236330 - Introduction to Optimization: Homework #1

April 2, 2020

Amit Rotner
123456789
Or Steiner
123456789

Analytical Differentiation:

Task 1:

Find the Gradient and Hessian of:

$$f_1(x) = \varphi(Ax)$$

Where,

$$\begin{gathered} f_1 : \mathbb{R}^n \to \mathbb{R} \\ x \in \mathbb{R}^n \\ A \in \mathbb{R}^{m \times n} \\ \varphi : \mathbb{R}^m \to \mathbb{R} \end{gathered}$$

Given $\nabla\varphi$ and $\nabla^2\varphi$.

Solution:

We start by finding the gradient:

Denoting $u = Ax$. Hence $du = dAx$.

From the external definition of gradient, we know that,

$$d\varphi = \langle \nabla\varphi, du \rangle = \nabla\varphi^T du = \nabla\varphi^T A dx$$

Since $df = \nabla f^T dx$ and $d\varphi = df$ (because $f(x) = \varphi(Ax)$) we get that:

$$g = \nabla f = A^T \nabla\varphi(Ax)$$

Now for the Hessian,

$$dg = d\nabla f = dA^T \nabla\varphi(u) = A^T d\nabla\varphi(u)$$

From the external definition of Hessian we know that $d\nabla\varphi(u) = H du$.

$$A^T d\nabla\varphi(u) = A^T H(u)\, du = A^T H(Ax)\, A dx$$

Since $H = \nabla^2\varphi$, it yields that:

$$H(x) = A^T \nabla^2\varphi(Ax)\, A$$

Task 2:

Develop the Gradient and Hessian of:

$$f_2(x) = h(\varphi(x))$$

Where,

$$\begin{gathered} f_2 : \mathbb{R}^n \to \mathbb{R} \\ x \in \mathbb{R}^n \\ \varphi : \mathbb{R}^m \to \mathbb{R} \\ h : \mathbb{R} \to \mathbb{R} \end{gathered}$$

Given $\nabla\varphi,\ \nabla^2\varphi,\ h',\ h''$.

Solution:

We start by finding the gradient:

Denoting $u = \varphi(x)$. From the external definition of the gradient, we know that:

$$du = \langle \nabla\varphi(x), dx \rangle$$

$$dh = h'(u)\, du$$

Combining these two together yields:

$$dh = h'(u)\langle \nabla\varphi(x), dx \rangle = \langle h'(u)\nabla\varphi(x), dx \rangle = \langle h'(\varphi(x))\nabla\varphi(x), dx \rangle$$

Since $df = \nabla f^T dx$ we get that:

$$g = \nabla f = \left(h'(\varphi(x))\nabla\varphi(x)\right)^T \underset{(1)}{=} h'(\varphi(x))\nabla\varphi(x)$$

(1) - Since $h'(\varphi(x))$ is a scalar and $\nabla\varphi(x) = \nabla\varphi^T(x)$.

Now for the Hessian,

$$
\begin{aligned}
dg &= d\nabla f = d\left(h'\left(u\right)\nabla\varphi\left(x\right)\right)\\
&= dh'\left(u\right)\nabla\varphi\left(x\right) + h'\left(u\right)d\nabla\varphi\left(x\right)\\
&= h''\left(u\right)du\nabla\varphi\left(x\right) + h'\left(u\right)\nabla^2\varphi\left(x\right)dx\\
&\underset{(1)}{=} h''\left(u\right)\nabla\varphi\left(x\right)du + h'\left(u\right)\nabla^2\varphi\left(x\right)dx\\
&= h''\left(u\right)\cdot\nabla\varphi\left(x\right)\cdot\left\langle\nabla\varphi\left(x\right),dx\right\rangle + h'\left(u\right)\nabla^2\varphi\left(x\right)dx\\
&= h''\left(u\right)\cdot\nabla\varphi\left(x\right)\cdot\nabla\varphi^T\left(x\right)dx + h'\left(u\right)\nabla^2\varphi\left(x\right)dx\\
&= \left(h''\left(u\right)\cdot\nabla\varphi\left(x\right)\cdot\nabla\varphi^T\left(x\right) + h'\left(u\right)\nabla^2\varphi\left(x\right)\right)\cdot dx
\end{aligned}
$$

(1) - Since $du = \left\langle\nabla\varphi\left(x\right),dx\right\rangle$ is a scalar.

From the external definition of Hessian we know that $dg = Hdx$, hence:

$$H = h''\left(u\right)\cdot\nabla\varphi\left(x\right)\cdot\nabla\varphi^T\left(x\right) + h'\left(u\right)\nabla^2\varphi\left(x\right)$$