מבוא למכונות לומדות - תרגיל 4

רפאל שקד גרינפלד - 305030868

11 ביוני 2018

1. צ"ל: $\frac{Z_{t+1}}{Z_t}=2\sqrt{\varepsilon_{t+1}\left(1-\varepsilon_{t+1}\right)}$. ראשית, נשים לב כי

$$D_i^{(t+1)}=\frac{D_i^{(t)}\exp\left(-y_iw_th_t\left(x_i\right)\right)}{\sum_{j=1}^{m}D_j^{(t)}\exp\left(-y_jw_th_t\left(x_j\right)\right)}=\frac{\frac{1}{m}\prod_{k=1}^{t}\exp\left(-y_iw_kh_k\left(x_i\right)\right)}{\underbrace{\prod_{k=1}^{t}\sum_{j=1}^{m}D_j^{(k)}\exp\left(-y_jw_kh_k\left(x_j\right)\right)}_{\text{we will mark this as K}}}=$$

$$=\frac{\exp\left(\sum_{k=1}^{t}-y_iw_kh_k\left(x_i\right)\right)}{m\cdot K}=\frac{\exp\left(-y_i\sum_{k=1}^{t}w_kh_k\left(x_i\right)\right)}{m\cdot K}=\frac{\exp\left(-y_if_t\left(x_i\right)\right)}{m\cdot K}$$

ולכן

$$\exp\left(-y_if_t\left(x_i\right)\right)=D_i^{(t+1)}\cdot m\cdot K$$

בנוסף,

$$f_{t+1}=\sum_{k=1}^{t+1}w_kh_k=\sum_{k=1}^{t}w_kh_k+w_{t+1}h_{t+1}=f_t+w_{t+1}h_{t+1}$$

כעת

$$\frac{Z_{t+1}}{Z_t}=\frac{\frac{1}{m}\sum_{i=1}^{m}\exp\left(-y_i f_{t+1}\left(x_i\right)\right)}{\frac{1}{m}\sum_{i=1}^{m}\exp\left(-y_i f_t\left(x_i\right)\right)}=\frac{\sum_{i=1}^{m}\exp\left(-y_i\left(f_t\left(x_i\right)+w_{t+1}h_{t+1}\left(x_i\right)\right)\right)}{\sum_{i=1}^{m}\exp\left(-y_i f_t\left(x_i\right)\right)}=$$
$$=\frac{\sum_{i=1}^{m}D_i^{(t+1)}\cdot m\cdot K\exp\left(-y_i w_{t+1}h_{t+1}\left(x_i\right)\right)}{\sum_{i=1}^{m}D_i^{(t+1)}\cdot m\cdot K}=\sum_{i=1}^{m}D_i^{(t+1)}\exp\left(-y_i w_{t+1}h_{t+1}\left(x_i\right)\right)$$
$$=\sum_{i=1}^{m}D_i^{(t+1)}\exp\left(-w_{t+1}\right)1_{[y_i=h_{t+1}(x_i)]}+\sum_{i=1}^{m}D_i^{(t+1)}\exp\left(w_{t+1}\right)1_{[y_i\neq h_{t+1}(x_i)]}$$
$$=\exp\left(-\frac{1}{2}\log\left(\frac{1}{\varepsilon_{t+1}}-1\right)\right)\sum_{i=1}^{m}D_i^{(t+1)}1_{[y_i=h_{t+1}(x_i)]}+\exp\left(\frac{1}{2}\log\left(\frac{1}{\varepsilon_{t+1}}-1\right)\right)\sum_{i=1}^{m}D_i^{(t+1)}1_{[y_i\neq h_{t+1}(x_i)]}$$
$$=\left(\frac{1}{\varepsilon_{t+1}}-1\right)^{-\frac{1}{2}}\cdot\left(1-\varepsilon_{t+1}\right)+\left(\frac{1}{\varepsilon_{t+1}}-1\right)^{\frac{1}{2}}\varepsilon_{t+1}$$
$$=\sqrt{\frac{\varepsilon_{t+1}}{1-\varepsilon_{t+1}}}\left(1-\varepsilon_{t+1}\right)+\sqrt{\frac{1-\varepsilon_{t+1}}{\varepsilon_{t+1}}}\varepsilon_{t+1}=2\sqrt{\varepsilon_{t+1}\left(1-\varepsilon_{t+1}\right)}$$

2. (א) תהי $h\in L\left(DS,T\right)$, $h=\operatorname{sign}\left(\sum_{t=1}^{T}w_t h_t\left(x\right)\right)$. לכל t, יהי $j_t\in[d]$ נקודת הפיצול של $h_t=\left(x_{j_t}-\theta_t\right)\cdot b_t$. נגדיר עץ $tree\left(x\right)$ מאוזן בעומק T כך שברמה t, נקודת הפיצול של העץ תהיה $x_{j_t}\geq\theta_t$. אז נשים לב כי בהינתן עלה l, לכל $x_1,x_2\in\mathbb{R}$ שמגיעים ב־$tree$ ל־l, מתקיים לכל t

$$\operatorname{sign}\left(x_{1j_t}-\theta_t\right)=\operatorname{sign}\left(x_{2j_t}-\theta_t\right)$$

בפרט נובע כי

$$h_t\left(x_1\right)=h_t\left(x_2\right)$$

ולכן

$$h\left(x_1\right)=\operatorname{sign}\sum_{t=1}^{T}w_t h_t\left(x_1\right)=\operatorname{sign}\sum_{t=1}^{T}w_t h_t\left(x_2\right)=h\left(x_2\right)$$

לכן נקבע כי $tree\left(x_1\right)=h\left(x_1\right)$ ונקבל שזה אכן מוגדר היטב.

(ב) נגדיר את העץ הבא: שעבור $X_i \leq 1$ מחזיר 1 אחרת גם 1. (עץ שמדמה היפוטזה נאיבית בצורה של עץ...)

נטען כי אין היפותזה ב־$L\left(DS,T\right)$ שמסווגת כמוהו לאף T. יהי T כלשהו ויהי $h \in L\left(DS,T\right)$.

נגדיר $sign(0) = 1$ ונניח בשלילה כי $\exists t \in \mathbb{N} :: t \in L\left(DS,T\right)$ כעת, נסדר את Θ_i לפי הקורדינטות המכריעות

ונראה כי קיים $X \in \mathbb{R} :: t(x) = -1$ והיות והעץ תמיד חיובי נסיק כי קיימת סתירה.

$$\forall x_i \in X$$

נבחר את Θבקור' Θ_i כך שמשקלה בהיפוטזה DS מקיים שהוא המקסימלי עבור Θ ואזי

$$if\left(b=1\right):we\,will\,define\left(X_i<\Theta\right)$$

$$if\left(b=-1\right):we\,will\,define\left(X_i>\Theta\right)$$

ואזי נקבל שלכל קורדינטה i בX סכום ההיפוטזות שמשתמשות בהיפוטזה יהיה בהכרח שלילי ולכן

$$t(x) = sign(\sum_{i=1}^{T} h_i(X)W_i) = -1$$

בסתירה.