



**בחינת אותות ומערכות - מועד ב תשע"ו 12.09.16**

**גרסה מספר: 582442784**

שם הסטודנט \_\_\_\_\_

ת.ז. \_\_\_\_\_

סטודנט/ית יקר/ה,

המבחן כולל: 20 שאלות -16 סגורות, 4 פתוחות.

אנא קרא/י בתשומת לב את ההוראות:

- הדבק/י את מדבקת הברקוד במקום המיועד לכך. (במידה ואין בידך מדבקה אנא רשום את מספר ת.ז. במקום המיועד)
- יש למלא את טופס התשובות בעט כדורי בלבד על ידי סימון  ברור!
- יש לסמן תשובה אחת בלבד, סימון שתי תשובות כנכונות תחשב כשגיאה גם אם אחת מהן נכונה.
- מצורף דף נוסחאות
- חומר עזר מותר: מחשבון
- את התשובות לשאלות הפתוחות יש לרשום על גבי טופס המבחן, ולא במחברת חיצונית. ניתן להשתמש בשאר חלקי הטופס בתור דפי טיוטה.

**משך הבחינה: שלוש שעות**

**ב ה צ ל ח ה !!!**

אני (שם ומשפחה) \_\_\_\_\_ ת.ז. \_\_\_\_\_ נבחן/ת בבחינה זו, מצהיר/ה כי לא אעתיק

ולא אשתמש באמצעים המנוגדים לטוהר הבחינות בטכניון.

## אותות ומערכות מועד ב 12.9.16

1. נתונה המערכת

$$\dot{x} = 2x - 4$$

$$x(0) = 5$$

מהו הפתרון האנליטי של המערכת?

