



המחלקה למתמטיקה
Department of Mathematics

פרויקט מסכם לתואר בוגר במדעים (B.Sc)
במתמטיקה שימושית

שימוש בהתמרות מחסום משופרות
לפתרון בעיות אופטימיזציה מאולצת לא לינארית

מריה צ'ברקוב

The Modified Barrier Functions for Nonlinear Constrained
Optimization Problems

Maria Chvarkov

Advisor:

Dr. Fiana Jacobzon

מנחה:

ד"ר פיאנה יעקובזון

Karmiel

כרמיאל

2014

תוכן העניינים:

פרק 1- מבוא	1
פרק 2- תיאור הבעיה	2
דוגמא 2.1 – עיכוב כללי מינימאלי של תעבורה ברשת	2
סימונים והגדרות 2.2	4
פרק 3- סקר ספרות - המרות מחסום ושיטת המ"מ	5
למה 3.1 (למת מחסום)	5
משפט 3.2 (ההתכנסות של שיטות המתבססות על המרות מחסום)	5
למה 3.3 (למת הכופלי Karush Kuhn Tucker)	6
הגדרה 3.4 (המרות מחסום משופרות של Polyak)	7
פרק 4- סוגיית קיום הפתרון – תנאים ההכרחיים ומספיקים	8
הגדרה 4.5 (אילוצים פעילים)	8
משפט 4.6 (תנאי הכרחי של Karush-Kuhn-Tucker)	8
משפט 4.7 (תנאי מספיק אופטימום)	8
פרק 5 – מחלקת המרות B	9
הגדרה 5.1 (מחלקת המרות B)	9
למה 5.3 (שקילות של בעיית אופטימיזציה מקורית ובעיה מוגדרת באמצעות המרת מחסום)	10
משפט 5.4 (ההתכנסות של שיטות המתבססות על המרות מחסום ממחלקת B)	11
פרק 6 – תיאור שיטה האיטרטיבית לפתרון	18
מסקנה 6.1 (התכנסות שיטה איטרטיבית מוצעת)	18
משפט 6.4 (קביעת ערכים התחלתיים עבור שיטת האיטרציות)	20
פרק 7 – דוגמאות והרצות	24
דוגמא 7.2 (פתרון איטרטיבי של דוגמא 2.1)	24
דוגמא 7.3	27
דוגמא 7.4	30
דוגמא 7.5	33
רשימת מקורות	36
נספחים	37
נספח 1 - הוכחת קמרות של בעיית אופטימיזציה בדוגמא 2.1	37
נספח 2 - ממוש שיטת המרות מחסום משופרות ע"י שימוש במתל"ב:	39
נספח 3 - דוגמא למבנה נתונים (עבור בעיה 7.2)	42
נספח 4 - הוכחה שהמרות משופרות של Polyak שייכות למחלקה B	43
נספח 5 - ציטוט תוצאות והוכחות	44

1. מבוא

בעיות אופטימיזציה מאולצת מהוות מודל מתמטי חשוב למגוון בעיות הנדסיות. מחקרים רבים עוסקים בשאלות סביב קיום ויחידות, שיטות לקירוב ואפיונים שונים של פתרון לבעיות אלו.

בשנות החמישים של המאה הקודם [5] Frish ו-[3] Carrol המליצו על פתרון של בעיות אופטימיזציה מאולצות באמצעות התמרות מחסום (ה"מ). יותר מאוחר המרות אלו נחקרו בהרחבה על ידי

[4] Fiacco & McCormick ועבורם פותחו שיטות פתרון שונות.

יחד עם זאת לאנליזה באמצעות ה"מ, כמו גם לשיטות הפתרון שמתבססות עליהן יש כמה חסרונות מובנים. עקרון השיטות שמתבססות על ה"מ הוא שימוש באי חסימות שלהן סביב הגבול של התחום האפשרי. זה מבטיח שחיפוש אופטימום המתחיל מנקודה פנימית כלשהיא של תחום האפשרי יישאר בתוכו בלא טיפול מיוחד באילוצים. יחד עם זאת נוצרת בעיה חישובית קשה כאשר התהליך מתקרב לגבולות של האילוצים האקטיביים, ובנוסף, ה"מ וכך גם הנגזרות שלהן אינן מוגדרות כלל בנקודת האופטימום. התהליך החישובי נהיה כבד יותר מצעד לצעד וגם קצב ההתכנסות נהיה איתי יותר.

לחלופין, הלגרנז'אן הקלאסי (ל"ק) שמהווה כלי בסיס לאופטימיזציה מאולצת: גם לאנליזה (תנאים הכרחיים ומספיקים לפתרון, התורה הדואלית) וגם לשיטות אופטימיזציה, גם הוא לוקה בחסרונות. למשל, עבור כופלי לגרנז' אופטימאליים קבועים, ייתכן והאופטימום לא קיים, אפילו אם תנאי אופטימאליות מסוג שני מתקיימים.

ב- [12] Polyak הציע שתי המרות מחסום משופרות (המ"מ) שמשלבות את היתרונות של הלגרנז'אן הקלאסי ושל המרות מחסום לבעיות אופטימיזציה מאולצות, אבל פטורות מהחסרונות הבולטים של שניהם. הן מוגדרות באופטימום, שומרות על חלקות מאותו הסדר כמו הבעיה המקורית בסביבת התחום האפשרי ולא עולות אינסופית בסביבת הגבול האקטיבי של התחום האפשרי.

יחד עם זאת הם מאפשרים אנליזה בדיוק כמו הלגרנז'אן הקלאסי, הן קמורות ממש בסביבת הפתרון והפתרון האופטימאלי תמיד קיים אם תנאי אופטימאליות מסדר שני מתקיימים. ב-[12] פותחו שלוש שיטות בהתבסס על המ"מ: השיטה שפותרת בעיה פרימאלית עם קצב התכנסות ליניארי, השיטה שפותרת בעיה דואלית עם קצב התכנסות ליניארי והשיטה הכללית לפתרון בו זמני של בעיות פרימאליות ודואלית עם קצב התכנסות ריבועי (אם בכל צד מקטינים פרמטר הכנס, קצב ההתכנסות של שתי השיטות הראשונות הופך להיות סופרליניארי). העיקרון של שיטות אלו תופס גם עבור בעיית מינימום-מקסימום כאשר מציגים אותה כבעיית אופטימיזציה מאולצת.

בעבודה זו הגדרתי מחלקה B של המרות מחסום, הכוללת גם את ההמרות של [12] Polyak. והוכחתי שלכל המרה ממחלקה B עובדת השיטה הפרימאלית שפותחה ב-[12]. שימוש בשיטה הודגם על דוגמאות שונות, בין היתר מתחומי הנדסה.

בפרק 2 הצגתי את הבעיית אופטימיזציה מאולצת שבה עוסקת עבודה זו, כולל דוגמא העוסקת בתעבורה ברשת שמהוו בעיית אופטימיזציה קמורה לא לינארית ומאולצת. בפרק 3 סקרתי ספרות והצגתי רקע מדעי וגישת מחסום כללית לפתרון בעיית אופטימיזציה מאולצת. הדגשתי תכונות עיקריות של המרות מחסום והצגתי המרות מחסום משופרות שהוצעו ע"י [12] Polyak, תוך התייחסות ליתרונות של המרות מחסום משופרות לעומת המרות מחסום קלאסיות של [5] Frish ו-[3] Carrol. בפרק 4 סקרתי מושגים בסיסיים ותוצאות שקשורות לקיום פתרון של בעיית אופטימיזציה מאולצת. בפרק 5 הגדרתי מחלקת המרות מחסום משופרות B , ניסחתי והוכחתי למה ומשפט מרכזיים המבססים התכנסות של שיטות איטרטיביות של [12] Polyak לכל המרה ממחלקה B של המרות מחסום משופרות. פרק 6 מציג את שיטת האיטרציות של [12] Polyak ודן ביישומה, כולל בחירת נקודה התחלתית. בפרק 7 הצגתי פתרון של הדוגמאות בעזרת שימוש בהמרות מחסום משופרות ושיטת איטרציות שפותחה ע"י [12] Polyak.

2. תיאור הבעיה:

אנו נעסוק בבעיית אופטימיזציה לא ליניארית עם אילוצים מהצורה

$$(*) \quad \begin{aligned} x^* &= \arg \min \{ f(x) / x \in \Omega \} \\ \Omega &= \{ x \in \mathbb{R}^n / g_i(x) \geq 0, \quad i = \overline{1, m} \} \end{aligned}$$

כאן:

$f: \mathbb{R}^n \rightarrow \mathbb{R}$ – היא פונקציית מטרה (פונקציה נתונה),

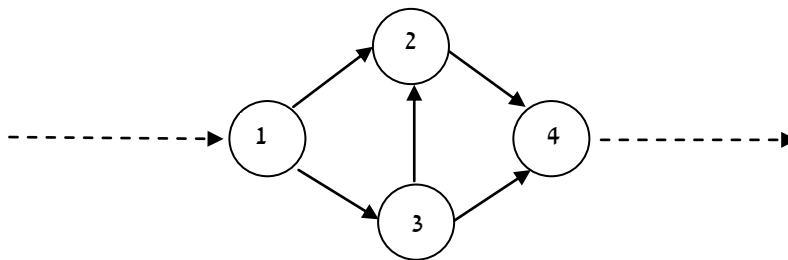
$g_i: \mathbb{R}^n \rightarrow \mathbb{R}$ – הן פונקציות אילוצים, $i = \overline{1, m}$.

(הבעיה לא ליניארית, כלומר לפחות אחת מהפונקציות (מטרה או אילוצים) היא פונקציה לא ליניארית).

דוגמא 2.1: עיכוב כללי מינימאלי של תעבורה ברשת.

ניתוח תעבורת הרשת הוא אחד המשימות החשובות בעולם המודרני. ניתוח מסלול אופטימאלי בין שתי תחנות ברשת יכול לפתור בעיות עומסים מיותרים ועיכובים במערכת הכבישים, למשל. דוגמא אחרת היא העברת נתונים ברשת מחשבים. גם כאן למהירות העברת הנתונים מנקודה אחת לנקודה אחרת ברשת יכולה להיות חשיבות מכרעת בהקשרים שונים. במקרים אלה פתרון של בעיית אופטימיזציה מציע לבחור מסלול מהיר ביותר המאפשר תעבורה ללא עיכובים מיותרים ברשת.

נניח שברצוננו לנתח תעבורה במערכת כבישים. נסתכל על מערכת כבישים פשוטה המתוארת ע"י גרף הבא:



איור 1. בעיית התעבורה ברשת

במערכת קיימים 5 כבישים ו-4 צמתים.

צומת כניסה למערכת הוא צומת מספר 1 וצומת יציאה הוא צומת מספר 4.

ידוע שבכל דקה עוברים במערכת F רכבים.

אנחנו מעוניינים למצוא מספר הרכבים העוברים בכל כביש בדקה כך שעכוב כללי של רכבים במערכת הכבישים יהיה מינימאלי.

נסמן:

$i = 1, 2, 3$ – צומת כניסה לכביש

$j = 2, 3, 4$ – צומת יציאה מהכביש

x_{ij} – מספר הרכבים בכביש המחבר בין צומת i לצומת j (מספים אי-שליליים)

c_{ij} – מספרים מקסימאליים של רכבים בכביש המחבר בין צומת i לצומת j

t_{ij} – זמן להגיע מצומת i לצומת j

F – מספר הרכבים במערכת כבישים

במקרה זה פונקציית המטרה $f: \mathbb{R}^5 \rightarrow \mathbb{R}$ היא פונקציית עיכוב כללי של רכבים במערכת כבישים:

$$f(x) = t_{12}x_{12} + t_{13}x_{13} + t_{32}x_{32} + t_{24}x_{24} + t_{34}x_{34}$$

$$x = (x_{12}, x_{13}, x_{32}, x_{24}, x_{34}) \in \mathbb{R}^5$$

נניח שנתון כי:

$$t_{12} = 5 + \frac{0.1x_{12}}{11 - x_{12}}, \quad t_{13} = \frac{x_{13}}{31 - x_{13}}, \quad t_{32} = 1 + \frac{x_{32}}{11 - x_{32}}, \quad t_{24} = \frac{x_{24}}{31 - x_{24}}, \quad t_{34} = 5 + \frac{0.1x_{34}}{11 - x_{34}}$$

ונניח כי ישנה הגבלה על מספר הרכבים בכל כביש,

$$c_{12} = 10, c_{13} = 30, c_{32} = 10, c_{24} = 30, c_{34} = 10$$

נגדיר אילוצים:

$$\begin{array}{lll} g_1(x) = x_{12} - c_{12} & g_6(x) = x_{12} & h_1(x) = x_{13} - x_{32} - x_{34} \\ g_2(x) = x_{13} - c_{13} & g_7(x) = x_{13} & h_2(x) = x_{12} + x_{32} - x_{24} \\ g_3(x) = x_{32} - c_{32} & g_8(x) = x_{32} & h_3(x) = x_{12} + x_{13} - F \\ g_4(x) = x_{24} - c_{24} & g_9(x) = x_{24} & x \in \mathbb{R}^5 \\ g_5(x) = x_{34} - c_{34} & g_{10}(x) = x_{34} & \end{array}$$

נקבל בעיית אופטימיזציה מאולצת לא לינארית:

$$(i) \quad x^* = \arg \min \{ f(x) / x \in \Omega \}$$

$$\Omega = \{ x \in \mathbb{R}^n / \quad g_i(x) \geq 0, \quad i = \overline{1, 10}, \quad h_j(x) = 0 \quad j = \overline{1, 3} \}$$

הערה:

אילוצי שוויון: $h_j(x) = 0, j = \overline{1, k}$ ניתן בקלות להפוך לאילוצי אי-שוויון:

$$\begin{cases} h_j(x) \leq 0; & j = \overline{1, k} \\ -h_j(x) \leq 0; & j = \overline{1, k} \end{cases}$$

אם $h_j(x); j = \overline{1, k}$ הן פונקציות לינאריות אז כל האילוצים הנ"ל הן פונקציות קמורות.

אחרת האילוצים המוכפלים במינוס אחד הופכים להיות פונקציות קעורות.

סימונים והגדרות 2.2:

$$g(x) = (g_1(x), \dots, g_m(x)) \quad - \quad \text{וקטור האילוצים של בעיה (*)}$$

$$i = \overline{1, m} \text{ , לכל } g_i(x) \geq 0 \Leftrightarrow g(x) \geq 0 \quad -$$

$$g'_i(x) = \nabla g_i(x) \quad -$$

$$f'(x) = \nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right) \quad - \quad \text{הגרדיינט של } f$$

$$g'(x) = \nabla g(x) = (\nabla g_1, \dots, \nabla g_m) = J[g(x)] = \left[\frac{\partial g_i(x)}{\partial x_j} \right]_{i=1}^m \bigg|_{j=1}^n \quad - \quad \text{יקוביאן של } g$$

$$f''(x) = H(f) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right]_{n \times n} \quad - \quad \text{ההסיין של } f$$

$$g_{r+1}(x^*) = \dots = g_m(x^*) > 0 \text{ ואילו } g_1(x^*) = \dots = g_r(x^*) = 0 \quad \text{נניח כי מתקיים}$$

$$I = \{1, 2, \dots, r\} \quad - \quad \text{תקרא קבוצת האינדקסים האקטיביים}$$

$$\sigma = \max \{g_{r+1}(x^*), \dots, g_m(x^*)\} > 0 \quad -$$

$$g^{(m-r)}(x) = (g_{r+1}(x), \dots, g_m(x)) \text{ ו- } g^{(r)}(x) = (g_1(x), \dots, g_r(x)) \quad -$$

$$y^{(m-r)}(x) = (y_{r+1}, \dots, y_m) \text{ ו- } y^{(r)} = (y_1, \dots, y_r) \quad -$$

$$g^{(m-r)'}(x) = J[g(x)] = \left[\frac{\partial g_i(x)}{\partial x_j} \right]_{i=r+1}^m \bigg|_{j=1}^n, \quad g^{(r)'}(x) = J[g(x)] = \left[\frac{\partial g_i(x)}{\partial x_j} \right]_{i=1}^r \bigg|_{j=1}^n \quad -$$

3. סקר ספרות - המרות מחסום ושיטת המ"מ:

לפתרון של בעיות אופטימיזציה מאולצות מקובל להשתמש בשיטות פנימיות (Interior Methods). הן נכנסו לשימוש נרחב במהלך שנות השישים של המאה הקודמת לבעיות מאולצות לא ליניאריות, השימוש בהם לבעיה הבסיסית של תכנות ליניארי לא עלה על הדעת בגלל הדומיננטיות המוחלטת שיטת סימפלקס לפתרון בעיות אלו. במהלך שנות השבעים של המאה הקודם השימוש בשיטות מחסום הפך ללא פופולרי בגלל פיתוחם של חלופות יותר יעילות (Sequential Quadratic Programming, Augmented Lagrangian).

עד שנות השמונים המוקדמות של המאה הקודם, שיטות מחסום נחשבו כפרק סגור בהיסטוריה של אופטימיזציה. תמונה זו השתנתה באופן דרמטי ב-1984, כאשר N. Karmarkar [7] פרסם שיטה פנימית המתכנסת בזמן פולינומיאלי לבעיית תכנות לינארי. ב-1985 החוקרים הראו קשר פורמלי בין השיטה שלו ושיטת המחסום הקלאסי. מאז לשיטות פנימיות תפקיד דומיננטי גם בתאוריה וגם בפרקטיקה של אופטימיזציה מאולצת. אנו נתייחס כאן לרעיון הכללי של שיטות המחסום.

שיטת מחסום מציעה כאלטרנטיבה לבעיה (*) בעיה חדשה מהסוג:

$$(B(\varepsilon)) \quad x_\varepsilon = \arg \min \{ f(x) + \varepsilon \beta(x) / x \in \Omega \}$$

$$\Omega_B = \{ x \in \mathbb{R}^n / g_i(x) > 0, \quad i = \overline{1, m} \}$$

כאשר הפונקציה $\beta: \mathbb{R}^n \rightarrow \mathbb{R}$ היא המרת המחסום שמקיימת:

- $\beta(x) \geq 0$ לכל $x \in \Omega_B$
- $\lim_{x \rightarrow \partial \Omega_B} \beta(x) = \infty$

לדוגמא, אפשר לבחור כהמרת מחסום את הפונקציה $\beta(x) = \sum_{i=1}^m \frac{1}{g_i(x)}$ Carroll [3]. היציע אותה לראשונה.

תוצאות הבאות שהתקבלו עבור בעיה $(B(\varepsilon))$ מבססות שימוש בשיטות מחסום לפתרון בעיות אופטימיזציה מאולצות.

למה 3.1 (למת המחסום):

יהיה x^* פתרון של בעיה (*). עבור בעיה $(B(\varepsilon))$, נסמן $r(\varepsilon, x) = f(x) + \varepsilon \beta(x)$.

נניח כי הסדרה $\{\varepsilon_k\}_{k=1}^\infty$ יורדת ומקיימת $\lim_{k \rightarrow \infty} \varepsilon_k = 0$. יהיה x^k פתרון של $(B(\varepsilon_k))$, אז מתקיים:

$$r(\varepsilon_k, x^k) \geq r(\varepsilon_{k+1}, x^{k+1}) \quad (1)$$

$$\beta(x^k) \leq \beta(x^{k+1}) \quad (2)$$

$$f(x^k) \leq f(x^{k+1}) \quad (3)$$

$$f(x^*) \leq f(x^k) \leq r(\varepsilon_k, x^k) \quad (4)$$

משפט 3.2 (משפט ההתכנסות של שיטות המתבססות על המרות מחסום):

נניח כי $\beta(x), f(x), g_i(x)$ לכל $1 \leq i \leq m$ הן פונקציות רציפות.

נניח כי לבעיה (*) קיים פתרון אופטימאלי x^* , כך שלכל $\delta > 0$ מתקיים $\{x \in \mathbb{R}^n / \|x - x^*\| < \delta\} \cap \Omega_B \neq \emptyset$.

תהי סדרת הפתרונות של $(B(\varepsilon_k))$, אז כל נקודה $\bar{x} = \lim_{k \rightarrow \infty} x^k$ היא פתרון של (*).

