

תרגיל בית 2

הנחיות להגשה:

- יש להגיש –

1. בזוגות

- 2. קובץ PDF עם גרפים, הסברים וחישובים. שם הקובץ יהיה מורכב ממספר ת.ז. של שני הסטודנטים ויתחיל בשם Wet1 לדוגמה, Wet1_987654321_123456789.pdf.
- 3. יש להגיש את כל קבצי ה-python עם השמות הנכונים, כפי שנדרשתם בתרגילים.

- אופן ההגשה: יש לקבץ לקובץ עם השם HW2_ID1_ID2 בסיומת zip את קבצי הקוד וקובץ התשובות.
- בשאלות שבהן התבקשתם להציג גרף רשמו הסבר קצר ומסקנות ממנו. יש לציין צירים וכותרת. ניתן להשתמש בפקציית basic_plot הנתונה בקובץ הקוד. יש להציג את הגרף ביחידות dB, את ציר התדר ביחידות Hz וציר הזמן ביחידות של שניות, ולהציג בציר התדר רק את חלקו החיובי $[0, \frac{f_s}{2}]$.
- יש לשים לב כי הורדתם לסביבת העבודה שלכם את כלל הספריות הנדרשות בהתאם ל imports בקבצי הקוד המצורפים, או להריץ דרך google colab.
- במהלך התרגיל אתם נדרשים ל"ספרה מזהה". חברו את הספרה הראשונה של מספרי תעודות הזהות שלכם, מצאו את השארית של חלוקת התוצאה ב-10 = זוהי הספרה המזהה שלכם. חשוב! יש לציין את הספרה המזהה בתחילת במסמך ה-PDF של פתרון תרגיל הבית באופן ברור. הגשה ללא כתיבת מהי הספרה המזהה תגרור ציון 0 בתרגיל בית זה.

שאלה 1 – קונבולוציה בזמן = כפל בתדר

נתון אות בזמן רציף המורכב מסכום של שלושה סינוסים :

סינוס בתדר 1 kHz, סינוס בתדר 3 kHz וסינוס שלישי בתדר 5 kHz. כאשר X הוא הספרה המזהה שלכם (לדוגמה: עבור ספרה המזהה 4 תדר הסינוס השלישי יהיה 4.5 kHz). האות נדגם בתדר דגימה של 20 kHz. בקובץ PART1.ipynb השתמשו בקטעי הקוד שמתחת לכותרות "Question2 - A" עד "Question2 - D" כדי לממש את סעיפי השאלה הבאים.

א. צרו ווקטור המייצג את 1000 הדגימות הראשונות של האות והציגו גרף שלו. צרו את הגרף באמצעות הפונקציה plot_basic שממומשת בקובץ PART1.ipynb. הפונקציה מקבלת, בין השאר, את וקטור הערכים של ציר הזמן: כדי לייצר אותו תצטרכו להשלים את הקוד בפונקציה vector_time_create שמופיעה בקובץ.

ב. חשבו את התמרת הפורייה של האות בעזרת הפונקציות np.fft.fft, np.abs ו-np.log10. צרו גרף של התמרת הפורייה והציגו אותו, בעזרת הפונקציה basic_plot. כדי לייצר את וקטור הערכים של ציר התדר, השלימו את הקוד בפונקציה create_freq_vector שמופיעה בקובץ PART1.ipynb. על מה מחשבים את ה-log-10? מהי רזולוציית התדר של ההתמרה? כלומר, מהו "רוחב הסרט" שמייצג כל תא תדר בסריג התדר שהתקבל? הסבירו את תמונת התדר המתקבלת.

ג. הגדירו ווקטור חדש המכיל רק את 256 הדגימות הראשונות של האות, וחשבו מחדש את התמרת הפורייה על דגימות אלו. באיזה ערכים מתקבלים השיאים כעת? מהי רזולוציית התדר של ההתמרה כעת? כלומר, מהו "רוחב הסרט" שמייצג כל תא תדר בסריג התדר שהתקבל? הסבירו את תמונת התדר המתקבלת בתרגיל זה נבדוק את התכונה שקונבולוציה בין שני אותות בציר הזמן שקולה לכפל בין התמרות פורייה של האותות, והתמרה הפוכה חזרה לציר הזמן.

