2/11/2016

תרגיל 2

גזירות:

פונקציה $f:\mathbb{R}^n \mapsto \mathbb{R}^k$ נקראת גזירה ב-$p \in \mathbb{R}^n$ אם קיים אופרטור ליניארי $L:\mathbb{R}^n \mapsto \mathbb{R}^k$ כך שמתקיים:

$$\lim_{\Delta p \to 0} \frac{\|f(p+\Delta p) - f(p) - L(\Delta p)\|}{\|\Delta p\|} = 0$$

האופרטור L נקרא הדיפרנציאל של f ומתקיים:

$$L_p = Df|_p = \left(\frac{\partial f_i}{\partial x_j}(p_i)\right)_{i,j}$$

אם $f:\mathbb{R}^n \mapsto \mathbb{R}$ אזי:

$$Df|_p = \left.\left(\frac{\partial f}{\partial x_j}\right)\right|_p = \nabla f(p)^T$$

ובמקרה זה מקבלים את הגדרת הדיפרנציאל כפי שהכרנו אותה מאינפי II.

נרחיב את הגדרת הגזירות בכתיב מטריציוני:

$$f:\mathbb{R}^n \mapsto \mathbb{R}^k \quad f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_k(x) \end{pmatrix}$$

אזי מתקיים:

$$R(\Delta p) = f(p+\Delta p) - f(p) - L(\Delta p)$$

$$= \begin{pmatrix} f_1(p+\Delta p) \\ \vdots \\ f_k(p+\Delta p) \end{pmatrix} - \begin{pmatrix} f_1(p) \\ \vdots \\ f_k(p) \end{pmatrix} - \begin{pmatrix} \nabla f_1(p)^T \\ \vdots \\ \nabla f_k(p)^T \end{pmatrix}(\Delta p) = \begin{pmatrix} f_1(p+\Delta p) - f_1(p) - \nabla f_1(p)^T \Delta p \\ \vdots \\ f_k(p+\Delta p) - f_k(p) - \nabla f_k(p)^T \Delta p \end{pmatrix}$$

כלומר:

$$\lim_{\Delta p \to 0} \left\|\frac{R(\Delta p)}{\|p\|}\right\| = 0 \Leftrightarrow \lim_{\Delta p \to 0} \frac{R(\Delta p)}{\|\Delta p\|} = 0$$

אך נשים לב, כי הנ"ל נכון אם ורק אם לכל $1 \le i \le k$ מתקיים:

$$\lim_{\Delta p \to 0} \frac{f_i(p+\Delta p) - f_i(p) - \nabla f_i(p)^T \Delta p}{\|\Delta p\|} = 0$$

2.1 מסקנה - $f:\mathbb{R}^n \mapsto \mathbb{R}^k$ גזירה אם ורק אם $f_i:\mathbb{R}^n \mapsto \mathbb{R}$ גזירה לכל $1 \le i \le k$.

2.1.2 מסקנה – בפרט, אם ל-f יש נ"ח רציפות, אזי f גזירה. ההיפך, לא בהכרח נכון.

<u>דוגמאות:</u>

א. נניח $f:\mathbb{R}^3 \mapsto \mathbb{R}^2$ נתונה על ידי:

$$f(x,y,z) = \left(\underbrace{x^2y - z^4}_{f_1}, \underbrace{xy^3z}_{f_2}\right)$$

במקרה זה קל לראות כי f בעלת נגזרות חלקיות רציפות ולכן היא גזירה. כמו כן, $Df \in \mathbb{R}^{2\times 3}$ ומתקיים:

$$Df|_p = \begin{pmatrix} 2xy & x^2 & -4z^3 \\ y^3z & 3xy^2z & xy^3 \end{pmatrix}\Bigg|_p$$

לדוגמה, עבור הנקודה $p = (1,4,2)$ מתקיים:

$$Df|_{(1,4,2)} = \begin{pmatrix} 8 & 1 & -32 \\ 128 & 96 & 64 \end{pmatrix}$$

ב. בהרצאה עסקנו בפונקציות האפיניות, כלומר פונקציות $f\colon \mathbb{R}^n \mapsto \mathbb{R}^k$ הנתונה על ידי:

$$f(x) = Ax + b \quad \begin{matrix} A \in \mathbb{R}^{k\times n} \\ b \in \mathbb{R}^k \end{matrix}$$

