

המחלקה למתמטיקה
Department of Mathematics

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

**רצולריזציה של בעיות לא מוגדרות היטב ושימושה
בבעית קרוב של פונקציה**

פולינה רויזמן

**Regularization of ill-posed problems with
application in function approximation**

Polina Roizman

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

**רצולריזציה של בעיות לא מוגדרות היטב ושימושה
בבעית קרוב של פונקציה**

פולינה רויזמן

**Regularization of ill-posed problems with
application in function approximation**

Polina Roizman

Advisor:

Prof. Vladimir
Turetsky

מנחה:

פרופסור ולדימיר
טורצקי

Karmiel

כרמיאל

2018

תוכן עניינים

4.....	הקדמה:
4.....	הגדרות כלליות:
6.....	הקרוב הטוב ביותר:
9.....	בעיה מוגדרת היטב ו conditional number:
13.....	הריגורויזציה של טיחונוב:
16.....	ישום התאוריה:
21.....	התוכנה בMATLAB:
23.....	ביבליוגרפיה:

הקדמה:

מערכות ליניאריות רבות "מזוהמות" מטעויות או סטיות. במצב זה קשה להגיע לפתרון מדויק כאשר הנתונים אינם כאלו.

בפרויקט זה בחרתי להתעמק ולהוכיח את הרגולריזציה של טיחונוב אשר הוצגה על ידי פיליפס בשנת 1962 במקביל לטיחונוב אשר הציג אותה בשנת 1963. ריגולריזציה זו תסייע בפתרון מקורב של בעיות אשר מוגדרות ill- conditioned.

.

הגדרות כלליות:

הגדרה: מרחב נורמי הינו מרחב וקטורי אשר מוגדרת בו נורמה.

דוגמה: הערך המוחלט הסטנדרטי הוא נורמה המוגדרת על הישר הממשי עצמו במרחב בעל מימד אחד.

הגדרה: מרחב טרום הילברט הינו מרחב אשר מוגרת בו מכפלה פנימית.

דוגמה: אחד המרחבים הידועים הינו $L^2(x)$ המוגדר להיות קבוצת כל הפונקציות $f: X \rightarrow \mathbb{R}$ שהן אינטגרביליות והוא מרחב הילברט עם מכפלה פנימית

$$(f, g) = \int_a^b f(x)g(x)dx$$

הגדרה: מרחב בנך הינו מרחב ליניארי נורמי שלם.

דוגמה: אוסף הסדרות של מספרים ממשיים $a_n, n \in \mathbb{N}$ שמקיימות $\sum_{n=1}^{\infty} |a_n| < \infty$ הוא מרחב בנך עם הנורמה $\|a_n\|_1 = \sum_{n=1}^{\infty} |a_n|$ מרחב זה מסומן l_1

הגדרה: A תת קבוצה U של מרחב נורמי X תיקרא חסומה אם קיים מספר חיובי C כזה שמקיים $\|x\| \leq C$ עבור כל $x \in U$.

הגדרה: יהיו B, C מרחבים נורמים, הנורמה של אופרטור ליניארי $T: B \rightarrow C$ המסומנת $\|T\|$ היא הסופרימום של הנורמות $\|T(x)\|_C$, כאשר $x \in B$ הוא וקטור היחידה.

הגדרה: מטריצה הרמיטית הינה מטריצה ריבועית אשר פעולת ההצמדה איננה משנה אותה

דוגמה: המטריצה $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ הינה מטריצה הרמיטית מכיוון שאחרי פעולת ההצמדה היא תישאר אותו הדבר.

הקרוב הטוב ביותר:

הגדרה 1.1: יהיה $U \subset X$ תת מרחב של מרחב נורמי X ויהיה $w \in X$. אלמנט $v \in U$ נקרא הקרוב הטוב ביותר ל w ביחס ל U אם

$$\|w - v\| = \inf_{u \in U} \|w - u\|$$

כלומר, אם ל $v \in U$ יש את המרחק הקטן ביותר מ w .

משפט 1.1: יהיה U תת מרחב ליניארי בעל מספר סופי של מימדים של מרחב טרום-הילברט X עם בסיס $u_1, \dots, u_n$. הקובמינציה הליניארית

$$v = \sum_{k=1}^n a_k u_k$$

היא הקרוב הטוב ביותר ל $w \in X$ ב U אם ורק אם המקדמים $a_1, \dots, a_n$ מקיימים את המערכת המשוואות ליניאריות אלגבריות הבאה:

$$\sum_{k=1}^n a_k (u_k, u_j) = (w, u_j), j = 1, \dots, n$$

הערה 1.1: כל סדרה מתכנסת הינה חסומה

למה 1.1: כל סדרה חסומה במרחב נורמי סופי X מכילה תת סדרה מתכנסת.

למה 1.2: יהיה U תת מרחב בעל מספר סופי של מימדים של המרחב הנורמי X , אזי עבור כל אלמנט $w \in X$ קיים הקרוב הטוב ביותר ביחס ל U .

הוכחה: יהי $w \in X$, נבחר בסדרה המינימאלית (u_n) עבור w כלומר $u_n \in U$ מקיים $\|w - u_n\| \rightarrow d := \inf_{u \in U} \|w - u\|$, $n \rightarrow \infty$. מכיוון ש $\|w - u_n\| \leq \|w - u_n\| + \|w\|$ אזי הסדרה (u_n) חסומה. לפי למה 1.1, הסדרה (u_n) מכילה תת מרחב מתכנס $(u_{n(i)})$ עם גבול $v \in U$. אזי $\|w - v\| = \lim_{i \rightarrow \infty} \|w - u_{n(i)}\| = d$.
מש"ל.

למה 1.3: יהיה U תת מרחב של מרחב טרום-הילברט X . אלמנט v הוא הקרוב הטוב ביותר ל $w \in X$ ב U אם ורק אם $(w - v, u) = 0$ עבור כל $u \in U$. כלומר אם ורק אם $w - v \perp U$. עבור כל $w \in X$ קיים לכל היותר קרוב טוב ביותר אחד ב U .

הוכחה: נציין את המשוואה

$$\|w - u\|^2 = \|w - v\|^2 + 2 \operatorname{Re}(w - v, v - u) + \|v - u\|^2$$

אשר קיימת עבור כל $u, v \in U$, מכאן הוכח כי התנאי $(w - v, u) = 0$ מספיק מכיוון ש U הוא תת מרחב ליניארי. כדי לבסס את הצורך אנו מניחים ש v הוא הקרוב הטוב ביותר וכי $(w - v, u_0) \neq 0$ עבור כמה $u_0 \in U$ לכן, מכיוון ש U הוא תת מרחב ליניארי אנו יכולים להניח כי $(w - v, u_0) \in \mathbb{R}$.
נבחר $u = v + \frac{(w - v, u_0)}{\|u_0\|^2} u_0$ ונגיע ל

$$\|w - u\|^2 = \|w - v\|^2 - \frac{(w - v, u_0)}{\|u_0\|^2} \leq \|w - v\|^2$$

אשר סותר את ההנחה ש v הוא הקרוב הטוב ביותר של w . לבסוף נניח כי v_1 ו v_2 הם הקרובים הטובים ביותר, מהתנאי $(w - v, u) = 0$ נובע כי

$$(w - v_1, v_1 - v_2) = 0 = (w - v_2, v_1 - v_2)$$

זה מצביע על כך ש $(v_1 - v_2, v_1 - v_2) = 0$ וכי $v_1 = v_2$.
מש"ל.

