

הרצאה 14: מיון נקודות קריטיות, שיטת כופלי לגרנז'

יוחאי מעין

הערות

- המסמך הזה מכיל הערכה לא סופית ולא מחייבת של החומר שיועבר בהרצאות.
- יתכנו חוסרים, שגיאות וכל מיני אי התאמות אחרות ביחס למה שיועבר בפועל.
- הוכחות ינתנו בהרצאות. יש להכיר ולהבין אותן. פירוט נוסף בדוגמאות ובכלל ינתן בהרצאות.
- הגדרות מרכזיות ומשפטים מרכזיים מובלטים במסגרת צבעונית. יש להכיר היטב גם את ההגדרות ואת המשפטים שאינם מרכזיים.
- החומר המחייב הוא החומר אשר מועבר בהרצאות בפועל.
- דיווחים על טעויות במייל yohaim@tauex.tau.ac.il יעזרו מאוד ויתקבלו בשמחה.

תוכן העניינים

2	פולינום טיילור ובעיות קיצון (המשך)	1
2	אפיון של מוגדרות חיובית ועוד	1.1
4	מיון של נקודות קריטיות	1.2
6	שיטות לפתרון בעיות קיצון תחת אילוץ	1.3

When you're trying to solve a problem using
Lagrange multipliers but it isn't working.



1 | פולינום טיילור ובעיות קיצון (המשך)

1.1 אפיון של מוגדרות חיובית ועוד

אפיון של תבנית מוגדרת חיובית

המשפט הבא יאפשר שני דברים.

ראשית, לקבוע בקלות אם ההסיאן מוגדר חיובית.

שנית, להוכיח שאז יש מינימום מקומי.

לכן בעתיד נצטרך רק לבדוק את התנאי הפשוט שמיד נציג במשפט, ואם הוא יתקיים - נדע שמצאנו נקודת מינימום מקומי.

משפט 1: תהי H מטריצה סימטרית ממשית. התנאים הבאים שקולים:

1. H מוגדרת חיובית.

2. כל הערכים העצמיים של H חיוביים.

3. קיים $c > 0$ כך ש- $H\vec{u} \cdot \vec{u} \geq c|\vec{u}|^2$ לכל $\vec{u}$.

4. קריטריון סילבסטר: המינורים הראשיים של H כולם חיוביים -

$$\Delta_1 > 0, \Delta_2 > 0, \dots, \Delta_n > 0$$

המינורים הראשיים הם הדטרמיננטות של תתי-המטריצות הבאות.

עבור מטריצה 2×2 , יש שני מינורים ראשיים Δ_1 ו- Δ_2 . אלו הם הדטרמיננטות של הבלוקים הצבועים בכחול:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \rightsquigarrow \Delta_1 = a_{11}, \quad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \rightsquigarrow \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

עבור מטריצה 3×3 , באופן דומה:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \rightsquigarrow \Delta_1 = a_{11}, \quad \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \rightsquigarrow \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \rightsquigarrow \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

נזכיר לבסוף כי את הקשרים השימושיים

$$\det(H) = \lambda_1 \cdots \lambda_n, \quad \text{tr}(H) = \lambda_1 + \cdots + \lambda_n$$

הערה: תנאי 2 עובד באופן דומה עבור מוגדרת שלילית (ערכים עצמיים שליליים), מוגדרת אי שלילית (ערכים עצמיים אי שליליים) ומוגדרת אי חיובית (ערכים עצמיים אי חיוביים).

תנאי 3 עובד באופן דומה עבור מוגדרת שלילית ($H\vec{u} \cdot \vec{u} \leq -c|u|^2$), אבל לא עבור מוגדרת אי שלילית או אי חיובית.

לגבי תנאי 4, נדון בכך מיד.

עוד על קריטריון סילבסטר

עבור מוגדרות שלילית, קריטריון סילבסטר גם עובד. H מוגדרת שלילית אם"ס למינורים הראשיים יש בדיוק את הסימנים המתחלפים הבאים:

$$\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0$$

(וכו' אם המטריצה היא $n \times n$, $n > 3$).

קל לזכור את התנאים הללו על ידי שחזור שלהם מתוך מטריצה אלכסונית מוגדרת חיובית (כל איברי האלכסון חיוביים) או מוגדרת שלילית (כל איברי האלכסון שליליים).

