שיעור 5 - 28-04-2025

סעיף 2 – וקטורים מקריים

בסעיף זה נעסוק בווקטורים מקריים $X\colon \Omega \to \mathbb{R}^n$ עבור $n \geq 1$. הערכים של הווקטור המקרי הם משתנים מקריים סקלרים $X_j\colon \Omega \to \mathbb{R}$.

2.1 התפלגות משותפת

התפלגות ההסתברות של ווקטור מקרי כלשהו מוגדרת על ידי ההתפלגות המצטברת שלו:

הגדרה 2.1

פונקציית ההתפלגות המצטברת של ווקטור מקרי X עם ערכים ב- $\mathbb{R}^n$ היא:

$$F_X(x) = \mathbb{P}(X \leq x), \;\; x \in \mathbb{R}^n .$$

כמו במקרה הסקלרי החד ממדי, פונקציית ההתפלגות המצטברת מקיימת מספר תכונות מוגדרות.

התכונה השלישית - ***רציפות מימין*** צריכה להתקיים לכל מרכיב של הווקטור, כאשר שאר המרכיבים מוחזקים קבועים, כלומר:

$$\lim_{\varepsilon \searrow 0} F_X(x + \varepsilon e_i) = F_X(x) \qquad \forall x \in \mathbb{R}^n, \qquad \forall i = 1, \dots, n$$

כאשר e_i הם ווקטורי היחידה של הבסיס הסטנדרטי ב- $\mathbb{R}^n$ (עם ערך 1 במקום ה-*i* ו- 0 בשאר המקומות) .

התכונה השנייה – ***נורמליזציה***, הופכת להיות

$$\lim_{R \to -\infty} F_X\,(x + Re_i) = 0 \qquad \forall x \in \mathbb{R}^n, \quad \forall i = 1, \dots, n$$

וכן

$$\lim_{x \to \infty} F_X(x) = 1$$

כאשר $x \to \infty$, משמעו שכל הרכיבים של x שואפים לאינסוף בדרך כלשהי.

בכדי לנסח תכונה המקבילה לתכונה הראשונה – *תכונת המונוטוניות* נגדיר את <u>*האופרטור ההפרש*</u> $\Delta^i_{r,q}$ אשר פועל על פונקציות $h\colon \mathbb{R}^n \to \mathbb{R}$:

$$\left(\Delta^i_{r,q} \circ h\right)(x) \coloneqq h\big((x_1, \dots, x_{i-1}, q, x_{i-1}, \dots, x_n)\big) - h\big((x_1, \dots, x_{i-1}, r, x_{i-1}, \dots, x_n)\big).$$

(נזכור שהסימן $g \circ f$ מתאר הרכבה של פונקציות).

התכונה המקבילה למונוטוניות היא <u>תכונת האי שליליות</u>:

$$\left(\Delta^1_{a_1,b_1} \circ \cdots \circ \Delta^n_{a_n,b_n} F_X\right)(x) \geq 0 \qquad \forall a, b \in \mathbb{R}^n$$

כלומר, ההסתברות של כל היפר מלבן היא אי שלילית.

כל פונקציה F המקיימת את שלוש התכונות שפרטנו למעלה היא פונקציית התפלגות מצטברת לגיטימית. דהיינו, ניתן לבנות מרחב הסתברות ולהגדיר עליו משתנה מקרי כך ש: $\mathbb{P}(\{\omega \in \Omega: X(\omega) \leq x\}) = F(x)$.

<u>למשל</u>, עבור *d=2*:

את ההסתברות של ההיפר מלבן הדו-ממדי $(a_1, b_1] \times (a_2, b_2]$, את $\mathbb{P}(a_1 < X_1 \leq b_1, a_2 < X_2 \leq b_2)$, אפשר למצוא על ידי הנוסחה:

$$F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) \geq 0$$

חישוב ההסתברות של המלבן הדו-ממדי נעשה על ידי הפרש של ערכי ארבעת הפינות תוך שימוש בטכניקת הכללה-החסרה (*inclusion-exclusion*) :

נתחיל בהסתברות בין הפינה השמאלית הנמוכה עד לפינה הימנית הגבוהה - $F(b_1, b_2)$.