הערה 1.2: יהיו X, Y, Z מרחבים נורמים ויהיו $A: X \rightarrow Y, B: Y \rightarrow Z$ אופרטורים ליניאריים חסומים, אזי $BA: X \rightarrow Z$ המוגדר על ידי $(BA)x := B(Ax)$ עבור כל $x \in X$, הוא אופרטור ליניארי חסום עם

$$\|BA\| \leq \|A\| \|B\|$$

למה 1.4: יהיה U תת מרחב ליניארי בעל מספר סופי של מימדים של מרחב טרום-הילברט X . אזי לכל אלמנט $w \in X$ קיים קרוב הטוב ביותר יחיד ב U . האופרטור $P: X \rightarrow U$ הממפה $w \in X$ לתוך הקרוב הטוב ביותר שלו הוא אופרטור ליניארי חסום עם התכונות $P^2 = P, \|P\| = 1$ והוא ניקרא היטל אורטוגונלי מ X ל U .

$$\|w - u_n\|^2 \leq d^2 + \frac{1}{n}, n \in \mathbb{N} \text{ כאשר } (u_n) \text{ נבחר סדרה}$$

אכן ש $d := \inf_{u \in U} \|w - u\|$ כאשר

$$\|(w - u_n) + (w - u_m)\|^2 + \|u_n - u_m\|^2 = 2\|w - u_n\|^2 = 2\|w - u_m\|^2 + 2\|w - u_m\|^2 \leq 4d^2 + \frac{2}{n} + \frac{2}{m}$$

עבור כל $n, m \in \mathbb{N}$ מכיון $\frac{1}{2}(u_n + u_m) \in U$ נובע כי

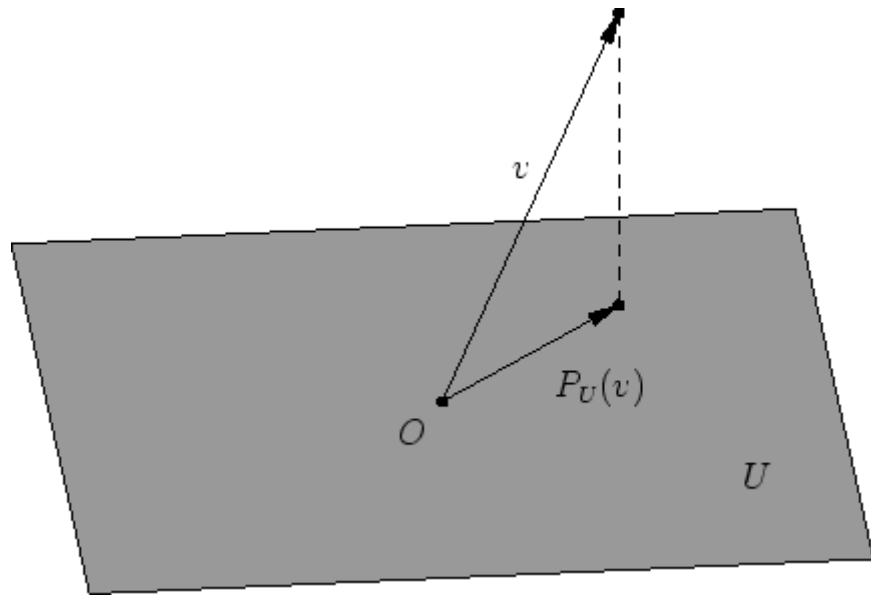
$$\|u_n - u_m\|^2 \leq 4d^2 + \frac{2}{n} + \frac{2}{m} - 4 \left\| w - \frac{1}{2}(u_n + u_m) \right\|^2 < \frac{2}{n} + \frac{2}{m}$$

מכאן נובע כי (u_n) היא סדרת קושי ומכיון ש U הוא מרחב שלם אזי קיים אלמנט $v \in U$ כזה ש $u_n \rightarrow v, n \rightarrow \infty$. כאשר $\|w - u_n\|^2 \leq d^2 + \frac{1}{n}, n \rightarrow \infty$ נקבל כי v הוא הקרוב הטוב ביותר ב U . יחידות הפתרון נובעת מלמה 1.3. באופן טריוויאלי נקבל כי $P_u = u$ עבור כל $u \in U$. מכאן נקבל כי $P^2 = P$. מהמשוואה $(w - v, v) = 0$ ניתן להסיק כי P הוא אופרטור ליניארי אשר מקיים

$$\|w\|^2 = \|Pw\|^2 + \|w - Pw\|^2 > \|Pw\|^2$$

עבור כל $w \in X$. לכן P הוא חסום עם $\|P\| \leq 1$, בשל הערה 1.2 וכי $P^2 = P$ נובע כי $\|P\| \geq 1$.
אשר סותר את ההוכחה לכן $\|P\| = 1$.

מש"ל.



הערה 1.3: המשוואה הנורלית של הקרוב הטוב ביותר בתוך מרחב טרום-הילברט הופכת את הפתרון לטריבויאלי אם $u_1, \dots, u_n$ אורתונורמאליים.

מסקנה 1.1: יהיה U תת מרחב ליניארי בעל מספר סופי של מימדים של מרחב טרום-הילברט X עם בסיס $u_1, \dots, u_n$ אזי ההיטל האורתונורמאלי נתון על ידי

$$P_w = \sum_{k=1}^n (w, u_k) u_k, \quad w \in X$$

וכך הוכחנו את משפט 1.1

בעיה מוגדרת היטב ו conditional number:

בעיה המוגדרת היטב צריכה לקיים 3 תנאים:

1. פתרון קיים
2. פתרון יחיד
3. הפתרון צריך להיות תלוי בנתונים באופן רציף

התנאי השלישי הוא הבעיית מביין השלושה מכיוון שהנתנים תמיד מזוהמים מטעויות.

בעיה אשר לא מקיימת את אחד או יותר מתנאים אלו תיקרא בעיה לא מוגדרת היטב או ill-posed.

דוגמה 2.1: אחד התרגילים הפשוטים הוא $y' = \sqrt{x}, y(0) = 0$. קיימים שני פתרונות לתרגיל זה והם $y = \left(\frac{x}{2}\right)^2, y \equiv 0$. הפתרון קיים אך הוא אינו יחיד ולכן זו בעיה אשר לא מוגדרת היטב.

מכיוון שאנו רוצים כי הטעויות יהיו מינימליות אנו צריכים להגדיר את הרמה שבה הבעיה מוגדרת היטב.

הגדרה 2.1: $A: X \rightarrow Y$ כך ש A אופרטור ליניארי וקיים $A^{-1}: Y \rightarrow X$ אזי

$$\text{cond}(A) = \|A\| \|A^{-1}\|$$

כך ש $\text{cond}(A)$ יקרא Condition number של A .

הערה 2.1: המספר תלוי בסוג הנורמה הנבחרת (נורמה $\infty, 1, 2$)
 $1 = \|I\| = \|AA^{-1}\| \leq \text{cond}_p(A) = \|A\| \|A^{-1}\|$ לכן $\text{cond}_p(A) \geq 1$. במקרה שנתונים המימד והתחום של מטריצה A נרשום $\text{cond}_p(A)$ $p = 1, 2, \infty$.

דוגמה 2.2: נבחר במטריצה הבאה: $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. נחשב את $\text{cond}_p(A)$ כאשר $p = 1, 2$ כך שנקבל $\text{cond}_1(A) = 21, \text{cond}_2(A) = 14.933$.

