

קורס 67800 תשע"א
בחינה מועד ב'

הנחיות:

- בבחינה חמש שאלות. עליכם לענות על שלוש מתוכם (משקל כל שאלה 33 נקודות).
- לא ייבדקו יותר משלוש שאלות כך שאם פתרתם שאלות נוספות עליכם לבחור אילו שאלות ייבדקו ולסמן זאת בצורה ברורה.
- עליכם לנמק היטב כל תשובה ולהוכיח באופן מדויק כל טענה שנדרשתם להוכיח. תשובה נכונה ללא נימוק ו/או דרך לא תזכה אתכם בנקודות.
- עליכם לענות על כל השאלות הבאות באופן עצמאי. תלמיד שיתפס מעתיק, או שיהיה קיים חשש כי העתיק יועבר לטיפול רשויות האוניברסיטה.
- אין להשתמש בחומר עזר מכל סוג שהוא.
- (1 נקודות) ודאו כי מספר הזהות שלכם רשום על המחברת.

משך הבחינה: שעתיים וחצי

בהצלחה !

1.

א. (10 נקודות) הראו ש $(X \perp Y, W|Z) \Rightarrow (X \perp Y|Z)$.

ב. (13 נקודות) הראו שאם פילוג הוא רשת מרקובית עבור גרף G אז הוא מקיים את התכונות $I_{sep}(G)$.

ג. (10 נקודות) אם ידוע שפילוג $p(x)$ מתפרק לפי הגרף המכוון $X_1 \rightarrow X_3 \leftarrow X_2$, האם תמיד קיים גרף **לא מכוון** G כך ש $I(p) = I_{sep}(G)$? אם כן, מה הוא G ? מה תהיה תשובתכם אם $p(x)$ מתפרק לפי הגרף המכוון $X_1 \rightarrow X_3 \rightarrow X_2$. נמקו תשובתכם בפירוט.

2. נתונה רשת בייסיאנית $X_1 \rightarrow X_3 \leftarrow X_2$ ו $x_i \in \{0,1\}$. נתון ש $p(x_1), p(x_2)$ אחידים. כמו כן הפילוג $p(x_3|x_1, x_2)$ תלוי בפרמטר יחיד θ והתלות נתונה ע"י: $p(x_3|x_1, x_2; \theta) \propto e^{\theta x_3(x_1+x_2)}$.

בהנתן מדגם מתוך פילוג זה נסמן ב $p_D(x_1, x_2, x_3)$ את הפילוג האמפירי.

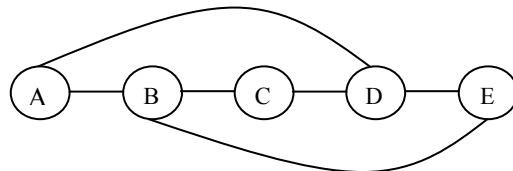
א. (10 נקודות) הראו שפונקציית הנראות תלויה רק בהסתברויות האמפיריות הבאות: $p_D(x_1, x_2)$ ו $p_D(x_1 = 1, x_3 = 1), p_D(x_2 = 1, x_3 = 1)$. כתבו תנאי עבור אומד ה ML התלוי רק בהסתברויות אלו.

ב. (11 נקודות) מהו אומד ה ML כאשר $p_D(x_1, x_2, x_3)$ מתפלג אחיד (תוכלו "לנחש" פיתרון. לא כדאי לנסות לפתור ישירות)? הראו שהוא מקיים את התנאי שרשמתם בסעיף הקודם.

ג. (12 נקודות) נתון מדגם כך ש $p_D(x_1 = 1, x_2 = 1) = 0$. מצאו ביטוי סגור עבור אומד ה ML לפרמטר θ .

3. בשאלה זו נבחן אסטרטגיות שונות לסדר אלימינציה. בדומה למה שראיתם בתרגול, אסטרטגיה תאופיין על ידי פונקציה $f(X_i)$ שתתן ציון לכל צומת שעוד לא עשינו לו אלימינציה. הפונקציה תהיה תלויה בגרף הנוצר אחרי מחיקה של המשתנים שעברו אלימינציה (כאשר בכל שלב מחברים את כל השכנים של צומת שעבר אלימינציה). הצומת שיבחר לאלימינציה בכל שלב הוא זה שעבורו $f(X_i)$ מינימלי. אם יותר מצומת אחד מחזיר את הערך המינימלי, נבחר אחד מהם באופן שרירותי.