נחסיר את ההסתברות של הרצועה השמאלית - $F(a_1, b_2)$

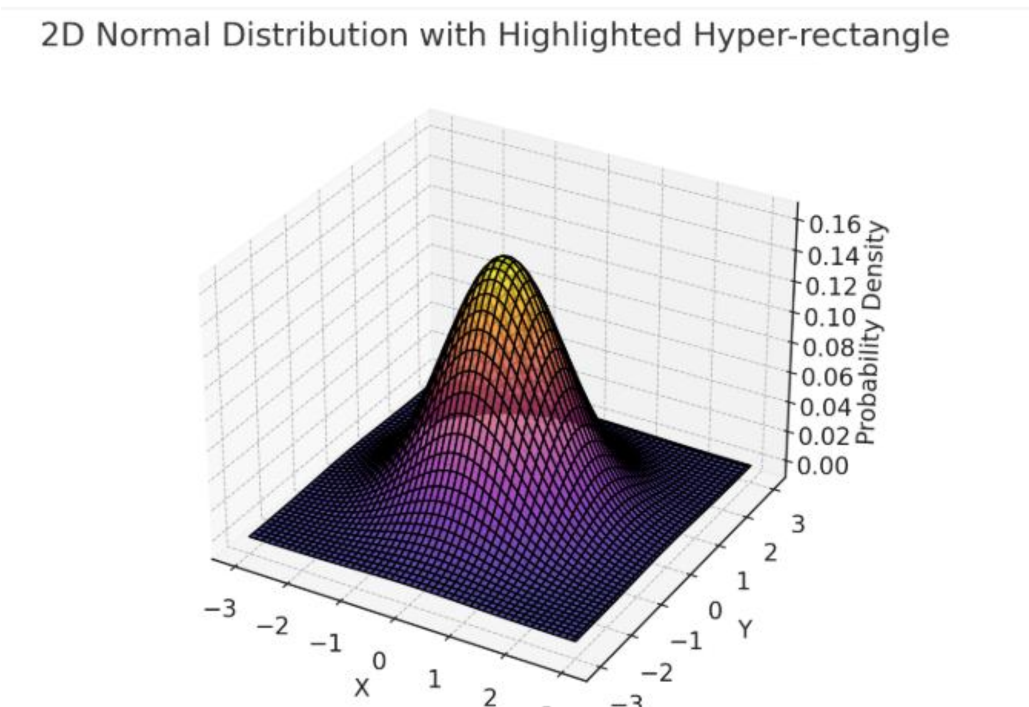
נחסיר את ההסתברות של הרצועה התחתונה - $F(b_1, a_2)$

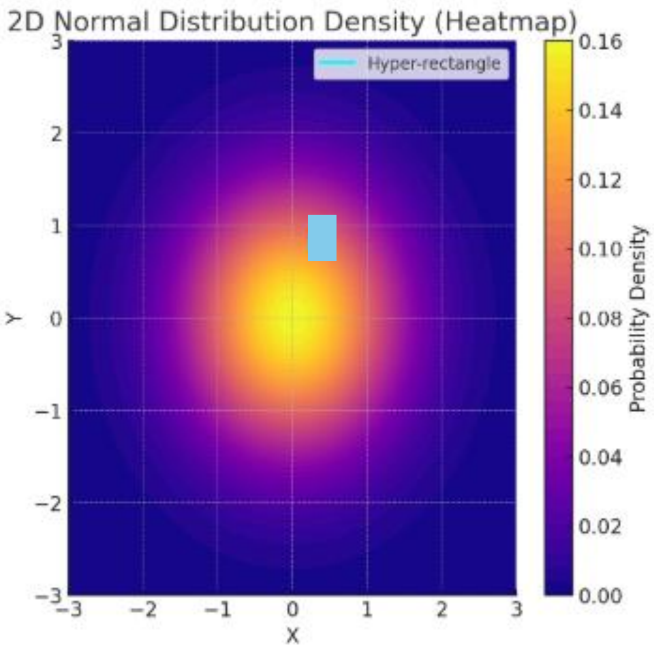
נוסיף את ההסתברות של הפינה הכפולה אשר הורדה פעמיים - $F(a_1, a_2)$

