למידה חישובית

תרגיל כיתה מספר 1 – רגרסיה ליניארית ואלגוריתם ה- Gradient Descent

אלגוריתם הגרדיאנט

1. עבור אלגוריתם ה- gradient descent, אם פונקציית המחיר היא:

$$J(\theta) = \frac{1}{2}\theta_1^2 + cos(\theta_2)$$

ידוע כי מקדם הלמידה $\alpha = 0.01$, והערך ההתחלתי של וקטור הפרמטרים הוא:

$$\theta = \begin{pmatrix} 2\pi \\ 0.5\pi \end{pmatrix}$$

אזי הגרדיאנט של $J(\theta)$ הוא: $\nabla J(\theta) = \_\_\_\_\_\_$

והצעד הראשון (האיטרציה הראשונה) של אלגוריתם ה- gradient descent הוא (רשמו את נוסחת העדכון בה השתמשתם):

א. $\theta = \begin{pmatrix} 0.5 \\ 2.5 \end{pmatrix}$

ב. $\theta = \begin{pmatrix} 1.98\pi \\ \pi/2 + 0.01 \end{pmatrix}$

ג. $\theta = \begin{pmatrix} 0.55 \\ 2.625 \end{pmatrix}$

ד. $\theta = \begin{pmatrix} 2.01\pi \\ \pi/2 \end{pmatrix}$

ה. $\theta = \begin{pmatrix} 0.45 \\ 2.375 \end{pmatrix}$

ו. $\theta = \begin{pmatrix} 0.01\pi \\ \pi/2 \end{pmatrix}$

ז. אף תשובה לא נכונה

נוסחת העדכון היא: $\theta = \_\_\_\_\_\_\_\_\_\_\_\_$

2. נתונה פונקציית המחיר $J(\theta): \Re \rightarrow \Re$

$$J(\theta) = \theta^2 + 5\,sin(\theta)$$

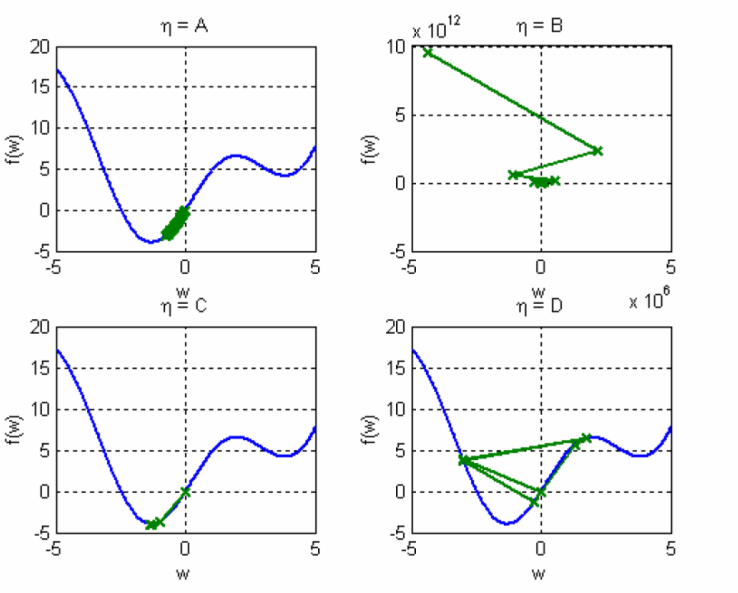
א. מהו תנאי הכרחי לנקודת מינימום?

ב. רשמו את אלגוריתם הגרדיאנט לבעייה זו.

ג. הדגימו שני צעדי לימוד, עבור הערך ההתחלתי $\theta = 0$, וצעד הלימוד $a = 0.2$

ד. הגרפים הבאים מציגים מספר איטרציות של אלגוריתם הגרדיאנט עבור ערכים שונים של צעד הלימוד.

$$\alpha = \ 0.01, 0.2, 0.6, 3$$



התאימו בין צעד הלימוד לגרף .