מסקנה:

- מלמה 3.1 נובע שאם התחלנו לעבוד בשיטת מינימיזציה מוכחת מנקודה פנימית בתחום Ω_B אז $\{x^k\}_{k=2}^\infty \subset \Omega_B$. כלומר, הרעיון של השיטה הוא להתחיל בנקודה המקיימת את האילוצים של בעיה $(B(\varepsilon))$.
- אם לבעיה (*) קיים פתרון, אז תהליך חיפוש שבו בכל צעד מוצאים את המינימום לבעיה לא מאולצת $x_k = \arg \min \{f(x) + \varepsilon_k \beta(x)\}$ יורדת ומקיימת $\lim_{k \rightarrow \infty} \varepsilon_k = 0$ ייצור סדרה שמתכנסת לפתרון של בעיה (*) בתחום Ω .
- אם המרת מחסום $\beta: \mathbb{R}^n \rightarrow \mathbb{R}$ גזירה ברציפות, אז התנאי $\nabla f(x) + \varepsilon \nabla \beta(x) = 0$ שקול לתנאי $\nabla f(x) + \varepsilon \sum_{i=1}^m \frac{\partial \beta(g_i(x))}{\partial g_i(x)} \nabla g_i(x) = 0$. אם הפתרון של $(B(\varepsilon_k))$ הוא x^k , נסמן $u_i^k = \frac{\partial \beta(g_i(x^k))}{\partial g_i(x^k)}$ ונקבל $\nabla f(x^k) + \varepsilon \sum_{i=1}^m u_i^k \nabla g_i(x^k) = 0$. כלומר ניתן להתייחס ל- $u^k = (u_1^k, u_2^k, \dots, u_m^k)$ כווקטור כופלי לאגראנז' (כופלי Karush-Kuhn-Tucker).

למה 3.3 (כופלי Karush-Kuhn-Tucker)

- נניח כי $\beta(x), f(x), g_i(x)$ לכל $1 \leq i \leq m$ הן פונקציות גזירות ברציפות. נניח כי לבעיה (*) קיים פתרון אופטימאלי x^* , כך שלכל $\delta > 0$ מתקיים $\{x \in \mathbb{R}^n \mid \|x - x^*\| < \delta\} \cap \Omega_B \neq \emptyset$.
- אז אם $\bar{x} = \lim_{k \rightarrow \infty} x^k$ ו- $\bar{x}$ מקיים את תנאי האי-תלות של גרדיינט האילוצים הפעילים, אז קיים הגבול $\bar{u} = \lim_{k \rightarrow \infty} u^k$ ו- $\bar{u}$ הינו ווקטור כופלי Karush-Kuhn-Tucker עבור פתרון אופטימאלי $\bar{x}$ של (*).

המרת המחסום הלוגריתמית הוצעה לראשונה על ידי Frish [5] נוכל לכתוב באמצעותה את פונקציית המטרה של

$$F^*(x, \varepsilon) = f(x) + \varepsilon \sum_{i=1}^m \log(g_i(x)) \quad \text{בעיה } (B(\varepsilon)) \text{ בצורה:}$$

בדוגמא קודמת עבור המרת מחסום שהוצעה על ידי Carroll [3] נוכל לכתוב באמצעותה את פונקציית המטרה של

$$C^*(x, \varepsilon) = f(x) - \varepsilon \sum_{i=1}^m (g_i(x))^{-1} \quad \text{בעיה } (B(\varepsilon)) \text{ בצורה:}$$

כאשר ε הוא פרמטר חיובי ידוע כפרמטר "קנס".

בהמשך McCormic & Fiacco [4] המשיכו לחקור פונקציות מחסום של Carroll ו-Frish והראו שתחת הנחות

טבעיות על בעיית אופטימיזציה (*) הסדרה $\{x^k\}_{k=1}^\infty$ מתכנסת לפתרון שלה כאשר סדרת פרמטרי "קנס" $\{\varepsilon_k\}_{k=1}^\infty$

שואפת לאפס ונורמת ההפרש בין פתרון האופטימאלי x^* לבין x^k בכל שלב הוא מסדר גודל של ε_k .

בתחילה סברו שכל שיטת מינימיזציה סטנדרטית (למשל, שיטת ניוטון, שיטת גרדיאנט) תתאים בלא קושי לפתרון

בעיות לא מאולצות $(B(\varepsilon_k))$ ובתהליך איטרטיבי תיווצר הסדרה $\{x^k\}_{k=1}^\infty$ שתתכנס לפתרון בעיית אופטימיזציה (*).

מסתבר שבפועל הדברים לא כל כך פשוטים. יהיו $\lambda_{\min}$ ו- $\lambda_{\max}$ ערכים עצמיים המקסימאלי והמינימאלי (בהתאמה) של מטריצת הסיאן (Hessian) של פונקציית המטרה. בשיטות נומריות סטנדרטיות מספר הספקטראלי $\lambda_{\max} / \lambda_{\min}$ קובע את הדיוק וקצב התכנסות לפתרון. [9] Lootsma ו- [10] Murray חקרו מטריצת הסיאן של פונקציות מחסום והסיקו, שברוב המקרים המספר הספקטראלי של מטריצת הסיאן עבור הבעיה $(B(\varepsilon))$ גדל מאד ככל ש- ε קטן. עובדה זו גורמת בעיות נומריות קשות במימוש השיטות מבוססות פונקציות מחסום. בנוסף, המרות מחסום של Frish ו- Carroll לא מוגדרות בפתרון לבעיית אופטימיזציה מקורית x^* . חוקרים רבים ניסו לפתח כל מיני שיטות על מנת לעקוף את הבעיות. אבל שיטות אלו התאימו למחלקות מצומצמות מאוד של בעיות אופטימיזציה. למשל, [11] Sofer & Nash הציגו שיטה שמתאימה לבעיות אופטימיזציה מאולצות בעלת אילוצים פשוטים (אילוצים הם גבולות).

[6] Osborne & Jittornrum הציעו לשפר המרות מחסום ידועות והציגו המרת מחסום לוגריתמית:

$$\lambda_i \geq 0, i = \overline{1, m}, \beta(x) = \sum_{i=1}^m \lambda_i \log(g_i(x))$$

הינה

$$f(x) + \varepsilon \sum_{i=1}^m \lambda_i \log(g_i(x))$$

כלומר היא בעצם תחליף לפונקציית לאגרנז'יאן, ו- $\lambda_i, i = \overline{1, m}$ הם כפלי לאגרנז'י

מתאימים. Osborne & Jittornrum הצליחו להראות שעבור בחירה מסוימת של כפלי לאגרנז'י, הסדרה $\{x^k\}_{k=1}^\infty$ מתכנסת לפתרון ככל ש- ε שואף ל-0, וקצב התכנסות הוא לינארי או אפילו יותר טוב מלינארי. המרה זו ונהית לא חסומה כאשר מתקרבים לגבול של התחום האפשרי ואינה מוגדרת בפתרון עצמו. התקדמות משמעותית בשיפור וניתוח פונקציות מחסום הופיעה בעבודות של [12] Polyak. הוא הציג שתי המרות מחסום שהן למעשה שיפור של המרות מחסום של Frish ו- Carrol.

הגדרה 3.4 (המרות מחסום משופרות):

$$\psi_F(\varepsilon, t) = \varepsilon \log\left(\frac{t}{\varepsilon} + 1\right)$$

$$\psi_C(\varepsilon, t) = -\varepsilon \left[\left(\frac{t}{\varepsilon} + 1 \right)^{-1} - 1 \right]$$

המרות מחסום משופרות הן פונקציות אי-שליליות, רציפות מונוטוניות וקמורות/קעורות במשתנה שני. במקרה זה פונקציות המטרה מתאימות הן:

$$F(x, \varepsilon, y) = f(x) - \varepsilon \sum_{i=1}^m y_i \log\left(\frac{g_i(x)}{\varepsilon} + 1\right)$$

$$C(x, \varepsilon, y) = f(x) + \varepsilon \sum_{i=1}^m y_i \left[\left(\frac{g_i(x)}{\varepsilon} + 1 \right)^{-1} - 1 \right]$$

כאשר $y_i, i = \overline{1, m}$ הם כפלי לאגרנז'י מתאימים.

פונקציות מטרה המבוססות על המרות מחסום משופרות של [12] Polyak הן בעצם תחליף ללגרנז'יאן הקלאסי. הן חלקות מאותו סדר כמו פונקציות המטרה ואילוצים בפנים התחום האפשרי ועל שפתו, ולכן גם מוגדרות וחלקות בנקודת האופטימום ושומרות על כל התכונות החשובות של לאגרנז'יאן.

Polyak הראה שעבור $\varepsilon > 0$ מספיק קטן תחת הנחות טבעיות סדרה $\{x^k\}_{k=1}^\infty$ מתכנסת לפתרון של בעיה המקורית.

מטרת עבודה זו להגדיר דרישות פורמאליות לפונקציה בכדי שתוכל לשרת כהמרת מחסום משופרת בשיטת Polyak במקום ψ_1 או ψ_2 .

4. סוגיית קיום הפתרון – תנאים ההכרחיים ומספיקים

בכדי שנוכל לדון בנושא זה נזדקק לכמה מושגים ותוצאות.

הגדרה 4.1:

בעיה (*) נקראת בעיית אופטימיזציה קמורה כאשר $f(x), g_i(x), i = \overline{1, m}$ הן פונקציות קמורות.

הערה: (i) היא בעיית אופטימיזציה קמורה (ראה נספח מס' 1).

משפט 4.2 תנאי מספיק והכרחי לקמירות הפונקציה

תהי נתונה פונקציה $f(x)$ גזירה פעמיים ברציפות בתחום קמור ופתוח D ב- $\mathbb{R}^n$,

אז $f(x)$ קמורה ב- D אם ורק אם מטריצת הסיאן שלה מוגדרת אי-שלילית $H(f) \geq 0$.

(כלומר, אם ורק אם לכל וקטור $\vec{a} \in D$ מתקיים $\vec{a}^T H(f) \vec{a} \geq 0$).

הגדרה 4.3 - פונקצית לגראנז'יאן:

תהי נתונה בעיית אופטימיזציה (*). ונגדיר פונקצית לגראנז'יאן $L: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ של (*) בצורה

$$(1) \quad L(x, y) = f(x) + y^T g(x)$$

הגדרה 4.4 - תנאי סלייטור Slater

תהי נתונה בעיית אופטימיזציה מאולצת (*). אם (*) בעיה קמורה, וקיימת נקודה $x_0 \in D$ כך ש-

$$(2) \quad g_i(x_0) > 0; \quad i = \overline{1, m}$$

אז אומרים ש- (*) מקיימת תנאי סלייטור.

הגדרה 4.5

תהי $I^* = \{1 \leq k_i \leq m \mid g_{k_i}(x^*) = 0\}$. נאמר שפונקציות $g_i(x), i \in I^*$ הן אילוצים פעילים לבעיה (*).

ללא הגבלת הכלליות נניח כי $g_i(x^*) = 0$ לכל $1 \leq i \leq r$ ב-(*), ואז $I^* = \{1, 2, \dots, r\}$.

(סמנו $g^{(r)}(x) = \nabla g^{(r)}(x) = H(g^{(r)}(x))$, $g^{(r)}(x) = (g_1(x), g_2(x), \dots, g_r(x))$)

משפט 4.6 - (תנאי הכרחי של KKT)

נניח כי עבור הבעיה (*) הפונקציות $f(x), g_i(x), i = \overline{1, m}$ גזירות ברציפות.

אם $x^* \in D$ מינימום לוקלי שמקיים

$$(3) \quad \text{rank } g^{(r)'}(x^*) = r \quad \text{כך ש-} i \in I^* \quad \text{לכל } g_i(x^*) = 0$$

אז קיים וקטור $y^* \in \mathbb{R}^m$ שעבורו מתקיים:

$$(4) \quad L'_x(x^*, y^*) = \nabla f(x^*) - \sum_{i=1}^m y_i^* \nabla g_i(x^*) = 0, \quad y_i^* g_i(x^*) = 0, \quad i = \overline{1, m}$$

משפט 4.7 - (תנאי מספיק)

תהי (*) בעיית אופטימיזציה גזירה פעמיים ברציפות.

נניח כי בנקודה $x^* \in D$ מתקיימים תנאים (3) ו-(4), כך ש- $y_i^* > 0, i \in I^*$. אם קיים $\lambda > 0$ כך ש-

$$(5) \quad \langle L''_{xx}(x^*, y^*) d, d \rangle \geq \lambda \|d\|^2; \quad \forall 0 \neq d \in \mathbb{R}^m, \quad g^{(r)'} d = 0, \quad g^{(m-r)'} d \geq 0$$

אז $x^* \in D$ היא נקודת מינימום מבודדת של (*)

5. מחלקת התמרות B:

נגדיר באופן כללי מחלקת התמרות מחסום B שמישה לשיטת המרות מחסום משופרות.

5.1 הגדרה:

יהיה $\varepsilon > 0$ כלשהו.

B היא קבוצת כל הפונקציות $\psi(\varepsilon, t): \mathbb{R}^{++} \times [-\varepsilon, \infty) \rightarrow \mathbb{R}$ ו- $\psi(\varepsilon, t) \in C^2$, אשר מקיימות:

$$\lim_{\varepsilon \rightarrow 0} \psi(\varepsilon, t) = \psi(0, t) = 0 \quad (\text{א})$$

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \psi'_\varepsilon(\varepsilon, t) = 0 \quad (\text{ב})$$

$$\psi(\varepsilon, 0) = 0 \quad (\text{ג})$$

(ד) $\psi(\varepsilon, t)$ היא פונקציה יורדת וקמורה או עולה וקעורה

(ה) עבור $\psi(\varepsilon, t)$ יורדת: אם $t < \infty$ אז $-\infty < \psi(\varepsilon, t)$. עבור $\psi(\varepsilon, t)$ עולה: אם $t < \infty$ אז $\psi(\varepsilon, t) > \infty$

$$\psi'_t(\varepsilon, t) = \begin{cases} O(\varepsilon), & t \neq 0 \\ 1, & t = 0 \end{cases} \quad \text{אם } \psi(\varepsilon, t) \text{ עולה, אז} \quad \psi'_t(\varepsilon, t) = \begin{cases} O(\varepsilon), & t \neq 0 \\ -1, & t = 0 \end{cases} \quad \text{אם } \psi(\varepsilon, t) \text{ יורדת, אז}$$

$$\lim_{\varepsilon \rightarrow 0} \psi''_{tt}(\varepsilon, t) = \begin{cases} 0, & t \neq 0 \\ -\infty, & t = 0 \end{cases} \quad \text{אם } \psi(\varepsilon, t) \text{ עולה, אז} \quad \lim_{\varepsilon \rightarrow 0} \psi''_{tt}(\varepsilon, t) = \begin{cases} 0, & t \neq 0 \\ \infty, & t = 0 \end{cases} \quad \text{אם } \psi(\varepsilon, t) \text{ יורדת, אז}$$

(ח) קיים $k \in \mathbb{N}$ כך ש- $\lim_{\varepsilon \rightarrow 0} \varepsilon^k \psi''_{tt}(\varepsilon, 0) = l$ ו- $l \in \mathbb{R}$

(ט) עבור $\psi(\varepsilon, t)$ יורדת: $\lim_{t \rightarrow -\varepsilon} \psi(\varepsilon, t) = \infty$. עבור $\psi(\varepsilon, t)$ עולה: $\lim_{t \rightarrow -\varepsilon} \psi(\varepsilon, t) = -\infty$

5.2 הערה:

המרות מחסום $\psi_F(\varepsilon, t) = \varepsilon \log(\frac{t}{\varepsilon} + 1)$ ו- $\psi_C(\varepsilon, t) = \varepsilon[(\frac{t}{\varepsilon} + 1)^{-1} - 1]$ שייכות למחלקה B (ראה נספח 4).

תהי $\psi \in B$ ו- $\varepsilon > 0$ כלשהו,

5.2 נתבונן בשתי בעיות אופטימיזציה:

$$(*) \quad x^* = \arg \min \{ f(x) / x \in \Omega \}, \quad \Omega = \{ x \in \mathbb{R}^n / g_i(x) \geq 0, \quad i = \overline{1, m} \}$$

-1

$$(**) \quad x^* = \arg \min \{ f(x) / x \in \Omega_\psi \}$$

כאן עבור ψ יורדת וקמורה $\Omega_\psi = \{ x \in \mathbb{R}^n / -\infty < \psi(\varepsilon, g_i(x)) \leq 0, \quad i = \overline{1, m} \}$

ואילו עבור ψ עולה וקעורה $\Omega_\psi = \{ x \in \mathbb{R}^n / 0 \leq \psi(\varepsilon, g_i(x)) < \infty, \quad i = \overline{1, m} \}$

כאן $f(x), -g_i(x), i = \overline{1, m}$ קמורות.

יהיה $s_\psi = -\text{sign}(\psi'_t(\varepsilon, t))$, הלאגראניאן של (***) הוא

$$F(\varepsilon, x, y) = f(x) + s_\psi \sum_{i=1}^m y_i \psi(\varepsilon, g_i(x))$$

נראה שאם (*) היא בעיית אופטימיזציה קמורה אז כך גם (**), ובנוסף, אם בעיה (*) מקיימת תנאים הכרחיים ומספיקים לקיום הפתרון, אז גם בעיה (**) ובאותה נקודה $(x^*, y^*) \in \mathbb{R}^{n+m}$, לכן שתי הבעיות שקולות זו לזו.

למה 5.3:

תהי נתונה המרה $\psi \in B$, לכל $\varepsilon > 0$ הבעיה (*) שקולה ל- (**).

הוכחה: ללא הגבלת הכלליות נניח כי ψ עולה וקעורה.

$$(1) \quad g_i(x) = 0 \Leftrightarrow \psi(\varepsilon, g_i(x)) = 0 \text{ וגם } g_i(x) > 0 \Leftrightarrow \psi(\varepsilon, g_i(x)) > 0$$

$$(2) \quad -g_i(x), i = \overline{1, m} \Leftrightarrow -\psi(\varepsilon, g_i(x)), i = \overline{1, m} \text{ קמורות.}$$

אכן, אם לכל $x \in \Omega$ מתקיים $g_i''(x) \leq 0$ לכל $i = \overline{1, m}$, אז כיוון ש- $\psi'_t(\varepsilon, t) > 0$ ו- $\psi''_{tt}(\varepsilon, t) \leq 0$

$$\text{לכל } x \in \Omega_{\psi} \text{ מתקיים } \psi''_{xx}(\varepsilon, g_i(x)) = \psi''_{gg}(\varepsilon, g_i(x))(g'_i(x))^2 + \psi'_{g_i}(\varepsilon, g_i(x))g''_i(x) \leq 0$$

(3) נניח כי בעיה (*) מקיימת את K-K-T, כלומר קיימת נקודה $(x^*, y^*) \in \Omega \times \mathbb{R}_+^m$ שמקיימת

$$(i) \quad L'_x(x^*, y^*) = f'(x^*) - \sum_{i=1}^m y_i^* g'_i(x^*) = 0; \quad y_i^* g_i(x^*) = 0, \quad i = 1 \dots m$$

נראה שגם בעיה (**) מקיימת את K-K-T.