<u>מטפורה ויזואלית</u> - אפשר לדמיין שאנחנו מערימים קופסאות בפינה של חדר, ואנו מעוניינים בחישוב נפח של קופסא מסוימת (המלבן). נתחיל תחילה ניקח את כל הקופסאות עד לפינה הרחוקה ביותר, ואז נחסיר את ה"אגף השמאלי" וה"אגף התחתון". אולם יש פינה שמוחסרת פעמיים לכן יש להוסיף אותה חזרה.

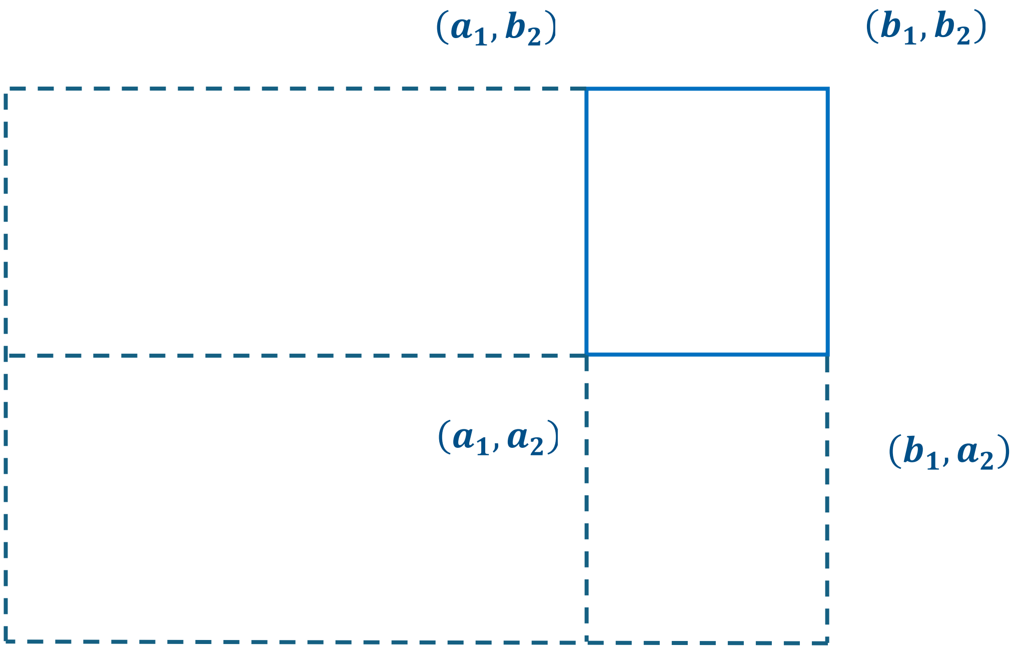
חישוב ההסתברויות של היפר מלבנים בממדים גבוהים יותר גם היא אפשרים באמצעות טכניקת הכללה-החסרה (התכונה 3 מבטיחה הסתברויות חיוביות).

<u>תרשימים להמחשה</u>:





אזהרה – צריך לדמיין את החדר בתלת-ממד!



כמו במקרה החד ממדי, לפונקציית התפלגות מצטברת רב ממדית יכול להיות חלק בדיד וחלק רציף.

מקרה נפוץ הוא כאשר <u>F_X היא רציפה לגמרי</u> ולפונקציית ההתפלגות המצטברת הרב ממדית <u>יש נגזרות חלקיות</u>

$$f_X(x) = \frac{\partial^n}{\partial x_1 \dots \partial x_n} F_X(x), \tag{2.1}$$

כך ש:

$$F_X(x) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f_X(u) du_1 \dots du_n \ .$$

הפונקציה $f_X(x \cdot)$ נקראת פונקציית הצפיפות של הווקטור המקרי, או פונקציית הצפיפות המשותפת של המשתנים המקריים $X_1, \dots X_n$ המרכיבים את הווקטור המקרי X.