$x(t) = 2 + 2e^{2t}$  .א

$x(t) = 2 + 3e^{2t}$  .ב

$x(t) = 4 + 1e^{2t}$  .ג

$x(t) = 4e^{2t}$  .ד

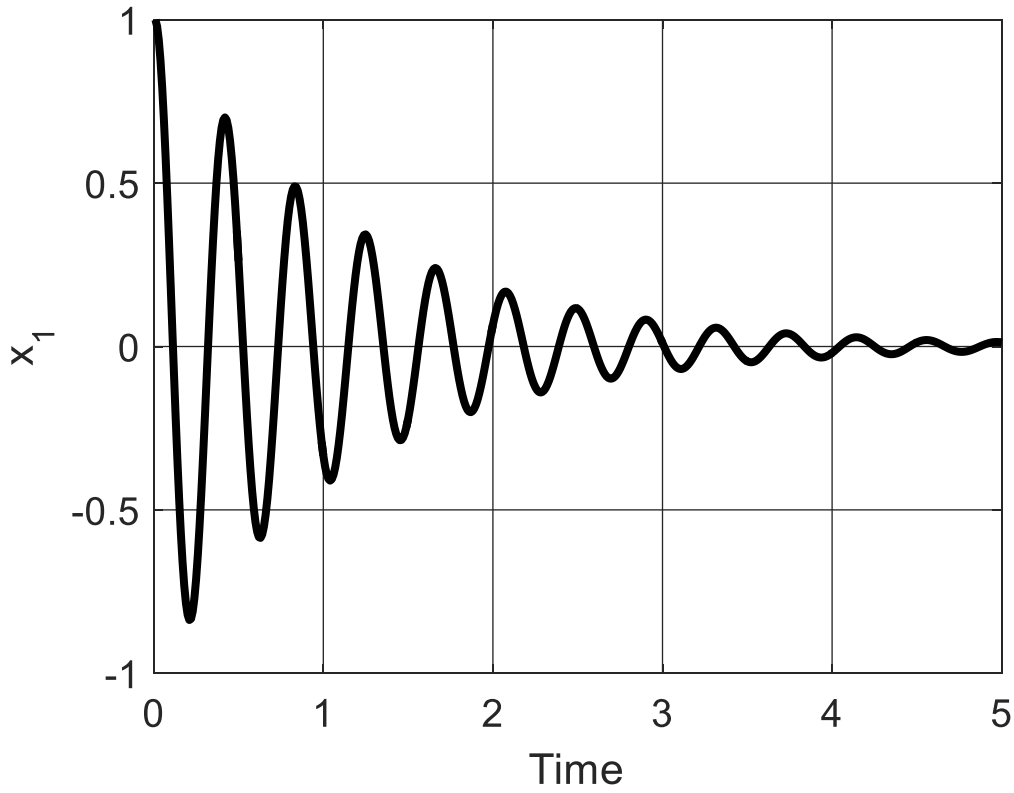
$x(t) = 5e^{2t}$  .ה

2. נתונה המערכת הבאה:

$$\dot{x} = Ax$$

$$x(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

כמו כן, נתון הגרף של  $x_1$  כפונקציה של זמן:



מיצאו את המטריצה A :

$$\begin{pmatrix} 15 & 2 \\ 2 & 15 \end{pmatrix}$$

א.

$$\begin{pmatrix} -15 & 2 \\ -2 & -15 \end{pmatrix}$$

ב.

$$\begin{pmatrix} 2 & 15 \\ 15 & 2 \end{pmatrix}$$

ג.

$$\begin{pmatrix} -2 & 15 \\ -15 & -2 \end{pmatrix}$$

ד.

$$\begin{pmatrix} 15 & 15 \\ -2 & 2 \end{pmatrix}$$

ה.

3. נתונה המשוואה הדיפרנציאלית הבאה:

$$\frac{dx}{dt} = 0.25x(4 - 4x)$$

למשוואה יש שתי נקודות שבת, כאשר אחת מהן יציבה. נסמן נקודת שבת זו ב- $x_0$

איזו משוואה מתארת את הקרוב הלינארי לדינמיקה ליד נקודה  $x_0$  כלומר עבור  $x = x_0 + z$  ?

א.  $\frac{dz}{dt} = -4z$

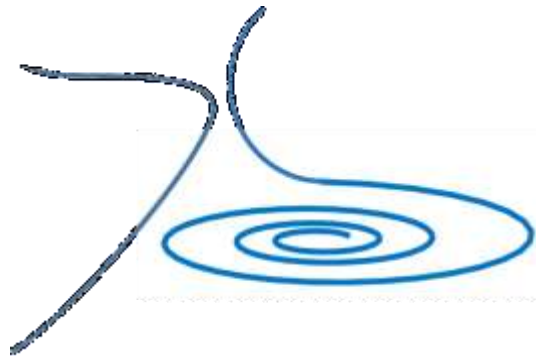
ב.  $\frac{dz}{dt} = 4z$

ג.  $\frac{dz}{dt} = -z$

ד.  $\frac{dz}{dt} = z$

ה.  $\frac{dz}{dt} = -0.25z$

4. נתבונן בשני מסלולים הבאים במרחב הפאזה של מערכת דינמית דו-מימדית.



ניתן לומר כי:

- א. זהו מצב לא חוקי - לא ייתכן שמסלול אחד הוא ספירלה והשני לא
- ב. זהו מצב אפשרי במערכת ליניארית שיש לה ערך עצמי מרוכב וגם ערך עצמי ממשי
- ג. במערכת הדינאמית הזאת כל המסלולים מתכנסים לנקודה אחת
- ד. המערכת הדינאמית שלפנינו בהכרח לא ליניארית
- ה. זהו מצב לא חוקי - לא ייתכן ששני מסלולים במרחב הפאזה יתקרבו ואז שוב יתרחקו

5. ניתן לראות כי הפתרונות של המערכת מסדר שני הנתונה ע"י:

$$\ddot{x} + 4\dot{x} + 5x = 0$$

הם תנודתיים, כאשר האמפליטודה (משרעת, מרחק בין מינימום למקסימום) של התנודות יורדת עם הזמן.

מהו זמן המחזור  $T$  של התנודות?

