דיפרנציאביליות - הוכחות נבחרות

חשבון אינפיניטסימלי (3) - 80415

מרצה: תמר (תמי) ציגלר

מתרגלים: אריה דויטש, קרן דניאל ואור קופרמן

סוכם ע"י שריה אנסבכר

סמסטר א' תשפ"ד, האוניברסיטה העברית

תוכן העניינים

סביר להניח שהסיכומים שלי מכילים טעויות רבות - אני מוצא כאלה כל יום (נפוצות טעויות רשימת),
אני מפציר בכם לעדכן אותי בכל טעות שאתם מוצאים (ממש כל טעות ללא יוצא מן הכלל);
אתם מוזמנים להגיב על המסמכים ב-Google Drive, דוא"ל לי לשלוח או באתר פנייה למלא.

לסיכומים נוספים היכנסו לאתר:
אקסיומת השלמות - סיכומי הרצאות במתמטיקה
https://srayaa.wixsite.com/math

1 התחלה

♣ בסיכומים של נושא זה נעבוד אך ורק עם הנורמה והמטריקה האוקלידיות על $\mathbb{R}^n$ ועם הנורמה האופרטורית ב-$\text{Hom}\left(\mathbb{R}^k,\mathbb{R}^m\right)$, כמו כן כמעט כל הפונקציות שנעבוד איתן תהיינה מהצורה $f:\mathbb{R}^k\to\mathbb{R}^m$ (יהיו $k,m\in\mathbb{R}$) ולא נזכיר את כל אלו בכל פעם מחדש.

משפט 1.1. גזירות גוררת רציפות

תהא f פונקציה גזירה בנקודה $a\in\mathbb{R}^k$, f גם רציפה ב-a.

משפט 1.2. יחידות הנגזרת

תהא f פונקציה גזירה בנקודה $a\in\mathbb{R}^k$ ותהיינה $T_1,T_2\in\text{Hom}\left(\mathbb{R}^k,\mathbb{R}^m\right)$ כך שמתקיים[1]:

$$\lim_{x\to 0}\frac{f\left(a+x\right)-\left(T_1\left(x\right)+f\left(a\right)\right)}{\left\|x\right\|}=\lim_{x\to 0}\frac{f\left(a+x\right)-\left(T_2\left(x\right)+f\left(a\right)\right)}{\left\|x\right\|}=0$$

מתקיים $T_1=T_2$.

טענה 1.3. תהא $T:\mathbb{R}^k\to\mathbb{R}^m$ העתקה ליניארית, יהי $b\in\mathbb{R}^m$ ותהא $f:\mathbb{R}^k\to\mathbb{R}^m$ פונקציה המוגדרת ע"י (לכל $v\in\mathbb{R}^k$):

$$f\left(v\right):=T\left(v\right)+b$$

f גזירה בכל נקודה ולכל $a\in\mathbb{R}^k$ מתקיים $Df_a=T$.

♣ בפרט הנגזרת של פונקציה קבועה היא העתקת האפס, והנגזרת של כל העתקה ליניארית היא אותה העתקה ליניארית.

משפט 1.4. תהא f פונקציה גזירה בנקודה $a\in\mathbb{R}^k$, לכל $v\in\mathbb{R}^k$ הנגזרת הכיוונית $D_vf\left(a\right)$ קיימת ומתקיים $\partial_vf\left(a\right)=Df_a\left(v\right)$.

♣ כלומר הנגזרת מעתיקה כל וקטור אל הנגזרת הכיוונית שלו[2].

♣ משפט זה מראה לנו שהגרף של $Df_a+f\left(a\right)$ הוא הישריה המשיקה לגרף של f בנקודה a שכן בכל כיוון שנבחר נראה שהישר המוכל בגרף של $Df_a+f\left(a\right)$ בכיוון זה משיק לגרף של f ב-a.

♣ אם היינו מגדירים את הנגזרת הכיוונית כך שתהיה קבועה לכל הווקטורים שכיוונם זהה גם אם גודלם שונה, אז היינו צריכים לשנות את ניסוח המשפט ולכתוב:

$$\partial_vf\left(a\right)=\frac{Df_a\left(v\right)}{\left\|v\right\|}$$

[1] כלומר T_1 ו-T_2 מקיימות את הנדרש כדי ש-f תהיה גזירה ב-a.

[2] **אם לא מגדירים נגזרת כיוונית עבור וקטור האפס אז יש לסייג ולומר ש-$v\neq 0$.**

מסקנה 1.5. תהא f פונקציה גזירה בנקודה $a\in\mathbb{R}^k$, העמודה ה-j של Df_a היא הנגזרת החלקית $\partial_j f(a)$ לכל $\mathbb{N}\ni j\leq k$.

♣ Df_a היא העתקה ליניארית ולא מטריצה, כשאנו אומרים "העמודה ה-j של Df_a" כוונתנו לעמודה ה-j של המטריצה המייצגת של Df_a בבסיס הסטנדרטי.

♣ כפי שכבר ראינו בליניארית אין שום דבר מיוחד בבסיס הסטנדרטי, באותה מידה ניתן לייצג את Df_a באמצעות כל בסיס $\mathcal{B}:=(v_1,v_2,\ldots,v_k)$ של $\mathbb{R}^k$ (התחום) ובסיס $\mathcal{C}$ של $\mathbb{R}^m$ (הטווח), ואז העמודה ה-j של $[Df_a]_{\mathcal{C}}^{\mathcal{B}}$ תהיה $\left[\partial_{v_j}f(a)\right]_{\mathcal{C}}$. הנקודה הזו תהיה שימושית במיוחד בהמשך כשנרצה ללכסן את הנגזרת השנייה של הפונקציה, אז נמצא בסיס מלכסן ונייצג אותה באמצעותו והעמודות של המטריצה המייצגת תוגדרנה לפי הבסיס המלכסן שאינו בהכרח הבסיס הסטנדרטי.

♣ כפי שראינו ניתן "לפרק" את f ל-m פונקציות $f_1,f_2,\ldots,f_m:\mathbb{R}^k\to\mathbb{R}$ כך שמתקיים:

$$f(x)=\begin{bmatrix}f_1(x)\\ f_2(x)\\ \vdots\\ f_m(x)\end{bmatrix}$$

ואז f גזירה בנקודה $a\in\mathbb{R}^k$ אם"ם כל אחת מהפונקציות הללו גזירה ב-a. כעת, מכיוון שהטווח שלהן הוא $\mathbb{R}$ הרי שהנגזרת של כל אחת מהן היא מטריצת שורה, והנגזרת של f היא מטריצה בגודל $m\times k$ שהשורה ה-i שלה היא הנגזרת של f_i. א"כ הקואורדינטה ה-ij של Df_a היא הנגזרת החלקית ה-j של f_i, כלומר לכל $\mathbb{N}\ni i\leq m$ ולכל $\mathbb{N}\ni j\leq k$ מתקיים[3]:

$$[Df_a]_{ij}=(\partial_j f(a))_i=\partial_j f_i(a)=\frac{\partial f_i}{\partial x_j}(a)$$

כלומר:

$$Df_a=\begin{bmatrix}| & | & & |\\ \partial_1 f(a) & \partial_2 f(a) & \cdots & \partial_k f(a)\\ | & | & & |\end{bmatrix}=\begin{bmatrix}\frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_k}(a)\\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_k}(a)\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \cdots & \frac{\partial f_m}{\partial x_k}(a)\end{bmatrix}$$

התוצאה הזו אינה מקרית: הרי הנגזרת מנסה "לחקות" את f באמצעות מטריצה (העתקה ליניארית) והעמודה ה-j היא זו שקולטת את המשתנה ה-j, כלומר העמודה ה-j בנגזרת היא בדיוק אותו חלק שמנה "לחקות" את אופן הפעולה של f לפי המשתנה ה-j; ובצורה דומה השורה ה-i של מטריצה היא בדיוק זו שקובעת את הקואורדינטה ה-i של תוצאת הכפל של מטריצה בווקטור, כלומר השורה ה-i של הנגזרת היא בדיוק אותו חלק בנגזרת שמנסה "לחקות" את f_i.

♣ המסקנה הזו מאפשרת לנו לבדוק בקלות יחסית אם פונקציה גזירה בנקודה, איננו צריכים לבדוק את כל ההעתקות הליניאריות - אם הפונקציה גזירה אז המטריצה המייצגת של הנגזרת חייבת להיות זו שנקבעת ע"פ הנגזרות החלקיות.

[3] בכל הסימונים הללו ה-j מסמן את הנגזרת החלקית ה-j וה-i מסמן את הקואורדינטה ה-i (של f בסימון f_i ושל הנגזרת החלקית בסימון $(D_j f(a))_i$).

מסקנה 1.6. תהיינה $A\subseteq\mathbb{R}^{k}$ ו-$f:A\to\mathbb{R}$ כך ש-f גזירה בנקודה פנימית $a\in A$, לכל $v\in\mathbb{R}^{k}$ מתקיים:

$$\partial_{v}f\left(a\right)=Df_{a}\left(v\right)=\left\langle \nabla f\left(a\right)\mid v\right\rangle$$

♣ מהגדרה נובע שהנגזרת הכיוונית בכיוון מאונך לגרדיאנט היא 0.

מסקנה 1.7. תהיינה $A\subseteq\mathbb{R}^{k}$ ו-$f:A\to\mathbb{R}$ כך ש-f גזירה בנקודה פנימית $a\in A$, לכל וקטור יחידה $u\in\mathbb{R}^{k}$ מתקיים $\left|Df_{a}\left(u\right)\right|\le\left\Vert \nabla f\left(a\right)\right\Vert$, ואם בנוסף $\nabla f\left(a\right)\ne0$ אז:

$$Df_{a}\left(\pm\frac{\nabla f\left(a\right)}{\left\Vert \nabla f\left(a\right)\right\Vert }\right)=\pm\left\Vert \nabla f\left(a\right)\right\Vert$$

♣ כלומר הגרדיאנט "מצביע" על הכיוון שבו הפונקציה תלולה יותר מבכל כיוון אחר, בכיוון הגרדיאנט נמצאת העלייה התלולה ובכיוון הנגדי הירידה.

משפט 1.8. תהא f פונקציה, אם כל הנגזרות החלקיות של f בנקודה $a\in\mathbb{R}^{k}$, קיימות בסביבה של a ורציפות ב-a, אז f גזירה ב-a.

הוכחה.