קריטריון סילבסטר **לא עובד** לבדיקה של מוגדרות אי שלילית או אי חיובית.

לא נוכיח את קריטריון סילבסטר.

דוגמאות לבדיקת התנאים

דוגמאות:

1. ניקח את התבנית $h(x, y) = x^2 + y^2$ עבורה

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

הערכים העצמיים הם 1, 1 (כלומר, 1 עם ריבוי אלגברי 2). שניהם חיוביים ולכן המטריצה והתבנית מוגדרות חיובית.

2. ניקח את התבנית $h(x, y) = x^2 - 2xy + y^2$ עבורה

$$H = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

נבדוק את קריטריון סילבסטר.

$$\begin{cases} \Delta_1 = 1 > 0 \\ \Delta_2 = \det(H) = 1 - 1 = 0 \end{cases}$$

זה כישלון של הקריטריון, כלומר לא ניתן להסיק דבר.

מצד שני, הערכים העצמיים הם 2, 0 ולכן התבנית והמטריצה מוגדרות אי שלילית.

3. עבור המטריצה $H = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ מתקיים

$$\Delta_1 = 2 > 0, \quad \Delta_2 = \begin{vmatrix} 2 & -1 \\ -1 & 1 \end{vmatrix} = 1 > 0$$

לכן H מוגדרת חיובית.

4. עבור המטריצה $H = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & -2 \\ 0 & -2 & -1 \end{pmatrix}$ מתקיים כנ"ל ש-

$$\Delta_1 = 2 > 0, \quad \Delta_2 = 1 > 0$$

אבל:

$$\Delta_3 = \begin{vmatrix} 2 & -1 & 0 \\ -1 & 1 & -2 \\ 0 & -2 & -1 \end{vmatrix} = -9 < 0$$

לכן המטריצה אינה מוגדרת חיובית.

5. עבור המטריצה $H = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 0 \\ -1 & -0 & -1 \end{pmatrix}$ מתקיים כנ"ל ש-

$$\Delta_1 = 1 > 0, \quad \Delta_2 = -3 < 0, \quad \Delta_3 = 2 > 0$$

לכן המטריצה אינה מוגדרת חיובית ואינה מוגדרת שלילית.

כאן $\det(H) \neq 0$ ולכן כל הערכים העצמיים שונים מאפס. לכן חלק מהערכים העצמיים חיוביים וחלקם שליליים.

1.2 מיון של נקודות קריטיות

תנאי מספיק לקיצון מקומי

נוכל כעת להסיק את המשפט הבא, שאליו חתרנו.

משפט 2. תהי f פונקציה גזירה פעמיים ברציפות בתחום D ותהי $P \in D$ נקודה קריטית של f . אם ההסיאן $H_f(P)$ מוגדר חיובית, אז P היא נקודת מינימום מקומי.

באופן דומה, אם $H_f(P)$ מוגדר שלילית, אז P נקודת מקסימום.

תנאי הכרחי לקיצון מקומי

החיסרון של משפט 2 הוא שהוא לא מהווה תנאי הכרחי. כלומר, אם H לא מוגדר חיובית, עדיין יתכן מינימום.

המשפט הבא נותן לנו תנאי הכרחי שהוא לא רחוק מאוד ממה שראינו: ההסיאן חייב להיות מוגדר אי-שלילית.

משפט 3. תהי f פונקציה גזירה פעמיים ברציפות בתחום D ותהי $P \in D$ נקודה קריטית של f . אם P היא נקודת מינימום מקומי של f , אז ההסיאן $H_f(P)$ מוגדר אי-שלילית.

לכן, אם להסיאן יש ערכים עצמיים עם סימנים שונים, אז הנקודה היא אוקף (ע"ע חיובי פוסל מוגדרות אי-חיובית ולהפך).

כדי להוכיח את המשפט, נזדקק לתרגיל הפשוט הבא שהוא יישום של כלל השרשרת: מכאן ניתן להסיק שאם ההסיאן אינו מוגדר אי שלילית ואינו מוגדר אי חיובית בנקודה קריטית נתונה, אז הנקודה אינה מינימום ואינה מקסימום ועל כן היא אוקף.

