leq \theta \leq 1$. Θέτουμε $\mathbf{x} = \theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2$.

Χρησιμοποιώντας την (4.4) δύο φορές, λαμβάνουμε

$$f(\mathbf{x}_1) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{x}_1 - \mathbf{x}), \tag{4.5}$$

$$f(\mathbf{x}_2) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{x}_2 - \mathbf{x}). \tag{4.6}$$

Πολλαπλασιάζοντας την (4.5) με θ και την (4.6) με $(1-\theta)$, και προσθέτοντας τις ανισότητες κατά μέλη, λαμβάνουμε

$$\begin{aligned} \theta f(\mathbf{x}_1) + (1-\theta) f(\mathbf{x}_2) &\geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T \left(\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2 - \mathbf{x}\right) \\ &= f(\mathbf{x}), \end{aligned}$$

αποδεικνύοντας ότι η $f(x)$ είναι κυρτή. $\square$

4.3 Συνθήκες δεύτερης τάξης

Θεώρημα 4.3.1. Έστω ανοιχτό σύνολο $\mathbf{dom} f \subseteq \mathbb{R}^n$ και $f : \mathbf{dom} f \to \mathbb{R}$ συνεχώς διπλά διαφορίσιμη συνάρτηση. Τότε, η f είναι κυρτή αν, και μόνο αν, το $\mathbf{dom} f$ είναι κυρτό σύνολο και

$$\nabla^2 f(\mathbf{x}) \succeq \mathbf{O} \tag{4.7}$$

για κάθε $\mathbf{x} \in \mathbf{dom} f$.

Απόδειξη. *(Ευθύ)* Έστω ότι η f είναι κυρτή συνάρτηση, $\mathbf{x} \in \mathbf{dom} f$, και $\mathbf{h} \in \mathbb{R}^n$.

Αφού το $\mathbf{dom} f$ είναι ανοιχτό σύνολο, έχουμε ότι $\mathbf{x}+\lambda\mathbf{h} \in \mathbf{dom} f$, για $|\lambda|$ αρκετά μικρό. Από την (4.4), λαμβάνουμε

$$f(\mathbf{x}+\lambda\mathbf{h}) \geq f(\mathbf{x}) + \lambda \nabla f(\mathbf{x})^T \mathbf{h}.$$

Από το ανάπτυγμα Taylor δεύτερης τάξης, έχουμε

$$f(\mathbf{x}+\lambda\mathbf{h}) = f(\mathbf{x}) + \lambda \nabla f(\mathbf{x})^T \mathbf{h} + \frac{\lambda^2}{2}\, \mathbf{h}^T \nabla^2 f(\mathbf{x}) \mathbf{h} + O(\lambda^3).$$

Συνδυάζοντας τις δύο παραπάνω σχέσεις, λαμβάνουμε ότι

$$\frac{\lambda^2}{2}\, \mathbf{h}^T \nabla^2 f(\mathbf{x}) \mathbf{h} + O(\lambda^3) \geq 0.$$

Διαιρώντας και τα δύο μέλη με λ^2 και παίρνοντας όριο για $\lambda \to 0$, λαμβάνουμε ότι

$$\mathbf{h}^T \nabla^2 f(\mathbf{x}) \mathbf{h} \geq 0$$

για κάθε $\mathbf{h} \in \mathbb{R}^n$, το οποίο δίνει ότι, για κάθε $\mathbf{x} \in \mathbf{dom} f$,

$$\nabla^2 f(\mathbf{x}) \succeq \mathbf{O}.$$

(Αντίστροφο) Θα πρέπει να αποδείξουμε ότι, αν $\nabla^2 f(\mathbf{x}) \succeq \mathbf{O}$, για κάθε $\mathbf{x} \in \mathbf{dom} f$, τότε η f είναι κυρτή.

Από το ανάπτυγμα Taylor δεύτερης τάξης, λαμβάνουμε

$$f(\mathbf{y}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y}-\mathbf{x}) + \frac{1}{2}\, (\mathbf{y}-\mathbf{x})^T \nabla^2 f(\mathbf{z}) (\mathbf{y}-\mathbf{x}),$$

με $\mathbf{z} = \lambda\mathbf{x}+(1-\lambda)\mathbf{y}$, για κάποιο $\lambda \in [0,1]$ (η απόδειξη γίνεται με χρήση του θεωρήματος Μέσης Τιμής).

