Το θεώρημα του Alexandrov

Γιώργος Γιανναράκης και Δαυιδούλα Δημοπούλου

Περίληψη

Το 1939, ο Alexandr Alexandrov απέδειξε το ακόλουθο θεώρημα: Έστω $C \subseteq \mathbb{R}^d$ ανοιχτό και κυρτό, $f : C \to \mathbb{R}$ μια κυρτή συνάρτηση. Τότε, η f είναι δύο φορές παραγωγίσιμη σχεδόν παντού στο C. Στην παρούσα εργασία παρουσιάζουμε μία απόδειξη του θεωρήματος.

1 Το θεώρημα του Alexandrov

Ένα από τα βασικά αποτελέσματα της θεωρίας της κυρτής ανάλυσης είναι το θεώρημα του Alexandrov σχετικά με την διαφορισιμότητα δευτέρου βαθμού των κυρτών συναρτήσεων.

Θεώρημα 1.1. *Έστω $C \subseteq \mathbb{R}^d$ ανοιχτό και κυρτό, $f : C \to \mathbb{R}$ μια κυρτή συνάρτηση. Τότε, η f είναι δύο φορές παραγωγίσιμη σχεδόν παντού στο C.*

Περιγράφουμε συνοπτικά τα εργαλεία που χρησιμοποιούνται για την απόδειξη του θεωρήματοσ:

(α) Το θεώρημα του Rademacher: οι Lipschitz συναρτήσεις είναι διαφορίσιμες σχεδόν παντού. Βασικό ρόλο στην απόδειξη αυτού του αποτελέσματος παίζει το θεώρημα του Lebesgue για την διαφορισιμότητα των απολύτως συνεχών συναρτήσεων $g : [a,b] \to \mathbb{R}$. Οι κυρτές συναρτήσεις $f : C \subseteq \mathbb{R}^d \to \mathbb{R}$ είναι τοπικά Lipschitz, άρα σχεδόν παντού διαφορίσιμες (θεώρημα Reidemeister). Περιγράφουμε αυτά τα αποτελέσματα στην Παράγραφο 2.

(β) Το γεγονός ότι η εικόνα του συνόλου των κρίσιμων σημείων μιας Lipschitz συνάρτησης έχει μηδενικό μέτρο. Για κάθε $K : \mathbb{R}^d \to \mathbb{R}^d$, το σύνολο X το σύνολο των κρίσιμων σημείων της K είναι το σύνολο των $x \in \mathbb{R}^d$ στα οποία η K δεν είναι διαφορίσιμη ή είναι διαφορίσιμη αλλά ο πίνακας Jacobi A_x της K στο x δεν είναι αντιστρέψιμος. Αν η $K : \mathbb{R}^d \to \mathbb{R}^d$ είναι Lipschitz συνάρτηση, τότε το σύνολο των x στα οποία η K δεν είναι διαφορίσιμη έχει μηδενικό μέτρο (από το θεώρημα του Rademacher) και το ίδιο ισχύει για την εικόνα του. Από την άλλη πλευρά, αν $N \subset \mathbb{R}^d$ είναι το σύνολο των x στα οποία η K είναι διαφορίσιμη με $\det A_x = 0$, θα δείξουμε ότι $\lambda(K(N)) = 0$ χρησιμοποιώντας το λήμμα κάλυψης του Besicovitch. Περιγράφουμε αυτά τα αποτελέσματα στην Παράγραφο 3.

(γ) Θεμελιώδη ρόλο στην απόδειξη παίζουν οι ιδιότητες του υποδιαφορικού ∂f μιας κυρτής συνάρτησης $f : C \to \mathbb{R}$. Στην Παράγραφο 4 θα αποδείξουμε ότι η ∂f είναι διαφορίσιμη συνάρτηση σχεδόν παντού.

Στην Παράγραφο 5 συνδυάζουμε όλα τα παραπάνω για την απόδειξη του θεωρήματος του Alexandrov.

2 Lipschitz συναρτήσεις

Έστω A ανοικτό υποσύνολο του $\mathbb{R}^d$. Μια συνάρτηση $f : A \to \mathbb{R}^m$ λέγεται Lipschitz αν υπάρχει σταθερά $M > 0$ τέτοια ώστε

$$\|f(x) - f(y)\| \leqslant M\|x - y\| \tag{2.1}$$

για κάθε $x, y \in A$, όπου $\|z\|$ είναι η Ευκλείδεια νόρμα του z.

Θεώρημα 2.1 (θεώρημα επέκτασης των McShane-Whitney)**.** *Έστω $f : A \to \mathbb{R}$, $A \subseteq \mathbb{R}^d$ μια M-Lipschitz συνάρτηση. Τότε, υπάρχει M-Lipschitz συνάρτηση $f : \mathbb{R}^d \to \mathbb{R}$ τέτοια ώστε $F\,|_A = f$.*

Απόδειξη. Για κάθε $a \in A$ ορίζουμε την συνάρτηση

$$f_a(x) = f(a) + M|x - a|, \qquad x \in \mathbb{R}^d. \tag{2.2}$$

Τώρα, ορίζουμε $F : \mathbb{R}^d \to \mathbb{R}$ με

$$F(x) = \inf_{a \in A} f_a(x). \tag{2.3}$$

Για τυχόν $x \in A$, έχουμε $|f(x) - f(a)| \leqslant M|x - a|$ άρα $f(a) + M|x - a| \geqslant f(x)$ με ισότητα αν $a = x$. Συνεπώς,

$$F(x) = \inf_{a \in A} f_a(x) = \inf_{a \in A}\{f(a) + M|x - a|\} = f(x). \tag{2.4}$$

Μένει να δείξουμε ότι η F είναι M-Lipschitz. Έστω $x, y \in \mathbb{R}^d$. Θα δείξουμε ότι $|F(x) - F(y)| \leqslant M|x - y|$. Ισοδύναμα ζητάμε $-M|x - y| \leqslant F(x) - F(y) \leqslant M|x - y|$, δηλαδή

$$F(y) - M|x - y| \leqslant F(x) \leqslant F(y) + M|x - y|. \tag{2.5}$$

Αρκεί να δείξουμε την δεξιά ανισότητα. Η άλλη, προκύπτει με εναλλαγή των x και y. Θα δείξουμε ότι

$$\begin{aligned} F(x) = \inf_{a \in A}\{f(a) + M|x - a|\} &\leqslant \inf_{a \in A}\{f(a) + M|y - a|\} + M|x - y| \\ &= F(y) + M|x - y|. \end{aligned}$$

