

# מודלים לא לינאריים בחקר ביצועים --- תרגיל בית 9

26 במאי 2024

## הנחיות להגשה

- יש להגיש שני קבצים בלבד:

- קובץ PDF עם תשובות לשאלות וצילום מסך של קוד ופלט.
- קובץ ZIP המכיל את כל קובצי הקוד אותם נדרשתם לכתוב.

## שאלה 1

בשאלה זו ננתח את התנהגותן של מניות הכלולות במדד האמריקאי S&A 500 בשנת 2016. נתמקד בניתוח מחיר הסגירה של כל מניה על פני שנה, שהוא מחיר המניה בסוף כל יום מסחר.

א. הורידו את הקובץ `prices.csv` המצורף לתרגיל הבית, בו נמצאים מחירי המניות בפורמט CSV. הריצו את קטע הקוד #1 בקובץ `script_HW9`, צרפו את הפלט שמתקבל והסבירו מה מבצעות השורה השנייה והשלישית בקוד.

ב. הריצו את קטע הקוד #2 בקובץ `script_HW9`, צרפו את התרשים שמתקבל, והסבירו מה מבצעת כל אחת משורות הקוד.

ג. הורידו את הקובץ `securities.csv` המצורף לתרגיל הבית וכתבו פונקציה המופעלת באופן הבא

```
symbols, prices, sectors = load_shares()
```

הפונקציה טוענת נתונים מהקבצים `prices.csv` ו-`securities.csv` מידע על מניות עליהן קיימים נתונים על כל ימי המסחר בשנת 2016, ומחזירה שלושה משתנים

- הרכיב `i` ב-`symbols` הוא מחרוזת עם שם המניה `i`.
- הרכיב `i` ב-`sectors` הוא מחרוזת עם שמו של המגזר העסקי של המניה `i`. המגזר העסקי של כל מניה מופיע בקובץ `securities.csv` בעמודה `GICSSector`.
- `prices[i,:]` הוא וקטור של מחירי הסגירה של המניה `i` בכל ימי המסחר בשנת 2016.

ד. כתבו פונקציה המופעלת באופן הבא

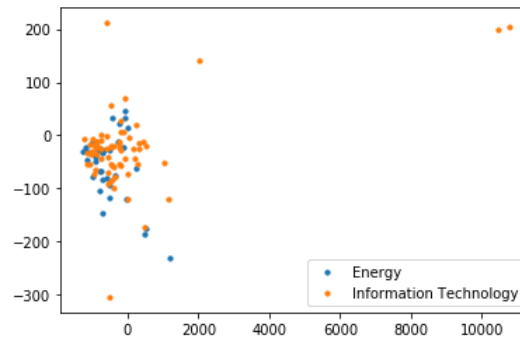
```
proj = pca_project(X, k)
```

הפונקציה מקבלת וקטורים בשורת המטריצה  $X \in \mathbb{R}^{n \times m}$  ומטילה אותה על  $k$  וקטורי הבסיס הראשונים המחושבים ע"י אלגוריתם PCA. הפונקציה מחזירה מטריצה כך ש-`proj[i,j]` היא ההטלה של `x[i,:]` על הציר `j`. אין להשתמש בסעיף זה בפונקציות מובנות למימוש PCA ויש לממש את האלגוריתם בעצמכם.

ה. כיתבו פונקציה המופעלת באופן הבא

```
plot_sectors(proj, sectors, sectors_to_plot)
```

הפונקציה מייצרת תרשים של נקודות במישור שנמצאות בשורות המטריצה `proj`, כאשר כל נקודה מייצגת מניה. המגזר העסקי של כל מניה נמצא במשתנה `sectors` והפונקציה יוצרת את התרשים רק עבור המניות במגזרים שנתונים במשתנה `sectors_to_plot`. לצורך בידקת הקוד עד כה, קטע הקוד #5 בקובץ `script_HW9` אמור ליצור את התרשים



1. בתרשים הקודם לא ניתן להבחין בהבדל בין התנהגותן של מניות בשני מגזרים אלה. הסיבה לכך היא ההבדל בקנה המידה של כל מניה. שתי חברות בעלות אותו שווי ואף התנהגות דומה יכולות להיראות שונה בתרשים, זאת כי אחת מהן בחרה לחלק את השווי שלה למעט מניות יקרות ואילו השנייה בחרה בהרבה מניות זולות. בעיה זו נקראת `Data Scaling`. אחת הדרכים להתגבר עליה כאשר מדובר בנתונים פיננסיים היא לא לייצג את המניה באמצעות וקטור המחירים  $p \in \mathbb{R}^m$  אלא באמצעות הווקטור

$$(\ln(p_{i+1}) - \ln(p_i))_{i=1}^{m-1}$$

חיזרו על הניסוי מהסעיף הקודם, רק שהפעם הזינו לפונקציה `pca_project` מטריצה שמורכבת מווקטורים בתצורה החדשה. צרפו את הקוד והתרשים שמתקבל.

2. אחת הדרכים הנפוצות להורדת סיכונים היא לפזר את ההשקעות במגזרים שונים, כך שמגמות במגזר מסויים ישפיעו במידה מועטה על מגמות במגזרים האחרים. תחת הנחה זו, האם כדאי להשקיע ב־זמנית בפיננסים ובטכנולוגיות מידע? מה לגבי השקעה ב־זמנית באנרגיה, טכנולוגיות מידע ונדל"ן? צרפו תרשימים שמגבים את החלטותיכם.

ח. כיתבו קטע קוד היוצר תרשים של כלל המניות מכלל המגזרים באמצעות הפונקציות אותן כתבתם בשאלה זו. בתרשים שיתקבל תבחינו במניה אחת בולטת. מיצאו את המניה (ה־symbol שלה) וצרפו תרשים של מחיריה לאורך השנה, לעומת מניה "טיפוסית" כלשהי. מה ניתן ללמוד מהתרשים? צרפו את שני התרשימים ואת קטעי הקוד.

## שאלה 2

תהי נקודה  $x^0 \in \mathbb{R}^n$  ויהי  $r > 0$ . נגדיר את הקבוצה  $B(x^0, r) = \{x \in \mathbb{R}^n : \|x - x^0\| \leq r\}$ . הראו ש־ $B(x^0, r)$  היא קמורה וסגורה ומיצאו נוסחה לחישוב הפונקציה  $P_{B(x^0, r)} : \mathbb{R}^n \rightarrow B(x^0, r)$ .