ד. כעת נשנה את תדר הסינוס השני ל-19 kHz במקום 3 kHz. הגדירו מחדש 256 דגימות של האות, בהנחה שהוא נדגם באותו תדר דגימה של 20 kHz כמו במקור. חשבו והציגו את ספקטרום האות. הסבירו את התוצאה המתקבלת באופן אנליטי ותארו סכמתית במישור התדר את תהליך הדגימה שהביא לתוצאה זו.

שאלה 2 –

כפי שהוצג לכם בתרגול, הציגו את 3 התדרים הנ"ל מסעיף 1ד' בתכנת praata הן בזמן והן בספקטרוגרמה. הוסיפו צילום מסך ל-PDF.

שאלה 3 – Pitch

- א. מהו תחום התדרים האופייני ל pitch של אות דיבור אנושי ?
- ב. חשבו את התחום האופייני למשך הזמן של מחזור pitch של אות דיבור אנושי. כלומר, המירו את התשובה שרשמתם בסעיף הקודם ליח' של שניות.
- ג. נניח שהאות דגום בתדר דגימה f_s ומשך המחזור הוא p דגימות. מהו תדר ה-pitch במקרה זה ? רשמו את הביטוי כתלות בפרמטרים p ו- f_s .
- ד. נניח ש- $p = 100$ samples , ו- $f_s = 8000$ Hz. מהו משך הזמן של מחזור ה-pitch? מהו התדר f_{pitch} המתקבל ? וודאו ששני הערכים (בזמן ובתדר) נמצאים בתוך תחומי הערכים האופייניים שחישבתם בסעיפים הראשונים.

שאלה 4 – הלם

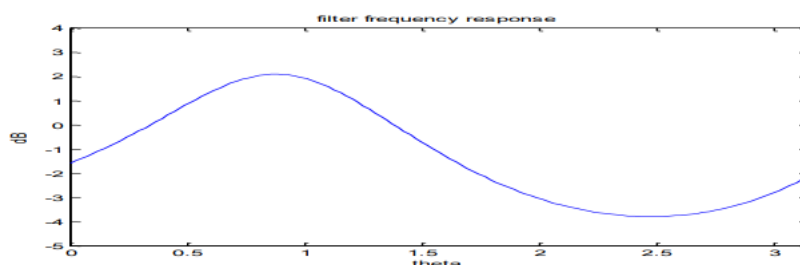


- א. מה נקבל אם נבצע התמרת פורייה על רכבת ההלמים הבאה:

$$\sum_{n=-\infty}^{\infty} \delta(t-nT) \xleftrightarrow{F} ?$$

- הציגו את התשובה כביטוי מתמטי והסבירו את משמעותו.
- אם אתם משתמשים בפרמטרים, ציינו את הקשר שלהם לפרמטרים המופיעים בביטוי של האות בתחום הזמן.

- ב. נתון מסנן H שצורת תגובת התדר שלו היא:



- אם נעביר את רכבת ההלמים דרך מסנן זה, מה תהיה צורת אות המוצא בתדר ? ציירו והסבירו איכותית

$$\sum_{n=-\infty}^{\infty} \delta(t-nT) \rightarrow \boxed{H} \rightarrow ?$$

שאלה 5 – אוטוקורלציה

פונקציית האוטוקורלציה מוגדרת כך:

$$r(k) = \sum_{n=0}^{N-k-1} x_n x_{n+k}, \quad k = 0, \dots, N-1$$

הניחו כי האות x מתאפס מחוץ לתחום שבין 0 ל- $(N-1)$.