- ראשית נשים לב לכך שניתן להניח בהג"כ שהטווח של f הוא $\mathbb{R}$: נפרק את f לרכיבים - אם היא הנגזרות החלקיות שלה רציפות ב-a אז הדבר נכון לכל רכיב הנפרד, ואם כל רכיב בנפרד גזיר אז גם f גזירה.
א"כ נניח שהטווח של f הוא $\mathbb{R}$, מכאן שאם f גזירה אז הנגזרת שלה היא מטריצת השורה (מסקנה 1.5):

$$\left[\begin{array}{cccc}\partial_{1}f\left(a\right) & \partial f_{2}\left(a\right) & \ldots & \partial f_{k}\left(a\right)\end{array}\right]$$

א"כ אנחנו רוצים להוכיח כי:

$$\lim_{x\to0}\frac{f\left(x\right)-f\left(a\right)-\sum_{j=1}^{k}\partial_{j}f\left(a\right)\cdot\left(x_{j}-a_{j}\right)}{\left\Vert x-a\right\Vert }=0$$

- יהי $y\in\mathbb{R}^{k}$ כך ש-y מוכל בסביבה של a; לכל $k\ge j\in\mathbb{N}$ נסמן ב-I_{j} את הקטע $\left[a_{j},y_{j}\right]$ (אם $y_{j}\ge a_{j}$) או את הקטע $\left[y_{j},a_{j}\right]$ (אם $y_{j}\le a_{j}$)[4], ונגדיר פונקציה $g_{j}:I_{j}\to\mathbb{R}$ ע"י (לכל $t\in I_{j}$):

$$g_{j}\left(t\right):=f\left(a+t\cdot e_{j}+\sum_{i=1}^{j-1}\left(y_{i}-a_{j}\right)\cdot e_{i}\right)$$

$$\begin{aligned}\Rightarrow f\left(y\right)-f\left(a\right)&=f\left(\sum_{j=1}^{k}y_{j}\cdot e_{j}\right)-f\left(a\right)=f\left(a+\sum_{j=1}^{k}\left(y_{j}-a_{j}\right)\cdot e_{j}\right)-f\left(a\right)\\&=\sum_{j=1}^{k}\left(f\left(a+\sum_{i=1}^{j}\left(y_{i}-a_{j}\right)\cdot e_{i}\right)-f\left(a+\sum_{i=1}^{j-1}\left(y_{i}-a_{j}\right)\cdot e_{i}\right)\right)\\&=\sum_{j=1}^{k}\left(g_{j}\left(y_{j}\right)-g_{j}\left(a_{j}\right)\right)\end{aligned}$$

♣ מה שעשינו בעצם הוא לפרק את הווקטור $y-a$ לרכיבים ולמדוד את השינוי של f בכל רכיב, כמובן שסכום השינויים הוא השינוי הכללי. נוכל כעת להציג את הביטוי שאנו רוצים לחשב את גבולו באופן הבא[5]:

$$\frac{f\left(x\right)-f\left(a\right)-\sum_{j=1}^{k}\partial_{j}f\left(a\right)\cdot\left(x_{j}-a_{j}\right)}{\left\Vert x-a\right\Vert }=\sum_{j=1}^{k}\left(\frac{g_{j}\left(x_{j}\right)-g_{j}\left(a_{j}\right)}{x_{j}-a_{j}}-\partial_{j}f\left(a\right)\right)\cdot\frac{x_{j}-a_{j}}{\left\Vert x-a\right\Vert }=0$$

---

[4] אם $y_{j}=a_{j}$ אז מהגדרה $I_{j}=\left\{ y_{j}\right\} =\left\{ a_{j}\right\}$.
[5] כפי שנראה בהמשך לא תהיה בעיה אם המקרה שבו $x_{j}=a_{j}$.

בשלב הבא נראה ש-g_j רציפה על הקטע הסגור וגזירה על פנים הקטע, ולכן ממשפט לגראנז' נקבל שהביטוי $\frac{g_j(x_j)-g_j(a_j)}{x_j-a_j}$ שווה לנגזרת של g בנקודה שבין x_j ל-a_j, ואז מהרציפות של הנגזרות החלקיות נקבל:

$$\lim_{x\to a}\sum_{j=1}^{k}\left(\frac{g_j\left(x_j\right)-g_j\left(a_j\right)}{x_j-a_j}-\partial_j f\left(a\right)\right)=0$$

וכל מה שנותר לנו הוא לשים לב לכך ש-$\frac{x_j-a_j}{\|x-a\|}$ הוא גודל חסום.

- מהגדרה לכל $k\geq j\in\mathbb{N}$ ולכל $t\in I_j^{\circ}$ מתקיים[6]:

$$\begin{aligned}g_j'\left(t\right)&=\lim_{h\to 0}\frac{g_j\left(t+h\right)-g_j\left(t\right)}{h}\\&=\lim_{h\to 0}\frac{f\left(h\cdot e_j+a+t\cdot e_j+\sum_{i=1}^{j-1}\left(y_i-a_j\right)\cdot e_i\right)-f\left(a+t\cdot e_j+\sum_{i=1}^{j-1}\left(y_i-a_j\right)\cdot e_i\right)}{h}\\&=\partial_j f\left(a+t\cdot e_j+\sum_{i=1}^{j-1}y_i\cdot e_i\right)\end{aligned}$$

כלומר g_j גזירה בפנים הקטע I_j והנגזרת שלה היא הנגזרת החלקית ה-j של f. מכאן שע"פ משפט הערך הממוצע (אינפי' 1), לכל $k\geq j\in\mathbb{N}$ קיים $c_j\in\left(0,1\right)$ כך שמתקיים[7]:

$$g_j\left(y_j\right)-g_j\left(a_j\right)=g'\left(c_j\cdot\left(y_j-a_j\right)\right)\cdot\left(y_j-a_j\right)$$

ולפיכך קיימים $c_1,c_2,\ldots,c_k\in\left(0,1\right)$ כך שמתקיים:

$$f\left(y\right)-f\left(a\right)=\sum_{j=1}^{k}\left(g_j\left(y_j\right)-g_j\left(a_j\right)\right)=\sum_{j=1}^{k}\partial_j f\left(a+c_j\cdot\left(y_j-a_j\right)\cdot e_j+\sum_{i=1}^{j-1}y_i\cdot e_i\right)\cdot\left(y_j-a_j\right)$$

- א"כ יהי $0<r\in\mathbb{R}$ כך ש-$B_r\left(a\right)$ מוכל בתחום ההגדרה של f, ולכל $k\geq j\in\mathbb{N}$ תהא $c_j:B_r\left(a\right)\to\left(0,1\right)$ פונקציה כך שלכל $x\in B_r\left(a\right)$ יתקיים:

$$f\left(x\right)-f\left(a\right)=\sum_{j=1}^{k}\partial_j f\left(a+c_j\left(x\right)\cdot\left(x_j-a_j\right)\cdot e_j+\sum_{i=1}^{j-1}x_i\cdot e_i\right)\cdot\left(x_j-a_j\right)$$

מכאן שלכל $a\neq x\in B_r\left(a\right)$ מתקיים:

$$\frac{f\left(x\right)-f\left(a\right)-\sum_{j=1}^{k}\partial_j f\left(a\right)\cdot\left(x_j-a_j\right)}{\|x-a\|}=\sum_{j=1}^{k}\left(\partial_j f\left(a+c_j\cdot\left(x_j-a_j\right)\cdot e_j+\sum_{i=1}^{j-1}x_i\cdot e_i\right)-\partial_j f\left(a\right)\right)\cdot\frac{x_j-a_j}{\|x-a\|}$$

כמו כן לכל $a\neq x\in B_r\left(a\right)$ ולכל $k\geq j\in\mathbb{N}$ מתקיים:

$$\left|\frac{x_j-a_j}{\|x-a\|}\right|\leq 1\qquad\qquad\lim_{x\to a}\left(\partial_j f\left(a+c_j\cdot\left(x_j-a_j\right)\cdot e_j+\sum_{i=1}^{j-1}x_i\cdot e_i\right)-\partial_j f\left(a\right)\right)=0$$

קיום הגבול הנ"ל וערכו נובעים מהרציפות של הנגזרות החלקיות ב-a, והא"ש שבצד ימין נובע ישירות מהגדרת הנורמה.

$$\Rightarrow\lim_{x\to a}\frac{f\left(x\right)-f\left(a\right)-\sum_{j=1}^{k}\partial_j f\left(a\right)\cdot\left(x_j-a_j\right)}{\|x-a\|}=0$$

כלומר f גזירה ב-a.

■

[6] אם קיים $k\geq j\in\mathbb{N}$ כך ש-$y_j=a_j$ אז $I_j^{\circ}=\emptyset$ וטענה זו מתקיימת באופן ריק, אנחנו נראה בהמשך שלא נזדקק לכך ש-$I_j^{\circ}\neq\emptyset$ לכל $k\geq j\in\mathbb{N}$.

[7] שוויון זה הוא טריוויאלי במקרה שבו $y_j=a_j$ ולכן אין צורך בכך ש-$I_j^{\circ}\neq\emptyset$.

טענה 1.9. תהא f פונקציה, f גזירה ברציפות בקבוצה פתוחה $U\subseteq\mathbb{R}^{k}$ אם"ם הנגזרות החלקיות שלה רציפות בכל נקודה ב-U.

2 כללי גזירה

משפט 2.1. גזירה היא פעולה ליניארית

תהיינה f ו-g פונקציות דיפרנציאביליות בנקודה $x\in\mathbb{R}^{k}$ ויהיו $\alpha,\beta\in\mathbb{R}$, מתקיים:

$$D_{\alpha\cdot f+\beta\cdot g}\left(x\right)=\alpha\cdot D_{f}\left(x\right)+\beta\cdot D_{g}\left(x\right)$$

משפט 2.2. כלל השרשרת

תהיינה $A\subseteq\mathbb{R}^{k}$ ו-$B\subseteq\mathbb{R}^{m}$ קבוצות פתוחות, ותהיינה $f:A\to B$ ו-$g:B\to\mathbb{R}^{n}$ פונקציות כך ש-f גזירה בנקודה $a\in A$ ו-g גזירה ב-$f\left(a\right)$. הפונקציה $g\circ f$ גזירה ב-a ומתקיים:

$$D\left(g\circ f\right)_{a}=Dg_{f\left(a\right)}\circ Df_{a}$$

♣ אם מניחים את עניין הדיפרנציאביליות של $g\circ f$ כלל השרשרת אינטואיטיבי למדי: אנחנו מקרבים את f ו-g באמצעות העתקות ליניאריות, היה זה מפתיע מאד אם הקירוב הליניארי של $g\circ f$ לא היה הרכבת הקירובים של f ו-g.