Αφού έχουμε υποθέσει ότι $\nabla^2 f(\mathbf{x}) \succeq \mathbf{O}$ για κάθε $\mathbf{x} \in \mathbf{dom} f$, τότε έχουμε ότι $\nabla^2 f(\mathbf{z}) \succeq \mathbf{O}$ και

$$(\mathbf{y}-\mathbf{x})^T \nabla^2 f(\mathbf{z}) (\mathbf{y}-\mathbf{x}) \geq 0.$$

Συνεπώς, για κάθε $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$,

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y}-\mathbf{x}),$$

και, βάσει του Θεωρήματος 4.4, συμπεραίνουμε ότι η f είναι κυρτή. □

4.4 Παραδείγματα κυρτών συναρτήσεων

Παράδειγμα 4.4.1. Ακολουθούν παραδείγματα βαθμωτών κυρτών και κοίλων συναρτήσεων.

1. Η $f(x) = e^{ax}$ είναι κυρτή στο $\mathbb{R}$, για κάθε $a \in \mathbb{R}$.
2. Η $f(x) = x^a$ είναι κυρτή στο $\mathbb{R}_{++}$ αν $a \geq 1$ ή $a \leq 0$, και κοίλη αν $0 \leq a \leq 1$.
3. Η $f(x) = |x|^p$ είναι κυρτή στο $\mathbb{R}$, για $p \geq 1$.
4. Η $f(x) = \log(x)$ είναι κοίλη στο $\mathbb{R}_{++}$.

Παράδειγμα 4.4.2. Τετραγωνικές συναρτήσεις. Έστω $f : \mathbb{R}^n \to \mathbb{R}$, με

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T\mathbf{x} + r$$

με $\mathbf{P} \in \mathbb{R}^{n\times n}$ και $\mathbf{P} = \mathbf{P}^T$, $\mathbf{q} \in \mathbb{R}^n$ και $r \in \mathbb{R}$.

Μπορεί να αποδειχθεί ότι (να το αποδείξεις...)

$$\nabla f(\mathbf{x}) = \mathbf{P}\mathbf{x} + \mathbf{q}$$

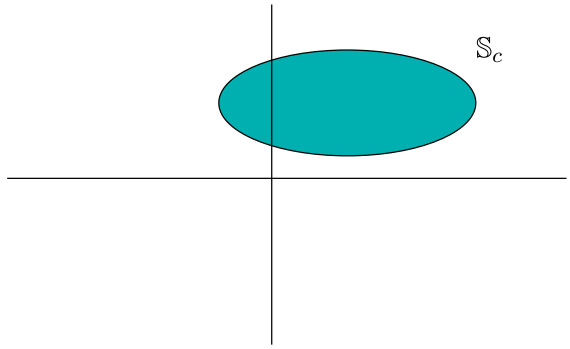
και

$$\nabla^2 f(\mathbf{x}) = \mathbf{P}.$$

Από τις συνθήκες δεύτερης τάξης, συμπεραίνουμε ότι η f είναι κυρτή (αυστηρά κυρτή) αν, και μόνο αν, $\mathbf{P} \succeq \mathbf{O}$ ($\mathbf{P} \succ \mathbf{O}$).

Παράδειγμα 4.4.3. Ακολουθούν παραδείγματα διανυσματικών κυρτών και κοίλων συναρτήσεων.

1. Η $f(\mathbf{x}) = \|\mathbf{x}\|$ είναι κυρτή στον $\mathbb{R}^n$, όπου $\|\cdot\|$ είναι οποιαδήποτε συνάρτηση μέτρου στον $\mathbb{R}^n$.
2. Η $f(\mathbf{x}) = \max\{x_1, \dots, x_n\}$ είναι κυρτή στον $\mathbb{R}^n$.
3. Η συνάρτηση $f(\mathbf{x}) = \log\left(e^{x_1} + \cdots + e^{x_n}\right)$ είναι κυρτή στον $\mathbb{R}^n$.
4. Η συνάρτηση $f(\mathbf{x}) = \left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}$ είναι κοίλη στον $\mathbb{R}^n_{++}$.