- א. הוכיחו כי $r(k)$ סימטרית. כלומר, הוכיחו כי: $r(k) = r(-k), k = 0, 1, 2, \dots, N-1$
- ב. הסבירו (אין צורך בהוכחה פורמלית) מדוע פונקציית האוטוקורלציה מקיימת את התכונות הבאות:
 - a. המקסימום הגלובלי שלה מתקבל ב- $k = 0$ והוא שווה לאנרגיית האות.
 - b. אם האות $x(n)$ מחזורי במחזור P , מתקבלות מקסימום מקומיות ב- $0, \pm P, \pm 2P, \pm 3P$.
 - c. העוצמה של המקסימום המקומיות הולכות ודועכת בהתאם למרחק מ- $k = 0$.

שאלה 6 – אוטוקורלציה

שאלה זו מתייחסת לשערוך מחזור ה-pitch בשיטת האוטוקורלציה. מטרת השאלה היא לממש את הפונקציה `pitch_detect_corr` ולבדוק את נכונות המימוש באמצעות אות סינתטי. לטובת השאלה, הניחו כי האות המחזורי מייצג מחזור של `pitch`, על אף שמדובר באות סינתטי שאינו מייצג אות דיבור אמיתי.

- א. נתון קטע אות $y[n]$ בן 256 דגימות, המוגדר כסינוס בתדר $12X$ Hz, כאשר X מייצג את ספרת האחדות של התדר וערכו הוא הספרה המזהה שלכם. האות נדגם בתדר 8kHz . הציגו את האות בציר הזמן ונסו לשערך מתוך הגרף מהו המחזור של האות ביח' של זמן. חשבו את התדר שאמור להתקבל מתוך הערך שקיבלתם (היזכרו בשאלה מס' 3). האם התקבל בדיוק התדר שאליו כיוונתם? אם לא, מה הסיבה להבדל בין המספרים?
- ב. הציגו את גרף הספקטרום של האות וציינו על הגרף את תדר ה-`pitch` (מומלץ לוודא ע"י זום לתמונה). השוו לערך התדר שקיבלתם בסעיף א'. האם התקבל בדיוק אותו מספר? אם לא, מה הסיבה להבדל בין המספרים (היזכרו בשאלה מס' 2)?
- ג. הציגו את פונקציית האוטוקורלציה של האות ונסו לזהות בעין את אורך מחזור ה-`pitch` על פי גרף הפונקציה. כדי לחשב את פונקציית האוטוקורלציה של הקטע, היעזרו בפונקציה `correlate` (לקוחה מחבילת `scipy`). ציינו את אורך מחזור ה-`pitch` בדגימות וחשבו גם את התדר שלו. האם התקבלה התוצאה לה ציפיתם? הסבירו.

ד. ממשו פונקציה בשם `pitch_detect_corr` המקבלת כקלט מסגרת באורך 256 דגימות שנדגמה ב- 8kHz ומחזירה כפלט ערך יחיד שהוא משך מחזור ה- `pitch` (ביחידות של מספר דגימות) המאפיין את המסגרת.

הנחיות:

- a. עיבדו לפי ההסבר המופיע במבוא התיאורטי (סעיף 2.3), והיעזרו לשם כך בפונקציה (`find_peaks`) לקוחה מחבילת `scipy`.
- b. הפונקציה מניחה כי הקלט הוא קטע קולי, ולכן מחפשת קיום של מחזור `pitch` רק בתחום הערכים האפשריים מבחינה פיזיולוגית.
- c. במקרה שהפונקציה אינה מאתרת שיא מתאים שעונה לדרישה הנ"ל היא תחזיר ערך "אפס".
- ה. לאחר המימוש, הריצו את קטע הקוד תחת הכותרת "`pitch_detect_corr` Test" כדי לבחון את תקינות הפונקציה שכתבתם. וודאו שבסיום ההרצה קיבלתם אישור על תקינות הפונקציה שלכם. אם קיבלתם הודעת שגיאה יש לתקן את הפונקציה בהתאם להודעה.

