cs-236330-optimization-hw1

Manor Zvi - 204030720

April 2020

1 Task 1

1.1 Full Derivation of the Gradient of $f_1(x) = \varphi(Ax)$

Let: $Ax = u$
Therefore, the differential: $du = Adx$
And then, by the definition of external differential: $df_1 = \langle \nabla\varphi, du \rangle$
by the definition of euclidean Inner Product:
$\langle \nabla\varphi, du \rangle = \nabla\varphi^T du = \nabla\varphi^T Adx = \left(A^T \nabla\varphi\left(Ax\right)\right)^T dx$
And to conclude:

$$\nabla f_1\left(x\right) = A^T \nabla\varphi\left(Ax\right) \tag{1}$$

1.2 Full Derivation of the Hessian of $f_1(x) = \varphi(Ax)$

Let: $g = \nabla f_1\left(x\right) = \nabla f_1\left(x\right) = A^T \nabla\varphi\left(Ax\right)$
By the external definition of the Hessian: $dg = Hdx$
Therefore:

$$d\left(A^T \nabla\varphi\left(Ax\right)\right) = A^T d\left(\nabla\varphi\left(u\right)\right) = A^T \langle \nabla\varphi, du \rangle = A^T \nabla^2 du = A^T \nabla^2 \varphi Adx \tag{2}$$

And then we conclude:

$$H = A^T \nabla^2 \varphi\left(Ax\right) A \tag{3}$$

2 Task 2

2.1 Full Derivation of the Gradient of $f_2(x) = h(\varphi(x))$

$df_2\left(x\right) = d\left(h'\left(\varphi\left(x\right)\right)\right) = \langle h'\left(\varphi\left(x\right)\right), d\varphi\left(x\right)\rangle = h'\left(\varphi\left(x\right)\right) d\varphi\left(x\right) (*)$
$dd\varphi\left(x\right) = \langle \nabla\varphi\left(x\right), dx \rangle = \nabla\varphi\left(x\right)^T dx$
$(*)\, df_2\left(x\right) = h'\left(\varphi\left(x\right)\right) \nabla\varphi\left(x\right)^T dx \Rightarrow \nabla f_2\left(x\right) = h'\left(\varphi\left(x\right)\right) \nabla\varphi\left(x\right)$

2.2 Full Derivation of the Hessian of $f_2(x) = h(\varphi(x))$

Let: $g\left(x\right) = \nabla f_2\left(x\right) = h'\left(\varphi\left(x\right)\right)\nabla\varphi\left(x\right)$
$dg = d\left(h'\left(\varphi\left(x\right)\right)\nabla\varphi\left(x\right)\right) =$
$dh'\left(\varphi\left(x\right)\right)\nabla f_2\left(x\right) + h'\left(\varphi\left(x\right)\right)d\nabla f_2\left(x\right) =$
$\langle h''\left(\varphi\left(x\right)\right), d\varphi(x)\rangle\nabla\varphi(x) + h'\left(\varphi\left(x\right)\right)\langle\nabla^2\varphi(x), dx\rangle =$
$h''\left(\varphi\left(x\right)\right)\nabla\varphi(x)^T dx\nabla\varphi(x) + h'\left(\varphi\left(x\right)\right)\nabla^2\varphi(x)dx =$
$\left(h''\left(\varphi\left(x\right)\right)\nabla\varphi(x)^T\nabla\varphi(x) + h'\left(\varphi\left(x\right)\right)\nabla^2\varphi(x)\right)dx$ And Therefore:

$$H = h''\left(\varphi\left(x\right)\right)\nabla\varphi(x)^T\nabla\varphi(x) + h'\left(\varphi\left(x\right)\right)\nabla^2\varphi(x) \tag{4}$$

3 Task 3

3.1 Full Derivation of the Gradient of:

$$\varphi\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right) = \sin(x_1x_2x_3) \tag{5}$$

$d\varphi\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right) = \langle\nabla\sin(x_1x_2x_3), dx\rangle = \nabla\sin(x_1x_2x_3)^T dx$
$\nabla\sin(x_1x_2x_3) = \left(\begin{bmatrix}x_2x_3\\x_1x_3\\x_1x_2\end{bmatrix}\right)\cos(x_1x_2x_3) \Rightarrow$
$\nabla\varphi\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right) = \left(\begin{bmatrix}x_2x_3\\x_1x_3\\x_1x_2\end{bmatrix}\right)\cos(x_1x_2x_3)$

3.2 Full Derivation of the Hessian

$g = \nabla\varphi\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right)$
$dg = Hdx$
$dg = \nabla^2\varphi\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right)^T dx$
$H = \nabla^2\varphi\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right)^T =$
$\begin{bmatrix}0 & x_3 & x_2\\x_3 & 0 & x_1\\x_2 & x_1 & 0\end{bmatrix}\cos(x_1x_2x_3) - \begin{bmatrix}x_2 & x_3\\x_1 & x_3\\x_1 & x_2\end{bmatrix}\begin{bmatrix}x_2 & x_1 & x_1\\x_3 & x_3 & x_2\end{bmatrix}\sin(x_1x_2x_3)$

3.3 Full Derivation of the First & Second Derivatives of:

$$h(x) = \exp(x) \tag{6}$$

$dh(x) = h'(x), dx\rangle = h'(x)dx$
$h'(x) = \exp(x)$
$dh''(x) = \exp(x)$

3.4 Analytical Calculations:

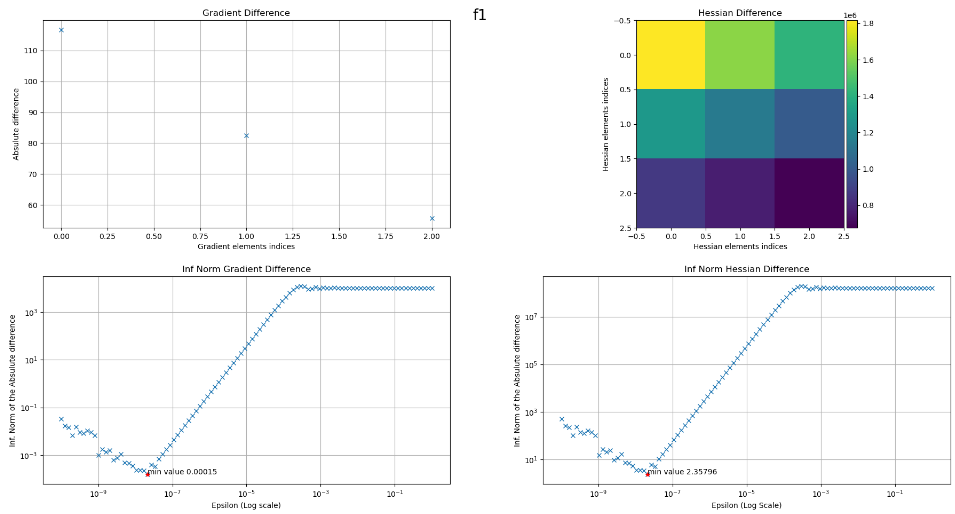
Run the functions run_f1 & run_f2 in the file hw1.py
Each function calls f1/f2 accordingly and receive a class called f1/f2_par which contains its parameters.
Inside f1/f2_par there are pointers to functions that do the calculations (called phi, gradient_phi and so on)

4 Task 4 - Numerical Calculations:

Run the function run_numdiff in the file hw1.py
Code structure is very similar to the code in section 3.

5 Task 5 - Comparisons

f1



f2