(מתוך תרגיל בקורס מבוא למערכות לומדות, הטכניון)

ה. ממשו את אלגוריתם הגרדיאנט ב Python עבור בעייה זו וציירו את הגרפים עבור צעדי הלימוד השונים מהסעיף הקודם.

ו. נניח כי הפונקציה היא $J(\theta) = \frac{1}{2}\theta^2 + 5\,sin(\theta)$: , וכן $\theta_0 = 5$ כתבו מה התוצאה ועל איזה בעייה היא עשויה להצביע?

ז. חזרו על התרגיל כאשר : $J(\theta) = 0.3\theta^2 + sin(3\theta)$, וכן $\theta_0 = -5, -2.5, 2$

מהי מסקנתכם?

רגרסיה ליניארית

3. עבור בעיית הרגרסיה הליניארית, נניח כי וקטור התכונות מכיל תכונה אחת בלבד (לדוגמא : שטח הבית בדוגמת מחירי הדירות, זמן ההתפרצות בדוגמת הגייזר הנאמן), כלומר ההיפותזה $h_\theta(x^{(i)})$ היא :

$$h_\theta(x^{(i)}) = \theta_0 + \theta_i x^{(i)},$$

וכן פונקציית המחיר

$$J(\theta) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2 = \frac{1}{2m}\sum_{i=1}^{m}\left(\theta_0 + \theta_i x^{(i)} - y^{(i)}\right)^2$$

חשבו מהם θ_0 ו- θ_1 האופטימליים הממזערים את $J\left(\theta\right)$.

4. Piarce (1948) מדד את תדירות הצרצור של צרצרי קרקע (striped ground cricket, מספר תנודות כנפיים לשניה או פולסי קול לשניה). וכן את טמפ׳ הקרקע (ראו טבלה 1). מאחר וצרצרים הם בעלי חיים אקזותרמיים (בעלי דם קר) קיים בסיס להשערה כי הפעילות הפיזיולוגית שלהם תהיה תלוייה בטמפ׳ החיצונית, ולכן לכך קשר בין תדירות התנודות לבין הטמפ׳.

באופן כללי נמצא כי הצרצרים אינם משמיעים קול בטמפ׳ הנמוכה מ- 60 מעלות או גבוהה מ- 100 מעלות פרנהייט (15.5 ו-37 מעלות צלזיוס בהתאמה).

בתרגיל זה נניח כי קיים קשר לינארי בין התדירות לבין הטמפ׳.

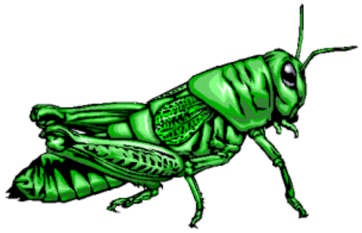
א. ציירו את הנתונים באמצעות ה- Python (ראו קובץ Xcricket.mat בתרגיל כיתה מספר 1 במודל).

ב. חשבו את הפרמטרים המתאימים והתאימו עקומה ליניארית לנתונים באמצעות חישוב אנליטי.

ג. ממשו את אלגוריתם ה- Gradient Descent וחשבו את המקדמים. השוו למקדמים אותם קיבלתם בסעיף ב׳.

ד. מהי תדירות הצרצור הצפויה עבור טמפ׳ של 95 מעלות? ועבור 65 מעלות פרנהייט?

ה. מה התדירות הצפויה עבור טמפ׳ של 32 מעלות פרנהייט (קיפאון).



Temperature (º F)	Chirps/Second
88.6	20.0
71.6	16.0
93.3	19.8
84.3	18.4
80.6	17.1
75.2	15.5
69.7	14.7
71.6	15.7
69.4	15.4
83.3	16.3
79.6	15.0
82.6	17.2
80.6	16.0
83.5	17.0
76.3	14.4

טבלה 1: מתוך *The Song of Insects* by George W. Pierce, 1948, page 20