הוכחה. תהיינה $r:A\to B$ ו-$s:B\to\mathbb{R}^{n}$ שתי פונקציות המוגדרות ע"י (לכל $x\in A$ ולכל $y\in B$):

$$\begin{aligned}r\left(x\right) & :=f\left(x\right)-f\left(a\right)-Df_{a}\left(x-a\right)\\ s\left(y\right) & :=g\left(y\right)-g\left(f\left(a\right)\right)-Dg_{f\left(a\right)}\left(y-f\left(a\right)\right)\end{aligned}$$

מהגדרת r ו-s נובע שלכל $x\in A$ מתקיים:

$$\begin{aligned}g\left(f\left(x\right)\right)-g\left(f\left(a\right)\right)-Dg_{f\left(a\right)}\left(Df_{a}\left(x-a\right)\right) & =g\left(f\left(x\right)\right)-g\left(f\left(a\right)\right)-Dg_{f\left(a\right)}\left(f\left(x\right)-f\left(a\right)-r\left(x\right)\right)\\ & =g\left(f\left(x\right)\right)-g\left(f\left(a\right)\right)-Dg_{f\left(a\right)}\left(f\left(x\right)-f\left(a\right)\right)+Dg_{f\left(a\right)}\left(r\left(x\right)\right)\\ & =s\left(f\left(x\right)\right)+Dg_{f\left(a\right)}\left(r\left(x\right)\right)\end{aligned}$$

ומכאן שגם:

$$\begin{aligned}\frac{\left\Vert g\left(f\left(x\right)\right)-g\left(f\left(a\right)\right)-Dg_{f\left(a\right)}\left(Df_{a}\left(x-a\right)\right)\right\Vert }{\left\Vert x-a\right\Vert } & \leq\frac{\left\Vert s\left(f\left(x\right)\right)\right\Vert +\left\Vert Dg_{f\left(a\right)}\left(r\left(x\right)\right)\right\Vert }{\left\Vert x-a\right\Vert }\\ & =\frac{\left\Vert s\left(f\left(x\right)\right)\right\Vert }{\left\Vert x-a\right\Vert }+\frac{\left\Vert Dg_{f\left(a\right)}\left(r\left(x\right)\right)\right\Vert }{\left\Vert x-a\right\Vert }\\ & =\frac{\left\Vert s\left(f\left(x\right)\right)\right\Vert }{\left\Vert x-a\right\Vert }+\left\Vert Dg_{f\left(a\right)}\left(\frac{r\left(x\right)}{\left\Vert x-a\right\Vert }\right)\right\Vert \end{aligned}$$

מהיות f ו-g גזירות ב-a וב-$f\left(a\right)$ (בהתאמה) נובע כי:

$$\lim_{x\to a}\frac{\left\Vert r\left(x\right)\right\Vert }{\left\Vert x-a\right\Vert }=\lim_{y\to f\left(a\right)}\frac{\left\Vert s\left(y\right)\right\Vert }{\left\Vert y-f\left(a\right)\right\Vert }=0$$

מכאן שהמחובר הימני באגף שימין של השוויון הקודם (מודגש באדום) שואף ל-0 כש-x שואף ל-a ($Dg_{f\left(a\right)}$ רציפה ב-0 ו-$Dg_{f\left(a\right)}\left(0\right)=0$), א"כ נותר לנו להוכיח רק שמתקיים:

$$\lim_{x\to a}\frac{\left\Vert s\left(f\left(x\right)\right)\right\Vert }{\left\Vert x-a\right\Vert }=0$$

תהא $\psi : B \to \mathbb{R}$ פונקציה המוגדרת ע"י (לכל $y \in B$):

$$\psi(y) := \begin{cases} \frac{\|s(y)\|}{\|y-f(a)\|} & y \neq f(a) \\ 0 & y = f(a) \end{cases}$$

נשים לב לכך ש-$s(f(a)) = 0$ ולכן לכל $x \in A$ מתקיים:

$$\psi(f(x)) \cdot \frac{\|f(x) - f(a)\|}{\|x-a\|} = \frac{\|s(f(x))\|}{\|x-a\|}$$

מאריתמטיקה של גבולות, מרציפות של ψ ב-$f(a)$, וממשפט ההצבה בגבולות, נובע כי:

$$\lim_{x\to a} \frac{\|s(f(x))\|}{\|x-a\|} = \lim_{x\to a} \psi(f(x)) \cdot \lim_{x\to a} \frac{\|f(x)-f(a)\|}{\|x-a\|} = \psi(f(a)) \cdot \|Df_a\|_{\text{op}} = 0$$

■

מסקנה 2.3 תהא f פונקציה גזירה בנקודה $a \in \mathbb{R}^k$ ותהא g פונקציה גזירה ב-$f(a)$, מתקיים:

$$\|D(g \circ f)_a\|_{op} \leq \left\|Dg_{f(a)}\right\|_{op} \cdot \|Df_a\|_{op}$$

למה 2.4 תהיינה $A \subseteq \mathbb{R}^k$ ו-$f : A \to \mathbb{R}^k$ פונקציה הפיכה, אם f גזירה בנקודה פנימית $a \in A$ ו-f^{-1} גזירה ב-$f(a)$ אז מכלל השרשרת נובע כי:

$$\text{Id}(a) = \text{Id}'(a) = \left(Df^{-1}\right)_{f(a)} \circ Df(a)$$

ומכאן ש-$(Df(a))^{-1} = \left(Df^{-1}\right)_{f(a)}$.[8]

מסקנה 2.5 תהיינה $A \subseteq \mathbb{R}^k$ ו-$f : A \to \mathbb{R}^k$ פונקציה הפיכה וגזירה בנקודה פנימית $a \in A$, אם Df_a אינה הפיכה אז f^{-1} אינה גזירה ב-$f(a)$.

♣ **האם המשפט עבור גזירה של פונקציה הופכית מאינפי' 1 נכון גם בממדים גבוהים? כנראה שלא. הנה התרגום שלו לממדים גבוהים:**

משפט. גזירת פונקציה הופכית

תהיינה $A, B \subseteq \mathbb{R}^k$ ו-$f : A \to B$ פונקציה הפיכה, לכל $b \in B$ כך ש-f גזירה ב-$f^{-1}(b)$ וגם $Df_{f^{-1}(b)}$ הפיכה מתקיים:

$$\left(Df^{-1}\right)_b = \left(Df_{f^{-1}(b)}\right)^{-1}$$

אנחנו נוכיח בפרק הבא את משפט הפונקציה ההפוכה שנותן את אותה תוצאה אבל דורש ש-f תהיה **גזירה ברציפות בסביבה** של $f^{-1}(b)$.

מצד שני משפט הפונקציה ההפוכה אינו דורש ש-f תהיה הפיכה, אלא מוכיח שאם $Df_{f^{-1}(b)}$ הפיכה אז קיימת סביבה של $f^{-1}(b)$ שבה f הפיכה.

למה 2.6 תהא $p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ פונקציית המכפלה הפנימית המוגדרת ע"י $p(x,y) := \langle x \mid y \rangle$ לכל $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$, p גזירה בכל נקודה ולכל $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$ מתקיים:

$$Dp_{(x,y)} = \left[\begin{array}{cccc|cccc} y_1 & y_2 & \cdots & y_n & x_1 & x_2 & \cdots & x_n \end{array}\right]$$

משפט 2.7. כלל לייבניץ

יהיו $A \subseteq \mathbb{R}^k$ ו-$f : \mathbb{R}^k \to \mathbb{R}^n$ כך ש-f גזירה בנקודה פנימית $a \in A$, ויהיו $B \subseteq \mathbb{R}^m$ ו-$g : \mathbb{R}^m \to \mathbb{R}^n$ כך ש-g גזירה בנקודה פנימית $b \in B$.

תהא $h : \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}$ פונקציה המוגדרת ע"י $h(x,y) := \langle f(x) \mid g(y) \rangle$ לכל $(x,y) \in \mathbb{R}^k \times \mathbb{R}^m$, h גזירה ב-(a,b) ולכל $(x,y) \in \mathbb{R}^k \times \mathbb{R}^m$ מתקיים:

$$Dh_{(a,b)}(x,y) = \langle g(b) \mid Df_a(x) \rangle + \langle f(a) \mid Dg_b(y) \rangle$$

[8] בגלל שמדובר בהעתקות ליניאריות על אותו מרחב מספיק להראות הפיכות בכיוון אחד בלבד.

3 יחסי הגומלין בין פונקציה לנגזרתה

תהא $A\subseteq\mathbb{R}^k$.

3.1 התחלה

סימון: לכל $a,b\in\mathbb{R}^k$ נסמן $[a,b]:=\{a+t\cdot(b-a)\mid t\in[0,1]\}$.

♣ מהגדרה $[a,b]=[b,a]$.

משפט 3.1. משפט הערך הממוצע
נניח ש-A היא קבוצה פתוחה, תהיינה $a,b\in A$ כך ש-$[a,b]\subseteq A$ ותהא $f:A\to\mathbb{R}$ פונקציה גזירה.
קיים $t\in(0,1)$ כך שמתקיים:
$$f(b)-f(a)=Df_{a+t\cdot(b-a)}(b-a)$$

♣ זוהי הכללה של משפט הערך הממוצע של לגראנז' מאינפי' 1 - אם $k=1$ אז $Df_{a+t\cdot(b-a)}(b-a)=f'(a+t\cdot(b-a))\cdot(b-a)$, כלומר:
$$\frac{f(b)-f(a)}{b-a}=f'(a+t\cdot(b-a))$$

למה 3.2. תהא $\gamma:[0,1]\to\mathbb{R}^m$ מסילה גזירה ב-$(0,1)$ ורציפה ב-$[0,1]$ כך שהקבוצה $\left\{\left\|D\gamma_t\right\|_{\text{op}}:t\in(0,1)\right\}$ חסומה ונסמן ב-M את החסם העליון שלה, מתקיים:
$$\left\|\gamma(1)-\gamma(0)\right\|\leq M$$

הוכחה. תהיינה $f:[0,1]\to\mathbb{R}$ פונקציות המוגדרות ע"י (לכל $t\in[0,1]$):
$$\begin{aligned}f(t)&:=\left\|\gamma(t)-\gamma(0)\right\|=\sqrt{\langle\gamma(t)-\gamma(0)\mid\gamma(t)-\gamma(0)\rangle}\\ g(t)&:=f^2(t)=\langle\gamma(t)-\gamma(0)\mid\gamma(t)-\gamma(0)\rangle\end{aligned}$$

נסמן $x:=\sup\{t\in[0,1]:f(t)=0\}$; x אכן מוגדר היטב מפני שהקבוצה הנ"ל חסומה מלעיל ע"פ הגדרתה, ואינה ריקה משום ש-0 שייך אליה.
נשים לב לכך שמהיות f רציפה נובע ש-$f(x)=0$[9], כלומר $x=\max\{t\in[0,1]:f(t)=0\}$. מכאן שאם $x=1$ אז קיבלנו את המבוקש:
$$\left\|\gamma(1)-\gamma(0)\right\|=f(1)=f(x)=0\leq M$$

לכן נניח בהג"כ ש-$x<1$. מכלל לייבניץ נובע כי (לכל $t\in(0,1)$):
$$g'(t)=2\cdot\langle\gamma(t)-\gamma(0)\mid\gamma'(t)\rangle$$

ומכלל השרשרת נובע כי (לכל $t\in(0,1)$ כך ש-$f(t)\neq 0$):
$$f'(t)=\frac{g'(t)}{2\cdot\sqrt{\langle\gamma(t)-\gamma(0)\mid\gamma(t)-\gamma(0)\rangle}}=\frac{g'(t)}{2\cdot f(t)}$$

מכאן שע"פ א"ש קושי-שוורץ מתקיים (לכל $t\in(0,1)$ כך ש-$f(t)\neq 0$):
$$2\cdot f(t)\cdot f'(t)=g'(t)\leq 2\cdot\left\|\gamma(t)-\gamma(0)\right\|\cdot\left\|\gamma'(t)\right\|=2\cdot f(t)\cdot\left\|\gamma'(t)\right\|$$

[9] מהאפיון השני של החסם העליון נובע שלכל $0<\varepsilon\in\mathbb{R}$ קיים $t\in(x-\varepsilon,x)$ כך ש-$f(t)=0$.

