Non-convex Optimization for Analyzing Big Data

Assignment 2

Zhiwei Han

Jun 10, 2016

Task 1

(a)

Chain rule:

$$D(f \circ g)(x) = Df(g(x)) \circ Dg(x) \tag{1}$$

The directional derivative of composition $f \circ g$ is:

$$\begin{aligned} D_v(f \circ g)(x) &= D_{Dg(x)[v]} f(g(x)) \\ &= Df(g(x))[Dg(x)[v]] \end{aligned} \tag{2}$$

Here given $f : \mathbb{R}^n \to \mathbb{R}, \mathbf{x} \mapsto \mathbf{x}^\intercal \mathbf{x}$ and $g : \mathbb{R}^n \to \mathbb{R}^n, \mathbf{z} \mapsto \mathbf{Az}$
Choose standard inner product as product in all the spaced used.

$$< \mathbf{x}, \mathbf{y} >= \mathbf{x}^\intercal \mathbf{y}$$

With the inner product chosen:

$$D_\mathbf{v} f(\mathbf{x}) = 2\mathbf{x}^\intercal \mathbf{v} \tag{3}$$

and

$$D_\mathbf{v} g(\mathbf{z}) = \mathbf{Av} \tag{4}$$

Plug into the fomular:

$$D_v(f \circ g)(x) = 2\mathbf{x}^\intercal \mathrm{Av} =< 2\mathbf{A}^\intercal \mathbf{x}, \mathbf{v} > \tag{5}$$

So,

$$\mathbf{g} = 2\mathbf{A}^\intercal \mathbf{x}$$

(b)

Product rule:

$$D_{\mathbf{H}}(f \cdot g)(\mathbf{X}) = D_{\mathbf{H}} f(\mathbf{X}) \cdot g(\mathbf{X}) + f(\mathbf{X}) D_{\mathbf{H}} \cdot g(\mathbf{X}) \tag{6}$$

Here given $f : \text{skew}_n \to \mathbb{R}^n, \mathbf{X} \mapsto \mathbf{X}^\intercal \mathbf{a}$ and $g : \text{skew}_n \to \mathbb{R}^n, \mathbf{X} \mapsto \mathbf{X}\mathbf{b}$
The Objectiv function is the multiplication of these two functions:

$$h(\mathbf{X}) = f(\mathbf{X})^\intercal \cdot g(\mathbf{X})$$

Choose standard inner product as product in all the spaced used.

$$< \mathbf{x}, \mathbf{y} >= \mathbf{x}^\intercal \mathbf{y}$$

Here we have:

$$D_H h((X)) = D_H f(\mathbf{X})^\intercal \cdot g(\mathbf{X}) + f(\mathbf{X})^\intercal \cdot D_H g(\mathbf{X}) \tag{7}$$

With the inner product chosen, we have the following directional derivative:

$$D_H f(\mathbf{X})^\intercal = \mathbf{a}^\intercal \mathbf{H} \tag{8}$$

and

$$D_H g(\mathbf{X}) = \mathbf{H}\mathbf{b} \tag{9}$$

Plug into the fomular:

$$\begin{aligned} D_H h((X)) &= \mathbf{a}^\intercal \mathbf{H}\mathbf{X}\mathbf{b} + \mathbf{a}^\intercal \mathbf{X}\mathbf{H}\mathbf{b} \\ &= - < \mathbf{H}\mathbf{a}, \mathbf{X}\mathbf{b} > - < \mathbf{X}\mathbf{a}, \mathbf{H}\mathbf{b} > \\ &=< -(\mathbf{X}\mathbf{b}\mathbf{a}^\intercal + \mathbf{X}\mathbf{a}\mathbf{b}^\intercal), \mathbf{H} > \end{aligned} \tag{10}$$

So,

$$\mathbf{g} = -(\mathbf{X}\mathbf{b}\mathbf{a}^\intercal + \mathbf{X}\mathbf{a}\mathbf{b}^\intercal)$$

Task 2

Please See the comment in uploaeded matlab file.