כיוון שבביטוי: $y_i^* \psi(\varepsilon, g_i(x^*))$ מתקיים

$$\text{אם } g_i(x^*) = 0 \text{ אז גם } \psi(\varepsilon, g_i(x^*)) = 0 \text{ ומתקיים } y_i^* \psi(\varepsilon, g_i(x^*)) = 0$$

$$\text{אם } y_i^* = 0 \text{ אז ומתקיים } y_i^* \psi(\varepsilon, g_i(x^*)) = 0$$

$$\text{לכן } y_i^* g_i(x^*) = 0 \Leftrightarrow y_i^* \psi(\varepsilon, g_i(x^*)) = 0 \text{ לכל } i = 1 \dots m$$

נניח כי תנאי (i) מתקיים,

אם $y_i^* \neq 0$ או $g_i(x^*) = 0$, לכן לפי תכונה (ו) מתקיים $\psi'_{g_i}(\varepsilon, g_i(x^*)) = 1$ ולמעשה מקבלים ש-

$$f'(x^*) + s_{\psi} \sum_{i=1}^m y_i^* g'_i(x^*) \psi'_{g_i}(\varepsilon, g_i(x^*)) = f'(x^*) - \sum_{i=1}^m y_i^* g'_i(x^*)$$

או, במילים אחרות,

$$F'_x(\varepsilon, x^*, y^*) = f'(x^*) + s_{\psi} \sum_{i=1}^m y_i^* g'_i(x^*) \psi'_{g_i}(\varepsilon, g_i(x^*)) = 0, \quad y_i^* g_i(x^*) = 0, \quad i = 1, \dots, m$$

(4) נניח כי בעיה (*) מקיימת $\text{rank } g^{(r)'} = r$, $y_i^* > 0$, $i = 1, \dots, r$

$$\text{נסמן } \psi^{(r)} = (\psi(\varepsilon, g_1(x)), \psi(\varepsilon, g_2(x)), \dots, \psi(\varepsilon, g_r(x)))$$

ובהתאמה ב- $\psi^{(r)'} \psi$ את מטריצת היעקוביאן שלו.

$$\text{עמודה } k\text{-ית של } \psi^{(r)'} \text{ היא מהצורה } \left[\psi^{(r)'} \right]_k = \psi'_{g_k}(\varepsilon, g_k(x^*)) \left[g^{(r)'} \right]_k$$

כיוון שהמרה ψ עולה לפי משתנה שני. לכן גם בעיה (2) מקיימת $\text{rank } \psi^{(r)'} = r$, $y_i^* > 0$, $i = 1, \dots, r$

(5) נניח כי בעיה (*) מקיימת תנאים מספיקים למינימום מסדר שני בנקודה (x^*, y^*) ,

כלומר, קיים $\lambda > 0$ כך ש- $\langle L''_{xx}(x^*, y^*) y, y \rangle \geq \lambda \|y\|^2$; $\forall y \neq 0$, $g^{(r)'} y = 0$ נראה שגם בעיה (**) מקיימת תנאים מתאימים.

$$\begin{aligned}
\langle F''_{xx}(\varepsilon, x^*, y^*) y, y \rangle &= \\
&= \left\langle \left(f''_{xx}(x^*) + s_\psi \sum_{i=1}^m y_i^* \psi''_{gg}(\varepsilon, g_i(x^*)) g_i^T(x^*) g'_i(x^*) + s_\psi \sum_{i=1}^m y_i^* \psi'_g(\varepsilon, g_i(x^*)) g''_i(x^*) \right) y, y \right\rangle \\
&= \left\langle \left(f''_{xx}(x^*) + s_\psi \sum_{i=1}^r y_i^* \psi'_g(\varepsilon, g_i(x^*)) g_i^T(x^*) g'_i(x^*) + s_\psi \sum_{i=1}^r y_i^* \psi''_{gg}(\varepsilon, g_i(x^*)) g_i^T(x^*) g'_i(x^*) \right) y, y \right\rangle \\
&= \left\langle \left(f''_{xx}(x^*) - \sum_{i=1}^r y_i^* g''_i(x^*) + s_\psi \sum_{i=1}^r y_i^* \psi''_{gg}(\varepsilon, g_i(x^*)) g_i^T(x^*) g'_i(x^*) \right) y, y \right\rangle
\end{aligned}$$

נסמן: $U = \text{diag}\{y_1, y_2, \dots, y_r\}$, נוכל לרשום כי

$$\langle F''_{xx}(\varepsilon, x^*, y^*) y, y \rangle = \langle L''_{xx}(x^*, y^*) y, y \rangle - \langle \psi^{(r)'} g^{(r)'} U^* g^{(r)'} y, y \rangle$$

לכן, לפי טענה 1 ב- [1] (ראה ציטוט בנספח 5) קיים $\varepsilon_0 > 0$, כך שלכל $\varepsilon > 0$ ו- $\lambda > \gamma > 0$ מתקיים

$$(ii) \quad \langle F''_{xx}(\varepsilon, x^*, y^*) y, y \rangle \geq \gamma \|y\|^2; \quad \gamma > 0, \forall y \neq 0, \quad \psi^{(r)'} y = 0$$

מש"ל

נתבונן בבעיית אופטימיזציה

$$\begin{aligned}
&\hat{x} = \arg \min \{ f(x) / x \in \Omega_\varepsilon \} \\
(***) \quad \Omega_\varepsilon &= \{ x \in \mathbb{R}^n / g_i(x) \geq -\varepsilon, \quad i = \overline{1, m} \}
\end{aligned}$$

נציין כי $\Omega \subset \Omega_\varepsilon$, בנוסף, אם Ω קומפקטית, אז גם Ω_ε קומפקטית. יהיו $\varepsilon_0, \delta, \mu > 0$ נגדיר תחום:

$$D(\varepsilon_0, y^*, \delta, \mu) = \left\{ (\varepsilon, y) \in \mathbb{R}^{m+1} \mid 0 < \varepsilon \leq \varepsilon_0 \text{ and } \begin{cases} y_i \geq \mu, & |y_i - y_i^*| \leq \frac{\delta}{\varepsilon}, & 1 \leq i \leq r \\ 0 \leq y_i \leq \frac{\delta}{\varepsilon}, & r+1 \leq i \leq m \end{cases} \right\}$$

כאשר ב- (x^*, y^*) מתקיימים תנאי אופטימאליות 3-5. של בעיה (*). נראה כי לבעיה (*) קיים פתרון $\hat{x}(\varepsilon, y)$

מעל $D(\varepsilon_0, y^*, \delta, \mu)$ ויתרה מזו $\hat{x}(\varepsilon, y^*) = x^*$

משפט 5.4

נניח כי בבעיה (*) מתקיים $f \in C^2$, $-g_i \in C^2$, $i = 1, \dots, m$ וכי בנקודה $x^* \in \Omega$ מתקיימים תנאים הכרחיים ומספיקים למינימום של בעיה (*). תהי נתונה המרה $\psi \in B$ (או על פי למה 1 לבעיה (**)) מתקיימים תנאים (3) – (5) (למעלה).

$$\text{תהי } F(\varepsilon, x, y) = f(x) + s_\psi \sum_{i=1}^m y_i \psi(\varepsilon, g_i(x)) \text{, כך ש- } s_\psi = -\text{sign}(\psi'_t(\varepsilon, t))$$

אז קיים $\varepsilon_0 > 0$ ו- $\delta > 0$ מספיק קטן, כך שלכל $0 < \mu < \min_{1 \leq i \leq m} y_i^*$ ו- $(\varepsilon, y) \in D(\varepsilon_0, y^*, \delta, \mu)$ מתקיים

$$1. \quad \text{קיים וקטור } \hat{x} = \hat{x}(\varepsilon, y) \text{ שעבורו מתקיים } \hat{x} = \arg \min \{ F(x, \varepsilon, y) / x \in \Omega_\varepsilon \} \text{, כך ש- } F'_x(\varepsilon, \hat{x}, y) = 0$$

$$2. \quad \text{לכל } \varepsilon > 0 \text{ מתקיים } \hat{x}(\varepsilon, y^*) = x^* \text{ ו- } \hat{y}(\varepsilon, y^*) = y^* \text{ (נקודת שבת של } \hat{y}(\varepsilon, y) \text{)}$$

$$3. \quad \text{לזוג וקטורים } \hat{x} = \hat{x}(\varepsilon, y) \text{ ו- } \hat{y} = \hat{y}(\varepsilon, y) = \text{diag}\{\psi'_g(\varepsilon, g_1(x)), \psi'_g(\varepsilon, g_2(x)), \dots, \psi'_g(\varepsilon, g_m(x))\} \cdot y$$

קיים קבוע $0 < c < 1$ שאינו תלוי ב- ε וכך ש-

$$(6) \quad \max \{ \|\hat{x} - x^*\|, \|\hat{y} - y^*\| \} < c\varepsilon \|y - y^*\|$$

$$4. \quad \text{הפונקציה } F(x, \varepsilon, y) \text{ קמורה ממש בסביבת } \hat{x}$$

הוכחה

1. בעיה (*) קמורה ו- Ω_ε קומפקטית. לפונקציה רציפה $F(\varepsilon, x, y) = f(x) + s_\psi \sum_{i=1}^m y_i \psi(\varepsilon, g_i(x))$

בתחום קומפקטי קיימת נקודת מינימום $\hat{x} = \hat{x}(y, \varepsilon) = \arg \min \{F(x) / x \in \Omega_\varepsilon\}$.

$\psi(\varepsilon, t) \in B$, לכן לפי דרישה (ט) של הגדרת B מתקיים כי $\lim_{t \rightarrow -\varepsilon} F(\varepsilon, x, t) = \infty$, כלומר המינימום שלה

לא מתקבל על השפה של Ω_ε אלא בנקודה פנימית. כיוון שהבעיה שלנו גזירה ברציפות מתקיים $F'_x(\varepsilon, \hat{x}, y) = 0$.

2. נסמן:

$$t = (t_1, t_2, \dots, t_m), \quad t_i = \varepsilon(y_i - y_i^*), \quad i = \overline{1, m}$$

$$B_\delta(0) = \{t \mid |t_i| < \delta, i = \overline{1, m}\}$$

ונגדיר פונקציה $y(\varepsilon, x, t): R^{n+m+1} \rightarrow R_+^m$ כך ש-

$$\hat{y}_i(\varepsilon, x, t) = y_i \psi'_{g_i}(\varepsilon, g_i(x)) = \begin{cases} \left(\frac{t_i}{\varepsilon} + y_i^*\right) \psi'_{g_i}(\varepsilon, g_i(x)), & i = \overline{1, r} \\ \frac{t_i}{\varepsilon} \psi'_{g_i}(\varepsilon, g_i(x)), & i = \overline{1+r, m} \end{cases}$$

מתקיים

$$\hat{y}'_{i\varepsilon}(\varepsilon, x, t) = -\frac{t_i}{\varepsilon^2} \psi'_{g_i}(\varepsilon, g_i(x)) + \frac{t_i}{\varepsilon} \psi''_{g_i\varepsilon}(\varepsilon, g_i(x)), \quad i = \overline{1, m}$$

ובפרט עבור $t = \vec{0}$ מתקיים

$$(7) \quad \hat{y}_{i\varepsilon}(\varepsilon, x^*, 0) = 0, \quad \hat{y}'_{i\varepsilon}(\varepsilon, x^*, 0) = 0, \quad i = \overline{1, m}$$

נגדיר פונקציה

$$(8) \quad h(\varepsilon, x, t) = \sum_{i=r+1}^m \hat{y}_i(\varepsilon, x, t) g_i'^T(x) = \frac{1}{\varepsilon} \sum_{i=r+1}^m t_i \psi'_{g_i}(\varepsilon, g_i(x)) g_i'^T(x)$$

לכל $\varepsilon > 0$ הפונקציה $h(\varepsilon, x, t)$ גזירה ברציפות עבור $x \in B_\mu(x^*)$ ו- $t \in B_\delta(0)$ ומקיימת

$$(9) \quad h'_{y^r}(\varepsilon, x^*, 0) = \vec{0} \in \mathbb{R}^{n \times r}, \quad h'_\varepsilon(\varepsilon, x^*, 0) = \vec{0} \in \mathbb{R}^{n \times n}, \quad h(\varepsilon, x^*, 0) = \vec{0} \in \mathbb{R}^n$$

($\hat{y}^{(r)}$) לא תלויה ברכיבים h .

נגדיר פונקציה $\Phi(\varepsilon, x, y, t): \mathbb{R}^{1+n+r+m} \rightarrow \mathbb{R}^{n+r}$ על תחום $(0, \infty) \times B_\mu(x^*) \times B_\mu(x^*) \times B_\delta(0)$ בצורה

$$\Phi(\varepsilon, x, \hat{y}^{(r)}, t) :=$$

$$\left(\nabla f(x) + s_\psi \sum_{i=1}^r \hat{y}_i \nabla g_i(x) + s_\psi h(\varepsilon, x, t); \quad \varepsilon^k s_\psi \left(\frac{t_i}{\varepsilon} + y_i^* \right) \psi'_{g_i}(\varepsilon, g_i(x)) - \varepsilon^k s_\psi \hat{y}_i, i = \overline{1, r} \right)$$

כך שמתקיים $\lim_{\varepsilon \rightarrow 0} \varepsilon^k \psi''_{g_i\varepsilon}(\varepsilon, 0) = l$ ו- $0 \neq l \in \mathbb{R}$ כפי שנדרש ב-(ח) של הגדרת מחלקת B .

נציב נקודה $(\varepsilon, x, y, t) = (\varepsilon, x^*, y^{(r)*}, 0)$ ונקבל לפי תנאי (i) לאופטימאליות של בעיה (*) בנקודה (x^*, y^*) :

$$\begin{aligned} \Phi(\varepsilon, x^*, y^{(r)*}, 0) &= \\ &= \left(f'(x^*) + s_\psi \sum_{i=1}^r y_i^* g'_i(x^*) + s_\psi h(\varepsilon, x^*, 0); \quad \varepsilon^k \left(\frac{0}{\varepsilon} + y_i^* \right) \psi'_{g_i}(\varepsilon, g_i(x^*)) + \varepsilon^k s_\psi y_i^*, i = \overline{1, r} \right) = \\ &= (L'_x(x^*, y^*); \varepsilon^k y_i^* (\psi'_{g_i}(\varepsilon, 0) + s_\psi), i = \overline{1, r}) = \vec{0} \in \mathbb{R}^{n+r} \end{aligned}$$

$$\left[\Phi(\varepsilon, x, \hat{y}^{(r)}, t) \right]_{xy}'' = \begin{pmatrix} f''(x) + s_\psi \sum_{i=1}^r \hat{y}_i g''_i(x) + s_\psi h'(\varepsilon, x, t); & s_\psi \left(\frac{t_i}{\varepsilon} + y_i^* \right) \varepsilon^k \psi''_{g_i g_i}(\varepsilon, g_i(x)) g'_i(x), i = \overline{1, r} \\ s_\psi \sum_{i=1}^r g'_i(x); & \varepsilon^k s_\psi, i = \overline{1, r} \end{pmatrix}$$

נציב נקודה $(\varepsilon, x, y, t) = (\varepsilon, x^*, y^{(r)*}, 0)$ ונקבל,

$$\left[\Phi(\varepsilon, x^*, y^{(r)*}, 0) \right]_{xy}'' = \begin{pmatrix} f''(x^*) + s_\psi \sum_{i=1}^r y_i^* g''_i(x^*) + s_\psi \underbrace{h'_x(\varepsilon, x^*, 0)}_{=0}; & s_\psi \varepsilon^k y_i^* \psi''_{g_i g_i}(\varepsilon, g_i(x^*)) g'_i(x^*), i = \overline{1, r} \\ s_\psi \sum_{i=1}^r g'_i(x^*); & \varepsilon^k s_\psi, i = \overline{1, r} \end{pmatrix}$$

אם נסמן:

$$\begin{aligned} \Phi'_{(\varepsilon)} &= \Phi'_{xy}(\varepsilon, x^*, y^{(r)*}, 0), \quad L''_{xx} = L''_{xx}(x^*, y^*), \\ U^* &= \text{diag} \{ y_i^* \varepsilon^k \psi''_{gg}(\varepsilon, g_i(x^*)), i = \overline{1, r} \} \end{aligned}$$

נציב נקודה $(\varepsilon, x, y, t) = (\varepsilon, x^*, y^*, 0)$ ונקבל

$$\Phi'_{(\varepsilon)} = \begin{pmatrix} L''_{xx} & s_\psi U^* g^{(r)'} \\ s_\psi g^{(r)'} & -\varepsilon^k s_\psi I^r \end{pmatrix} \Rightarrow \Phi'_{(0)} = \begin{pmatrix} L''_{xx} & s_\psi U^* g^{(r)'} \\ s_\psi g^{(r)'} & 0^{r \times r} \end{pmatrix}$$

מלמת עזר 2 (הוכחה בנספח 6) נובע כי $\Phi'_{(0)}$ היא מטריצה לא סינגולארית, לכן נורמה שלה חסומה, כלומר קיים

סקלר $\lambda_0 > 0$ שאינו תלוי ב- ε כך ש- $\|\Phi'_{(0)}\| < \lambda_0$. בנוסף קיים סקלר $\lambda > 0$, כך שלכל $w \in \mathbb{R}^{n+r}$ מתקיים

$$\langle \Phi'^T_{(0)} \Phi'_{(0)} w, w \rangle > \lambda \|w\|^2$$

בגלל הרציפות לפי ε מתקיים $\lim_{\varepsilon \rightarrow 0} \langle \Phi'^T_{(\varepsilon)} \Phi'_{(\varepsilon)} w, w \rangle = \langle \Phi'^T_{(0)} \Phi'_{(0)} w, w \rangle \geq \lambda \|w\|^2$ לכן קיים $\varepsilon_0 > 0$ כך

שלכל $\varepsilon_0 > \varepsilon > 0$ מתקיים $\langle \Phi'^T_{(\varepsilon)} \Phi'_{(\varepsilon)} w, w \rangle \geq \frac{1}{2} \lambda \|w\|^2$ לכל $w \in \mathbb{R}^{n+r}$.

כלומר, קיים $\varepsilon_0 > 0$ כך שלכל $\varepsilon_0 > \varepsilon > 0$ המטריצה $\Phi'_{(\varepsilon)} \Phi'^T_{(\varepsilon)}$ אינה סגולרית ולכן יש לה מטריצה הפוכה

לא סינגולרית. כלומר, קיים סקלר $\rho > 0$ שאינו תלוי ב- $\varepsilon_0 > \varepsilon > 0$ כך ש- $\|\Phi'^{-1}_{(\varepsilon)}\| < \rho$.

נזכיר כי לכל $\varepsilon_0 > \varepsilon > 0$ מתקיים $\Phi(\varepsilon, x^*, y^{(r)*}, 0) = \vec{0} \in \mathbb{R}^{n+r}$, $f \in C^2$, $g_i \in C^2$, $i = 1, \dots, m$.

יהיה $\varepsilon_0 > \varepsilon_1 > 0$, ו- $E = [\varepsilon_1, \varepsilon_0] \times \{\vec{0} \in \mathbb{R}^m\} = \{(\varepsilon, \vec{0}) \mid \varepsilon_1 \leq \varepsilon \leq \varepsilon_0, \vec{0} \in \mathbb{R}^m\}$ ממשפט הפונקציות

הסתומות 2 (ראה נספח 7) נובע שקיים $\delta > 0$ ווקטורי פונקציות חלקים

$$\hat{y}(\varepsilon, t) = (\hat{y}_1(\varepsilon, t), \hat{y}_2(\varepsilon, t), \dots, \hat{y}_r(\varepsilon, t)), \quad \hat{x}(\varepsilon, t) = (\hat{x}_1(\varepsilon, t), \hat{x}_2(\varepsilon, t), \dots, \hat{x}_n(\varepsilon, t))$$

המוגדרים באופן יחיד בסביבה $B_\delta(E) = \{(\varepsilon, s) \in \mathbb{R}^{m+1} \mid |s_i| \leq \delta; \varepsilon \in [\varepsilon_1, \varepsilon_0]; i = \overline{1, m}\}$ של

קבוצה קומפקטית E כך שמתקיים $\hat{y}(\varepsilon, 0) = (y_1^*, y_2^*, \dots, y_r^*)$, $\hat{x}(\varepsilon, 0) = (x_1^*, x_2^*, \dots, x_n^*)$ לכל $\varepsilon_0 > \varepsilon > 0$.

(במילים אחרות, $\vec{0} = t = y - y^*$) והוכחנו את הטענה 2. של המשפט.