וממילא גם:

$$f'(t) \leq \|\gamma'(t)\| \leq M$$

מהגדרה לכל $t \in (x,0)$ מתקיים $f(t) \neq 0$ ולכן ממשפט לגראנז' (אינפי' 1) נובע שמתקיים (עבור $t \in (x,1)$ כלשהו):

$$\|\gamma(1)-\gamma(0)\| < \frac{\|\gamma(1)-\gamma(0)\|}{1-x} = \frac{\|\gamma(1)-\gamma(0)\|-0}{1-x} = \frac{f(1)-f(x)}{1-x} = f'(t) \leq \|\gamma'(t)\| \leq M$$

■

משפט 3.3. נניח ש-A היא קבוצה פתוחה תהיינה $a,b \in A$ כך ש-$[a,b] \subseteq A$ ותהא $f: A \to \mathbb{R}^m$ פונקציה גזירה.
נניח שהקבוצה $\left\{\|Df_c\|_{\text{op}} : c \in [a,b]\right\}$ חסומה ונסמן ב-M את החסם העליון שלה, מתקיים:

$$\|f(b)-f(a)\| \leq M \cdot \|b-a\|$$

כלומר אם הנגזרת של f חסומה אז f רציפה לפי ליפשיץ והחסם העליון על קבוצת הנורמות האופרטוריות של הנגזרות הוא קבוע ליפשיץ של f.

הוכחה. תהא $\gamma: [0,1] \to A$ המסילה המוגדרת ע"י $\gamma(t) := a + t \cdot (b-a)$ לכל $t \in [0,1]$, מהגדרה לכל $t \in (0,1)$ מתקיים $\|D\gamma_t\|_{\text{op}} = \|b-a\|$ ולכן לכל $t \in (0,1)$ מתקיים:

$$\|D(f \circ \gamma)_t\|_{\text{op}} \leq \left\|Df_{\gamma(t)}\right\|_{\text{op}} \cdot \|D\gamma_t\|_{\text{op}} = M \cdot \|b-a\|$$

מהלמה (3.2) נובע כי:

$$\|f(b)-f(a)\| = \|(f \circ \gamma)(1) - (f \circ \gamma)(0)\| \leq M \cdot \|b-a\|$$

■

3.2 נגזרות גבוהות

טענה 3.4. יהיו $A \subseteq \mathbb{R}^k$ ו-$f: A \to \mathbb{R}^m$ כך ש-f גזירה פעמיים בנקודה פנימית $a \in A$.
$D^2 f_a$ היא פונקציה ביליניארית, כלומר לכל $v,w,u \in \mathbb{R}^k$ ולכל $\alpha, \beta \in \mathbb{R}$ מתקיים:

$$D^2 f_a(v)(\alpha \cdot w + \beta \cdot u) = \alpha \cdot D^2 f_a(v)(w) + \beta \cdot D^2 f_a(v)(u)$$
$$D^2 f_a(\alpha \cdot v + \beta \cdot u)(w) = \alpha \cdot D^2 f_a(v)(w) + \beta \cdot D^2 f_a(u)(w)$$

♣ למעשה הטענה נכונה גם עבור נגזרות מסדר גבוה יותר - נגזרת בנקודה, מכל סדר שהוא, היא פונקציה מולטי-ליניארית.

הוכחה. השוויון הראשון נובע מהעובדה ש-$D^2 f_a$ היא העתקה ליניארית מ-$\mathbb{R}^k$ ל-$\text{Hom}\left(\mathbb{R}^k, \mathbb{R}^m\right)$, ולכן $D^2 f_a(v)$ היא העתקה ליניארית. השוויון השני גם הוא נובע מהעובדה ש-$D^2 f_a$ היא העתקה ליניארית מ-$\mathbb{R}^k$ ל-$\text{Hom}\left(\mathbb{R}^k, \mathbb{R}^m\right)$, כלומר $D^2 f_a(\alpha \cdot v + \beta \cdot u) = \alpha \cdot D^2 f_a(v) + \beta \cdot D^2 f_a(u)$ ולכן גם:

$$D^2 f_a(\alpha \cdot v + \beta \cdot u)(w) = \left(\alpha \cdot D^2 f_a(v) + \beta \cdot D^2 f_a(u)\right)(w) = \alpha \cdot D^2 f_a(v)(w) + \beta \cdot D^2 f_a(u)(w)$$

■

מסקנה 3.5. תהא f פונקציה גזירה פעמיים בנקודה $a \in \mathbb{R}^k$, כל הנגזרות החלקיות מסדר שני של f ב-a קיימות ולכל $k \geq i,j \in \mathbb{N}$ מתקיים:

$$D^2 f_a(e_i, e_j) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial_j f}{\partial x_i}(a)$$

מסקנה 3.6. תהא f פונקציה גזירה פעמיים בנקודה $a \in \mathbb{R}^k$, כל הנגזרות החלקיות מסדר שני של f ב-a קיימות ולכל $x,y \in \mathbb{R}^k$ מתקיים:

$$D^2 f_a(x,y) = \sum_{i=1}^{k} \sum_{j=1}^{k} x_i \cdot y_j \cdot \frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \sum_{i=1}^{k} \sum_{j=1}^{k} x_i \cdot y_j \cdot \frac{\partial_j f}{\partial x_i}(a)$$

הוכחה. יהיו $x,y\in\mathbb{R}^{k}$ מהמסקנה הקודמת (3.5) ומטענה 3.4 נובע כי:

$$\begin{aligned}D^{2}f_{a}\left(x,y\right) & =D^{2}f_{a}\left(\sum_{i=1}^{k}x_{i}\cdot e_{i},y\right)=\sum_{i=1}^{k}x_{i}\cdot D^{2}f_{a}\left(e_{i},y\right)=\sum_{i=1}^{k}x_{i}\cdot D^{2}f_{a}\left(e_{i},\sum_{j=1}^{k}y_{j}\cdot e_{j}\right)\\ & =\sum_{i=1}^{k}x_{i}\cdot\left(\sum_{j=1}^{k}y_{j}\cdot D^{2}f_{a}\left(e_{i},e_{j}\right)\right)=\sum_{i=1}^{k}\sum_{j=1}^{k}x_{i}\cdot y_{j}\cdot D^{2}f_{a}\left(e_{i},e_{j}\right)=\sum_{i=1}^{k}\sum_{j=1}^{k}x_{i}\cdot y_{j}\cdot\frac{\partial^{2}f}{\partial x_{i}\partial x_{j}}\left(a\right)\end{aligned}$$

■

מסקנה 3.7. תהא $f:A\to\mathbb{R}$ פונקציה גזירה פעמיים בנקודה פנימית $a\in A$, לכל $x,y\in\mathbb{R}^{k}$ מתקיים:

$$D^{2}f_{a}\left(x,y\right)=y^{t}\cdot H\left(f\right)\cdot x=\left\langle y\mid H\left(f\right)\cdot x\right\rangle$$

כאשר $H\left(f\right)$ היא מטריצת ההסיאן של f ב-a.

משפט 3.8. תהא f פונקציה, אם כל הנגזרות החלקיות מסדר שני של f בנקודה $a\in\mathbb{R}^{k}$ קיימות בסביבה של a ורציפות ב-a, אז f גזירה פעמיים ב-a.

משפט 3.9. משפט שוורץ[10] (או **משפט קלרו**[11] **על שוויון פונקציות מעורבות**)
תהא f פונקציה גזירה פעמיים בסביבה של נקודה $a\in\mathbb{R}^{k}$, אם $D^{2}f$ רציפה ב-a אז לכל $k\geq i,j\in\mathbb{N}$ מתקיים:

$$\frac{\partial^{2}f}{\partial x_{i}\partial x_{j}}\left(a\right)=\frac{\partial^{2}f}{\partial x_{j}\partial x_{i}}\left(a\right)$$

♣ נשים לב לכך שמתקיים:

$$\begin{aligned}\frac{\partial^{2}f}{\partial x_{i}\partial x_{j}}\left(a\right) & =\lim_{t\to0}\frac{\partial_{j}f\left(a+t\cdot e_{i}\right)-\partial_{j}f\left(a\right)}{t}=\lim_{t\to0}\frac{\lim_{s\to0}\frac{f\left(a+t\cdot e_{i}+s\cdot e_{j}\right)-f\left(a+t\cdot e_{i}\right)}{s}-\lim_{s\to0}\frac{f\left(a+s\cdot e_{j}\right)-f\left(a\right)}{s}}{t}\\ & =\lim_{t\to0}\frac{\lim_{s\to0}\frac{\left(f\left(a+t\cdot e_{i}+s\cdot e_{j}\right)-f\left(a+t\cdot e_{i}\right)\right)-\left(f\left(a+s\cdot e_{j}\right)-f\left(a\right)\right)}{s}}{t}\\ & =\lim_{t\to0}\left(\lim_{s\to0}\frac{\left(f\left(a+t\cdot e_{i}+s\cdot e_{j}\right)-f\left(a+t\cdot e_{i}\right)\right)-\left(f\left(a+s\cdot e_{j}\right)-f\left(a\right)\right)}{s\cdot t}\right)\\ & =\lim_{t\to0}\left(\lim_{s\to0}\frac{\left(f\left(a+t\cdot e_{i}+s\cdot e_{j}\right)-f\left(a+s\cdot e_{j}\right)\right)-\left(f\left(a+t\cdot e_{i}\right)-f\left(a\right)\right)}{s\cdot t}\right)\end{aligned}$$

ובאותו אופן גם:

$$\frac{\partial^{2}f}{\partial x_{j}\partial x_{i}}\left(a\right)=\lim_{s\to0}\left(\lim_{t\to0}\frac{\left(f\left(a+t\cdot e_{i}+s\cdot e_{j}\right)-f\left(a+s\cdot e_{j}\right)\right)-\left(f\left(a+t\cdot e_{i}\right)-f\left(a\right)\right)}{s\cdot t}\right)$$

כלומר כל שעלינו לעשות הוא "להפוך" את סדר הגבולות.

הוכחה. יהי $0<r\in\mathbb{R}$ כך ש-f גזירה ב-$B_{r}\left(a\right)$, נניח ש-$D^{2}f$ רציפה ב-a, ויהיו $k\geq i,j\in\mathbb{N}$.

- ניתן להוכיח את המשפט עבור כל קואורדינטה בנפרד, א"כ נניח בהג"כ שהטווח של f הוא $\mathbb{R}$.
תהיינה $g,h:\left(-r,r\right)^{2}\to\mathbb{R}$ פונקציות המוגדרות ע"י (לכל $s,t\in\left(-r,r\right)$):

$$\begin{aligned}g\left(s,t\right) & :=f\left(a+s\cdot e_{i}+t\cdot e_{j}\right)-f\left(a+s\cdot e_{i}\right)\\ h\left(s,t\right) & :=f\left(a+s\cdot e_{i}+t\cdot e_{j}\right)-f\left(a+t\cdot e_{j}\right)\end{aligned}$$

[10] ערך בוויקיפדיה שוורץ הרמן.
[11] ערך בוויקיפדיה: קלרו אלכסיס.