נגדיר כי erroneous של צד ימין של המשוואה $Ax = y$ הוא y^δ עם שגיאה ברמה $\|y^\delta - y\|_2 \leq \delta$

משפט 2.1: יהיו X ו Y מימדי בנך. $A: X \rightarrow Y$ אופרטור ליניארי שמקיים $A^{-1}: Y \rightarrow X$ ויהי $A^\delta: X \rightarrow Y$ אופרטור ליניארי כך שמתקיים $\|A^{-1}\| \|A^\delta - A\| < 1$ בהנחה ש x ו x^δ פתרונות של המשוואות $Ax = y, A^\delta x^\delta = y^\delta$ אזי מתקיים:

$$\frac{\|x^\delta - x\|}{\|x\|} \leq \frac{\text{cond}(A)}{1 - \text{cond}(A) \frac{\|A^\delta - A\|}{\|A\|}} \left\{ \frac{\|y^\delta - y\|}{\|y\|} + \frac{\|A^\delta - A\|}{\|A\|} \right\}$$

הוכחה: נרשום כי $A^\delta = A^\delta + A - A = A + A A^{-1} (A^\delta - A) = A [I + A^{-1} (A^\delta - A)]$
 $(A^\delta)^{-1} = [I + A^{-1} (A^\delta - A)]^{-1} [A]^{-1}$

$$\|[A^\delta]^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|A^\delta - A\|}$$

$$A^\delta (x^\delta - x) = A^\delta x^\delta - Ax - A^\delta x + Ax = y^\delta - y - (A^\delta - A)x$$

$$x^\delta - x = [A^\delta]^{-1} \{y^\delta - y - (A^\delta - A)x\}$$

$$\|x^\delta - x\| \leq \|A^\delta\|^{-1} \{\|y^\delta - y\| + \|A^\delta - A\| \|x\|\} \quad \because \|x\|$$

$$\|A\| \|x\| \geq \|y\|$$

$$\begin{aligned} \frac{\|x^\delta - x\|}{\|x\|} &\leq \frac{\|A^\delta\|^{-1}}{\|x\|} \{ \|y^\delta - y\| + \|A^\delta - A\| \|x\| \} \\ \frac{\|x^\delta - x\|}{\|x\|} &\leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|A^\delta - A\|} \left\{ \frac{\|y^\delta - y\|}{\|x\|} + \|A^\delta - A\| \right\} \\ &\quad \|A\| \text{ נכפול ונחלק את צד ימין ב} \\ \frac{\|x^\delta - x\|}{\|x\|} &\leq \frac{\|A\| \|A^{-1}\|}{1 - \|A^{-1}\| \|A^\delta - A\|} \left\{ \frac{\|y^\delta - y\|}{\|A\| \|x\|} + \frac{\|A^\delta - A\|}{\|A\|} \right\} \\ 1 - \|A^{-1}\| \|A^\delta - A\| &= 1 - \text{cond}(A) \frac{\|A^\delta - A\|}{\|A\|} \\ \frac{\|x^\delta - x\|}{\|x\|} &\leq \frac{\text{cond}(A)}{1 - \text{cond}(A) \frac{\|A^\delta - A\|}{\|A\|}} \left\{ \frac{\|y^\delta - y\|}{\|y\|} + \frac{\|A^\delta - A\|}{\|A\|} \right\} \end{aligned}$$

מש"ל.

מכאן ניתן להבין כי $\text{cond}(A)$ יכול לשמש כמידת היציבות של האופרטור הליניארי A וביחוד של מערכת משוואות.

במערכת של משוואות ליניאריות ככל שה $\text{cond}(A)$ קטן יותר כך שהמערכת יציבה יותר ולהפך.

הערה 2.4: כאשר A הינה מטריצה הרמיטית אזי $\text{cond}_2(A) = \frac{|\lambda_{\max}|}{|\lambda_{\min}|}$ כאשר λ הינו ערך עצמי

דוגמה: נבחר במטריצה $A = \begin{pmatrix} 7 & 3 \\ 3 & -1 \end{pmatrix}$ הערכים העצמיים שלה הם $\lambda_1 = 8, \lambda_2 = -2$

$$\text{cond}_2(A) = \frac{|\lambda_{\max}|}{|\lambda_{\min}|} = \frac{8}{2} = 4 \text{ נחשב את}$$

משפט 2.2: יהי A מטריצה $m \times n$ עם $\text{rank}(A) = r$ אזי קיימים מספרים אי-שלילים

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_r > \mu_{r+1} = \dots = \mu_n$$

וקטורים אורתוגונליים $u_1, \dots, u_n \in \mathbb{C}^n$ ו $v_1, \dots, v_m \in \mathbb{C}^m$ כך ש

$$\begin{cases} Au_j = \mu_j v_j, & A^* v_j = \mu_j u_j, & j = 1, \dots, r \\ Au_j = 0, & j = r+1, \dots, n \\ A^* v_j = 0, & j = r+1, \dots, m \end{cases}$$

ויהי $x \in \mathbb{R}^n$ אזי $Ax = \sum_{j=1}^r \mu_j (x, u_j) v_j$. כל המערכת (μ_j, u_j, v_j) עם התכונות הללו תיקרא המערכת העצמית של מטריצה A .

הוכחה: לכל מטריצה הרמיטית $A^* A$ עם $\text{rank}(A) = r$ יש n וקטורים אורתוגונליים $u_1, \dots, u_n$ השייכים לערכים עצמיים אי-שלילים שונים כך שמתקיים: $A^* A u_j = \mu_j^2 u_j$, $j = 1, \dots, n$ והם מסודרים כך:

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_r > 0 \text{ ו } \mu_{r+1} = \mu_{r+2} = \dots = \mu_n = 0$$

$$v_j = \frac{1}{\mu_j} A u_j, \quad j = 1, \dots, r$$

$$(v_j, v_k) = \frac{1}{\mu_j \mu_k} (Au_j, Au_k) = \frac{1}{\mu_j \mu_k} (u_j, A^* Au_k) = \delta_{j,k} \quad , \quad j, k = 1, \dots, r$$

$$A^* v_j = \mu_j u_j \quad , \quad j = 1, \dots, r \quad \text{לכן } \delta_{j,k} = \begin{cases} 1, & k = j \\ 0, & k \neq j \end{cases}$$

ידוע כי $Nul(A^*A) = Nul(A)$ ולכן $Au_j = 0$, $j = r+1, \dots, n$ אם $r < m$, לפי תהליך גראם-שמידט אנו יכולים להרחיב את $v_1, \dots, v_r$ לבסיס אורתונורמלי $v_1, \dots, v_m$ של $\mathbb{R}^m$. מכיוון שמתקיים $rank(A^*) = r$ אזי $\dim(N(A^*)) = m - r$ לכן $A^* v_j = 0$, $j = r+1, \dots, m$. מש"ל.

מכיוון ש $u_1, \dots, u_n$ יוצרים בסיס אורתונורמלי ל $\mathbb{R}^n$, נוכל לרשום $x \in \mathbb{R}^n$ כך:

$$x = \sum_{j=1}^n (x, u_j) u_j$$

משפט 2.3: תהי A מטריצה $m \times n$ עם $rank(A) = r$ ומערכת עצמית (μ_j, u_j, v_j) . המערכת הליניארית $Ax = y$ פתירה אם ורק אם $(y, z) = 0$ עבור כל $z \in \mathbb{R}^m$ עם $A^* z = 0$. במקרה זה הפתרון יהיה מהצורה

$$(2) \quad x_0 = \sum_{j=1}^r \frac{1}{\mu_j} (y, v_j) u_j$$

הוכחה: נניח כי x הינו פתרון וכי מתקיים $A^* z = 0$ אזי $(y, z) = (Ax, z) = (x, A^* z) = 0$. כאן ניתן לראות את נחיצות התנאי $(y, z) = 0$ עבור הפתרון של $Ax = y$. נניח כי תנאי זה מתקיים.