שאלה 7 – הכרת אות הדיבור ותכונותיו

בתרגיל בית זה נבנה מערכת מקודד-מפענח פשוטה לדחיסת אות דיבור, המבוססת על חיזוי ליניארי. לשם כך נכיר את מאפייני אות הדיבור ונממש מספר בלוקים בסיסיים שירכיבו את המקודד המלא. בלוקים אלו הם:

1. סיווג אות כ-Unvoiced / Voiced
2. שערור מחזור ה- `pitch`.
3. שערור האנרגיה הממוצעת של אות העירור.
4. שערור מקדמי החזאי הלינארי (מקדמי ה- LPC).

כעת נעסוק בהכרת המאפיינים העיקריים של אות הדיבור. לאחר מכן נבנה את שני הבלוקים הראשונים. את שאר הבלוקים ואת הרכבת מערכת המקודד-מפענח ובדיקתו על אותות דיבור "אמיתיים" נבצע בתרגילי הבית הבאים.

1. בחירת קובץ הקלט המתאים

- קובץ הקלט שלכם ייקבע לפי המספר המזהה שלכם. לדוגמה: עבור מספר מזהה שערכו 1, קובץ הקלט יהיה `speech1.wav`. הקובץ דגום בקצב של 8 KHz.
- א. עדכנו את הספרה המזהה שלכם בקובץ `PART2.ipynb` וקראו את קובץ הקלט ואת תדר הדגימה ע"י שימוש בפונקציה `read_wav_file` הממומשת בקובץ
 - ב. אנחנו נעבוד עם מסגרות באורך 256 דגימות. האריכו את האות ע"י שרשור אפסים נוספים בסופו, כך שהאורך הכולל יהיה מספר שלם של מסגרות
 - ג. השמיעו את האות ע"י שימוש בפונקציה `Audio`. מהו המשפט הנאמר?

2. בחירת קובץ הקלט המתאים

- א. העתיקו את המימוש שלכם לפונקציית `create_time_vector` לקובץ. השתמשו בה וב- `basic_plot` כדי להציג את האות בציר הזמן.
- ב. הסתכלו בגרף האות בציר הזמן, ונסו להבין איפה מתחילה ומסתיימת כל מילה במשפט הנאמר. ניתן להשתמש בפקודה `Audio` כדי להאזין לקטעים ספציפיים באות ע"י הזנת זמן התחלה וסיום. לאחר שהחלטתם על זמן התחלה וסיום של מילה מסוימת, רשמו את המילה ואת הזמנים המתאימים לה בקטע הקוד המצורף בהמשך, ואז הריצו את הקטע שאחריו ליצירת גרף עם אזורים המסומנים בצבעים שונים. אם יש שתי מילים שמתחברות, הגדירו כמיטב הבנתכם איפה מסתיימת המילה הראשונה ומתחילה המילה השנייה, אין צורך לדייק מאד. הציגו את הגרף שמתקבל עם סימוני המילים. ניתן במקום זאת להשתמש גם בתכנת ה- `PRAAT` לשם כך. לפי מה שנחלצו לכם.
3. ניתוח אזורים קוליים ו-א-קוליים
- הציגו שוב את הגרף של האות בציר הזמן והשתמשו באופציית ה"זום" כדי לבחון חלקים מהאות במבט מקרוב. האם אתם מזהים בעין אזורים הנראים כ"קוליים"? האם אתם מצליחים לזהות אזורים שנראים כ"א-קוליים"?
- אתרו לפחות שני אזורים "קוליים" ושני אזורים "א-קוליים" וצינו את זמן ההתחלה והסוף שלהם בקוד. הריצו את הקוד שבהמשך כדי לסמן אותם על הגרף. הסבירו את הקריטריונים שלפיהם קבעתם את האזורים: ציינו תכונות המאפיינות אותות קוליים וא-קוליים שניתן להבחין בהן בגרף. התייחסו לצורת האות, עוצמת האות וכוד'.
4. בחירת מסגרת קולית ומסגרת א-קולית
- א. חלקו את האות למסגרות באורך 256 דגימות כל אחת (ניתן להשתמש ב `np.reshape`, שימו לב שבתוצר שמתקבל האיברים מסודרים לאורך השורות).
- ב. בחרו מסגרת של 256 דגימות מתוך אחד מהאזורים הקוליים שסימנתם בסעיף 1.2 והציגו אותה (לשם פשטות, בגרף זה ציר ה- x יכול לייצג דגימות מ-0 עד אורך המסגרת פחות 1). ציינו ב-PDF את אינדקס המסגרת, ומאיזו מילה והברה נלקחה המסגרת.
- מסגרת זו תיקרא בהמשך מסגרת הדוגמה הקולית.
- ג. בחרו מסגרת של 256 דגימות מתוך אחד מהאזורים הא-קוליים שסימנתם בסעיף 1.2 והציגו אותה (לשם פשטות, בגרף זה ציר ה- x יכול לייצג דגימות מ-0 עד אורך המסגרת פחות 1). ציינו ב-PDF את אינדקס המסגרת, ומאיזו מילה והברה נלקחה המסגרת.
- מסגרת זו תיקרא בהמשך מסגרת הדוגמה הא-קולית.

