חשבון אינפיניטסימלי מתקדם 1

יובל קפלן

סיכום הרצאות פרופ׳ מתניה בן־ארצי בקורס ״חשבון אינפיניטסימלי מתקדם 1״ (80315) באוניברסיטה העברית, 8–2007.

תוכן מחברת זו הוקלד ונערך על־ידי יובל קפלן. אין המרצה אחראי לכל טעות שנפלה בו. סודר באמצעות $\LaTeX\,2_\varepsilon$ ב־27 ביוני 2008.

עדכונים ותיקונים יופיעו ב־/http://www.limsoup.net. לתגובות, לתיקונים ובכל עניין אחר, אנא כתבו ל־yuvak@gmx.net.

סיכומים נוספים בסדרה:

2006–7	חשבון אינפיניטסימלי 1	אלגברה לינארית 1
	חשבון אינפיניטסימלי 2	אלגברה לינארית 2
	תורת הקבוצות	
2007–8	מבנים אלגבריים 1	תורת ההסתברות 1
	חשבון אינפי' מתקדם 1	
בקרוב	חשבון אינפי' מתקדם 2	מבוא לטופולוגיה
	תורת המספרים וקריפטו'	מבנים אלגבריים 2
	תולדות המתמטיקה	

תוכן עניינים

1 מבוא לטופולוגיה של $\mathbb{R}^n$

1.1 נורמה ומרחק

הנורמה האוקלידית ב־$\mathbb{R}^n$ מוגדרת על־ידי $\|x\| = (x_1^2 + \ldots + x_n^2)^{\frac{1}{2}}$. מתקיימות התכונות הבאות: 22.1.2008

1. $\|x\| = 0$ אס"ם $x = 0$;
2. לכל $t \in \mathbb{R}^n$, $\|tx\| = |t|\|x\|$ (הומוגניות), ובפרט $\|x\| = \|-x\|$;
3. $\|x+y\| \le \|x\| + \|y\| \quad \forall x, y \in \mathbb{R}^n$ (אי־שוויון המשולש).

המרחק בין שתי נקודות מוגדר על־ידי $d(x,y) = \|x-y\|$. מתקיימות התכונות הבאות:

1. $d(x,y) = 0$ אס"ם $x = y$;
2. $d(x,y) = d(y,x)$;
3. לכל $x, y, z \in \mathbb{R}^n$, $d(x,z) \le d(x,y) + d(y,z)$.

הוכחה. $d(x,z) = \|x-z\| = \|x-y+y-z\| \le \|x-y\| + \|y-z\|$.

1.2 קבוצות פתוחות, סגורות וחסומות

הגדרה. יהיו $x \in \mathbb{R}^n$ ו־$r > 0$. הקבוצה $B_r(x) = \{y \mid \|y-x\| < r\}$ היא **הכדור הפתוח** מרדיוס r שמרכזו x. הקבוצה $\overline{B}_r(x) = \{y \mid \|y-x\| \le r\}$ היא **הכדור הסגור** מרדיוס r שמרכזו x. כדור פתוח כדור סגור

נעיר שלכל $s < r$, $\overline{B}_s(x) \subseteq B_r(x) \subseteq \overline{B}_r(x)$.

הגדרה. תהי $\varnothing \ne A \subseteq \mathbb{R}^n$. **הקוטר** של A הוא $\delta(A) = \sup\{d(x,y) \mid x,y \in A\}$. קוטר

דוגמה. עבור כדור כלשהו מרדיוס r: $\delta(B_r(x)) = 2r$. עבור $A = \{x\}$ (כלומר, נקודה): $\delta(A) = 0$.

הגדרה. קבוצה $A \subseteq \mathbb{R}^n$ נקראת **חסומה** אם קיים $r > 0$ כך ש־$A \subseteq B_r(0)$. קבוצה חסומה

טענה 1: התנאים הבאים שקולים:

1. A חסומה;
2. $\delta(A) < \infty$;
3. קיימים $x \in \mathbb{R}^n$ ו־$r > 0$ כך ש־$A \subseteq B_r(x)$;
4. לכל $x \in \mathbb{R}^n$ קיים $r > 0$ כך ש־$A \subseteq B_r(x)$.

הגדרה. $A \subseteq \mathbb{R}^n$ נקראת **פתוחה** אם לכל $x \in A$ קיים $\varepsilon > 0$ כך ש־$B_\varepsilon(x) \subseteq A$.[1] קבוצה פתוחה

[1] זה שקול לכך שקיים $\eta > 0$ כך ש־$\overline{B}_\eta(x) \subseteq A$.

דוגמה. $B_r(x)$ פתוחה: יהי $y \in B_r(x)$. אז $r > \|y - x\| = \delta$. ברור ש־$B_{r-\delta}(y) \subseteq B_r(x)$, כי אם $z \in B_{r-\delta}(y)$ אז $\|x - z\| \le \|x - y\| + \|y - z\| < \delta + (r - \delta) < r$.

דוגמה. תהי $f : (a, b) \to (0, \infty)$ פונקציה רציפה. אזי

$$D_f = \{(x, y) \mid a < x < b,\ 0 < y < f(x)\}$$

פתוחה.

קבוצה פתוחה **הגדרה.** $A \subseteq \mathbb{R}^n$ נקראת **סגורה** אם A^C קבוצה פתוחה.

דוגמה. $A = \overline{B}_r(x)$ סגורה: $A^C = \{y \mid \|y - x\| > r\}$ וניקח $B_{\|y-x\|-r}(y)$. בפרט, כל נקודה היא קבוצה סגורה: היא כדור סגור ברדיוס אפס.

טענה 2: איחוד כל משפחה של קבוצות פתוחות הינו פתוח: כלומר, אם U_α עבור $\alpha \in I$ קבוצה פתוחה, אזי $U = \bigcup_{\alpha \in I} U_\alpha$ פתוחה.

הוכחה. יהי $x \in U$. צריך להראות שקיים $r > 0$ כך ש־$B_r(x) \subseteq U$: ואכן, קיים $\alpha \in I$ כך ש־$x \in U_\alpha$, ולכן קיים $r > 0$ כך ש־$B_r(x) \subseteq U_\alpha \subseteq U$.

טענה 3: חיתוך מספר סופי של קבוצות פתוחות הינו פתוח: כלומר, אם $U_1, \dots, U_n$ קבוצות פתוחות, אזי $V = \bigcap_{j=1}^n U_j$ פתוחה.

הוכחה. יהי $x \in V$. לכל $1 \le j \le n$, $x \in U_j$. יהי $r_j > 0$ כך ש־$B_{r_j}(x) \subseteq U_j$. ניקח $r = \min(r_1, \dots, r_n)$, ואז $B_r(x) \subseteq B_{r_j}(x)$ לכל j.

דוגמה (הטענה איננה נכונה לחיתוך כלשהו). $U_j = B_{\frac{1}{j}}(0)$ לכל $j \in \mathbb{N}$. אז $\bigcap U_j = \{0\}$. כלומר, חיתוך בן־מניה של קבוצות פתוחות יכול לא להיות פתוח.

דוגמה. $\varnothing$ ו־$\mathbb{R}^n$ גם פתוחות וגם סגורות. (למעשה, אין ב־$\mathbb{R}^n$ קבוצות נוספות שהן פתוחות וגם סגורות.) יתר על כן, קבוצות לא חייבות להיות פתוחות או סגורות.

טענה 4: חיתוך כל משפחה של קבוצות סגורות הינו סגור.

טענה 5: איחוד מספר סופי של קבוצות סגורות הינו סגור.[2]

טענות אלה נובעות מהקודמות מכיוון שמתקיים כי $A = (A^C)^C$, $(\bigcup A_\alpha)^C = \bigcap A_\alpha^C$ ו־$(\bigcap A_\alpha)^C) = \bigcup A_\alpha^C$.

סגור **הגדרה.** $A \subseteq \mathbb{R}^n$. **הסגור של A** הוא הקבוצה הסגורה המינימלית (ביחס ההכלה) שמכילה את A: מסמנים $\overline{A} = \bigcap_{B \text{ סגורה}, A \subseteq B} B$. **הפנים של A** הוא הקבוצה הפתוחה המקסימלית המוכלת פנים ב־A: מסמנים $A^\circ = \operatorname{int} A = \bigcup_{U \text{ פתוחה}, U \subseteq A} U$.

[2] גם כאן, איחוד בן־מניה של קבוצות סגורות לא חייב להיות סגור: הוא יכול, למשל, להיות $\mathbb{R}^n$ פחות נקודה.

דוגמה. $\overline{B_r(x)} = \overline{B}_r(x)$.

הגדרה. תהי $\{x^{(k)}\}_{k=1}^\infty$. $x \in \mathbb{R}^n$ הוא **גבול** של $\{x^{(k)}\}_{k=1}^\infty$ אם לכל $\varepsilon > 0$ קיים K כך ש-$x^{(k)} \in B_\varepsilon(x)$ לכל $k > K$. מסמנים $x = \lim_{k\to\infty} x^{(k)}$, $x^{(k)} \to x$. גבול של סדרה

לא לכל סדרה יש גבול, אך אם הוא קיים הוא יחיד: אם $x = \lim x^{(k)}$ ו-$y = \lim x^{(k)}$, לכל $\varepsilon > 0$ קיים k כך ש-$\|x - x^{(k)}\| < \varepsilon$ וגם $\|y - x^{(k)}\| < \varepsilon$, לכן $\|x-y\| \le \|x - x^{(k)}\| + \|y - x^{(k)}\| < 2\varepsilon$. זה נכון לכל ε, לכן $\|x-y\| = 0$ ו-$x = y$.

טענה 6: A קבוצה סגורה אס"ם $\{x^{(k)}\}_{k=1}^\infty \subseteq A$ מתכנסת אזי $\lim_{k\to\infty} x^{(k)} \in A$.

הוכחה. ($\Leftarrow$) נניח ש-A סגורה. תהי $\{x^{(k)}\} \subseteq A$, $x^{(k)} \to x$. צריך להראות כי $x \in A$. נניח בשלילה כי $x \in A^C$. A^C פתוח, לכן קיים $\varepsilon > 0$ כך ש-$B_\varepsilon(x) \subseteq A^C$. אבל $x^{(k)} \to x$, לכן קיים k כך ש-$B_\varepsilon(x) \ni x^{(k)} \in A$ - סתירה.

($\Rightarrow$) צריך להראות ש-A^C פתוחה. נניח בשלילה שקיים $x \in A^C$ כך שלכל $\varepsilon > 0$, $B_\varepsilon(x) \not\subseteq A^C$. בפרט, לכל $k \in \mathbb{N}$ קיים $x^{(k)} \in A \cap B_{\frac{1}{k}}(x)$. ברור ש-$A \ni x^{(k)} \to x$, ולפי ההנחה גם $x \in A$ - סתירה.

23.1.2008 **הגדרה.** תהי $A \subseteq \mathbb{R}^n$. אומרים שקבוצה $B \subseteq A$ **צפופה** ב-A אם $A \subseteq \overline{B}$. קבוצה צפופה

דוגמה. $\mathbb{Q}$ צפופה ב-$\mathbb{R}$.

טענה 7: B צפופה ב-A אס"ם לכל $x \in A$ ולכל $r > 0$, $B_r(x) \cap B \neq \varnothing$.

הוכחה. ($\Leftarrow$) בדרך השלילה. נניח שיש $x \in A$ ו-$r > 0$ כך ש-$B_r(x) \cap B = \varnothing$. כלומר, $B \subseteq B_r(x)^C$ ו-$B_r(x)^C$ סגורה, לכן $\overline{B} \subseteq B_r(x)^C$. אזי גם $\overline{B}_r(x) \cap B = \varnothing$ ובפרט $x \notin \overline{B}$ - סתירה.

($\Rightarrow$) יהי $x \in A$. לפי ההנחה, לכל $k \in \mathbb{N}$ קיים $x^{(k)} \in B_{\frac{1}{k}}(x) \cap B$. ברור ש-$B \ni x^{(k)} \to x$, לכן $x \in \overline{B}$.

דוגמה. $\{(\frac{m}{2^k}, \frac{p}{2^k}) \mid m, p \in \mathbb{Z},\ k \in \mathbb{N}\}$ היא קבוצה צפופה ב-$\mathbb{R}^2$: לכל (a,b), מתקיים $0 \le a - \frac{m}{2^k} < 2^{-k}$, $0 \le b - \frac{p}{2^k} < 2^{-k}$ עבור m, p ו-k כלשהם; לכן מתקיים התנאי שבטענה.

הגדרה. **השפה** של A היא $\partial A = \overline{A} \cap \overline{A^C}$. שפה

טענה 8: $\partial A = \overline{A} \setminus A^\circ = \overline{A} \cap (A^\circ)^C$.

הוכחה. נראה כי $A^\circ = (\overline{A^C})^C$:

($\subseteq$) מספיק להוכיח כי $A^C \subseteq (A^\circ)^C$, אך זה שקול לכך ש-$A^\circ \subseteq A$, שנכון לפי הגדרה.

($\supseteq$) מספיק להוכיח ש-$(\overline{A^C})^C \subseteq A$, אך זה שקול לכך ש-$A^C \subseteq \overline{A^C}$, וזה ברור.

1.3 קבוצות קומפקטיות

כיסוי
כיסוי פתוח

הגדרה. $A \subseteq \mathbb{R}^n$. משפחת קבוצות $\{U_k\}$ נקראת **כיסוי** של A אם $A \subseteq \bigcup_\alpha U_\alpha$. כיסוי נקרא **פתוח** אם כל הקבוצות U_α פתוחות.

דוגמה. $\{\mathbb{R}^n\}$ מהווה כיסוי לכל $A \subseteq \mathbb{R}^n$.

דוגמה. נניח שלכל $x \in A$ מותאם מספר $r(x) > 0$; אזי $\{B_{r(x)}(x)\}_{x\in A}$ כיסוי פתוח. אם $A = B_r(x)$, ניתן לעשות זאת על־ידי בחירת $r(y) = r - \|x-y\|$.

תת־כיסוי

הגדרה. יהי $\{U_\alpha\}_{\alpha\in I}$ כיסוי של A. המשפחה $\{U_\alpha\}_{\alpha\in J}$, $J \subseteq I$ נקראת **תת־כיסוי** של $\{U_\alpha\}_{\alpha\in I}$ אם גם היא כיסוי של A.

דוגמה. $\{(n-\frac{1}{2}, n+\frac{1}{2}) \mid n \in \mathbb{Z}\}$ הוא כיסוי של $\mathbb{Z}$ שאין לו תת־כיסוי־ממש.

דוגמה. נסמן $\mathbb{Q} = \{q_n\}_{n=1}^\infty$ ונגדיר $U_n = (q_n - 2^{-n}, q_n + 2^{-n})$. אז $\mathbb{Q} \subseteq \bigcup_{n=1}^\infty U_n$. כיסוי זה מקיים $|U_n| = 2\cdot 2^{-n}$. אם נגדיר לכל ε מחדש $U_n = (q_n - 2^{-n}\varepsilon, q_n + 2^{-n}\varepsilon)$, נקבל $\sum |U_n| = 2\varepsilon$ - כלומר, ניתן להקטין את נפח הכיסוי כרצוננו.

קבוצה קומפקטית

הגדרה. $A \subseteq \mathbb{R}^n$ נקראת **קומפקטית** אם לכל כיסוי פתוח של A יש תת־כיסוי סופי.

דוגמה. $\mathbb{Z}$, לפי הדוגמה למעלה, אינה קומפקטית.

טענה 9: A קומפקטית $\Longleftarrow$ A חסומה.

הוכחה. המשפחה $\{B_r(0)\}_{r>0}$ היא כיסוי פתוח של A. מקומפקטיות A, קיימים $r_1, \dots, r_m$ כך ש־$A \subseteq \bigcup_{j=1}^m B_{r_j}(0)$. ניקח $r = \max(r_1, \dots, r_m)$. אז $\bigcup_{j=1}^m B_{r_j}(0) = B_r(0)$, ולכן A חסומה.

טענה 10: A קומפקטית $\Longleftarrow$ A סגורה.

הוכחה. צריך להראות ש־A^C פתוח, כלומר לכל $x \in A^C$ צריך להראות שקיים $r > 0$ כך ש־$A \cap B_r(x) = \varnothing$. עבור $y \in A$ נגדיר $r(y) = \frac{1}{2}\|y-x\|$. כמובן, $\{U_y = B_{r(y)}(y)\}_{y\in A}$ הוא כיסוי פתוח של A. מקומפקטיות A, קיימים $y_1, \dots, y_m \in A$ כך ש־$A \subseteq \bigcup_{j=1}^m U_{y_j}$. ניקח $r = \min(r(y_1), \dots, r(y_m))$. אז $A \cap B_r(x) = \varnothing$, כי $U_{y_j} \cap B_r(x) = \varnothing$ לפי בחירת r, ולכן $\bigcup_{j=1}^m U_{y_j} \cap B_r(x) = \varnothing$.

טענה 11: אם $B \subseteq \mathbb{R}^n$ קומפקטית ו־$A \subseteq B$ סגורה אזי A גם קומפקטית.

הוכחה. יהי $\{U_\alpha\}$ כיסוי פתוח של A. ברור ש־$\{U_\alpha\}\bigcup\{A^C\}$ הוא כיסוי פתוח של $\mathbb{R}^n$, לכן בפרט של B. מקומפקטיות B, קיים תת־כיסוי סופי של B. נניח שזהו $A^C, U_{\alpha_1}, \dots, U_{\alpha_m}$. ברור ש־$U_{\alpha_1}, \dots, U_{\alpha_m}$ הוא כיסוי של A, שהרי $B \subseteq A^C \cup \bigcup_{j=1}^m U_{\alpha_j}$.