- א.  $2\pi$
- ב. 1
- ג.  $\pi/2$
- ד.  $2\pi/\sqrt{5}$
- ה.  $\sqrt{5}$

6. חוקר סרטן גילה כי לאחר מתן תרופה מסוימת, קצב שינוי גודלו של גידול סרטני תלוי בגודל של אותו גידול באופן הבא:

$$\frac{dx}{dt} = x(x-1)(x-3)$$

כאשר  $x$  מציין את גודל הגידול בסמ"ק.

איזו מסקנה נובעת מקשר זה:

- א. אם התרופה תוזרק כאשר הגידול בגודל של 0.5 סמ"ק, הגידול יתכווץ ויעלם
- ב. אם התרופה תוזרק כאשר הגידול בגודל של 0.5 סמ"ק, הגידול ישתנה קצת, וייתיצב על גודל של בערך 0.5 סמ"ק
- ג. אם התרופה תוזרק כאשר הגידול בגודל של 1.5 סמ"ק, הגידול יגדל וייתיצב על גודל של 3 סמ"ק
- ד. אם התרופה תוזרק כאשר הגידול בגודל של 4 סמ"ק, הגידול יתכווץ וייתיצב על גודל של בערך 2 סמ"ק
- ה. אם התרופה תוזרק כאשר הגידול בגודל של 4 סמ"ק, הגידול ימשיך לגדול

7. נתונה משוואה דיפרנציאלית ליניארית  $\dot{x} = x$  עם תנאי התחלה  $x(0) = 1$ :

נסמן פתרון נומרי עם  $\Delta t = 1$  ע"י  $x_{num1}(t)$ ,

נסמן פתרון נומרי עם  $\Delta t = 0.25$  ע"י  $x_{num2}(t)$

מה ניתן לומר על הקשר בין שני הפתרונות:

א.  $x_{num1}(1) < x_{num2}(1)$

ב.  $x_{num1}(1) > x_{num2}(1)$

ג.  $x_{num1}(1) = x_{num2}(1)$

ד.  $x_{num1}(4) = x_{num2}(1)$

ה.  $x_{num1}(1) = x_{num2}(4)$

\* אם  $\Delta t$  קטן יותר, הגדל יגדל.

כי ה"ו עולה כי  $x_{num}(1)$

כי ה"ו עולה כי  $\Delta t$  קטן יותר, ולכן

$t = 1$

8. תרופה כימותרפית מוזרמת דרך הדם לאיבר מסויים.

נסמן ב

$N$  את כמות התאים הסרטניים

$C$  את ריכוז התרופה באיבר

נתון כי  $a, b > 0$

בשאלה זו נתעלם מהדינמיקה של ריכוז התרופה, ונתמקד בכמות התאים הסרטניים. ידוע כי תאים אלו מתרבים בקצב מסוים ללא תרופה, מינון נמוך של התרופה מוריד את קצב ההתרבות, ושיעילות התרופה יורדת עבור ריכוזים גבוהים של תרופה.

מהי המשוואה המתארת את הדינמיקה של מספר התאים?

א.  $\frac{dN}{dt} = -aC + bN$

ב.  $\frac{dN}{dt} = -C(a - C)N + bN$

ג.  $\frac{dN}{dt} = N(a - N) - C$

ד.  $\frac{dN}{dt} = a(b - C)N$

ה.  $\frac{dN}{dt} = -(a - C) + bN$

9. חוקר חילק את מאגר המידע שבידו לקבוצת אימון ולקבוצת בוחן והחליט לבנות את המודל המתאר את המידע ע"י רגרסיית לאסו. במהלך הקורס למדנו כי פונקציית המטרה של רגרסיית

לאסו היא  $\min imize(\text{error} + \lambda \cdot \sum_{j=0}^p |a_j|)$  כאשר  $P$  הוא מימד הדגימות.

מבין המשפטים הבאים, מהו המשפט הנכון ביותר?