שאלה 8 – שערך תדר היסוד (Pitch) של הדיבור

1. שערך מתוך מסגרת קולית

- א. הציגו את מסגרת הדוגמה הקולית בזמן כאשר ציר ה- x הוא ביחידות של דגימות. שערך את מחזור ה- `pitch` מתוך התבוננות בגרף (השתמשו באופציית ה- `zoom` על מנת לקבל שערך טוב יותר). ציינו את אורך מחזור ה- `pitch` ביח' של דגימות, בזמן ובתדר.

- ב. הציגו את מסגרת הדוגמה הקולית בתדר כאשר ציר ה-x הוא ביחידות של הרץ. כדי לקבל תמונה טובה יותר נשתמש בטריק הבא (רלוונטי לסעיף זה בלבד): שרשרו למסגרת אפסים בסופה כך שאורכה הכולל יהיה 512, ואז בצעו התמרת פורייה למסגרת ה"מורחבת" ע"י שימוש בפונקציה
- מה היא רזולוציית התדר שמתקבלת?
- שערכו את תדר מחזור ה-pitch מתוך התבוננות בגרף התמרת ה-DFT שקיבלתם. האם המספר שקיבלתם תואם את התוצאה שקיבלתם בסעיף א? הסבירו.
- ג. סמנו על גבי הגרף שתיים מההרמוניות של תדר ה-pitch כמה הרמוניות אתם מצליחים לראות בבירור?
- ד. האם ניתן להבחין במעטפת הספקטרלית של מסנן מערכת הדיבור בתמונת ההתמרה? (רמז: היזכרו בשאלת 4) סמנו את עקומת המעטפת על גבי הגרף.
- אתם יכולים לצייר את העקומה בעזרת קוד, ע"י הצבת קואורדנאטות x ו-y מתאימות, או לצייר ידנית על גבי snapshot של תמונת הגרף. שני הסוגים של התשובות יתקבלו.
2. שערוך מתוך פונקציית אוטוקורלציה
- א. חשבו את פונקציית האוטוקורלציה של מסגרת הדוגמה הקולית. העריכו את אורך מחזור ה-pitch בדגימות מתוך הסתכלות בגרף הפונקציה. הסבירו את תשובתכם. חשבו גם את התדר שלו.
- ב. הפעילו את הפונקציה `pitch_detect_corr` שכתבתם בבית על מסגרת הדוגמה הקולית המקורית (ללא ריפוד באפסים). חשבו את התדר המתקבל והשוו לתוצאות הקודמות.
3. שערוך במסגרת א-קולית
- א. בצעו התמרת DFT למסגרת הדוגמה הא-קולית, והציגו אותה בתחום הזמן ובתחום התדר. האם גם כאן ניתן להבחין במחזור pitch? הסבירו את תשובתכם.
- ב. הציגו את פונקציית האוטוקורלציה של מסגרת הדוגמה הא-קולית. מהם ההבדלים העיקריים בין הגרף המתקבל, לגרף שהתקבל עבור מסגרת הדוגמה הקולית?