מכיוון ש V הינו בסיס אורתונורמלי של $\mathbb{R}^m$ נוכל להציג את y כך: $y = \sum_{j=1}^r (y, v_j) v_j$. מכיוון שאנו יודעים כי $A^* v_j = 0$ עבור כל $j = r+1, \dots, m$. עבור הוקטור x_0 המוגדר לפי (2) אנו נקבל כי

$$Ax_0 = y \quad \text{ומכאן ניתן להבין גם } Ax_0 = \sum_{j=1}^r (y, v_j) v_j.$$

מש"ל.

מכיוון ש $N(A) = \text{span}\{u_{r+1}, \dots, u_n\}$ אזי הוקטור x_0 (הפתרון) מקבל יכולת של $(x_0, x) = 0$ עבור כל $x \in N(A)$. במקרה זה למשוואה $Ax = y$ יש יותר מפתרון אחד. מוסיפים לפתרון x_0 את הפתרון של המשוואה ההומוגנית $Ax = 0$ כך ש x שרירותי והינו הפתרון של המשוואה ההומוגנית

$$\|x_0 + x\|_2^2 = \|x_0\|_2^2 + 2\text{Re}(x_0, x) + \|x\|_2^2 = \|x_0\|_2^2 + \|x\|_2^2$$

ניתן לראות כי הפתרון x_0 הינו הפתרון הקטן ביותר בנורמה אוקלידית. במקרה ולמערכת $Ax = y$ לא קיים פתרון נציג את y כך: $y = \sum_{j=1}^r (y, v_j) v_j$.

נניח כי x_0 נתון לפי (2) ו $x \in \mathbb{R}^n$ שרירותי אזי $(Ax - Ax_0, Ax_0 - y) = 0$ מכיוון

ש $Ax - Ax_0 \in \text{span}\{v_1, \dots, v_r\}$ ו $Ax_0 - y \in \text{span}\{v_{r+1}, \dots, v_m\}$ אזי נקבל כי

$$\|Ax - y\|_2^2 = \|(Ax - Ax_0) + (Ax_0 - y)\|_2^2 = \|Ax - Ax_0\|_2^2 + \|Ax_0 - y\|_2^2$$

הדומה ברובו לפתרון בשיטת הריבועים המזעריים. ושוב נראה כי (2) הינו הפתרון בעל הנורמה הקטנה ביותר.

על כן ניתן להגדיר אופרטור ליניארי $A^\dagger: \mathbb{R}^m \rightarrow \mathbb{R}^n$ כך:

$$A^\dagger y := \sum_{j=1}^r \frac{1}{\mu_j} (y, v_j) u_j \quad , \quad y \in \mathbb{R}^m$$

משפט 2.4: תהי A מטריצה $m \times n$ עם מערכת עצמית (μ_j, u_j, v_j) ויהי $y \in \mathbb{R}^n$, $y^\delta \in \mathbb{R}^m$ כך שמתקיים $\|y^\delta - y\|_2 \leq \delta \leq \|y^\delta\|_2$ לכל $\delta > 0$. אזי קיים מספר שלם קטן יותר $p = p(\delta)$ כך שיתקיים

$$\|Ax_p - y^\delta\|_2 \leq \delta \quad \text{אזי } x_p \rightarrow A^\dagger y \text{ כאשר } \delta \rightarrow 0.$$

$$F: \{0, 1, \dots, r\} \rightarrow \mathbb{R}, F(p) = \|Ax_p - y^\delta\|_2^2 - \delta^2 \quad \text{הוכחה:}$$

$$(3) \quad F(p) = \sum_{j=p+1}^m |(y^\delta, v_j)|^2 - \delta^2$$

על כן F מונוטונית לא גדלה עם $F(0) = \|y^\delta\|_2^2 - \delta^2 \geq 0$ ו $F(r) = -\delta^2 < 0$ אם $rank(A) = r = m$. אם $r < m$ אז בעזרת $(y, v_j) = 0$, $j = r+1, \dots, m$ נקבל כי

$$F(r) = \sum_{j=r+1}^m |(y^\delta - y, v_j)|^2 - \delta^2 \leq \|y^\delta - y\|_2^2 - \delta^2 \leq 0$$

לכן קיים מספר שלם קטן יותר $p = p(\delta)$ כך ש $F(p) \leq 0$ נשים לב כי $p \leq r$. בתהליך עצמו פרמטר העצירה p נקבע על ידי סיום הסכום של x_p כאשר הצד הימני של $F(p)$ בסכום (3) נהיה קטן או שווה לאפס בפעם הראשונה. על מנת להראות את התהליך נשים לב כי $\|Ax_p - y^\delta\|_2 \leq \delta$ ולכן

$$\|Ax_p - y\|_2 \leq \|Ax_p - y^\delta\|_2 + \|y^\delta - y\|_2 \leq 2\delta \rightarrow 0$$

$$Ax_p \rightarrow y, \quad \delta \rightarrow 0$$

מכך, מכיוון ש $A^\dagger Au = v$ לכל $v \in span\{v_1, \dots, v_r\}$, ניתן להסיק כי $x_p \rightarrow A^\dagger y$, $\delta \rightarrow 0$ מש"ל.

הריגולריזציה של טיחונוב:

רעיון הריגולריזציה אומר בעצם לכפול את $\frac{1}{\mu_j}$ בגורם הנקרא $damping factor$:

$$\frac{\mu_j^2}{\alpha + \mu_j^2}$$

כאשר α נקראת פרמטר הריגולריזציה והוא חיובי.

משפט 3.1: תהי A מטריצה $m \times n$ בעלת $rank(A) = r$ עם מערכת מעצמית (μ_j, u_j, v_j) ויהי $\alpha > 0$, אזי עבור כל $y \in \mathbb{R}^m$ המערכת הליניארית $\alpha x_\alpha + A^* A x_\alpha = A^* y$ הינה בעלת פתרון יחיד והוא:

$$x_\alpha = \sum_{j=1}^r \frac{\mu_j}{\alpha + \mu_j^2} (y, v_j) u_j$$

הוכחה: עבור כל $\alpha > 0$ המטריצה $\alpha I + A^* A$ הינה חיובית ולכן איננה סינגולרית. מכיוון ש $\alpha I + A^* A u_j = (\alpha + \mu_j^2) u_j$ מערכת סינגולרית עבור המטריצה $\alpha I + A^* A$ ונתונה על ידי $(\alpha + \mu_j^2, u_j, u_j)$ כאשר $j = 1, \dots, n$. ממשפט 2.3 ומכך ש $(A^* y, u_j) = (y, A u_j)$ נובע כי הפתרון באמת מהצורה

$$x_\alpha = \sum_{j=1}^r \frac{\mu_j}{\alpha + \mu_j^2} (y, v_j) u_j$$

מש"ל.

מסקנה 3.1: תחת ההנחות של משפט 3.1 ניתן להסיק כי

$$\lim_{\alpha \rightarrow 0} (\alpha I + A^* A)^{-1} A^* y = A^\dagger y$$

המשפט הבא ידבר על האינטרפוטציה של $Tikhonov Regularization$ על שיטת הריבועים המזעריים.