- א. ככל ש  $\lambda$ -יהיה יותר גבוה, יהיו יותר מקדמים מאופסים, שגיאת האימון תעלה, אך לא בהכרח שגיאת הבוחן תרד לאורך כל ההעלאה של  $\lambda$
- ב. ככל ש  $\lambda$ -יהיה יותר גבוה, יהיו יותר מקדמים מאופסים, שגיאת האימון תעלה, ובהכרח שגיאת הבוחן תרד לאורך כל ההעלאה של  $\lambda$
- ג. ככל ש  $\lambda$ -יהיה יותר גבוה, יהיו יותר מקדמים מאופסים, שגיאת האימון תרד, אך לא בהכרח שגיאת הבוחן תרד לאורך כל ההעלאה של  $\lambda$
- ד. ככל ש  $\lambda$ -יהיה יותר נמוך, יהיו פחות מקדמים מאופסים, שגיאת האימון תרד, ובהכרח שגיאת הבוחן תרד לאורך כל ההורדה של  $\lambda$

10. איזו מהשיטות הבאות מחייבת שבנוסף לסט הכניסות יהיה גם סט יציאות אשר מכיל את היציאה המתאימה לכל כניסה?

- א. PCA
- ב. Lasso regression
- ג. K-means
- ד. כל השיטות הנ"ל הינן שיטות להורדת מימדיות הכניסה, ולכן באף אחת מהן אין חובה לקבל בנוסף את סט היציאות

11. איזה מהמשפטים הבאים נכון עבור שיטות לתיקון P-values בעקבות השוואות מרובות?

- א. הסיכוי לטעות מסוג 1 נעשה קטן יותר
- ב. מבחנים שהיו סיגניפיקנטים קודם עלולים לא להיות סיגניפיקנטים לאחר התיקון
- ג. מבחנים שהיו לא סיגניפיקנטים קודם יכולים להיות סיגניפיקנטים לאחר התיקון
- ד. התיקון הופך את הדוגמאות לבלתי תלויות זו בזו

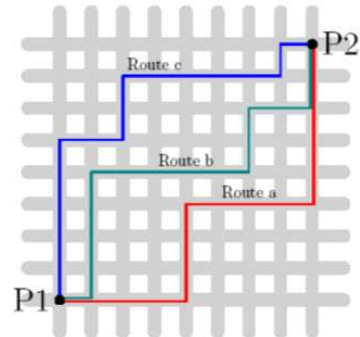
12. מהו המשפט הנכון ביותר המתאר את שיטות ה-clustering?

- א. גם בשיטת K-means וגם בשיטת hierarchical clustering המשתמש לא צריך לקבוע מראש את מספר ה-clusters
- ב. גם בשיטת K-means וגם בשיטת hierarchical clustering ניתן לייצר דנדוגרמה
- ג. בהנחה שביצענו clustering על אותו ה-data בשיטת K-means ובשיטת hierarchical clustering והתקבלו אותם מספר clusters - ה-clusters יהיו זהים
- ד. בשיטת k-means יש אלמנט רנדומלי באלגוריתם שיכול להשפיע על התוצאות הסופיות



13. מרחק מנהטן הוא מרחק הנמדד על פני רשת (בדומה לרחובות מנהטן בהן ניתן ללכת רק על פני הרשת, ראה ציור להמחשה).

איזה מבין המשפטים הבאים נכון?



- א. מרחק אוקלידי בין שתי נקודות תמיד יהיה קטן ממש ממרחק מנהטן בין שתי נקודות
- ב. גם במרחק מנהטן וגם במרחק אוקלידי יש יותר ממסלול אחד שהוא המסלול הקצר ביותר בין שתי נקודות
- ג. קיים מקרה שבו מרחק אוקלידי ומרחק מנהטן בין שתי נקודות יהיו שווים
- ד. רק מרחק אוקלידי יכול להיות מחושב עבור מימדים מרובים

14. חוקר ביצע 20 מדידות של microarray gene expression של 10 אנשים חולים ו-10 של אנשים בריאים והוא רוצה למצוא גנים השונים בין החולים לבריאים.

איזו אנליזה היא המתאימה ביותר למקרה זה ?