מכאן שלכל $s,t\in\left(-r,r\right)$ מתקיים:

$$\begin{aligned}g\left(s,t\right)-g\left(0,t\right) & =\left(f\left(a+s\cdot e_{i}+t\cdot e_{j}\right)-f\left(a+s\cdot e_{i}\right)\right)-\left(f\left(a+t\cdot e_{j}\right)-f\left(a\right)\right)\\ & =\left(f\left(a+s\cdot e_{i}+t\cdot e_{j}\right)-f\left(a+t\cdot e_{j}\right)\right)-\left(f\left(a+s\cdot e_{i}\right)-f\left(a\right)\right)=h\left(s,t\right)-h\left(s,0\right)\end{aligned}$$

- מהגזירות של f פעם ראשונה נובע שכאשר מקבעים את אחד הרכיבים של g מקבלים פונקציה רציפה וגזירה על $\left(-r,r\right)$, מכאן שע"פ משפט הערך הממוצע (אינפי' 1) לכל $s\in\left(-r,r\right)$ קיים $\lambda_{s}\in\left(0,1\right)$ כך שמתקיים:

$$g\left(s,t\right)-g\left(0,t\right)=s\cdot\partial_{1}g\left(\lambda_{s}\cdot s,t\right)=s\cdot\left(\partial_{i}f\left(a+\lambda_{s}\cdot s\cdot e_{i}+t\cdot e_{j}\right)-\partial_{i}f\left(a+\lambda_{s}\cdot s\cdot e_{i}\right)\right)$$

ובאותו אופן מהגזירות של f פעם שנייה נובע שלכל $s,t\in\left(-r,r\right)$ קיימים $\lambda_{s},\lambda_{t}\in\left(0,1\right)$ כך שמתקיים:

$$\begin{aligned}g\left(s,t\right)-g\left(0,t\right) & =s\cdot\left(\partial_{i}f\left(a+\lambda_{s}\cdot s\cdot e_{i}+t\cdot e_{j}\right)-\partial_{i}f\left(a+\lambda_{s}\cdot s\cdot e_{i}\right)\right)\\ & =s\cdot t\cdot\partial_{j}\partial_{i}f\left(a+\lambda_{s}\cdot s\cdot e_{i}+\lambda_{t}\cdot t\cdot e_{j}\right)\end{aligned}$$

- מהרציפות של $\partial_{j}\partial_{i}f$ ב-a נקבל שמתקיים:

$$\lim_{(s,t)\to(0,0)}\frac{g\left(s,t\right)-g\left(0,t\right)}{s\cdot t}=\lim_{x\to a}\partial_{j}\partial_{i}f\left(x\right)=\frac{\partial f}{\partial x_{j}\partial x_{i}}\left(a\right)$$

מצד שני באותה צורה נקבל שמתקיים:

$$\lim_{(s,t)\to(0,0)}\frac{h\left(s,t\right)-h\left(s,0\right)}{s\cdot t}=\lim_{x\to a}\partial_{i}\partial_{j}f\left(x\right)=\frac{\partial f}{\partial x_{i}\partial x_{j}}\left(a\right)$$

והרי כפי שהזכרנו לכל $s,t\in\left(-r,r\right)$ מתקיים $g\left(s,t\right)-g\left(0,t\right)=h\left(s,t\right)-h\left(s,0\right)$, כלומר מתקיים:

$$\frac{\partial f}{\partial x_{i}\partial x_{j}}\left(a\right)=\lim_{(s,t)\to(0,0)}\frac{h\left(s,t\right)-h\left(s,0\right)}{s\cdot t}=\lim_{(s,t)\to(0,0)}\frac{g\left(s,t\right)-g\left(0,t\right)}{s\cdot t}=\frac{\partial f}{\partial x_{j}\partial x_{i}}\left(a\right)$$

■

מסקנה 3.10. תהא f פונקציה גזירה פעמיים בסביבה של נקודה $a\in\mathbb{R}^{k}$, אם $D^{2}f$ רציפה ב-a אז מטריצת ההסיאן של f ב-a היא מטריצה סימטרית ולכן גם לכסינה.

3.3 נקודות קיצון

משפט 3.11. משפט פרמה

תהא $f:A\to\mathbb{R}$ פונקציה גזירה בנקודה פנימית $a\in A$, אם a היא נקודת קיצון מקומית של f אז $Df_{a}=0$, כלומר a היא נקודה קריטית.

משפט 3.12. משפט טיילור[12]

תהא f פונקציה גזירה n פעמים בנקודה $a\in\mathbb{R}^{k}$, מתקיים:

$$\lim_{x\to a}\frac{f\left(x\right)-P_{f,n,a}\left(x\right)}{\left\Vert x-a\right\Vert }=0$$

משפט טיילור אינו בחומר למבחן ולכן עוד לא כתבתי לו הוכחה.

תזכורת: יהי V מ"ו מעל $\mathbb{R}$ ותהא $B:V\times V\to\mathbb{R}$ תבנית בילינארית.

- B תיקרא חיובית בהחלט אם לכל $0_{V}\neq v\in V$ מתקיים $B\left(v,v\right)>0$.
- B תיקרא חיובית למחצה אם לכל $0_{V}\neq v\in V$ מתקיים $B\left(v,v\right)\geq0$.

[12] ערך בוויקיפדיה: טיילור ברוק.

- B תיקרא <u>שלילית בהחלט</u> אם לכל $0_V \neq v \in V$ מתקיים $B(v,v) < 0$.
- B תיקרא <u>שלילית למחצה</u> אם לכל $0_V \neq v \in V$ מתקיים $B(v,v) \leq 0$.

מסקנה 3.13. תהא $f : A \to \mathbb{R}$ פונקציה גזירה פעמיים בנקודה פנימית $a \in A$ כך ש-$Df_a = 0$ (כלומר a היא נקודה קריטית), מתקיימים כל הפסוקים הבאים:

- אם D^2f חיובית בהחלט אז a היא נקודת מינימום מקומית של f.
- אם D^2f שלילית בהחלט אז a היא נקודת מקסימום מקומית של f.
- אם a היא נקודת מינימום מקומית של f אז D^2f חיובית למחצה.
- אם a היא נקודת מקסימום מקומית של f אז D^2f שלילית למחצה.

מסקנה 3.14. תהא $f : A \to \mathbb{R}$ פונקציה גזירה פעמיים בסביבה של נקודה פנימית $a \in A$ כך ש-$Df_a = 0$ (כלומר a היא נקודה קריטית) ו-D^2f רציפה ב-a (כלומר ההסיאן סימטרית), מתקיימים כל הפסוקים הבאים:

- אם כל הערכים העצמיים של $H(f)$ חיוביים אז a היא נקודת מינימום מקומית של f.
- אם כל הערכים העצמיים של $H(f)$ שליליים אז a היא נקודת מקסימום מקומית של f.
- אם a היא נקודת מינימום מקומית של f אז כל הערכים העצמיים של $H(f)$ אי-שליליים.
- אם a היא נקודת מקסימום מקומית של f אז כל הערכים העצמיים של $H(f)$ אי-חיוביים.

4 משפט הפונקציה ההפוכה ומסקנותיו

4.1 משפט הפונקציה ההפוכה

למה 4.1. הדטרמיננטה היא פונקציה רציפה.

♣ כוונתנו כאן היא שהפונקציה $\det : \mathrm{Hom}\left(\mathbb{R}^k, \mathbb{R}^k\right) \to \mathbb{R}$ היא פונקציה רציפה, כזכור הגדרנו את הדטרמיננטה של העתקה ליניארית ע"י הדטרמיננטה של מטריצה מייצגת שלה.

מסקנה 4.2. לכל $T \in \mathrm{Hom}\left(\mathbb{R}^k, \mathbb{R}^k\right)$ הפיכה קיים $0 < r \in \mathbb{R}$ כך שלכל $S \in B_r(T)$[13] גם S הפיכה.

למה 4.3. תהא $A \subseteq \mathbb{R}^k$ קבוצה פתוחה, תהא $f : A \to \mathbb{R}^m$ פונקציה גזירה ברציפות, ויהיו $a, b, c \in A$ כך ש-$[b,c] \subseteq A$[14]; מתקיים:

$$\|f(c) - f(b) - Df_a(c-b)\| \leq \|c-b\| \cdot \max_{x \in [b,c]} \|Df_x - Df_a\|_{\text{op}}$$

הוכחה. תהא $g : A \to \mathbb{R}^m$ פונקציה המוגדרת ע"י $g(x) := f(x) - Df_a(x)$ לכל $x \in A$, מכאן שלכל $x \in A$ מתקיים:

$$\|Dg_x\|_{\text{op}} = \|Df_x - Df_a\|_{\text{op}}$$

ממשפט 3.3 נובע כי:

$$\begin{aligned}\|f(c) - f(b) - Df_a(c-b)\| &= \|(f(c) - Df_a(c)) - (f(b) - Df_a(b))\| = \|g(c) - g(b)\| \\ &\leq \|c-b\| \cdot \sup_{x\in[b,c]} \|Dg_x\|_{\text{op}} = \|c-b\| \cdot \sup_{x\in[b,c]} \|Df_x - Df_a\|_{\text{op}}\end{aligned}$$

■

[13] נזכיר ש-$B_r(T)$ הוא כדור פתוח ב-$\mathrm{Hom}\left(\mathbb{R}^k, \mathbb{R}^k\right)$ סביב T.
[14] הדרישה שהנגזרת תהיה רציפה באה להבטיח שהיא חסומה על $[b,c]$.

למה 4.4. תהא $A\subseteq\mathbb{R}^{k}$ קבוצה פתוחה ותהא $g:A\to\mathbb{R}^{k}$ פונקציה גזירה, אם קיים $\varepsilon\in(0,1)$ כך שלכל $x\in A$ מתקיים:
$$\left\Vert Dg_{x}-\text{Id}\right\Vert _{\text{op}}\le\varepsilon$$
אז לכל $a\in A$ ולכל $0<r\in\mathbb{R}$ כך ש-$B_{r}\left(a\right)\subseteq A$ מתקיים (עבור אותו ε):
$$B_{(1-\varepsilon)r}\left(g\left(a\right)\right)\subseteq g\left(B_{r}\left(a\right)\right)\subseteq B_{(1+\varepsilon)r}\left(g\left(a\right)\right)$$

הוכחה. נניח שקיים $\varepsilon\in(0,1)$ כך שלכל $a\in A$ מתקיים $\left\Vert Dg_{a}-\text{Id}\right\Vert _{\text{op}}\le\varepsilon$, ויהי ε כנ"ל.

- יהיו $a\in A$ ו-$0<r\in\mathbb{R}$ כך ש-$B_{r}\left(a\right)\subseteq A$, מא"ש המשולש ההפוך נובע כי:
$$\left|\left\Vert Dg_{a}\right\Vert _{\text{op}}-1\right|=\left|\left\Vert Dg_{a}\right\Vert _{\text{op}}-\left\Vert \text{Id}\right\Vert _{\text{op}}\right|\le\left\Vert Dg_{a}-\text{Id}\right\Vert _{\text{op}}\le\varepsilon$$
כלומר $\left\Vert Dg_{a}\right\Vert _{\text{op}}\le1+\varepsilon$, ולכן ע"פ משפט 3.3 לכל $x\in B_{r}\left(a\right)$ מתקיים:
$$\left\Vert g\left(x\right)-g\left(a\right)\right\Vert \le\left(1+\varepsilon\right)\cdot\left\Vert x-a\right\Vert <\left(1+\varepsilon\right)\cdot r$$
כלומר $g\left(B_{r}\left(a\right)\right)\subseteq B_{(1+\varepsilon)r}\left(g\left(a\right)\right)$.