משפט 3.2: תהי A מטריצה $m \times n$ ויהיה $\alpha > 0$, אזי עבור כל $y \in \mathbb{R}^m$ קיים פתרון $x_\alpha \in \mathbb{R}^n$ והוא יחיד כך שמתקיים

$$(4) \quad \|Ax_\alpha - y\|_2^2 + \alpha \|x_\alpha\|_2^2 = \inf_{x \in \mathbb{C}^n} \{\|Ax - y\|_2^2 + \alpha \|x\|_2^2\}$$

הפתרון המינימאלי x_α ניתן על ידי פתרון המערכת הליניארית הנ"ל.

הוכחה:

$$\|Ax - y\|_2^2 + \alpha \|x\|_2^2 = \|Ax_{\alpha\alpha} - y\|_2^2 + \alpha \|x_\alpha\|_2^2 + 2\operatorname{Re}(x - x_\alpha, \alpha x_\alpha + A^* A x_\alpha - A^* y) + \|Ax - Ax_\alpha\|_2^2 + \alpha \|x - x_\alpha\|_2^2$$

כאשר $x, x_\alpha \in \mathbb{R}^n$. ניתן לראות כי הפתרון x_α של משפט 4.1 מרצה את הפתרון של (4).

מצד שני נניח כי x_α הינו פתרון של (4) אך $\alpha x_\alpha + A^* A x_\alpha \neq A^* y$ אזי נגדיר

$z = \alpha x_\alpha + A^* A x_\alpha - A^* y$ עבור $z \in \mathbb{R}$, $\epsilon \in \mathbb{R}$ ואז עבור

$$\|Ax - y\|_2^2 + \alpha \|x\|_2^2 = \alpha \|Ax_\alpha - y\|_2^2 + \alpha \|x_\alpha\|_2^2 - 2\epsilon \alpha + \epsilon^2 b$$

כאשר $a := \|z\|_2^2$, $b := \|Az\|_2^2 + \alpha\|z\|_2^2$ חיוביים. על ידי בחירת $\varepsilon = \frac{a}{b}$ נקבל

$$\|Ax - y\|_2^2 + \alpha\|x\|_2^2 < \|Ax_\alpha - y\|_2^2 + \alpha\|x_\alpha\|_2^2$$

אשר נותן סתירה.

מש"ל.

במשפט זה מראים כיצד מחזיקים את $\|Ax_\alpha - y\|_2^2$ קטן ויציב על ידי איסור על x_α להפוך לגדול במהלך הפעולה $\alpha\|x_\alpha\|_2^2$.

מהוכחת המשפט ניתן להבין כי הערכים העצמיים של המטריצה $\alpha I + A^*A$ הינם $\alpha + \mu_j^2$, $j = 1, \dots, n$ ולכן $cond_2(\alpha I + A^*A) = \frac{\alpha + \mu_1^2}{\alpha + \mu_n^2} \leq \frac{2\mu_1^2}{\alpha}$, $0 < \alpha < \mu_1^2$

יציבות המערכת תלויה בפרמטר הריגורליזציה α שיהיה גדול מספיק. מצד שני על מנת להחזיק את השארית של המערכת

$\alpha x_\alpha + A^*Ax_\alpha = A^*y$ קרובה ככל האפשר למערכת מהקורית $Ax = y$ אנו מצפים מ α להיות קטן.

Tikhonov Regularization מקרבת את הפתרון $x = A^\dagger y$ של הבעיה $Ax = y$ על ידי הפתרון x_α

של מערכת משוואות ליניארית רגולרית $\alpha x_\alpha + A^*Ax_\alpha = A^*y^\delta$, אזי עבור ההשגיאה הכללית

$$x_\alpha - x = (\alpha I + A^*A)^{-1}A^*(y^\delta - y) + (\alpha I + A^*A)^{-1}A^*y - A^\dagger y$$

ולפי אי-שוויון המשולש נקבל כי

$$\|x_\alpha - x\|_2 \leq \|(\alpha I + A^*A)^{-1}A^*\|_2 \delta + \|(\alpha I + A^*A)^{-1}A^*y - A^\dagger y\|_2$$

כך ניתן לראות כי השגיאה הכללית תלויה ובנויה משני משתנים $E_{total} \leq E_{data} + E_{approx}$.

נעריך את השגיאה של הנתונים

$$E_{data} = \|(\alpha I + A^*A)^{-1}A^*\|_2 \delta > \frac{\mu_r}{\alpha + \mu_r} \delta$$

זה משקף את ההשפעה של נתונים שגואים, עבור δ מתוקן, ונהיה גדול יותר ככול ש $\alpha \rightarrow 0$, אם הערך הסינגולרי החיובי הקטן ביותר μ_r קרוב לאפס. התנאי השני

$$E_{approx} = \|(\alpha I + A^*A)^{-1}A^*y - A^\dagger y\|_2$$

מתאר את טעות הקירוב מכיוון שהחלפנו את $Ax = y$ במערכת רגולרית $\alpha x_\alpha + A^*Ax_\alpha = A^*y^\delta$ וממסקנה 3.1 ניתן לראות כי זה שואף לאפס ככל ש $\alpha \rightarrow 0$.

מצד אחד אם נסתכל על

$$cond_2(\alpha I + A^*A) = \frac{\alpha + \mu_1^2}{\alpha + \mu_n^2} \leq \frac{2\mu_1^2}{\alpha} , \quad 0 < \alpha < \mu_1^2$$

אזי היציבות של המערכת דורשת פרמטר ריגורליזציה α גדול כדי להשאיר את E_{data} קטן וכך להשאיר את ההשפעה של טעות הנתונים $\|y^\delta - y\|_2$ קטנה. מצד שני על מנת להשאיר את E_{approx} קטן צריך לדרוש שהפרמטר α יהיה קטן.

כמובן שבחירת הפרמטר α תצטרך להתבצע על ידי התפשרויות בין רמת הדיוק לבין רמת היציבות.

דרך אופטימאלית להשגת מתנה זו נתונה על ידי *Discrepancy principle*.

משפט 3.3: נניח כי $Ax = y$ פתירה. תהי A מטריצה $m \times n$ ו $y^\delta \in \mathbb{C}^m, y \in A(\mathbb{R}^n)$ המקיימים $\|y^\delta - y\|_2 \leq \delta < \|y^\delta\|_2$ עבור $\delta > 0$, אזי קיים $\alpha = \alpha(\delta) > 0$ אחד ויחיד כך שהפתרון x_α של המערכת $\alpha x_\alpha + A^*Ax_\alpha = A^*y^\delta$ מקיים $\|Ax_\alpha - y^\delta\|_2 = \delta$ *Discrepancy principle*. *Tikhonov Regularization* הינו רגולרי במובן בו אם הרמה של השגיאה δ נוטה לאפס אזי $x_\alpha \rightarrow A^\dagger y, \delta \rightarrow 0$.

