# פתרון מטלה 5 — חישוביות וקוגניציה, 6119

8 בדצמבר 2025

## שאלת הכנה

### סעיף א׳

נבדוק מה נכון לומר על מקורות מקריים בבעיות למידת חיזוק.
**פתרון** הפלט של המערכת חייב להיות פונקציה דטרמניסטית של הקלט (תשובה ב'), אחרת לא נוכל לבצע הליך למידה לאחר סף מסוים (כתלות בשונות). פונקציית הגמול יכולה להיות דטרמניסטית (תשובה ג'), שהרי נוכל להגדירה בעצמנו. תשובה ה', שכן אחרת לא נוכל לנצל את כוחו האמתי של מנגנון למידת החיזוק.

### סעיף ב׳

נניח שמערכת לומדת בעזת למידת חיזוק, קיבלה קלט והוציאה פלט מקרי בהינתן הקלט והפרמטר $w$ וקיבלה גמול שלילי. נבדוק מה אפשרי בהינתן מצב זה.
**פתרון** ערך הזכאות יכול להיות חיובי או שלילי, אך בהנחה שאנו מחשבים באלגוריתם reinforce נקבל שבהתאמה כלל הלמידה יהיה $w$ יהיה שלילי או חיובי (תשובות ב' וג').

## שאלה 1

נדון בחישוב שמבצע טורף במהלך ניסיון לחזות את מיקום טרפו. נסמן את מיקום הטרף $y$ ומחשבת הטורף $\hat{y}$, נניח כי שניהם $\in \mathbb{R}$. נניח ש־$\mathbb{E}(y) =$ $m, \mathrm{Var}(y) = s^2$ עבור ערכים קבועים. נניח ש־$\hat{y}$ משתנה מקרי נורמלי ו־$\hat{y} \sim N(\mu, \sigma^2)$ כאשר $\mu, \sigma$ נלמדים באמצעות אלגוריתם reinforce.
לכל ניסוי הטורף מקבל את התגמול $r = -(\hat{y} - y)^2$.

### סעיף א׳

נגדיר במצב הנתון מה היא המערכת הלומדת, מה הפרמטרים הפנימיים, מה הקלט הפלט והגמול. נבין גם מה מפת התלויות במקרה הנתון.

**פתרון** במקרה זה המערכת הלומדת היא $\hat{y}$ כפונקציה של $\mu, \sigma$, לכן בהתאם הפרמטרים הפנימיים הם $\mu, \sigma$. הקלט הוא $y$ הפלט הוא $\hat{y}$ כערך והתלויות הן $\hat{y}(y)$ כאשר $\sigma, \mu$ נלמדים כחלק מהליך הלמידה ותלויים ב־$s, m$.

### סעיף ב׳

נמצא את כלל העדכון המפורש של $\mu, \sigma$ כתלות ב־$r, \eta$ כפונקציית גמול וקבוע אימון.

**פתרון** באלגוריתם למידת reinforce מוגדר $\Delta w_i = \eta e_i r$ וכן מהגדרה $e_i = \frac{\partial}{\partial w_i} \ln \hat{y}(y; w)$ כאשר $w = (\mu, \sigma)$. אבל נתון $\hat{y} \sim N(\mu, \sigma^2)$ ולכן,

$$\hat{y}(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$$

כפונקציית צפיפות,

$$e_\mu = \frac{\partial}{\partial \mu} \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)\right) = \frac{\partial}{\partial \mu}\left(-\frac{(y-\mu)^2}{2\sigma^2} + \sqrt{2\pi\sigma^2}\right) = -\frac{\partial}{\partial \mu}\frac{(y-\mu)^2}{2\sigma^2} = \frac{y-\mu}{2\sigma^2}$$

וכן,

$$e_\sigma = \frac{\partial}{\partial \sigma}\left(-\frac{(y-\mu)^2}{2\sigma^2} + \sqrt{2\pi\sigma^2}\right) = \frac{(y-\mu)^2}{\sigma^3} + \sqrt{2\pi}\,\mathrm{sign}(\sigma)$$

### סעיף ג׳

נחשב את השינוי הממוצע של $\sigma, \mu$.