Σχήμα 4.3: Σύνολο υποστάθμης κυρτής συνάρτησης.

4.5 Σύνολα υποστάθμης

Ορισμός 4.5.1. Έστω $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$. Το σύνολο

$$\mathbb{S}_c = \{\mathbf{x} \in \mathbf{dom} f \mid f(\mathbf{x}) \leq c\} \tag{4.8}$$

καλείται **σύνολο c-υποστάθμης** (c-sublevel set)της f.

Θεώρημα 4.5.1. Έστω $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ κυρτή συνάρτηση. Τότε, το σύνολο c-υποστάθμης $\mathbb{S}_c$ είναι κυρτό, για κάθε $c \in \mathbb{R}$.

Απόδειξη. *[Θα πρέπει να αποδείξουμε ότι αν $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{S}_c$ και $0 \leq \theta \leq 1$, τότε $\mathbf{x} = \theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathbb{S}_c$].*

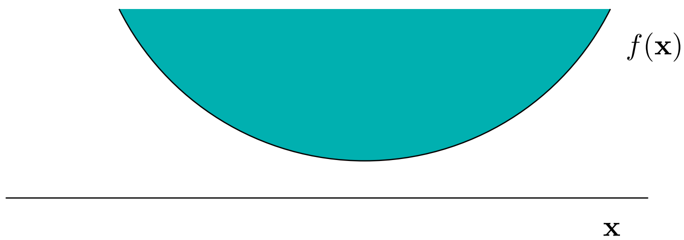
Έστω, ότι $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{S}_c$ και $0 \leq \theta \leq 1$.

Τότε, $f(\mathbf{x}_1) \leq c$ και $f(\mathbf{x}_2) \leq c$, και $\mathbf{x} = \theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathbf{dom} f$. Επιπλέον,

$$\begin{aligned} f(\mathbf{x}) = f(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2) &\leq \theta f(\mathbf{x}_1) + (1-\theta) f(\mathbf{x}_2) \\ &\leq \theta\, c + (1-\theta)\, c \\ &= c. \end{aligned}$$

Συνεπώς, $\mathbf{x} = \theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathbb{S}_c$. □

Πρέπει να επισημάνουμε ότι η αντίστροφη πρόταση δεν ισχύει. Για παράδειγμα, τα σύνολα υποστάθμης της συνάρτησης $f : \mathbb{R}_+ \to \mathbb{R}$, με $f(x) = \log(x)$, είναι κυρτά. Όμως, η συνάρτηση f είναι κοίλη.



Σχήμα 4.4: Επιγράφημα κυρτής συνάρτησης.

4.6 Επιγράφημα

Ορισμός 4.6.1. Έστω $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$. Το σύνολο

$$\mathbf{epi} f = \{(\mathbf{x}, t) \,|\, \mathbf{x} \in \mathbf{dom} f, f(\mathbf{x}) \leq t\}$$

καλείται **επιγράφημα** (epigraph)της f.

Θεώρημα 4.6.1. Έστω συνάρτηση $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$. Η f είναι κυρτή αν, και μόνο αν, το σύνολο $\mathbf{epi} f$ είναι κυρτό.

Απόδειξη. (Ευθύ). [Αν η f είναι κυρτή συνάρτηση, τότε το σύνολο $\mathbf{epi} f$ είναι κυρτό.]

Υποθέτουμε ότι η f είναι κυρτή.

