CSI436/536 Gabriel Falcones

Homework 1

1. $f(x,y) = \dfrac{1}{1+e^{-x-2y}}$

a. Gradient

$$\frac{df(x,y)}{dx} = \frac{-1}{(1+e^{-x-2y})^2} \cdot \left(e^{-x-2y}\right) \cdot (-1) = \frac{e^{-x-2y}}{(1+e^{-x-2y})^2}$$

$$\frac{df(x,y)}{dy} = \frac{-1}{(1+e^{-x-2y})^2} \left(e^{-x-2y}\right) \cdot (-2) = \frac{2e^{-x-2y}}{(1+e^{-x-2y})^2}$$

$$\nabla f(x,y) = \begin{bmatrix} \dfrac{e^{-x-2y}}{(1+e^{-x-2y})^2} \\ \dfrac{2e^{-x-2y}}{(1+e^{-x-2y})^2} \end{bmatrix}$$

b) Hessian matrix

$$* \frac{d^2f(x,y)}{dx^2} = \frac{df'(x,y)}{dx} = \frac{d\left(e^{-x-2y}/(1+e^{-x-2y})^2\right)}{dx}$$

$$g(x,y) = e^{-x-2y}$$

$$h(x,y) = (1+e^{-x-2y})^2$$

$$\frac{df(x,y)}{dx^2} = \frac{\frac{dg(x,y)}{dx} \cdot h(x,y) - \frac{dh(x,y)}{dx} \cdot g(x,y)}{(h(x,y))^2}$$

$$\frac{dg(x,y)}{dx} = -e^{-x-2y}$$

$$\frac{dh(x,y)}{dx} = 2(1+e^{-x-2y})(e^{-x-2y})(-1)$$

$$= -2e^{-x-2y}(1+e^{-x-2y})$$

$$\frac{df(x,y)}{dx^2} = \frac{(-e^{-x-2y})(1+e^{-x-2y})^{\cancel{2}} - (-2e^{-x-2y})\cancel{(1+e^{-x-2y})}(e^{-x-2y})}{(1+e^{-x-2y})^{\cancel{4}\,3}}$$

$$\frac{\partial f(x,y)}{\partial x^2} = \frac{(2e^{-x-2y})(e^{-x-2y}) - e^{-x-2y}(1+e^{-x-2y})}{(1+e^{-x-2y})^3} = \frac{e^{-x-2y}(2e^{-x-2y} - 1 - e^{-x-2y})}{(1+e^{-x-2y})^3} = \frac{e^{-x-2y}(e^{-x-2y}-1)}{(1+e^{-x-2y})^3}$$

$$* \frac{d^2 f(x,y)}{dy\,dx} = \frac{\partial\left(e^{-x-2y}/(1+e^{-x-2y})^2\right)}{\partial y}$$

$g(x,y) = e^{-x-2y}$

$h(x,y) = (1+e^{-x-2y})^2$

$$\frac{d^2 f(x,y)}{\delta y\,\delta x} = \frac{\frac{\delta g(x,y)}{\delta y} \cdot h(x,y) - \frac{\delta h(x,y)}{\delta y} \cdot g(x,y)}{h(x,y)^2}$$

$$\frac{\partial g(x,y)}{\partial y} = -2e^{-x-2y}$$

$$\frac{\partial h(x,y)}{\partial y} = -4e^{-x-2y}(1+e^{-x-2y})$$

$$\frac{\delta f(x,y)}{\delta y\,\delta x} = \frac{(-2e^{-x-2y})(1+e^{-x-2y})^{\cancel{2}} - (-4e^{-x-2y})\cancel{(1+e^{-x-2y})}(e^{-x-2y})}{(1+e^{-x-2y})^{\cancel{4}\,3}}$$

$$\frac{\delta f(x,y)}{\delta y\,\delta x} = \frac{(4e^{-x-2y})(e^{-x-2y}) - 2e^{-x-2y}(1+e^{-x-2y})}{(1+e^{-x-2y})^3}$$

$$\frac{\delta f(x,y)}{\delta y\,\delta x} = \frac{2e^{-x-2y}(2e^{-x-2y} - 1 - e^{-x-2y})}{(1+e^{-x-2y})^3} = \frac{2e^{-x-2y}(e^{-x-2y} - 1)}{(1+e^{-x-2y})^3} \quad //$$

$$* \frac{\delta f(x,y)}{\delta y^2} = \frac{\delta\left(\frac{2e^{-x-2y}}{(1+e^{-x-2y})^2}\right)}{\delta y} = \frac{2\,\delta\left(\frac{e^{-x-2y}}{(1+e^{-x-2y})^2}\right)}{\delta y} = 2\,\frac{\delta f(x,y)}{\delta y\,\delta x}$$

$$\frac{\delta f(x,y)}{\delta y^2} = \frac{4e^{-x-2y}(e^{-x-2y} - 1)}{(1+e^{-x-2y})^3} \quad //$$

$$* \frac{\delta f(x,y)}{\delta x\,\delta y} = \frac{\delta\left(2e^{-x-2y}/(1+e^{-x-2y})^2\right)}{\delta x} = \frac{2 \cdot \delta\left(e^{-x-2y}/(1+e^{-x-2y})^2\right)}{\delta x}$$

$$= 2\,\frac{\delta f(x,y)}{\delta x^2}$$

$$= \frac{2e^{-x-2y}(e^{-x-2y} - 1)}{(1+e^{-x-2y})^3} \quad //$$

$$H = \begin{bmatrix} \frac{\delta^2 f(x,y)}{\delta x^2} & \frac{\delta^2 f(x,y)}{\delta x \delta y} \\ \frac{\delta^2 f(x,y)}{\delta y \delta x} & \frac{\delta^2 f(x,y)}{\delta y^2} \end{bmatrix}$$

$$H = \begin{bmatrix} \frac{e^{-x-2y}(e^{-x-2y}-1)}{(1+e^{-x-2y})^3} & \frac{2e^{-x-2y}(e^{-y-2y}-1)}{(1+e^{-x-2y})^3} \\ \frac{2e^{-x-2y}(e^{-x-2y}-1)}{(1+e^{-x-2y})^3} & \frac{4e^{-x-2y}(e^{-x-2y}-1)}{(1+e^{-x-2y})^3} \end{bmatrix}$$

c. Is this function a convex function, why?