Για κάθε $a \in A$ έχουμε $|x-a| \leqslant |y-a| + |x-y|$, άρα $f(a)+M|x-a| \leqslant f(a)+M|y-a| + M|x-y|$, άρα

$$\begin{aligned}\inf_{a\in A}\{f(a)+M|x-a|\} &\leqslant \inf_{a\in A}\{f(a)+M|y-a|+M|x-y|\} \\ &= \inf_{a\in A}\{f(a)+M|y-a|\}+M|x-y|,\end{aligned}$$

αφού, γενικά, ισχύει

$$\inf_{x\in X}\{f(x)+\kappa\} = \inf_{x\in X}\{f(x)\}+\kappa. \tag{2.6}$$

Για την απόδειξη της τελευταίας ισότητας δείχνουμε τις δύο ανισότητες. Ενδεικτικά, για την

$$\inf\{f(x)+\kappa : x\in X\} \geqslant \inf\{f(x) : x\in X\}+\kappa \tag{2.7}$$

αρκεί να δείξουμε πως ο αριθμός $\inf\{f(x) : x\in X\}+\kappa$ είναι κάτω φράγμα του $\{f(x)+\kappa : x\in X\}$. Έστω $x\in X$. Τότε, $f(x)\geqslant \inf\{f(x) : x\in X\}$ άρα $f(x)+\kappa \geqslant \inf\{f(x) : x\in X\}+\kappa$. Το x ήταν τυχόν, άρα το $\inf\{f(x) : x\in X\}+\kappa$ είναι κάτω φράγμα του $\{f(x)+\kappa : x\in X\}$. □

Σημείωση. Το Θεώρημα 2.1 γενικεύεται στην περίπτωση που η f είναι διανυσματική συνάρτηση, δηλαδή, $f : A\to\mathbb{R}^n$, όπου $A\subseteq\mathbb{R}^m$. Η γενίκευση αυτή είναι γνωστή ως το θεώρημα του Kirszbraunn.

Σύμφωνα με ένα κλασσικό αποτέλεσμα του Lebesgue, αν $g : [a,b]\to\mathbb{R}$ είναι μια απολύτως συνεχής συνάρτηση τότε η παράγωγός της ορίζεται σχεδόν παντού στο $[a,b]$, είναι Lebesgue ολοκληρώσιμη, και για κάθε $y<x$ στο $[a,b]$ ισχύει

$$(y) = g(x)+\int_x^y g'(t)dt. \tag{2.8}$$

Οι Lipschitz συναρτήσεις $f : [a,b]\to\mathbb{R}$ είναι απολύτως συνεχείς, συνεπώς το θεώρημα του Lebesgue εφαρμόζεται γι¨ αυτές:

Θεώρημα 2.2. *Έστω $g : [a,b]\to\mathbb{R}$ μια Lipschitz συνάρτηση με σταθερά M. Τότε, η παράγωγος της g ορίζεται σχεδόν παντού στο $[a,b]$ και*

$$|g'(x)|\leqslant M \tag{2.9}$$

σχεδόν για κάθε $x\in[a,b]$. Επίσης, η g είναι Lebesgue ολοκληρώσιμη, και για κάθε $y<x$ στο $[a,b]$ ισχύει

$$(y) = g(x)+\int_x^y g'(t)dt. \tag{2.10}$$

Από την κυρτή ανάλυση γνωρίζουμε πως κάθε κυρτή συνάρτηση $f : A \to \mathbb{R}$, όπου A είναι ένα ανοικτό υποσύνολο του $\mathbb{R}^d$, είναι τοπικά Lipschitz: είναι Lipschitz σε κάθε συμπαγές υποσύνολο του A. Άρα, θα μπορούμε να εφαρμόσουμε το Θεώρημα 2.2 για κάθε ευθύγραμμο τμήμα που περιέχεται στο A.

Υπενθυμίζουμε τώρα τον ορισμό της διαφορισιμότητας για μια απεικόνιση $K : C \to \mathbb{R}^d$.

Ορισμός 2.3. Η απεικόνιση $K : C \to \mathbb{R}^d$, $C \subseteq \mathbb{R}^d$ ανοιχτό, λέγεται *διαφορίσιμη* σε κάποιο $x \in C$ αν υπάρχει $d \times d$ πίνακας A (ο πίνακας Jacobi της K στο x) τέτοιος ώστε

$$K(y) = K(x) + A(y - x) + o(\|x - y\|) \text{ για } y \to x, y \in C. \tag{2.11}$$

Το θεώρημα διαφόρισης του Rademacher εξασφαλίζει ότι οι Lipschitz συναρτήσεις είναι σχεδόν παντού διαφορίσιμες:

Θεώρημα 2.4 (θεώρημα διαφόρισης του Rademacher)**.** *Έστω $C \subseteq \mathbb{R}^d$ ανοιχτό, και $f : C \to \mathbb{R}^m$ μια Lipschitz συνεχής συνάρτηση. Τότε, η f είναι διαφορίσιμη σχεδόν παντού στο C.*

Η απόδειξη του Θεωρήματος 2.4 παραλείπεται. Φυσικά, μας ενδιαφέρει η περίπτωση $m = 1$. Δεδομένου ότι οι κυρτές συναρτήσεις είναι τοπικά Lipschitz, το θεώρημα του Rademacher εφαρμόζεται γι¨ αυτές. Συγκεκριμένα, θα χρησιμοποιήσουμε το θεώρημα του Reidemeister.