ההסתברות של מאורע $\{\omega \in \Omega : X(\omega) \in \Gamma\}$ כאשר $\Gamma \subseteq \mathbb{R}^n$ תת קבוצה של מרחב הממשי ה-n ממדי, מחושבת על ידי אינטגרציה רב ממדית:

$$\mathbb{P}(X \in \Gamma) = \int \dots \int_\Gamma \ f_X(u)\, du.$$

כל פונקציה אי שלילית שהאינטגרל שלה מעבר ל- $\mathbb{R}^n$ שווה ל-1 היא פונקציית צפיפות לגיטימית.

<u>דוגמא 2.2</u>

נסתכל על ניסוי בו מודדים בנקודת זמן מסוימת את המיקום של פריט הנע באקראיות בתיבה. לאור האקראיות, התפלגות הווקטור המקרי של מיקום הפריט בעת המדידה היא התפלגות אחידה *3D* (תלת-ממדית).

יהי $X = (X_1, X_2, X_3)$ התפלגות אחידה על היפר- מלבן $(a_1, b_1] \times (a_2, b_2] \times (a_3, b_3]$:

$$f(x_1, x_2, x_3) = \begin{cases} \dfrac{1}{(b_1 - a_1)(b_2 - a_2)(b_3 - a_3)} & a_1 \le x_1 \le b_1,\ \ a_2 \le x_2 \le b_2,\ \ a_3 \le x_3 \le b_3 \\ 0 & \text{אחרת} \end{cases}$$

דוגמא ליישום –רובוטים שדוגמים מיקום או אוריינטציות באורח אחיד בתוך מרחב עבודה תלת ממדי עבור תיכנון תנועה, משימות חקירה או משימות כיסוי.

מקרה חשוב אחר הוא כאשר <u>F_X היא בדידה לגמרי</u>, כלומר היא משתנה רק בנקודות השייכות לקבוצה בת מניה של וקטורים $\{x_1, x_2, \dots\}$. במקרה כזה נאמר כי ל-X יש התפלגות הסתברויות *(p.m.f.)*, ולכל $\Gamma \subseteq \mathbb{R}^n$

$$\mathbb{P}(X \in \Gamma) = \sum_{x_j \in \Gamma} P_X(x)$$

<u>דוגמא 2.3</u>

נסתכל על **ההתפלגות המולטינומית** - $X \sim Mult(n, p)$, עם פרמטר שלם n ו-p הוא ווקטור של פרמטרים כך ש- $p \in S^{k-1}$.

S^{k-1} הוא **סימפלקס** של הסתברויות ממימד $k-1$ כלומר: $S^{k-1} = \left\{x \in \mathbb{R}^k : x_j \ge 0\ \forall j\ ,\ \sum_{i=1}^k x_i = 1\right\}$.

ההתפלגות המולטינומית היא התפלגות בדידה לגמרי, ומהווה הכללה של ההתפלגות הבינומית. נניח מבצעים ניסוי עם k תוצאות אפשריות. הסתברות לקבל כל תוצאה היא p_j . חוזרים על הניסוי n פעמים בלתי תלויות, והווקטור המקרי X מציג את מספר הפעמים שהתקבלה כל תוצאה j .

פונקציית התפלגות ההסתברויות היא

$$P_X(x) = \frac{n!}{x_1!\ldots x_n!} p_1^{x_1} \ldots p_k^{x_k}\ ,\ x_i \in \mathbb{N} \cup \{0\}\ ,\ \sum_{i=1}^{n} x_i = n. \qquad (2.2)$$

ההתפלגות הבינומית מתקבלת כאשר $n = 2$.

כמובן, יחסי הגומלין בין החלקים הבדידים והרציפים עשויים להיות מורכבים הרבה יותר לאור מבנה גיאומטרי עשיר יותר של $\mathbb{R}^n$ עם ממד n הגדול מ-1. בפרט, ייתכן שיהיו לו אטומים בממדים נמוכים, כפי שמדגימה הדוגמה הבאה.