3. כיוון שהפונקציות $\hat{y}(\varepsilon, t)$, $\hat{x}(\varepsilon, t)$ חלקות בתחום קומפקטי $[\varepsilon_1, \varepsilon_0] \times \Omega_\varepsilon$ קיים $\eta > 0$ כך שמתקיים

$$\max \{\|\hat{x}(\varepsilon, t) - x^*\|, \|\hat{y}(\varepsilon, t) - y^*\|\} \leq \eta_0,$$

$$(10) \quad \Phi(\varepsilon, x(\varepsilon, t), \hat{y}^r(\varepsilon, t), t) \equiv \vec{0} \in \mathbb{R}^{n+r}$$

וגם

$$\|\Phi'_{x\hat{y}^r}(\varepsilon, x(\varepsilon, t), \hat{y}^{(r)}(\varepsilon, t), t)^{-1}\| \leq 2\rho, \quad \forall (\varepsilon, t) \in B_\delta(E)$$

נרשום את (6) בצורה מפורשת,

$$(11) \quad f'(x(\varepsilon, t)) + s_\psi \sum_{i=1}^r \hat{y}_i(\varepsilon, t) g'_i(x(\varepsilon, t)) + s_\psi h(\varepsilon, x(\varepsilon, t), t) = 0;$$

$$(12) \quad \hat{y}_i(\varepsilon, t) = \left(\frac{t_i}{\varepsilon} + y_i^* \right) \psi'_{g_i}(\varepsilon, g_i(x(\varepsilon, t))), i = \overline{1, r}$$

ונבחר (כמו קודם)

$$(13) \quad \hat{y}_i(\varepsilon, t) = \frac{t_i}{\varepsilon} \psi'_{g_i}(\varepsilon, g_i(x(\varepsilon, t))), i = \overline{r+1, m}$$

ראשית נעריך את $\|\hat{y}^{(m-r)}(\varepsilon, t)\|$,

$$\hat{y}^{(m-r)}(\varepsilon, t) = \frac{1}{\varepsilon} \text{diag} \{ \psi'_g(\varepsilon, g_r(x)), \psi'_g(\varepsilon, g_2(x)), \dots, \psi'_g(\varepsilon, g_m(x)) \} \cdot y^{(m-r)}$$

יהיה $\delta > 0$ מספיק קטן, אז לכל $(\varepsilon, t) \in B_\delta(E) \subseteq [\varepsilon_1, \varepsilon_0] \times \Omega_\varepsilon$ מתקיים

$$\|\hat{x}(\varepsilon, t) - x(\varepsilon, 0)\| = \|\hat{x}(\varepsilon, t) - x^*\| \leq \eta$$

ולפי סעיף 1. בהגדרת מחלקה B , קיים קבוע $\mu > 0$ כך ש-

$$(14) \quad \max_{r+1 \leq i \leq m} \left| \frac{1}{\varepsilon} \psi'_{g_i}(\varepsilon, g_i(x)) \right| < \mu$$

לכן

$$|\hat{y}_i(\varepsilon, t)| = \left| \frac{t_i}{\varepsilon} \psi'_{g_i}(\varepsilon, g_i(x(\varepsilon, t))) \right| \leq \mu |t_i| = \mu \varepsilon |y_i - y^*_i| = \mu \varepsilon |y_i| = \mu \varepsilon y_i, \quad i = \overline{r+1, m}$$

$$|\hat{y}_i(x, \varepsilon) - y^*_i(x, \varepsilon)| = \mu \varepsilon |y_i - y^*_i|, \quad i = \overline{r+1, m} \quad \text{למעשה קיבלנו כי}$$

או גם ש-

$$\|\hat{y}(x, \varepsilon)\| = \|\hat{y}(x, \varepsilon) - y^*(x, \varepsilon)\| \leq \mu \varepsilon \|y - y^*\|$$

כעת נגזור את (7) ביחס ל- t :

$$(15) \quad f''_{xx}(x(\varepsilon, t)) x'_t(\varepsilon, t) + s_\psi \sum_{i=1}^r (\hat{y}'_{it}(\varepsilon, t) g'_i(x(\varepsilon, t)) + \hat{y}_i(\varepsilon, t) g''_{ix}(x(\varepsilon, t)) x'_t(\varepsilon, t)) \\ + s_\psi h'_x(\varepsilon, x(\varepsilon, t), t) x'_t(\varepsilon, t) + s_\psi h'_t(\varepsilon, x(\varepsilon, t), t) = 0;$$

כלומר:

$$\left(f''_{xx}(x(\varepsilon, t)) + s_\psi \sum_{i=1}^r (\hat{y}_i(\varepsilon, t) g''_{ix}(x(\varepsilon, t)) + h'_x(\varepsilon, x(\varepsilon, t), t)) \right) x'_t(\varepsilon, t) \\ = -s_\psi \left(\sum_{i=1}^r \hat{y}'_{it}(\varepsilon, t) g'_i(x(\varepsilon, t)) + h'_t(\varepsilon, x(\varepsilon, t), t) \right);$$

במילים אחרות:

$$(16) \quad L''_{xx}(x(\varepsilon, t), \hat{y}(\varepsilon, t)) x'_t(\varepsilon, t) + s_\psi g^{(r)'}(x(\varepsilon, t))^T \hat{y}^{(r)'}_t(\varepsilon, t) = s_\psi h'_t(\varepsilon, x(\varepsilon, t), t)$$

כאשר:

$$(17) \quad L''_{xx}(x(\varepsilon, t), \hat{y}(\varepsilon, t)) = f''_{xx}(x(\varepsilon, t)) + s_\psi \sum_{i=1}^r (\hat{y}_i(\varepsilon, t) g''_{ix}(x(\varepsilon, t)) + h'_x(\varepsilon, x(\varepsilon, t), t))$$

$$x'_t(\varepsilon, t) = J_t x(\varepsilon, t), \quad y'_t(\varepsilon, t) = J_t y(\varepsilon, t)$$

נגזור את (12) לפי t ונקבל:

$$\hat{y}'_{it_i}(\varepsilon, t) = \frac{1}{\varepsilon} \psi'_{g_i}(\varepsilon, g_i(x(\varepsilon, t))) + \left(\frac{t_i}{\varepsilon} + y_i^* \right) \psi''_{g_i g_i}(\varepsilon, g_i(x(\varepsilon, t))) g'_{ix}(x(\varepsilon, t)) x'_t(\varepsilon, t), \quad i = \overline{1, r}$$

נוכל לרשום זאת בצורה:

$$-\varepsilon \left(\frac{t_i}{\varepsilon} + y_i^* \right) \psi''_{g_i g_i}(\varepsilon, g_i(x(\varepsilon, t))) g'_{ix}(x(\varepsilon, t)) x'_t(\varepsilon, t) + \varepsilon \hat{y}'_{it_i}(\varepsilon, t) = \psi'_{g_i}(\varepsilon, g_i(x(\varepsilon, t))), \quad i = \overline{1, r}$$

נסמן:

$$(18) \quad D_r^{-1}(\varepsilon, t) = \text{diag} \left[\psi'_{g_i}(\varepsilon, g_i(x(\varepsilon, t))) \right]_{i=1}^r \in \mathbb{R}^{r \times r}$$

$$(19) \quad U(\varepsilon, t) = \text{diag} \left\{ \hat{y}_i(\varepsilon, t) \varepsilon^k \psi''_{gg}(\varepsilon, g_i(x(\varepsilon, t))), i = \overline{1, r} \right\} \in \mathbb{R}^{r \times r}$$

כך שמתקיים $\lim_{\varepsilon \rightarrow 0} \varepsilon^k \psi''_{tt}(\varepsilon, 0) = l$ ו- $0 \neq l \in \mathbb{R}$ כפי שנדרש ב-(ח) של הגדרת מחלקת B .

ואז עבור $\delta > 0$ מספיק קטן ו- $0 < t < \delta$, נוכל לרשום באופן וקטורי:

$$(20) \quad -U(\varepsilon, t) g^{(r)'}(x(\varepsilon, t)) x'_t(\varepsilon, t) + \varepsilon^k \hat{y}'_t(\varepsilon, t) = \left[\varepsilon^{k-1} D_r^{-1}(\varepsilon, x); 0^{r, m-r} \right]$$

נגדיר פונקציה :

$$P'(\varepsilon, t) = \begin{pmatrix} L''_{xx}(\varepsilon, x(\varepsilon, t), t) & s_\psi g^{(r)'}(x(\varepsilon, t)) \\ s_\psi U(\varepsilon, t) g^{(r)'}(x(\varepsilon, t)) & -s_\psi \varepsilon^k I^r \end{pmatrix}$$

נשים לב כי $\Phi'_{(\varepsilon)}{}^T(x(\varepsilon, 0), y(\varepsilon, 0)) = P'_{xy}(x(\varepsilon, 0), y(\varepsilon, 0))$ ולכן היא מטריצה הפיכה לכל $(\varepsilon, t) \in B_\delta(E)$.

נשלב את (12) ו-(16) ונקבל :

$$(21) \quad w'_t(\cdot) = \begin{bmatrix} x'_t(\cdot) \\ \hat{y}'_t(\cdot) \end{bmatrix} = P'^{-1}(\cdot) \begin{bmatrix} h'_t(x(\cdot), \cdot) \\ [-s_\psi \varepsilon^{k-1} D_r(\varepsilon, t); 0^{r, m-r}] \end{bmatrix} = P'^{-1}(\cdot) R(\cdot)$$

נסמן :

$$g^{m-r}(x(\varepsilon, t)) = (g_i(x(\varepsilon, t)), i = \overline{r+1, m})$$

$$\hat{y}^{m-r}(x(\varepsilon, t)) = (y_i, i = \overline{r+1, m})$$

ואז נוכל לרשום :

$$h(\varepsilon, x, t) = \sum_{i=r+1}^m \hat{y}_i(\varepsilon, x, t) g_i'^T(x) = (\hat{y}^{m-r}(\varepsilon, x, t))^T g^{(m-r)'}(x)$$

$$h'_t(\varepsilon, x, t) = (\hat{y}^{(m-r)'}(\varepsilon, x, t))^T g_i^{(m-r)'}(x) + (\hat{y}^{(m-r)}(\varepsilon, x, t))^T g_i^{(m-r)''}(x) x'_t(\cdot)$$

כאשר

$$\hat{y}_t^{(m-r)'}(x(\varepsilon, t), \varepsilon) = (\hat{y}'_{it}(\varepsilon, t), i = \overline{r+1, m})$$

$$= \left(\frac{1}{\varepsilon} \left\{ \psi_{g_i}'(\varepsilon, g_i(x(\varepsilon, t))) + t_i \psi''_{g_i g_i}(\varepsilon, g_i(x(\varepsilon, t))) g_i'(x(\varepsilon, t)) x'(\varepsilon, t) \right\}, i = \overline{r+1, m} \right)$$

הראנו ב-2. כי עבור כל $(\varepsilon, \vec{0}) \in \text{int } B_\delta(E)$ מתקיים :

$$\hat{y}(\varepsilon, 0) = (y_1^*, y_2^*, \dots, y_r^*) = y^{(r)*} > 0, \quad \hat{x}(\varepsilon, 0) = (x_1^*, x_2^*, \dots, x_n^*) = x^*,$$

$$\hat{y}^{(m-r)}(x(\varepsilon, 0)) = 0 \in \mathbb{R}^{m-r}$$

לכן :

$$\hat{y}_t^{(m-r)'}(x(\varepsilon, 0), \varepsilon) = \left(\frac{1}{\varepsilon} \psi_{g_i}'(\varepsilon, g_i(x(\varepsilon, 0))), i = \overline{r+1, m} \right)$$

לפי (14) :

$$\hat{y}_t^{(m-r)'}(x(\varepsilon, 0), \varepsilon) = \left(O^{(m-r) \times r}; \frac{1}{\varepsilon} D_{m-r}^{-1}(x^*) \right) \leq \left(O^{(m-r) \times r}; \mu I^{m-r} \right)$$

לכן

$$\|h'_t(x(\varepsilon, 0), \varepsilon, 0)\| = \left\| \left(g_i^{(m-r)'}(x(\varepsilon, 0)) \right)^T \hat{y}^{(m-r)'}(x(\varepsilon, 0), \varepsilon, 0) \right\|$$

$$= \left\| \left(g_i^{(m-r)'}(x^*) \right)^T \left[0^{m-r, r}; \frac{1}{\varepsilon} D_{m-r}^{-1}(x(\varepsilon, 0)) \right] \right\| \leq \mu \|g_i^{(m-r)'}(x^*)\|$$

כמו כן, מ-(18) נובע כי $D_r(\varepsilon, 0) = s_\psi I^r$, לכן

$$w'_t(x(\varepsilon, 0), y(\varepsilon, 0)) = \begin{bmatrix} x'_t(\varepsilon, 0) \\ \hat{y}'_t(\varepsilon, 0) \end{bmatrix} = P'^{-1} \begin{bmatrix} h'_t(x(\varepsilon, 0), y(\varepsilon, 0)) \\ [s_\psi \varepsilon^{k-1} I^r; 0^{r, m-r}] \end{bmatrix} = \Phi'^{-1T}_\varepsilon R_0$$

אם ניקח בחשבון ש- $\rho < \|\Phi'^{-1}_\varepsilon\|$ ו- $\|h'_t(x(\varepsilon, 0), y(\varepsilon, 0))\| \leq \mu \|g_i^{(m-r)'}(x^*)\|$, נקבל

$$\max \{\|x'_t(\varepsilon, 0)\|, \|\hat{y}'_t(\varepsilon, 0)\|\} \leq \rho \left(\mu \|g_i^{(m-r)'}(x^*)\| + \|\varepsilon^{k-1} I^r\| \right) \leq \rho \left(\mu \|g_i^{(m-r)'}(x^*)\| + 1 \right) := c$$

לכן מחלקות של הפונקציות נובע שעבור $\delta > 0$ מספיק קטן לכל $(\varepsilon, t) \in B_\delta(E)$ ו- $0 < \tau < 1$ מתקיים

$$\|\Phi'^{-1T}(x(\varepsilon, \tau t), \hat{y}(\varepsilon, \tau t)) \cdot R(x(\varepsilon, t), \hat{y}(\varepsilon, t))\| \leq c$$

ומכאן נקבל:

$$\begin{bmatrix} x(\varepsilon, t) - x^* \\ \hat{y}(\varepsilon, t) - y^* \end{bmatrix} = \begin{bmatrix} x(\varepsilon, t) - x(\varepsilon, 0) \\ \hat{y}(\varepsilon, t) - \hat{y}(\varepsilon, 0) \end{bmatrix} = \int_0^1 \Phi'^{-1T}(x(\varepsilon, \tau t), \hat{y}(\varepsilon, \tau t)) \cdot R(x(\varepsilon, t), \hat{y}(\varepsilon, t)) d\tau$$

כלומר,

$$\|x(\varepsilon, t) - x^*, \hat{y}(\varepsilon, t) - y^*\| \leq \int_0^1 \|\Phi'^{-1T}(x(\varepsilon, \tau t), \hat{y}(\varepsilon, \tau t)) \cdot R(x(\varepsilon, t), \hat{y}(\varepsilon, t))\| t d\tau \leq c \|t\|$$

כלומר,

$$\|x(\varepsilon, t) - x^*\|^2 + \|\hat{y}(\varepsilon, t) - y^*\|^2 \leq c^2 \varepsilon^2 \|y - y^*\|^2$$

ולכן הוכחנו כי

$$\|x(\varepsilon, t) - x^*\| \leq c\varepsilon \|y - y^*\|, \quad \|\hat{y}(\varepsilon, t) - y^*\| \leq c\varepsilon \|y - y^*\|$$

4. הראנו בלמה 1 סעיף 6, כי אם עבור בעיה (*) קיים $\lambda > 0$ כך ש-

$$\langle L''_{xx}(x^*, y^*) y, y \rangle \geq \lambda \|y\|^2; \quad \forall y \neq 0, \quad g^{(r)'} y = 0$$

$$\varepsilon_0 > \varepsilon > 0 \text{ ו- } \lambda > \gamma > 0 \text{ מתקיים}$$

$$F(\varepsilon, x, y) \text{ הפונקציה } \langle F''_{xx}(\varepsilon, x^*, y^*) y, y \rangle \geq \gamma \|y\|^2; \quad \gamma > 0, \forall y \neq 0, \quad g^{(r)'} y = 0$$

קמורה ממש בסביבת $\hat{x}$.

מסקנה 5.5:

המשוואות (5), (7) יחד עם קמורות ממש של $F(x, \varepsilon, y)$ בסביבת $\hat{x}$ גוררות ש- $\hat{x}$ מקיים תנאי אופטימאליות למינימום לוקאלי.

6. תיאור השיטה האיטרטיבית לפתרון

בפרק זה נציג שיטת האיטרציות לבעיית אופטימיזציה (*) שהציע Polyak [12] נסביר את השיטה בעזרת התוצאות של משפט 5.4 ולמה 5.3.

ונניח ש $(x^k, y^k) \in \mathbb{R}^{n+m}$ הוא פתרון של איטרציה $k \in \mathbb{N}$. באיטרציה $k+1$ נחשב:

$$(A) \quad \begin{cases} x^{k+1} = \arg \min(F(\varepsilon, x, y^k)) \\ y^{k+1} = [\text{diag} \psi'_g(\varepsilon, g_i(x^{k+1}))]_{i=1}^m y^k \end{cases}$$

כתוצאה מכך נוצרת סדרה $\{(x^k, y^k)\}$.

למימוש השיטה פותרים בכל שלב את בעיית האופטימיזציה $x^{k+1} = \arg \min(F(\varepsilon, x, y^k))$ לצורך זה ניתן להשתמש באחת משיטות סטנדרטיות של אופטימיזציה לא מאולצת.

כדי לאתחל את האיטרציות צריך לבחור את $(\varepsilon, y^0) \in D(\varepsilon_0, y^*, \delta, \mu)$.

6.1 מסקנה

נניח כי בבעיה (*) הפונקציות $f \in C^2$, $-g_i \in C^2$, $i=1, \dots, m$ הן קמורות וכי בנקודה $x^* \in \Omega$ מתקיימים תנאים הכרחיים ומספיקים למינימום של בעיה (*). לפי למה 5.3 פונקציה גם הבעיה (**) מקיימת תנאי אופטימאליות (3) – (5) בנקודה $x^* \in \Omega$. ואז לפי משפט 5.4 קיים זוג וקטורים (x^{k+1}, y^{k+1}) , כך ש-
 $x^{k+1} = \arg \min \{F(\varepsilon, x, y^{k+1}) / x \in \Omega_\varepsilon\}$ ו- $y^{k+1} = [\text{diag} \psi'_g(\varepsilon, g_i(x^{k+1}))]_{i=1}^m y^k$ עבורו לכל $\varepsilon > 0$ קיים $0 < c < 1$ שאינו תלוי ב- ε כך שמתקיים $\max \{\|x^{k+1} - x^*\|, \|y^{k+1} - y^*\|\} < c\varepsilon \|y^k - y^*\|$.
 כלומר, בכל איטרציה של (A) מתקרבים לפתרון (x^*, y^*) , וקצב ההתכנסות הינו ליניארי.

6.2 הערה

את הבחירה של נקודה התחלתית $(\varepsilon, y^0) \in D(\varepsilon_0, y^*, \delta, \mu)$ נעשה באופן הבא:

נציב $y^0 = e = (1, \dots, 1) \in \mathbb{R}^m$ בפונקציה $F(\varepsilon, x, y)$ ונגדיר פונקציה חדשה

$$M(\varepsilon, x) = F(\varepsilon, x, e) = f(x) + \sum_{i=1}^m e_i \psi(\varepsilon, g_i(x))$$

נניח גם כאן, ללא הגבלת הכלליות כי $\psi(\varepsilon, x) \in B$ היא פונקציה עולה וקעורה (מתקיים $\psi(\varepsilon, g_i(x)) \leq 0$ לפי (1) של למה 5.3). אז הבעיה:

$$\begin{aligned} (B_M(\varepsilon)) \quad x_\varepsilon &= \arg \min \{M(\varepsilon, x) / x \in \Omega\} \\ \Omega_B &= \{x \in \mathbb{R}^n / g_i(x) > 0, \quad i = \overline{1, m}\} \end{aligned}$$

מקיימת את הדרישות של למה 3.1 ושל משפט 3.2, לכן ניתן להוכיח (ראה משפט 6.2) שאם לבעיה (*) יש פתרון אז הוא גם המינימום של הפונקציה $M(x, \varepsilon)$.