הגדרה. **תיבה** ב־$\mathbb{R}^n$ היא קבוצה מן הצורה תיבה
$$I = [a_1, b_1] \times \ldots \times [a_n, b_n] = \{(x_1, \ldots, x_n) \mid a_1 \le x_1 \le b_1, \ldots, a_n \le x_n \le b_n\}$$
כאשר $a_1 \le b_1, \ldots, a_n \le b_n$.

למה 12 (קנטור): יהיו $c_k \le d_k$ כך ש־$\ldots \le c_2 \le c_1$ ו־$\ldots \ge d_2 \ge d_1$. אזי מתקיים $[\sup c_k, \inf d_k] = \bigcap_{k=1}^{\infty} [c_k, d_k] \neq \varnothing$.

למה 13: תהיינה $I_1, I_2, \ldots$ תיבות לא ריקות ב־$\mathbb{R}^n$ כך ש־$I_1 \supseteq I_2 \supseteq \ldots$. אזי $\bigcap_{j=1}^{\infty} I_j \neq \varnothing$.

הוכחה. נובע מהלמה של קנטור במימד אחד עבור כל אחת מהקואורדינטות.

טענה 14: תיבה היא קבוצה קומפקטית.

הוכחה. בדרך השלילה. תהי I תיבה ויהי $\{U_\alpha\}$ כיסוי פתוח שלה שאין לו תת־כיסוי סופי. ברור שיש 2^n תיבות $Q_1, \ldots, Q_{2^n}$ ב־I כך ש־$I = \bigcup_{j=1}^{2^n} Q_j$ ו־$\delta(Q_j) = \frac{1}{2}\delta(I)$.[3] לפחות עבור j אחד, אין ל־U_α תת־כיסוי סופי של Q_j. נקרא לתיבה זו I_1. באינדוקציה, נמצא תיבות $\ldots \subseteq I_2 \subseteq I_1 \subseteq I$ כך ש־$\delta(I_k) = \frac{1}{2^k}\delta(I)$ ול־I_k אין תת־כיסוי סופי מ־$\{U_j\}$. לפי הלמה של קנטור, $B = \bigcap I_k \neq \varnothing$ ו־$\delta(B) = 0$; אז $B = \{x\}$. יהי α כך ש־$x \in U_\alpha$. קיים $r > 0$ כך ש־$B_r(x) \subseteq U_\alpha$. יהי k כך ש־$\delta(I_k) = \frac{1}{2^k}\delta(I) < r$. אז $x \in I_k \subseteq B_r(x) \subseteq U_\alpha$. כלומר, קיבלנו כיסוי של I_k על־ידי איבר אחד של $\{U_\beta\}$ – סתירה.

משפט 15: $A \subseteq \mathbb{R}^n$ קומפקטית אס"ם A סגורה וחסומה.

הוכחה. ($\Leftarrow$) כבר הוכחנו.

($\Rightarrow$) A חסומה, לכן ברור שקיימת תיבה I המכילה את A.[4] קיבלנו ש־A קבוצה סגורה המוכלת בקבוצה קומפקטית, ולפי טענה קודמת, A קומפקטית בעצמה.

1.4 קבוצות קשירות

29.1.2008 **הגדרה.** תהי $C \subseteq \mathbb{R}^n$. נאמר כי C **קשירה** אם עבור U ו־V פתוחות כך ש־$U \cap C$ זרה ל־$V \cap C$ ומתקיים $C \subseteq U \cup V$ אזי $U \cap C = \varnothing$ או $V \cap C = \varnothing$. קבוצה קשירה

דוגמה. הקטע $[0,1]$ קשיר.

הוכחה. נניח בשלילה כי קיימות קבוצות פתוחות $U, V \subseteq \mathbb{R}^n$ כך שמתקיים $I \subseteq U \cup V$ ו־$(U \cap I) \cap (V \cap I) = \varnothing$ אבל $U \cap I \neq \varnothing$ ו־$V \cap I \neq \varnothing$.

ניקח נקודה $x \in U \cap I$ ונקודה $y \in V \cap I$. בלי הגבלת הכלליות, $x < y$. נסתכל בקבוצה $U_{x,y} = U \cap [x, y]$. נקבוצה הזו מוכלת ב־$[x, y - \delta]$ עבור $\delta > 0$ מסויים, כי $y \in V$ ו־V פתוחה. נתבונן ב־$z = \sup U_{x,y}$. מתקיים $x < z \le y - \delta$.

[3] אנו בעצם חוצים את התיבה I בכל מימד.

[4] אם $A \subseteq B_r(0)$ אז $A \subseteq [-r, r]^n$.

האם $z \in U$? לא, כי אחרת קיים $\eta > 0$ כך ש־$U \supseteq (z-\eta, z+\eta)$ ונובע כי $z \neq \sup U_{x,y}$.

האם $z \in U$? לא, כי אם כן אזי קיים $\eta > 0$ כך ש־$V \supseteq (z-\eta, z+\eta)$ ולכן $z \neq \sup U_{x,y}$.

מצאנו, אם כן, נקודה $z \in I$ אשר אינה שייכת ל־U או ל־V, בסתירה לכך ש־$I \subseteq U \cup V$.

2 פונקציות וקטוריות על $\mathbb{R}^n$

2.1 פונקציות ממשיות

הגדרה. **פונקציה ממשית** f על קבוצה $D \subseteq \mathbb{R}^n$ היא התאמה חד־ערכית של מספר ממשי $f(x)$ לכל $x \in D$.

הגדרה. תהי $f : D \to \mathbb{R}$ ו־$x_0 \in D$. נאמר כי f **רציפה** ב־x_0 אם לכל $\varepsilon > 0$ קיים $\delta > 0$ כך ש־$f(B(x_0, \delta) \cap D) \subseteq (f(x_0) - \varepsilon, f(x_0) + \varepsilon)$. רציפות
אם f רציפה בכל נקודה ב־D, נאמר כי f רציפה על D.

טענה 16: f רציפה ב־$x_0 \in D$ אס"ם לכל סדרה $\{x^{(k)}\}_{k=1}^{\infty} \subseteq D$ המתכנסת ל־x_0 מתקיים $f(x^{(k)}) \to f(x_0)$.

הוכחה. ($\Leftarrow$) תהי $x^{(k)} \to x_0$ סדרה ב־D. יש להראות כי $f(x^{(k)}) \stackrel{k\to\infty}{\longrightarrow} f(x_0)$. נניח בשלילה כי $f(x^{(k)}) \not\to f(x_0)$. אזי קיים $\eta > 0$ כך ש־$(f(x_0) - \eta, f(x_0) + \eta)$ לא מכיל את הערכים $\{f(x^{(k_1)}), \ldots, f(x^{(k_j)}), \ldots\}$ עבור תת־סדרה $\ldots < k_j < \ldots < k_2 < k_1$. אבל מהגדרת הרציפות, קיים $\delta_\eta > 0$ כך ש־$f(B(x_0, \delta_\eta) \cap D) \subseteq (f(x_0) - \eta, f(x_0) + \eta)$, וקיבלנו סתירה לעובדה שאם $x^{(k)} \to x_0$ אז החל מ־k מסויים, $x^{(k)} \in B(x_0, \delta_\eta)$.

($\Rightarrow$) נניח שמתקיים תנאי הסדרות ונוכיח רציפות ב־x_0. נניח בשלילה כי f איננה רציפה ב־x_0, כלומר קיים $\eta > 0$ כך שלכל $\delta > 0$, $f(B(x_0, \delta) \cap D) \nsubseteq (f(x_0) - \eta, f(x_0) + \eta)$. בפרט, ניקח סדרה $\delta_k = \frac{1}{k}$ ונמצא $x^{(k)} \in D$, $|x^{(k)} - x_0| < \frac{1}{k}$ כך ש־$f(x^{(k)}) \notin (f(x_0) - \eta, f(x_0) + \eta)$, כלומר $|f(x^{(k)}) - f(x_0)| \geq \eta$ לכל $k \in \mathbb{N}$. אבל אז מתקיים $x^{(k)} \to x_0$ ו־$f(x^{(k)}) \not\to f(x_0)$, בסתירה לתנאי הסדרות.

טענה 17: תהי $C \subseteq \mathbb{R}^n$ קבוצה קומפקטית ותהי $f : C \to \mathbb{R}$ פונקציה רציפה. אזי f חסומה ומקבלת את המינימום והמקסימום שלה על C. במילים אחרות, קיימים $m \leq M$ כך ש־(א) $\forall x \in C \quad m \leq f(x) \leq M$ ו־(ב) $\exists x_m, x_M \in C \ \mid \ f(x_m) = m,\ f(x_M) = M$.

הוכחה. (א) ראשית נראה כי קיים M כזה. בשלילה, קיימת סדרה $\{x^{(k)}\}_{k=1}^{\infty} \subseteq C$ כך ש־$f(x^{(k)}) \geq k$. לפי המשפט לגבי קבוצות קומפקטיות, קיימת תת־סדרה $x^{(k_j)} \stackrel{j\to\infty}{\longrightarrow} x_0 \in C$. f רציפה ב־x_0 ולכן מתקיים $f(x^{(k_j)}) \to f(x_0)$. אבל $f(x^{(k_j)}) \geq k_j$ – סתירה.

(ב) נראה קיום x_M. ראשית, ניקח $M = \sup_{x \in C}\{f(x)\} < \infty$. קיימת סדרה $\{x^{(m)}\} \subseteq C$ כך ש־$f(x^{(k)}) \to M$; מקומפקטיות, לסדרה זו יש תת־סדרה $x^{(k_j)} \to x_M \in C$; מרציפות, $f(x^{(k_j)}) \to f(x_M)$. אז $M = f(x_M)$ (מכיוון שגבול סדרה מתכנסת יחיד).

המקרה של m: כתרגיל. (הדרך הקצרה: להתבונן ב־$-f$.)

מסקנה 18: תהי f רציפה על קבוצה קומפקטית C. נניח כי $f(x) > 0$ לכל $x \in C$. אזי קיים $\alpha > 0$ כך ש־$\forall x \in C \quad f(x) \geq \alpha$.[5]

[5] כמסקנה ממסקנה זו, הקטע $(0, 1]$ איננו קומפקטי.

הוכחה. יהי $0 \le m = \inf_{x\in C}\{f(x)\}$. קיים $x_m \in C$ כך ש־$f(x_m) = m$, אז $m = \alpha > 0$.

דוגמה. C אינה קומפקטית, f רציפה ולא מקבלת אקסטרמום: $f(x) = x$ על $C = (0,1)$. 5.2.2008

משפט 19: $f : C \to \mathbb{R}$ רציפה, $C \subseteq \mathbb{R}^n$ קומפקטית $\Longleftarrow$ f רציפה במידה שווה, כלומר לכל $\varepsilon > 0$ קיים $\delta > 0$ כך שלכל $x, y \in C$ כך ש־$|x-y| < \delta$ מתקיים $|f(x)-f(y)| < \varepsilon$.

הוכחה. נניח שלא, ותהיינה $\{x^k\},\{y^k\} \subseteq C$ שתי סדרות כך ש־(א) $|x^k - y^k| < \frac{1}{k}$ ו־(ב) $|f(x^k) - f(y^k)| > \varepsilon_0$ עבור $\varepsilon_0 > 0$ קבוע.

קיימת תת־סדרה $x^{k_j} \to \bar{x} \in C$, ואז גם $y^{k_j} \to \bar{x}$.[6] כיוון ש־f רציפה ב־$\bar{x} \in C$, קיים $\delta = \delta(\varepsilon_0)$ כך שלכל $z \in C$ כך ש־$|z - \bar{x}| < \delta$ מתקיים $|f(z) - f(\bar{x})| < \frac{1}{2}\varepsilon_0$.

עכשיו, עבור j מספיק גדול, $|x^{k_j} - \bar{x}| < \delta$ ו־$|y^{k_j} - \bar{x}| < \delta$ ולכן $|f(x^{k_j}) - f(\bar{x})| < \frac{1}{2}\varepsilon_0$ ו־$|f(y^{k_j}) - f(\bar{x})| < \frac{1}{2}\varepsilon_0$. אז $|f(x^{k_j}) - f(y^{k_j})| < \varepsilon_0$, בסתירה ל־(ב).

2.2 שקילות הנורמות ב־$\mathbb{R}^n$

משפט 20: תהי $\|\cdot\|$ נורמה ב־$\mathbb{R}^n$. אזי קיים $\lambda > 1$ כך שלכל $x \in \mathbb{R}^n$, $\lambda^{-1}|x| \le \|x\| \le \lambda|x|$, כאשר $|x|$ היא הנורמה האוקלידית.

הוכחה. נרשום $x = \sum_{i=1}^n x_i e_i$ כאשר $\{e_i\}_{i=1}^n$ בסיס ב־$\mathbb{R}^n$. מאי־שוויון המשולש ואי־שוויון קושי־שוורץ, נקבל

$$\|x\| \le \sum_{i=1}^n |x|\|e_i\| \le (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}(\sum_{i=1}^n \|e_i\|^2)^{\frac{1}{2}} = \mu|x|$$

כאשר $\mu = (\sum_{i=1}^n \|e_i\|^2)^{\frac{1}{2}}$.

מצד שני, עלינו להראות כי קיים $\theta > 0$ כך ש־$\theta|x| \le \|x\|$ לכל $x \in \mathbb{R}^n$. נסתכל בספירת (שפת כדור) היחידה $S = \{|x| = 1\}$. זוהי קבוצה קומפקטית: היא חסומה וסגורה.[7]

למה 1.20: $0 < g(x) = \|x\|$ רציפה על S.

הוכחה. לכל $x, y \in S$, $|g(x) - g(y)| = |\|x\| - \|y\|| \le \|x - y\| \le \mu|x-y|$. כלומר, אם $\varepsilon > 0$ נבחר $\delta = \frac{\varepsilon}{\mu}$ ואז $|x - y| < \delta$, לכן $\mu|x-y| < \varepsilon$ ו־$|g(x) - g(y)| < \varepsilon$.

על־פי המשפט הקודם, קיים $\alpha > 0$ כך ש־$g(x) = \|x\| \ge \alpha$ לכל $x \in S$, כלומר לכל $|x| = 1$. יהי $0 \ne z \in \mathbb{R}^n$ כלשהו. נתבונן ב־$x = \frac{z}{|z|} \in S$. אז $\|x\| = \|\frac{z}{|z|}\| \ge \alpha$ ולכן $\|z\| \ge \alpha|z|$. ברור ש־$\|z\| \ge \alpha|z|$ גם עבור $z = 0$, ולכן נוכל לקחת $\theta = \alpha$.

לסיכום, $\theta|x| \le \|x\| \le \mu|x|$ לכל $x \in \mathbb{R}^n$, ונבחר $\lambda > 1$ כך ש־$\lambda \ge \mu$ ו־$\lambda^{-1} \le \theta$.

מסקנה 21: תהיינה $\|\cdot\|_1$ ו־$\|\cdot\|_2$ שתי נורמות על $\mathbb{R}^n$. אזי קיים $\lambda_{12} > 1$ כך שלכל $x \in \mathbb{R}^n$, $\lambda_{12}^{-1}\|x\|_1 \le \|x\|_2 \le \lambda_{12}\|x\|_1$.

---

[6] כי $|y^{k_j} - \bar{x}| \le |y^{k_j} - x^{k_j}| + |x^{k_j} - \bar{x}| \xrightarrow{j\to\infty} 0$.

[7] S סגורה כי היא שפת כדור וכל שפה היא קבוצה סגורה, וכמו־כן גבול כל סדרת נקודות מתכנסת ב־S הוא ב־S והקבוצה המשלימה ל־S פתוחה.

דוגמה. ניקח $\|x\| = |x_1| + |x_2|$ ($n=2$). נסתכל בקבוצה $\|x\| = 1$, שהיא ספירת היחידה בנורמה הזו: זהו ריבוע סביב הראשית. משקילות נורמה זו לנורמה האוקלידית נסיק שיש לספירת יחידה זו מעגל חוסם (לא בהכרח צמוד) ומעגל חסום.

מהדיון לעיל נובע שכל המושגים שלמדנו עד כה אינם תלויים בנורמה: למשל,

1. התכנסות סדרה ב־$\mathbb{R}^n$: אמרנו כי שאיפה כזו פירושה שלכל $\varepsilon > 0$ קיים $K(\varepsilon)$ כך שלכל $k > K(\varepsilon)$ מתקיים $|x^k - x| < \varepsilon$; כעת נוכל להגדיר זאת כך: קיימת נורמה $\|\cdot\|$ כך שלכל $\varepsilon > 0$ קיים $\tilde{K}(\varepsilon)$ כך שלכל $k > \tilde{K}(\varepsilon)$ מתקיים $\|x^k - x\| < \varepsilon$ (ואז $|x^k - x| < \lambda\varepsilon$). ניקח, למשל, את הנורמה $\|x\|_\infty = \max_{1\le i\le n}|x_i|$: אז $\|x^k - x\|_\infty \to 0$ אם"ם $\max_{1\le i\le n}|x_i^k - x_i| \to 0$, ומשקילות נורמה זו לנורמה האוקלידית, נקבל שהתכנסות שקולה להתכנסות רכיב־רכיב.
2. רציפות פונקציות: אמרנו כי $f : D \to \mathbb{R}$ רציפה ב־x_0 אם לכל $\varepsilon > 0$ קיים $\delta > 0$ כך ש־$f(y \in D \mid |y - x_0| < \delta) \subseteq (f(x_0) - \varepsilon, f(x_0) + \varepsilon)$; כעת נוכל להגדיר זאת כך: קיימת נורמה $\|\cdot\|$ כך שלכל $\varepsilon > 0$ קיים $\delta > 0$ כך שלכל $y \in D$ כך ש־$\|y - x_0\| < \delta$ מתקיים $\|f(y) - f(x_0)\| < \varepsilon$.
3. תכונות קבוצות פתוחות, סגורות וחסומות אינן תלויות בנורמה. (בדיקה כתרגיל.)