נתונה הרשת המרקובית; הבאה:



למשתנים מספר ערכים כדלהלן: $|A|=2, |B|=10, |C|=2, |D|=9, |E|=2$. הפונקציה $f_C(X_i)$ מוגדרת להיות מספר השכנים של X_i .

א. (5 נקודות) עבור הרשת הנתונה, תארו סדר אלימינציה המשתמש בפונקציה $f_C(X_i)$ וישו B הוא המשתנה האחרון.

- ב. (6 נקודות) עבור סדר האלימינציה שהצעתם ב (א), מהו מספר פעולות הכפל ומספר פעולות החיבור הנדרש על מנת לחשב את הפילוג השולי של B.
- ג. (11 נקודות) הציגו אסטרטגיה אחרת (יש להגדיר פונקציה f מתאימה) אשר הינה יותר יעילה עבור החישוב שמתואר ב (ב). עליכם להסביר מדוע האסטרטגיה החדשה יעילה יותר אבל אין צורך לחשב את מספר הפעולות הנדרשות באופן מדויק.
- ד. (11 נקודות) תנו דוגמא לגרף קשיר ובו לפחות 5 צמתים כך ש $f_C(X_i)$ לא יותר גרועה מהאסטרטגיה שהצעתם ב (ג). גם כאן הניחו שכאשר יש תיקו הבחירה היא שרירותית.
4. בשאלה זו נעסוק בלימוד המבנה של מודל ה-TAN שפגשתם בתרגול. להזכירכם, במודל זה משתנה C אשר הוא הורה של משתנים X_1, \dots, X_n . בנוסף בין המשתנים האלו אנו מאפשרים להוסיף קשתות כך שלא נוצרים מעגלים. בתרגול ראינו כי ניתן ללמוד את העץ האופטימלי מעל X_1, \dots, X_n (כאשר C הורה של כולם) ביעילות, תחת ציון הנראות. בשאלה זו נרצה להשתמש בפונקציית ה-BIC על מנת ללמוד קשתות מעל המשתנים X_1, \dots, X_n .
- א. (11 נקודות) הראו בפרוט כיצד ללמוד באופן יעיל את העץ הטוב ביותר תחת פונקציית הציון הנ"ל. הוכיחו את נכונות הבנייה שלכם.
- ב. (10 נקודות) הסבירו מדוע יתכן שהמבנה ללא מעגלים הממקסם את פונקציית הציון מעל המשתנים X_1, \dots, X_n אינו העץ שמצאתם ב (א), אלא יער (אוסף של עצים).
- ג. (12 נקודות) ידוע כעת שלהתפלגות ממנה נדגמו הנתונים מבנה של Naïve Bayes, ללא כל קשתות בין המשתנים X_1, \dots, X_n . בהינתן מדגם, האם ייתכן כי המבנה הממקסם את פונקציית הציון אינו Naïve Bayes? הוכיחו את טענתכם.
5. נניח כי אנו במצב בו אנו מעוניינים בהתפלגות $P(X)$ שאיננו יכולים לחשבה או לדגום ממנה. בידינו $P'(X) \propto P(X)$ אשר אנו יכולים לחשב אותה עבור כל ערך של X אך איננו יכולים לדגום ממנה. כמו כן בידינו Q(X) אשר ניתנת לחישוב לכל ערך של X ואשר ניתן לדגום ממנה.
- א. (11 נקודות) הסבירו כיצד ניתן להשתמש ב-Q על מנת לחשב $P(X=x)$ בשיטת ה-Normalized Importance Sampling והוכיחו את נכונות השיטה.
- ב. (10 נקודות) תארו במדויק את אלגוריתם Metropolis-Hastings לדגימה מפילוג $P(x)$ עם proposal distribution הנתונה על ידי $Q(x'|x)$. מה ניתן להגיד על סדרת הדגימות אם הדגימה הראשונה מוגרלת מ $P(x)$?
- ג. (12 נקודות) נתמקד במשתנה n ממדי $x = x_1, \dots, x_n$. נתונה המוגרלת באופן הבא: תחילה מוגרל אינדקס i באופן אחיד בין אחד ל n ואז:
- $$Q(x'|x) = \begin{cases} p(x'_i|x_{-i}) & x'_i = x_{-i} \\ 0 & \text{Otherwise} \end{cases}$$
- הראו שהאלגוריתם MH שמתקבל שקול ל Gibbs-Sampling עם בחירה אחידה של המשתנים המעודכנים.