

תעבורה אקולוגית - תרגיל 6

לשם קן זכור

316163260

1. בלע

$$\Delta G = \eta \cdot R \cdot e_G \quad , \quad g(y; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(y-\mu)^2}{2\sigma^2}} \quad , \quad y \sim N(\mu, \sigma^2) \quad \text{מק"מ} \quad (1)$$

$$\Delta G = \eta \cdot R \cdot \frac{\partial \log g(y; \mu, \sigma)}{\partial \sigma}$$

שאלה: מהי הנגזרת?

$$\frac{\partial \log g(y; \mu, \sigma)}{\partial \sigma} = \frac{\partial \left(\log \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{(y-\mu)^2}{2\sigma^2} \right)}{\partial \sigma} = \frac{\partial \left(-\log(\sqrt{2\pi\sigma^2}) - \frac{(y-\mu)^2}{2\sigma^2} \right)}{\partial \sigma}$$

הנגזרת היא הנגזרת

$$\frac{\partial}{\partial \sigma} \left(-\frac{1}{2} \log(2\pi\sigma^2) - \frac{(y-\mu)^2}{2\sigma^2} \right) = -\frac{1}{\sigma} + \frac{(y-\mu)^2}{\sigma^3}$$

$$\Delta G = \eta \cdot R \cdot \left(-\frac{1}{\sigma} + \frac{(y-\mu)^2}{\sigma^3} \right) \quad \text{שאלה}$$

הסבר: אילוץ: אילוץ: יש שני מקרים: (I) $(y-\mu)^2 < \sigma^2$ כלומר $R < 0$ כלומר

אנחנו רוצים שהשונות תהיה קטנה יותר מ-1. כלומר $R > 0$ כלומר אנחנו רוצים שהשונות תהיה גדולה יותר מ-1.

(II) $(y-\mu)^2 > \sigma^2$ כלומר $R < 0$ כלומר אנחנו רוצים שהשונות תהיה קטנה יותר מ-1.

כלומר $R > 0$ כלומר אנחנו רוצים שהשונות תהיה גדולה יותר מ-1.

$$\Delta G = \eta \cdot (-(y-m)^2) \cdot \left(\frac{(y-m)^2}{\sigma^3} - \frac{1}{\sigma} \right) \quad \text{זכור} \quad E[\Delta G] \quad \text{הערות: 3.1.7} \quad (2)$$

REINFORCE או כוונה של

$$\frac{\partial}{\partial \sigma} E[R] = E \left[R \cdot \frac{\partial \log g(y, \mu, \sigma)}{\partial \sigma} \right] = \frac{1}{\eta} E[\Delta G]$$

(זכור 3.1.7)

$E[R]$ זכור

$$E[R] = -\mu^2 - \sigma^2 + 2m\mu - m^2$$

↑
זכור 3.1.7

זכור

$$\frac{\partial E[R]}{\partial \sigma} = -2\sigma$$

$$\frac{1}{\eta} E[\Delta G] = -2\sigma$$

זכור

$$\boxed{E[\Delta G] = -2\eta\sigma}$$

כדי להקטין את הריבועים, קטנו את הממוצע, נרצה להקטין את הממוצע

$$E[\Delta G] = 0$$

$$\left(E[\Delta G] = -\eta G \quad \text{כ} \quad G=0 \right)$$

זה אומר שרפסיק לא מיוז כאשר קשורה היא 0, וזה אומר שפנץ באותה נקודה, והשונו
 לא תשתנה יתר, אומר נפנץ בנקודה הזו "לנצח", אומר מנצחן אל המסרה,
 זה קורה כי מנצחן שהתחל שאלו היא המסרה (בתחילתו) ואם העלו אליה נצח אהישאר
 בו = שהשונו תישאר 0.