6.3 הערה

נציין שאם Ω היא קבוצה קומפקטית אזי קיים $\varepsilon_0 > 0$ כך שמתקיים אי-שוויון הבא:

$$(22) \quad f(x) \geq f(x) + \sum_{i=1}^m \psi(\varepsilon, g_i(x)) \geq f(x) + s_\psi O(\varepsilon); \quad \forall x \in \Omega, \quad 0 < \varepsilon < \varepsilon_0$$

המשפט הבא משתמש ב-(22) למציאת $(\varepsilon, y^0) \in D(\varepsilon_0, y^*, \delta, \mu)$.

משפט 6.4:

תהי $\psi(\varepsilon, t) \in B$

א. אם הבעיה (*) מקיימת תנאי Slater, הפונקציות $f(x), g_i(x), i=1, \dots, m$ גזירות וקיים $\varepsilon_0 > 0$

שעבורו Ω_{ε_0} היא קבוצה קומפקטית, אז לכל $\varepsilon > 0$

• קיים $x(\varepsilon) = \arg \min \{M(x, \varepsilon) / x \in \Omega_\varepsilon\}$ כך ש- $M'_x(x(\varepsilon), \varepsilon) = 0$

• $\lim_{\varepsilon \rightarrow 0} f(x(\varepsilon)) = \lim_{\varepsilon \rightarrow 0} M(x(\varepsilon), \varepsilon) = f(x^*)$

ב. אם $f \in C^2, g_i \in C^2, i=1, \dots, m$ ומתקיימים תנאים (4) ו-(5) אזי קיים $\varepsilon_0 > 0$ כך שלכל $\varepsilon > 0$

מתקיים: $x(\varepsilon) = \arg \min \{M(x, \varepsilon) / x \in \Omega_\varepsilon\}$ כך ש- $M'_x(x(\varepsilon), \varepsilon) = 0$

ועבור $y(\varepsilon) = y(x(\varepsilon)) = [\text{diag } \psi'_g(\varepsilon, g_i(x))]_{i=1}^m e$ לזוג $(x(\varepsilon), y(\varepsilon))$ מתקיים

$$(23) \quad \|x(\varepsilon) - x^*\| \leq c\varepsilon, \quad \|y(\varepsilon) - y^*\| \leq c\varepsilon$$

המספר $0 < c < 1$ אינו תלוי ב- $\varepsilon_0 > 0$.

ג. תחת הנחות מ-ב. הפונקציה $M(x, \varepsilon)$ קמורה במובן חזק בסביבת הנקודה $x(\varepsilon)$.

הוכחה של משפט 6.4:

א. מתקיים $\lim_{x \rightarrow \partial \Omega} M(\varepsilon, x) = \infty$ ו- $\Omega_\varepsilon \xrightarrow{\varepsilon \rightarrow 0} \Omega$

$M(x, \varepsilon)$ רציפה בקבוצה קומפקטית Ω_{ε_0} וגדלה בצורה לא חסומה ככל שמתקרבים לשפה של Ω .

כלומר, המינימום של $x(\varepsilon)$ של $M(x, \varepsilon)$ מתקבל ב- $\text{int } \Omega_{\varepsilon_0}$ ולכן $M'_x(x(\varepsilon), \varepsilon) = 0$.

תהי $\{x(\varepsilon_k)\}_{k=1}^\infty$ סידרה כלשהי של $\{x(\varepsilon)\}_{0 < \varepsilon \leq \varepsilon_0} \subset \text{int } \Omega_{\varepsilon_0}$, כך ש- $\lim_{k \rightarrow \infty} \varepsilon_k = 0$ ולכל מתקיים $\varepsilon_0 > \varepsilon_k > 0$. לפי למה 3.1 ומשפט 3.2 $\{f(x(\varepsilon_k))\}_{k=1}^\infty$ סדרה יורדת וחסומה (f רציפה בתחום

קומפקטי), לכן קיים $\bar{x} = \lim_{k \rightarrow \infty} x(\varepsilon_k) \in \text{int } \Omega_{\varepsilon_0}$ ו- $\bar{x} = \arg \min_{x \in \Omega_\varepsilon} f(x)$.

על פי נוסחה (22) מתקיים: $f(\bar{x}) \leq M(\bar{x}, \varepsilon_k) - s_\psi O(\varepsilon_k)$

לכן לכל $\delta > 0$ מספיק קטן, קיים $k_0 \in \mathbb{N}$ כך שלכל $k > k_0$ אפשר לרשום כי

$$f(x(\varepsilon_k)) - \delta < f(\bar{x}) < f(x(\varepsilon_k)) + \delta \quad \text{וגם} \quad f(\bar{x}) \leq M(\bar{x}, \varepsilon_k) + \delta$$

$$\text{אז} \quad f(x(\varepsilon_k)) - \delta < f(\bar{x}) \leq M(\bar{x}, \varepsilon_k) + \delta \leq M(x^*, \varepsilon_k) + \delta \leq f(x^*) + \delta$$

כלומר, לכל $\delta > 0$ מספיק קטן מתקיים $f(x(\varepsilon_k)) < f(x^*) + 2\delta$

ולכן $\lim_{k \rightarrow \infty} f(x(\varepsilon_k)) = f(\bar{x}) \leq f(x^*)$, אך מצד שני $f(\bar{x}) \geq f(x^*)$ לכן $f(\bar{x}) = f(x^*)$.

ואז לפי (22) נקבל $\lim_{\varepsilon \rightarrow 0} f(x(\varepsilon_k)) = \lim_{\varepsilon \rightarrow 0} M(x(\varepsilon_k), \varepsilon_k) = f(x^*)$ ולפי היינה גם

$$\lim_{\varepsilon \rightarrow 0} f(x(\varepsilon)) = \lim_{\varepsilon \rightarrow 0} M(x(\varepsilon), \varepsilon) = f(x^*)$$

ב. אם (4) ו-(5) מתקיימים אזי (x^*, y^*) הוא זוג KKT היחיד ולכן:

$$\lim_{\varepsilon \rightarrow 0} x(\varepsilon) = x^*, \lim_{\varepsilon \rightarrow 0} y(\varepsilon) = y^* \quad \text{و} \quad \lim_{\varepsilon \rightarrow 0} f(x(\varepsilon)) = \lim_{\varepsilon \rightarrow 0} M(x(\varepsilon), \varepsilon) = f(x^*)$$

בנוסף אם $x \rightarrow \partial\Omega_\varepsilon$, הרי שלפי דרישה (ט) של הגדרה 5.1 (הגדרת מחלקת B) מתקיים כי

$\psi(\varepsilon, g_i(x)) \rightarrow \infty$ ולכן גם $M(x(\varepsilon), \varepsilon) \rightarrow \infty$. כלומר, המינימום שלה לא מתקבל על השפה של Ω_ε

אלה בנקודה פנימית, אבל אז חייב להתקיים:

$$M'_x(x(\varepsilon), \varepsilon) = f(x(\varepsilon)) + \sum_{i=1}^m \psi'_g(\varepsilon, g_i(x(\varepsilon))) g'_{x_i}(x(\varepsilon)) = f(x(\varepsilon)) + \sum_{i=1}^m y_i(\varepsilon) g'_{x_i}(x(\varepsilon)) = 0$$

כעת נעריך את $\|\Delta x\| = \|x(\varepsilon) - x^*\|$, $\|\Delta y\| = \|y(\varepsilon) - y^*\|$

לכן, $f(x) \in C^2$

$$f'(x(\varepsilon)) = f'(x^*) + f''(x^*)\Delta x + h^f(\Delta x)$$

כך ש-

$$h^f(0) = 0; \quad \|h^f(\Delta x)\| \leq \alpha(\Delta x)\Delta x; \quad \lim_{\Delta x \rightarrow 0} \alpha(\Delta x) = 0$$

באותו אופן לכל $i = \overline{1, m}$, מתקיים:

$$g_j'(x(\varepsilon)) = g_j'(x^*) + g_j''(x^*)\Delta x + h_j^g(\Delta x)$$

כך ש-

$$h_i^g(0) = 0; \quad \|h_i^g(\Delta x)\| \leq \alpha_i(\Delta x)\Delta x; \quad \lim_{\Delta x \rightarrow 0} \alpha_i(\Delta x) = 0$$

$y(x(\varepsilon))$ היא פונקציה גזירה פעמיים ברציפות, ולכן לכל $i = \overline{1, m}$:

$$y_i(\varepsilon) = y_i(x^*, \varepsilon) + y_i'(x^*, \varepsilon)\Delta x + h_i^y(\Delta x);$$

כך ש-

$$h_i^y(0) = 0; \quad \|h_i^y(\Delta x)\| \leq \beta_i(\Delta x)\Delta x; \quad \lim_{\Delta x \rightarrow 0} \beta_i(\Delta x) = 0$$

$$\psi(\varepsilon, g_i(x))$$
 היא פונקציה יורדת וקמורה, לכן:

$$y_i(x^*, \varepsilon) = \begin{cases} \psi'_g(\varepsilon, g_i(x^*)) = -1, & g_i(x^*) = 0 \Leftarrow i \in I \\ O(\varepsilon), & g_i(x^*) \geq \sigma \Leftarrow i \notin I \end{cases}$$

לפי (ו) של מחלקת B .

בנוסף

$$y'_{ix}(x^*, \varepsilon) = \psi''_{qq}(\varepsilon, g_i(x^*))g'_i(x^*)$$

(לפי (ח) של מחלקה B קיים $k \in \mathbb{N}$ ו- $l \in \mathbb{R}$ כך ש- $0 \neq l$ ו- $(\psi''_{gg}(\varepsilon, g_i(x^*))) = \frac{l}{\varepsilon^k}$)

נציב בביטוי $M'_x(x(\varepsilon), \varepsilon) = f(x(\varepsilon)) + \sum_{i=1}^m y_i(\varepsilon) g'_i(x(\varepsilon))$ את $f(x), g_i(x), y_i(\varepsilon)$ כך $i \in I$

ניקח בחשבון ש- $y_i^* = \Delta y_i(\varepsilon) + y_i$, ונקבל:

$$\begin{aligned}
0 &= f'(x^*) + f''(x^*)\Delta x + h^f(\Delta x) + \sum_{i=1}^r (\Delta y_i + y_i^*)(g'_i(x^*) + g''_i(x^*)\Delta x + h_i^g(\Delta x)) + \sum_{i=r+1}^m y_i(\varepsilon)g'_i(x(\varepsilon)) = \\
&= [f'(x^*) + \sum_{i=1}^r y_i^* g'_i(x^*)] + [f''(x^*) + \sum_{i=1}^r y_i^* g''_i(x^*)]\Delta x + \sum_{i=1}^r \Delta y_i g'_i(x^*) + [h^f(\Delta x) + \sum_{i=1}^r \Delta y_i (g''_i(x^*)\Delta x + \\
&+ h_i^g(\Delta x)) + \sum_{i=r+1}^m y_i(\varepsilon)g'_i(x(\varepsilon))]
\end{aligned}$$

אם נסמן :

$$h^{fg}(\Delta x) = [h^f(\Delta x) + \sum_{i=1}^r \Delta y_i (g''_i(x^*)\Delta x + h_i^g(\Delta x)) + \sum_{i=r+1}^m y_i(\varepsilon)g'_i(x(\varepsilon))]$$

הרי שעבור $\|\Delta x\| \rightarrow 0$ נקבל $\alpha(\Delta x) \rightarrow 0$ ו- $\|h^{fg}(\Delta x)\| \leq \alpha(\Delta x)\|\Delta x\|$; $h^{fg}(0) = 0$

מתקיים תנאי (3) של KKT לכן נרשום את הביטוי בצורה אחרת : $L''_{xx}\Delta x + g^{(r)'}\Delta y + h^{fg}(\Delta x) = 0$

ונקבל,

$$y_i(\varepsilon) = y_i(x^*, \varepsilon) + y'_i(x^*, \varepsilon)\Delta x + h_i^y(\Delta x) = -1 + \psi''(\varepsilon, g_i(x^*))g'_i(x^*)\Delta x + h_i^y(\Delta x)$$

נסמן :

$$\bar{e} = (1, \dots, 1) \in R^r, \quad y^* = (y_1^*, \dots, y_r^*), \quad h^y(\Delta x) = (h_1^y(\Delta x), \dots, h_r^y(\Delta x)), \quad v^* = -\bar{e} - y^*$$

אזי :

$$\Delta y = y - y^* = -\bar{e} + \psi''_{gg}(\varepsilon, g_i(x^*))g^{(r)'}(x^*)\Delta x + h^y(\frac{1}{2}\Delta x) - y^*$$

$$\Delta y = v^* + \psi''_{gg}(\varepsilon, g_i(x^*))g^{(r)'}(x^*)\Delta x$$

נקח בחשבון שמתקיים : $\beta(\Delta x) \rightarrow 0$ כאשר $\|\Delta x\| \rightarrow 0$, $\|h^y(\Delta x)\| \leq \beta(\Delta x)\|\Delta x\|$, $h^y(0) = 0$, אזי

$$D \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} L''_{xx} & g^{(r)'} \\ -\varepsilon \psi''_{gg}(\varepsilon, g(x^*))g^{(r)'} & \varepsilon I^r \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} 0 \\ \varepsilon v \end{bmatrix} + \begin{bmatrix} h^{fg}(\Delta x) \\ \varepsilon h^y(\Delta x) \end{bmatrix}$$

בהוכחת משפט 5.4 הראנו שמטריצה D אינה סינגולרית.

ועבור $\varepsilon > 0$ קיים קבוע c שאינו תלוי ב- ε כך ש- $\|D^{-1}\| \leq c$.

נסמן $p = (0, \varepsilon v^*)$, $q(\Delta x) = (-h^{fg}(\Delta x), \varepsilon h^y(\Delta x))$

אזי מ- $\|\Delta x\| \rightarrow 0$ נובע ש- $\|q(\Delta x)\| \leq k(\Delta x)\|\Delta x\|$, $k(\Delta x) \rightarrow 0$

ואם נשאיף את $x(\varepsilon) \rightarrow x^*, y(\varepsilon) \rightarrow y^*$ ונסמן : $\Delta z = (\Delta x, \Delta y)$, נקבל $\Delta z = D^{-1}p + D^{-1}q(\Delta x)$, כלומר מתקיימת הערכה (23).
וקיים קבוע c הבלתי תלוי ב- ε כך ש- $\|\Delta z\| \leq c\varepsilon$,

ג. נראה ש- $M(x(\varepsilon), \varepsilon)$ קמורה במובן חזק סביב $x(\varepsilon)$:

$$M(x(\varepsilon), \varepsilon) = f(x(\varepsilon)) + \sum_{i=1}^m \psi(\varepsilon, g_i(x(\varepsilon)))$$

$$M'(x(\varepsilon), \varepsilon) = f'(x(\varepsilon)) + \sum_{i=1}^m \psi'_g(\varepsilon, g_i(x(\varepsilon)))g'_i(x(\varepsilon))$$

$$M''_{xx}(x(\varepsilon), \varepsilon) = f''(x(\varepsilon)) + \sum_{i=1}^m \psi''_{gg}(\varepsilon, g_i(x(\varepsilon)))g'_i(x(\varepsilon))g'_i(x(\varepsilon)) + \sum_{i=1}^m \psi''_g(\varepsilon, g_i(x(\varepsilon)))g'_i(x(\varepsilon))$$

אם $y_i(\varepsilon) \rightarrow y^*$ הרי ש- $g_i(x(\varepsilon)) \rightarrow g_i(x^*)$, $\forall i = \overline{1, r}$,

וקיים $\sigma > 0$ כך ש- $g_i(x(\varepsilon)) \rightarrow g_i(x^*) \geq \sigma > 0$ לכל $i = \overline{r+1, m}$.

וכמו כן מתקיים: $\psi''_{gg}(\varepsilon, g_i(x(\varepsilon))) = O(\varepsilon)$ ויותר מכך $y_i(\varepsilon) = O(\varepsilon)$ לכל $i = \overline{r+1, m}$.

לכן

$$f''(x(\varepsilon)) + \sum_{i=1}^m y_i(\varepsilon) g''_i(x(\varepsilon)) \rightarrow L''_{xx}(x^*, y^*)$$

כלומר,

$$\sum_{i=1}^r \psi''_{gg}(\varepsilon, g_i(x(\varepsilon))) g'_i(x(\varepsilon)) g'^T_i(x(\varepsilon)) \rightarrow \psi''_{gg} g^{(r)T} g^{(r)'}$$

$$\left\| \sum_{i=r+1}^m \psi''_{gg}(\varepsilon, g_i(x(\varepsilon))) g'_i(x(\varepsilon)) g'^T_i(x(\varepsilon)) \right\| = O(\varepsilon) \quad \text{וגם} \quad \left\| \sum_{i=r+1}^m y_i(\varepsilon) g''_i(x(\varepsilon)) \right\| = O(\varepsilon)$$

לכן עבור $\varepsilon > 0$ מספיק קטן מתקיים

$$M''_{xx}(x(\varepsilon), \varepsilon) \cong L''_{xx}(x^*, y^*) + \psi''_{gg}(\varepsilon, g_i(x^*)) g^{(r)T}(x^*) g^{(r)'}(x^*)$$

$g_i(x) \in C^2$, לכן עבור $\varepsilon > 0$ מספיק קטן, $M(x(\varepsilon), \varepsilon)$ פונקציה קמורה במובן חזק בסביבת $x(\varepsilon)$.

מש"ל.

נחזור לשיטת האיטרציות של [12] Polyak :

אם נציב בפונקציה $F(\varepsilon, x, y) = f(x) + s_\psi \sum_{i=1}^m y_i \psi(\varepsilon, g_i(x))$ את $y^0 = (1, \dots, 1) \in R^m$ אז לפי משפט 6.4

קיים $\varepsilon_0 > 0$ כך שלכל $\varepsilon > \varepsilon_0$ קיים זוג (x_1, y_1) כך ש-

$$x^1 = x(\varepsilon) = \arg \min \{ F_0(x, \varepsilon) / x \in \Omega_\varepsilon \}$$

$$y^1 = y(x(\varepsilon)) = [\text{diag} \psi'_g(\varepsilon, g_i(x))]_{i=1}^m$$

$$F_0(x, \varepsilon) = f(x) + s_\psi \sum_{i=1}^m \psi(\varepsilon, g_i(x))$$

$$\|x_1 - x^*\| \leq c\varepsilon, \|y_1 - y^*\| \leq c\varepsilon \quad \text{ובנוסף}$$

באופן כללי נציג את שיטת האיטרציות:

$$(A^*) \quad \begin{cases} y^0 = (1, 1, \dots, 1) \in \mathbb{R}^m \\ x^{k+1} = \arg \min (F(\varepsilon, x, y^k)) \\ y^{k+1} = [\text{diag} \psi'_g(\varepsilon, g_i(x^{k+1}))]_{i=1}^m y^k \end{cases}$$

נניח $\forall i = \overline{1, m}$, $g(x), f(x) \in C^2$, ותנאים (3)-(5) מתקיימים,

אז קיים $\varepsilon > 0$ כך ששיטת איטרציות יוצרת סידרה $\{(x^s, y^s)\}_{s=0}^\infty$ כך

$\max\{\|x^s - x^*\|, \|y^s - y^*\|\} \leq c\varepsilon$ ו- $c > 0$ אינו תלוי ב- ε . כלומר קצב התכנסות לאופטימום הוא לינארי.

7. דוגמאות והרצות

בפרק זה נדגים פתרון עבור דוגמא 2.1 בעזרת שימוש המרות מחסום משופרות בשיטה האיטרטיבית (A^*) . ובנוסף נציג עוד כמה דוגמאות אופטימיזציה מאולצות וקמורות ונציג את הפתרון שקיבלנו עבור שימוש בשיטה הזאת.

הערה 7.1:

אלגוריתם ממומש ע"י פונקציות מתל"ב (ראה נספח 2). בכל איטרציה צריך לפתור בעיה אופטימיזציה לא מאולצת על מנת למצוא $x^{k+1} = \arg \min(F(\varepsilon, x, y^k))$. השתמשנו בפונקציה מבנית של מתל"ב fminsearch. פונקציה עובדת לפי שיטה downhill simplex method, שהיא אחת השיטות המקובלות לפתרון בעיות אופטימיזציה לא מאולצת ולא לינארית.