Θεώρημα 2.5 (Reidemeister)**.** *Έστω C ανοιχτό και κυρτό υποσύνολο του $\mathbb{R}^d$, και $f : C \to \mathbb{R}$ κυρτή συνάρτηση. Τότε, η f είναι διαφορίσιμη σχεδόν παντού στο C.*

Απόδειξη. Το θεώρημα προκύπτει άμεσα από το θεώρημα του Rademacher και από το γεγονός ότι κάθε κυρτή $f : C \to \mathbb{R}$ είναι Lipschitz στα συμπαγή υποσύνολα του C. □

Θα χρειαστούμε επίσης την εξής πρόταση:

Πρόταση 2.6. *Έστω $K : \mathbb{R}^d \to \mathbb{R}^d$ μια Lipschitz συνεχής συνάρτηση, και έστω $N \subseteq \mathbb{R}^d$ με $\lambda(N) = 0$. Τότε, $\lambda(K(N)) = 0$.*

Απόδειξη. Έστω $\varepsilon > 0$. Αφού $\lambda(N) = 0$, υπάρχουν $B_k = B(a_k, r_k)$, $k \in \mathbb{N}$, ώστε $N \subseteq \bigcup_{k=1}^{\infty} B_k$ και $\sum_{k=1}^{\infty} \lambda(B_k) < \varepsilon$. Είναι: $K(B_k) = \{K(x) : x \in B_k\}$. Η K είναι L-Lipschitz, άρα $\|K(x) - K(a_k)\| \leqslant L\|x - a_k\| \leqslant Lr_k$ για κάθε $x \in B_k$. Δηλαδή, το $K(B_k)$ περιέχεται σε μπάλα ακτίνας Lr_k. Έτσι, $\lambda^*(K(B_k)) \leqslant L^d \lambda(B_k)$. Τελικά,

$$\lambda^*(K(N)) \leqslant \lambda^* \left(\bigcup_{k=1}^{\infty} K(B_k) \right) \leqslant L^d \sum_{k=1}^{\infty} \lambda(B_k) < L^d \cdot \varepsilon. \tag{2.12}$$

Αφήνοντας το $\varepsilon \to 0$, το ζητούμενο έπεται. □

3 Το σύνολο των κρίσιμων σημείων μιας Lipschitz συνάρτησης

Έστω $f : \mathbb{R}^d \to \mathbb{R}^d$, και έστω X το σύνολο των κρίσιμων σημείων της f, δηλαδή το σύνολο των $x \in \mathbb{R}^d$ στα οποία η f δεν είναι διαφορίσιμη ή είναι διαφορίσιμη αλλά ο πίνακας Jacobi της f έχει τάξη μικρότερη από d (δεν είναι αντιστρέψιμος). Αν το x είναι κρίσιμο σημείο της f που ανήκει στην δεύτερη κατηγορία, λέμε ότι η παράγωγος της f στο x είναι ιδιάζουσα. Εμάς, μας ενδιαφέρει το εξής θεώρημα για την κλάση των Lipschitz συναρτήσεων:

Θεώρημα 3.1. *Έστω $K : \mathbb{R}^d \to \mathbb{R}^d$ μια Lipschitz συνεχής συνάρτηση, και έστω $N \subset \mathbb{R}^d$ το σύνολο των σημείων του $\mathbb{R}^d$ στα οποία η K είναι διαφορίσιμη με ιδιάζουσα παράγωγο, δηλαδή $\det A_x = 0$, όπου A_x ο $d \times d$ πίνακας Jacobi της K στο x. Τότε, $\lambda(K(N)) = 0$.*

Το θεώρημα 3.1 προκύπτει από το εξής λήμμα:

Λήμμα 3.2. *Έστω $K : \mathbb{R}^d \to \mathbb{R}^d$ μια Lipschitz συνεχής συνάρτηση, και έστω $N_M \subset \mathbb{R}^d$ το σύνολο των σημείων του $\mathbb{R}^d$ στα οποία η K είναι διαφορίσιμη και $|\det A_x| \leqslant M$, όπου $M > 0$. Τότε,*

$$\lambda(K(N_M)) \leqslant C_1(d) M \lambda(N_M), \tag{3.1}$$

όπου $C_1(d)$ είναι μια θετική σταθερά που εξαρτάται μόνο από τη διάσταση.

Απόδειξη. Η απόδειξη βασίζεται σε ένα κλασσικό λήμμα κάλυψης του Besicovitch:

Λήμμα 3.3 (Besicovitch, 1945)**.** *Υπάρχει σταθερά $C(d)$ ώστε: για κάθε $A \subseteq \mathbb{R}^d$ και για κάθε συνάρτηση $r : A \to (0, \infty)$ με*

$$\sup\{r(x) : x \in A\} < \infty \tag{3.2}$$

μπορούμε να βρούμε ακολουθία $\{a_k\}_{k=1}^{\infty}$ σημείων του A τέτοια ώστε

$$A \subseteq \bigcup_{k=1}^{\infty} \overline{B}(a_k, r(a_k)) \tag{3.3}$$

και, για κάθε $x \in \mathbb{R}^d$, το πλήθος των k για τους οποίους $x \in \overline{B}(a_k, r(a_k))$ είναι το πολύ ίσο με $C(d)$.

Συνεχίζουμε με την απόδειξη του Λήμματος 3.2. Μπορούμε να υποθέσουμε ότι $\lambda(N_M) < \infty$ (αλλιώς, δεν έχουμε τίποτα να δείξουμε). Μπορούμε να βρούμε ανοικτό σύνολο U τέτοιο ώστε $N_M \subseteq U$ και $\lambda(U) \leqslant 2\lambda(N_M)$.