<u>דוגמא 2.4</u>

יהי ξ משתנה מיקרי עם ערכים ב- $\mathbb{R}$ ופונקציית צפיפות $f_\xi(\cdot)$. נגדיר ווקטור מקרי $X = (\xi, \xi)$ פונקציית ההתפלגות המצטברת של הווקטור היא

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\xi \leq x_1\ ,\ \xi \ll x_2) = F_\xi(x_1 \wedge x_2) \qquad (2.3)$$

כאשר כאן נשתמש $x_1 \wedge x_2 = \min\,(x_1, x_2)$.

אם לפונקציית התפלגות זו הייתה צפיפות, נניח $f(\cdot)$

$$\mathbb{P}(X_1 = X_2) = \iint_{\{x \in \mathbb{R}^2 : x_1 = x_2\}} f_X(x) dx = 0$$

אולם זה סותר את הברור מאליו כי $\mathbb{P}(X_1 = X_2) = \mathbb{P}(\xi = \xi) = 1$. לכן לא קיימת פונקציית צפיפות.

למעשה נוסחה (2.1) של הנגזרות החלקיות שווה 0 בכל מקום למעט על האלכסון בו הנגזרת לא קיימת.

ההתפלגות המוגדרת על ידי (2.3) מרוכזת על האלכסון, והמשתנה המקרי X הוא <u>לא בדיד</u> מכיוון שההתפלגות השולית רציפה אבל הוא גם <u>לא רציף</u> כפי שראינו קודם.

<u>ההתפלגויות השוליות</u> של התפלגויות רב ממדיות הן התפלגויות תקפות מממד קטן יותר. אם X הוא ווקטור מקרי ב-$\mathbb{R}^n$ עם פונקציית התפלגות מצטברת $F_X(x)$, אזי פונקציית ההתפלגות המצטברת של הווקטור המקרי

$$X_J \coloneqq \left(X_{i_1}, \ldots, X_{i_k}\right), \qquad J = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$$

מתקבלת על ידי לקיחת הרכיב $x_i\ i \notin J$ לאינסוף, תוך שמירה על $x_i\ i \in J$ קבועים. פונקציית ההתפלגות המצטברת $F_{X_J}(\cdot)$ היא המצטברת השולית ה-k ממדית . ובפרט אם ל- $F_X(x)$ יש צפיפות אזי גם ל- $F_{X_J}(\cdot)$ יש צפיפות המתקבלת מהצפיפות המקורית על ידי אינטגרציה "החוצה" של כל $x_i\ i \notin J$. פונקציית הצפיפות זו $f_{X_J}(\cdot)$ היא הצפיפות השולית ה- k ממדית של פונקציית הצפיפות $f_X(\cdot)$.

כל הנאמר כאן נכון גם עבור פונקציות הסתברות שוליות. בתרגיל תקבלו שאלות העוסקות בדוגמאות שונות של התפלגויות רב ממדיות והתפלגויות שוליות שלהן.

דוגמא 2.5 (לא הוצגה בשיעור)

נסתכל על זוג המשתנים המקריים X ו- Y עם הצפיפות

$$f_{X,Y}(x,y\,;\alpha)=f_x(x)f_Y(y)(1+\alpha[2F_X(x)-1)(2F_Y(y)-1)]$$

כאשר $\alpha\in[-1,1]$ פרמטר של הצפיפות, f_x ו- f_Y פונקציות צפיפות חד ממדיות ו- F_X ו- F_Y הן פונקציות ההתפלגות המצטברת המתאימות.

קל לראות כי הצפיפות הדו-ממדית אי שלילית. כמו כן

$$\int_{\mathbb{R}} f_x(x)\,2F_X(x)-1)dx=\int_{\mathbb{R}}\frac{d}{dx}\big(F_X(x)\big)^2\,dx-1=0$$

לכן האינטגרל של $f_{X,Y}$ שווה ל- 1 לכל $|\alpha|\le 1$.