שאלה 9 – סיווג סגמנט לקולי / א-קולי

1. חישוב קצב חציות אפס
- א. ממשו בקוד פונקציה בשם `zero_cross` המקבלת כקלט וקטור דגימות באורך כלשהו ומחשבת את מספר חציות האפס שהתרחשו במסגרת, לפי הביטוי שמופיע בסעיף 2.5 של חומר הרקע
- ב. בידקו את תקינות הפונקציה שלכם על ידי הרצתה על שני ווקטורים נתונים המכילים מידע בעל אופי שונה:
1. הראשון מכיל סינוס בתדר 100 הרץ, שדגום בתדר 8000 הרץ.
 2. השני מכיל רעש לבן עם תוחלת 0 ושונות 1, ניתן להשתמש בפקודה `np.random.randn`.
- הריצו את הפונקציה `zero_cross` שכתבתם על שני הווקטורים. השוו את התוצאות.
2. חישוב אנרגיה לזמן קצר
- ממשו בקוד פונקציה בשם `short_time_energy` המקבלת כקלט וקטור דגימות באורך כלשהו ומחשבת את האנרגיה הממוצעת שלו לפי הנוסחה המופיעה בסעיף 2.5 של חלק המבוא.

3. מימוש הפונקציה `vu_classify`

א. ממשו קוד פונקציה בשם `vu_classify` ומהמקבלת כקלט מסגרת באורך 256 דגימות של אות הדיבור. השתמשו באותות שיצרתם בסעיף 1 בשאלה זו (הסינוס והרעש) כדי לבדוק את התקינות של הפונקציה.

להלן ההנחיות לכתיבת הפונקציה:

הפונקציה תקבל מסגרת של דגימות ותחזיר כפלט ערך בוליאני:

1 – המסגרת היא קולית או 0 – המסגרת אינה קולית (א -קולית או שקט)

בשלב ראשון הפונקציה תבדוק את ערך האנרגיה הממוצעת של המסגרת על פני חלון מלבני באורך המסגרת (השתמשו בפונקציה שמימשתם בסעיף הקודם).

אם ערך זה קטן מ-0.00001 (כלומר $1e-5$) נניח כי מדובר במסגרת המכילה "שקט" ונחזיר ערך '0'.

אחרת, הפונקציה תחליט שהמסגרת היא א- קולית אם מתקיימים **לפחות שניים** מתוך ארבעת

התנאים הבאים:

a. מספר חציות האפס במסגרת גדול מ- 33% (שליש) מאורך המסגרת.

b. ערך האנרגיה הממוצעת במסגרת (ששוערך קודם) קטן מאחוז אחד מהערך המוחלט המקסימלי במסגרת.

c. בחישוב פונקציית האוטוקורלציה, השיא המקומי הגדול ביותר שאינו ממוקם באפס, ממוקם באזור שאינו יכול להכיל ערכי pitch חוקיים. ניתן לאתר את האינדקסים של המקסימות מקומיות ע"י שימוש בפונקציות `np.argmax` ו-`find_peaks`.

d. בחישוב פונקציית האוטוקורלציה, היחס בין ערך השיא המקומי הגדול ביותר שאינו ממוקם באפס לבין ערך הפונקציה באפס, קטן מ-0.35.

4. ההבדלים בין אות קולי לאות א-קולי

א. חשבו את האנרגיה הממוצעת של המסגרת הקולית ושל המסגרת הא-קולית. השוו בניהן. למי אנרגיה ממוצעת גדולה יותר? יש לצרף צילום מסך של התוצאות.

ב. בדקו את הפונקציה `zero_cross` שכתבתם על שתי מסגרות הדוגמה. למי קצב חציות אפס גדול יותר? יש לצרף צילום מסך של התוצאות.

ג. בדקו את הפונקציה `vu_classify` שכתבתם על שתי מסגרות הדוגמה. האם היא סיווגה את המסגרות נכון? יש לצרף צילום מסך של התוצאות.