Σταθεροποιούμε $\varepsilon > 0$. Για κάθε $x \in N_M$, χρησιμοποιώντας την υπόθεση ότι $|\det A_x| \leqslant M$ και ένα επιχείρημα παρόμοιο με αυτό της απόδειξης της Πρότασης 2.6, μπορούμε να βρούμε αρκετά μικρό $r_x = r_x(\varepsilon) > 0$ με την εξής ιδιότητα: η κλειστή μπάλα $\overline{B}(x, r_x)$ περιέχεται στο U και

$$\lambda(K(B(x, r_x))) \leqslant (M + \varepsilon)\lambda(B(x, r_x)). \tag{3.4}$$

Από το γεγονός ότι $\overline{B}(x,r_x) \subseteq U$ για κάθε $x \in N_M$ και $\lambda(U) < \infty$, συμπεραίνουμε ότι

$$\sup\{r_x : x \in N_M\} < \infty. \tag{3.5}$$

Μπορούμε τότε να χρησιμοποιήσουμε το λήμμα κάλυψης του Besicovitch και να βρούμε ακολουθία $\{x_k\}_{k=1}^{\infty}$ σημείων του N_M τέτοια ώστε

$$N_M \subseteq \bigcup_{k=1}^{\infty} \overline{B}(x_k,r_k), \tag{3.6}$$

(όπου $r_k := r_{x_k}$) και κάθε $x \in N_M$ ανήκει σε $C(d)$ το πολύ από τις μπάλες $\overline{B}(x_k,r_k)$. Η τελευταία ιδιότητα της κάλυψης μας επιτρέπει να γράψουμε

$$\sum_{k=1}^{\infty} \chi_{\overline{B}(x_k,r_k)}(x) \leqslant C(d)\chi_U(x), \qquad x \in \mathbb{R}^d. \tag{3.7}$$

Συνεπώς,

$$\begin{aligned}\sum_{k=1}^{\infty} \lambda(\overline{B}(x_k,r_k)) &= \int_{\mathbb{R}^d} \sum_{k=1}^{\infty} \chi_{\overline{B}(x_k,r_k)}(x)dx \\ &\leqslant \int_{\mathbb{R}^d} C(d)\chi_U(x)dx \\ &= C(d)\lambda(U) \leqslant 2C(d)\lambda(N_M).\end{aligned}$$

Τώρα, γράφουμε

$$\begin{aligned}\lambda(K(N_M)) &\leqslant \lambda\left(K\Big(\bigcup_{k=1}^{\infty} B(x_k,r_k)\Big)\right) \leqslant \sum_{k=1}^{\infty} \lambda(K(B(x_k,r_k))) \\ &\leqslant (M+\varepsilon)\sum_{k=1}^{\infty} \lambda(B(x_k,r_k)) \leqslant 2C(d)(M+\varepsilon)\lambda(N_M).\end{aligned}$$

Εφόσον το $\varepsilon > 0$ ήταν τυχόν, παίρνουμε το συμπέρασμα με $C_1(d) = 2C(d)$. □

Απόδειξη του Θεωρήματος 3.1. Αν $\lambda(N) < \infty$ τότε ο ισχυρισμός του θεωρήματος προκύπτει άμεσα από το Λήμμα 3.2. Για κάθε $M > 0$ ισχύει $N \subseteq N_M$, άρα

$$\lambda(K(N)) \leqslant \lambda(K(N_M)) \leqslant C_1(d)M\lambda(N_M). \tag{3.8}$$

Αφήνοντας το $M \to 0$ βλέπουμε ότι $\lambda(K(N)) = 0$.

Αν $\lambda(N) = \infty$, γράφουμε $N = \bigcup_{m=1}^{\infty}(N \cap B(0,m))$ και παρατηρούμε ότι $\lambda(K(N \cap B(0,m))) = 0$ εφαρμόζοντας το προηγούμενο αποτέλεσμα για το σύνολο $N \cap B(0,m)$ το οποίο έχει πεπερασμένο μέτρο. Τέλος, από την

$$\lambda(K(N)) = \lambda\left(K\Big(\bigcup_{m=1}^{\infty}(N \cap B(0,m))\Big)\right) \leqslant \sum_{m=1}^{\infty} \lambda(K(N \cap B(0,m))) = 0 \tag{3.9}$$

έπεται το συμπέρασμα. □

4 Υποδιαφορικό κυρτής συνάρτησης

Θεμελιώδη ρόλο στην απόδειξη, έχει το υποδιαφορικό κυρτής συνάρτησης.

Ορισμός 4.1. Αν $C \subseteq \mathbb{R}^d$ είναι ανοικτό κυρτό σύνολο και $f : C \to \mathbb{R}$ είναι μια κυρτή συνάρτηση, τότε για κάθε $x \in C$ το *υποδιαφορικό* της f στο x είναι το σύνολο των $u \in \mathbb{R}^d$ για τα οποία ισχύει

$$f(y) \geqslant f(x) + \langle u, y - x \rangle, \qquad y \in C. \tag{4.1}$$

Γράφουμε

$$\partial f(x) := \{u \in \mathbb{R}^d : f(y) \geqslant f(x) + \langle u, y - x \rangle,\ y \in C\}. \tag{4.2}$$

Παρατηρούμε πως το υποδιαφορικό ορίζει μια απεικόνιση $x \mapsto \partial f(x)$ της οποίας οι τιμές είναι υποσύνολα του $\mathbb{R}^d$. Από την κυρτή ανάλυση γνωρίζουμε πως το $\partial f(x)$ είναι συμπαγές, μη κενό και πως είναι μονοσύνολο, έστω $\partial f(x) = \{u\}$, αν και μόνο αν η f είναι διαφορίσιμη στο x. Σε αυτή την περίπτωση, ταυτίζουμε το $\partial f(x)$ με το u, και τότε το υποδιαφορικό δεν είναι παρά η κλίση της f στο x.

Πρόταση 4.2. *Η απεικόνιση ∂f είναι μονότονη, με την εξής έννοια: αν $x, y \in C$, $u \in \partial f(x)$ και $v \in \partial f(y)$, τότε*

$$\langle x - y, u - v \rangle \geqslant 0. \tag{4.3}$$

Απόδειξη. Από τον ορισμό του υποδιαφορικού έχουμε

$$f(y) \geqslant f(x) + \langle u, y - x \rangle \quad \text{και} \quad f(x) \geqslant f(y) + \langle v, x - y \rangle. \tag{4.4}$$

Προσθέτοντας τις δύο ανισότητες παίρνουμε

$$0 \geqslant \langle u, y - x \rangle + \langle v, x - y \rangle = \langle u, y - x \rangle - \langle v, y - x \rangle = \langle u - v, y - x \rangle, \tag{4.5}$$

άρα

$$\langle u - v, x - y \rangle = -\langle u - v, y - x \rangle \geqslant 0. \tag{4.6}$$

□

Σε αυτό το σημείο, ορίζουμε μια απεικόνιση η οποία θα βοηθήσει στην απόδειξη χρήσιμων ιδιοτήτων του υποδιαφορικού. Στα επόμενα, I είναι η ταυτοτική απεικόνιση. Θεωρούμε το σύνολο

$$D = \bigcup_{x \in C} (x + \partial f(x)). \tag{4.7}$$

Πρόταση 4.3. *Η απεικόνιση (με τιμές σύνολα) $K = (I + \partial f)^{-1}$, με πεδίο ορισμού το D, είναι Lipschitz. Ειδικότερα, οι τιμές της είναι μονοσύνολα.*

Επεξηγήσεις.