6.2.2008 **הגדרה.** לכל $\varnothing \neq E \subseteq \mathbb{R}^n$ ולכל נקודה $x \in \mathbb{R}^n$ נגדיר את **המרחק** של x מ־E על־ידי $dist(x, E) = \inf_{y\in E}\{|x-y|\}$. מרחק נקודה מקבוצה

המרחק תמיד קיים (אינפימום מספרים אי־שליליים). כמו כן, אם $x \in E$ אז $dist(x,E) = 0$.[8]

משפט 22: לכל $E \subseteq \mathbb{R}^n$, הפונקציה $d(x) = dist(x, E)$ רציפה עבור $x \in \mathbb{R}^n$.

הוכחה. יהיו $x^1, x^2 \in \mathbb{R}^n$. אז $|x^1 - y| \ge |x^2 - y| - |x^1 - x^2|$. ניקח $\inf$ באגף ימין:

$$|x^1 - y| \ge dist(x^2, E) - |x^1 - x^2|$$

ניקח $\inf$ באגף שמאל:

$$dist(x^1, E) \ge dist(x^2, E) - |x^1 - x^2|$$

וכן, באופן סימטרי, מתקיים

$$dist(x^2, E) \ge dist(x^1, E) - |x^1 - x^2|$$

הפיתוח הפורמאלי - כתרגיל; מתקבל $|d(x^1) - d(x^2)| \le |x^1 - x^2|$.[9]

משפט 23: אם $C \subseteq \mathbb{R}^n$ סגורה ו־$x \in \mathbb{R}^n$, קיימת נקודה $x_c \in C$ כך ש־$dist(x, C) = |x - x_c|$.

הוכחה. קיים $r > 0$ כך ש־$C_r = C \cap \overline{B(x,r)} \neq \varnothing$ ($r > |x-y|$ עבור איזושהי נקודה $y \in C$). C_r קומפקטית.[10] יתר על כן, $dist(x, C_r) = dist(x, C)$, כי כל הנקודות ב־$C \setminus C_r$ מרחקן מ־x גדול מ־r. לכן להוכחת המשפט מספיק לחפש נקודה $x_c \in C_r$.

[8] יתר על כן, הקבוצה E סגורה אם"ם ($x \in E \iff d(x) = 0$).
[9] קיבלנו, למעשה, יותר מסתם רציפות: רציפות במידה שווה, ליפשיציות.
[10] היא חסומה וסגורה.

יהי $\bar{d} = dist(x, C_r)$. כיוון ש־$dist$ הוגדר על־ידי inf הרי קיימת סדרה $\{x^k\}_{k=1}^{\infty} \subseteq C_r$ כך ש־$\bar{d} \searrow |x - x^k|$. מצד שני, C_r קומפקטית ולכן קיימת תת־סדרה $x^{k_j} \to x_c \in C_r$. אז $|x - x^{k_j}| \to |x - x_c|$, $\bar{d} \leftarrow$ ולכן $\bar{d} = |x - x_c|$.

באופן חלופי, נסתכל בפונקציה $d(y) = |x - y|$, $y \in C_r$. $d(y)$ רציפה ב־y, ולכן מקבלת מינימום על C_r.[11] אז קיימת $x_c \in C_r$ כך ש־$|x - x_c| = \min_{y \in C_r} |x - y| = d(x, C_r)$.

2.3 פונקציות וקטוריות

$f : D \to \mathbb{R}^m$ ‏($D \subseteq \mathbb{R}^n$); לכל $x \in D$, $f(x) = (f_1(x), f_2(x), \dots, f_m(x)) \in \mathbb{R}^m$.

רציפות **הגדרה.** נאמר כי f **רציפה** בנקודה $x_0 \in D$ אם לכל $\varepsilon > 0$ קיים $\delta > 0$ כך שמתקיים $f(B(x_0, \delta) \cap D) \subseteq B(f(x_0), \varepsilon)$.[12] (ההגדרה אינה תלויית נורמה.)

נאמר ש־f רציפה על D אם היא רציפה בכל נקודה של D.

טענה 24: f רציפה ב־x_0 אם"ם $\{f_i(x)\}_{i=1}^m$ רציפות ב־x_0 כפונקציות ממשיות.

הוכחה. ניקח ב־$\mathbb{R}^m$ את הנורמה $\|y\|_\infty = \max_{1 \le i \le m} |y_i|$. אם f רציפה ב־x_0, מתקיים $f(B(x_0, \delta) \cap D) \subseteq \{y \mid \max_{1 \le i \le m} |y_i - f_i(x_0)| < \varepsilon\}$, כלומר אם $x \in B(x_0, \delta) \cap D$ אז $|f_i(x) - f_i(x_0)| < \varepsilon$ לכל $i = 1, \dots, m$.

טיעון זה עובד בשני הכיוונים, כי אם כל הפונקציות $f_i(x)$ רציפות ב־x_0, אזי עבור $\varepsilon > 0$ קיים $\delta_i > 0$ כך שאם $x \in B(x_0, \delta_i) \cap D$ אז $|f_i(x) - f_i(x_0)| < \varepsilon$. נבחר $\delta = \min_{1 \le i \le m} \delta_i$, ואז עבור $x \in B(x_0, \delta) \cap D$ נקבל $|f_i(x) - f_i(x_0)| < \varepsilon$ לכל $1 \le i \le m$. במילים אחרות, $f(B(x_0, \delta) \cap D) \subseteq B(f(x_0), \varepsilon)$ (כדור ב־$\|\cdot\|_\infty$).

טענה 25 (רציפות לפי סדרות): $f : D \to \mathbb{R}^m$ רציפה בנקודה $x_0 \in D$ אם"ם לכל סדרה $D \supseteq \{x^k\} \to x_0$, $f(x^k) \to f(x_0)$.

הוכחה. f רציפה ב־x_0 אם"ם $f_i(x)$ רציפות ב־x_0 אם"ם $f_i(x^k) \to f(x_0)$ לכל $1 \le i \le m$ אם"ם $f(x^k) \to f(x_0)$.

משפט 26: תהי $C \subseteq \mathbb{R}^n$ קומפקטית ו־$f : C \to \mathbb{R}^m$ רציפה. אזי $f(C) \subseteq \mathbb{R}^m$ קומפקטית. במילים אחרות, תמונה רציפה של קבוצה קומפקטית היא קומפקטית.

הוכחה. כדי להראות ש־$f(C)$ קומפקטית, מספיק להראות כי לכל סדרה ב־$f(C)$ ישנה תת־סדרה המתכנסת בקבוצה. ניקח סדרה כזאת, $\{f(x^k)\}_{k=1}^{\infty}$ כאשר $\{x^k\} \subseteq C$. קיימת תת־סדרה $\{x^{k_j}\}_{j=1}^{\infty}$ המתכנסת ב־C (הקומפקטית): $x^{k_j} \to \overline{x} \in C$. אבל f רציפה, ולכן מתקבל $f(x^{k_j}) \to f(\overline{x}) \in f(C)$. לכן מצאנו תת־סדרה מתכנסת ב־$f(C)$, כנדרש.

[11] C_r, כזכור, קומפקטית.

[12] מבינים היכן הכדור לפי הקונטקסט.

דוגמה. $C \subseteq \mathbb{R}^n$ קומפקטית ו־$f : C \to \mathbb{R}$ רציפה ($m = 1$). אז $f(C) \subseteq \mathbb{R}$ קומפקטית. מכאן, קבוצה קומפקטית ממשית מכילה את הנקודה העליונה שלה ואת הנקודה התחתונה שלה.

משפט 27: תהי $C \subseteq \mathbb{R}^n$ קשירה ו־$f : C \to \mathbb{R}^m$ רציפה. אזי $f(C)$ קשירה. במילים אחרות, תמונה רציפה של קבוצה קשירה היא קשירה.

הוכחה. בשלילה, נניח ש־$f(C)$ לא קשירה. אז ניתן למצוא קבוצות פתוחות $\tilde{U}, \tilde{V} \subseteq \mathbb{R}^m$ כך ש־(א) $\tilde{U} \cap f(C)$ זרה ל־$\tilde{V} \cap f(c)$; (ב) $(\tilde{U} \cap f(C)) \cup (\tilde{V} \cap f(C)) = f(C)$; (ג) $\tilde{U} \cap f(C) \neq \varnothing$, $\tilde{V} \cap f(C) \neq \varnothing$.

תהי $x_0 \in C$ נקודה כך ש־$f(x_0) \in \tilde{U}$. קיים $\varepsilon > 0$ כך ש־$B(f(x_0), \varepsilon) \subseteq \tilde{U}$. בגלל רציפות f ב־x_0 קיים $\delta > 0$ כך ש־$f(B(x_0, \delta) \cap C) \subseteq B(f(x_0), \varepsilon) \subseteq \tilde{U} \cap f(C)$. זאת אומרת, אם $f(x_0) \in \tilde{U}$ (ולכן בהכרח לא ב־$\tilde{V}$), אזי כדור שלם סביב x_0 ($\cap\, C$) נשלח על־ידי f ל־$\tilde{U}$.

מסקנה: קבוצת הנקודות ב־C הנשלחות על־ידי f ל־$\tilde{U}$ היא $U \cap C$, כאשר U פתוחה (במילים אחרות, היא איחוד כל הכדורים $B(x_0, \delta)$ כמו לעיל).

באותו אופן, קבוצת כל הנקודות ב־C הנשלחות על־ידי f ל־$\tilde{V}$ היא $V \cap C$, כאשר V פתוחה (ב־$\mathbb{R}^n$).

$U \cap C$ זרה ל־$V \cap C$, כי התמונות זרות ($\tilde{U} \cap C$ זרה ל־$\tilde{V} \cap C$); $(U \cap C) \cup (V \cap C) = C$, כי כל נקודה $y \in C$ מקיימת $f(y) \in \tilde{U}$ או $f(y) \in \tilde{V}$ (כי $f(C) \subseteq \tilde{U} \cup \tilde{V}$); $U \cap C \neq \varnothing$ כי קיימת נקודה $y \in C$ כך ש־$f(y) \in \tilde{U}$, ו־$V \cap C \neq \varnothing$ מאותה סיבה. אז C לא קשירה, בסתירה להנחה.

12.2.2008 תמונה רציפה של קבוצה קשירה היא קשירה (לדוגמה, עקומת Peano – פונקציה מהישר למרובע). קבוצה קשירה על הישר היא בהכרח קטע (אם $C \subseteq \mathbb{R}$ קשירה, כל הנקודות בין שתי נקודות נתונות ב־C חייבות אף הן להיות ב־C). מכך נובע שתמונה רציפה של קטע $[a, b]$ היא קטע: כלומר, אם $f(a) \leq \xi \leq f(b)$, אז יש $c \in [a, b]$ כך ש־$f(c) = \xi$. (זהו משפט ערך הביניים.)

דוגמה. נניח כי $D \subseteq \mathbb{R}^n$ (פתוחה ו)קשירה ו־$f : D \to \mathbb{R}$ רציפה. נניח שעבור $x_0, x_1 \in D$ מתקיים $f(x_0) < 0 < f(x_1)$. אם קיימת $\gamma(t) : [a, b] \to D$ רציפה כך ש־$\gamma(b) = x_1$ ו־$\gamma(a) = x_0$, אזי f מתאפסת בנקודה כלשהי על $\gamma(t)$ (כי $f(\gamma(t))$ רציפה).[13]

טענה 28: נניח כי $f : D_1 \to \mathbb{R}^m$ ($D_1 \subseteq \mathbb{R}^n$) רציפה ב־$x_0 \in D_1$ ו־$g : D_2 \to \mathbb{R}^k$ ($D_2 \subseteq \mathbb{R}^m$) כך ש־$f(B(x_0, \delta) \cap D_1) \subseteq D_2$ עבור $\delta > 0$ מסויים. נניח עוד כי g רציפה ב־$f(x_0)$. אזי $g \circ f$ רציפה ב־x_0.

הוכחה. צריך להראות שאם $\mathbb{R}^k \ni y_0 = g \circ f(x_0)$ אזי לכל $\varepsilon > 0$ ניתן למצוא $\eta > 0$ כך שמתקיים $(g \circ f)(B(x_0, \eta) \cap D_1) \subseteq B(y_0, \varepsilon)$.

רציפות g ב־$f(x_0)$ אומרת כי קיים $\theta > 0$ כך ש־$g(B(f(x_0), \theta) \cap D_2) \subseteq B(y_0, \varepsilon)$.

[13] יכולנו גם לומר זאת בלי דיון במסילות.

לפי הנתון, קיים $\delta > 0$ כך ש־$f(B(x_0,\delta)\cap D_1) \subseteq B(f(x_0),\theta)$, ועבור $\delta > 0$ מספיק קטן, אם $f(B(x_0,\delta)\cap D_1) \subseteq D_2$ אז $f(B(x_0,\delta)\cap D_1) \subseteq B(f(x_0),\theta)\cap D_2$.
בסך הכול נקבל $g(f(B(x_0,\delta)\cap D_1)) \subseteq g(B(f(x_0),\theta)\cap D_2) \subseteq B(y_0,\varepsilon)$.

2.3.1 פונקציות לינאריות

פונקציה לינארית

הגדרה. פונקציה היא **לינארית** אם ניתן לייצג אותה כ־$f(x) = Ax \in \mathbb{R}^m$, כאשר $A \in M_{m\times n}$ (מטריצה בת m שורות ו־n עמודות).

ברור שפונקציה לינארית רציפה בכל נקודה (ניתן להראות זאת לפי סדרות או לפי כך שכל רכיב הוא פונקציה לינארית ולכן רציפה).

טענה 29: תהי $\|\cdot\|_1$ נורמה על $\mathbb{R}^n$ ו־$\|\cdot\|_2$ נורמה על $\mathbb{R}^m$. אזי קיים
$$\infty > \|A\|_{1,2} \overset{\text{def}}{=} \sup_{\|x\|_1=1} \|Ax\|_2 = \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_1}$$
וה־sup הוא max.

הוכחה. Ax היא פונקציה רציפה ולכן $\|Ax\|_2$ היא פונקציה רציפה על $x \in \mathbb{R}^n$ (כי כל נורמה היא פונקציה רציפה). מצד שני, ספירת היחידה לפי $\|\cdot\|_1$ – $\|x\|_1 = 1$ – היא קומפקטית. לכן לפונקציה הרציפה $\|Ax\|_2$ יש מקסימום על $\|x\|_1 = 1$: כלומר, קיימת נקודה $\|x_0\|_1 = 1$ כך ש־$\|Ax_0\|_2 \geq \|Ax\|_2 \quad \forall \|x\|_1 = 1$. את הגודל $\|Ax_0\|_2$ נסמן ב־$\|A\|_{1,2}$.
יהי $x \neq 0$ כלשהו. אזי $\frac{x}{\|x\|_1}$ שייך לספירת היחידה לפי $\|\cdot\|_1$. נקבל
$$\frac{1}{\|x\|_1}\|Ax\|_2 = \|A(\frac{x}{\|x\|_1})\|_2 \leq \|A\|_{1,2}$$
לכן $\|A\|_{1,2} = \max_{x\neq 0}\frac{\|Ax\|_2}{\|x\|_1}$.

הערות:

1. המספר $\|A\|_{1,2}$ "מודד" את מידת העיוות של תלמונת ספירת היחידה של $\|\cdot\|_1$: הוא נותן את הנורמה המקסימלית של תמונת נקודה על ספירת היחידה בטרנספורמציה מ־$\|\cdot\|_1$ ל־$\|\cdot\|_2$.
2. $\|A\|_{1,2}$ הוא בעצם נורמה על $M_{m\times n}$ כי מתקיימות התכונות הנדרשות: $\|A\|_{1,2} = 0$ אס"ם $A = 0$ כי $\|A\|_{1,2} = 0$ אס"ם $Ax = 0$ לכל $x \in \mathbb{R}^n$; קיום אי־שוויון המשולש: כתרגיל.

נורמה אופרטורית

נורמה זו נקראת **הנורמה האופרטורית** של A.

3. הנורמה $\|A\|_{1,2}$ היא גם נורמה על $\mathbb{R}^{mn}$, כי המרחב הוקטורי $M_{m\times n}$ איזומורפי ל־$\mathbb{R}^{mn}$.
4. בהינתן נורמה $\|\cdot\|$ על $M_{m\times n} \cong \mathbb{R}^{mn}$, קיימות נורמות $\|\cdot\|_1$ על $\mathbb{R}^n$ ו־$\|\cdot\|_2$ על $\mathbb{R}^m$ כך ש־$\|A\| = \|A\|_{1,2} \quad \forall A \in M_{m\times n}$. (הוכחה ב־Summary 3 באתר הקורס.)
5. ניקח $m = n$. נחפש נורמה $\|\cdot\|_1$ כך ש־$\|A\| = \|A\|_{1,1}$, כלומר נתונה נורמה $\|\cdot\|$ על $\mathbb{R}^n$ ומחפשים נורמה $\|\cdot\|_1$ על $\mathbb{R}^n$ כך ש־$\|A\| = \max_{\|x\|_1=1}\|Ax\|_1$ עבור כל $A \in M_{m\times n}$.
נניח של־A ערך עצמי ממשי λ ווקטור עצמי $x \neq 0$ כך ש־$\|x\|_1 = 1$. אז $\|Ax\|_1 =$

$|\lambda| = |\lambda| \|x\|_1$ ומכאן $|\lambda| \geq \|A\|_{1,1}$. לכן, בכל נורמה אופרטורית, $\|A\|_{1,1} \leq |\lambda|$ לכל λ ערך עצמי ממשי של A. בפרט, $\|I\|_{1,1} \geq 1$.

