# Introduction to Numerical Optimization– HW1

Noam Chen and Ze′ev Sheleg

April 27, 2021

## Section 1.1.1

We will calculate the gradient of $f_1(\bar{x})$:
In order to calculate the gradient of $f_1(x)$ we will use the external definition of the gradient which is derived from the differential:

$$df(x) = \Sigma_i \frac{\partial f}{\partial x_i} \partial x_i$$

when $f(x) = \phi(u)$, when $u = Ax \in \mathbb{R}^m,$ thus

$$df(x) = \Sigma_i \frac{\partial \phi(u)}{\partial u_i} \partial u_i$$

when

$$du = d(Ax) = Adx$$

therefore

$$df(x) = \Sigma_i \frac{\partial \phi(u)}{\partial u_i} \partial u_i = (\nabla \phi(u))^T \cdot du = (\nabla \phi(u)^T Adx =< A^T \cdot \nabla \phi, dx >$$

hence the gradient of $f_1(\bar{x})$ is

$$\nabla f_1(x) = A^T \cdot \nabla \phi \, \textbf{.}$$

We will calculate the hessian of $f_1(\bar{x})$:
we will define the following function:

$$g(x) = \nabla f_1(x) = A^T \cdot \nabla \phi$$

We know by definition that:

$$dg = Hdx$$

when H is the hessian matrix of $f_1(x)$ and therfore:

$$dg = d(\nabla f_1(x)) = d(A^T \cdot \nabla \phi)$$

And since $A^T$ is a scalar then:

$$dg = A^T \cdot d(\nabla \phi(Ax))$$

And according to what we have already proof in the lecture we can conclude:

$$dg \underbrace{=}_{in\,respect\,to\,u} A^T \nabla^2 \phi du = A^T \nabla^2 \phi A \cdot dx$$

And by definition we can infer that:

$$H = A^T \nabla^2 \phi A \, \textbf{.}$$

**Section** 1.1.2

We will calculate the gradient of $f_2(x)$:
Since $h$ is scalar function using the chain rule we can infer that:

$$df_2(x) = d(h(\underbrace{\phi(x)}_{t}))$$

When $t \in \mathbb{R}$, therefore:

$$df_2(x) = dh(t) = h'(t)dt$$

And after substitution we get:

$$df_2(x) = h'(\phi(x)) \cdot d(\phi(x)) = h'(\phi(x)) \cdot \nabla\phi \cdot dx$$

The above equation can be written as follows:

$$df_2(x) = (h'(\phi(x)) \cdot \nabla\phi)^T = h'(\phi(x)) \cdot (\nabla\phi)^T \cdot dx =< h'(\phi(x)) \cdot \nabla\phi, dx >$$

Hence the gradient is:

$$\nabla f_2(x) = h'(\phi(x)) \cdot \nabla\phi \text{.}$$

Similarly to the previous section we will calclate the hessian:
We will define the following function:

$$k(x) = \nabla f_2(x) = h'(\phi(x)) \cdot \nabla\phi$$

By definition we know:

$$dk(x) = Hdx$$

When H is the hessian of $f_2(x)$ and therefore:

$$dk(x) = d(h'(\phi(x)) \cdot \nabla\phi) = \frac{\partial^2 h(\phi(x))}{\partial^2 x} \cdot d(\nabla\phi) = \frac{\partial^2 h(\phi(x))}{\partial^2 x} \cdot \nabla^2\phi \cdot dx = \frac{\partial^2 h(\phi(x))}{\partial^2 x} \cdot H_\phi \cdot dx$$

Hence the Hessian follows the following expression:

$$H = \frac{\partial^2 h(\phi(x))}{\partial^2 x} \cdot H_\phi \text{ .}$$

## **Section** 1.1.3

We will calculate the gradient of $\phi$:
As we saw in section 1.1.2:

$$\nabla f(x) = h'(\phi(x)) \cdot \nabla\phi$$

Under the assumption:

$$h(x) = sinx,\ \phi(x_1, x_2, x_3) = x_1x_2x_3$$

Then we get:

$$\nabla\phi(x) = cos(x_1x_2x_3) \cdot (\frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_2}, \frac{\partial\phi}{\partial x_3})^T = cos(x_1x_2x_3) \cdot (x_2x_3,\ x_1x_3,\ x_2x_3)^T$$

The finding of the Hessian follows to the previous section as well:

$$H = \frac{\partial h(\phi(x))^2}{\partial^2 x} \cdot H_\phi = -sin(x_1x_2x_3) \cdot H_\phi = -sin(x_1x_2x_3) \cdot \begin{pmatrix} 0 & x_3 & x_2 \\ x_3 & 0 & x_1 \\ x_2 & x_1 & 0 \end{pmatrix} \text{.}$$

## **Section** 1.1.4

The 1st derivative of $h$:

$$\frac{\partial h(x)}{\partial x} = \frac{1}{2\sqrt{1+x^2}} \cdot 2x = \frac{x}{\sqrt{1+x^2}}$$

The 2nd derivative of $h$:

$$\frac{\partial h(x)^2}{\partial^2 x} = \frac{1 \cdot \sqrt{1+x^2} + \frac{x}{\sqrt{1+x^2}} \cdot x}{1+x^2} = \frac{1+2x^2}{(1+x^2)^{\frac{3}{2}}}$$