בס"ד

<u>סדנת מומחים 88-654</u>
<u>עבודה, תש"ף</u>

תאריך הגשה: יום ו' 20.3 עד לכניסת השבת. לשלוח במייל ל- `schiff@math.biu.ac.il`
מה יש להגיש: מסמך אחד בפורמט pdf עם התשובות לשאלות! יש לסכם מה עשית ואת התוצאות שקבלת. האורך של המסמך לא יעלה על 10 עמודים בפונט נורמלי. יש לדאוג להסברים ברורים והצגה יעילה של התוצאות.
הנתונים: שאלות 1,2 מתייחסות לנתונים על 50 מניות מהבורסה בתל-אביב. ניתן להוריד את הנתונים כקובץ zip מ- `http://u.math.biu.ac.il/~schiff/sharedata.zip`. בתוך קובץ זה יש 50 קבצי csv, כל אחד עם נתונים יומיים על מניה אחת במשך השנים 2013 עד 2018. הקבצים האלה הורדו מאתר הבורסה.
שאלה 3 מתייחסת לנתונים על שערי ריבית לשנת 2019 מאתר האוצר בארה"ב. לניתן לראות/להוריד את הקובץ בפורמט txt או בפרומט pdf מ- `http://u.math.biu.ac.il/~schiff/tcd.txt` או `http://u.math.biu.ac.il/~schiff/tcd.pdf`

1. אלגוריתם k-means הוא אלגוריתם למידה לא מפוקחת, עם המשימה למיין n נקודות נתונים (data points) במרחב ממימד d ל-k אשכולות (clusters).

   ניתן לתאר את האתגוריתם באופן הבא:

   (א) בחר k מרכזים התחלתיים, $a_1, \ldots, a_k$ (לדוגמה: בחר באופן אקראי k נקודות מתוך נקודות הנתונים).

   (ב) לכל אחת מנקודות הנתונים, $x_1, \ldots, x_n$, שייך אותה למרכז a_i הקרוב אליה ביותר (לדוגמה: במרחק האוקלידי).

   (ג) חשב מחדש כל מרכז $a_1, \ldots, a_k$ כממוצע הנקודות ששויכו אליו.

   על פעולות אלה חוזרים עד שבמהלך צעד מסוים המרכזים לא זזו (כלומר, שיוך הנקודות לא השתנה).

   השתמשו באלגוריתם k-means למיין את 50 המניות ל-5 אשכולות על סמך הנתונים הבאים:

   (א) ($d = 2$) הממוצע וסטיית התקן של התשואות היומיות של המניות.

   (ב) ($d = 2$) ממוצע התשואות היומיות וה-beta של המניות. (על מנת להגדיר את ה-beta יש להגדיר את השוק להיות העולם של כל 50 המניות, כאשר נותנים משקל סביר לכל מנייה ומנייה.)

   (ג) ($d = 6$) התשואות השנתיות של המניות.

   (ד) ($d = 72$) התשואות החודשיות של המניות.

   (ה) ($d = 72$) התשואות החודשיות המנורמלות של המניות. אם התשואה החודשית היא X, התשואה המנורמלת היא $\frac{X-\mu}{\sigma}$ כאשר μ ו- σ הם הממוצע וסטיית התקן של התשואות חודשיות.

   בכל מקרה יש לנסות לתת תיאור גרפי ו/או מילולי של התוצאות. שימו לב שהרצות שונות של האלגוריתם עלולות לתת תוצאות שונות (בגלל האקראיות בבחירות המרכזים ההתחלתיים). יציבות התוצאה היא אחת המדדים של ההצלחה של האלגוריתם.

2. בהנתן שתי סדרות זמניות $X_1,\ldots,X_N$ ו- $Y_1,\ldots,Y_N$, ה- r-lagged correlation coefficient של הסדרות הוא מקדם המתאם בין הסדרה $X_1,\ldots,X_{N-r}$ ובין הסדרה $Y_{r+1},\ldots,Y_N$. כאן $r=0,1,2,\ldots$. כששתי הסדרות הן אותה סדרה מקבלים את מקדמי המתאם העצמי (the autocorrelation coefficients) של הסדרה. כש-$r-0$ מקבלים את מקדם המתאם הרגיל. יש לשים לב שסדר הסדרות הוא חשוב!