**פתרון** ידוע ש־$y \sim N(m, s)$, ומצאנו בסעיף הקודם את השינוי בסעיף הקודם. נחשב,

$$\begin{aligned}
\mathbb{E}(\Delta\mu) &= \eta\mathbb{E}(e_\mu \cdot r) \\
&= -\eta\mathbb{E}(\frac{y-\mu}{2\sigma^2}(\hat{y}-y)^2) \\
&= -\frac{\eta}{2\sigma^2}\mathbb{E}((y-\mu)(\hat{y}^2 - 2y\hat{y} + y^2)) \\
&= -\frac{\eta}{2\sigma^2}(\mathbb{E}(\hat{y}^2 y) - 2\mathbb{E}(y^2\hat{y}) + \mathbb{E}(y^3) - \mu\mathbb{E}(\hat{y}^2) + 2\mu\mathbb{E}(y\hat{y}) - \mu\mathbb{E}(y^2)) \\
&= -\frac{\eta}{2\sigma^2}(\hat{y}^2 m - 2m^2\hat{y} + \mathbb{E}(y^3) - \mu\hat{y}^2 + 2m\mu\hat{y} - \mu\mathbb{E}(y^2)) \\
&= -\frac{\eta}{2\sigma^2}(\sigma^2 m + \mu^2 m - 2m^2\sigma^2 - 2m^2\mu^2 + m^3 + 3\mu\sigma^2 - \mu\sigma^2 - \mu^3 + 2m\mu^2 - \mu m^2 - \mu s^2)
\end{aligned}$$

החישוב עבור $\mathbb{E}(\Delta\sigma)$ דומה.

### סעיף ד׳

נבדוק לאילו ערכים הפרמטרים מתכנסים בממוצע.

**פתרון** אפשר לראות (על־ידי פתרון המשוואה הדיפרנציאלית) שההתכנסות היא $\mu \to m, \sigma \to s$, הגיוני שכן הטורף לומד את התנהגות הטרף.

## שאלה 2

נתון פרספטרון בינארי $y = H(wx)$ עבור $w, x \in \mathbb{R}^N$ כך ש־$y$ משתנה מקרי המקיים,

$$\mathbb{P}(y = 1 \mid x, w) = \frac{1}{1 + \exp(-w^t x)}$$

נסמן $R$ פונקציית גמול על תוצאת $y$ בלמידת חיזוק.

### סעיף א׳

נתאר את חלקי הבעיה.

**פתרון** המערכת הלומדת היא $y(x \mid w)$, הפרמטרים הפנימיים הם $w$, הקלט הוא $x$ הפלט הוא המשתנה המקרי $y$ והגמול הוא $R$. $R$ תלוי ב־$y$, והוא בתורו תלוי ב־$w, x$.

### סעיף ב׳

נחשב את ערך הזכאות כתלות ב־$y$.

**פתרון** כאשר $y = 0$ ועבור $1 \le i \le N$,

$$e_i = \frac{\partial}{\partial w_i} \ln \mathbb{P}(y \mid x, w) = -\frac{\partial}{\partial w_i} \ln(1 + \exp(-w^t x)) = \frac{x_i \exp(-w^t x)}{1 + \exp(-w^t x)}$$

עבור $y = 1$ נקבל $\mathbb{P}(y = 1 \mid x, w) = 1 - \mathbb{P}(y = 0 \mid x, w)$ וכן,

$$e_i = \frac{\partial}{\partial w_i} \ln(\exp(-w^t x)) - \frac{\partial}{\partial w_i} \ln(1 + \exp(-w^t x)) = -x_i + \frac{x_i \exp(-w^t x)}{1 + \exp(-w^t x)}$$

### סעיף ג׳

נמצא ביטוי כללי ל־$e_i$.

**פתרון** מהסעיף הקודם,

$$e_i = -x_i y + \frac{x_i \exp(-w^t x)}{1 + \exp(-w^t x)}$$

### סעיף ד׳

נתון שבניסוי מסוים הגמול הוא $R$, נמצא את כלל העדכון של $w_i$ ונבדוק אם הכלל הוא מקומי.

**פתרון** נתון שקיבלנו את תוצאת ניסוי בודד, ולכן נוכל לבחון את שינוי כלל הלמידה רק בהקשר של למידת אונליין, כלומר ב־reinforce. במקרה זה נתון $\Delta w_i = \eta e_i R$, כלומר,

$$\Delta w_i = \eta R \left( -x_i y + \frac{x_i \exp(-w^t x)}{1 + \exp(-w^t x)} \right) = \eta R x_i \frac{-y + (1 - y) \exp(-w^t x)}{1 + \exp(-w^t x)}$$

נשים לב שלצורך כלל העדכון אנו צריכים את התוצאה, את $R$ ואת תוצאת כלל הנוירונים באותה שכבה, ולכן הוא לא מקומי.

**שאלה 3**