- א. חישוב T test עבור כל גן כדי למצוא האם הדוגמאות של החולים שונות מהדוגמאות של הבריאים, ולאחר מכן ביצוע תיקון להשוואות מרובות לכל הגנים
- ב. ביצוע PCA על החולים בנפרד ועל הבריאים בנפרד, ומציאת הגנים החופפים בין התוצאות
- ג. ביצוע clustering על החולים והבריאים ביחד ומציאת הגנים שנמצאים ב-cluster אחד ושונים בין הבריאים לחולים
- ד. רגרסיית לאסו

15. ביצעת clustering פעם עם centroid linkage ופעם עם average linkage. בשלב מסוים בדנדרוגרמה שנבנתה בעזרת centroid linkage הקלאסטר המורכב מדוגמאות {1,2} חובר לקלאסטר המורכב מדוגמאות {3,4}. באופן דומה, גם בדנדרוגרמה שנבנתה בעזרת average linkage אותם קלאסטרים חוברו בנקודה כלשהי.

מבין המשפטים הבאים, מהו המשפט הנכון?

- א. החבירות בין שני הקלאסטרים יתבצעו בשתי הדנדרוגרמות במיקום זהה
- ב. החבירות בין שני הקלאסטרים יתבצעו בגובה נמוך או שווה בדנדרוגרמה של ה-centroid linkage מאשר ב-average linkage
- ג. החבירות בין שני הקלאסטרים יתבצעו בגובה נמוך או שווה בדנדרוגרמה של ה-average linkage מאשר ב-centroid linkage
- ד. אין מספיק מידע בשביל לדעת על ההבדלים בחבירות בין שתי שיטות ה-linkage

16. נתונים נתוני gene expression של 20,000 גנים הנמדדו ב-20 חולים המשתייכים לשתי קבוצות: A ו-B. חוקר בדק שינויים ב-gene expression בין החולים המשתייכים לשתי הקבוצות. על כל גן התבצע מבחן t לבדיקת הבדלים ברמות הביטוי. החוקר קבע את הרף לדחיית השערת האפס (אלפה) להיות 0.01. בהתאם לרף שנבחר, החוקר גילה 8,000 גנים שנמצאו שונים באופן סיגניפיקנטי בין שתי הקבוצות, אך לא התחשב במספר הבדיקות שביצע.

כיצד ניתן לתקן שגיאה זו?

- א. פרמוטציה של דוגמאות בתוך הקבוצות והערכת FDR
- ב. החוקר השתמש ברף (אלפה) נמוך מדי, עליו להגדיל את אלפה עד שיקבל כמות סבירה של גנים שונים
- ג. החוקר השתמש ברף (אלפה) גבוה מדי, עליו לחלק את האלפא ב-8000
- ד. פרמוטציות של כל הדוגמאות והערכת FDR

## פתוחה 1

נתונה המערכת הבאה:

$$\begin{aligned} \dot{x} &= y - 2x \\ \dot{y} &= x^2 - 2x + y - 1 \end{aligned}$$

א. מיצאו את כל נקודות שיווי המשקל של המערכת.

ב. עבור כל נקודת שיווי משקל, בצעו לינאריזציה. רישמו את המטריצה המתקבלת, וסווגו את נקודת שיווי המשקל (צומת, אוסף, ספירלה, יציב/לא יציב, ...)

$$y = 2x$$

$$x^2 - 2x + 2x - 1 = 0$$

$$x^2 = 1 \Rightarrow x = \pm 1$$

$$A = (1, 2) \quad B = (-1, -2) \quad \text{נקודות יציב}$$

(X)

$$\begin{pmatrix} -2 & 1 \\ 2x-2 & 1 \end{pmatrix}$$

(I) יציב

$$A: \begin{pmatrix} -2 & 1 \\ 0 & 1 \end{pmatrix} \quad \tau = -1 \quad \Delta = -2 \quad \lambda = -2, 1 \quad \text{אוסף}$$

$$B: \begin{pmatrix} -2 & 1 \\ -4 & 1 \end{pmatrix} \quad \tau = -1 \quad \Delta = 2 \quad \lambda = \frac{-1 \pm i\sqrt{7}}{2} \quad \text{ספירלה יציב}$$