3 דיפרנציאביליות

3.1 הגדרה

הגדרה. תהי $f : B(x_0,\delta)\setminus\{x_0\}\to\mathbb{R}^m$ כאשר $\delta>0$ מסויים ו־$B(x_0,\delta)\subseteq\mathbb{R}^n$. נאמר כי $\mathbb{R}^m\ni\xi=\lim_{x\to x_0}f(x)$ אם לכל $\varepsilon>0$ קיים $\eta>0$ כך ש־$f(B(x_0,\eta)\setminus\{x_0\})\subseteq B(\xi,\varepsilon)$ ($\eta<\delta$). בפרט, f **רציפה** ב־x_0 אם״ם $\lim_{x\to x_0}f(x)=f(x_0)$.

גבול של פונקציה בנקודה

רציפות פונקציה

הגדרה. תהי $h : B(x_0,\delta)\setminus\{x_0\}\to\mathbb{R}^m$. נאמר כי h **מתאפסת ב־**x_0 **מהר יותר מ־**$|x-x_0|$ אם $\lim_{x_0\neq x\to x_0}\frac{h(x)}{|x-x_0|}=0\in\mathbb{R}^m$. מסמנים $h(x)=o(|x-x_0|)$.

תכונת הקירוב הלינארי

הגדרה. תהי $f : B(x_0,\delta)\to\mathbb{R}^m$. נאמר כי ל־$f$ **תכונת הקירוב הלינארי** ב־x_0 אם קיימת העתקה לינארית $T:\mathbb{R}^n\to\mathbb{R}^m$ כך ש־$f(x)-f(x_0)=T(x-x_0)+o(|x-x_0|)$; במילים אחרות: $\lim_{x\to x_0}\frac{f(x)-f(x_0)-T(x-x_0)}{|x-x_0|}=0\in\mathbb{R}^m$.

דיפרנציאביליות

13.2.2008

אומרים ש־f **דיפרנציאבילית** ב־x_0.

טענה 30: אם f דיפרניצאבילית ב־x_0, אזי T נקבע באופן יחיד.

הוכחה. נניח כי

$$\begin{aligned} f(x)-f(x_0) &= T_1(x-x_0)+o(|x-x_0|) \\ f(x)-f(x_0) &= T_2(x-x_0)+o(|x-x_0|) \end{aligned}$$

צריך להראות $T_1=T_2$. על־ידי חיסור, $(T_1-T_2)(x-x_0)=o(|x-x_0|)$. ניקח $x=y+x_0$. $(T_1-T_2)(y)=o(|y|)$. ניקח את הנורמה האוקלידית ונקבל $\frac{|(T_1-T_2)(y)|}{|y|}\xrightarrow{|y|\to 0}0$. מליניאריות, $|(T_1-T_2)(\frac{y}{|y|})|=\frac{|(T_1-T_2)(y)|}{|y|}$, לכן הנורמה האופרטורית $\|T_1-T_2\|=0$ (כי $(T_1-T_2)(z)=0\quad\forall|z|=1$. או: לכל $\lambda>0$, $|(T_1-T_2)(z)|=|\frac{(T_1-T_2)(\lambda z)}{\lambda}|\to 0$).
לכן $T_1=T_2$.

דיפרנציאל

הגדרה. הטרנספורמציה T נקראת **הדיפרנציאל** של f ב־x_0. מסמנים $Df(X_0)(\in M_{m\times n})$.

טענה 31: אם $f,g : B(x_0,\delta)\to\mathbb{R}^n$ דיפרנציאביליות ב־x_0, אזי כך גם $f+g$ ומתקיים $D(f+g)(x_0)=Df(x_0)+Dg(x_0)$.

הוכחה.
$$\begin{aligned} f(x)-f(x_0) &= Df(x_0)(x-x_0)+o(|x-x_0|) \\ g(x)-g(x_0) &= Dg(x_0)(x-x_0)+o(|x-x_0|) \end{aligned}$$

אז $(f+g)(x)-(f+g)(x_0)=(Df(x_0)+Dg(x_0))(x-x_0)+o(|x-x_0|)$. הראינו קיום טרנספורמציה כנדרש, ומהטענה הקודמת שמבטיחה יחידות, מתקבל הדרוש.

טענה 32: אם f כנ״ל דיפרנציאבילית ב־x_0, אזי היא רציפה ב־x_0.

הוכחה. $|f(x)-f(x_0)|\le\|Df(x_0)\||x-x_0|+c|x-x_0|$ (כי $\|A\|=\max_{x\neq 0}\frac{|Ax|}{|x|}$ ולכן $|Ax|\le\|A\||x|\quad\forall x$, וכי $\frac{R_{x_0}(x)}{|x-x_0|}\to 0$ גורר $|R_{x_0}(x)|\le\varepsilon|x-x_0|$).

נתרכז במקרה $m=1$, כלומר $f:B(x_0,\delta)\to\mathbb{R}$.

הלמה של Riesz: תהי $L:\mathbb{R}^n\to\mathbb{R}$ העתקה לינארית; אזי קיים וקטור $l\in\mathbb{R}^n$ כך ש־$Ly=\langle l,y\rangle$ לכל $y\in\mathbb{R}^n$.

לפי למה זו, נוכל לכתוב $f(x)-f(x_0)=\langle Df(x_0),x-x_0\rangle+o(|x-x_0|)$.

הגדרה. הווקטור $Df(x_0)$ נקרא **הגרדיאנט** של f ב־x_0, ומסומן $\nabla f(x_0)$. גרדיאנט

טענה 33: $f:B(x_0,\delta)\to\mathbb{R}$ דיפרנציאבילית ב־x_0; אזי $Df(x_0)y=\sum_{i=1}^n a_iy_i$, כאשר $y=\sum_{i=1}^n y_ie_i$ לגבי הבסיס הסטנדרטי ו־$a_i=\frac{\partial f}{\partial x_i}(x_0)=\lim_{h\to0}\frac{f(x_0+he_i)-f(x_0)}{h}$.[14]

הוכחה. היות שנתון כי f דיפרנציאבילית ב־x_0, הרי שאם נבחר בסיס $\{e_i\}_{i=1}^n$, נקבל עבור $y\in\mathbb{R}^n$ כי $Df(x_0)(y)=\sum_{i=1}^n a_iy_i$ כאשר $y=\sum_{i=1}^n y_ie_i$, $a_1,\dots,a_n$ קבועים ונקבעים באופן חד־ערכי על־ידי $Df(x_0)$. נמצא את a_i: נבחר $y=he_j$ עבור j מסויים, ואז עבור $x=x_0+y$,

$$\begin{aligned} f(x)-f(x_0) &= Df(x_0)(y)+o(|x-x_0|)\\ &= a_jy_j+o(|x-x_0|)\\ &= a_jh+o(|h|)\end{aligned}$$

אם כן, $\frac{f(x)-f(x_0)}{h}=\frac{f(x_0+he_j)-f(x_0)}{h}=a_j+\frac{o(|h|)}{h}\xrightarrow{h\to0}a_j$.

מכאן, $f(x)-f(x_0)=\sum_{i=1}^n\frac{\partial f}{\partial x_i}(x_0)\cdot(x-x_0)_i+o(|x-x_0|)$ אם f דיפרנציאבילית ב־x_0. הנגזרת החלקית תלויית־בסיס; למעשה, $\frac{\partial f}{\partial x_i}(x_i)$ הוא השיפוע של הפונקציה במשתנה יחיד $t\mapsto f(x_0+te_i)$.

דוגמה. $g(x,y)=\frac{xy}{x^2+y^2}$ עבור $(x,y)\neq(0,0)$, $g(0,0)=0$. הנגזרות החלקיות קיימות ב־0, אבל הפונקציה אינה רציפה שם.

טענה 34: תהי $f:B(x_0,\delta)\to\mathbb{R}^m = (f_1,\dots,f_m)$. אזי f דיפרנציאבילית ב־x_0 אם"ם $f_1,\dots,f_m$ דיפרנציאביליות ב־x_0.

הוכחה. ($\Leftarrow$) נניח כי f דיפרנציאבילית ב־x_0, כלומר

$$f(x)-f(x_0)=Df(x_0)(x-x_0)+o(|x-x_0|)$$

(ב־$\mathbb{R}^m$). ניקח רכיב i:

$$f_i(x)-f_i(x_0)=\left[\text{השורה ה־}i\text{ של }Df(x_0)\right](x-x_0)+o(|x-x_0|)$$

(ב־$\mathbb{R}^n$). אז לפי הגדרה, f_i דיפרנציאבילית ב־x_0.

($\Rightarrow$) נניח כי f_i ($i=1,\dots,m$) דיפרנציאביליות ב־x_0.

$$f_i(x)-f_i(x_0)=L_i(x-x_0)+o(|x-x_0|)$$

כאשר L_i שורה באורך n: $L_i=(\frac{\partial f_1}{\partial x_1},\dots,\frac{\partial f_n}{\partial x_n})$. אז

$$f(x)-f(x_0)=L(x-x_0)+o(|x-x_0|)$$

כאשר L_i היא השורה ה־i של $L=Df(x_0)\in M_{m\times n}$.

[14] טענה זו מגדירה $\nabla f(x_0)=(a_1,\dots,a_n)$.

מסקנה 35: בהינתן בסיס $\{e_1,\dots,e_m\} \subseteq \mathbb{R}^m$ ובסיס $\{e'_1,\dots,e'_n\} \subseteq \mathbb{R}^n$, נקבל, כאשר $f : B(x_0,\delta) \to \mathbb{R}^m$ דיפרנציאבילית ב־x_0,

$$Df(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \dots & \frac{\partial f_m}{\partial x_1}(x_0) \end{pmatrix} = J(x_0)$$

זהו **היעקוביאן** של f ב־x_0. יעקוביאן

משפט 36 (כלל השרשרת): $f : B(x_0,\delta_1) \to \mathbb{R}^m$ דיפרנציאבילית ב־x_0. נסמן $y_0 = f(x_0)$. תהי $g : B(y_0,\delta_2) \to \mathbb{R}^k$ דיפרנציאבילית ב־y_0. אזי $g \circ f$ מוגדרת ב־$B(x_0,\delta_3)$ (עבור $0 < \delta_3 < \delta_1$ כלשהו) ודיפרנציאבילית ב־x_0. בנוסף, $D(g \circ f)(x_0) = Dg(y_0) \circ Df(x_0)$.[15]

הוכחה. היות ש־f דיפרנציאבילית ב־x_0, היא רציפה שם; אז עבור $0 < \delta_3 < \delta_1$, מתקיים $f(B(x_0,\delta_3)) \subseteq B(y_0,\delta_2)$, כלומר באמת $g \circ f$ מוגדרת ב־$B(x_0,\delta_3)$.

$$\begin{aligned} f(x) - f(x_0) &= Df(x_0)(x-x_0) + R_{x_0}(x) \\ g(y) - g(y_0) &= Dg(y_0)(y-y_0) + R_{y_0}(y) \end{aligned}$$

עבור $R_{x_0}(x) = o(|x-x_0|)$ ו־$R_{y_0}(y) = o(|y-y_0|)$. נציב $y = f(x)$, $y_0 = f(x_0)$:

$$g(f(x)) - g(f(x_0)) = Dg(y_0)[Df(x_0)(x-x_0) + R_{x_0}(x)] + R_{y_0}(y)$$

20.2.2008 נותר להראות כי $Dg(y_0)R_{x_0}(x) + R_{y_0}(y) = o(|x-x_0|)$. כזכור, $h(x) = o(|x-x_0|)$ אם"ם $\forall\varepsilon > 0 \quad \exists\delta > 0 : |x-x_0| < \delta \Longrightarrow |h(x)| < \varepsilon|x-x_0|$ וכמו כן, $|Ax| \le \|A\||x|$. אז עבור $\varepsilon > 0$, כאשר $|x-x_0| < \delta$,

$$\begin{aligned} |Dg(y_0)R_{x_0}(x)| &\le \|Dg(y_0)\||R_{x_0}(x)| \\ &\le \|Dg(y_0)\| \cdot \varepsilon|x-x_0| \end{aligned}$$

כיוון ש־f רציפה ב־x_0, קיים $\eta > 0$ כך שאם $|x-x_0| < \eta$ אז $|f(x)-f(x_0)| < \delta_2$. אז כאשר $|x-x_0| < \min(\delta,\eta)$,

$$\begin{aligned} |R_{y_0}(y)| &< \varepsilon|f(x)-f(x_0)| \\ &< \varepsilon[|Df(x_0)(x-x_0)| + |R_{x_0}(x)|] \\ &< \varepsilon[\|Df(x_0)\||x-x_0| + \varepsilon|x-x_0| \\ &< \varepsilon \cdot C|x-x_0| \end{aligned}$$

הראינו ש־$(g \circ f)(x) - (g \circ f)(x_0) = Dg(f(x_0)) \circ Df(x_0) + o(|x-x_0|)$, כלומר ל־$g \circ f$ תכונת הקירוב הלינארי ב־x_0 ו־$D(g \circ f)(x_0) = Dg(f(x_0)) \circ Df(x_0)$. לכן בהינתן קואורדינטות ב־$\mathbb{R}^n$, $\mathbb{R}^m$, $\mathbb{R}^k$, מתקיים $J(g \circ f)(x_0) = Jg(f(x_0))Jf(x_0)$. כלומר, $\frac{\partial}{\partial x_1}(g \circ f)_1 = \sum_{j=1}^m \frac{\partial g_1}{\partial y_j}(f(x_0)) \cdot \frac{\partial f_j}{\partial x_1}(x_0)$.

[15] הפעולה היא הרכבת העתקות לינאריות.

3.2 נגזרות כיווניות

תהי $f : B(x_0,\delta) \to \mathbb{R}^m$ דיפרנציאבילית ב־x_0 ותהי $\gamma : (a,b) \to B(x_0,\delta)$ מסילה כך ש־$x_0 = \gamma(c)$. נסמן $\gamma(t) = (\gamma_1(t),\dots,\gamma_n(t))$. דיפרנציאביליות פירושה קיום $\gamma_i'(t)$, וזו השורה ה־i של $D\gamma(t)$. נתבונן ב־$f(\gamma(t)) : (a,b) \to \mathbb{R}^m$: לפי כלל השרשרת, מתקיים $D(f(\gamma(t)))\mid_{t=c} = Df(x_0)\circ D\gamma(c)$ וכן $\frac{df_1(\gamma(t))}{dt}\mid_{t=c} = \sum_{i=1}^n \frac{\partial f_1}{\partial x_i}(x_0)\cdot\gamma_i'(t)\mid_{t=c}$.

הגדרה. הנגזרת $\frac{d}{dt}f(\gamma(t))$ נקראת **הנגזרת של f לאורך המסילה** $\gamma(t)$.

מקרה פרטי מתקבל כאשר $\gamma(t) = l(t) = x_0 + (t-c)\vec{u}$, עבור $\vec{u} \in S^{n-1}$ (כלומר, $\vec{u}$ וקטור בספירת היחידה [האוקלידית] ב־$\mathbb{R}^n$). אם נסמן $\vec{u} = (u_1,\dots,u_n)$, לכל $1 \le j \le m$ נקבל $\frac{d}{dt}f_j(l(t))\mid_{t=c} = \sum_{i=1}^n \quad (x_0)u_i$. מצד שני, לפי הגדרת הנגזרת נקבל שנגזרת זו היא $\frac{d}{dt}f_j(l(t))\mid_{t=c} = \lim_{h\to 0}\frac{f_j(l(c+h))-f_j(l(c))}{h} = \lim_{h\to 0}\frac{f_j(x_0+h\vec{u})-f_j(x_0)}{h}$.

הגדרה. נגזרת זו נקראת **הנגזרת הכיוונית של f** (או של f_j) **בכיוון $\vec{u}$ בנקודה** x_0. נגזרת כיוונית

כאשר $m=1$, $f : B(x_0,\delta)\to\mathbb{R}$ פונקציה ממשית דיפרנציאבילית ב־x_0, הנגזרת הכיוונית בכיוון $\vec{u}$ של f בנקודה x_0 היא $\nabla f(x_0)\cdot\vec{u} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0)u_i$. מסמנים גם $\frac{\partial f}{\partial l_i}(x_0) = \frac{\partial f}{\partial x_i}(x_0) = \frac{\partial f}{\partial u}(x_0) = D_u f$.

מסקנה 37: כאשר f דיפרנציאבילית ב־x_0, ידיעת כל $\frac{\partial f}{\partial x_i}(x_0)$ עבור $1 \le i \le n$ $\Longleftarrow$ ידיעת כל הנגזרות הכיווניות.

אם f דיפרנציאבילית ב־x_0, אזי קיימות הנגזרות החלקיות ב־x_0.

משפט 38: $B(x_0,\delta)\subseteq\mathbb{R}^n$. אם $f : B(x_0,\delta)\to\mathbb{R}$ בעלת נגזרות חלקיות בכדור. אם הנגזרות החלקיות רציפות ב־x_0, אזי f דיפרנציאבילית ב־x_0.