דוגמא 7.2 (הדגמת פתרון עבור דוגמא 2.1 – עיכוב כללי מינימאלי של תעבורה ברשת):

בעיה נלקחה ממדריך משתמש של GINO ראה ב-[8].

פתרון מדויק לבעיה מתקבל בנקודה $(2.5, 2.5, 0, 2.5, 2.5)$, כאשר ערך של הפונקציה בנקודה זאת 40.3030.

תוצאות שהתקבלו ע"י שימוש בשיטת המ"מ:

נבחר: $\varepsilon = 0.1$

פונקציה מחסום	נקודת מינימום	ערך של פונקציה	מספר איטרציות
$\psi_1(\varepsilon, t) = \varepsilon \log(\frac{t}{\varepsilon} + 1)$	$x_{12} = 2.5357$ $x_{13} = 2.4599$ $x_{32} = -0.0705$ $x_{24} = 2.4567$ $x_{34} = 2.5271$	40.1300	35
$\psi_2(\varepsilon, t) = \varepsilon[(\frac{t}{\varepsilon} + 1)^{-1} - 1]$	$x_{12} = 2.5311$ $x_{13} = 2.4676$ $x_{32} = -0.0611$ $x_{24} = 2.4671$ $x_{34} = 2.5290$	40.2234	25

הערה: חישובים התבצעו בעזרת תכנית מתל"ב (ראה נספח מספר 3).

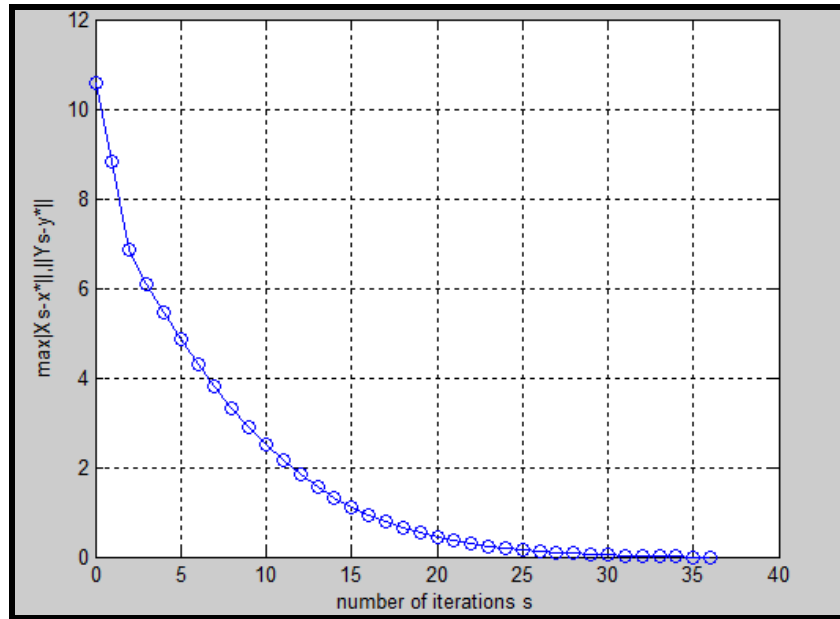
השוואת תוצאות שקיבלנו מול פתרון מדויק:

נסמן:

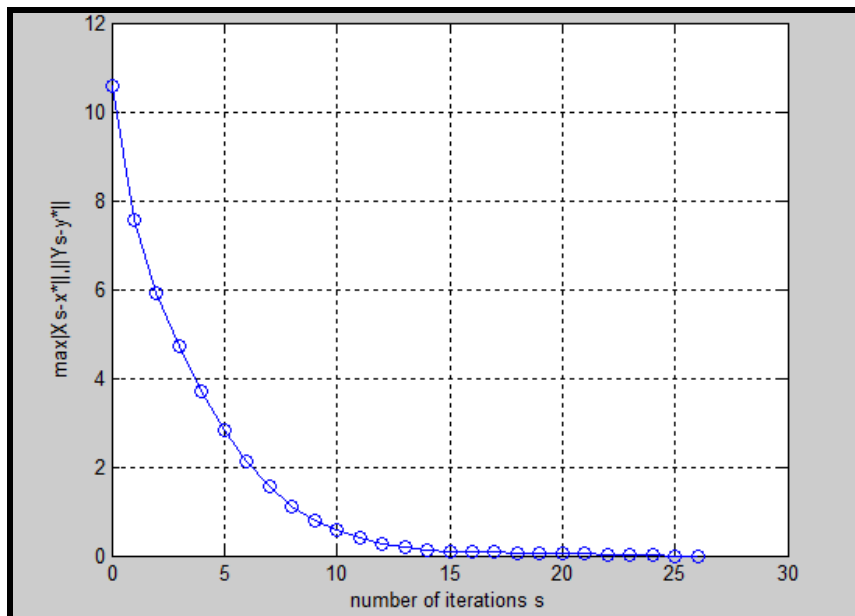
$x^*, f(x^*)$ ערך של הפונקציה בנקודת מינימום ונקודת מינימום עבור פתרון אמיתי

$\hat{x}, f(\hat{x})$ ערך של הפונקציה בנקודת מינימום ונקודת מינימום שקיבלו עבור שימוש בשיטת המ"מ.

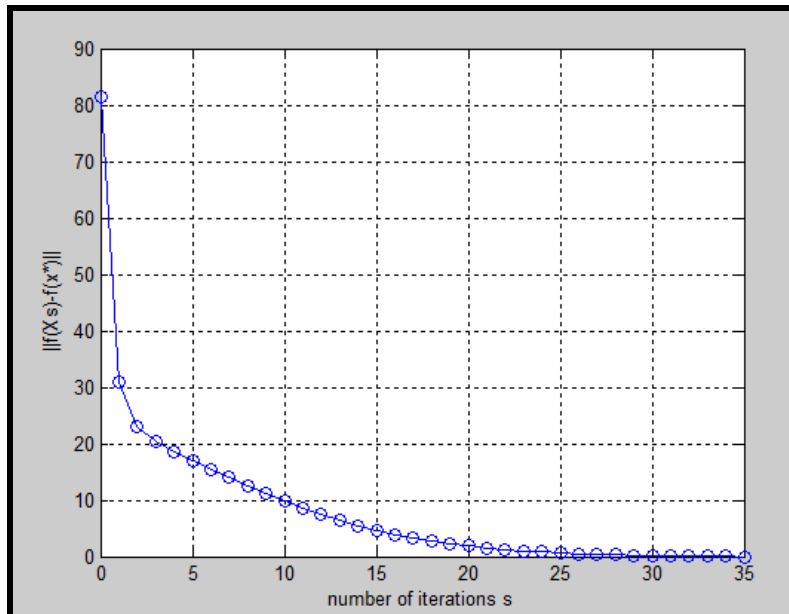
פונקציה מחסום	$ f(\hat{x}) - f(x^*) $	$\ \hat{x} - x^*\ $
$\psi_1(\varepsilon, t) = \varepsilon \log(\frac{t}{\varepsilon} + 1)$	0.0934	0.0705
$\psi_2(\varepsilon, t) = \varepsilon[(\frac{t}{\varepsilon} + 1)^{-1} - 1]$	0.0796	0.0611



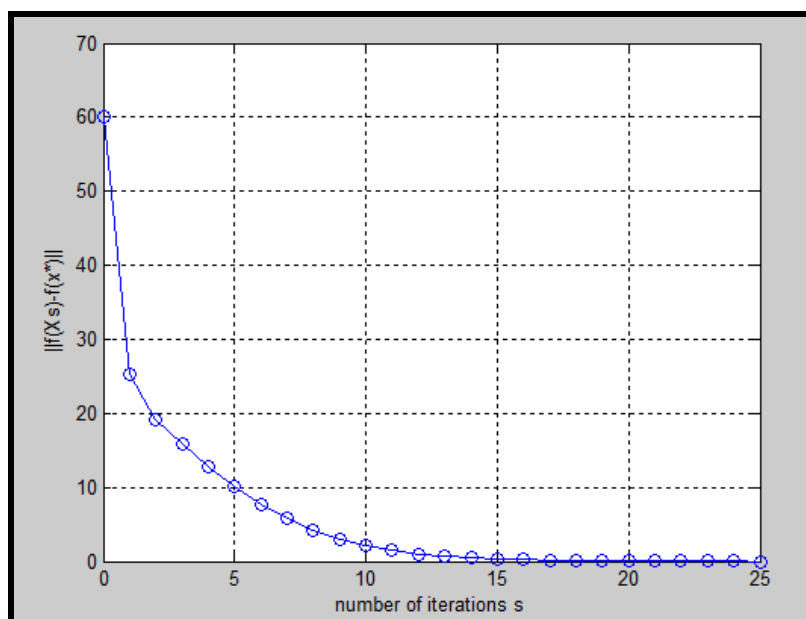
איור 2. התכנסות בנורמה לפתרון של דוגמא 7.2 עבור המרה : $\psi_1(\varepsilon, t) = \varepsilon \log\left(\frac{t}{\varepsilon} + 1\right)$



איור 3. התכנסות בנורמה לפתרון של דוגמא 7.2 עבור המרה : $\psi_2(\varepsilon, t) = \varepsilon \left[\left(\frac{t}{\varepsilon} + 1 \right)^{-1} - 1 \right]$



איור 4. התכנסות לערך מינימאלי של הפונקציה בדוגמא 7.2 עבור המרה : $\psi_1(\varepsilon, t) = \varepsilon \log\left(\frac{t}{\varepsilon} + 1\right)$



איור 5. התכנסות לערך מינימאלי של הפונקציה בדוגמא 7.2 עבור המרה : $\psi_2(\varepsilon, t) = \varepsilon \left[\left(\frac{t}{\varepsilon} + 1 \right)^{-1} - 1 \right]$

דוגמא 7.3:

תהי נתונה בעיית אופטימיזציה מאולצת קמורה (לא ליניארית)

$$\min(x_1^2 + e^{x_2})$$

$$64 - x_1^2 - x_2^2 \geq 0$$

$$9 - x_1 - x_2 \geq 0$$

$$25 - (x_1 - 10)^2 - x_2^2 \geq 0$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$

$$(x_1, x_2) \in R^2$$

פתרון מדויק לבעיה מתקבל בנקודה $(5, 0)$ כאשר ערך של הפונקציה בנקודה זאת 26.

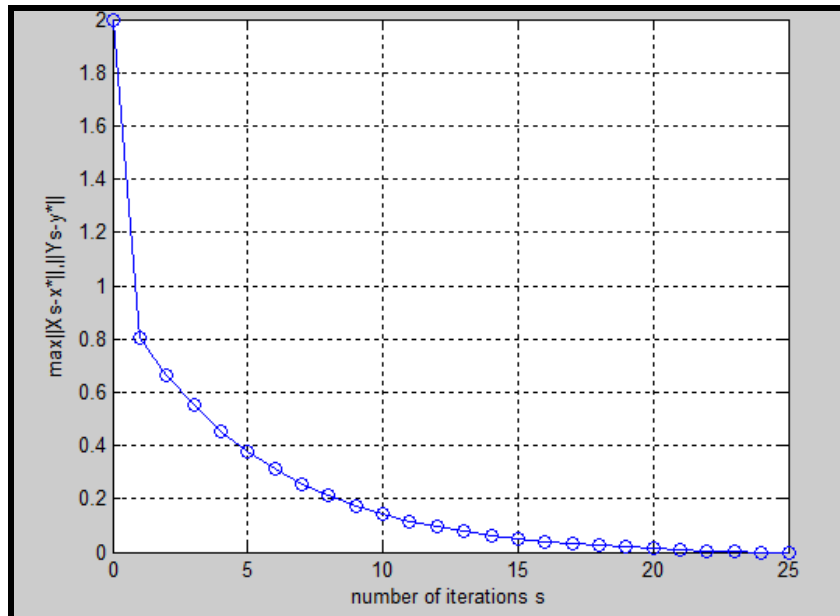
תוצאות שהתקבלו ע"י שימוש בשיטת המ"מ:

נבחר: $\varepsilon = 0.05$

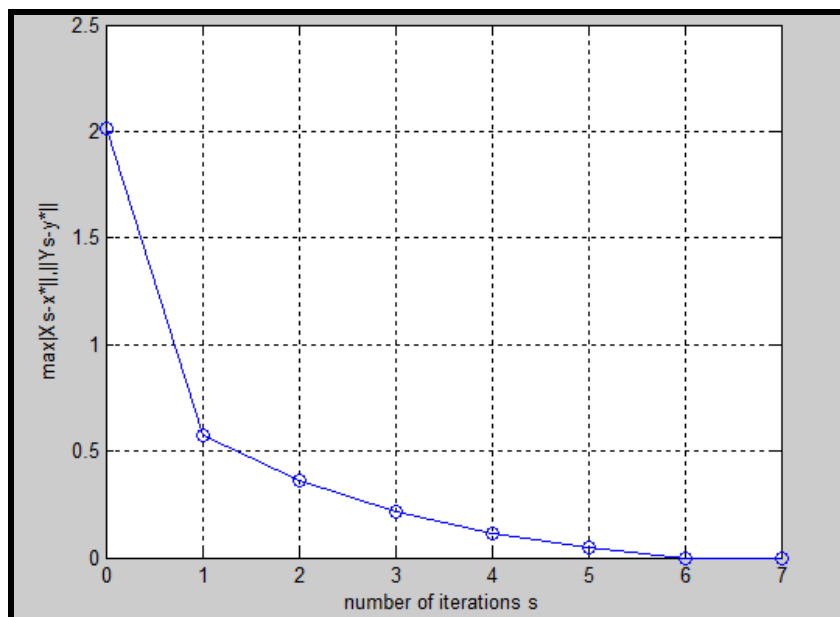
פונקציה מחסום	נקודת מינימום	ערך של פונקציה	מספר איטרציות
$\psi_1(\varepsilon, t) = \varepsilon \log(\frac{t}{\varepsilon} + 1)$	$x_1 = 5.0003$ $x_2 = 0.0149$	26.0175	24
$\psi_2(\varepsilon, t) = \varepsilon[(\frac{t}{\varepsilon} + 1)^{-1} - 1]$	$x_1 = 5.0010$ $x_2 = -0.0057$	26.0044	6

השוואת תוצאות שקיבלנו מול פתרון אמיתי:

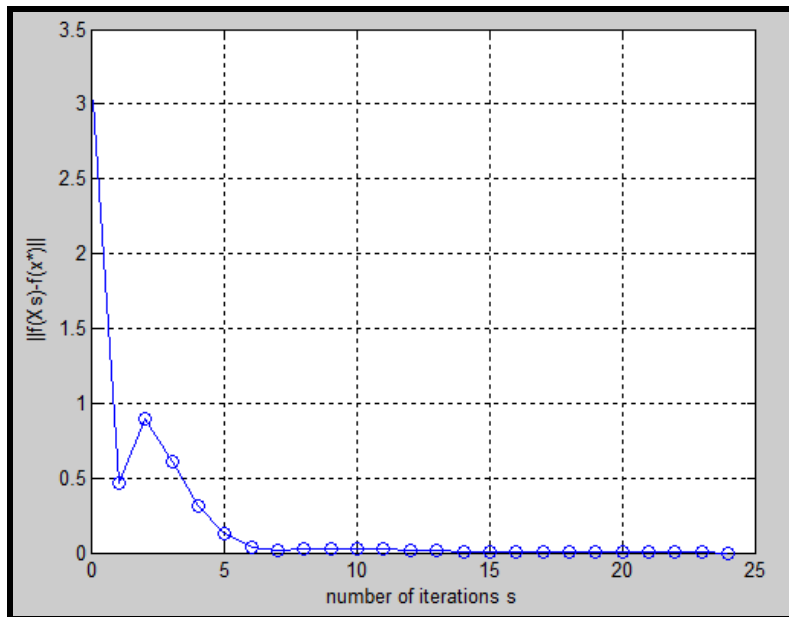
פונקציה מחסום	$ f(\hat{x}) - f(x^*) $	$\ \hat{x} - x^*\ $
$\psi_1(\varepsilon, t) = \varepsilon \log(\frac{t}{\varepsilon} + 1)$	0.0175	0.0059
$\psi_2(\varepsilon, t) = \varepsilon[(\frac{t}{\varepsilon} + 1)^{-1} - 1]$	0.0044	0.0611



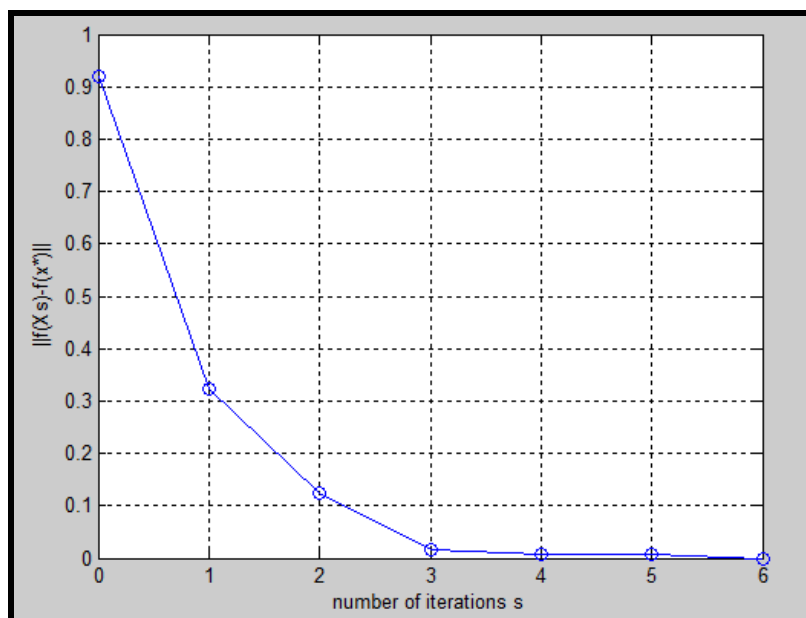
איור 6. התכנסות לפתרון בנורמה של דוגמא 7.3 עבור המרה : $\psi_1(\varepsilon, t) = \varepsilon \log\left(\frac{t}{\varepsilon} + 1\right)$



איור 7. התכנסות לפתרון בנורמה של דוגמא 7.3 עבור המרה : $\psi_2(\varepsilon, t) = \varepsilon \left[\left(\frac{t}{\varepsilon} + 1 \right)^{-1} - 1 \right]$



איור 6. התכנסות לערך מינימאלי של הפונקציה בדוגמא 7.3 עבור המרה : $\psi_1(\varepsilon, t) = \varepsilon \log\left(\frac{t}{\varepsilon} + 1\right)$



איור 7. התכנסות לערך מינימאלי של הפונקציה בדוגמא 7.3 עבור המרה : $\psi_2(\varepsilon, t) = \varepsilon \left[\left(\frac{t}{\varepsilon} + 1 \right)^{-1} - 1 \right]$

דוגמא 7.4:

תהי נתונה בעיית אופטימיזציה מאולצת קמורה (לא ליניארית)

$$\min(x_1^2 + x_2^2)$$

$$x_1^2 - 2 \geq 0$$

$$x_2^2 - 1 \geq 0$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$

$$(x_1, x_2) \in R^2$$

פתרון מדויק לבעיה מתקבל בנקודה $(\sqrt{2}, 1)$, כאשר ערך של הפונקציה בנקודה זאת 3.

