# Problem 1 : Gradients & More

1.

a) This function is convex because according to operations which preserve convexity, the pointwise maximum of convex functions are convex. If for $w \in \mathbb{R}^n$ and fixed $y_1, y_M \in \mathbb{R}$ and $x^M \in \mathbb{R}^n$ are all convex, then $f(w)$ will be convex.

$$\frac{\partial}{\partial w}((-yw^Tx)^2) = 2(-yw^Tx)\frac{\partial}{\partial w}(-yw^Tx)$$
$$= 2(-yw^Tx)(-yxTw^{T-1}) = 2y^2Tx^2w^{2T-1}$$

c) This function is convex because as stated in the powerpoint mentioning operations preserving convexity, when discussing the pointwise maximum of convex functions, the pointwise maximum of convex functions are convex if $[a^{(m)T}x + b_m]$ for $x \in \mathbb{R}^n$ and constants $a^{(1)}, \ldots, a^{(m)} \in \mathbb{R}^n$ and $b_1, \ldots, b_m \in \mathbb{R}$ are all convex.

$$\frac{\partial}{\partial a}(a^Tx + b) = Txa^{T-1}, \quad \frac{\partial}{\partial b}(a^Tx+b) = 1$$

b) This function is convex because with $f(x) = \log[\sum_{i=1}^n \exp(a_ix_i)]$ for $x \in \mathbb{R}^n$ and constants $a \in \mathbb{R}^n$, we have $\nabla^2 f(x) = \frac{1}{1^Tz}\operatorname{diag}(z) - \frac{1}{(1^Tz)^2}zz^T \quad (z_i = \exp a_ix_i)$

To show $\nabla^2 f(x) \geq 0$, we need to verify $v^T\nabla^2 f(x)v \geq 0$ for all $v$:

$$v^T\nabla^2 f(x)v = \frac{(\sum_i z_iv_i^2)(\sum_i z_i) - (\sum_i v_iz_i)^2}{(\sum_i z_i)^2} \geq 0$$

since $(\sum_i v_iz_i)^2 \leq (\sum_i z_iv_i^2)(\sum_i z_i)$, $\frac{\partial}{\partial a}(\log_{10}(e^{a_ix_i})) = \frac{\partial}{\partial a}\left(\frac{\ln(e^{ax})}{\ln(10)}\right)$

Scanned with CamScanner

$$\pm \frac{x}{\ln(10)}$$

Problem 3: Numerical Gradients

1.

a) $\frac{\partial}{\partial a}\left[\frac{\exp(a)}{1+\exp(a)} \ln\left(\frac{\exp(a)}{1+\exp(a)}\right) + \frac{1}{1+\exp(a)} \ln\left(\frac{1}{1+\exp(a)}\right)\right]$

$$\frac{\partial}{\partial a}\left(\frac{e^a}{1+e^a} \ln\left(\frac{e^a}{1+e^a}\right)\right) = \frac{e^a(a - \ln(1+e^a)) + e^a}{(e^a+1)^2}$$

$$\frac{\partial}{\partial a}\left(\frac{1}{1+e^a} \ln\left(\frac{1}{1+e^a}\right)\right) = \frac{e^a \ln(1+e^a) - e^a}{(1+e^a)^2}$$

$$\frac{e^a(a-\ln(1+e^a))+e^a}{(e^a+1)^2} + \frac{e^a \ln(1+e^a) - e^a}{(1+e^a)^2} : \frac{ae^a}{(e^a+1)}$$

$\frac{ae^a}{(e^a+1)} > 0$, thus this is a convex optimization problem

Scanned with CamScanner