דוגמאות למיון של נקודות קיצון

1. נתונה $f(x, y) = x^2 + xy$ אז

$$\vec{\nabla} f = (2x + y, x) = \vec{0} \iff x = y = 0$$

נחשב את ההסיאן בנקודה הקריטית:

$$H_f(0, 0) = \begin{pmatrix} 2 & 1 \\ 2 & 0 \end{pmatrix}$$

לכן

$$\Delta_1 = 2 > 0, \quad \Delta_2 = -1 < 0$$

נסיק שההסיאן הפיך, כלומר, כל הערכים העצמיים שלו שונים מאפס. מצד שני, הדטרמיננטה שלילית, ולכן יש לשני הערכים העצמיים סימנים הפוכים. לכן מדובר באוקף.

2. נתונה $f(x, y, z) = x^2y - x^3 + 2x^2 + xy - 3y^2 - 2xz + 6yz + z^2$ אז

$$\vec{\nabla} f = (2xy - 3x^2 + 4x + y - 2z, x^2 + x - 6y + 6z, -2x + 6y + 2z)$$

ולכן הנקודות הקריטיות הן הפתרונות של מערכת המשוואות

$$\begin{cases} 2xy - 3x^2 + 4x + y - 2z = 0 \\ x^2 + x - 6y + 6z = 0 \\ -2x + 6y + 2z = 0 \end{cases}$$

לא נמצא את כולן, אבל נשים לב כי $(0, 0)$ היא נקודה קריטית ובה

$$H_f(0, 0, 0) = \begin{pmatrix} 4 & 1 & -2 \\ 1 & -6 & 6 \\ -2 & 6 & 2 \end{pmatrix}$$

לכן

$$\Delta_1 = 4 > 0, \quad \Delta_2 = -25 < 0, \quad \Delta_3 = -194 \neq 0$$

לכן מדובר באוקף.

נשים לב שחשוב לוודא כי $\Delta_3 \neq 0$, אחרת קריטריון סילבסטר אינו בתוקף ויש לבדוק ישירות מהם הערכים העצמיים האחרים (בנוסף ל- $\lambda = 0$ במקרה זה).

סיכום המקרים של מיון נקודות קריטיות

האינטואיציה היא שערכים עצמיים ווקטורים מצביעים על כיווני עליה וירידה, אבל ערך עצמי 0 הוא בבחינת קלף לא ידוע.

בהתאם לכך, אם ל- $f(x, y)$ יש נקודה קריטית (x_0, y_0) ו- $H = H_f(x_0, y_0)$, אז:

1. אם $\det H > 0$, הנקודה היא קיצון מקומי, וניתן לקבוע מינימום/מקסימום לפי הסימן של

$$H_{11} = f_{xx}(x_0, y_0)$$

(או, באופן שקול, לפי $f_{yy}(x_0, y_0)$).

2. אם $\det H < 0$, הנקודה היא אוקף.

3. אם $\det H = 0$, יש ע"ע 0, ולכן לא ניתן לקבוע - אבל ניתן לפסול מקרה לפי הסימן של הע"ע השני.

עבור פונקציה בשלושה משתנים, אם ל- $f(x, y, z)$ יש נקודה קריטית (x_0, y_0, z_0) ו- $H = H_f(x_0, y_0, z_0)$:

1. אם $\det H \neq 0$, ניתן לקבוע את סוג הנקודה לפי קריטריון סילבסטר (אם זה לא מתאים בדיוק לאחד משני המקרים, אז אוקף).

2. אם $\det H = 0$, יש למצוא את שני הערכים העצמיים האחרים. אין להשתמש בקריטריון סילבסטר.

• אם שני הע"ע האחרים בעלי אותו סימן, ניתן רק לפסול את אחד המקרים.

• אם יש להם סימנים שונים, הנקודה היא אוקף.

מוגדרות ההסיאן	ערכים עצמיים	סילבסטר	קיצון מקומי
חיובית	חיוביים	$\Delta_1 > 0, \Delta_2 > 0, \Delta_3 > 0, \dots$	מינימום
שלילית	שליליים	$\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots$	מקסימום
אי שלילית	אי שליליים	אין	לא מקסימום
אי חיובית	אי חיוביים	אין	לא מינימום
אין	לפחות אחד שלילי ואחד חיובי	לפעמים עוזר	אוקף

1.3 שיטות לפתרון בעיות קיצון תחת אילוץ

הצגת הבעיה של מציאת קיצון תחת אילוץ

נתונה $f(x, y)$, ונניח שאנחנו רוצים למצוא את $\max_{(x,y) \in D} f(x, y)$ כאשר D תחום.