(א) בחרו 4 מתוך 50 המניות שהן באותו סקטור (לדוגמה: בנקים, נדל"ן, היי-טק). חשבו עבור כל זוג מתוך 4 המניות (כולל הזוגות של כל מניה עם עצמה) את ה-r-lagged correlation coefficients של התשואות היומיות, עבור $r=0,1,2,3,4$. (לכל r יש להציג מטריצה של תוצאות, כאשר רכיב i,j נותן את המקדם הרלוונטי עבור הזוג מניה i עם מניה j.)

(ב) חזרו על אותו התליך, אבל עם ערך מוחלט של התשואות.

(ג) חזרו על התהליך מסעיפים (א) ו-(ב) אבל עם 4 מניות מסקטורים מגוונים.
כאן המקום להערות על כל התוצאות של 3 הסעיפים עד כאן: איפה יש קורלציות משמעותיות ואיפה אין?

(ד) בשיעור דברנו על תהליך GARCH(1,1) כתהליך מתאים לתשואות של מניה. התהליך $X(t)$ מוגדר על ידי המשוואות

$$X(t)=\sigma(t)Z(t), \qquad \sigma(t)^2=v+\alpha\sigma(t-1)^2+\beta X(t-1)^2$$

כאן v,α,β הם מקדמים, ו- $Z(t)\sim N(0,1)$ הוא רעש גאוסיאני לבן. הסברנו בשיעור שבמצב הסטציונרי

$$\mathbf{E}[X(t)]=0\,, \qquad \mathrm{Var}(X(t))=\frac{v}{1-\alpha-\beta}$$

ושיש צורך לקחת $\alpha,\beta>0$ עם $\alpha+\beta<1$.

ההכללה הפשוטה ביותר לתהליכי רב מימדי הוא (constant correlation coefficient GARCH), או, בקצרה, CCC-GARCH. לוקחים כל רכיב $X_i(t)$ להתנהג לפי GARCH, עם תהליך שונות $\sigma_i(t)$ משלו, אבל הרעשים הגאוסיאניים הלבנים $Z_i(t)$ הם תלויים זה בזה, עם מקדם מתאם ρ.

על ידי סימולציה של זוג של תהליכים שמקיים CCC-GARCH תארו את התלות של ה- r-lagged correlation coefficients של הסדרות $X_i(t)$ ו- $|X_i(t)|$ על הפרמטר ρ. מספיק כאן לקבוע את כל הפרמטרים האחרים $v_1,v_2,\alpha_1,\alpha_2,\beta_1,\beta_2$ ולשנות את ρ.

3. (א) חשבו את ה- annualized swap rates לתקופות של 5 ו-10 שנים המתאימים לשערי הריבית של האוצר, לכל יום מסחר של שנת 2019. יש להציג את התוצאות כגרף, ולציין את השערים הגבוהים והנמוכים ביותר במשך השנה. שימו לב: המספרים שמופיעים באתר של האוצר הם continuous compounding rates, אבל יש לתת את התשובה בצורה של annualized rates.

(ב) במשך שנת 2019 עקום הריבית עבר היפוך (inversion). בדרך כלל הריבית ל-10 שנה היא יותר גבוהה מהריבית לחודש, אבל באמצע השנה המצב התהפך. האם קרה שה- swap rate ל-5 שנים עלה מעל ה- swap rate ל-10 שנים? האם ייתכן מצב כזה?

(ג) באתר SEB (חברת פיננסים שוודית) מצאתי את הרשימה הבאה של swap rates במטבעות שונים:

Years	1	2	3	4	5	6	7	8	9	10
Rate SEK	0.14	0.13	0.14	0.16	0.18	0.21	0.25	0.29	0.33	0.37
Rate NOK	1.80	1.83	1.79	1.76	1.75	1.75	1.74	1.74	1.75	1.75
Rate DKK	-0.24	-0.23	-0.21	-0.19	-0.16	-0.13	-0.10	-0.06	-0.01	0.03

מצאו zero coupon rates מתאימים.
מהי הסיבה לשוני בין השערים במטבעות השונים?

עבודה נעימה!