במקרה זה, מתקיים:

$$Df|_p = A$$

2.2 הערה – פונקציה ריבועית בכתיב מטריציוני ניתנת לייצוג על ידי:

$$\boldsymbol{f}\colon \mathbb{R}^{\boldsymbol{n}} \mapsto \mathbb{R} \quad \boldsymbol{f(x) = x^TQx + b^tx + c}$$

כאשר $\boldsymbol{Q} \in \mathbb{R}^{\boldsymbol{n\times n}}$ סימטרית, וכן $\boldsymbol{b} \in \mathbb{R}^{\boldsymbol{n}}, \boldsymbol{c} \in \mathbb{R}$.

דוגמה:

$$f(x,y) = x^2 + 2xy + 4y^2 + 3x + 5y + 1 = (x\ y)\begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + (3\ 5)\begin{pmatrix} x \\ y \end{pmatrix} + 1$$

נחשב נגזרות חלקיות:

$$b^Tx = \sum_{i=1}^n b_ix_i \quad x^T(Qx) = \sum_{i=1}^n x_i(Qx_i) = \sum_{i=1}^n x_i\left(\sum_{j=1}^n q_{ij}x_j\right)$$

$$= \sum_{i,j=1}^n q_{ii}x_i^2 + \sum_{\substack{i,j=1 \\ i<j}}^n 2q_{ij}x_ix_j$$

וסה"כ נקבל כי:

$$f(x) = \sum_{i=1}^n q_{ii}\,x_i^2 + \sum_{\substack{i,j=1 \\ i<j}}^n 2q_{ij}x_ix_j + \sum_{i=1}^n b_ix_i + C$$

בהצגה כזו נוכל לחשב נגזרות חלקיות ולקבל כי:

$$\frac{\partial f}{\partial x_i} = 2q_{ii}x_i + \sum_{\substack{j=1 \\ j\neq i}}^n 2q_{ij}x_j + b_i = 2\sum_{j=1}^n q_{ij}x_j + b_i = 2(Qx)_i + b_i = (2Qx + b)_i$$

ולכן נקבל כי:

$$\boxed{\vec{\nabla} f(p) = 2Qp + b}$$

ג. תהא $f: \mathbb{R}^n \setminus \{0\} \mapsto \mathbb{R}^n$ הנתונה על ידי $f(x) = \frac{x}{\|x\|_2}$. נחשב את $Df \in \mathbb{R}^{n\times n}$
נחשב נגזרות חלקיות ונקבל:

$$f(x) = \left(\frac{x_1}{\|x\|_2}, \cdots, \frac{x_n}{\|x\|_2}\right)$$

עבור $j \neq i$ נקבל:

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\frac{x_i}{\|x\|_2}\right) = -\frac{x_i}{\|x\|_2^2}\frac{\partial}{\partial x_j}(\|x\|_2) = -\frac{x_i}{\|x\|_2^2}\frac{1}{2\|x\|_2}2x_j = -\frac{x_i x_j}{\|x\|_2^3}$$

אך עבור $i = j$ נקבל כי:

$$\frac{\partial f_i}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\frac{x_i}{\|x\|_2}\right) = \frac{\|x\|_2 - x_i\frac{x_i}{\|x\|_2}}{\|x\|_2^2} = \frac{\|x\|_2^2 - x_i^2}{\|x\|_2^3}$$

וכאמור, כל הנגזרות החלקיות שקיבלנו רציפות על $\mathbb{R}^n \setminus \{0\}$ ולכן ממסקנה 2.1.1, נסיק כי הפונקציה גזירה כנדרש, ומתקיים:

$$\boxed{Df|_x = -\frac{1}{\|x\|_2^3}(x_i x_j)_{i,j} + \frac{1}{\|x\|_2} I = \frac{1}{\|x\|_2} - \frac{1}{\|x\|_2^3}(x \cdot x^T)}$$