1 - נמצאה נק' A או B

1 - נמצאו שתיים, ולא נקודות אחרות

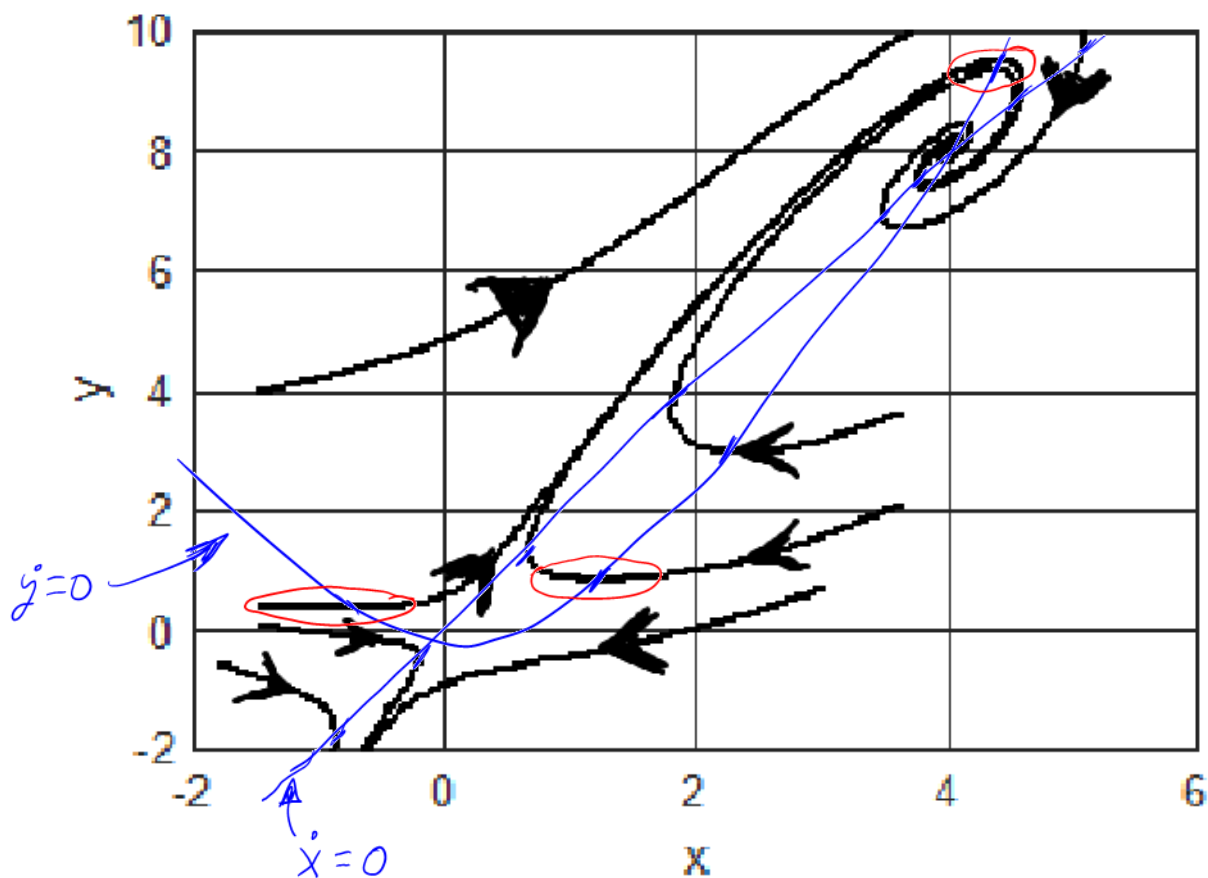
1 - מטריצה נכונה של A

$\left[ \begin{matrix} \text{מטריצה נכונה של A} \\ \text{קיימת מטריצה נכונה של A} \end{matrix} \right. \left. \begin{matrix} \text{אם } \frac{1}{2} \text{ אז צורב כללי} \\ \text{של } \begin{pmatrix} -2 & 1 \\ 2x-2 & 1 \end{pmatrix}$  \end{matrix} \right] \left. \vphantom{\left[ \begin{matrix} \text{מטריצה נכונה של A} \\ \text{קיימת מטריצה נכונה של A} \end{matrix} \right.}} \right\} \frac{1}{2}