הוכחה. נסמן $x = x_0 + h_1e_1 + \ldots + h_ne_n$ ($e_1,\dots,e_n$ בסיס סטנדרטי ו־$x \in B(x_0,\delta)$).
אז

$$\begin{aligned} f(x)-f(x_0) &= f(x) - f(x_0+h_1e_1+\ldots+h_{n-1}e_{n-1}) \\ &\quad + f(x_0+\textstyle\sum_{i=1}^{n-1}h_ie_i) - f(x_0+\sum_{i=1}^{n-2}h_ie_i) \\ &\quad \vdots \\ &\quad + f(x_0+h_1e_1) - f(x_0) \\ &= \frac{\partial f}{\partial x_n}(y_n)h_n + \ldots + \frac{\partial f}{\partial x_1}(y_1)h_1 \end{aligned}$$

כאשר y_j היא נקודה על הקטע המחבר את הנקודה $x_0 + h_1e_1 + \ldots + h_je_j$ אל הנקודה $x_0 + h_1e_1 + \ldots + h_{j-1}e_{j-1}$ (לפי משפט ערך הביניים בכל קואורדינטה). אז

$$f(x)-f(x_0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0)h_i + \sum_{i=1}^n\left(\frac{\partial f}{\partial x_i}(y_i) - \frac{\partial f}{\partial x_i}(x_0)\right)h_i$$

לכל $\varepsilon > 0$, רציפות $\frac{\partial f}{\partial x_i}$ ב־x_0 גוררת קיום $\eta > 0$ כך שאם $|h| < \eta$, $|\frac{\partial f}{\partial x_i}(y_i) - \frac{\partial f}{\partial x_i}(x_0)| < \varepsilon$, שהרי $|y_i - x_0| < |h|$. אז $|R_{x_0}(x)| = |\sum_{i=1}^n(\frac{\partial f}{\partial x_i}(y_i) - \frac{\partial f}{\partial x_i}(x_0))h_i| \le \varepsilon|h|$ ולכן $R_{x_0}(x) = o(|x-x_0|)$. נובע של־f תכונת הקירוב הלינארי ב־x_0.

נקודה קריטית **הגדרה.** אם ב־x_0 מתקיים $\nabla f(x_0)=0$, נאמר כי x_0 היא **נקודה קריטית** של f.

מקסימום לוקאלי **הגדרה.** $f:D\to\mathbb{R}$, $D\subseteq\mathbb{R}^n$. נאמר כי ל־f **מקסימום לוקאלי** ב־$x\in D$ אם קיים כדור 26.2.2008
$B(x,r)$, $r>0$ כך שעבור $y\in B(x,r)\cap D$ מתקיים $f(y)\le f(x)$.
מקסימום גלובאלי מתקיים אם $\forall y\in D \quad f(y)\le f(x)$.

משפט 39: תהי f דיפרנציאבילית ב־x_0 ובעלת אקסטרמום לוקאלי שם. אזי $\nabla f(x_0)=0$.[16]

הוכחה. בפרט, x_0 היא אקסטרמום לוקאלי של הפונקציה במשתנה יחיד $t\mapsto f(x_0+te_i)$ (כאשר e_i וקטור היחידה בכיוון x_i). אזי ב־$t=0$ הנגזרת מתאפסת, ולכן $\frac{\partial f}{\partial x_i}(x_0)=0$.

מההוכחה נקבל שכאשר f דיפרנציאבילית ב־x_0 נקודת אקסטרמום לוקאלי, כל הנגזרות הכיווניות מתאפסות; אבל אם ידוע רק שהנגזרות החלקיות קיימות, מההוכחה ברור שהן מתאפסות, אולם אפשר שהנגזרות הכיווניות האחרות אינן קיימות כלל.

3.3 על־משטחים ומשטחי גובה

על־משטח **הגדרה.** נתונה קבוצה $S\subseteq B(x_0,\delta)\subseteq\mathbb{R}^n$ כך ש־$x_0\in S$. נאמר כי S היא **על־משטח** אם לכל נקודה $x\in S$ קיים כדור פתוח $B(x,\eta)\subseteq B(x_0,\delta)$ כך ש־$S\cap B(x,\eta)$ הוא גרף של פונקציה $x_j=f(x_1,\dots,x_{j-1},x_{j+1},x_n)$.

דוגמה. $x^2+y^2+z^2=1$ לא מגדיר גרף יחיד, אבל בסביבת כל נקודה המשטח הוא גרף.

דוגמה. כל גרף הוא על־משטח ב־$\mathbb{R}^{n+1}$; כלומר, אם $y=f(x_1,\dots,x_n)$ אזי ניתן לייצג $\mathbb{R}^{n+1}\supseteq S=\{(x,y)\in\mathbb{R}^{n+1}\mid x\in D,\mathbb{R}\ni y=f(x)\}$.

הגדרה. תהי $\varphi:D\subseteq\mathbb{R}^{n+1}\to\mathbb{R}$ ותהי $x_0\in D$ נקודה שבה $\varphi(x_0)=0$. נאמר כי המשוואה
משטח גובה $\varphi=0$ מגדירה **משטח גובה** בסביבה של x_0 אם קיים כדור $B(x_0,\delta)$ כך ש־$B(x_0,\delta)\cap D$ קבוצת הנקודות המקיימות $\varphi(x)=0$ היא על־משטח.

דוגמה. נסתכל ב־$x^2+y^2=1$, $\varphi(x,y)=x^2+y^2-1$. אם $\overline{x}^2+\overline{y}^2=1$ אזי קיים כדור (למעשה, עיגול) סביב $(\overline{x},\overline{y})$ שבו המשוואה נפתרת באופן חד־ערכי ל־x או ל־y: למשל, $y=\pm\sqrt{1-x^2}$ מתארת את y כפונקציה של x בסביבת כל נקודה פרט לשתי נקודות האפס, שם נוכל באופן דומה לכתוב את x כפונקציה של y בסביבה.

דוגמה. $f(x,y)=\frac{xy}{x^2+y^2}$ (ונגדיר $f(0,0)=0$). האם אפשר לכתוב את y כפונקציה של x או את x כפונקציה של y בסביבת $(0,0)$? לא, כי קבוצת האפסים היא שני הצירים, ולכן אפילו בעיגול קטן סביב $(0,0)$ היא איננה גרף.

[16] זהו תנאי הכרחי לקיום אקסטרמום לוקאלי בנקודה פנימית, כאשר f דיפרנציאבילית שם.

דוגמה. לחשב את המקסימום של הפונקציה $f(x,y)=x^2+17y^2$ תחת האילוץ $x^2+y^2=1$ (מעגל הוא קבוצה קומפקטית ו־f רציפה בסביבתו, לכן קיים מקסימום כזה). גיאומטרית, נסתכל בקווי הגובה של f, האליפסות $x^2+17y^2=c$ עבור $c>0$. הערך הגבוה יתקבל כאשר קו הגובה ישיק לקו האילוץ. (בנקודת ההשקה הנורמלים מקבילים.)

מקסימום יתקבל כאשר יש משקל מקסימלי ל־y – בנקודה $x=0,y=1$, שאז $f(0,1)=17$ (וגם ב־$x=0,y=-1$).

נניח כי $f:B(x_0,\delta)\subseteq\mathbb{R}^{n+1}\to\mathbb{R}$ דיפרנציאבילית ב־x_0, $\varphi(x_0)=0$ ו־$\varphi=0$ הוא משטח גובה S ב־$B(x_0,\delta)$. כיוון ש־φ דיפרנציאבילית ב־x_0, מתקיים (כאשר $x,\nabla\varphi(x_0)\in\mathbb{R}^{n+1}$) $\varphi(x)-\varphi(x_0)=\nabla\varphi(x_0)(x-x_0)+o(|x-x_0|)$. בפרט, אם $x\in S$, מתקיים $\varphi(x)=0$ ואז $\nabla\varphi(x_0)(x-x_0)=o(|x-x_0|)$.

בפרט, אם $\gamma(t):(-\eta,eta)\to S$ היא מסילה דיפרנציאבילית על S כאשר $\gamma(0)=x_0$, אזי $\varphi(\gamma(t))\equiv 0$. במילים אחרות, $\frac{d}{dt}\varphi(\gamma(t))\mid_{t=0}=\nabla\varphi(x_0)\cdot\gamma(0)=0$. זאת אומרת, כל המשיקים הנמצאים ב־S מאונכים ל־$\nabla\varphi(x_0)$. כלומר, $\nabla\varphi(x_0)$ הוא **הנורמל** למשטח הגובה $\varphi=0$ בנקודה x_0.

דוגמה. $\varphi(x,y)=x^2+y^2$. נסתכל בקו הגובה $\varphi(x,y)=17$ כקו גובה ב־$\mathbb{R}^2$. נקבל $\nabla\varphi(x_0,y_0)=(2x_0,2y_0)\|(x_0,y_0)\perp S$.

27.2.2008 **הגדרה.** $\varphi:B(x_0,\delta)\to\mathbb{R}$ דיפרנציאבילית ב־x_0, ונניח כי הקבוצה $S=\{\varphi(x)=\varphi(x_0)\}$ היא משטח גובה. העל־מישור $\nabla\varphi(x_0)(x-x_0)=0$ ($x\in\mathbb{R}^{n+1}$) נקרא **המישור המשיק** ל־S ב־x_0. מישור משיק

דוגמה. ניקח את האליפסואיד $E=\{x^2+17y^2+7z^2=3\}$ ותהי $(x_0,y_0,z_0)\in E$. המישור המשיק ב־(x_0,y_0,z_0) ניתן על־ידי

$$\begin{aligned}(2x_0,34y_0,14z_0)(x-x_0,y-y_0,(z-z_0)) &= 0\\ 2x_0(x-x_0)+34y_0(y-y_0)+14z_0(z-z_0) &= 0\\ 2x_0x+34y_0y+14z_0z &= 2x_0^2+34y_0^2+14z_0^2=6\\ x_0x+17y_0y+7z_0z &= 3\end{aligned}$$

דוגמה. גרף $y=f(x)$ של פונקציה $f:B(x_0,\delta)\to\mathbb{R}$. אם f דיפרנציאבילית ב־x_0, הגרף הוא משטח הגובה ה־0 של $\varphi(x,y)=f(x)-y$. אז ב־$(x_0,y_0=f(x_0))\in\mathbb{R}^{n+1}$, $\nabla\varphi(x_0,y_0)=(\nabla f(x_0),-1)$ (הנורמל לגרף בנקודה). המישור המשיק הוא

$$\begin{aligned}(x-x_0,y-y_0)(\nabla f(x_0),-1) &= 0\\ (x-x_0)\nabla f(x_0)-(y-y_0) &= 0\\ \nabla f(x_0)(x-x_0) &= y-y_0\end{aligned}$$

(זהו גרף של מישור מעל $\mathbb{R}^n$).

הגרדיאנט מאונך למשטח הגובה; במילים אחרות, $\nabla\varphi$ ב־x_0 הוא כיוון העלייה המקסימלי של φ: $\frac{\partial\varphi}{\partial u} = \nabla\varphi(x_0)\vec{u}$ מקסימלי כאשר $\vec{u}\|\nabla\varphi(x_0)$, כלומר $\nabla\varphi(x_0) = \lambda\vec{u}$, $\lambda > 0$. אז ניקח $\vec{u} = \frac{\nabla\varphi(x_0)}{|\nabla\varphi(x_0)|} \in S^n \subseteq \mathbb{R}^{n+1}$ (כאשר $\nabla\varphi(x_0) \neq 0$).

3.4 הערכת הפרשים באמצעות הגרדיאנט

$D \subseteq \mathbb{R}^n$, f דיפרנציאבילית ב־D, $x, y \in D$ וכן גם $[x,y] \in D$ (הקטע הישר ביניהן). נסתכל בפונקציה $g(t) = f(tx+(1-t)y)$ עבור $t \in [0,1]$. לפי כלל השרשרת נקבל $g'(t) = \nabla f(tx-(1-t)y)(x-y)$. לפי משפט הערך הממוצע, $g(1)-g(0) = g'(t_0)\cdot(1-0)$, כלומר $f(x) - f(y) = \nabla f(t_0x+(1-t_0)y)(x-y)$, ומקושי־שוורץ נקבל

$$\begin{aligned} |f(x)-f(y)| &\leq |\nabla f(t_0x+(1-t_0)y)||x-y| \\ &\leq \sup_{z\in[x,y]} |\nabla f(z)||x-y| \end{aligned}$$

(כאשר יש דיפרנציאביליות ברציפות, זהו $\max$; אחרת, ה־$\sup$ יכול להיות אינסוף).

נעיר כי במימדים גבוהים יותר לא מתקיים משפט Rolle, וייתכן ש־$f(x) = f(y)$ אך ~~$\exists\nabla f(z) = 0$~~.

3.5 כופלי לגראנז׳: מקרה לא כללי

יהיו $f : B(x_0,\delta) \to \mathbb{R}$ ו־$\varphi : B(x_0,\delta) \to \mathbb{R}$ דיפרנציאביליות ב־x_0 כך ש־$\nabla\varphi(x_0) \neq 0$ ו־$S = \{\varphi(x) = \varphi(x_0)\}$ משטח גובה. נניח כי x_0 היא נקודה אקסטרימלית מקומית של f על S (ביחס לאילוץ $\varphi(x) = \varphi(x_0)$).

טענה 40 (תנאי הכרחי לאקסטרמום עם אילוץ): $\exists\mu \in \mathbb{R} : \nabla f(x_0) = \mu\nabla\varphi(x_0)$.

הוכחה. לשם ההוכחה, נניח כי על S קיימות מסילות $\gamma_1(t),\dots,\gamma_{n-1}(t)$[17] דיפרנציאביליות כך ש־$\gamma_1(0) = \dots = \gamma_{n-1}(t) = x_0$ ו־$\gamma_1'(0),\dots,\gamma_{n-1}'(0)$ וקטורים בלתי־תלויים ב־$\mathbb{R}^n$.[18]

נסתכל בפונקציות $g_i(t) = f(\gamma_i(t))$ עבור $i = 1,\dots,n-1$. אלה פונקציות גזירות ב־$t=0$ ובעלות אקסטרמום לוקאלי שם: $g_i'(t)\mid_{t=0} = \nabla f(x_0)\cdot\gamma_i'(t)\mid_{t=0} = 0$. אבל $\nabla\varphi(x_0)$ מאונך ל־S: $\nabla\varphi(x_0)\cdot\gamma_i'(t)\mid_{t=0} = 0$. לכן הן $\nabla f(x_0)$ והן $\nabla\varphi(x_0)$ הם וקטורים ממימד n המאונכים ל־$n-1$ הוקטורים הבלתי־תלויים $\gamma_i'(t)\mid_{t=0}$. כיוון שלפי ההנחה $\nabla\varphi(x_0) \neq 0$, חייב $\nabla f(x_0) = \mu\nabla\varphi(x_0)$.

הקבוע μ נקרא **כופל לגראנז׳**.

[17] אלו פונקציות $\gamma_i(t) : (-\eta,\eta) \to B(x_0,\delta) \subseteq \mathbb{R}^n$.

[18] קיום נובע ממשפט הפונקציות הסתומות.

4 משפט הפונקציות הסתומות

4.1 משפט הפונקציות ההפוכות

$f : D \to \mathbb{R}^m$, D קבוצה פתוחה וקשירה (תחום). נניח כי f דיפרנציאבילית ו־$Df(x)$ פונקציה רציפה מ־$x \in D$ לתוך מרחב ההעתקות הלינאריות מ־$\mathbb{R}^n$ ל־$\mathbb{R}^m$ ($\mathbb{R}^{nm} \cong \hom(\mathbb{R}^n, \mathbb{R}^m)$).

$Df(x)$ רציפה אס"ם כל הרכיבים של המטריצה $Jf(x)$ הן פונקציות רציפות (ממשיות), כלומר $\frac{\partial f_i}{\partial x_j}(x)$ הן פונקציות ממשיות רציפות ב־$x \in D$ לכל $1 \le i \le m$, $1 \le j \le n$.

אם תנאי זה מתקיים, נאמר כי $f \in C^1(D)$ – אוסף הפונקציות הגזירות ברציפות על D.

תזכורת: תהי $\mathbb{R} \to f(a,b) \ni C^1(a,b)$ ונניח כי $f'(t) \neq 0$ עבור $t \in (a,b)$. אזי בהכרח $\operatorname{sgn} f'(t)$ קבוע, כלומר תמיד $+$ או תמיד $-$, ו־$f(t)$ מונוטונית ממש ומקבלת כל ערך בין $f(c)$ ו־$f(d)$ בדיוק פעם אחת ב־(c,d).

4.3.2008 **הגדרה.** נקודה $y \in D$ תיקרא **נקודה רגולרית** (ביחס ל־f) אם $Df(y)$ רגולרית, כלומר מתקיים $\det Jf(x) \neq 0$. נקודה רגולרית

משפט 41: אם $y \in D$ נקודה רגולרית, קיים $\delta > 0$ כך שתמונה V של $B(y,\delta)$ על־ידי f היא קבוצה פתוחה, וההעתקה $f : B(y,\delta) \to V$ חד־חד ערכית.

דוגמה. אם f לינארית, $f(x) = Ax$ ל־$A \in M_{n\times n}$, אזי $Jf = A$; במצב זה, אם $\det A \neq 0$, f מעבירה כדור פתוח לקבוצה פתוחה.

הוכחה. בדרך הלמות.

למה 1.41: קיים $\delta_1 > 0$ כך שכל נקודות $\overline{B(y,\delta_1)}$ רגולריות לגבי f: אם $x \in \overline{B(y,\delta_1)}$ אז $\det Jf(z) \neq 0$.

הוכחה. הפונקציה $z \mapsto \det Jf(z)$ רציפה (סכום מכפלות של פונקציות רציפות) ושונה מ־0 ב־y. לכן בכדור מספיק קטן $\overline{B(y,\delta_1)}$ היא שונה מ־0.