יתרה מזאת, פונקציות הציפות השוליות הן:

$$\int_{\mathbb{R}} f_{X,Y}(x,y\,;\alpha))\,dx=f_Y(y)\quad,\quad\int_{\mathbb{R}} f_{X,Y}(x,y\,;\alpha))\,dy=f_x(x)$$

כלומר, לפונקציות צפיפות משותפת שונות המתאימות לערכי α שונים יש אותן פונקציות צפיפות שוליות.

2.2 תוחלת

תוחלת ביחס לווקטור מקרי מוגדרת בדומה למקרה הסקלר החד ממדי. יהיה X ווקטור מקרי רציף לגמרי עם פונקציית צפיפות $f_X(\cdot)$ המוגדרת על $\mathbb{R}^n$, אזי לכל פונקציה $h:\mathbb{R}^n\to\mathbb{R}^m$

$$\mathbb{E}h(X):=\begin{pmatrix}\mathbb{E}h_1(X)\\ \vdots\\ \mathbb{E}h_m(X)\end{pmatrix}$$

כאשר

$$\mathbb{E}h_j(X)=\int_{\mathbb{R}^n} h_j\,(x)f_X(x)dx$$

כל אימת שהאינטגרל מוגדר היטב.

עבור X ווקטור מקרי שהוא בדיד לגמרי, נחליף את האינטגרל בסכום משוקלל בהסתברויות.

נסתכל עתה על המקרה הפרטי הפשוט $\mathbb{E}(X)=(\mathbb{E}(X_1),\dots,\mathbb{E}(X_n))^T$. מכל מה שלמדנו על מרחבים ווקטורים ברור כי תכונות התוחלת החד ממדית תקפות גם עבור התוחלת הרב ממדית, כאשר השוויון והפעולות נעשות על מרכיבי הווקטור פי כללי המרחב הווקטורי מעל שדה הממשיים:

1. לכל X ו-Y ווקטורים מקריים ו- A ו-B מטריצות קבועים בממדים מתאימים קיים

$$\mathbb{E}(AX+BY)=A\mathbb{E}(X)+B\mathbb{E}(Y)\tag{2.4}$$

2. מטריצת השונות המשותפת של הווקטורים $X \in \mathbb{R}^n,\ Y \in \mathbb{R}^m$ היא

$$Cov(X,Y) = \mathbb{E}((X-\mathbb{E}X)(Y-\mathbb{E}Y)^T = \mathbb{E}XY^T - \mathbb{E}X\mathbb{E}Y^T$$

רכיבי המטריצה הם $Cov\left(X_i, Y_j\right)$

3. בפרט מטריצת ה-Cov של ווקטור מקרי $X \in \mathbb{R}^n$ היא

$$Cov(X,X) = \mathbb{E}(X-\mathbb{E}X)(X-EX)^T$$

נשים לב כי המטריצה $Cov(X,X)$ מוגדרת אי שלילית, כלומר $\forall\ \mathrm{v} \in \mathbb{R}^n$

$$\mathrm{v}^T Cov(X,X)\mathrm{v} = \mathbb{E}[\mathrm{v}^T(X-\mathbb{E}X)(X-EX)^T\mathrm{v}] = \mathbb{E}\left(\mathrm{v}^T(X-\mathbb{E}X)\right)^2 \geq 0$$

רכיבי האלכסון של $Cov(X,X)$ הם השונות של הסקלרים $var(X_j)$, וכן מתקיים:

$$tr(Cov(X,X)) = tr(\mathbb{E}(X-\mathbb{E}X)(X-\mathbb{E}X)^T) = \mathbb{E}\left(tr(X-\mathbb{E}X)^T(X-\mathbb{E}X)\right) = \mathbb{E}\|X-\mathbb{E}X\|^2$$

מטריצת השונות המשותפת מקודדת את התלות הליניארית הזוגית בין הערכים של וקטור X, כפי שיוסבר בפירוט בהמשך.

2.3 פונקציה יוצרת מומנטים בשלב זה עקב אילוצי זמן לא נלמד בשיעור – ישלח בהמשך כחומר קריאה)