Για να αποδείξουμε ότι το $\mathbf{epi} f$ είναι κυρτό σύνολο θα πρέπει να αποδείξουμε ότι αν

$$(\mathbf{x}_1, t_1), (\mathbf{x}_2, t_2) \in \mathbf{epi} f, \quad 0 \leq \theta \leq 1,$$

τότε

$$\theta(\mathbf{x}_1, t_1) + (1 - \theta)(\mathbf{x}_2, t_2) \in \mathbf{epi} f.$$

Έστω, λοιπόν, ότι $(\mathbf{x}_1, t_1), (\mathbf{x}_2, t_2) \in \mathbf{epi} f$, δηλαδή,

$$f(\mathbf{x}_1) \leq t_1, \quad f(\mathbf{x}_2) \leq t_2. \tag{4.9}$$

Εξαιτίας του ότι η f είναι κυρτή, έχουμε, για $0 \leq \theta \leq 1$,

$$\begin{aligned} f(\theta\mathbf{x}_1 + (1 - \theta)\mathbf{x}_2) &\leq \theta f(\mathbf{x}_1) + (1 - \theta) f(\mathbf{x}_2) \\ &\overset{(4.9)}{\leq} \theta t_1 + (1 - \theta) t_2, \end{aligned}$$

δηλαδή, $\theta(\mathbf{x}_1, t_1) + (1-\theta)(\mathbf{x}_2, t_2) \in \mathbf{epi}f$.

(Αντίστροφο). [Αν το σύνολο $\mathbf{epi}f$ *είναι κυρτό, τότε η* f *είναι κυρτή.]*

Υποθέτουμε ότι το σύνολο $\mathbf{epi}f$ είναι κυρτό.

Αρχικά, γνωρίζουμε ότι η προβολή του $\mathbf{epi}f$ πάνω στις πρώτες n συντεταγμένες του είναι κυρτό σύνολο.

Συνεπώς, το σύνολο $\mathbf{dom}f$ είναι κυρτό.

Επιπλέον, αν $(\mathbf{x}_1, t_1), (\mathbf{x}_2, t_2) \in \mathbf{epi}f$ και $0 \leq \theta \leq 1$, τότε $\theta(\mathbf{x}_1, t_1) + (1-\theta)(\mathbf{x}_2, t_2) \in \mathbf{epi}f$.

Παρατηρούμε ότι μπορούμε να θέσουμε $t_1 = f(\mathbf{x}_1)$ και $t_2 = f(\mathbf{x}_2)$ (γιατί;).

Τότε, εξαιτίας της κυρτότητας του $\mathbf{epi}f$, έχουμε ότι $\theta(\mathbf{x}_1, f(\mathbf{x}_1)) + (1-\theta)(\mathbf{x}_2, f(\mathbf{x}_2)) \in \mathbf{epi}f$, δηλαδή,

$$f(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2) \leq \theta f(\mathbf{x}_1) + (1-\theta) f(\mathbf{x}_2),$$

δηλαδή, η f είναι κυρτή. □

4.7 Πράξεις συναρτήσεων οι οποίες διατηρούν την κυρτότητα

4.8 Μη-αρνητικά σταθμισμένα αθροίσματα

Θεώρημα 4.8.1. Έστω $f_k : \mathbb{R}^n \to \mathbb{R}$, για $k = 1, \ldots, K$, κυρτές συναρτήσεις και $w_k \geq 0$, για $k = 1, \ldots, K$. Τότε η συνάρτηση

$$f(\mathbf{x}) = \sum_{k=1}^{K} w_k f_k(\mathbf{x})$$

είναι κυρτή.

Απόδειξη. Από την κυρτότητα των f_k, για $k = 1, \ldots, K$, έχουμε ότι, αν $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ και $0 \leq \theta \leq 1$, τότε

$$f_k(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2) \leq \theta f_k(\mathbf{x}_1) + (1-\theta) f_k(\mathbf{x}_2), \text{ για } k = 1, \ldots, K.$$

Πολλαπλασιάζοντας τα δύο μέλη κάθε ανισότητας με w_k, για $k = 1, \ldots, K$, και αθροίζοντας, λαμβάνουμε

$$\begin{aligned}\sum_{k=1}^{K} w_k f_k(\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2) &\le \sum_{k=1}^{K} w_k \left(\theta f_k(\mathbf{x}_1) + (1-\theta) f_k(\mathbf{x}_2)\right) \\ &= \theta \sum_{k=1}^{K} w_k f_k(\mathbf{x}_1) + (1-\theta) \sum_{k=1}^{K} w_k f_k(\mathbf{x}_2)\end{aligned}$$