היות שבאופן כללי, גם אם f רציפה, אין בהכרח מקסימום או מינימום בתחום (שהוא קבוצה פתוחה), הבעיה הנכונה שתמיד יש לה פתרון אם f רציפה ו- D חסום היא

$$\max / \min f(x, y), \quad (x, y) \in \overline{D}$$

דוגמאות:

1. אם $D = \{x^2 + y^2 < 1\}$, במעבר לסגור $\bar{D} = \{x^2 + y^2 \leq 1\}$ נוספת השפה

$$x^2 + y^2 = 1$$

2. אם $D = \{x^2 + y^2 + z^2 < 1\}$, במעבר לסגור $\bar{D} = \{x^2 + y^2 + z^2 \leq 1\}$ נוספת השפה

$$x^2 + y^2 + z^2 = 1$$

3. אם $D = \{(x, y) \mid y > 0\}$, במעבר לסגור $\bar{D} = \{(x, y) \mid y \geq 0\}$ נוספת השפה

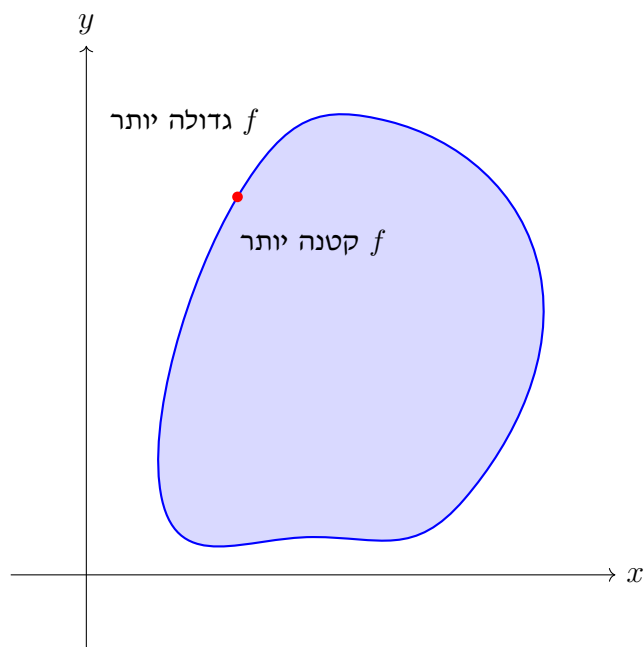
$$y = 0$$

כלומר, ציר ה- x .

שיטת הפתרון: אוספים נקודות חשודות, מציבים אותן ב- $f(x, y)$ אחת-אחת, ובודקים מתי f הכי גדולה/קטנה.

הבעיה: אנחנו יודעים באילו נקודות פנימיות יתכן קיצון - אלו הנקודות הקריטיות. אבל נכון לכרגע, בכל נקודה על השפה יתכן קיצון. אין לנו אפשרות להציב את כולן ל- $f(x, y)$ ולהשוות. לכן עלינו למצוא דרך למצוא נקודות חשודות על שפת התחום ∂D , שהיא לא תחום בפני עצמו (לא פתוחה).

הנקודה החשובה היא שיתכן מינימום על השפה מבלי ש- $\vec{\nabla} f = \vec{0}$. לכן אנחנו צריכים תנאי אחר שיאתר נקודות קיצון.



כדי להמשיך, נניח ש- ∂D ניתנת לתיאור על ידי משוואה: $g(x, y) = 0$ (וכנ"ל במקרה של שלושה משתנים).

דוגמה: אם $D = \{x^2 + y^2 < 1\}$, עיגול היחידה, אז השפה היא מעגל היחידה. היא ניתנת לתיאור על ידי המשוואה $x^2 + y^2 = 1$. כלומר, $g(x, y) = x^2 + y^2 - 1$.