הוכחה: אנו צריכים להראות שלפונקציה $F: (0, \infty) \rightarrow \mathbb{R}$ המוגדרת על ידי $F(\alpha) := \|Ax_\alpha - y^\delta\|_2^2 - \delta^2$

יש אפס יחודי. במובנים של מערכת סינגולרית אנו יודעים כי $x_\alpha = \sum_{j=1}^r \frac{\mu_j}{\alpha + \mu_j^2} (y, v_j) u_j$ כך שנקבל

$$F(\alpha) = \sum_{j=1}^m \frac{\alpha^2}{(\alpha + \mu_j^2)^2} |(y^\delta, v_j)|^2 - \delta^2$$

על כן F הינה מונוטונית עולה בעלת הגבול $F(\alpha) \rightarrow -\delta^2, \alpha \rightarrow 0$ ו $F(\alpha) \rightarrow \|y^\delta\|_2^2 - \delta^2, \alpha \rightarrow \infty$. נשים לב כי התנאי $\|y^\delta - y\|_2 \leq \delta < \|y^\delta\|_2$ מרמז על כך $y \neq 0$. לכן בעזרת $\alpha x_\alpha + A^*Ax_\alpha = A^*y^\delta$ ונקבל $\|Ax_\alpha - y^\delta\|_2 = \delta$

$$\|y^\delta\|_2 - \delta = \|y^\delta\|_2 - \|Ax_\alpha - y^\delta\|_2 \leq \|Ax_\alpha\|_2$$

$$\alpha \|Ax_\alpha\|_2 = \|AA^*(y^\delta - Ax_\alpha)\|_2 \leq \|AA^*\|_2 \delta$$

שילוב של שתי המשוואות הנ"ל ובתוספת $\|y^\delta\|_2 \geq \|y\|_2 - \delta$ נקבל כי

$$\alpha \leq \frac{\|AA^*\|_2 \delta}{\|y\|_2 - 2\delta}$$

מכאן ניתן להבין כי $\alpha \rightarrow 0$ וגם $\delta \rightarrow 0$. התנאי $x_\alpha \rightarrow A^\dagger y, \delta \rightarrow 0$ נובע מכך שמבצעים הצגה מחודשת כך

$$A^\dagger y = \sum_{j=1}^r \frac{1}{\mu_j} (y^\delta, v_j) u_j, \quad y^\delta \in \mathbb{R}^m$$

וכך

$$x_\alpha = \sum_{j=1}^r \frac{\mu_j}{\alpha + \mu_j^2} (y^\delta, v_j) u_j$$

וכמו כן גם מהעובדה כי $\delta \rightarrow 0, \|y^\delta - y\|_2 \rightarrow 0$.

מש"ל.

מנקודת מבט פראקטית אנו לא נבחר α שיתאים במדויק עבור $\|Ax_\alpha - y^\delta\|_2 = \delta$ אלא נבצע תהליך פשוט יותר, נבחר גודל מסויים ל α ואז ננחלק אותה כל פעם בגודל קבוע y עד ש $F(\alpha)$ יהפוך להיות שלילי.

ישום התאוריה:

נתונה פונקציה $f(x)$ על קטע $[a, b]$. נרצה לקרב את פונקציה זו בעזרת פולינום בתת מרחב סופי של L^2 בעל n מימדים כאשר הבסיס של תת מרחב זה הן הפונקציות $1, x, x^2, \dots, x^n$

אזי נרצה כי

$$\int_a^b (f(x) - a_0 - a_1x - \dots - a_nx^n)^2 dx \rightarrow \min$$

במילים אחרות נרצה את קרוב הטוב ביותר לפונקציה הנתונה, כפי שהוכח במשפט 1.1 .

נסביר את המכפלה הפנימית במרחב L^2 :

עבור פונקציות $f, g \in L^2(x)$, מכפלה פנימית הינה

$$(f, g) = \int_a^b f(x)g(x)dx$$

אזי ממשפט 1.1 המוכיח את קיום הקרוב הטוב ביותר בהשלכה על תת מרחב סופי של L^2 מתקבל כי הוא מהצורה:

$$\sum_{k=1}^n a_k x^k$$

וזאת רק אם מתקיים עבור $a_1, \dots, a_n$

$$\sum_{k=1}^n a_k (x^k, x^j) = (f(x), x^j), j = 0, \dots, n$$

כלומר:

$$\sum_{k=0}^n \int_a^b x^k x^j a_k dx = \int_a^b f(x) x^j dx, \quad j = 0, \dots, n$$

וכך מקבלים את מטריצה A ואת y .

כאשר $[a, b] = [0, 1]$ כל אחד מערכי מטריצה A יהיה מהצורה:

$$a_{i,j} = \frac{1}{i+j+1}$$

כך שמטריצה A תיראה כך כאשר $n = 8$

$$A = \begin{pmatrix} 1.000 & 0.5000 & 0.3333 & 0.2500 & 0.2000 & 0.1667 & 0.1429 & 0.1250 \\ 0.5000 & 0.3333 & 0.2500 & 0.2000 & 0.1667 & 0.1429 & 0.1250 & 0.1111 \\ 0.3333 & 0.2500 & 0.2000 & 0.1667 & 0.1429 & 0.1250 & 0.1111 & 0.1000 \\ 0.2500 & 0.2000 & 0.1667 & 0.1429 & 0.1250 & 0.1111 & 0.1000 & 0.9090 \\ 0.2000 & 0.1667 & 0.1429 & 0.1250 & 0.1111 & 0.1000 & 0.9090 & 0.0833 \\ 0.1667 & 0.1429 & 0.1250 & 0.1111 & 0.1000 & 0.9090 & 0.0833 & 0.0769 \\ 0.1429 & 0.1250 & 0.1111 & 0.1000 & 0.9090 & 0.0833 & 0.0769 & 0.0714 \\ 0.1250 & 0.1111 & 0.1000 & 0.0909 & 0.0833 & 0.0769 & 0.0714 & 0.0667 \end{pmatrix}$$

לאחר שבנינו את מטריצה A נבדוק את היציבות שלה על ידי חישוב ה $condition number$ שלה:
המטריצה הינה הרמטית לכן נחשב לפי הנוסחה :

$$cond_2(A) = \frac{|\lambda_{max}|}{|\lambda_{min}|} = \frac{|1.6959|}{|1.111 * 10^{-10}|} = 1.5257 * 10^{10}$$

כפי שרואים המספר הוא ענק לכן אנו מבינים כי מערכת זו הינה מערכת ill -conditioned.
נתבונן בהשפעה של n על $condition number$ של A :

n	3	4	5	6	7	8
$cond(A)$	524.0568	$1.5 * 10^4$	$4.7 * 10^5$	$1.5 * 10^7$	$4.8 * 10^8$	$1.5 * 10^{10}$

אנו יכולים להבין כי פתרון בשיטה ברגילה של היפוך מטריצות לא יעבוד המקרה זה מיכוון שכל שינוי הקטן ביותר בצד ימין של המשוואה יוביל לשינוי ענק בפתרון.

על כן נשתמש בגולריזציה של טיחונוב על מנת לפתור את בעיה זו!

טיחונוב מציע כי נכפול את הגורם $\frac{1}{\mu_j}$ בגורם הנקרא $damping factor$

$$\frac{\mu_j^2}{\alpha + \mu_j^2}$$

כך שהיתקבל הפתרון מהצורה:

$$x_\alpha = \sum_{j=1}^r \frac{\mu_j}{\alpha + \mu_j^2} (y, v_j) u_j$$

כעת כל שנשאר לעשות הוא לחשב את α .

במשפט 3.3 הוכח כי הדרך הטובה ביותר לחשב את α היא בעזרת $discrepancy principle$ ומהוכחת המשפט בנינו את הפונקציה

$$F(\alpha) = \sum_{j=1}^m \frac{\alpha^2}{(\alpha + \mu_j^2)^2} |(y^\delta, v_j)|^2 - \delta^2$$

מבחינה תאורטית α צריכה לעמוד בתנאי $\|Ax_\alpha - y^\delta\|_2 = \delta$. בצורה הישומית נמצא אותה בתהליך

פשוט יותר אשר בו אנו בוחרים α ראנדומלית ואז מחלקים אותה בקבוע γ לבחירתנו עד שהפונקציה $F(\alpha)$ לא הופכת להיות שלילית.