το οποίο δίνει ότι

$$f(\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2) \le \theta f(\mathbf{x}_1) + (1-\theta) f(\mathbf{x}_2)$$

□

4.9 Σύνθεση με αφφινική συνάρτηση

Θεώρημα 4.9.1. Έστω $f : \mathbb{R}^n \to \mathbb{R}$, $\mathbf{A} \in \mathbb{R}^{n \times m}$, και $\mathbf{b} \in \mathbb{R}^n$. Έστω $g : \mathbb{R}^m \to \mathbb{R}$, με $g(\mathbf{x}) = f(\mathbf{A}\mathbf{x} + \mathbf{b})$ και πεδίο ορισμού $\mathbf{dom}\, g = \{\mathbf{x} \in \mathbb{R}^m \,|\, \mathbf{A}\mathbf{x} + \mathbf{b} \in \mathbf{dom}\, f\}$. Αν η f είναι κυρτή συνάρτηση, τότε και η g είναι κυρτή. Αν η f είναι κοίλη συνάρτηση, τότε και η g είναι κοίλη.

Απόδειξη. *Αφήνεται στον αναγνώστη.* □

4.10 Σημειακά μέγιστα και ελάχιστα

Θεώρημα 4.10.1. Έστω $f_k : \mathbb{R}^n \to \mathbb{R}$, για $k = 1, \ldots, K$, κυρτές συναρτήσεις. Τότε η συνάρτηση

$$f(\mathbf{x}) = \max\{f_1(\mathbf{x}), \ldots, f_K(\mathbf{x})\}$$

είναι κυρτή.

Απόδειξη. Από την κυρτότητα των f_k, για $k = 1, \ldots, K$, έχουμε ότι, αν $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ και $0 \le \theta \le 1$, τότε

$$f_k(\theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2) \le \theta f_k(\mathbf{x}_1) + (1-\theta) f_k(\mathbf{x}_2), \text{ για } k = 1, \ldots, K.$$

Συνεπώς,

$$\begin{aligned}
\max_{k=1,\ldots,K}\{f_k(\theta\mathbf{x}_1+(1-\theta)\mathbf{x}_2)\} &\le \max_{k=1,\ldots,K}\{\theta f_k(\mathbf{x}_1)+(1-\theta)f_k(\mathbf{x}_2)\}\\
&\overset{(!)}{\le} \max_{k=1,\ldots,K}\{\theta f_k(\mathbf{x}_1)\}+\max_{k=1,\ldots,K}\{(1-\theta)f_k(\mathbf{x}_2)\}\\
&= \theta \max_{k=1,\ldots,K}\{f_k(\mathbf{x}_1)\}+(1-\theta)\max_{k=1,\ldots,K}\{f_k(\mathbf{x}_2)\}\\
&= \theta f(\mathbf{x}_1)+(1-\theta)f(\mathbf{x}_2).
\end{aligned}$$

Η μόνη μη-τετριμμένη ανισότητα είναι η (!), η οποία, ουσιαστικά, είναι ισοδύναμη με την ανισότητα

$$\max_{k=1,\ldots,K}\{a_k+b_k\}\le \max_{k=1,\ldots,K}\{a_k\}+\max_{k=1,\ldots,K}\{b_k\},$$

για οποιαδήποτε βαθμωτά a_k, b_k, για $k=1,\ldots,K$ (προσπαθήστε να αποδείξετε αυτή την ανισότητα). $\square$

4.11 Σύνθεση συναρτήσεων

Έστω $f,h:\mathbb{R}\to\mathbb{R}$ και $g(t)=(f\circ h)(t)$. Η g είναι κυρτή αν, και μόνο αν, $D^2g(t)\ge 0$ για κάθε $t\in\mathbf{dom}g$.

Γνωρίζουμε ότι $Dg(t)=Df(h(t))\,Dh(t)$ και

$$D^2g(t)=D^2f(h(t))\,(Dh(t))^2+Df(h(t))D^2h(t).$$

Με κατάλληλη επιλογή των f και h, μπορούμε να έχουμε $D^2g(t)\ge 0$ για κάθε $t\in\mathbf{dom}g$.