ד. תהא $f: \mathbb{R}^{n\times n} \mapsto \mathbb{R}^{n\times n}$ המוגדרת על ידי $f(A) = A^2$. נחשוב על f כעל $f: \mathbb{R}^{n^2} \mapsto \mathbb{R}^{n^2}$. נראה כי f גזירה. נשתמש בכך שאם L קיים אז הוא יחיד וננסה לכתוב:

$$f(A+H) - f(A) = L(H) + o(h)$$

כאשר, כאמור, $\lim_{H\to 0}\frac{o(H)}{\|H\|} = 0$. אם הצלחנו אזי קיבלנו את ה-L הדרוש.

$$f(A+H) - f(A) = A^2 + A\cdot H = H \cdot A + H^2 - A^2 = (A\cdot H + H \cdot A) + H^2$$

קל לראות כי H^2 הינו אופרטור שמתאים[1] להיות $o(H)$ ואכן הסוגריים הינם אופרטור ליניארי ב-H. מכאן שקיבלנו את הדרוש, ו-L הוא האופרטור:

$$L = \mathbb{R}^{n\times n} \mapsto \mathbb{R}^{n\times n} \quad L(H) = A\cdot H + H \cdot A$$

<u>תרגיל:</u>

תהא $f: \mathbb{R}^n \mapsto \mathbb{R}$ פונקציה סקלארית בעלת נגזרת חלקית בכדור:

$$B(p,r) = \{p' | \|p - p'\|_\infty < r\}$$

הראו כי לכל $p + \Delta p \in B(p,r)$ קיימות $q_1, q_2, \cdots, q_n \in B(p,r)$ כך שמתקיים:

$$f(p+\Delta p) - f(p) = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(q^i)\Delta p_i$$

<u>פתרון:</u>

נפתור עבור $n = 2$:

$$p = (x,y) \quad p + \Delta p = (x + \Delta x, y + \Delta y)$$

[1] מתקיים, באופן כללי $\left\|\frac{H^2}{\|H\|}\right\| = \frac{\|H^2\|}{\|H\|} = \|H\| \to 0$ ולכן H^2 הוא אכן $o(H)$ כנדרש.

<u>הרעיון</u> – נקבע ערך של y ונתבונן על f כפונקציה במשתנה יחיד, $f(\cdot, y_0): \mathbb{R} \mapsto \mathbb{R}$.

במקרה כזה, נרצה להשתמש במשפט לגרנז' עבור הקטע $[x, x+\Delta x]$. לשם כך נזכור שאמנם f אינה גזירה, אך הנגזרות החלקיות שלה אכן קיימות, ולכן הפונקציה החדשה, במשתנה יחיד, כן גזירה ולכן רציפה בתחום, כלומר תנאי משפט לגרנז' אכן מתקיים ונסיק כי קיימת נקודה $x^\star \in [x, x+\Delta x]$ כך שמתקיים:

$$f(x+\Delta x, y_0) - f(x, y_0) = \frac{\partial f}{\partial x}(x^*, y_0)\Delta x$$

נחזור לפונקציה המקורית שלנו, ונשים לב כי:

$$f(p+\Delta p) - f(p) = f(x+\Delta x, y+\Delta y) - f(x,y) \star$$

נוכל לכתוב סכום זה באופן הבא:

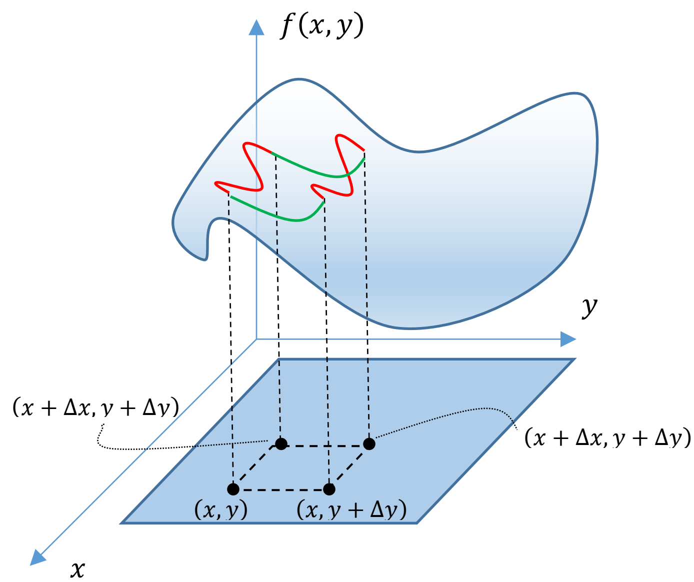
$$\star = f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y) + f(x, y+\Delta y) - f(x,y)$$