- $x + \partial f(x) := \{x + u : u \in \partial f(x)\}$.
- Αν $x + u \in D$ τότε το $K(x + u)$ είναι το σύνολο όλων των $t \in C$ με $x + u \in t + \partial f(t)$.
- Όταν γράφουμε ότι η K είναι Lipschitz, εννοούμε το εξής: αν $w \in K(x + u)$ και $z \in K(y + v)$, τότε

$$\|z - w\| \leqslant \|(y + v) - (x + u)\|. \tag{4.8}$$

Απόδειξη. Έστω $x + u, y + v \in D$ και $w \in K(x + u)$, $z \in K(y + v)$. Υποθέτουμε ότι $z \neq w$, αλλιώς η (4.8) ισχύει τετριμμένα. Έχουμε $x + u \in w + \partial f(w)$ και $y + v \in z + \partial f(z)$. Από την Πρόταση 4.2,

$$\langle z - w, y + v - z - x - u + w\rangle \geqslant 0. \tag{4.9}$$

[Πράγματι, έχουμε $z, w \in C$ και $x + u \in w + \partial f(w)$, άρα $x + u - w \in \partial f(w)$ και $y + v \in z + \partial f(z)$, άρα $y + v - z \in \partial f(z)$.]

Ξαναγράφουμε αυτήν την ανισότητα στη μορφή

$$\langle z - w, y + v - x - u\rangle + \langle z - w, w - z\rangle \geqslant 0, \tag{4.10}$$

απ¨ όπου παίρνουμε

$$\langle z - w, y + v - x - u\rangle \geqslant \langle z - w, z - w\rangle = \|z - w\|^2. \tag{4.11}$$

Έπεται ότι

$$\|z - w\|^2 \leqslant \langle z - w, y + v - x - u\rangle \leqslant \|z - w\| \, \|y + v - x - u\|, \tag{4.12}$$

άρα

$$\|z - w\| \leqslant \|y + v - x - u\|. \tag{4.13}$$

Δείχνουμε τώρα ότι οι τιμές της K είναι μονοσύνολα. Έστω ότι για κάποιο $x \in C$ υπάρχουν $w, z \in K(x + u)$ για κάποιο $u \in \partial f(x)$, και $w \neq z$. Τότε,

$$0 < \|z - w\| \leqslant \|x + u - x - u\| = 0, \tag{4.14}$$

το οποίο είναι άτοπο. □

Στα επόμενα θα ασχοληθούμε με την συνέχεια της απεικόνισης του υποδιαφορικού. Τονίζουμε πως αν το $\partial f(x)$ είναι μονοσύνολο τότε ο ορισμός της συνέχειας είναι ο κλασικός: αν $\{u\} = \partial f(x)$, λέμε ότι η ∂f είναι συνεχής στο x αν για κάθε περιοχή N του x το σύνολο $\partial f(y)$ περιέχεται στην N για κάθε $y \in C$ που η απόστασή του $\|y - x\|$ από το x είναι αρκούντως μικρή.

Εφαρμόζοντας αυτό το θεώρημα για την f έχουμε:

Λήμμα 4.4. *Η f είναι διαφορίσιμη στο $C \setminus Z$, όπου $Z \subseteq C$ είναι ένα σύνολο μέτρου μηδέν.*

Θεώρημα 4.5. *Έστω C ανοιχτό και κυρτό υποσύνολο του $\mathbb{R}^d$, και $f : C \to \mathbb{R}$ κυρτή συνάρτηση. Τότε, η ∂f είναι συνεχής σχεδόν παντού στο C.*

Απόδειξη. Δείχνουμε ότι η ∂f είναι συνεχής στο $C \setminus Z$, όπου Z το σύνολο στο Λήμμα 4.4. Θεωρούμε $x \in C \setminus Z$ και τυχούσα ακολουθία (x_n) διανυσμάτων του $C \subseteq \mathbb{R}^d$ με $x_n \to x$. Θα δείξουμε ότι $\partial f(x_n) \to \partial f(x)$.

Επειδή η f είναι διαφορίσιμη στο x, έχουμε $\partial f(x) = \{u\}$ για κάποιο $u \in \mathbb{R}^d$. Τα $\partial f(x_n)$ όμως, αφού $x_n \in C$, είναι υποσύνολα του $\mathbb{R}^d$. Τονίζουμε πως ο χώρος στον οποίον ανήκουν, είναι αυτός των συμπαγών υποσυνόλων του $\mathbb{R}^d$, τον οποίο συμβολίζουμε με $\mathcal{H}(\mathbb{R}^d)$, εφοδιασμένος με τη μετρική Hausdorff.

Για να δείξουμε την $\partial f(x_n) \to \partial f(x)$, θεωρούμε τυχόντα $u_n \in \partial f(x_n)$ και θα δείξουμε ότι $u_n \to u$. Έχουμε: επειδή $u_n \in \partial f(x_n)$, για κάθε $y \in C$ ισχύει

$$f(y) \geqslant f(x_n) + \langle u_n, y - x_n \rangle. \tag{4.15}$$

Θεωρούμε μια περιοχή $N \subseteq C$ του x τέτοια ώστε η f να είναι L-Lipschitz εκεί. Επειδή $x_n \to x \in N$, υπάρχουν πεπερασμένοι όροι της ακολουθίας έξω απ¨ το N. Απαλοίφοντάς τους, και αλλάζοντας αν χρειάζεται δείκτη (η σύγκλιση δεν επηρεάζεται) θεωρούμε πως $x_n \in N$, $n = 1, 2, \ldots$. Για κάθε $n \in \mathbb{N}$ διαλέγουμε $y_n \in N$ τέτοιο ώστε: $y_n - x_n \neq 0$ και το $y_n - x_n$ έχει την ίδια κατεύθυνση με το διάνυσμα u_n. Από την (4.15) είναι

$$L\|y_n - x_n\| \geqslant f(y_n) - f(x_n) \geqslant \langle u_n, y_n - x_n \rangle = \|u_n\| \, \|y_n - x_n\|, \tag{4.16}$$

άρα $\|u_n\| \leqslant L$. Δηλαδή, η (u_n) είναι φραγμένη.