Για παράδειγμα, αν η f είναι κυρτή και μή-φθίνουσα, το οποίο σημαίνει ότι $Df(t)\ge 0$ και $D^2f(t)\ge 0$, και η h είναι κυρτή, το οποίο σημαίνει ότι $D^2h(t)\ge 0$, τότε η g είναι κυρτή (γιατί;).

Για περισσότερα παραδείγματα, δείτε στις σελίδες 83–87 του Boyd–Vandenberghe.

4.12 Ελαχιστοποίηση

Θεώρημα 4.12.1. Έστω $f:\mathbb{R}^{n+m}\to\mathbb{R}$, $\mathbf{x}\in\mathbb{R}^n$ και $\mathbf{y}\in\mathbb{R}^m$. Έστω ότι η $f(\mathbf{x},\mathbf{y})$ είναι κυρτή ως προς $(\mathbf{x},\mathbf{y})$ και έστω ότι το $\mathbb{C}\subseteq\mathbb{R}^m$ είναι κυρτό σύνολο. Τότε η

$g(\mathbf{x}) = \min_{\mathbf{y}\in\mathbb{C}} f(\mathbf{x},\mathbf{y})$ είναι κυρτή, ως προς $\mathbf{x}$, αν $g(\mathbf{x}) > -\infty$ για κάθε $\mathbf{x}$ στο

$$\mathbf{dom}\, g = \{\mathbf{x} \,|\, (\mathbf{x},\mathbf{y}) \in \mathbf{dom} f \text{ για κάποιο } \mathbf{y} \in \mathbb{C}\}$$

και $g(\mathbf{x}) < \infty$, για κάποιο $\mathbf{x} \in \mathbf{dom} g$.

Απόδειξη. Θα πρέπει να αποδείξουμε ότι αν $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{dom} g$ και $0 \le \theta \le 1$, τότε $\mathbf{x} = \theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathbf{dom} g$ και $g(\mathbf{x}) \le \theta g(\mathbf{x}_1) + (1-\theta) g(\mathbf{x}_2$.

Για κάθε $\epsilon > 0$, μπορούμε να βρούμε $\mathbf{y}_1$ και $\mathbf{y}_2$ τέτοια ώστε

$$(\mathbf{x}_1,\mathbf{y}_1), (\mathbf{x}_2,\mathbf{y}_2) \in \mathbf{dom} f$$

$$g(\mathbf{x}_1) + \epsilon \ge f(\mathbf{x}_1,\mathbf{y}_1), \quad g(\mathbf{x}_2) + \epsilon \ge f(\mathbf{x}_2,\mathbf{y}_2).$$

Πολλαπλασιάζοντας την πρώτη ανισότητα με θ και τη δεύτερη με $(1-\theta)$ και προσθέτοντας, λαμβάνουμε

$$\begin{aligned}\theta g(\mathbf{x}_1) + (1-\theta) g(\mathbf{x}_2) + \epsilon &\ge \theta f(\mathbf{x}_1,\mathbf{y}_1) + (1-\theta) f(\mathbf{x}_2,\mathbf{y}_2)\\ &\ge f(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2, \theta\mathbf{y}_1 + (1-\theta)\mathbf{y}_2)\\ &= f(\mathbf{x}, \theta\mathbf{y}_1 + (1-\theta)\mathbf{y}_2)\\ &\ge g(\mathbf{x}).\end{aligned} \tag{4.10}$$

Συνεπώς, $\mathbf{x} \in \mathbf{dom} g$, και επειδή η ανισότητα ισχύει για κάθε $\epsilon > 0$, έχουμε ότι $g(\mathbf{x}) \le \theta g(\mathbf{x}_1) + (1-\theta) g(\mathbf{x}_2)$. Συνεπώς, η g είναι κυρτή συνάρτηση (δες, ΟΠΤΙΙΙ, σελ. 68). □

Παράρτημα

Σκιαγράφηση απόδειξης Θεωρήματος 4.1.1, υποθέτοντας $f : \mathbb{R}^n \to \mathbb{R}$. (Να δουλέψετε τις λεπτομέρειες για $f : \mathbf{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$, με $\mathbf{dom} f$ κυρτό σύνολο).

