# CS 4442b Assignment 1

Danilo Vladicic - Dvladici@uwo.ca - 250861339

February 9, 2020

## Question 1

Let $\omega \epsilon \mathbb{R}^n$ be a n-dimensional column vector Let $f(\omega)\epsilon\mathbb{R}^n$ be a function of w

### Part A

Let $f(\omega) = \omega^T x$ where $x\epsilon\mathbb{R}^n$ is a n dimensional vector Compute $\nabla f(\omega)$ So,

$$\begin{aligned} f(\omega) &= \omega^T x \\ &= [\omega_1, \cdots, \omega_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= x_1\omega_1 + x_2\omega_2 + \cdots + x_n\omega_n \end{aligned}$$

So,

$$\frac{\delta f}{\delta \omega_i} = x_i \quad \forall i$$

Then,

$$\frac{\delta f}{\delta \omega_i} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x$$

### Part B

Let $f(\omega) = tr(\omega\omega^T A)$ where $A\epsilon\mathbb{R}^n xn$ is a n dimensional matrix Compute $\nabla f(\omega)$

$$
\begin{aligned}
f(\omega) &= tr([\omega_1, \cdots, \omega_n] \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}) \\
&= tr(\begin{bmatrix} \omega_1^2 & \omega_1\omega_2 & \cdots & \omega_1 w_n \\ \omega_1\omega_2 & \omega_2^2 & \cdots & \omega_2\omega_1 \\ \vdots & \vdots & \ddots & \vdots \\ \omega_1\omega_n & \omega_2\omega_n & \cdots & \omega_n^2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}) \\
&= \Sigma_{i=0}^{n} c_{ii}
\end{aligned}
$$

Where $c_{ii}$ is the i-th row of $\omega\omega^T$ times the i-th column of A then,

$$
\begin{aligned}
\Sigma_{i=0}^{n} c_{ii} &= ([\omega_1^2, \cdots, \omega_1\omega_n] \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix}) + \cdots + ([\omega_1\omega_n, \cdots, \omega_n^2] \begin{bmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{bmatrix}) \\
&= (\omega_1^2 a_{11} + \omega_1\omega_2 a_{21} + \cdots + \omega_1\omega_n a_{n1}) \\
&\quad + (\omega_2\omega_1 a_{12} + \omega_2^2 a_{22} + \cdots + \omega_2\omega_n a_{n2}) + \cdots + (\omega_n\omega_1 a_{1n} + \cdots + \omega_n^2 a_{nn})
\end{aligned}
$$

Then each $\frac{\delta f}{\delta \omega_i}$ can be found by grouping the terms containing $\omega_i$ and then taking the partial derivative. Doing this for n we can find a general equation:

$$
= (\omega_n\omega_1 a_{1n} + \cdots + \omega_n^2 a_{nn}) + (\omega_n\omega_1 a_{n1} + \omega_n\omega_2 a_{n2} + \cdots + \omega_n\omega_{n-1} a_{nn-1})
$$

Then

$$
\begin{aligned}
\frac{\delta f}{\delta \omega_n} &= (\omega_1 a_{1n} + \cdots + \omega_{n-1} a_{n-1n} + 2\omega_n a_{nn}) + (\omega_1 a_{n1} + \omega_2 a_{n2} + \cdots + \omega_{n-1} a_{nn-1}) \\
&= 2\omega_n a_{nn} + \omega_{n-1}(a_{n-1n} + a_{nn-1}) + \cdots + \omega_1(a_{1n} + a_{n1}) \\
&= \omega_n(a_{nn} + a_{nn}) + \omega_{n-1}(a_{n-1n} + a_{nn-1}) + \cdots + \omega_1(a_{1n} + a_{n1}) \\
&= \Sigma_{j=1}^{k} \omega_j (a_{jn} + a_{nj})
\end{aligned}
$$

From this we find

$$
\frac{\delta f}{\delta \omega_i} = \Sigma_{j=1}^{n} \omega_j (a_{ji} + a_{ij})
$$

and finally

$$
\nabla f = \begin{bmatrix} \Sigma_{j=1}^{n} \omega_j (a_{j1} + a_{1j}) \\ \Sigma_{j=1}^{n} \omega_j (a_{j2} + a_{2j}) \\ \vdots \\ \Sigma_{j=1}^{n} \omega_j (a_{ji} + a_{ij}) \end{bmatrix}
$$

## Part C

From before, we computed

$$\frac{\delta f}{\delta \omega_i} = \Sigma_{j=1}^{n} \omega_j (a_{ji} + a_{ij})$$

We see that $\frac{\delta^2 f}{\delta \omega_i^2}$ exists only when $i = j$

$$\frac{\delta^2 f}{\delta \omega_i^2} = 2a_{ii}$$

For the remaining partials

$$\frac{\delta f}{\delta \omega_i \delta \omega_j} = \frac{\delta f}{\delta \omega_j \delta \omega_i} = a_{ij} + a_{ji}$$

Then we define the Hessian Matrix H as

$$H = \begin{bmatrix} 2a_{11} & a_{12} + a_{21} & \cdots & a_{1n} + a_{n1} \\ a_{21} + a_{12} & 2a_{22} & \cdots & a_{2n} + a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} + a_{1n} & a_{n2} + a_{2n} & \cdots & 2a_{nn} \end{bmatrix}$$

## Part D

Let $\alpha = \omega^T x = \Sigma_{i=1}^{n} \omega_i x_i$ then,

$$\frac{\delta \alpha}{\delta \omega_i} = x_i \tag{1}$$

Let $\sigma(\alpha) = \frac{1}{1+e^{-\alpha}}$ then,

$$\frac{\delta \sigma}{\delta \alpha} = \frac{e^{-\alpha}}{(1+e^{-\alpha})^2} \tag{2}$$

Let $f(\sigma) = \ln(\sigma)$ then,

$$\frac{\delta f}{\delta \sigma} = \frac{1}{\sigma} \tag{3}$$

Then joining equations 1, 2, and 3

$$\begin{aligned} \frac{\delta f}{\delta \omega_i} &= (\frac{\delta f}{\delta \sigma})(\frac{\delta \sigma}{\delta \alpha})(\frac{\delta \alpha}{\delta \omega_i}) \\ &= (1 + e^{-\omega^T x})(\frac{e^{-\omega^T x}}{(1 + e^{-\omega^T x})^2})(x_i) \\ &= \frac{e^{-\omega^T x} \times x_i}{1 + e^{-\omega^T x}} \times \frac{e^{\omega^T x}}{e^{\omega^T x}} \\ &= \frac{x_i}{e^{\omega^T x} + 1} \end{aligned}$$

Let $\lambda = \frac{1}{e^{\omega^T x}+1}$
Then,

$$\begin{aligned}
\nabla f &= \begin{bmatrix} \frac{\delta f}{\delta w_1} \\ \vdots \\ \frac{\delta f}{\delta w_i} \end{bmatrix} \\
&= \begin{bmatrix} \frac{x_1}{e^{\omega^T x}+1} \\ \vdots \\ \frac{x_i}{e^{\omega^T x}+1} \end{bmatrix} \\
&= \lambda \begin{bmatrix} x_1 \\ \vdots \\ x_i \end{bmatrix} \\
&= \lambda x
\end{aligned}$$