Για την απόδειξη της $u_n \to u$ αρκεί να δείξουμε ότι για κάθε συγκλίνουσα υπακολουθία (u_{k_n}) της (u_n) έχουμε

$$u_{k_n} \to u. \tag{4.17}$$

Τότε, η $u_n \to u$ έπεται ως εξής: έστω ότι $u_n \not\to u$. Τότε, υπάρχει $\varepsilon > 0$ ώστε $\|u_n - u\| \geqslant \varepsilon$ για άπειρα $n \in \mathbb{N}$. Θεωρούμε την υπακολουθία (u_{k_n}) αυτών των όρων. Είναι φραγμένη, άρα έχει συγκλινουσα υπακολουθία $u_{k_{\lambda_n}} \to v \neq u$. Αφού η $(u_{k_{\lambda_n}})$ είναι υπακολουθία της (u_n) οδηγούμαστε σε άτοπο.

Μένει να δείξουμε την (4.17). Έστω (u_{k_n}) τυχούσα συγκλίνουσα υπακολουθία της (u_n), με $u_{k_n} \to v$. Έχουμε $x_{k_n} \to x$ και η f είναι συνεχής. Από την (4.15) έχουμε

$$f(y) \geqslant f(x_{k_n}) + \langle u_{k_n}, y - x_{k_n} \rangle \tag{4.18}$$

για κάθε $y \in C$. Παίρνοντας όρια, έχουμε

$$f(y) \geqslant f(x) + \langle v, y - x \rangle \tag{4.19}$$

για κάθε $y \in C$. Άρα, $v \in \partial f(x) = \{u\}$, το οποίο αποδεικνύει ότι $v = u$. □

Για $x \in C$, ο ορισμός της διαφορισιμότητας της ∂f είναι αυτός που αναφέρθηκε προηγουμένως για την K, με την προϋπόθεση ότι το $\partial f(x)$ είναι μονοσύνολο, έστω $\partial f(x) = \{u\} \equiv u$.

Ισοδύναμα, μπορούμε να πούμε πως η ∂f είναι διαφορίσιμη στο $x \in C$ αν

$$\partial f(y) = u + H(y - x) + o(\|y - x\|) \tag{4.20}$$

καθώς $y \to x$, $y \in C$, ή αλλιώς,

$$v = u + H(y - x) + o(\|y - x\|) \tag{4.21}$$

καθώς $y \to x$, $y \in C$ ομοιόμορφα ως προς $v \in \partial f(y)$. Θα χρησιμοποιήσουμε την τελευταία περιγραφή της διαφορισιμότητας για να δείξουμε την ακόλουθη πρόταση:

Πρόταση 4.6. *Έστω ότι η ∂f είναι συνεχής στο $x \in C$, και ότι η απεικόνιση K, όπως ορίστηκε προηγουμένως, είναι διαφορίσιμη στο $w = x + u$, όπου $u \in \partial f(x)$, με μη-ιδιάζουσα παράγωγο. Τότε η ∂f είναι διαφορίσιμη στο x. Ιδιαίτερα, η παράγωγός της είναι $H = A^{-1} - I$.*

Απόδειξη. Παρατηρούμε ότι μεγάλο μέρος της απόδειξης είναι αποτέλεσμα της διαφορισιμότητας της K στο w: Δείχνουμε πρώτα ότι

$$K(z) - K(w) = A(z - w) + r, \qquad z \in D, \tag{4.22}$$

όπου το $\frac{\|r\|}{\|z-w\|}$ είναι αυθαίρετα μικρό αν το $\|z - w\|$ είναι αρκούντως μικρό. Με άλλα λόγια, για κάθε $\varepsilon > 0$ υπάρχει $\delta(\varepsilon) > 0$: αν $\|z - w\| < \delta$ τότε $\frac{\|r\|}{\|z-w\|} < \varepsilon$.

Πράγματι, έστω $w = x + u \in D$, $u = \partial f(x)$, $z = y + v \in D$, $y \in C$, $v \in \partial f(y)$ και επομένως $x = K(w)$, $y = K(z)$ από τον ορισμό της K. Έχουμε

$$\begin{aligned} y - x = A(z - w) + r &\Longleftrightarrow A^{-1}(y - x) = z - w + A^{-1}r \\ &\Longleftrightarrow y + v - x - u + A^{-1}r = A^{-1}(y - x) \end{aligned}$$

άρα

$$v = u + (A^{-1} - I)(y - x) - A^{-1}r. \tag{4.23}$$

Πρέπει όμως να δείξουμε ότι το $\frac{\|A^{-1}r\|}{\|y-x\|}$ είναι αυθαίρετα μικρό, ομοιόμορφα ως προς $v \in \partial f(y)$, αν το $\|y-x\|$ είναι αρκούντως μικρό. Αυτό θα γίνει ως εξής:

Αρχικά, ορίζουμε μια νόρμα για πραγματικούς $d \times d$ πίνακες ως εξής: αν $B = (b_{ik}) \in \mathbb{R}^{d \times d}$, τότε

$$\|B\| := \left(\sum_{i,k=1}^{d} b_{ik}^2 \right)^{1/2}. \tag{4.24}$$

Χρησιμοποιώντας την ανισότητα Cauchy-Schwarz εύκολα δείχνουμε το ακόλουθο: για κάθε $p \in \mathbb{R}^d$ έχουμε

$$\|Bp\|_2 \leqslant \|B\| \, \|p\|_2, \tag{4.25}$$

άρα

$$\|p\| = \|B^1 Bp\| \leqslant \|B^{-1}\| \, \|Bp\|. \tag{4.26}$$

Δηλαδή,

$$\|Bp\| \geqslant \frac{\|p\|}{\|B^{-1}\|} \tag{4.27}$$

για κάθε $p \in \mathbb{R}^d$ και για κάθε $B \in \mathbb{R}^{d \times d}$ με $\det B \neq 0$.