הפונקציה המקורבת היא $f(x) = e^x$.

רמת הדיוק של e בתוכנת המטלב ניתנת לשינוי על ידי פרמטר m

בדוגמה זו $m=3$.

ווקטור y יהיה :

$$y = \begin{pmatrix} 1.7182 \\ 0.9999 \\ 0.7182 \\ 0.5634 \\ 0.4645 \\ 0.3956 \\ 0.3447 \\ 0.3055 \end{pmatrix}$$

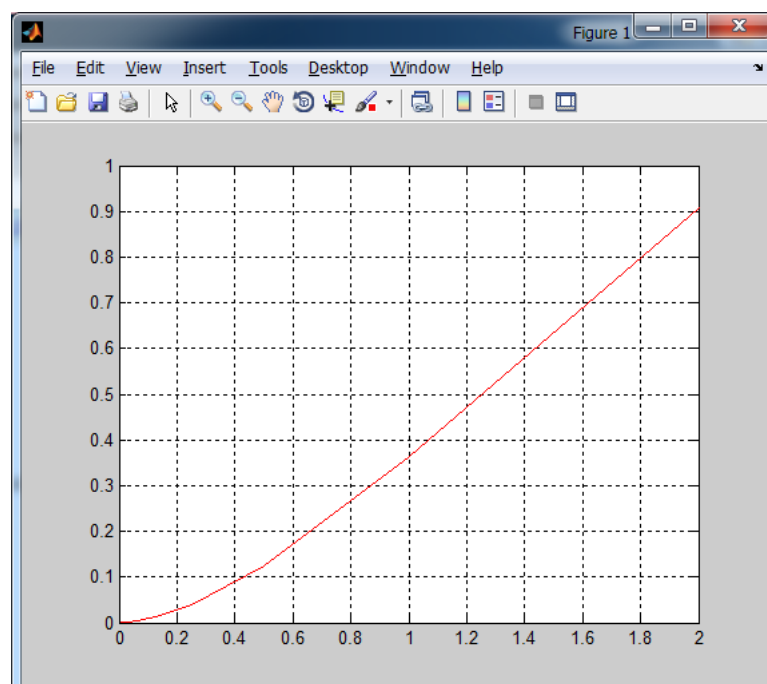
את וקטור ה y "זיהמנו" בשגאות על מנת לקבל תוצאות הדומות יותר למציאות היום יומית על ידי הוספה של מספר ראנדומאלי δ בתחום $[-0.0001, 0.0001]$. לכן הוא יראה כך:

$$y^\delta = \begin{pmatrix} 1.7183 \\ 0.9999 \\ 0.7181 \\ 0.5634 \\ 0.4645 \\ 0.3955 \\ 0.3446 \\ 0.3054 \end{pmatrix}$$

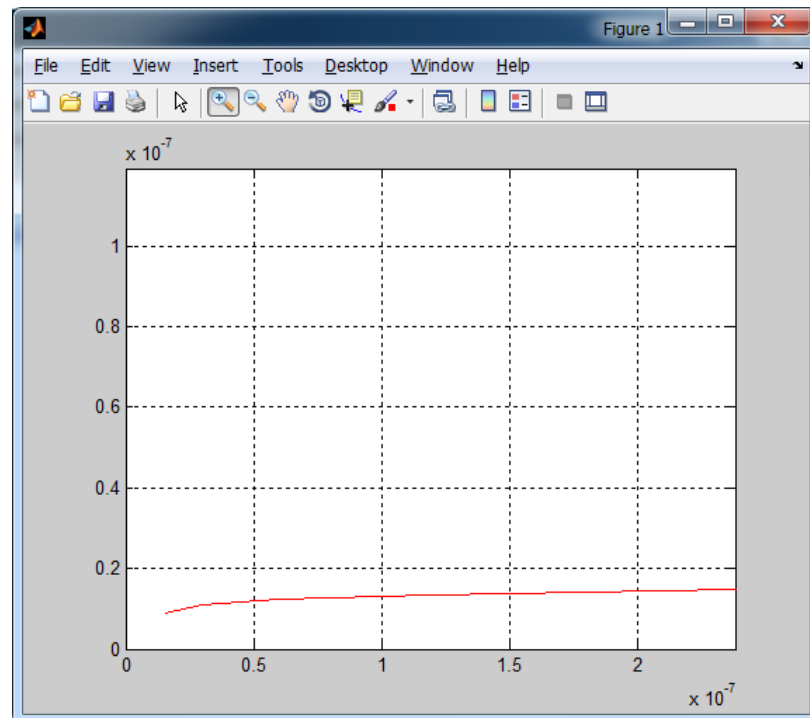
בדוגמה שלי בחרתי ב $\alpha = 2$ וב $\gamma = 0.5$.

כמו כן מכיוון ותוכנת המטלב עושה חישובים מדוייקים הגעה למצב שבו הפונקציה $F(\alpha) < 0$ איננו אפשרי ולא נותנת פתרונות הגיוניים על כן הוסכם להגדיר את ה "אפס המוחלט" על 10^{-8} .

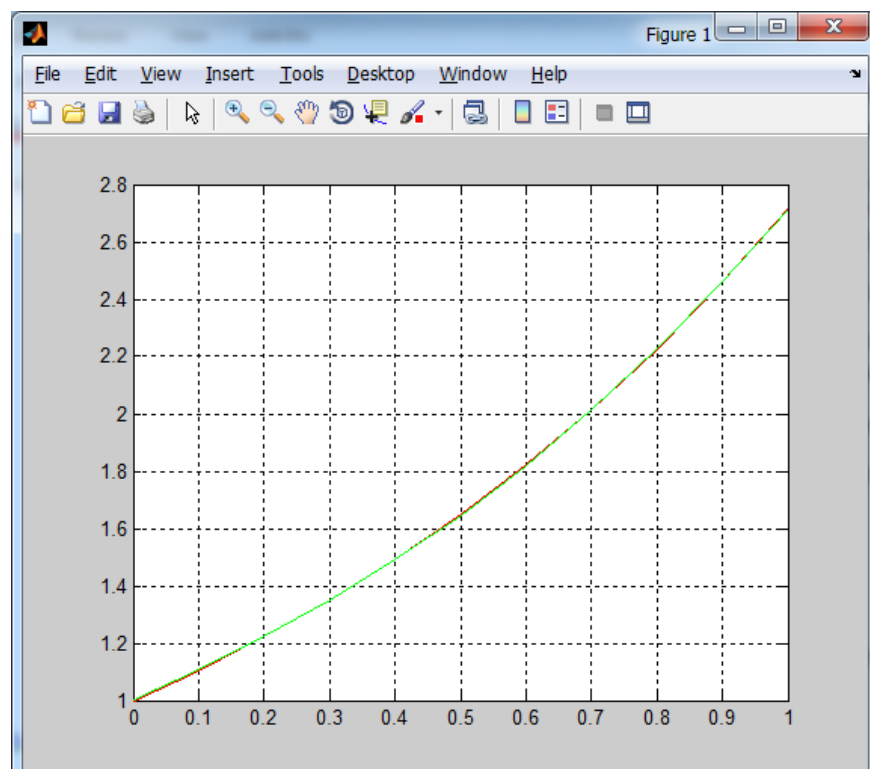
נציג גרף המראה את התלות בין $F(\alpha)$ ל α :



כאשר נתקרב לנקודת האפס נראה כי אנו לא באמת מגיעים למצב שבו $F(\alpha) < 0$ וזאת מכיוון שכאשר נבקש ממטלב לחשב את המספר מדוייק, הקרוב שינתן על ידו לאחר מכן יהיה גרוע ביחס לקירוב שינתן אם נגיד את נקודת ה"אפס המוחלט" כ 10^{-8} .