$\frac{1}{2}$  - אוסף נגד B

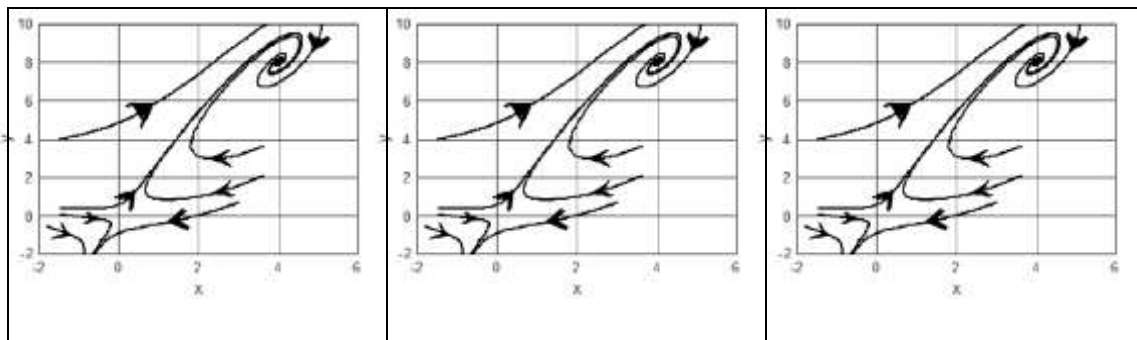
## פתוחה 2

לפניכם מספר מסלולים של מערכת דינמית דו מימדית.



עליכם לצייר על גרף זה את שני עקומי האפס של המערכת.

1. נתון כי כל עקום אפס (של  $x$  ושל  $y$ ) מורכב מעקום אחד בלבד.
2. על הגרף צריכים להופיע שני עקומים, כאשר ברור איזה שייך לא ואיזה שייך ל $y$ .
3. לרשותכם טיוטות של הגרף למטה – נא לכתוב תשובה סופית בלבד על הגרף הגדול.



1 -  $\dot{x}=0$  חוגק רמור 2 מסוליר אינכ"ז

1 -  $\dot{y}=0$  חוגק רמור 2 מסוליר אינכ"ז

1 - קי"ח חוגק ל"ז אונ"ז

1 - קי"ח חוגק ל"ז סטול"ז

1 - הרקוק ס"ז  $\dot{y}=0$  רמור באוס - האנכ"ז האנכ"ז

### פתוחה 3

נתונה טבלת המרחקים הבאה:

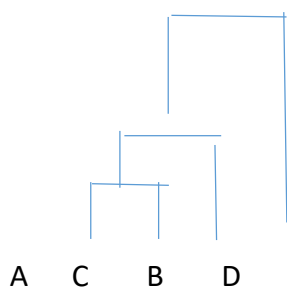
	D	C	B	A	
				0	A
			0	2.2	B
		0	2.2	2	C
	0	3.6	5.6	5.3	D

(א) בצע 2 צעדים ראשונים של clustering בשיטת single linkage וצייר את הצמתים שחוברו. נחבר את A ו-C ונחשב את המרחק ביניהם לשאר:

2 מקודות  
 $\frac{1}{2}$  לכל ערך נכון מחוץ לאלכסון.  
 $\frac{1}{2}$  על אפסים באלכסון

D	A-C	B	
		0	B
	0	2.2	A-C
0	3.6	5.6	D

הזוג הבא שנחבר הם C-A ו-D, ולכן הדנדוגרמה כרגע כך:



(ב) נניח שרוצים לבצע clustering בשיטת הצנטרואידים ובוחרים באקראי את ה-clusters ההתחלתיים. מה ההבדל בשלב הראשון של ה-clustering אם הקבוצות הן A-B / C-D או A-C/B-D לחילופין

