CSE 250B Winter 2019 - Quiz 4

Feb 15, 2019 - 20 Minutes **Name:** **PID:**

---

(1) This is a closed book, closed notes exam. Switch off your cell phone and do not communicate with anyone other than an exam proctor.
(2) Start writing when instructed. Stop writing when your time is up.
(3) Remember that your work is graded on the quality of your writing and explanation as well as the validity of the mathematics.

You are given below two functions f and g that map d-dimensional vectors x into scalars. For each of these functions, calculate the gradient and the Hessian. Recall that a function is convex if its Hessian is positive semi-definite at all inputs. Use the Hessian to determine whether each function is convex, and justify your answer.

(1) (5 points) $f(x) = e^{-\frac{1}{2}x^\top x}$.
$\nabla f(x) = -x \cdot e^{-\frac{1}{2}x^\top x}$.
$\nabla^2 f(x) = (xx^\top - I) \cdot e^{-\frac{1}{2}x^\top x}$.
$f(x)$ is not convex as the Hessian is not PSD at all x. In particular, if $x = [0, 0, \ldots, 0]$, then the Hessian at x is the matrix $diag(-1, -1, -1, \ldots, -1)$, which is not PSD – if $z = [1, 0, \ldots, 0]$, then $z^\top \nabla^2 f(x) z = -1 < 0$.

(2) (5 Points) Suppose $z^{(i)} \in \mathbb{R}^d$, for $i = 1, \ldots, n$. $g(x) = \sum_{i=1}^n (e^{x^\top z^{(i)}} - x^\top z^{(i)})$.
$\nabla g(x) = \sum_{i=1}^n z^{(i)} \cdot e^{x^\top z^{(i)}} - z^{(i)}$.
$\nabla^2 g(x) = \sum_{i=1}^n z^{(i)} \cdot (z^{(i)})^\top \cdot e^{x^\top z^{(i)}}$.

g is convex as the Hessian is PSD at all x. We prove it as follows.

For any x and $z^{(i)}$, $e^{x^\top z^{(i)}} > 0$; moreover, $z^{(i)} \cdot (z^{(i)})^\top$ is PSD – as for any vector $w \in \mathbb{R}^d$, we have $w^\top z^{(i)} \cdot (z^{(i)})^\top w = \|w^\top z^{(i)}\|^2 \geq 0$.

If c_i are scalars that are > 0 and if A_i are PSD matrices, then $\sum_i c_i A_i$ is also PSD; this is because for any vector w, $w^\top (\sum_i c_i A_i) w = \sum_i c_i w^\top A_i w \geq 0$ as each individual term $c_i w^\top A_i w \geq 0$. Plugging in $c_i = e^{x^\top z^{(i)}}$, and $A_i = z^{(i)} \cdot (z^{(i)})^\top$, we get that the Hessian is PSD at all x.