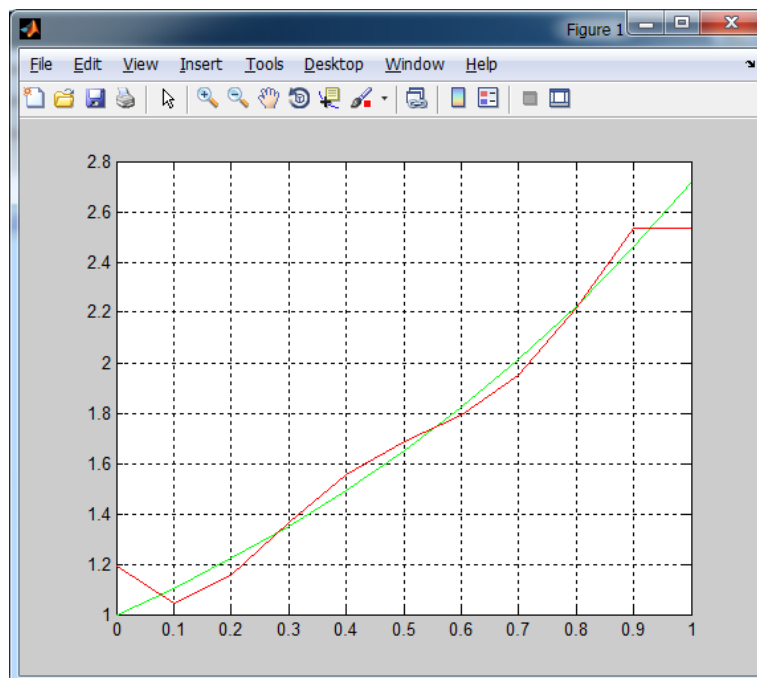


אם כך מצאנו את ה α שלנו והיא שווה ל $1.1921 * 10^{-7}$.
לאחר ההצבה נשווה את פונקציית המטרה שלנו לפתרון שהתקבל בצורה גרפית:

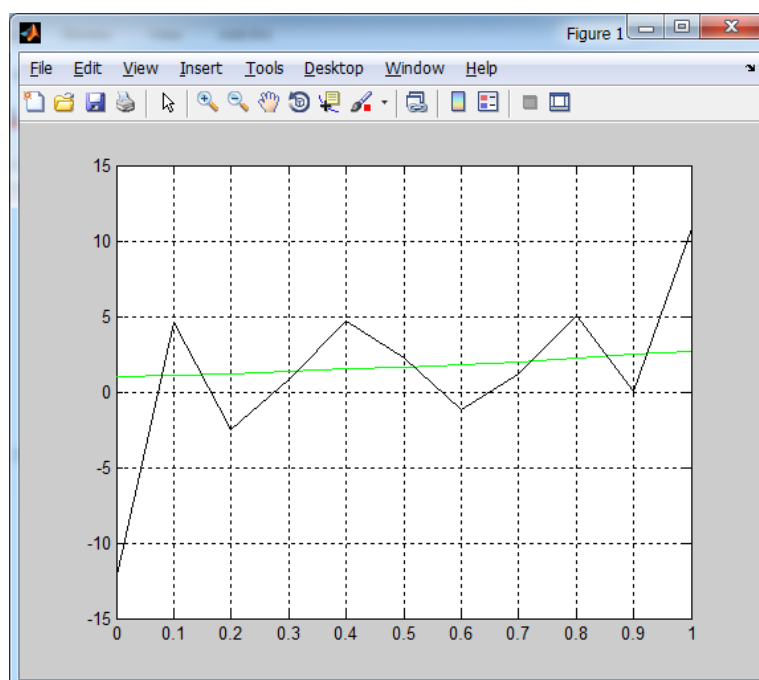


כאשר הגרף הירוק הוא הפונקציה המקורבת שהיא $f(x) = e^x$ והגרף האדום הוא הפתרון בעזרת הרגולריזציה של טיחונוב.

כאשר לא נגדיר את נקודת ה"אפס המוחלט" כ 10^{-8} תוצאת הקרוב תיראה כך:



לשם ההשוואה נראה פתרון בצורה גרפית של פתרון המערכת בצורה הישרה של היפוך מטריצות :



כאשר הגרף השחור הוא הפתרון בצורה של היפוך מטריצות והגרף הירוק הוא פונקציית המטרה.

```

function Sol3= ex1(n,m)
    for i=1:n+1
        for j=1:n+1
            A(i,j)=1/(i+j-1);
        end
    end

    cond(A)
    expo= round (exp(1)*10^m)/10^m;
    for k=1:n+1
        fun=@(x)expo.^x.*(x.^(k-1));
        y(k)=integral(fun,0,1);
    end

    DError= 0.0001;
    for i=1:length(y)
        YError(i)=y(i)-DError+(DError-(-DError)).*rand(1,1);
    end

    Sol1=inv(A)*YError';

    %%%%%%%%% SVD solution%%%
    [u,d,v]=svd(A);
    sing=svd(A);
    z=u'*YError';
    w=pinv(d)*z;
    Sol2=v*w;

    %%%%%%%%% Tikhonov regularizayion %%%%
    alfa = 2;
    gama=0.5;
    Falfa=0;
    for i=1:length(y)
        Falfa= Falfa + (((alfa^2)/((alfa+sing(i)^2)^2))* (abs(dot(YError,v(:,i)))^2 ));
    end
    Falfa= Falfa - DError^2;
    while Falfa>10^-8
        alfa=alfa * gama;
        Falfa=0;
        for i=1:length(y)
            Falfa= Falfa + (((alfa^2)/((alfa+sing(i)^2)^2))* (abs(dot(YError,v(:,i)))^2 ));
        end

        Falfa= Falfa - DError^2;
    end
    alfa;

```

```

Falfa;
for i=1:length(sing)
    for j=1:length(sing)
        if i==j
            dalfa(i,j)= d(i,j)/(alfa +d(i,j)^2);
        end
    end
end
z1=u'*YError';
w1=dalfa*z1;
Sol3=v*w1;

x=0:0.1:1;
for i=1:length(x)
    s1(i)=0;
    for j=1:n+1
        s1(i)=s1(i)+Sol1(j)*x(i)^(j-1);
    end
end
for i=1:length(x)
    s2(i)=0;
    for j=1:n+1
        s2(i)=s2(i)+Sol2(j)*x(i)^(j-1);
    end
end
for i=1:length(x)
    s3(i)=0;
    for j=1:n+1
        s3(i)=s3(i)+Sol3(j)*x(i)^(j-1);
    end
end
plot(x,exp(x),'-r',x,s3,'-g',x,s2,'--b',x,s1,'-k')
grid on
end

```

1. Phillips, D. L., "A Technique for the Numerical Solution of Certain Integral Equations of the First Kind". Journal of the ACM. Vol. 9, No. 84, 1962
2. A. N. Tikhonov and V. Y. Arsenin, Solution of Ill-posed Problems, ser. Scripta series in mathematics. Washington, D.C.: V. H. Winston & Sons, 1977.
3. R. Kress, Numerical Analysis, ser. Graduate texts in mathematics, Vol 181, Springer, 1991