נשים לב, עתה, כי כל זוג מחוסרים בביטוי זה מהווה בדיוק את המקרה שבו דנו קודם לכן עבור המרת הפונקציה לפונקציה במשתנה יחיד. לכן נסיק כי קיימות $q_1, q_2 \in \mathbb{R}$ כך שמתקיים:

$$f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y) = \frac{\partial f}{\partial x}(q_1)\Delta x \quad f(x, y+\Delta y) - f(x,y) = \frac{\partial f}{\partial y}(q_2)\Delta y$$

ולכן סה"כ נקבל כי:

$$f(p+\Delta p) = \sum_{i=1}^{2} \frac{\partial f}{\partial x_i}(q_i)\Delta x_i$$



איור 1 – תיאור סכמטי של הבעיה

<u>הערה</u> – אם f גזירה, אזי קיימת נקודה q (יחידה) בקטע המחבר את $p, p+\Delta p$ כך שמתקיים:

$$f(p+\Delta p) - f(p) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(q)\Delta \mathrm{p_i}$$

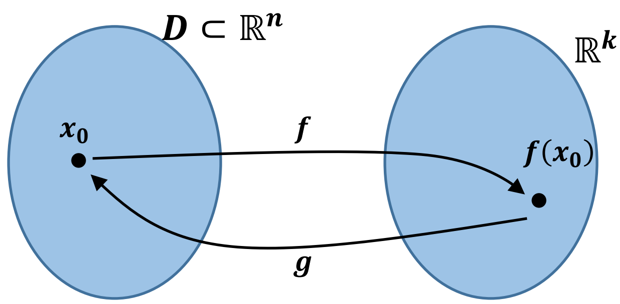
2.3 משפט – כלל השרשרת – בהנתן $f:\mathbb{R}^n \mapsto \mathbb{R}^k$ גזירה ב-p_0, וכן $g:\mathbb{R}^k \mapsto \mathbb{R}^m$ גזירה ב-$f(p_0)$, אזי מסיקים כי $g \circ f:\mathbb{R}^n \mapsto \mathbb{R}^m$ גזירה ב-p_0 ומתקיים:

$$\boldsymbol{D(g \circ f)|_{p_0} = Dg|_{f(p_0)} \cdot Df|_{p_0} \in \mathbb{R}^{m \times n}}$$

<u>תרגיל:</u>

תהא $f:D \mapsto \mathbb{R}^k$ כאשר נתון כי $D \subset \mathbb{R}^n$ קבוצה פתוחה וכן ש- f גזירה והפיכה. תהא $x_0 \in D$ כך שההופכית גזירה בסביבת $f(x_0)$. הראו כי $n = k$.

<u>פתרון:</u>



נסמן $g:\mathbb{R}^k \mapsto D$ את ההופכית של f.

נשים לב כי מתקיים, אם כן:

$$f \circ g = (id)_{\mathbb{R}^k}$$
$$g \circ f = (id)_D$$

כמו כן, $g \circ f$ היא הרכבה של פונקציות גזירות ולכן גזירה בסביבת $f(x_0)$.

$$I_k = D(id)_{\mathbb{R}_k} = D(f \circ g)|_{f(x_0)} = Df|_{x_0} \cdot Dg|_{f(x_0)}$$

נסמן, אם כן:

$$k = \operatorname{rank} I_k = \operatorname{rank}\left(Df|_{x_0} \cdot Dg|_{f(x_0)}\right) \overset{\text{דרגת כפל מטריציות}}{\le} \min\left\{\operatorname{rank} Df|_{x_0}, \operatorname{rank} Dg|_{f(x_0)}\right\}$$

מצד שני, נקבל כי הנ"ל נכון גם במקרה של ההרכבה ההפוכה, כלומר על דרגת $g \circ f$, ולכן נסיק כל כל אחת מהדרגות קטנה או שווה למינימום בין שתיהן ובפרט שוות אחת לשניה, כנדרש.