- כעת יהי $y\in B_{(1-\varepsilon)r}\left(g\left(a\right)\right)$, ותהא $h:A\to\mathbb{R}^{k}$ פונקציה המוגדרת ע"י:
$$h\left(x\right):=-g\left(x\right)+x+y$$
  - מהגדרה לכל $x\in A$ מתקיים $Dh_{x}=-Dg_{x}+\text{Id}$, ולכן גם:
$$\left\Vert Dh_{x}\right\Vert _{\text{op}}=\left\Vert -Dg_{x}+\text{Id}\right\Vert _{\text{op}}=\left\Vert Dg_{x}-\text{Id}\right\Vert _{\text{op}}\le\varepsilon$$
ומכאן שע"פ משפט 3.3 לכל $x_{1},x_{2}\in A$ מתקיים:
$$\left\Vert h\left(x_{1}\right)-h\left(x_{2}\right)\right\Vert \le\varepsilon\cdot\left\Vert x_{1}-x_{2}\right\Vert$$
  - יהי $0<s\in\mathbb{R}$ כך ש-$s<r$ ומתקיים $y\in B_{(1-\varepsilon)s}^{\hat{}}\left(g\left(a\right)\right)$, מכאן שלכל $x\in\hat{B}_{s}\left(a\right)$ מתקיים:
$$\begin{aligned}\left\Vert h\left(x\right)-a\right\Vert & =\left\Vert \left(h\left(x\right)-h\left(a\right)\right)+\left(h\left(a\right)-a\right)\right\Vert \le\left\Vert h\left(x\right)-h\left(a\right)\right\Vert +\left\Vert h\left(a\right)-a\right\Vert \\
 & \le\varepsilon\cdot s+\left\Vert \left(-g\left(a\right)+a+y\right)-a\right\Vert =\varepsilon\cdot s+\left\Vert y-g\left(a\right)\right\Vert <\varepsilon\cdot s+\left(1-\varepsilon\right)\cdot s=x
\end{aligned}$$
  - מכל אלו יחד נובע שהצמצום של h ל-$\hat{B}_{s}\left(a\right)$ הוא העתקה מכווצת על מרחב שלם, ולכן ע"פ משפט ההעתקה המכווצת יש לו נקודת שבת, כלומר קיים $x\in B_{s}^{\hat{}}\left(x\right)$ כך שמתקיים:
$$0=\left\Vert h\left(x\right)-x\right\Vert =\left\Vert y-g\left(x\right)\right\Vert$$
מהגדרה $s<r$ ולכן זה אומר שקיים $x\in B_{r}\left(a\right)$ כך ש-$g\left(x\right)=y$.
y הנ"ל היה שרירותי ולכן לכל $y\in B_{(1-\varepsilon)r}\left(a\right)$ קיים $x\in A$ כך ש-$g\left(x\right)=y$, כלומר $B_{(1-\varepsilon)r}\left(a\right)\subseteq g\left(B_{r}\left(a\right)\right)$.

■

למה 4.5. תהיינה $A\subseteq\mathbb{R}^{k}$ קבוצה פתוחה ו-$a\in A$, ותהא $f:A\to\mathbb{R}^{k}$ פונקציה חח"ע וגזירה ברציפות כך ש-Df_{a} הפיכה. אם קיים $\varepsilon\in(0,1)$ כך שלכל $x\in A$ מתקיים:
$$\left\Vert \left(Df_{a}\right)^{-1}\circ Df_{x}-\text{Id}\right\Vert _{\text{op}}\le\varepsilon$$
אז לכל $a\in A$ ולכל $0<r\in\mathbb{R}$ כך ש-$B_{r}\left(a\right)\subseteq A$ מתקיים (עבור אותו ε):
$$Df_{a}\left(B_{(1-\varepsilon)r}\left(\left(Df_{a}\right)^{-1}\left(f\left(a\right)\right)\right)\right)\subseteq f\left(B_{r}\left(a\right)\right)\subseteq Df_{a}\left(B_{(1+\varepsilon)r}\left(\left(Df_{a}\right)^{-1}\left(f\left(a\right)\right)\right)\right)$$

הוכחה. נניח שקיים $\varepsilon\in\left(0,\frac{1}{2}\right)$ כך שלכל $a\in A$ מתקיים $\left\Vert \left(Df_{a}\right)^{-1}\circ Df_{x}-\mathrm{Id}\right\Vert _{\mathrm{op}}\leq\varepsilon$, ויהי ε כנ"ל.
יהיו $a\in A$ ו-$0<r\in\mathbb{R}$ כך ש-$B_{r}\left(a\right)\subseteq A$, ונסמן $g:=\left(Df_{a}\right)^{-1}\circ f$; מכלל השרשרת נובע כי (לכל $x\in A$):
$$Dg_{x}=\left(Df_{a}\right)^{-1}\circ Df_{x}$$
ולכן לפי ההנחה מתקיים (לכל $x\in A$):
$$\left\Vert Dg_{x}-\mathrm{Id}\right\Vert _{\mathrm{op}}\leq\varepsilon$$
ומהלמה הקודמת (4.4) נובע כי:
$$B_{(1-\varepsilon)r}\left(g\left(a\right)\right)\subseteq g\left(B_{r}\left(a\right)\right)\subseteq B_{(1+\varepsilon)r}\left(g\left(a\right)\right)$$
ע"פ הגדרת g מתקיים:
$$f\left(B_{r}\left(a\right)\right)=Df_{a}\left(g\left(B_{r}\left(a\right)\right)\right)$$
ומכאן נובע כי:
$$\begin{aligned}Df_{a}\left(B_{(1-\varepsilon)r}\left(\left(Df_{a}\right)^{-1}\left(f\left(a\right)\right)\right)\right) & =Df_{a}\left(B_{(1-\varepsilon)r}\left(g\left(a\right)\right)\right)\subseteq f\left(B_{r}\left(a\right)\right)\\
 & \subseteq Df_{a}\left(B_{(1+\varepsilon)r}\left(g\left(a\right)\right)\right)=Df_{a}\left(B_{(1+\varepsilon)r}\left(\left(Df_{a}\right)^{-1}\left(f\left(a\right)\right)\right)\right)
\end{aligned}$$

■

משפט 4.6. משפט הפונקציה ההפוכה

תהא $A\subseteq\mathbb{R}^{k}$ קבוצה פתוחה, ותהא $f:A\to\mathbb{R}^{k}$ פונקציה גזירה ברציפות ב-A, ויהי $a\in A$.
אם Df_{a} הפיכה אז קיימת סביבה $U\subseteq A$ של a כך ש-$f\mid_{U}$ חח"ע ו-$f\left(U\right)$ פתוחה, ובנוסף f^{-1}[15] גזירה ברציפות ב-$f\left(U\right)$ ולכל $y\in f\left(U\right)$ מתקיים $\left(Df^{-1}\right)_{y}=\left(Df_{f^{-1}\left(y\right)}\right)^{-1}$.

הוכחה. נניח ש-Df_{a} הפיכה, נסמן $M:=2\cdot\left\Vert \left(Df_{a}\right)^{-1}\right\Vert _{\text{op}}$, ויהי $0<r\in\mathbb{R}$ כך שלכל $x\in B_{r}\left(a\right)$ הנגזרת Df_{x} תהיה הפיכה ובנוסף:

$$\left\Vert Df_{x}-Df_{a}\right\Vert _{\text{op}}\leq\frac{1}{2\cdot\left\Vert \left(Df_{a}\right)^{-1}\right\Vert _{\text{op}}}=\frac{1}{M}$$

מהיות f גזירה ברציפות, וממסקנה 4.2 נובע שאכן קיים r כזה.

- נוכיח שהצמצום של f ל-$B_{r}\left(a\right)$ הוא חח"ע.
  יהיו $x_{1},x_{2}\in B_{r}\left(a\right)$, ע"פ למה 4.3 מתקיים:

$$\left\Vert f\left(x_{1}\right)-f\left(x_{2}\right)-Df_{a}\left(x_{1}-x_{2}\right)\right\Vert \leq\left\Vert x_{1}-x_{2}\right\Vert \cdot\max_{x\in\left[x_{1},x_{2}\right]}\left\Vert Df_{x}-Df_{a}\right\Vert _{\text{op}}\leq\frac{\left\Vert x_{1}-x_{2}\right\Vert }{2\cdot\left\Vert \left(Df_{a}\right)^{-1}\right\Vert _{\text{op}}}$$

$$\begin{aligned}\Rightarrow\left\Vert f\left(x_{1}\right)-f\left(x_{2}\right)\right\Vert & \geq\left\Vert Df_{a}\left(x_{1}-x_{2}\right)\right\Vert -\frac{\left\Vert x_{1}-x_{2}\right\Vert }{2\cdot\left\Vert \left(Df_{a}\right)^{-1}\right\Vert _{\text{op}}}\\
 & =\frac{1}{\left\Vert \left(Df_{a}\right)^{-1}\right\Vert _{\text{op}}}\cdot\left(\left\Vert \left(Df_{a}\right)^{-1}\right\Vert _{\text{op}}\cdot\left\Vert Df_{a}\left(x_{1}-x_{2}\right)\right\Vert -\frac{1}{2}\cdot\left\Vert x_{1}-x_{2}\right\Vert \right)\\
 & \geq\frac{2}{M}\cdot\left(\left\Vert \left(Df_{a}\right)^{-1}\left(Df_{a}\left(x_{1}-x_{2}\right)\right)\right\Vert -\frac{1}{2}\cdot\left\Vert x_{1}-x_{2}\right\Vert \right)\\
 & \frac{2}{M}\cdot\left(\left\Vert x_{1}-x_{2}\right\Vert -\frac{1}{2}\cdot\left\Vert x_{1}-x_{2}\right\Vert \right)=\frac{\left\Vert x_{1}-x_{2}\right\Vert }{M}
\end{aligned}$$

  מכאן ש-$x_{1}=x_{2}\iff f\left(x_{1}\right)=f\left(x_{2}\right)$, כלומר הצמצום של f ל-$B_{r}\left(a\right)$ הוא חח"ע.
  בנוסף נובע מכאן ש-f^{-1}[16] רציפה לפי ליפשיץ ובפרט רציפה - לכל $y_{1},y_{2}\in f\left(B_{r}\left(a\right)\right)$ מתקיים:

$$\left\Vert f\left(f^{-1}\left(y_{1}\right)\right)-f\left(f^{-1}\left(y_{2}\right)\right)\right\Vert \geq\frac{\left\Vert f^{-1}\left(y_{1}\right)-f^{-1}\left(y_{2}\right)\right\Vert }{M}$$

  וממילא גם:

$$\left\Vert f^{-1}\left(y_{1}\right)-f^{-1}\left(y_{2}\right)\right\Vert \leq M\cdot\left\Vert f\left(f^{-1}\left(y_{1}\right)\right)-f\left(f^{-1}\left(y_{2}\right)\right)\right\Vert =M\cdot\left\Vert y_{1}-y_{2}\right\Vert $$

- נשים לב שאי אפשר לקחת את $B_{r}\left(a\right)$ בתור הסביבה U שמבטיח המשפט מפני שאולי $f\left(B_{r}\left(a\right)\right)$ אינה פתוחה.
  ע"פ למה 4.5 מתקיים:

$$V:=Df_{a}\left(B_{\frac{r}{2}}\left(\left(Df_{a}\right)^{-1}\left(f\left(a\right)\right)\right)\right)\subseteq f\left(B_{r}\left(a\right)\right)$$

  V היא קבוצה פתוחה מפני ש-Df_{a} היא הומיאומורפיזם, ולכן אם נסמן $U:=f^{-1}\left(V\right)\cap B_{r}\left(a\right)$ נקבל סביבה של a[17] כך ש-$f\left(U\right)=V$ פתוחה; בנוסף f חח"ע ב-U מפני ש-$U\subseteq B_{r}\left(a\right)$ ולכן יש ל-$f\mid_{U}$ הופכית.