למה 2.41: קיים $m > 0$ כך ש־$|Df(z)u| \ge m$ לכל $z \in \overline{B(y,\delta_1)}$ ו־$u \in S^{n-1}$.

הוכחה. לפי למה 1, לכל $z \in \overline{B(y,\delta_1)}$ ו־$u \in S^{n-1}$, $Df(z)u \neq 0$. אך $Df(z)u$ (ולכן גם $|Df(z)u|$) פונקציה רציפה על $K = \overline{B(y,\delta_1)} \times S^{n-1}$, שהיא קבוצה סגורה וחסומה ולכן קומפקטית; לכן מתקבל $\min |Df(z)u| := m > 0$.

למה 3.41: קיים $0 < \delta_2 < \delta_1$ כך שלכל $z_1, z_2 \in \overline{B(y,\delta_2)}$,
$$|f(z_2) - f(z_1) - Df(z_1)(z_2 - z_1)| < \frac{m}{3}|z_2 - z_1|$$

הוכחה. ל־$1 \le i \le n$,
$$\begin{aligned}
f_i(z_2) - f_i(z_1) &= \int_0^1 \frac{d}{dt} f_i(tz_2 + (1-t)z_1)dt \\
&= \int_0^1 \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(tz_2 + (1-t)z_1)(z_2 - z_1)_j dt \\
&= \left(\int_0^1 \nabla f_i(tz_2 + (1-t)z_1)dt\right) \cdot (z_2 - z_1)
\end{aligned}$$

"נצבור" את הרכיבים, כאשר ∇f_i הוא השורה ה־i:

$$\begin{aligned} f(z_2)-f(z_1) &= (\int_0^1 Df(tz_2+(1-t)z_1)dt)(z_2-z_1) \\ &= \int_0^1 (Df(tz_2+(1-t)z_1)-Df(z_2))dt(z_2-z_1)+ \\ &\quad +Df(z_2)(z_2-z_1) \end{aligned}$$

נקבל

$$\begin{aligned} f(z_2)-f(z_1)-Df(z_2)(z_2-z_1) &= \int_0^1 (Df(tz_2+(1-t)z_1)-Df(z_2))dt(z_2-z_1) \\ |f(z_2)-f(z_1)-Df(z_2)(z_2-z_1)| &\le \int_0^1 \|Df(tz_2+(1-t)z_1)-Df(z_2)\|dt|z_2-z_1| \end{aligned}$$

כיוון ש־$f \in C^1(D)$, הדיפרנציאל $Df(z)$ הוא פונקציה רציפה של z ב־$B(y,\delta_1)$. לכן היא רציפה במידה שווה על כל קבוצה קומפקטית חלקית. בפרט, קיים $0 < \delta_2 < \delta_1$ כך ש־$\|Df(\alpha)-Df(\beta)\| < \frac{m}{3}$ אם $\alpha,\beta \in \overline{B(y,\delta_2)}$, ולכן ל־$z_1,z_2 \in \overline{B(y,\delta_2)}$, $|f(z_2) - f(z_1)-Df(z_1)(z_2-z_1)| < \frac{m}{3}|z_2-z_1|$.

למה 4.41: ל־$z_1,z_2 \in \overline{B(y,\delta_2)}$, $|f(z_2)-f(z_1)| \ge \frac{2m}{3}|z_2-z_1|$. בפרט, f חד־חד ערכית ב־$\overline{B(y,\delta_2)}$.

הוכחה.

$$\begin{aligned} |f(z_2)-f(z_1)| &\ge |Df(z_1)(z_2-z_1)|- \\ &\quad -|f(z_2)-f(z_1)-Df(z_1)(z_2-z_1)| \\ &\ge m|z_2-z_1|-\frac{m}{3}|z_2-z_1| \\ &= \frac{2m}{3}|z_2-z_1| \end{aligned}$$

למה 5.41: יהי $0 < k = \min_{|z-y|=\delta_2}|f(z)-f(y)|$. תהי $\xi \in \mathbb{R}^n$ כך ש־$|\xi-f(y)| < \frac{k}{3}$. אזי קיימת נקודה $x \in B(y,\delta_2)$ כך ש־$f(x)=\xi$.

5.3.2008 **הוכחה.** נסתכל בפונקציה $\varphi(x) = |f(x)-\xi|^2 = \sum_{i=1}^n (f_i(x)-\xi_i)^2$. היא ממשית, רציפה ב־$\overline{B(y,\delta_2)}$ וגזירה ברציפות ב־$B(y,\delta_2)$. לכן ל־$\varphi(x)$ קיים מינימום x_m ב־$\overline{B(y,\delta_2)}$. אבל אם $|x-y|=\delta_2$,

$$\begin{aligned} |f(x)-\xi| &= |f(x)-f(y)+f(y)-\xi| \\ &\ge |f(x)-f(y)|-|f(y)-\xi| \\ &\ge k-\frac{k}{3}=\frac{2}{3}k \end{aligned}$$

כלומר, אם $|x-y|=\delta_2$, $\varphi(x) \ge (\frac{2}{3}k)^2$. לעומת זאת, $\varphi(y) \le (\frac{k}{3})^2$. לכן המינימום של φ לא מתקבל על השפה $\partial\overline{B(y,\delta_2)}$. לכן $x_m \in B(y,\delta_2)$, וכיוון ש־φ גזירה ברציפות, x_m נקודה קריטית.

$$\frac{\partial}{\partial x_1}\varphi(x) = 2\sum_{i=1}^n (f_i(x)-\xi_i)\frac{\partial f_i(x)}{\partial x_1} = 2(f(x)-\xi)\begin{pmatrix} \frac{\partial f_1}{\partial x_1} \\ \vdots \\ \frac{\partial f_n}{\partial x_1} \end{pmatrix}$$

אז $\nabla\varphi(x) = 2(f(x)-\xi)Jf(x)$, ולכן $0=\nabla\varphi(x_m)=(f(x_m)-\xi)Jf(x_m)$ בנקודה x_m. אך $Jf(x_m)$ מטריצה רגולרית, ולכן $f(x_m)-\xi=0$, כלומר $x_m=x_\xi$.

ראינו כי $f : B(y, \delta_2) \to \mathbb{R}^n$ חד־חד ערכית, ונשאר להראות כי $V = f(B(y, \delta_2))$ היא פתוחה. ראינו כי $f(y)$ נקודה פנימית של V (למה 5); $y \neq \bar{y} \in B(y, \delta_2)$ אזי סביב $\bar{y}$ נחזור על הטיעונים הקודמים (כי גם $\bar{y}$ רגולרית), ולכן גם $f(\bar{y})$ פנימית ב־V, ולכן V פתוחה.

מסקנה 42 (משפט ההעתקה הפתוחה): $f : D \to \mathbb{R}^n$, $f \in C^1(D)$. נניח כי כל נקודה $y \in D$ היא רגולרית לגבי f. אזי $f(D)$ פתוחה ב־$\mathbb{R}^n$.

הוכחה. $f(D)$ היא איחוד של תמונות של כדורים כמו במשפט הקודם.

נחזור למצב של הכדור $f : B(y, \delta) \to V$ ($\delta = \delta_2$).

טענה 43: הפונקציה ההפוכה $f^{-1} : V \to B(y, \delta)$ דיפרנציאבילית ב־$f(y)$.

הוכחה. $f(z) - f(y) = Df(y) \cdot (z - y) + R_y(z)$. לכן
$$\alpha - \beta = Df(y)(f^{-1}(\alpha) - f^{-1}(\beta)) + R_y(z)$$
ומכאן $f^{-1}(\alpha) - f^{-1}(\beta) = [Df(y)]^{-1}(\alpha - \beta) - [Df(y)]^{-1}R_y(z)$. כאשר $\alpha \to \beta$, מתקבל $|\alpha - \beta| \geq \frac{2m}{3}|f^{-1}(\alpha) - f^{-1}(\beta)|$, $f^{-1}(\alpha) \to f^{-1}(\beta)$. כלומר, f^{-1} רציפה ב־β.
$|[Df(y)]^{-1}R_y(z)| \leq \|[Df(y)]^{-1}\||R_y(z)|$. אבל $R_y(z) = o(|y - z|)$, כלומר עבור $\varepsilon > 0$ נתון, קיים $\eta > 0$ כך ש־$|y - z| < \eta \Longrightarrow |R_y(z)| < \varepsilon|y - z|$: ניקח η $|\alpha - \beta| < \frac{2m}{3}\eta$ $\Longleftarrow |f^{-1}(\alpha) - f^{-1}(\beta)| < \eta \Longleftarrow |R_y(z)| < \varepsilon|y - z|$.
$|R_y(z)| < \varepsilon|y - z| \leq \varepsilon\frac{3}{2m}|f(y) - f(z)| = \frac{3\varepsilon}{2m}|\alpha - \beta|$. כלומר, ל־$\eta$ $|\alpha - \beta| < \frac{2m}{3}\eta \Longleftarrow$ $|[Df(y)]^{-1}R_y(z)| \leq \|[Df(y)]^{-1}\|\frac{3\varepsilon}{2m}|\alpha - \beta|$, כלומר $[Df(y)]^{-1}R_y(z) = o(|\alpha - \beta|)$.

משפט 44 (דיפרנציאביליות הפונקציה ההפוכה): תהי $f : D \to \mathbb{R}^n$, $f \in C^1(D)$. תהי $y \in D$ נקודה רגולרית. אזי עבור $\delta > 0$ מספיק קטן, ההעתקה $f : B(y, \delta) \to V$ (א) היא חח״ע; (ב) V פתוחה; (ג) $f^{-1} : V \to B(y, \delta)$ דיפרנציאבילית, ועבור $\alpha \in V$, $Df^{-1}(\alpha) = [Df(z)]^{-1}$, $z = f^{-1}(\alpha)$. יתר על כן, $f^{-1} \in C^1(V)$.

הערות: (א) אם $n = 1$, רגולריות פירושה $f'(y) > 0$, ואז f מונוטונית ו־$\frac{d}{dz}f^{-1}(z) = \frac{1}{f'(y)}$, $z = f(y)$; (ב) במונחי המטריצות היעקוביאניות, $Jf^{-1}(\alpha) = [Jf(z)]^{-1}$, $z = f(\alpha)$, שוויון מטריצות $n \times n$.

הוכחה. הכול הוכח במסגרת הטענה, פרט לעובדה ש־$f^{-1} \in C^1(V)$. העובדה הזאת נובעת מרציפות הפעולה של היפוך מטריצות.

אנו יודעים כי אם $\alpha^k \to \alpha \in V$ אזי $z^k \to z \in B(y, \delta)$ ולכן $Df(z^k) \to Df(z)$ (הנחת $f \in C^1(B(y, \delta))$).

מרציפות פעולת היפוך מטריצה, $Df^{-1}(\alpha^k) = (Df(z_k))^{-1} \xrightarrow{?} (Df(z))^{-1} = Df^{-1}(\alpha)$.

4.2 טרנספורמציות גיאומטריות וקואורדינטות

ב־$n = 2$, נכתוב $f(x) = (f_1(x), f_2(x))$, כש־f_1 ו־f_2 פונקציות ממשיות. נתון כדור $B(y, \delta)$ שבו $\det\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{pmatrix} \neq 0$. בלי הגבלת הכלליות, $\frac{\partial f_1}{\partial x_1}(z) > 0$. אז f_1 מונוטונית ממש ב־x_1 לכל

x_2 קבוע.

11.3.2008 [חסר בינתיים]

12.3.2008 [חסר בינתיים]

4.3 משפט הפונקציות הסתומות

19.3.2008 $D \subseteq \mathbb{R}^{n+m}$ פתוחה, $f \in C^1(D)$; ל־$D \in (x_0,y_0)$ ($x_0 \in \mathbb{R}^n$, $y_0 \in \mathbb{R}^m$), נתבונן ב־$f(x_0,y_0) = 0 \in \mathbb{R}^m$. במקרה $m=1$, ראינו שאם $\nabla f(x_0,y_0) \neq 0$ (ובה"כ $\frac{\partial f}{\partial y}(x_0,y_0) > 0$), אזי בסביבת (x_0,y_0) משוואה זו מגדירה משטח גובה: ניתן לחלץ קואורדינטה אחת כפונקציה של n אחרות, ולכתוב $y = h(x)$, כלומר $f(x,h(x)) \equiv 0$.

משפט 45 (הפונקציות הסתומות): אם (א) $f(x_0,y_0) = 0$ ו־(ב) ההעתקה $Df(x_0,y_0)(0,\cdot)$ ב־$\hom(\mathbb{R}^m,\mathbb{R}^n)$ רגולרית, אזי קיים $\delta > 0$ ו־$h : B(x_0,\delta) \to \mathbb{R}^m \in C^1(B(x_0,\delta))$ כך ש־(א) $h(x_0) = y_0$ ו־(ב) $f(x,h(x)) = 0$ לכל $x \in B(x_0,\delta)$.[19]

הוכחה. נגדיר העתקה $\tilde{f} : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ על־ידי $\tilde{f}(x,y) = \binom{x}{f(x,y)}$. התעקה זו ודאי מוגדרת ו־$C^1(B((x_0,y_0),\delta))$.

$$D\tilde{f}(x_0,y_0) = \begin{pmatrix} I & 0 \\ & Df(x_0,y_0) \end{pmatrix}$$

בכתיב בלוקים (I מגודל $n \times n$, 0 מגודל $n \times m$); לפי הנחה (ב), הבלוק בגודל $m \times m$ בפינה הימנית התחתונה רגולרי. לכן גם $D\tilde{f}(x_0,y_0)$ רגולרי, וממשפט הפונקציה ההפוכה, עבור $\delta > 0$ מספיק קטן, $\tilde{f}(x,y)$ ניתנת להיפוך על $V = \tilde{f}(B((x_0,y_0),\delta))$.

נסתכל ב־$\tilde{f}^{-1}(V)$ ובפרט ב־$(x,h(x)) = \tilde{f}^{-1}(x,0) \in C^1(V)$. ל־$x \in B(x_0,\delta)$, $f(x,h(x)) = 0$, כי $\tilde{f}(x,h(x) = \binom{x}{f(x,0)} = (x,0)$. אז $\tilde{f}(x_0,h(x_0)) = \binom{x_0}{0}$ ו־$\tilde{f}(x_0,y_0) = \binom{x_0}{0}$, ומכאן $h(x_0) = y_0$, כי מחד־חד ערכיות לא ייתכן $f(x,h_1(x)) = f(x,h_2(x))$ ל־$h_1 \neq h_2$.

התנאי במונחי המטריצה היעקוביאנית:

$$\begin{pmatrix} \frac{\partial f_1}{\partial y_1}(x_0,y_0) & \dots & \frac{\partial f_1}{\partial y_m}(x_0,y_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(x_0,y_0) & \dots & \frac{\partial f_m}{\partial y_m}(x_0,y_0) \end{pmatrix}$$

רגולרית.

דוגמה (פרמטריזציה של משטחים ב־$\mathbb{R}^3$). תהיינה $x = f(u,v)$, $y = g(u,v)$, $z = h(u,v)$ שלוש פונקציות ממשיות גזירות ברציפות בסביבת (u_0,v_0). נניח כי היעקוביאן של f,g ביחס ל־u,v $\left(\begin{smallmatrix} \frac{\partial f}{\partial u} & \\ & \frac{\partial g}{\partial v}\end{smallmatrix}\right) =|_{(u_0,v_0)} \frac{\partial(f,g)}{(u,v)}$ רגולרי. אזי קיימת סביבה $B((x_0,y_0),\delta)$ כך שההתעקה $\binom{u}{v} \mapsto \binom{x}{y}$ (ההפוכה ל־$x = f(u,v)$, $y = g(u,v)$) חד־חד ערכית וגזירה ברציפות. עבור $(x,y) \in B((x_0,y_0),\delta)$, גם $z = h(u,v) = h(u(x,y),v(x,y)) = \tilde{h}(x,y)$ היא C. לכן קבוצת הנקודות $(f(u,v),g(u,v),h(u,v))$ מגדירה משטח בסביבת (x_0,y_0).

[19] כלומר, למשוואה $f(x,y) = 0$ יש פתרון בעל n דרגות חופש: לכל $x \in \mathbb{R}^n$ שנבחר ייקבע $y \in \mathbb{R}^m$.

ניתן להסתכל על שלוש המשוואות כ־ $F(x,y,z,u,v) = \{x - f(u,v) = 0 \quad y - g(u,v) = 0 \quad z - h(u,v) = 0\}$, כאשר $(u,v) \in \mathbb{R}^2$, $(x,y,z) \in \mathbb{R}^3$ ואנחנו בסביבת $(u_0, v_0, x_0, y_0, z_0)$. $m = 3$, $n = 2$;

$$DF = \begin{pmatrix} 1 & 0 & 0 & -\frac{\partial f}{\partial u} & -\frac{\partial f}{\partial v} \\ 0 & 1 & 0 & -\frac{\partial g}{\partial u} & -\frac{\partial g}{\partial v} \\ 0 & 0 & 1 & -\frac{\partial h}{\partial u} & -\frac{\partial h}{\partial v} \end{pmatrix}$$

ממשפט הפונקציות הסתומות, $v = v(x,y)$, $u = u(x,y)$, $z = z(x,y)$ בסביבת (x_0, y_0).

בסימוני המשפט, $h : \mathbb{R}^2 \to \mathbb{R}^3$; u ו־v נקראים **הפרמטרים של המשטח**.