Στη συνέχεια αποδεικνύουμε ότι: το $\|z-w\|$ είναι αυθαίρετα μικρό, ομοιόμορφα ως προς $v \in \partial f(y)$, αν το $\|y-x\|$ είναι αρκούντως μικρό. Με άλλα λόγια,

$$\text{Για κάθε } \varepsilon > 0 \text{ υπάρχει } \delta(\varepsilon) > 0 : \text{ αν } \|y-x\| < \delta, \text{ τότε } \|z-w\| < \varepsilon \tag{4.28}$$

ομοιόμορφα ως προς $v \in \partial f(y)$.

Απόδειξη. Χρησιμοποιούμε τη συνέχεια του ∂f στο x. Έστω $\varepsilon > 0$. Γνωρίζουμε πως υπάρχει $\delta(\varepsilon) > 0$ τέτοιο ώστε, αν $y \in C$ και $\|y-x\| < \delta$ τότε το $\partial f(y)$ περιέχεται στην $B(u, \varepsilon)$, όπου $u = \partial f(x)$. Έτσι, για κάθε $v \in \partial f(y)$ έχουμε $\|u-v\| < \varepsilon$ και από την τριγωνική ανισότητα

$$\|z-w\| = \|y+v-x-u\| \leqslant \|v-u\| + \|y-x\|, \tag{4.29}$$

άρα, αφήνοντας το $\|y-x\| \to 0$ (τελικά, θα έχουμε $\|y-x\| < \delta$) παίρνουμε $\|z-w\| < \varepsilon$. □

Τώρα, χρησιμοποιώντας τους ισχυρισμούς (4.22), (4.27) και (4.28) θα δείξουμε ότι

$$\begin{aligned} \|y-x\| &= \|K(z)-K(w)\| \geqslant \|A(z-w)\| - \|r\| \\ &\geqslant \frac{\|z-w\|}{\|A^{-1}\|} - \frac{\|z-w\|}{2\|A^{-1}\|} = \frac{\|z-w\|}{2\|A^{-1}\|}. \end{aligned}$$

Απόδειξη. Από τον ισχυρισμό (4.22) έχουμε

$$K(z) - K(w) = A(z-w) + r, \tag{4.30}$$

άρα

$$\|A(z-w)\| = \|K(z) - K(w) - r\| \leqslant \|K(z) - K(w)\| + \|r\|, \tag{4.31}$$

απ¨ όπου έπεται ότι

$$\|A(z-w)\| - \|r\| \leqslant \|K(z) - K(w)\| = \|y - x\}. \tag{4.32}$$

Από την (4.27),

$$\|A(z-w)\| \geqslant \frac{\|z-w\|}{\|A^{-1}\|}. \tag{4.33}$$

Επειδή $\|r\| > 0$, από την (4.28), για αρκετά μικρό $\|y-x\|$ έχουμε $\frac{\|z-w\|}{2\|A^{-1}\|} \leqslant \|r\|$, άρα

$$\|A(z-w)\| - \|r\| \geqslant \frac{\|z-w\|}{\|A^{-1}\|} - \frac{\|z-w\|}{2\|A^{-1}\|} = \frac{\|z-w\|}{2\|A^{-1}\|}. \tag{4.34}$$

□

Επιστρέφοντας στην (4.23), έχουμε από την τελευταία πρόταση,

$$\frac{2\|A^{-1}\|}{\|z-w\|} \geqslant \frac{1}{\|y-x\|}, \quad \text{άρα} \quad \frac{\|A^{-1}\| \cdot \|r\|}{\|y-x\|} \leqslant \frac{2\|(A^{-1})^2\| \cdot \|r\|}{\|z-w\|}. \tag{4.35}$$

Για $y \to x$, δηλαδή $\|y-x\| \to 0$, από την (4.28) έχουμε ότι και το $\|z-w\| \to 0$, άρα

$$\frac{\|r\|}{\|z-w\|} \to 0 \tag{4.36}$$

από την (4.22). Άρα,

$$\frac{2\|(A^{-1})^2\| \cdot \|r\|}{\|z-w\|} \to 0, \quad \text{άρα} \quad \frac{\|A^{-1}r\|}{\|y-x\|} \to 0 \tag{4.37}$$

καθώς το $y \to x$, ομοιόμορφα ως προς $v \in \partial f(y)$. □

Συνδυάζοντας τα προηγούμενα αποτελέσματα, δείχνουμε το εξής.

Πρόταση 4.7. *Η ∂f είναι διαφορίσιμη σχεδόν παντού στο C.*

Απόδειξη. Από την Πρόταση 4.3, η $K = (I + \partial f)^{-1} : D \to \mathbb{R}^d$ είναι συνάρτηση Lipschitz, άρα επεκτείνεται σε μια συνάρτηση Lipschitz ορισμένη στον $\mathbb{R}^d$, από το θεώρημα επέκτασης των McShane-Whitney (Θεώρημα 2.1). Συμβολίζουμε πάλι με K αυτήν την επέκταση.

Ορίζουμε M το σύνολο των σημείων του $\mathbb{R}^d$ στα οποία η K δεν είναι διαφορίσιμη, και N το σύνολο των σημείων στα οποία είναι διαφορίσιμη αλλά έχει ιδιάζουσα παράγωγο.

Από το Θεώρημα 2.5, την Πρόταση 2.6 και το Θεώρημα 3.1, το σύνολο $K(M \cup N)$ έχει μηδενικό μέτρο. Από το Λήμμα 4.4 και το Θεώρημα 4.5, το σύνολο L έχει κι αυτό μηδενικό μέτρο, άρα το $K(M \cup N) \cup L$ έχει μηδενικό μέτρο.