סעיף מבוטל, 3 מקודות זכות

#### פתוחה 4

חוקרת קיבלה סט דגימות, כאשר כל דגימה היא של מדידת ביטוי של 20000 גנים בחולה סרטן. על מנת להתמודד עם רב הממדיות, החוקרת החליטה להסתכל על הנתונים על ידי הפעלת PCA. שרטוט PC1 מול PC2 הראה בברור הפרדת הדוגמאות ל-שתי תתי קבוצות, בעיקר על סמך הציר הראשון. לפי אבחנה זו החוקרת ביצעה מבחן-T להבדלים בין ביטוי גנים בין שתי תתי הקבוצות וזיהתה 5000 גנים שונים ב-FDR של 20%.

1. האם ל- KMEANS עם  $k=2$  בהכרח יוביל לחלוקה של הדוגמאות לאותם 2- קבוצות שהחוקרת זיהתה? נמק את תשובתך.
2. האם ביצוע PCA רק על 15000 הגנים שלא זהו כשונים ב-T-test יכול לגרום לירידה במספר תתי הקבוצות שצפויה החוקרת לזהות ב-PCA? נמק את תשובתך
3. תאר את טיב היחס בין הגנים שנמצאו שונים לבין המקדמים במטריצת ה-coefficients של הציר הראשון?

1  
לא. אין הבטחה לשיוך זהה. הן בגלל שאיננו יודעים את פריסת המקודות, והן בגלל האתחול האקראי שיחדיו יכולים להוביל לחלוקה שונה לגמרי.

1 נק

2  
כן. סביר מאוד להניח שהחסרת הגנים השונים תקטין את השונות בין הדוגמאות ואף לגרום להעלמות אותה חלוקה ראשונית שהחוקרת זיהתה.

2 נק

3  
כיוון שהציר הראשון הוא זה שלפיו נצפתה ההפרדה בין תתי הקבוצות שביניהם נמצאו ההבדלים, ניתן לצפות שלגנים אלו יש מקדמים גדולים (בערך מוחלט) במטריצה זו.

2 נק

סעיף א: 1) מקודה)  
1/2 - יש לפחות נימוק אחד נכון  
1/2 - אין נימוקים לא נכונים

סעיף ב) שתי מקודות)  
1/2 - תשובה נכונה ללא קשר לנימוק  
1 - יש לפחות נימוק אחד נכון.  
1/2 - אין נימוקים לא נכונים.

סעיף ג) שתי מקודות)  
מצויין שהמקדמים של הגנים שנמצאו יהיו גדולים

### נוסחאות

ערכים עצמיים במטריצה  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  הם

$$\lambda_{1,2} = \frac{1}{2} \left[ \tau \pm \sqrt{\tau^2 - 4\Delta} \right]$$

כאשר

$$\Delta = ad - bc$$

$$\tau = a + d$$

כמו כן מתקיים:

$$\lambda_1 + \lambda_2 = \tau$$

$$\lambda_1 \lambda_2 = \Delta$$

הקשר בין ערך עצמי מרוכב לתדר וזמן מחזור:

$$\lambda = \alpha \pm i\omega$$

$$\omega = 2\pi f = \frac{2\pi}{T}$$

נוסחת אוילר

$$e^{i\theta} = \cos \theta + i \sin \theta$$

הוא היעקוביאן של המערכת

$$\begin{cases} \dot{x} = f(x, y) \\ \dot{y} = g(x, y) \end{cases}$$

$$\begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{pmatrix}$$

מבחן t לשתי קבוצות:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{X_1 X_2} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

כאשר שונות המדגם מוגדרת:

$$s_{X_1 X_2} = \sqrt{\frac{(n_1 - 1)s_{X_1}^2 + (n_2 - 1)s_{X_2}^2}{n_1 + n_2 - 2}} \quad s^2 = \frac{1}{N - 1} \sum_{i=1}^N (x_i - \bar{x})^2.$$

מרחק אוקלידי בין נקודות x, y עבור n מימדים:



$$d = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$