5 נגזרות מסדר גבוה ומשפט טיילור

5.1 נגזרות מסדר גבוה

$f : D \to \mathbb{R}$ כאשר $D \subseteq \mathbb{R}^n$ פתוחה, $f \in C^1(D)$. נניח כי $\frac{\partial f}{\partial x_j}$ היא בעלת נגזרת חלקית לפי 25.3.2008
x_k ב־D $x_0 \in D$. נסמנה $\frac{\partial}{\partial x_k}(\frac{\partial f}{\partial x_j})(x_0) = \frac{\partial^2 f}{\partial x_k \partial x_j}(x_0)$. באופן אנלוגי, הנגזרת לפי x_j של $\frac{\partial f}{\partial x_k}$ ב־x_0, אם קיימת, תסומן $\frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_k})(x_0) = \frac{\partial^2 f}{\partial x_j \partial x_k}(x_0)$.

משפט 46 (קושי): אם $\frac{\partial^2 f}{\partial x_j \partial x_k}$ ו־$\frac{\partial^2 f}{\partial x_k \partial x_j}$ קיימות בכדור $B(p,\delta) \subseteq D$ ואם הן רציפות ב־p, אזי הן שוות: $\frac{\partial^2 f}{\partial x_k \partial x_j}(p) = \frac{\partial^2 f}{\partial x_j \partial x_k}(p)$.

הוכחה. נסמן $x_j = x$, $x_k = y$. $(x,y) \in \mathbb{R}^2$. נסמן $p = (x_0, y_0)$. נסתכל בביטוי

$$A = f(x_0+h, y_0+k) - f(x_0, y_0+k) - f(x_0+h, y_0) - f(x_0, y_0)$$

נגדיר את הפונקציה $g(y) = f(x_0+h, y) - f(x_0, y)$. אז לפי משפט הערך הממוצע, עבור $0 \le \theta_1 \le 1$ מתקיים

$$\begin{aligned}
A = g(y_0+k) - g(y_0) &= g'(y_0+\theta_1 k)k \\
&= [\frac{\partial f}{\partial y}(x_0+h, y_0+\theta_1 k) + \frac{\partial f}{\partial y}(x_0, y_0+\theta_1 k)]k \\
&= k\frac{\partial^2 f}{\partial x \partial y}(x_0+\theta_2, y_0+\theta_1 k)h \\
&= hk\frac{\partial^2 f}{\partial x \partial y}(x_0+\theta_2 h, y_0+\theta_1 k)
\end{aligned}$$

באותו אופן, $A = hk\frac{\partial^2 y}{\partial y \partial x}(x_0+\theta_3 h, y_0+\theta_4 k)$. (לכל $1 \le i \le 4$, $0 \le \theta_i \le 1$.)
מכאן, לפי הנחת הרציפות ב־(x_0,y_0), נקבל מחד $\lim_{h,k\to 0}(\frac{A}{hk}) = \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)$ ומאידך $\lim_{h,k\to 0}(\frac{A}{hk}) = \frac{\partial^2 f}{\partial y \partial x}(x_0,y_0)$. מיחידות הגבול מתקבל הדרוש.

למעשה, תנאי מספיק לשוויון הנגזרות החלקיות הוא רציפות אחת מהן (ומכאן נובעת גם רציפות השנייה).

באופן דומה מוגדרות הנגזרות מסדר גבוה $\frac{\partial^l f}{\partial x_{i_1} \dots \partial x_{i_l}}$. (מניחים שכל הנגזרות מסדר עד $l-1$ קיימות ורציפות וכן שהן גזירות (ברציפות) לפי כל אחד מהמשתנים $x_1, \dots, x_n$. לדוגמה, $\frac{\partial^3 f}{\partial x_1^2 \partial x_2}$ אומר ש־$\frac{\partial^2 f}{\partial x_1 \partial x_2}$ קיימת ורציפה ב־D וכן שהיא גזירה לפי x_1. מסמנים $\frac{\partial^3 f}{\partial x_1^2 \partial x_2} = \frac{\partial^3 f}{\partial x_1(\partial x_1 \partial x_2)}$.)

אם הנגזרות רציפות (עד סדר l, כולל), סדר הגזירה אינו משנה; לדוגמה, $\frac{\partial^3 f}{\partial x_2 \partial x_1^2} = \frac{\partial^3 f}{\partial x_1^2 \partial x_2}$, כי ממשפט קושי על f מתקבל $\frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial^2 f}{\partial x_1 \partial x_2}$. אז

$$\begin{aligned}
\frac{\partial}{\partial x_1}(\frac{\partial^2 f}{\partial x_1 \partial x_2}) &= \frac{\partial}{\partial x_1}(\frac{\partial^2 f}{\partial x_2 \partial x_1}) \\
&= \frac{\partial^2}{\partial x_1 \partial x_2}(\frac{\partial f}{\partial x_1}) \\
&= \frac{\partial^2}{\partial x_2 \partial x_1}(\frac{\partial f}{\partial x_1}) = \frac{\partial^3 f}{\partial x_2 \partial x_1^2}
\end{aligned}$$

ממשפט קושי על $\frac{\partial f}{\partial x_1}$. (הנחנו שכל הנגזרות עד סדר 3 קיימות ורציפות כדי שנוכל להפעיל את משפט קושי על $\frac{\partial f}{\partial x_1}$.)

באופן כללי, ניתן להוכיח (באינדוקציה) שאם כל הנגזרות החלקיות עד הסדר הדרוש קיימות, רציפות וגזירות ברציפות, ניתן להחליף את סדר הגזירה. את מרחב הפונקציות הגזירות ברציפות

עד סדר l מסמנים ב־$C^l(D)$.

5.2 משפט טיילור

אם $g(t)$ מוגדרת על הקטע $(0,1)\subseteq\mathbb{R}$ וגזירה $l+1$ פעמים ברציפות, אזי

$$g(t)=g(t_0)+g'(t_0)(t-t_0)+\ldots+\frac{g^{(l)}(t_0)}{l!}(t-t_0)^l+\frac{g^{(l+1)}(\xi)}{(l+1)!}(t-t_0)^{l+1}$$

עבור $\xi\in I(t,t_0)$.

כעת נניח כי $f\in C^{l+1}(D)$, $D\subseteq\mathbb{R}^n$, $B(x,\delta)\subseteq D$, $y\in B(x,\delta)$. נסתכל בפונקציה $g(t)=f(ty+(1-t)x)=f(x+t(y-x))$. כיוון ש־$f\in C^{l+1}(D)$, גם $g\in C^{l+1}(-\eta,1+\eta)$ ($\eta>0$), בגלל כלל השרשרת.

$$\begin{aligned}
g'(t)&=\nabla f(x+t(y-x))\cdot(y-x)\\
g''(t)&=\textstyle\sum_{i,j=1}^n\frac{\partial^2 f}{\partial x_j\partial x_i}(x+t(y-x))\cdot(y_i-x_i)\cdot(y_j-x_j)\\
g^{(l)}(t)&=\textstyle\sum_{i_1,\ldots,i_l=1}^n\frac{\partial^l f}{\partial x_{i_1}\ldots x_{i_l}}(x+t(y-x))\cdot(y_{i_1}-x_{i_1})\cdot\ldots\cdot(y_{i_l}-x_{i_l})
\end{aligned}$$

ומנוסחת טיילור, כאשר $t_0=0$ ו־$t=1$, נקבל

$$\begin{aligned}
f(y)&=f(x)+\textstyle\sum_{i=1}^n\frac{\partial f}{\partial x_i}(x)(y_i-x_i)+\ldots+\\
&\quad+\tfrac{1}{l!}\textstyle\sum_{i_1,\ldots,i_l=1}^n\frac{\partial^l f(x)}{\partial x_{i_1}\ldots x_{i_l}}\prod_{j=1}^l(y_{i_j}-x_{i_j})\\
&\quad+\tfrac{1}{(l+1)!}\textstyle\sum_{i_1,\ldots,i_{l+1}=1}^n\frac{\partial^{l+1} f(\xi)}{\partial x_{i_1}\ldots x_{i_{l+1}}}\prod_{j=1}^{l+1}(y_{i_j}-x_{i_j})
\end{aligned}$$

($\xi\in B(x,\delta)$, ולמעשה $\xi=x+t(y-x)$ עבור $0\le t\le 1$) – **נוסחת טיילור** ל־$f\in C^{l+1}(D)$.

הביטוי ללא השארית נקרא **פולינום טיילור** ב־y מסדר l. פולינום טיילור

פולינום טיילור מסדר 1 הוא תכונת הקירוב הלינארי; לכן הוא יחיד.

5.3 מטריצת ההסיאן ואקסטרימה

נתבונן במקרה $l=1$:

$$f(y)=f(x)+\nabla f(x)(y-x)+\frac{1}{2!}\sum_{i,j=1}^n\frac{\partial^2 f}{\partial x_j\partial x_i}(\xi)(y_i-x_i)(y_j-x_j)$$

הביטוי $R_x(y)=o(|y-x|)=\frac{1}{2}\sum_{i,j=1}^n\frac{\partial^2 f}{\partial x_j\partial x_i}(\xi)(y_i-x_i)(y_j-x_j)$ הוא ביטוי השארית כאשר הנגזרות השניות רציפות.

הגדרה. המטריצה הסימטרית $H(x)=(\frac{\partial^2 f}{\partial x_i\partial x_j})_{1\le i,j\le n}$ נקראת **ה־Hessian** של f. הסיאן

תזכורת: תבניות ריבועיות

תהי $A=(A_{ij})_{1\le i,j\le n}$ מטריצה סימטרית ($A_{ij}=A_{ji}$), ונגדיר עבור $u\in\mathbb{R}^n$

$$Q(u)=\sum_{i,j=1}^n A_{ij}u_iu_j=\sum_{j=1}^n(\sum_{i=1}^n A_{ij}u_i)u_j)=\langle Au,u\rangle=u^TAu$$

אומרים ש־$A > 0$ / Q **תבנית ריבועית חיובית לחלוטין**[20] אם $\langle Au, u\rangle > 0$ $\quad \forall u \in \mathbb{R}^n$.
במקרה כזה, $\langle Au, v\rangle$ מכפלה סקלרית, ובפרט $\langle Au, u\rangle = \|u\|^2$; כיוון שכל הנורמות שקולות, $\langle Au, u\rangle = \|u\|^2 \geq c|u|^2$. דרך אחרת להראות זאת היא שעל ספירת היחידה $\langle Au, u\rangle \geq C$, ולכן נובע שלכל $u \neq 0$, $\langle A\frac{u}{|u|}, \frac{u}{|u|}\rangle \geq C$ ומכאן $\langle Au, u\rangle \geq C|u|^2$.

אם כן, ביטוי השארית הוא תבנית ריבועית המתאימה למטריצה $H(\xi)$:

$$f(y) = f(x) + \nabla f(x)(y-x) + \frac{(y-x)H(\xi)(y-x)}{2}$$

משפט 47: תהי $f \in C^2(D)$ ותהי $x \in D$ נקודה קריטית של f. נניח כי $H(x)$ חיובית לחלוטין. אזי x היא נקודת מינימום לוקאלי של f.

הוכחה. עבור $\xi, y \in B(x,\delta)$, $f(y) = f(x) + \nabla f(x)(y-x) + \frac{1}{2}(y-x)H(\xi)(y-x)$.
נניח כי $H(x) > 0$. כלומר, $\frac{1}{2}(y-x)H(x)(y-x) \geq C|y-x|^2$. היות ש־$H(y)$ רציפה (שהרי כל איבר של $H(y)$ הוא נגזרת שנייה כלשהי ולכן פונקציה רציפה), לכל $\varepsilon > 0$ קיים $\delta > 0$ כך שאם $|y-x| < \delta$ אז $\|H(y) - H(x)\| < \varepsilon$. אז $|\frac{1}{2}(y-x)(H(\xi)-H(x))(y-x)| \leq \frac{1}{2}\varepsilon|y-x|^2$. אז 26.3.2008

$$\begin{aligned}
\tfrac{1}{2}(y-x)H(\xi)(y-x) &= \tfrac{1}{2}(y-x)H(x)(y-x) + \tfrac{1}{2}(y-x)(H(\xi)-H(x))(y-x) \\
&\geq C|y-x|^2 - \tfrac{1}{2}\varepsilon|y-x|^2 \\
&\geq \tfrac{1}{2}C|y-x|^2
\end{aligned}$$

השארית תמיד חיובית בתנאים הנתונים: כלומר, הגרף נמצא מעל המישור המשיק.

משפט 48: אם $H(x) < 0$ אזי הנקודה הקריטית y היא נקודת מקסימום לוקאלי.

[20] באנגלית, המונח הוא positive definite.

6 אינטגרציה רימנית

6.1 הגדרה

תהי $Q = [a_1, b_1] \times \ldots \times [a_n, b_n]$ ותהי $f : Q \to \mathbb{R}$ חסומה.

הגדרה. **חלוקה** $P_i = \{t_0, \ldots, t_k\}$ של $[a_i, b_i]$ היא אוסף סופי של נקודות כך שמתקיים $a_1 = t_0 < \ldots < t_k = b_1$. (לקטע $I_j = [t_j, t_{j+1}]$ נקרא **קטע של החלוקה**.) חלוקה

חלוקה P של Q היא $P = P_1 \times \ldots P_n$, כאשר P_j חלוקה של $[a_j, b_j]$.

איבר של P הוא n־ייה $(t_{j_1}, \ldots, t_{j_n})$, כאשר $t_{j_k} \in P_k$. תיבה $S \in P$ היא תיבה מהצורה $S = I_{j_1} \times \ldots I_{j_n}$, כאשר $S_{j_k} \in P_k$. נפח תיבה כזו מוגדר על־ידי $V(S) = |I_{j_1}||I_{j_2}| \cdots |I_{j_n}|$.

הגדרה. P חלוקה. נגדיר $L(f, P) = \sum_{S \in P} m_S(f) V(S)$ **סכום תחתון** עם $m_s(f) = \inf_S f$, $U(f, P) = \sum_{S \in P} M_S(f) V(S)$ **סכום עליון** עם $M_S(f) = \sup_S f$. סכום (תחתון/עליון)

הגדרה. חלוקה P' נקראת **עידון** של P ($P' < P$) אם $P' \supseteq P$. בפרט, אם $S' \in P'$ תיבה, אזי $S' \subseteq S$ עבור $S \in P$ כלשהו.

טענה 49: אם P' עידון של P, אזי $L(f, P') \geq L(f, P)$ ו־$U(f, P') \leq U(f, P)$.

למה 50: לכל שתי חלוקות P ו־P', מתקיים $L(f, P) \leq U(f, P')$.

הוכחה. תהי P'' חלוקת עידון ל־P, P'. אזי $L(f, P) \leq L(f, P'') \leq U(f, P'') \leq U(f, P)$.

מסקנה 51: האוסף $\{L(f, P)\}_P$ נמצא "משמאלי" לאוסף $\{U(f, P)\}_P$.

הגדרה. f תיקרא **אינטגרבילית** אם $I = \sup_P L(f, P) = \inf_P U(f, P)$. I נקרא **האינטגרל** של f ($I_Q(f)$ או $I(f)$). אינטגרביליות

טענה 52: מרחב הפונקציות האינטגרביליות הוא מרחב לינארי, ו־I הוא פונקציונל לינארי עליו.

הוכחה. f ו־g אינטגרביליות. יהי $\varepsilon > 0$. קיימת חלוקה P_1 כך ש־$U(f, P_1) - L(f, P_1) < \frac{\varepsilon}{2}$. באותו אופן, $U(g, P_2) - L(g, P_2) < \frac{\varepsilon}{2}$. אם P היא עידון של P_1 ו־P_2, נקבל

$$\begin{array}{rcl}
U(f,P) - L(f,P) & < & \frac{\varepsilon}{2} \\
U(g,P) - L(g,P) & < & \frac{\varepsilon}{2} \\
U(f+g,P) & \leq & U(f,P) + U(g,P) \\
L(f+g,P) & \geq & L(f,P) + L(g,P) \\
\hline
U(f+g,P) - L(f+g,P) & < & \varepsilon
\end{array}$$

לכן $f + g$ אינטגרבילית. מתקיים $I(f + g) = I(f) + I(g)$ כי לכל $\varepsilon > 0$,

$$\begin{aligned}
L(f+g,P) &\geq L(f,P) + L(g,P) \\
&\geq I(f) - \tfrac{\varepsilon}{2} + I(g) - \tfrac{\varepsilon}{2} \\
&= I(f) + I(g) - \varepsilon
\end{aligned}$$

ובאותו אופן $U(f+g,P) \le I(f)+I(g)+\varepsilon$, ולכן $-\varepsilon \le I(f+g)-(I(f)+I(g)) \le \varepsilon$.
באופן דומה, אם $\lambda \in \mathbb{R}$ אזי $I(\lambda f) = \lambda I(f)$.

טענה 53: אם f רציפה על Q, אזי f אינטגרבילית.

הוכחה. בגלל רציפות במידה שווה, אם $\varepsilon > 0$ ניתן למצוא חלוקה P כך שלכל $S \in P$, $M_S(f) - m_S(f) < \varepsilon$. אז $U(f,P) - L(f,P) \le \varepsilon \sum_{S\in P} V(S) = \varepsilon V(Q)$.

פרמטר של חלוקה

הגדרה. תהי P חלוקה. נגדיר $\delta(P) = \max_{S\in P} diam(S)$ **פרמטר החלוקה**.

תיבת שפה

הגדרה. תהי P חלוקה, ותהי P' חלוקה אחרת. נאמר כי $S' \in P'$ היא **תיבת שפה** ביחס ל־P אם לא קיימת $S \in P$ כך ש־$S' \subseteq S$.

טענה 54: תהי P חלוקה נתונה. יהי $\varepsilon > 0$. אזי קיים $\eta > 0$ כך שעבור כל חלוקה P' כך ש־$\delta(P') < \eta$ מתקיים $\sum_{S'\in P' \text{ תיבת שפה}} V(S') < \varepsilon$.