Θα δείξουμε ότι το ∂f είναι διαφορίσιμο σε κάθε $x \in C \setminus (K(M \cup N) \cup L)$: Για κάθε τέτοιο x, από το Θεώρημα 4.5 έχουμε ότι το ∂f είναι συνεχές στο x, και από τον ορισμό των M και N συμπεραίνουμε πως η K είναι διαφορίσιμη στο $x + u$, με μη-ιδιάζουσα παράγωγο. Πράγματι,

$$x + u = (I + \partial f)(x) = K^{-1}(x) \notin M \cup N. \tag{4.38}$$

Από την Πρόταση 4.6 έπεται το συμπέρασμα. □

5 Απόδειξη του θεωρήματος

Θα δείξουμε ότι κάθε κυρτή συνάρτηση $f : C \to \mathbb{R}$, όπου C κυρτό και ανοιχτό υποσύνολο του $\mathbb{R}^d$, είναι δύο φορές παραγωγίσιμη σχεδόν παντού στο C. Δηλαδή:

Θεώρημα 5.1. *Σχεδόν για κάθε $x \in C$, υπάρχουν $u \in \mathbb{R}^d$ (η κλίση της f στο x) και ένας $H \in \mathbb{R}^{d\times d}$ (ο Εσσιανός πίνακας της f στο x) ώστε:*

$$f(y) = f(x) + \langle u, y - x\rangle + \frac{1}{2}(y - x)^T H(y - x) + o(\|y - x\|^2) \tag{5.1}$$

καθώς το $y \to x$, $y \in C$. Ιδιαίτερα, $u = \partial f(x)$ και H είναι η παράγωγος της ∂f στο x.

Απόδειξη. Η ιδέα είναι η εξής: θα περιορίσουμε την f σε ένα ευθύγραμμο τμήμα του πεδίου ορισμού της, στο οποίο θα είναι παραγωγίσιμη σχεδόν παντού. Ύστερα, θα την γράψουμε ως το ολοκλήρωμα της παραγώγου της.

Από το Λήμμα 4.4 και την Πρόταση 4.7 γνωρίζουμε πως οι f και ∂f είναι διαφορίσιμες σχεδόν παντού στο C. Έστω $x \in C$ στο οποίο είναι και οι δύο διαφορίσιμες. Από την διαφορισιμότητα της ∂f στο x έχουμε

$$\partial f(y) = u + H(y - x) + o(\|y - x\|) \tag{5.2}$$

καθώς το $y \to x$, $y \in C$, όπου $u = \partial f(x)$ και H είναι η παράγωγος της ∂f στο x. Το Λήμμα 4.4 δείχνει και κάτι άλλο: Η f είναι διαφορίσιμη στο $x + th$ για όλα σχεδόν τα μοναδιαία διανύσματα h και για όλα σχεδόν τα $t \geqslant 0$ για τα οποία ισχύει $x + th \in C$.

Θεωρούμε λοιπόν μοναδιαίο $h \in \mathbb{R}^d$ και $t \geqslant 0$ ώστε: $x+th \in C$ και η f είναι διαφορίσιμη στο $x+th$. Έχουμε

$$f(y) = f(x+th) + \langle v, y-(x+th)\rangle + o(\|y-(x+th)\|) \tag{5.3}$$

καθώς το $y \to x+th$, $y \in C$, όπου $v = \partial f(x+th)$. Ειδικότερα, προσεγγίζοντας το $x+th$ από την ίδια ευθεία (θέτουμε $y = x+sh$) έχουμε

$$\begin{aligned} f(x+sh) &= f(x+th) + \langle v, sh-th\rangle + o(\|h\| \cdot |s-t|) \\ &= f(x+th) + \langle v, h\rangle(s-t) + o(|s-t|), \end{aligned}$$

για $x+sh \to x+th$, δηλαδή για $s \to t$ (με $x+th \in C$). Έχουμε δηλαδή ότι η κυρτή συνάρτηση $t \mapsto f(x+th)$ *μιάς* μεταβλητής είναι διαφορίσιμη σχεδόν για όλα τα $t \geqslant 0$ για τα οποία $x+th \in C$. Η παράγωγός της είναι $\langle v, h\rangle$, όπου $v = \partial f(x+th)$.

Εφαρμόζοντας το Θεώρημα 2.2 για την $s \mapsto f(x+sh)$ στο $[0,t]$ έχουμε (από την (5.2))

$$\begin{aligned} f(x+th) &= f(x) + \int_0^t \langle \partial f(x+sh), h\rangle \, ds \\ &= f(x) + \int_0^t \langle u + H(sh) + o(\|sh\|), h\rangle \, ds \\ &= f(x) + \int_0^t (\langle u, h\rangle + s\langle H(h), h\rangle + \langle o(|s|), h\rangle) \, ds \\ &= f(x) + \int_0^t (\langle u, h\rangle + s \cdot h^T H h + o(|s|)) \, ds \\ &= f(x) + t\langle v, h\rangle + \frac{1}{2}(h^T H h)t^2 + o(t^2) \end{aligned}$$

καθώς το $t \to 0$, ομοιόμορφα ως προς h. Άρα,

$$f(y) = f(x) + \langle u, y-x\rangle + \frac{1}{2}(y-x)^T H(y-x) + o(\|y-x\|^2) \tag{5.4}$$

για $y \to x$, $y \in C$, ομοιόμορφα ως προς y της μορφής $x+th$, άρα για σχεδόν όλα τα $y \in C$. Επειδή τα $f(y)$ και $f(x) + \langle u, y-x\rangle + \frac{1}{2}(y-x)^T H(y-x)$ εξαρτώνται συνεχώς από το y, η (5.4) ισχύει για όλα τα $y \in C$.

Έτσι, δείξαμε την (5.1), δηλαδή το θεώρημα του Alexandrov. □

Αναφορές

[1] P. M. Gruber, *Convex anf discrete geometry.*

[2] J. Heinonen, *Lectures on Lipschitz analysis.*

[3] R. Howard, *Alexandrov's theorem on the second derivatives of convex functions vi Rademacher's theorem on the first derivatives of Lipschitz functions.*

[4] Τ. Χατζηαφράτης, *Απειροστικός Λογισμός σε πολλές μεταβλητές.*