למעשה, אנחנו נתעניין בבעיה קצת יותר כללית: $g(x, y) = 0$ לא תהיה חייבת לתאר את השפה של תחום, אלא איזשהו אילוף על הנקודות (x, y) :

$$\begin{cases} \max / \min & f(x, y) \\ \text{s.t.} & g(x, y) = 0 \end{cases} \quad (1)$$

הערה: אם $g(x, y)$ פונקציה רציפה, אז הקבוצה $g(x, y) = 0$ תמיד סגורה אבל לאו דווקא חסומה. זה חשוב מבחינת שימוש במשפט וירטואלי.

שיטת ההצבה לפתרון בעיות קיצון תחת אילוף

דוגמה: שיטת ההצבה: נפתור את הבעיה

$$\begin{cases} \max / \min & x^2 + y^2 \\ \text{s.t.} & x + y = 0 \end{cases}$$

נצמצם את הפונקציה לקבוצה שהאילוף מתאר.

כלומר, נחלק $y = -x$ מתוך תנאי האילוף ונציב לפונקציה:

$$h(x) := f(x, -x) = 2x^2, \quad x \in \mathbb{R}$$

לפונקציה $h(x)$ (שהיא f תחת האילוף) אין מקסימום $(h(x) \xrightarrow{x \rightarrow \infty} \infty)$ ויש לה מינימום גלובלי ב- $x = 0$. בנוסף,

$$h'(x) = 4x = 0 \iff x = 0$$

ולכן אין נקודות קיצון מקומי מלבד $x = 0$.

הערה: בניגוד למצב הפשוט יותר שבו $g(x, y) = 0$ מתאר שפה של תחום חסום ולכן קבוצה קומפקטית, באופן כללי לא מובטח שקיים מקסימום או מינימום בבעיה (1). לכן יש להוכיח קיום של כל אחד מהמקרים לפני שמסיקים שהנקודה שנתנה את התוצאה הגבוהה/נמוכה ביותר מתוך החשודות היא המקסימום/מינימום בהתאמה.

שיטת ההצבה טובה כאשר ניתן לחלק את אחד המשתנים מהאילוף $g(x, y) = 0$ (או $g(x, y, z) = 0$) כפונקציה של האחרים, אבל לעיתים קרובות זה לא אפשרי (או מאוד מסובך) מבחינה מעשית. הבעיה מחמירה כאשר רוצים לחקור קיצון תחת זוג אילוצים במקרה התלת-ממדי.

משפט כופלי לגרנז'

נניח ש- $P(x_0, y_0)$ היא מינימום של $f(x, y)$ תחת האילוף $g(x, y) = 0$.

נסמן ב- $\hat{u}$ כיוון משיק לעקום $g(x, y) = 0$. מבחינה אינטואיטיבית, הפונקציה f צריכה לגדול כאשר זזים מ- P בכיוון $\hat{u}$ (כי זה כיוון התנועה על הישר, ועל הישר P היא מינימום מקומי).

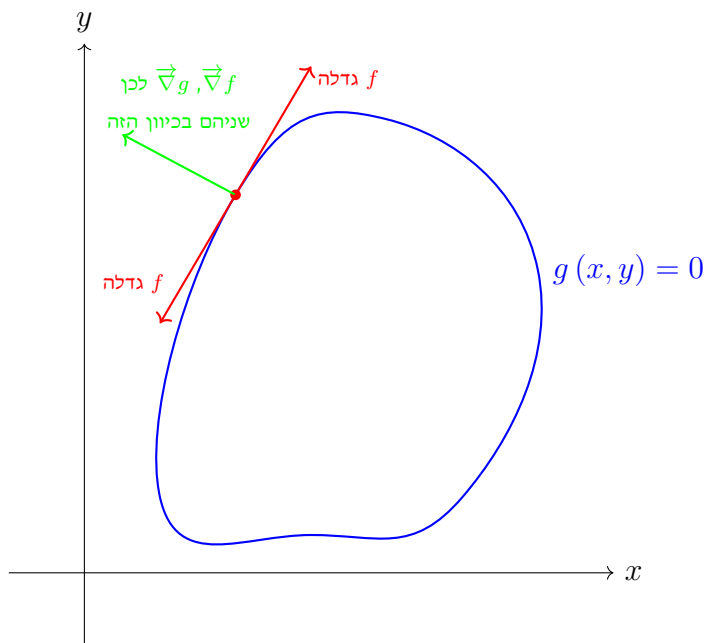
$$\frac{\partial f}{\partial \hat{u}} \geq 0 \quad \text{לכן צריך להתקיים}$$