הוכחה. יהי K סכום נפחי הפיאות של תיבות P. אז $\sum_{S'\in P' \text{ תיבת שפה}} V(S') \le \delta(P')K$.

6.2 תנאים לאינטגרביליות

6.2.1 תנאי רימן

סכום רימן — 1.4.2008

הגדרה. **סכום רימן** הוא $R(f;P,x_s) = \sum_{S\in P} f(x_S)V(S)$ ($x_S \in S$).

תנאי רימן

תנאי רימן: בהינתן f חסומה ממשית, קיים מספר $\overline{I}$ כך שלכל $\varepsilon > 0$ קיים $\eta > 0$ המקיים

$$\forall P \ : \ \delta(P) < \eta \quad |\overline{I} - R(f;P,x_s)| < \varepsilon \quad \forall x_S \in S \in P$$

משפט 55: f אינטגרבילית אם״ם f מקיימת את תנאי רימן.

הוכחה. ($\Leftarrow$) יהי $\varepsilon > 0$. קיימת חלוקה P, $U(f,P) - L(f,P) < \varepsilon$. נסמן $M = \sup_Q |f|$. יהי $\eta > 0$ כך שלכל חלוקה P' כך ש־$\delta(P') < \eta$ מתקיים $4M \sum_{S'\in P' \text{ תיבת שפה לפי } P} V(S') < \varepsilon$ (קיים לפי הטענה הקודמת). נקבל

$$\begin{aligned}
R(f;P',x_{S'}) &= \textstyle\sum_{S'\in P' \text{ לא שפה}} + \sum_{S'\in P' \text{ שפה}} f(x_{S'})V(S') \\
&\ge \textstyle\sum_{S'\in P' \text{ לא שפה}} m_S(f)V(S') - M\sum_{S'\in P' \text{ שפה}} V(S') \\
&\ge L(f,P) - \textstyle\sum_{S\in P} V(S \setminus \bigcup_{S'\subseteq S} S') - M\sum_{S'\in P' \text{ שפה}} V(S') \\
&\ge L(f,P) - M(\tfrac{\varepsilon}{4M} + \tfrac{\varepsilon}{4M}) \\
&= L(f,P) - \tfrac{\varepsilon}{2}
\end{aligned}$$

ובאופן דומה נקבל $R(f;P',x_{S'}) \le U(f,P) + \frac{\varepsilon}{2}$. לכן לכל $\varepsilon > 0$ קיים $\eta > 0$ כך שאם $\delta(P') < \eta$ אז $|R(f;P',x_{S'}) - I| \le \varepsilon + \frac{\varepsilon}{2}$ (כי $U(f,P) < I + \varepsilon$), ולפיכך מתקיים תנאי רימן עם $\overline{I} = I$.

($\Rightarrow$) להיפך, נניח שמתקיים תנאי רימן. יהי נתון $\varepsilon > 0$. אז קיים $\eta > 0$ כך שלכל חלוקה P', אם $\delta(P') < \eta$ אז $R(f;P',x_{S'}) < \overline{I} + \varepsilon$. על־ידי הזזה של $x_{S'}$ נקבל $U(f,P') \le \overline{I} + \varepsilon$

ו־$L(f, P') \geq \overline{I} - \varepsilon$. כלומר, לכל $\varepsilon > 0$ קיימת חלוקה P' כך ש־$U(f, P') - L(f, P') < \varepsilon$, ולכן f אינטגרבילית, ו־$\int_Q f = \overline{I}$.

טענה 56: אם f אינטגרבילית, אזי גם f^2 אינטגרבילית.

הוכחה. $U(f, P) - L(f, P) < \varepsilon$. נסמן $M = \sup |f|$. אז

$$\begin{aligned} M_S(f^2) - m_S(f^2) &= \sup(f^2) - \inf(f^2) \\ &\leq (\sup f)^2 - (\inf f)^2 \\ &\leq 2M(\sup f - \inf f) \\ &= 2M(M_S(f) - m_S(f)) \end{aligned}$$

(כי $a^2 - b^2 = (a-b)(a+b)$, ו־$a + b \leq 2M$ כאשר $a = \sup f$ ו־$b = \inf f$). לכן $U(f^2, P) - L(f^2 - P) \leq 2M\varepsilon$.

מסקנה 57: אם f ו־g אינטגרביליות, אזי fg אינטגרבילית.

הוכחה. $(f+g)^2 - (f-g)^2 = 4fg$.

6.2.2 במונחי רציפות

הגדרה. תהי f פונקציה ממשית חסומה המוגדרת על $D \subseteq \mathbb{R}^n$. אז **התנודה** של f ב־D היא $\operatorname{osc}(f, D) = \sup_D f - \inf_D f$.

הגדרה. **התנודה** של f בנקודה $x \in D$ היא תנודה

$$\operatorname{osc}(f, x) = \inf_{\delta > 0} \operatorname{osc}(f, B(x, \delta) \cap D) = \lim_{\delta \to 0} \operatorname{osc}(f, B(x, \delta) \cap D)$$

טענה 58: f רציפה ב־x אס"ם $\operatorname{osc}(f, x) = 0$.

הגדרה. תהי $A \subseteq Q$ (Q תיבה). נאמר כי ל־A יש **תכולה אפס** אם לכל $\varepsilon > 0$ A מוכלת באיחוד סופי של תיבות[21] בנפח כולל קטן מ־ε. תכולה אפס

טענה 59: f ממשית חסומה על Q (סגורה). אזי לכל $\eta > 0$, $A_\eta = \{x \in Q \mid \operatorname{osc}(f, x) \geq \eta\}$ סגורה (ולכן קומפקטית).

הוכחה. תהי $A_\eta \supseteq \{x^k\} \to x \in Q$. ניקח $\delta > 0$ ונתבונן ב־$B(x, \delta)$. עבור k מספיק גדול, $x^k \in B(x, \delta) \cap Q$. לכן קיים $\delta' > 0$ כך ש־$B(x^k, \delta') \subseteq B(x, \delta)$. אז

$$\operatorname{osc}(f, B(x, \delta) \cap Q) \geq \operatorname{osc}(f, B(x^k, \delta') \cap Q) \geq \operatorname{osc}(f, x^k) \geq \eta$$

ולכן $\operatorname{osc}(f, B(x, \delta) \cap Q) \geq \eta$ לכל $\delta > 0$, ומכאן $\operatorname{osc}(f, x) \geq \eta$.

משפט 60: f אינטגרבילית אס"ם לכל $\eta > 0$ הקבוצה A_η בעלת תכולה אפס.

[21] סגורות או פתוחות. זה שקול, כי אם נתונות תיבות פתוחות ניתן לקחת את הסגור שלהן, ואם נתונות תיבות סגורות אפשר להוסיף $\frac{\varepsilon}{2n}$ לגבולות הקטעים שמגדירים את התיבה ולהפכם לפתוחים.

הוכחה. ($\Leftarrow$) f אינטגרבילית. יהיו $\varepsilon > 0$ ו־$\eta > 0$. קיימת חלוקה P כך ש־

$$\eta \cdot \sum_{\substack{\operatorname{osc}(f,S)\geq\eta \\ S\in P}} V(S) \leq \sum (M_S - m_S)V(S) = U(f,P) - L(f,P) < \eta\varepsilon$$

ולכן $\sum_{S\in P, \operatorname{osc}(f,S)\geq\eta} V(S) < \varepsilon$.

אם $S \in x$ כאשר $\operatorname{osc}(f,S) < \eta$, אז $\operatorname{osc}(f,x) < \eta$. לכן $A_\eta \subseteq \bigcup_{S\in P, \operatorname{osc}(f,S)\geq\eta} S$ ו־$V(\bigcup S) < \varepsilon$: כלומר, A_η בעלת תכולה אפס (היא מוכלת באיחוד סופי של תיבות שנפחן הכולל קטן מ־ε).

2.4.2008 ($\Rightarrow$) נניח שמתקיים התנאי ונראה ש־f אינטגרבילית.

יהי נתון $\varepsilon > 0$. ניקח $\eta = \frac{\varepsilon}{4M}$, כאשר $M = \sup_Q |f|$. ידוע כי ניתן לכסות את A_η על־ידי אוסף סופי של תיבות פתוחות בנפח כולל קטן מ־η. ניתן להוסיף לתיבות פתוחות אלה אוסף סופי של תיבות פתוחות המכסות את ∂Q כך שהנפח הכולל של שני האוספים יהיה קטן מ־η. נקרא לאיחוד אוספים אלה $\{S_\alpha\}$. אז $\sum_\alpha V(S_\alpha) < \eta$ ו־$\bigcup_\alpha S_\alpha \supseteq \partial Q \cup A_\eta$. נסתכל ב־$G = Q \setminus \bigcup_\alpha S_\alpha$ סגורה, $G \subseteq Q^\circ$. לכל נקודה $x \in G$ ידוע כי $\operatorname{osc}(f,x) < \eta$. קיימת תיבה $\tilde{Q}_x$ שמרכזה x כך ש־$\operatorname{osc}(f,\tilde{Q}_x) < \eta$. תהי Q_x התיבה הפתוחה שמרכזה x וצלעותיה הן חצי מצלעות $\tilde{Q}_x$.[22] מספר סופי $Q_{x_1}, \ldots, Q_{x_m}$ מכסה את Q, בשל קומפקטיות G. לכן ניתן, על־ידי הפחתות מתאימות, להגיע לכיסוי של G על־ידי אוסף סופי של תיבות סגורות $\{T_\beta\}$ שאינן נחתכות בנקודות פנימיות: כלומר, אם $\beta_1 \neq \beta_2$ אז $T^\circ_{\beta_1} \cap T^\circ_{\beta_2} = \varnothing$. נסתכל על $H = \{S_\alpha, T_\beta\}$: מתקיים

$$\begin{aligned}
\sum_{S\in H}(M_S(f) - m_S(f))V(S) &\leq \sum_{S_\alpha} + \sum_{T_\beta} \\
&\leq 2M \cdot \frac{\varepsilon}{4M} + \eta \cdot V(Q) \\
&= \frac{\varepsilon}{2} + \frac{\varepsilon V(Q)}{4M}
\end{aligned}$$

למה 1.60: אם לכל $\varepsilon > 0$ ל־Q יש כיסוי סופי על־ידי תיבות סגורות H כך שמתקיים

$$\sum_{S\in H}(M_S(f) - m_S(f))V(S) < \varepsilon$$

אזי f אינטגרבילית.

הוכחה. נוכל למצוא חלוקה P כך שכל $T \in P$ מוכל ב־$S \in H$, ואז

$$\sum_{T\in P}(M_T(f) - m_T(f))V(T) \leq \sum_{S\in H}(M_S(f) - m_S(f))V(S) < \varepsilon$$

(השוויון מתקבל על־ידי איחוד כל ה־T המוכלים ב־S מסויים).

לכן מתקבל הדרוש.

הגדרה. קבוצה $A \subseteq Q$ תיקרא **בעלת מידה אפס** אם לכל $\varepsilon > 0$ קיים אוסף בן־מניה של תיבות (פתוחות או סגורות) $\{D_i\}_{i=1}^\infty$ כך ש־(א) $A \subseteq \bigcup_{i=1}^n D_i$; (ב) $\sum_{i=1}^\infty V(D_i) < \varepsilon$. מידה אפס

למה 61: איחוד בן־מניה של קבוצות בעלות מידה אפס הוא בעל מידה אפס.

[22] נשים לב שגם $\operatorname{osc}(f,\tilde{Q}_x \geq \operatorname{osc}(f,\overline{Q_x}) < \eta$.

הוכחה. בהינתן $\varepsilon > 0$, נכסה את הראשונה על־ידי אוסף בן־מניה בנפח כולל קטן מ־$\frac{\varepsilon}{2}$, את השנייה על־ידי אוסף בן־מניה בנפח כולל קטן מ־$\frac{\varepsilon}{4}$ וכו׳. איחוד כל הכיסויים הוא בן־מניה, בנפח כולל קטן מ־ε.

למה 62: קבוצה סגורה $A \subseteq Q$ (ולכן קומפקטית) היא בעלת מידה אפס אם״ם היא בעלת תכולה אפס.

הוכחה. תכולה אפס גוררת מידה אפס. להיפך, אם הקבוצה היא ממידה אפס, נכסה על־ידי אוסף בן־מניה של תיבות פתוחות בנפח קטן מ־ε. אבל תת־כיסוי סופי של זה מספיק.

מסקנה 63: **משפט:** f אינטגרבילית אס״ם A_η בעלת מידה אפס לכל $\eta > 0$.

הוכחה. A_η סגורה.

משפט 64: f אינטגרבילית אס״ם קבוצת נקודות אי־הרציפות של f היא בעלת מידה אפס.

הוכחה. $A_{\eta_2} \supseteq A_{\eta_1}$ אם $\eta_2 > \eta_1$. כלומר, מספיק לדבר על כך שלכל $k \in \mathbb{N}$, $A_{\frac{1}{k}}$ בעלת מידה אפס: כלומר, f אינטגרבילית אס״ם $\bigcup_{k=1}^{\infty} A_{\frac{1}{k}}$ בעלת מידה אפס. אבל $\bigcup_{k=1}^{\infty} A_{\frac{1}{k}} = \bigcup_{\eta>0} A_\eta$ הוא אוסף נקודות אי־הרציפות של f, שהרי x נקודת רציפות אס״ם $\operatorname{osc}(f,x) = 0$.

דוגמה. תהי $f \geq 0$ אינטגרבילית על Q. נניח כי $\int_Q f = 0$. אזי קבוצת הנקודות x בהן $f(x) > 0$ היא בעלת מידה אפס.

הוכחה. תהי x_0 נקודת רציפות של f ונניח $f(x_0) > 0$. ניתן לבנות תיבה $A \subseteq Q$ שמרכזה x_0 כך ש־$\frac{f(x_0)}{2} < f\ |_A$. תהי χ_A הפונקציה האפיינית של A.[23] אז $f \geq \frac{f(x_0)}{2}\chi_A$, ולכן מתקיים $\int_Q f \geq \frac{1}{2}f(x_0)\int_Q \chi_A = \frac{1}{2}f(x_0)V(A) > 0$. לכן כל נקודת רציפות מקיימת $f(x_0) = 0$.

6.3 משפט פוביני

תהי $Q \subseteq \mathbb{R}^{n+m}$ תיבה סגורה. נכתוב $Q = Q' \times Q''$, כאשר Q' תיבה סגורה ב־$\mathbb{R}^n$, Q'' תיבה סגורה ב־$\mathbb{R}^m$. ($Q \ni x = (x', x'') \in Q' \times Q''$), ו־$P = P' \times P''$ חלוקה ב־Q כאשר P' חלוקה ב־Q' ו־P'' חלוקה ב־Q''.)

תהי f אינטגרבילית ב־Q. עבור $x' \in Q'$, נסמן $\inf_{P''} U(f(x',*), P'') = \psi(x')$ ו־$\sup_{P''} L(x',*), P'') = \varphi(x')$. לו היינו יודעים כי $f(x',*)$ אינטגרבילית על Q'', היינו מקבלים $\psi(x') = \varphi(x') = \int_{Q''} f(x',x'')dx''$. בדרך־כלל זה לא נכון, אבל ברור ש־$\varphi(x') \leq \psi(x')$ תמיד (שהרי תמיד $U \geq L$).

משפט 65 (פוביני): תהי f אינטגרבילית ב־Q. אז הפונקציות $\psi(x')$ ו־$\varphi(x')$ אינטגרביליות על Q', ומתקיים $\int_{Q'} \varphi(x') = \int_{Q'} \psi(x') = \int_Q f = I$. יתר על כן, $\psi(x') = \varphi(x')$ פרט (אולי) לקבוצה בעלת מידה אפס.

[23] זוהי הפונקציה המקיימת $\chi_A(x) = 1$ עבור $x \in A$, $\chi_A(x) = 0$ אחרת. נקראת גם ״פונקציה מציינת״.

הוכחה. יהי $\varepsilon > 0$. ניקח חלוקה P של Q כך שכל סכום רימן $|R(f; P, x_S) - I| < \varepsilon$. נכתוב את תיבות P בצורה $S_{i,j} = S_i \times S_j$. בכל S_i ניקח $x_i \in S_i$ ו־$x_{j,i} \in S_j$. מתקיים, אם כן,

$$|\sum_{S \in P} f((x_i, x_{j,i}))V(S_i)V(S_j) - I| = |\sum_i (\sum_j f((x_i, x_{j,i}))V(S_j))V(S_i) - I| < \varepsilon$$

על־ידי לקיחת $\sup_j$ ו־$\inf_j$ (כלומר, על כל הבחירות $x_{j,i} \in S_j$), נקבל

$$|\sum_i U(f(x_i, *), P'')V(S_i) - I| \le \varepsilon$$

וכן

$$|\sum_i L(f(x_i, *), P'')V(S_i) - I| \le \varepsilon$$

לכן

$$|\sum_i \varphi(x_i)V(S_i) - I| = |R(\varphi; P', x_i) - I| \le \varepsilon$$

וכן

$$|\sum_i \psi(x_i)V(S_i) - I| = |R(\psi; P', x_i) - I| \le \varepsilon$$

ולכן φ ו־ψ אינטגרביליות, ו־$\int_Q f = \int_{Q'} \psi = \int_{Q'} \varphi$.

מכאן נקבל שגם $\int_{Q'} (\psi - \varphi) = 0$. אבל $\psi - \varphi \ge 0$, והדרוש מתקבל לפי הדוגמה למעלה.

מהמשפט מקבלים ש־$\int_Q f(x', x'')(dx''dx') = \int_{Q'} (\int_{Q''} f(x', x'')dx'')dx'$.