- יהי $b\in V$ ונסמן $c:=f^{-1}\left(b\right)$, לכל $y\in V$ קיים $x\in U$ כך ש-$f\left(x\right)=y$ ולכן גם:

$$\begin{aligned}f^{-1}\left(y\right)-f^{-1}\left(b\right)-\left(Df_{f^{-1}\left(b\right)}\right)^{-1}\left(y-b\right) & =x-c-\left(Df_{c}\right)^{-1}\left(f\left(x\right)-f\left(c\right)\right)\\
 & =-\left(Df_{c}\right)^{-1}\left(f\left(x\right)-f\left(c\right)-Df_{c}\left(x-c\right)\right)
\end{aligned}$$

---

[15] למעשה f אינה בהכרח הפיכה ומדובר בהופכית של $f\mid_{U}$.
[16] ההופכית של הצמצום של f ל-$B_{r}\left(a\right)$.
[17] מהגדרה $V=Df_{a}\left(B_{\frac{r}{2}}\left(\left(Df_{a}\right)^{-1}\left(f\left(a\right)\right)\right)\right)\in Df_{a}\left(\left(Df_{a}\right)^{-1}\left(f\left(a\right)\right)\right)=a$ ולכן $a\in f^{-1}\left(V\right)\cap B_{r}\left(a\right)=U$.

וממילא (נזכור ש-f היא פונקציה רציפה לפי ליפשיץ עם קבוע ליפשיץ M):

$$\begin{aligned}\frac{\left\|f^{-1}(y)-f^{-1}(b)-\left(Df_{f^{-1}(b)}\right)^{-1}(y-b)\right\|}{\|y-b\|}&\leq M\cdot\frac{\left\|-\left(Df_c\right)^{-1}\left(f(x)-f(c)-Df_c(x-c)\right)\right\|}{\left\|f^{-1}(x)-f^{-1}(c)\right\|}\\&\leq\left\|\left(Df_c\right)^{-1}\right\|_{\text{op}}\cdot M\cdot\frac{\|f(x)-f(c)-Df_c(x-c)\|}{\|x-c\|}\xrightarrow[x\to c]{}0\end{aligned}$$

מהרציפות של f^{-1} נובע כי:

$$\lim_{y\to b}\frac{\left\|f^{-1}(y)-f^{-1}(b)-\left(Df_{f^{-1}(b)}\right)^{-1}(y-b)\right\|}{\|y-b\|}=\lim_{x\to c}\left(\left\|\left(Df_c\right)^{-1}\right\|_{\text{op}}\cdot M\cdot\frac{\|f(x)-f(c)-Df_c(x-c)\|}{\|x-c\|}\right)=0$$

מכאן ש-f^{-1} גזירה בכל נקודה $b\in V$ והנגזרת היא אכן זו שציפינו לה.

ע"פ המשפט המרכזי של המצורפת המטריצה הפונקציה המעתיקה מטריצה להופכית שלה היא פונקציה רציפה, שכן הדטרמיננטה רציפה (למה 4.1), וכמו כן ההעתקות $x\mapsto Df_x$ ו-$y\mapsto f^{-1}(y)$ גם הן רציפות; מכאן שההעתקה $y\mapsto\left(Df^{-1}\right)_y$ גם היא העתקה רציפה משום שהיא הרכבה של שלוש האחרונות (בסדר הפוך מזה שבו הצגנו אותן).

■

מסקנה 4.7. תהא $A\subseteq\mathbb{R}^k$ קבוצה פתוחה ותהא $f:A\to\mathbb{R}^m$ פונקציה גזירה ברציפות, אם קיימת נקודה $a\in A$ כך ש-Df_a חח"ע אז $k\geq m$ ובנוסף קיימת סביבה $U\subseteq A$ של a כך ש-$f|_U$ חח"ע.

4.2 משפט ההעתקה הפתוחה

משפט 4.8. משפט ההעתקה הפתוחה

יהיו $k,m\in\mathbb{N}$ כך ש-$m\leq k$, תהא $A\subseteq\mathbb{R}^k$ קבוצה פתוחה ותהא $f:A\to\mathbb{R}^m$ פונקציה גזירה ברציפות. אם $\text{rk}\left(Df_a\right)=m$ לכל $a\in A$, אז f היא העתקה פתוחה.

הוכחה. תהא $U\subseteq A$ פתוחה, ותהא $a\in U$; נרצה להוכיח ש-$f(a)$ היא נקודה פנימית של $f(U)$.

מהעובדה ש-$\text{rk}\left(Df_a\right)=m$ נובע שקיימת העתקה ליניארית $T:\mathbb{R}^m\to\mathbb{R}^k$ כך ש-$\text{rk}\left(Df_a\circ T\right)=m$[18], כלומר $Df_a\circ T$ הפיכה.

א"כ תהא T כנ"ל, ותהא $g:\mathbb{R}^m\to\mathbb{R}^k$ פונקציה המוגדרת ע"י $g(x):=a+T(x)$ לכל $x\in\mathbb{R}^m$. ע"פ השרשרת $f\circ g$ גזירה בכל נקודה ונגזרתה בנקודה $x\in\mathbb{R}^m$ היא $Df_{a+T(x)}\circ T$, מכאן ש-$f\circ g$ גזירה ברציפות.

g רציפה ולכן $g^{-1}(U)$ היא קבוצה פתוחה, מהגדרה $0\in g^{-1}(U)$ ולכן ממשפט הפונקציה ההפוכה נובע שקיימת סביבה $\tilde{U}\subseteq g^{-1}(U)$ של 0 כך ש-$f\left(g\left(\tilde{U}\right)\right)$ פתוחה, תהא $\tilde{U}$ כנ"ל. מהגדרה מתקיים $g\left(\tilde{U}\right)\subseteq g\left(g^{-1}(U)\right)\subseteq U$, ולכן גם $f\left(g\left(\tilde{U}\right)\right)\subseteq f(U)$.

נשים לב לכך ש-$f(a)=f(g(0))$ ולכן $f(a)\in f\left(g\left(\tilde{U}\right)\right)$, ומהיות $f\left(g\left(\tilde{U}\right)\right)$ פתוחה ומוכלת ב-$f(U)$ נובע ש-$f(a)$ היא נקודה פנימית של $f(U)$. ■

[18] לדוגמה: T תעתיק בסיס של $\mathbb{R}^m$ לבסיס של $\left(\ker\left(Df_a\right)\right)^{\perp}$.

4.3 משפט הפונקציה הסתומה

משפט הפונקציה הסתומה אינו בחומר למבחן ולכן לא כתבתי לו הוכחה.

♣ מהו השיפוע של הישר המשיק למעגל היחידה בנקודה $\left(\frac{1}{2},\frac{\sqrt{3}}{2}\right)$?
היינו רוצים לגזור את הגאומטרי המקום $\left\{ (x,y)\in\mathbb{R}^{2}\mid x^{2}+y^{2}=1\right\}$ אלא שהוא אינו מהווה גרף של פונקציה, ולכן לא נוכל לגזור אותו.
הפתרון הטבעי כמובן הוא לחלק את המקום הגאומטרי לשני חלקים:

$$\left\{ (x,y)\in\mathbb{R}^{2}\mid x^{2}+y^{2}=1\right\} =\left\{ (x,y)\in\mathbb{R}^{2}\mid x^{2}+y^{2}=1,\ y>0\right\} \cup\left\{ (x,y)\in\mathbb{R}^{2}\mid x^{2}+y^{2}=1,\ y\leq0\right\}$$

את שתי הקבוצות הללו ניתן להציג כגרפים של פונקציות:

$$\left\{ (x,y)\in\mathbb{R}^{2}\mid x^{2}+y^{2}=1,\ y\geq0\right\} =\left\{ \ (x,y)\in\mathbb{R}^{2}\ \middle|\ y=\sqrt{1-x^{2}}\ \right\}$$
$$\left\{ (x,y)\in\mathbb{R}^{2}\mid x^{2}+y^{2}=1,\ y<0\right\} =\left\{ \ (x,y)\in\mathbb{R}^{2}\ \middle|\ y=-\sqrt{1-x^{2}}\ \right\}$$

ולכן אותן ניתן לגזור. מכללי גזירה נובע שהנגזרת של $\sqrt{1-x^{2}}$ היא $\frac{-2x}{2\sqrt{1-x^{2}}}=-\frac{x}{\sqrt{1-x^{2}}}$, ולכן השיפוע של המשיק למעגל היחידה בנקודה $\left(\frac{1}{2},\frac{\sqrt{3}}{2}\right)$ הוא $-\frac{1}{\sqrt{3}}=-\frac{\frac{1}{2}}{\sqrt{1-\left(\frac{1}{2}\right)^{2}}}$, וזה כמובן מתאים לכך שהמשיק למעגל מאונך לרדיוס בנקודה[19].

♣ המקרה שראינו לעיל הוא מקרה פשוט מאד, מה אם נרצה להשתמש בכלים החזקים שפיתחנו עבור פונקציות כדי לחקור מקומות גאומטריים המוגדרים באמצעות נוסחאות מסובכות יותר?
במישור ניתן להביא כל מקום גאומטרי כזה לצורה $\left\{ (x,y)\in\mathbb{R}^{2}\mid G\left(x,y\right)=0\right\}$ עבור פונקציה $g:\mathbb{R}^{2}\to\mathbb{R}$, במקרה הנ"ל G זו תהיה הפונקציה המוגדרת ע"י $G\left(x,y\right):=x^{2}+y^{2}-1$ לכל $(x,y)\in\mathbb{R}^{2}$. כעת נשים לב לכך שמה שעשינו הוא למצוא סביבה U של $\left(\frac{1}{2},\frac{\sqrt{3}}{2}\right)$[20] וסביבה V של $\frac{1}{2}$[21], כך שקיימת פונקציה $f:V\to B$[22] המקיימת $y=f\left(x\right)\iff G\left(x,y\right)=0$ לכל $(x,y)\in U$; כלומר לקחנו רק חלק מהמקום הגאומטרי כך שאותו חלק הוא אכן גרף של פונקציה.