$$\frac{d}{dt} \left[\frac{E[\Delta G]}{G} \right] = \frac{E[\Delta G]}{G^2} \cdot \frac{dG}{dt} = \frac{E[\Delta G]}{G^2} \cdot (-\eta G) = -\frac{\eta E[\Delta G]}{G}$$

$$\frac{d}{dt} \left[\frac{E[\Delta G]}{G} \right] = \frac{E[\Delta G]}{G^2} \cdot \frac{dG}{dt} = \frac{E[\Delta G]}{G^2} \cdot (-\eta G) = -\frac{\eta E[\Delta G]}{G}$$

$$\frac{d}{dt} \left[\frac{E[\Delta G]}{G} \right] = \frac{E[\Delta G]}{G^2} \cdot \frac{dG}{dt} = \frac{E[\Delta G]}{G^2} \cdot (-\eta G) = -\frac{\eta E[\Delta G]}{G}$$

$$\frac{d}{dt} \left[\frac{E[\Delta G]}{G} \right] = \frac{E[\Delta G]}{G^2} \cdot \frac{dG}{dt} = \frac{E[\Delta G]}{G^2} \cdot (-\eta G) = -\frac{\eta E[\Delta G]}{G}$$

$$\frac{d}{dt} \left[\frac{E[\Delta G]}{G} \right] = \frac{E[\Delta G]}{G^2} \cdot \frac{dG}{dt} = \frac{E[\Delta G]}{G^2} \cdot (-\eta G) = -\frac{\eta E[\Delta G]}{G}$$

2. a/c

$y=1$ 1st (1)

$$\frac{\partial \log P(y|x; w)}{\partial w_i} = \frac{\partial \log P(y=1|x; w)}{\partial w_i} =$$

$$\frac{\partial \log \left(\frac{1}{1+e^{-w^T x}} \right)}{\partial w_i} = \frac{\partial \left(-\log(1+e^{-w^T x}) \right)}{\partial w_i} =$$

$$\frac{-1}{1+e^{-w^T x}} \cdot e^{-w^T x} \cdot (-x_i) = \boxed{\frac{x_i \cdot e^{-w^T x}}{1+e^{-w^T x}}}$$

$y=0$ 2nd

$$\frac{\partial \log P(y|x; w)}{\partial w_i} = \frac{\partial \log P(y=0|x; w)}{\partial w_i} = \frac{\partial \log \left(\frac{e^{-w^T x}}{1+e^{-w^T x}} \right)}{\partial w_i} =$$

$$\frac{\partial (\log(e^{-w^T x}) - \log(1+e^{-w^T x}))}{\partial w_i} = \frac{\partial (-w^T x - \log(1+e^{-w^T x}))}{\partial w_i} =$$

$$\boxed{-x_i + \frac{x_i \cdot e^{-w^T x}}{1+e^{-w^T x}}}$$

②) נגזרת של $y=0$ נגזרת:

~~$$e_i = -x_i \cdot \frac{x_i \cdot e^{-w_i x_i}}{1 + e^{-w_i x_i}}$$~~

$$e_i = -x_i \cdot \frac{x_i \cdot e^{-w_i x_i}}{1 + e^{-w_i x_i}} = -x_i \left(1 - \frac{e^{-w_i x_i}}{1 + e^{-w_i x_i}} \right)$$

נגזרת של $y=1$ נגזרת:

$$e_i = -x_i (P[y=1|x;w])$$

נגזרת של $y=1$ נגזרת:

$$e_i = x_i (1 - P[y=1|x;w])$$

נגזרת של $y=1$ נגזרת:

$$e_i = x_i (y - P[y=1|x;w])$$

③) נגזרת של w_i הוא:

$$\Delta w_i = \eta \cdot R \cdot x_i (y - P[y=1])$$

נגזרת של w_i הוא Δw_i נגזרת של w_i נגזרת של x_i הנגזרת, η הוא גודל הלימוד, R הוא גודל המדגם, x_i הוא המאפיין, y הוא התוצאה, $P[y=1]$ הוא ההסתברות.

שאלה 3

5

6 W ו- W הוא הקולט כמה X_i תנים אציקוי הסברה.
הפיקס'ים שיקצוק נמצאים קמנד, פלומי אלה לנכים של W שמנצאים במנד, והם קילאו
אין גרוב כי הם מתאימים לזורה של הסברה 1.
הפיקס'ים שנקח, שאלם אחר החוק הקצוק, הם לנכים של W שקילאו אין נמוך, כי הם
מתאימים לזורה של הסברה 0.

$$P[y=1/x;w] = \frac{1}{1+e^{-w^T x}}$$

זה מתאים להסתברות אקראית הסברה 1:

כי אם נציג לנכים גרובים של W אז ההסתברות אקראית 1 תהיה.