دانشگاه صنعتی شریف
دانشکده‌ی مهندسی کامپیوتر

هوش مصنوعی

پاییز ۱۳۹۹
استاد: محمدحسین رهبان

تمرین دوم، بخش دوم

پاسخ سوالات نظری

۱. (۳۰+۱۰ نمره)

(آ) نیست

با استفاده از مشتق دوم، محدب بودن این تابع بررسی می‌شود. با توجه به این که در بازه‌ی $-\frac{۳}{۱۰}$ تا ۰ مشتق دوم منفی است، تابع محدب نیست.

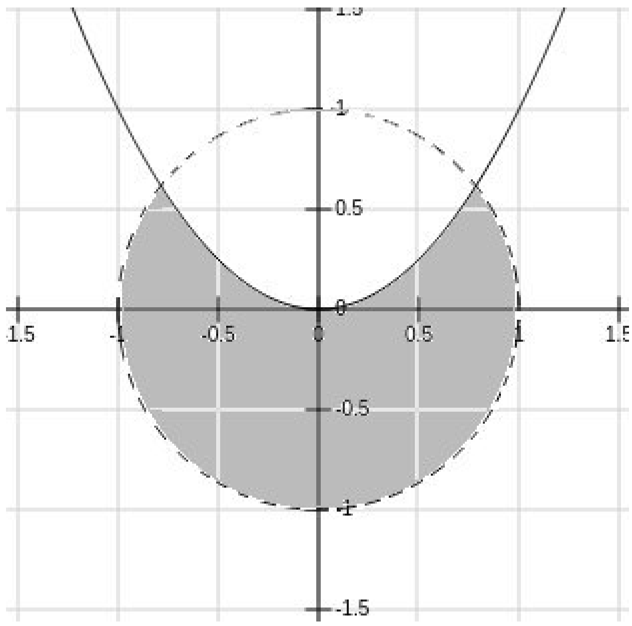
$$f''(x) = ۶۰x^{۲} + ۱۸x, \quad -\frac{۳}{۱۰} < x < ۰ \Rightarrow f''(x) < ۰$$

(ب) نیست

با یک مثال نقض می‌توان محدب نبودن این تابع را نشان داد. بایستی دو نقطه از مجموعه‌ی S را یافت که خط واصل آن‌ها به طور کامل درون مجموعه S قرار نمی‌گیرد.

$$\left.\begin{array}{l} x_{۱} : (-\frac{\sqrt{۲}}{۲}, \frac{۱}{۲}) \in S \\ x_{۲} : (\frac{\sqrt{۲}}{۲}, \frac{۱}{۲}) \in S \end{array}\right\} \Rightarrow \frac{x_{۱} + x_{۲}}{۲} : (۰, \frac{۱}{۲}) \notin S$$

همچنین شکل زیر نشان‌دهنده محدب‌نبودن مجموعه S می‌باشد.



(ج) هست
بایستی ثابت شود هر نقطه روی خط واصل بین دو نقطه از مجموعه‌ی S نیز در مجموعه موردنظر قرار می‌گیرد.

$$\left.\begin{array}{l}\forall\alpha : ۰ \leq \alpha \leq ۱ \\ \forall x : x \in S \\ \forall y : y \in S\end{array}\right\} \Rightarrow \left.\begin{array}{l}\alpha \sum_i |x_i| \leq \alpha \\ (۱-\alpha) \sum_i |y_i| \leq ۱-\alpha\end{array}\right\} \stackrel{+}{\Rightarrow} \sum_i |\alpha x_i| + |(۱-\alpha)y_i| \leq ۱$$

$$\xrightarrow{|a+b| \leq |a|+|b|} \sum_i |\alpha x_i + (۱-\alpha)y_i| \leq ۱ \Rightarrow \alpha x + (۱-\alpha)y \in S$$

هر نقطه‌ای که روی خط بین x و y وجود دارد نیز جزء مجموعه S می‌باشد.

(د) (امتیازی) هست
با توجه به این که A طبق صورت سوال Positive semi-definite است، تابع Hessian همواره مثبت بوده و تابع محدب است.

$$\nabla f = ۲Ax \Rightarrow \nabla^۲ f = ۲A \geq ۰$$

۲. (۲۰ نمره)

(آ)

$$W = \begin{bmatrix} w_۲ \\ w_۱ \\ w_۰ \end{bmatrix}, \; X^{(i)} = \begin{bmatrix} {x_i}^۲ \\ x_i \\ ۱ \end{bmatrix} \Rightarrow f(X;W) = \sum_i (W^T X^{(i)} - y^{(i)})^۲$$

$$\frac{\partial f}{\partial W} = \sum_i ۲(W^T X^{(i)} - y^{(i)}) X^{(i)}$$

(ب)

$$W_{i+۱} = W_i - \alpha \frac{\partial f}{\partial W_i}$$

با افزایش آلفا، سرعت حرکت به سمت مینیمم افزایش پیدا می‌کند، اما در صورتی که به مینیمم نزدیک باشد، زیاد بودن مقدار آلفا، ممکن است باعث به وجود آمدن حرکات زیگزاگی حول مینیمم می‌شود. از طرفی کم بودن مقدار آلفا باعث کاهش سرعت می‌شود. بنابراین بهتر است در ابتدا با مقدار آلفای بزرگ اجرای الگوریتم را آغاز کرده و با نزدیک شدن به مینیمم، آن را کاهش داد.