משפט 4.9. משפט הפונקציה הסתומה
תהיינה $A\subseteq\mathbb{R}^{k}$ ו-$B\subseteq\mathbb{R}^{m}$ קבוצות פתוחות ותהא $G:A\times B\to\mathbb{R}^{m}$ פונקציה גזירה ברציפות.
תהא $(a,b)\in A\times B$ כך ש-$g\left(a,b\right)=0$ ונסמן ב-M את תת-המטריצה של $DG_{(a,b)}$ הכוללת את m העמודות הימניות שלה (כלומר העמודה ה-j ב-M היא הנגזרת החלקית של G לפי המשתנה ה-$k+j$ שהם כל המשתנים ב-B):

$$M:=\left[\begin{array}{cccc}
| & | & & |\\
\partial_{k+1}G\left(a,b\right) & \partial_{k+2}G\left(a,b\right) & \cdots & \partial_{k+m}G\left(a,b\right)\\
| & | & & |
\end{array}\right]$$

אם M הפיכה אז קיימות: קבוצה פתוחה $U\subseteq A\times B$ כך ש-$(a,b)\in U$, קבוצה פתוחה $V\subseteq\mathbb{R}^{k}$ ופונקציה גזירה ברציפות $f:V\to B$ כך שמתקיים (לכל $(x,y)\in U$):

$$G\left(x,y\right)=0\iff y=f\left(x\right)$$

---

[19] השיפוע של הרדיוס לנקודה $\left(\frac{1}{2},\frac{\sqrt{3}}{2}\right)$ הוא $\sqrt{3}$ ואנחנו יודעים ששני ישרים במישור הם מאונכים אם"ם מכפלת השיפועים שלהם היא -1 (למעט עבור ישרים המקבילים לציר ה-y).

[20] $\left\{ (x,y)\in\mathbb{R}^{2}\mid y>0\right\}$.

[21] הקטע $(-1,1)$, במקרה זה לא היה צורך להצטמצם הרבה בציר ה-x אבל אם היה מדובר במקום גאומטרי מסובך יותר ייתכן שהיינו צריכים להצטמצם גם בו.

[22] $f\left(x\right):=\sqrt{1-x^{2}}$.

♣ בד"כ כשנשתמש במשפט הפונקציה הסתומה נצטרך להסיק את G מתוך המקום הגאומטרי הנתון לנו ולהוכיח שהיא אכן מקיימת את התנאים.

♣ לכאורה לא הרווחנו דבר - אין לנו שום מושג מיהי אותה f שקיבלנו מהמשפט[23]!
למרות זאת אנחנו יכולים לחשב את הנגזרת שלה בכל נקודה ע"י כלל השרשרת; לכל $x\in V$ מתקיים $G\left(x,f\left(x\right)\right)=0$, א"כ תהא $g:V\to\mathbb{R}^{k}\times\mathbb{R}^{m}$ המוגדרת ע"י $g\left(x\right):=\left(x,f\left(x\right)\right)$ לכל $x\in V$ ומכלל השרשרת נובע כי:

$$0=D\left(G\circ g\right)_{a}=DG_{\left(a,f\left(a\right)\right)}\circ Dg_{a}=DG_{\left(a,b\right)}\circ Dg_{a}$$

כלומר:

$$\begin{aligned}0= & \left[\begin{array}{cccc|c}| & | & & | \\ \partial_{1}G\left(a,b\right) & \partial_{2}G\left(a,b\right) & \cdots & \partial_{k}G\left(a,b\right) & M\\ | & | & & | \end{array}\right]\cdot\left[\begin{array}{c} I_{k}\\ \hline Df_{a} \end{array}\right]\\ = & \left[\begin{array}{cccc} | & | & & |\\ \partial_{1}G\left(a,b\right) & \partial_{2}G\left(a,b\right) & \cdots & \partial_{k}G\left(a,b\right)\\ | & | & & | \end{array}\right]+M\cdot Df_{a}\\ \Rightarrow & Df_{a}=-M^{-1}\cdot\left[\begin{array}{cccc} | & | & & |\\ \partial_{1}G\left(a,b\right) & \partial_{2}G\left(a,b\right) & \cdots & \partial_{k}G\left(a,b\right)\\ | & | & & | \end{array}\right]\end{aligned}$$

♣ במקרה שהבאנו לעיל (גזירת מעגל היחידה) נקבל (לכל $x\in\left(-1,1\right)$):

$$0=D\left(G\circ g\right)_{x}=DG_{\left(x,f\left(x\right)\right)}\circ Dg_{x}=DG_{\left(x,\sqrt{1-x^{2}}\right)}\circ Dg_{x}=\left[\begin{array}{c}2x\end{array}\right]+\left[\begin{array}{c}2\cdot\sqrt{1-x^{2}}\end{array}\right]\cdot Df_{x}$$
$$\Rightarrow Df_{x}=-\left[\begin{array}{c}\frac{1}{2\cdot\sqrt{1-x^{2}}}\end{array}\right]\cdot\left[\begin{array}{c}2x\end{array}\right]=\left[\begin{array}{c}-\frac{x}{2\cdot\sqrt{1-x^{2}}}\end{array}\right]$$

[23]לו היינו יודעים מיהי היה זה משום שהיינו יכולים לצמצם בעצמנו את התחום ולהסיק את ההופכית המקומית.

4.4 כופלי לגראנז'

♣ בפרק הקודם ראינו שכמו באינפי' 1 ניתן למצוא נקודות קיצון מקומיות בקבוצה פתוחה ע"י גזירת הפונקציה והשוואת הנגזרת ל-0, וכן ראינו שניתן למיין אותן ע"פ הנגזרת השנייה.
באינפי' 1 היו לנו לכל היותר שתי נקודות שפה שקל היה לבדוק אם הן מהוות נקודות קיצון, אבל כעת ייתכן שיש לנו קבוצה אין-סופית (ואפילו לא בת-מנייה) של נקודות שפה - ייתכן אפילו שהקבוצה כולה היא נקודות שפה!
כיצד נוכל למצוא את נקודות הקיצון במצב כזה?
הפתרון הוא לחלק את הקבוצה לשני חלקים: את הנקודות החשודות לקיצון בפנים הקבוצה נמצא בדרך הרגילה, ואת הנקודות החשודות לקיצון על השפה נמצא ע"י המשפט הבא; אחרי שמצאנו את כל הנקודות החשודות נצטרך להציב ולבדוק מי מהן היא אכן נקודת קיצון.

משפט 4.10. כופלי לגראנז'[24]

תהא $U\subseteq\mathbb{R}^{k}$ קבוצה פתוחה, יהי $k>m\in\mathbb{N}$ ותהיינה $f,g_{1},g_{2},\ldots,g_{m}\in C^{1}\left(U,\mathbb{R}\right)$.
נסמן:
$$A:=\left\{ \;x\in U\;\middle|\;g_{1}\left(x\right)=g_{2}\left(x\right)=\ldots=g_{m}\left(x\right)=0\;\right\}$$
ותהא $F:A\rightarrow\mathbb{R}^{m+1}$ פונקציה המוגדרת ע"י (לכל $x\in U$):
$$F\left(x\right):=\begin{bmatrix}g_{1}\left(x\right)\\ g_{2}\left(x\right)\\ \vdots\\ g_{m}\left(x\right)\\ f\left(x\right)\end{bmatrix}$$
מהגדרה F היא פונקציה גזירה ברציפות ולכל $x\in A$ מתקיים[25]:
$$DF_{x}=\begin{bmatrix}\nabla g_{1}\left(x\right)\\ \nabla g_{2}\left(x\right)\\ \vdots\\ \nabla g_{m}\left(x\right)\\ \nabla f\left(x\right)\end{bmatrix}$$
לכל נקודה $a\in A$, כך ש-a היא נקודת קיצון מקומית של f ב-A, מתקיים $\text{rk}\left(DF_{a}\right)<m+1$; כלומר הסדרה $\left(\nabla g_{1}\left(a\right),\nabla g_{2}\left(a\right),\ldots,\nabla g_{m}\left(a\right),\nabla f\left(a\right)\right)$ תלויה ליניארית.

♣ בד"כ הפונקציות $g_{1},g_{2},\ldots,g_{m}$ אינן נתונות לנו אלא אנו נצטרך להסיק אותן מתוך הקבוצה שבה אנו רוצים למצוא את נקודות הקיצון, מסיבה זו נוכל לדאוג לכך שהסדרה $\left(\nabla g_{1}\left(a\right),\nabla g_{2}\left(a\right),\ldots,\nabla g_{n}\left(a\right)\right)$ לא תהיה תלויה ליניארית (אם היא כן אז נדלל אותה לסדרה בת"ל), ואז המשפט אומר שאם $x\in A$ היא נקודת קיצון של f ב-A אז קיימים $\lambda_{1},\lambda_{2},\ldots,\lambda_{n}\in\mathbb{R}$[26] כך שמתקיים:
$$\nabla f\left(x\right)=\sum_{i=1}^{m}\lambda_{i}\cdot\nabla g_{i}\left(x\right)$$
כלומר קיבלנו מערכת של $n+k$ משוואות ב-$k+m$ נעלמים[27], ואם נצליח לפתור את המשוואות[28] נוכל לקבוע אלו נקודות ב-A חשודות בהיותן נקודות קיצון.

---

[24] ערך בוויקיפדיה: לגראנז' ז'וזף-לואי.
[25] יש לשחלף את הגרדיאנטים כדי לקבל שאלה הן השורות של DF_{x}.
[26] סקלרים אלה הם הנקראים "כופלי לגראנז'" ועל שמם נקרא המשפט.
[27] k הקואורדינטות של x ו-m ה-λ-ים.
[28] שימו לב שאלה אינן בהכרח משוואות ליניאריות, ולכן אין לנו אלגוריתם לפתירתן.

הוכחה. תהא $a\in A$ נקודת קיצון מקומית של f ב-A ונניח בשלילה ש-$\operatorname{rk}\left(DF_a\right)\geq m+1$, כלומר $\operatorname{rk}\left(DF_a\right)=m+1$. מהיות F גזירה ברציפות נובע שקיימת סביבה של a שבה דרגת הנגזרת מלאה בכל נקודה, תהא B סביבה כזו. ממשפט ההעתקה הפתוחה נובע ש-$F\mid_B$ היא העתקה פתוחה, ובפרט $F\left(B\right)$ היא קבוצה פתוחה. מכאן שקיים $0<\varepsilon\in\mathbb{R}$ כך שמתקיים:

$$\begin{bmatrix}0\\0\\\vdots\\0\\f\left(a\right)\pm\varepsilon\end{bmatrix}=\begin{bmatrix}g_1\left(a\right)\\g_2\left(a\right)\\\vdots\\g_m\left(a\right)\\f\left(a\right)\pm\varepsilon\end{bmatrix}=F\left(a\right)\pm\begin{bmatrix}0\\0\\\vdots\\0\\\varepsilon\end{bmatrix}\in F\left(B\right)$$

כלומר קיים $x\in B\cap A$ כך ש-$f\left(x\right)=f\left(a\right)\pm\varepsilon$ בסתירה לכך ש-a היא נקודת מקסימום/מינימום מקומית של f ב-A.
מכאן שהנחת השלילה אינה נכונה ו-$\operatorname{rk}\left(DF_a\right)<m+1$. ■

קיימת הוכחה נוספת המסתמכת על משפט הפונקציה הסתומה.