באופן דומה, f צריכה לגדול גם כאשר זזים מ- P בכיוון $-\hat{u}$ ולכן $-\frac{\partial f}{\partial \hat{u}} \geq 0$. בסך הכל,

$$0 = \frac{\partial f}{\partial \hat{u}}(P) = \vec{\nabla} f(P) \cdot \hat{u}$$

כלומר, $\vec{\nabla} f(P)$ ניצב לוקטור המשיק לעקום $g(x, y) = 0$.

מצד שני, ידוע ש- $\vec{\nabla} g(P)$ ניצב לעקום הנקודה P . לכן $\vec{\nabla} f(P) \parallel \vec{\nabla} g(P)$.



זו תכלית המשפט הבא, עד כדי תנאים טכניים: **משפט כופלי לגרנז'.**

משפט 4. נניח כי $f, g \in C^1$ בסביבת הנקודה P , ונניח ש- P היא נקודת קיצון מקומי של f בכפוף לאילוץ $g(x, y) = 0$. אם $\vec{\nabla} g(P) \neq \vec{0}$ אז קיים $\lambda \in \mathbb{R}$ כך ש-

$$\vec{\nabla} f(P) = \lambda \vec{\nabla} g(P)$$

לכן, נקודות חשודות שמקיימות את האילוץ:

$$1. \text{ קיים } \lambda \in \mathbb{R} \text{ עבורו } \vec{\nabla} f = \lambda \vec{\nabla} g$$

$$2. \vec{\nabla} g = 0$$

המשפט גם נכון במרחב, בדיוק עם אותו נוסח:

משפט 5: נניח כי $f, g \in C^1$ בסביבת הנקודה P , ונניח ש- P היא נקודת קיצון מקומי של f בכפוף לאילוץ $g(x, y, z) = 0$. אם $\vec{\nabla} g(P) \neq \vec{0}$ אז קיים $\lambda \in \mathbb{R}$ כך ש-

$$\vec{\nabla} f(P) = \lambda \vec{\nabla} g(P)$$

משפט כופלי לגרנז' עבור שני אילוצים ופונקציות לגרנז'

משפט כופלי לגרנז' ניתן להכללה לשני אילוצים במרחב:

משפט 6: נניח כי $f, g, h \in C^1$, בסביבת הנקודה P , ונניח ש- P היא נקודת קיצון מקומי של f בכפוף לאילוץ

$$\begin{cases} g(x, y, z) = 0 \\ h(x, y, z) = 0 \end{cases}$$

אם $\vec{\nabla} g(P) \nparallel \vec{\nabla} h(P)$ אז קיימים $\lambda, \mu \in \mathbb{R}$ כך ש-

$$\vec{\nabla} f(P) = \lambda \vec{\nabla} g(P) + \mu \vec{\nabla} h(P)$$

הערה: נהוג להגדיר פונקציית עזר, לפי המקרה:

$$\begin{aligned} \text{אילוץ אחד במישור} \quad F(x, y, \lambda) &= f(x, y) - \lambda g(x, y) \\ \text{אילוץ אחד במרחב} \quad F(x, y, z, \lambda) &= f(x, y, z) - \lambda g(x, y, z) \\ \text{שני אילוצים במרחב} \quad F(x, y, z, \lambda, \mu) &= f(x, y, z) - \lambda g(x, y, z) - \mu h(x, y, z) \end{aligned}$$

ואז התנאי על קיום λ (ו- μ) מסתכם ל- $\vec{\nabla} F = \vec{0}$ (הגרדיאנט לפי כל המשתנים של F).

גם כאן, אין להשתמש בהסיאן כדי למיין את הנקודות החשודות לפי התנהגות מקומית.

לפעם הבאה

מומלץ לעשות חזרה קצרה על המושגים הבאים לפני ההרצאה הבאה:

1. חלוקות וסכומי רימן עבור הגדרה של אינטגרל רימן בקטע $\int_a^b f(x) dx$.

“ Knowledge isn't free. You have to pay attention.

”

Richard P. Feynman