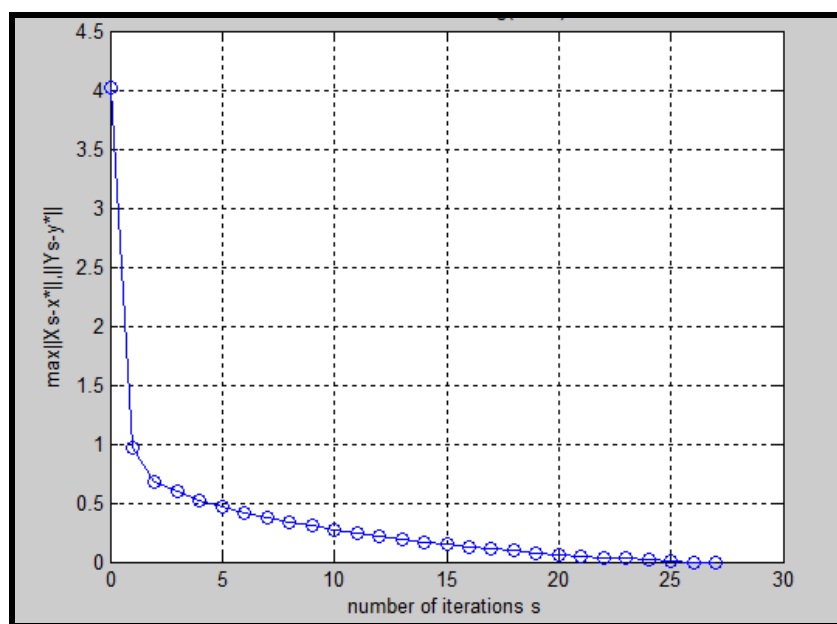
תוצאות שהתקבלו ע"י שימוש בשיטת המ"מ:

נבחר: $\varepsilon = 0.1$

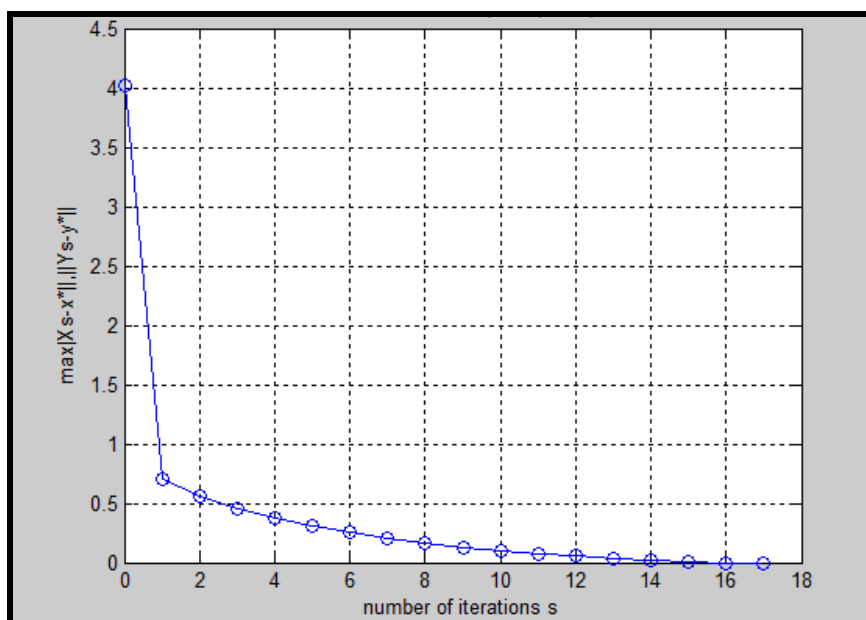
פונקציה מחסום	נקודת מינימום	ערך של פונקציה	מספר איטרציות
$\psi_1(\varepsilon, t) = \varepsilon \log\left(\frac{t}{\varepsilon} + 1\right)$	$x_1 = 0.9790$ $x_2 = 1.4091$	2.9442	16
$\psi_2(\varepsilon, t) = \varepsilon \left[\left(\frac{t}{\varepsilon} + 1\right)^{-1} - 1\right]$	$x_1 = 0.9891$ $x_2 = 1.4121$	2.9723	26

השוואת תוצאות שקיבלנו מול פתרון אמיתי:

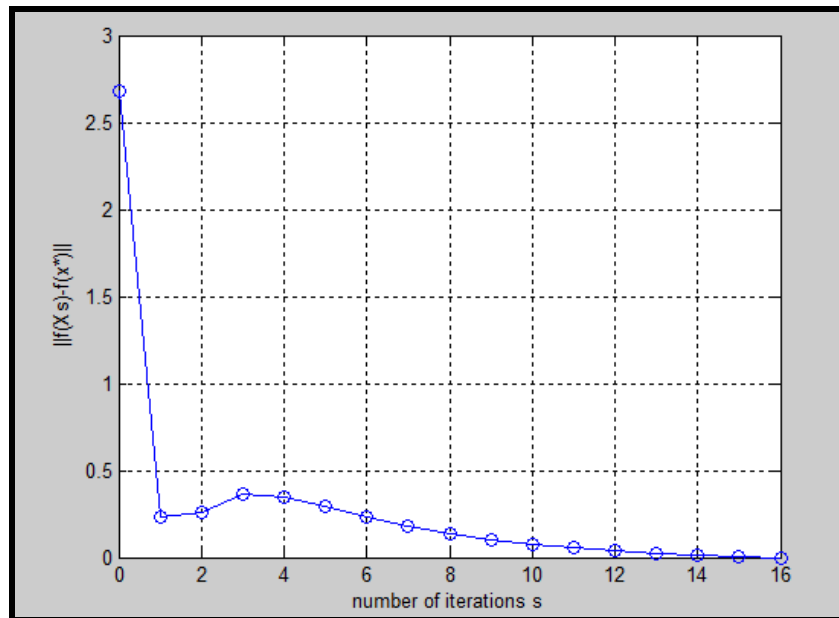
פונקציה מחסום	$ f(\hat{x}) - f(x^*) $	$\ \hat{x} - x^*\ $
$\psi_1(\varepsilon, t) = \varepsilon \log\left(\frac{t}{\varepsilon} + 1\right)$	0.0558	0.0210
$\psi_2(\varepsilon, t) = \varepsilon \left[\left(\frac{t}{\varepsilon} + 1\right)^{-1} - 1\right]$	0.0277	0.0109



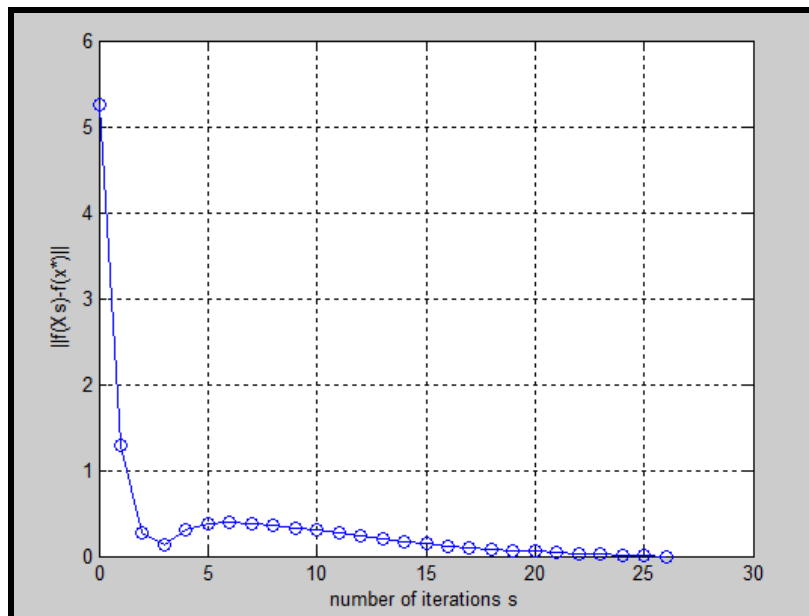
איור 8. התכנסות לפתרון בנורמה של דוגמא 7.4 עבור המרה : $\psi_1(\varepsilon, t) = \varepsilon \log\left(\frac{t}{\varepsilon} + 1\right)$



איור 9. התכנסות לפתרון בנורמה של דוגמא 7.4 עבור המרה : עבור : $\psi_2(\varepsilon, t) = \varepsilon \left[\left(\frac{t}{\varepsilon} + 1 \right)^{-1} - 1 \right]$



איור 10. התכנסות לערך מינימאלי של הפונקציה בדוגמא 7.4 עבור המרה : $\psi_1(\varepsilon, t) = \varepsilon \log\left(\frac{t}{\varepsilon} + 1\right)$



איור 11. התכנסות לערך מינימאלי של הפונקציה בדוגמא 7.4 עבור המרה : $\psi_2(\varepsilon, t) = \varepsilon \left[\left(\frac{t}{\varepsilon} + 1 \right)^{-1} - 1 \right]$

דוגמא 7.5:

תהי נתונה בעיית אופטימיזציה מאולצת קמורה (לא ליניארית)

$$\min(x_1 + x_2)^2 - 10*(x_1 + x_2)$$

$$8 - x_1^2 - x_2^2 \geq 0$$

$$8 - 3*x_1 - x_2 \geq 0$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$

$$(x_1, x_2) \in R^2$$

פתרון מדויק לבעיה מתקבל בנקודה $(2, 2)$ כאשר ערך של הפונקציה בנקודה זאת 24.-.

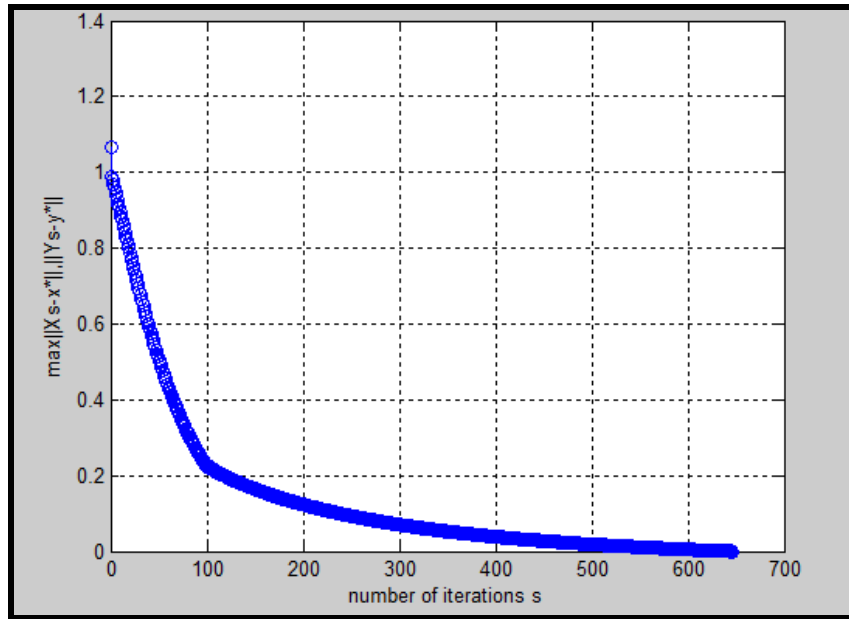
תוצאות שהתקבלו ע"י שימוש בשיטת המ"מ:

נבחר: $\varepsilon = 0.01$

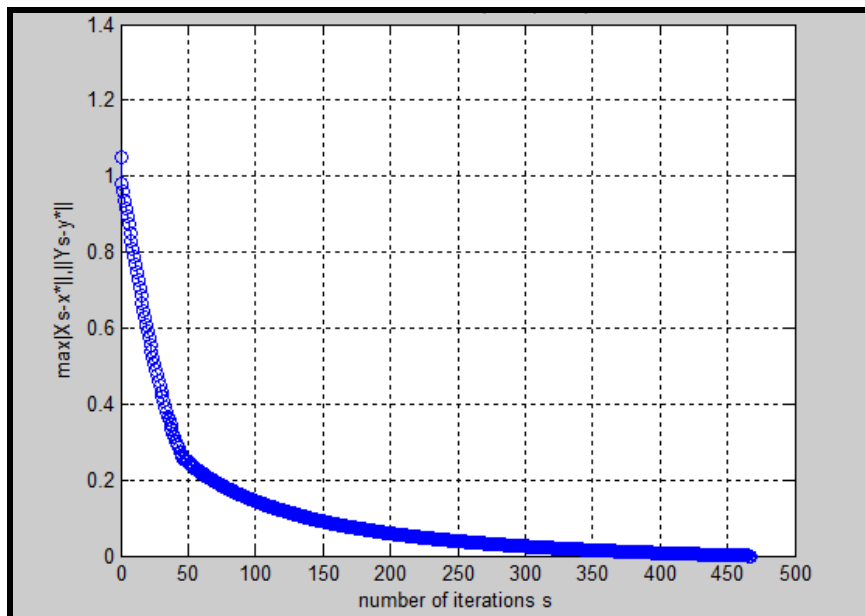
מספר איטרציות	ערך של פונקציה	נקודת מינימום	פונקציה מחסום
645	-24.0000	$x_1 = 2.0683$ $x_2 = 1.9317$	$\psi_1(\varepsilon, t) = \varepsilon \log(\frac{t}{\varepsilon} + 1)$
466	-24.0000	$x_1 = 2.0483$ $x_2 = 1.9517$	$\psi_2(\varepsilon, t) = \varepsilon[(\frac{t}{\varepsilon} + 1)^{-1} - 1]$

השוואת תוצאות שקיבלנו מול פתרון אמיתי:

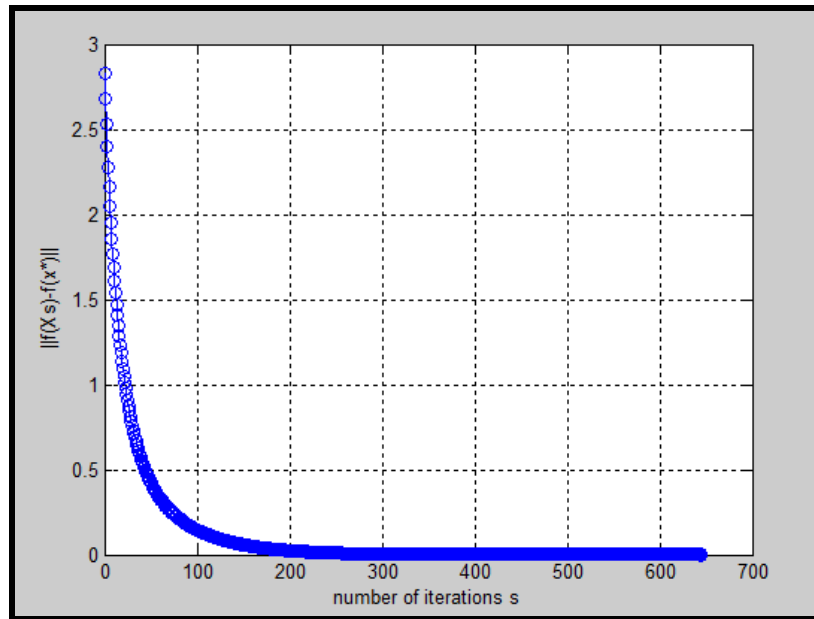
$\ \hat{x} - x^*\ $	$ f(\hat{x}) - f(x^*) $	פונקציה מחסום
0.0683	0	$\psi_1(\varepsilon, t) = \varepsilon \log(\frac{t}{\varepsilon} + 1)$
0.0483	0	$\psi_2(\varepsilon, t) = \varepsilon[(\frac{t}{\varepsilon} + 1)^{-1} - 1]$



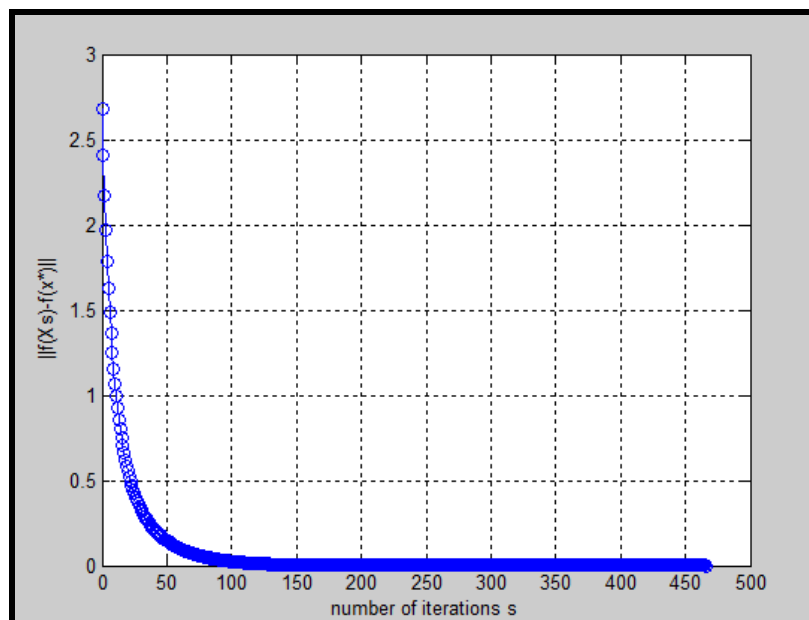
איור 12. התכנסות לפתרון בנורמה של דוגמא 7.5 עבור המרה : $\psi_1(\varepsilon, t) = \varepsilon \log\left(\frac{t}{\varepsilon} + 1\right)$



איור 13. התכנסות לפתרון בנורמה של דוגמא 7.5 עבור המרה : $\psi_2(\varepsilon, t) = \varepsilon \left[\left(\frac{t}{\varepsilon} + 1\right)^{-1} - 1\right]$



איור 12. התכנסות לערך מינימאלי של הפונקציה בדוגמא 7.5 עבור המרה : $\psi_1(\varepsilon, t) = \varepsilon \log\left(\frac{t}{\varepsilon} + 1\right)$



איור 13. התכנסות לערך מינימאלי של הפונקציה בדוגמא 7.5 עבור המרה : $\psi_2(\varepsilon, t) = \varepsilon \left[\left(\frac{t}{\varepsilon} + 1 \right)^{-1} - 1 \right]$

- [1] K. Arrow, L. Hurwicz, H. Uzawa,
Studies on Linear and Nonlinear Programming (Stanford University Press, Stanford, CA, 1958.
- [2] D. Bertsekas,
Constrained Optimization and Lagrange Multiplier Method. Academic Press, New York, 1982.
- [3] C.W. Carroll,
The created response surface technique for optimizing nonlinear restrained systems. 1969.
- [4] A. V. Fiacco and G. P. McCormick.
Nonlinear Programming: Sequential Unconstrained Minimization Techniques. J. Wiley and Sons, New York.(1968).
Reprinted as Classics in Applied Mathematics 4, SIAM, 1990.
- [5] K. R. Frisch.
The logarithmic potential method of convex programming.
Memorandum of May13, University Institute of Economics, Oslo, Norway. 1955.
- [6] K. Jittorntrum and M. Osborne.
A modified barrier function method with improved rate of
convergence for degenerate problems. Journal of the Australian Mathematical Society (SeriesB), 1980.
- [7] N. Karmarkar.
A new polynomial-time algorithm for linear programming. Combinatorica, 1984.
- [8] Liebman, J.; Lasdon, L.; Schrage, L.; and Waren,
A. Modeling and Optimization with GINO, California: The Scientific Pres, 1986.
- [9] F. A. Lootsma.
Hessian matrices of penalty functions for solving constrained optimization problems. Philips Research Reports. 1969.
- [10] W. Murray.
Analytical expressions for eigenvalues and eigenvectors of the Hessian matrices of barrier and penalty functions.
Journal of Optimization Theory and Applications, 1971.
- [11] S. G. Nash and A. Sofer.
A barrier method for large-scale constrained optimization. ORSA. Journal on Computing, 1993.
- [12] R. A. Polyak ,
Modified barrier functions. Mathematical sciences department,
IBM research division, T.J. Watson Research Center, 1989.

נספח 1 – הוכחת קמורות של בעיית אופטימיזציה בדוגמא 2.1

נראה כי בעיית אופטימיזציה (i) בדוגמא 2.1 היא בעיית אופטימיזציה קמורה.