Έστω $f : \mathbb{R}^n \to \mathbb{R}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, και $g_{\mathbf{x},\mathbf{y}} : \mathbb{R} \to \mathbb{R}$, με $g_{\mathbf{x},\mathbf{y}}(t) = f(t\mathbf{x} + (1-t)\mathbf{y})$. Στη συνέχεια, για απλούστευση του συμβολισμού, η $g_{\mathbf{x},\mathbf{y}}$ θα συμβολίζεται ως g.

Ευθύ: *Αν η f είναι κυρτή συνάρτηση, τότε η g είναι κυρτή συνάρτηση.*

Αν $f : \mathbb{R}^n \to \mathbb{R}$ κυρτή συνάρτηση και $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, τότε η $g_{\mathbf{x},\mathbf{y}} : \mathbb{R} \to \mathbb{R}$, με $g_{\mathbf{x},\mathbf{y}}(t) = f(t\mathbf{x} + (1-t)\mathbf{y})$, είναι κυρτή συνάρτηση.

Για να αποδείξουμε ότι η g είναι κυρτή συνάρτηση θα πρέπει να αποδείξουμε ότι για κάθε $t_1, t_2 \in \mathbb{R}$ και $0 \leq \lambda \leq 1$, θα πρέπει να ισχύει ότι

$$g(\lambda t_1 + (1-\lambda)t_2) \leq \lambda g(t_1) + (1-\lambda)g(t_2). \tag{4.11}$$

Ισοδύναμα, θα πρέπει να αποδείξουμε ότι

$$\begin{aligned} f\left((\lambda t_1 + (1-\lambda)t_2)\mathbf{x} + (1-(\lambda t_1 + (1-\lambda)t_2))\,\mathbf{y}\right) \leq \\ \lambda f\left(t_1\mathbf{x} + (1-t_1)\mathbf{y}\right) + (1-\lambda) f\left(t_2\mathbf{x} + (1-t_2)\mathbf{y}\right). \end{aligned} \tag{4.12}$$

Μετά από απλές αλγεβρικές πράξεις, λαμβάνουμε

$$\begin{aligned} (\lambda t_1 + (1-\lambda)t_2)\mathbf{x} + (1-(\lambda t_1 + (1-\lambda)t_2))\,\mathbf{y} = \\ \lambda\left(t_1\mathbf{x} + (1-t_1)\mathbf{y}\right) + (1-\lambda)\left(t_2\mathbf{x} + (1-t_2)\mathbf{y}\right). \end{aligned} \tag{4.13}$$

Συνεπώς, η (4.12), ισοδύναμα γράφεται ως

$$\begin{aligned} f\left(\lambda\left(t_1\mathbf{x} + (1-t_1)\mathbf{y}\right) + (1-\lambda)\left(t_2\mathbf{x} + (1-t_2)\mathbf{y}\right)\right) \leq \\ \lambda f\left(t_1\mathbf{x} + (1-t_1)\mathbf{y}\right) + (1-\lambda) f\left(t_2\mathbf{x} + (1-t_2)\mathbf{y}\right), \end{aligned} \tag{4.14}$$

η οποία είναι αληθής διότι έχουμε υποθέσει ότι η f είναι κυρτή συνάρτηση.

Αντίστροφο: *Αν η g είναι κυρτή συνάρτηση, τότε η f είναι κυρτή συνάρτηση.*

Θα πρέπει να αποδείξουμε ότι, αν $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ και $0 \leq \lambda \leq 1$, τότε

$$f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}). \tag{4.15}$$

Από τον ορισμό της g, έχουμε

$$g(\lambda) = f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}),\ g(1) = f(\mathbf{x}),\ g(0) = f(\mathbf{y}). \tag{4.16}$$

Από την κυρτότητα της g, λαμβάνουμε

$$g(\lambda) = g(\lambda \cdot 1 + (1-\lambda)\cdot 0) \leq \lambda g(1) + (1-\lambda) g(0), \tag{4.17}$$

η οποία είναι ισοδύναμη με την (4.15), ολοκληρώνοντας την απόδειξη.