$$\text{determinant}(H) = \frac{\delta^2 f(x,y)}{\delta x^2} \cdot \frac{\delta^2 f(x,y)}{\delta y^2} - \left(\frac{\delta^2 f(x,y)}{\delta x \delta y}\right)^2$$

$$= 4\left(\frac{e^{-x-2y}(e^{-x2y}-1)}{(1+e^{-x-2y})^3}\right)^2 - \left(\frac{2e^{-x-2y}(e^{-x-2y}-1)}{(1+e^{-x-2y})^3}\right)^2$$

$$= 4\left(\frac{e^{-x-2y}(e^{-x-2y}-1)}{(1+e^{-x-2y})^3}\right)^2 - 4\left(\frac{e^{-x-2y}(e^{-x-2y}-1)}{(1+e^{-x-2y})^3}\right)^2$$

$\det(H) = 0 \Rightarrow$ <u>Indefinite matrix</u>, it's not convex

2. Consider $f(x) = \frac{1}{1+e^{-w^T x}}$

a. Compute its gradient

$$\nabla f = \left(\frac{-1}{(1+e^{-w^T x})^2}\right) \cdot \left(e^{-w^T x}\right)(-w)$$ // Chain rule $\frac{\delta}{\delta u}\left(\frac{1}{u}\right) \cdot \frac{\delta}{\delta v}(e^{v}) \cdot \frac{\delta}{\delta x}(-w^T x)$

$U = 1+e^{-w^T x}$

$V = -w^T x$

$$\nabla f = \left(\frac{e^{-w^T x}}{(1+e^{-w^T x})^2}\right) w$$

b. compute Hessian matrix

$$H = \frac{\delta\left(\frac{e^{-w^T x}}{(1+e^{-w^T x})^2}\right) \cdot w^T}{\delta x}$$

$g(x) = e^{-w^T x}$ $\qquad \frac{\delta g(x)}{\delta x} = -e^{-w^T x} w$

$h(x) = (1+e^{-w^T x})^2$ $\qquad \frac{\delta h(x)}{\delta x} = 2(1+e^{-w^T x})(e^{-w^T x})(-w) = -2(1+e^{-w^T x})(e^{-w^T x}) w$

$$H = \left[\frac{g'(x)h(x) - g(x)h'(x)}{h(x)^2}\right] \cdot w^T = \left[\frac{-e^{-w^T x} w (1+e^{-w^T x})^{\cancel{2}} - e^{-w^T x}(-2)\cancel{(1+e^{-w^T x})}(e^{-w^T x}) w}{(1+e^{-w^T x})^{\cancel{4}\,3}}\right] \cdot w^T$$

$$H = \left[\frac{-e^{-w^T x}(1+e^{-w^T x}) w + 2e^{-w^T x}(e^{-w^T x}) w}{(1+e^{-w^T x})^3}\right] \cdot w^T$$

$$H = \left[\frac{e^{-w^T x}(-1 - e^{-w^T x} + 2e^{-w^T x}) w}{(1+e^{-w^T x})^3}\right] \cdot w^T$$

$$H = \frac{e^{-w^T x}(e^{-w^T x} - 1)}{(1+e^{-w^T x})^3} w w^T$$ //

3. Consider quadratic function

$f(x,y,z) = 2x^2 + 3y^2 - z^2 + 3xy - 4xz + 6yz - 2x + 4y - 7z + 4$

equivalent to the function $f(x) = \frac{1}{2} v^T A v + b^T v + c$ with $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

a. what is matrix A

$\frac{1}{2} v^T A v = 2x^2 + 3y^2 - z^2 + 3xy - 4xz + 6yz$

$(x \; y \; z) \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 4x^2 + 6y^2 - 2z^2 + 6xy - 8xz + 12yz$

$(ax + dy + gz)x + (bx + ey + hz)y + (cx + fy + iz)z$

$ax^2 + dxy + gxz + bxy + ey^2 + hyz + cxz + fyz + iz^2 = 4x^2 + 6y^2 - 2z^2 + 6xy - 8x$

$ax^2 = 4x^2$	$ey^2 = 6y^2$	$iz^2 = -2z^2$	$dxy + bxy = 6xy$	$gxz + cxz = 8xz$
$a = 4$	$e = 6$	$i = -2$	$d + b = 6$	$g + c = -8$
			$d = b = 3$	$g = c = -4$

$hyz + fyz = 12yz$

$h + f = 12$

$h = f = 6$

$A = \begin{bmatrix} 4 & 3 & -4 \\ 3 & 6 & 6 \\ -4 & 6 & -2 \end{bmatrix}$ //

b. what is vector b.

$b^T v = -2x + 4y - 7z$

$(a \; b \; c) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = ax + by + cz = -2x + 4y - 7z$

$ax = -2x$	$by = 4y$	$cz = -7z$
$a = -2$	$b = 4$	$c = -7$

$b = \begin{pmatrix} -2 \\ 4 \\ -7 \end{pmatrix}$ //