פונקציות אילוצים הן פונקציות לינאריות, לכן קמורות.

$$\Omega = \{x \in \mathbb{R}^5 / g_i(x) \geq 0, \quad i = \overline{1,10}, \quad h_j(x) = 0 \quad j = \overline{1,3}\}$$

נראה שפונקציה $f(x)$ היא פונקציה קמורה:

$$f(x) = 5 * x_{12} + \frac{0.1 * x_{12}^2}{11 - x_{12}} + \frac{x_{13}^2}{31 - x_{13}} + x_{32} + \frac{x_{32}^2}{11 - x_{32}} + \frac{x_{24}^2}{31 - x_{24}} + 5 * x_{34} + \frac{x_{34}^2}{11 - x_{34}}$$

נחשב מטריצת הסיאן שלה

$$H(f) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_{12}^2} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{\partial^2 f}{\partial x_{34}^2} \end{pmatrix}$$

מקבלים:

$$\begin{aligned} \frac{\partial^2 f}{\partial x_{12}^2} &= \frac{0.2}{11 - x_{12}} + \frac{0.4 * x_{12}}{(11 - x_{12})^2} + \frac{0.2 * x_{12}^2}{(11 - x_{12})^3} \\ \frac{\partial^2 f}{\partial x_{13}^2} &= \frac{2}{31 - x_{13}} + \frac{4 * x_{13}}{(31 - x_{13})^2} + \frac{2 * x_{13}^2}{(31 - x_{13})^3} \\ \frac{\partial^2 f}{\partial x_{32}^2} &= \frac{2}{11 - x_{32}} + \frac{4 * x_{32}}{(11 - x_{32})^2} + \frac{2 * x_{32}^2}{(11 - x_{32})^3} \\ \frac{\partial^2 f}{\partial x_{24}^2} &= \frac{2}{31 - x_{24}} + \frac{4 * x_{24}}{(31 - x_{24})^2} + \frac{2 * x_{24}^2}{(31 - x_{24})^3} \\ \frac{\partial^2 f}{\partial x_{34}^2} &= \frac{0.2}{11 - x_{34}} + \frac{0.4 * x_{34}}{(11 - x_{34})^2} + \frac{0.2 * x_{34}^2}{(11 - x_{34})^3} \end{aligned}$$

יהיה $a = (x_{12}, x_{13}, x_{24}, x_{32}, x_{34}) \in \Omega$, כלומר $0 \leq x_{12}, x_{32}, x_{34} \leq 10, \quad 0 \leq x_{13}, x_{24} \leq 30$

ונראה ש- $H(f)$ מטריצה מוגדרת אי-שלילית, כלומר שמתקיים $\langle H(f)a, a \rangle \geq 0$

לכן:

$$\begin{aligned}
a^T H(f) a = & a_1^2 \left(\frac{0.2}{11-x_{12}} + \frac{0.4 * x_{12}}{(11-x_{12})^2} + \frac{0.2 * x_{12}^2}{(11-x_{12})^3} \right) + \\
& + a_2^2 \frac{2}{31-x_{13}} + \frac{4 * x_{13}}{(31-x_{13})^2} + \frac{2 * x_{13}^2}{(31-x_{13})^3} + \\
& + a_3^2 \frac{2}{11-x_{32}} + \frac{4 * x_{32}}{(11-x_{32})^2} + \frac{2 * x_{32}^2}{(11-x_{32})^3} + \\
& + a_4^2 \frac{2}{31-x_{24}} + \frac{4 * x_{24}}{(31-x_{24})^2} + \frac{2 * x_{24}^2}{(31-x_{24})^3} + \\
& + a_5^2 \frac{0.2}{11-x_{34}} + \frac{0.4 * x_{34}}{(11-x_{34})^2} + \frac{0.2 * x_{34}^2}{(11-x_{34})^3} \geq 0
\end{aligned}$$

לכל $a \in \Omega$.

לכן $f(x)$ היא פונקציה קמורה ובעיה (2) היא בעיית אופטימיזציה קמורה.

נבדוק תנאי סלייטר:

נבחר נקודה $x^0 = (x_{12}^0, x_{13}^0, x_{24}^0, x_{32}^0, x_{34}^0)$ כך ש- $x_{12}^0 = 1, x_{13}^0 = 4, x_{24}^0 = 4, x_{32}^0 = 3, x_{34}^0 = 1$

עבור נקודה זאת אילוצי שיוויון מתקיימים, ואילוצי אי-שיוויון מתקיימים באופן חריף.

נספח 2 - ממוש שיטת המרות מחסום משופרות ע"י שימוש במתל"ב:

שיטת המרות מחסום משופרות ממומשת ע"י פונקציות מתל"ב. השתמשנו בשתי המרות מחסום משופרות של Polyak [12].

לפי שיטת המרות מחסום בכל איטרציה צריך לפתור בעיית אופטימיזציה לא מאולצת על מנת למצוא $x^{k+1} = \arg \min(F(\varepsilon, x, y^k))$. השתמשנו בפונקציה מבנית של מתל"ב `fminsearch`. פונקציה עובדת לפי שיטה `downhill simplex method`, שהיא אחת השיטות המקובלות לפתרון בעיות אופטימיזציה לא מאולצת ולא לינארית.

בכל שלב חישבנו ערך: $\max\{\|x^{s+1} - x^s\|, \|y^{s+1} - y^s\|\}$. הערך הזה קטן מאיטרציה לאיטרציה. לכן תנאי עצירה הוא כאשר הערך $\max\{\|x^{s+1} - x^s\|, \|y^{s+1} - y^s\|\}$ קטן מ- ε . כי אז נהיה קרובים לפתרון לכל היותר עם שגיאה של ε .

נציג את הפונקציה מתל"ב `BarrierMethodSol`:

```
function BarrierMethodSol(e,x0,y0,func,st)
%e - epsilon
%x0 - initial x value
%y0 - initial y value
%func - objective function handle
%st - constraints function handle

%solution with use of first barrier function

dfrn=1;
y1=y0;
x1=x0;
x2=x0;
res(1,:)=[x1 y1];
l=2;
while(dfrn>e)
    x1=fminsearch(@(x) Brfunc(x,y1,e,func,st),x2);
    g=st(x1);

    for j=1:length(y0)
        y2(j)=y1(j)*( g(j)/e+1)^(-2);
    end
    res(1,:)=[x1 y2];
    norm1(l)=max((sum((y2-y1).^2))^0.5,(sum((x1-x2).^2))^0.5);
    dfrn=norm1(l);
    fnc(l)=objectivef(x1);
    x2=x1;
    y1=y2;
    obj(l)=objectivef(x1);
    l=l+1;
end

%plots the graph
for a=1:l-2
    b=res(a,:);
    c=res(l-2,:);
    dfr(a)=max(abs(b-c));
    fnc_dfr(a)=max(abs(fnc(a)-fnc(l-1)));
end
dfr(l-1)=0;
fnc_dfr(l-1)=0;
norm1(l)=0;
plot(0:l-2,dfr,'o-');grid
xlabel('number of iterations s');
ylabel('max{||Xs-x*||,||Ys-y*||}');
title('barrier function e*(t/e+1)^(^-1^-1)-1');
figure
plot(0:l-2,fnc_dfr(1:end),'o-');grid
xlabel('number of iterations s');
ylabel('||f(Xs)-f(x*)||');
figure

%number of iterations
l-1
```

```

%solution with use of the second barrier function
dfrn=1;
y1_log=y0;
x1_log=x0;
x2=x0;
res2(1,:)=[x1_log y1_log];
l=2;
fnc(1)=objectivef(x0);
while(dfrn>e)
    x1_log=fminsearch(@ (x) Brfunc2(x,y1_log,e,func,st),x2);
    g=st(x1_log);

    for j=1:length(y0)
        y2(j)=y1_log(j)*(g(j)/e+1)^(-1);
    end
    res2(1,:)=[x1_log y2];
    norm1(1)=max((sum((y2-y1_log).^2))^0.5,(sum((x1_log-x2).^2))^0.5);
    fnc(1)=objectivef(x1_log);
    dfrn=norm1(1);
    x2=x1_log;
    y1_log=y2;
    l=l+1;
end

%the number of iterations
l-1
for a=1:l-2
    b=res2(a,:);
    c=res2(l-2,:);
    dfr2(a)=max(abs(b-c));
    fnc_dfr2(a)=max(abs(fnc(a)-fnc(l-1)));
end

dfr2(l-1)=0;
fnc_dfr2(l-1)=0;
plot(0:l-2,dfr2,'o-');grid
xlabel('number of iterations s');
ylabel('max{||Xs-x*||,||Ys-y*||}');
title('barrier function e*log(t/e+1)');
figure
plot(0:l-2,fnc_dfr2(1:end),'o-');grid
xlabel('number of iterations s');
ylabel('||f(Xs)-f(x*)||');

%the solutions
x1
x1_log
objectivef(x1)
objectivef(x1_log)

%Lagrangian function for barrier function1
function A=Brfunc(x,y,e,func,st)
g=st(x);
f=func(x);
A=f;
for i=1:length(g)
    A=A+y(i)*bf1(g(i),e);
end

%Lagrangian function for barrier function2
function A=Brfunc2(x,y,e,func,st)
g=st(x);
f=func(x);
A=f;
for i=1:length(g)
    A=A-y(i)*bf2(g(i),e);
end

%barrier function number 1
function f=bf1(t,e)

```

```
f=e*((t/e+1)^(-1)-1);  
  
%barrier function number 2  
function f=bf2(t,e)  
f=e*log(t/e+1);
```

עבור בעיה 7.2 פונקציה BarrierMethodSol מקבלת נתונים הבאים:

```
e=0.1;
y0=[1 1 1 1 1 1 1 1 1 1 1 1];
x0=[1 4 3 1 4];
BarrierMethodSol(e,x0,y0,@objectivef,@constraintsf);
```

```
function f=objectivef(x)
f=5*x(1)+x(1)^2/(10-x(1))+
30*x(2)^2/(30-x(2))+x(3)+
10*x(3)^2/(10-x(3))+
30*x(4)^2/(30-x(4))+
5*x(5)+x(5)^2/(10-x(5));
```

```
function g=constraintsf(x)
g1=x(1);
g2=x(2);
g3=x(3);
g4=x(4);
g5=x(5);
g6=(x(2)-x(3)-x(5));
g7=(x(1)+x(3)-x(4));
g8=(x(4)+x(5)-5);
g9=10-x(1);
g10=30-x(2);
g11=10-x(3);
g12=30-x(4);
g13=10-x(5);
g=[g1 g2 g3 g4 g5 g6 g7 g8 g9 g10 g11 g12 g13];
```

נראה שהמרות מחסום $\psi_1(\varepsilon, t) = \varepsilon \log(\frac{t}{\varepsilon} + 1)$ & $\psi_2(\varepsilon, t) = \varepsilon[(\frac{t}{\varepsilon} + 1)^{-1} - 1]$ שייכות למחלקה B.

$$\lim_{\varepsilon \rightarrow 0} \psi_1(\varepsilon, t) = \lim_{\varepsilon \rightarrow 0} \varepsilon \log(\frac{t}{\varepsilon} + 1) = \lim_{\varepsilon \rightarrow 0} \frac{\log(\frac{t}{\varepsilon} + 1)}{\frac{1}{\varepsilon}} = \lim_{\varepsilon \rightarrow 0} \frac{\left[\log(\frac{t}{\varepsilon} + 1) \right]_{\varepsilon}}{\left[\frac{1}{\varepsilon} \right]_{\varepsilon}} = 0 \quad \text{א.}$$

$$\lim_{\varepsilon \rightarrow 0} \psi_2(\varepsilon, t) = \lim_{\varepsilon \rightarrow 0} \varepsilon[(\frac{t}{\varepsilon} + 1)^{-1} - 1] = \lim_{\varepsilon \rightarrow 0} \varepsilon \left(\frac{\varepsilon}{\varepsilon + t} - 1 \right) = \psi_2(0, t) = 0$$

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \psi_{1\varepsilon}'(\varepsilon, t) = \lim_{\varepsilon \rightarrow 0} \left[\varepsilon \log(\frac{t}{\varepsilon} + 1) - \frac{\varepsilon}{t + \varepsilon} \right] = 0 \quad \text{ב.}$$

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \psi_{2\varepsilon}'(\varepsilon, t) = \lim_{\varepsilon \rightarrow 0} \varepsilon \left(\frac{1}{t + \varepsilon} - \frac{\varepsilon}{(t + \varepsilon)^2} - \varepsilon \right) = 0$$

$$\psi_1(\varepsilon, 0) = \varepsilon \log(\frac{0}{\varepsilon} + 1) = 0 \quad \text{ג.}$$

$$\psi_2(\varepsilon, 0) = \varepsilon((\frac{0}{\varepsilon} + 1)^{-1} - 1) = 0$$

ד. ψ_1 - עולה וקעורה, ψ_2 - יורדת וקמורה

$$-\infty < \frac{-t\varepsilon}{t + \varepsilon} = \varepsilon \left[\left(\frac{t}{\varepsilon} + 1 \right)^{-1} - 1 \right]; \quad \varepsilon \log(\frac{t}{\varepsilon} + 1) < \infty \quad \text{אם } t < \infty \quad \text{ה.}$$

$$-1 \leq \psi_{2t}'(\varepsilon, t) = -\frac{\varepsilon^2}{(t + \varepsilon)^2} \leq 0; \quad 0 \leq \psi_{1t}'(\varepsilon, t) = \frac{\varepsilon}{t + \varepsilon} \leq 1; \quad \forall t \geq 0 \quad \text{ו.}$$

$$\lim_{\varepsilon \rightarrow 0} \psi_{2tt}''(\varepsilon, 0) = \lim_{\varepsilon \rightarrow 0} \frac{2\varepsilon^2}{(t + \varepsilon)^3} = \infty; \quad \lim_{\varepsilon \rightarrow 0} \psi_{1tt}''(\varepsilon, 0) = \lim_{\varepsilon \rightarrow 0} -\frac{\varepsilon}{(t + \varepsilon)^2} = -\infty \quad \text{ז.}$$

$$\lim_{\varepsilon \rightarrow 0} \psi_{2tt}''(\varepsilon, t) = \lim_{\varepsilon \rightarrow 0} \frac{2\varepsilon^2}{(t + \varepsilon)^3} = 0; \quad \lim_{\varepsilon \rightarrow 0} \psi_{1tt}''(\varepsilon, t) = \lim_{\varepsilon \rightarrow 0} -\frac{\varepsilon}{(t + \varepsilon)^2} = 0; \quad \forall t \neq 0$$

$$\lim_{\varepsilon \rightarrow 0} \psi_{2tt}''(\varepsilon, 0) = \lim_{\varepsilon \rightarrow 0} \varepsilon^1 \frac{2\varepsilon^2}{(t + \varepsilon)^3} = 2; \quad \lim_{\varepsilon \rightarrow 0} \psi_{1tt}''(\varepsilon, 0) = \lim_{\varepsilon \rightarrow 0} -\varepsilon^1 \frac{\varepsilon}{(t + \varepsilon)^2} = -1 \quad \text{ח.}$$

$$\lim_{t \rightarrow -\varepsilon} \psi_1(\varepsilon, t) = \lim_{\varepsilon \rightarrow 0} \varepsilon \log(\frac{t}{\varepsilon} + 1) = -\infty; \quad \lim_{t \rightarrow -\varepsilon} \psi_2(\varepsilon, t) = \lim_{\varepsilon \rightarrow 0} \varepsilon[(\frac{t}{\varepsilon} + 1)^{-1} - 1] = \infty; \quad \text{ט.}$$

משפט 1 (The modified Debreu Theorem)

יהיו A מטריצה סימטרית מסדר $n \times n$, B היא מטריצה מסדר $n \times r$ ו- $U = [\text{diag } u_i]_{i=1}^r$ מטריצה אלכסונית מסדר $r \times r$.

אם התנאי $u_i > 0$ לכל $i = \overline{1, r}$ ו- $By = 0$ גורר שקיים $\lambda > 0$; $\langle Ay, y \rangle \geq \lambda \|y\|^2$;

אז קיים $\varepsilon_0 > 0$ כך שלכל $0 < \mu < \lambda$ ו- $\varepsilon \leq \varepsilon_0$ מתקיים $\langle (A + \frac{1}{\varepsilon} B^T U B)x, x \rangle \geq \mu \langle x, x \rangle, \forall x \in R^n$

ראה הוכחה ב-[1].

מסקנה:

אם נסמן $A = L''_{xx}$, $B = g^{(r)'} = U^*$ ו- $U = U^*$, אז תנאי משפט Debreu מתקיימים עבור מטריצות הנ"ל.

למת עזר: אם תנאי משפט 5.4 מתקיימים, אז $\Phi'_{(0)}$ היא מטריצה לא סינגולרית.

(מטריצה $\Phi'_{(0)}$ מוגדרת בעמוד 15)

הוכחה:

$\Phi'_{(0)} w = 0$ היה $w \in R^{2r}$ וקטור כלשהו הפותר את המשוואה $\Phi'_{(0)} w = 0$. $\Phi'_{(0)} = \begin{pmatrix} L''_{xx} & g^{(r)T} \\ U^* g^{(r)} & 0^{r \times r} \end{pmatrix}$

נסמן $w = \begin{bmatrix} u \\ v \end{bmatrix}$, ונוכל לרשום $\Phi'_{(0)} w = \begin{pmatrix} L''_{xx} & g^{(r)T} \\ U^* g^{(r)} & 0^{r \times r} \end{pmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{pmatrix} L''_{xx} u + g^{(r)T} v \\ U^* g^{(r)} u + 0 \end{pmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ מכאן נובע $\begin{cases} L''_{xx} u + g^{(r)T} v = 0 \\ U^* g^{(r)} u = 0 \end{cases}$

כעת, $U^* > 0$ $\Leftrightarrow g^{(r)} u = 0$ לכן $\langle L''_{xx} u, u \rangle = 0 \Leftrightarrow u = 0$ (לפי (5) של למה 5.3).

מצד שני אם $u = 0$ אז משוויין הראשון נובע שגם $g^{(r)T} v = 0$ כאשר $g^{(r)}$ מטריצה בעלת r עמודות בת"ל ולכן חייב להתקיים $v = 0$.

כלומר חייב להתקיים $\Phi'_{(0)} w = 0$ אם ורק אם $w = 0$ ולכן המטריצה $\Phi'_{(0)}$ אינה סינגולרית.

משפט פונקציות סתומות 2 (ראה [2], עמוד 12):

תהי $E = \{0 \in R^m\} \times [\varepsilon_1, \varepsilon_0]$ קבוצה קומפקטית. ותהי $\Phi(x, \varepsilon, \hat{y}, t) : R^{n+r+m+1} \rightarrow R^{m+r}$ פונקציה גזירה פעמיים ברצפות. נניח שווקטור $(\bar{x}, \bar{y}, \varepsilon, t) \in R^{n+r+m+1}$ מקיים: $\Phi(\bar{x}, \bar{y}, \varepsilon, t) = 0$ כך שמטריצה $\Phi'_{(x, \hat{y})}(\bar{x}, \bar{y}, \varepsilon, t)$ אינה סינגולרית לכל $(\varepsilon, t) \in E$.

אזי קיים סקלר $\alpha > 0$ ושתי פונקציות ווקטוריות גזירות פעמיים ברציפות

$$x(\varepsilon, t) = (x_1(\varepsilon, t), \dots, x_n(\varepsilon, t)), \quad \hat{y}(\varepsilon, t) = (y_1(\varepsilon, t), \dots, y_r(\varepsilon, t))$$

מוגדרות באופן יחיד בסביבה $S(E, \alpha) = \{(\varepsilon, t) \mid |t_i| \leq \alpha, i = \overline{1, m}; \varepsilon \in [\varepsilon_1, \varepsilon_0]\}$

כך ש- $\hat{y}(\varepsilon, 0) = \bar{y}$, $x(\varepsilon, 0) = \bar{x}$ לכל $\varepsilon \in [\varepsilon_1, \varepsilon_0]$ וכך ש- $\Phi(x(\varepsilon, t), y(\varepsilon, t), \varepsilon, t) \equiv 0$ ובנוסף:

$$(\dagger) \quad \begin{cases} \nabla_\varepsilon x(\varepsilon, t) = -\Phi'_{(\varepsilon)}(x(\varepsilon, t), \hat{y}(\varepsilon, t), \varepsilon, t) \cdot [\Phi'_{(x)}(x(\varepsilon, t), \hat{y}(\varepsilon, t), \varepsilon, t)] \\ \nabla_\varepsilon y(\varepsilon, t) = -\Phi'_{(\varepsilon)}(x(\varepsilon, t), \hat{y}(\varepsilon, t), \varepsilon, t) \cdot [\Phi'_{(\hat{y})}(x(\varepsilon, t), \hat{y}(\varepsilon, t), \varepsilon, t)] \end{cases}$$

מסקנה: תוצאה נוספת שמתקבלת מהביטוי $(\dagger)$ היא שקיים $\beta > 0$ כך שמתקיים

$$(\|x(\varepsilon, t) - \bar{x}\|^2 + \|y(\varepsilon, t) - \bar{y}\|^2)^{\frac{1}{2}} \leq \beta$$