HW4

2020年11月15日 16:23

HW4(2)

MATH 3332 Data Analytic Tools
Homework 4

Due date: 16 November, 6pm, Monday

1. Find the gradient of the following functions $f : V \mapsto \mathbb{R}$, where V is a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$.

   (a) $f(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{a}\|$ for a given $\boldsymbol{a} \in V$, and $\boldsymbol{x} \neq \boldsymbol{a}$.

   (b) $f(\boldsymbol{x}) = \|2\boldsymbol{x} - \boldsymbol{a}\|^2$ for a given $\boldsymbol{a} \in V$.

   (c) $f(\boldsymbol{x}) = \frac{1}{\|\boldsymbol{x}\|}$, where $\boldsymbol{x} \neq \boldsymbol{0}$.

   (d) $f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2 + \lambda\|\boldsymbol{B}\boldsymbol{x}\|^2$, where $\lambda > 0$, $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{p\times n}$.

   (e) $V = \mathbb{R}^n$ and $f(\boldsymbol{x}) = \sum_{i=1}^n \sqrt{x_i^2 + c}$ for some $c > 0$.

   (f) Let $f : \mathbb{R}^m \to \mathbb{R}$ be differentiable everywhere. Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ be a matrix. Let $g : \mathbb{R}^m \to \mathbb{R}$ be defined by
   $$g(\boldsymbol{x}) = f(\boldsymbol{A}\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \mathbb{R}^n.$$
   Express $\nabla g(\boldsymbol{x})$ in terms of $\boldsymbol{A}$ and gradient of f.

2. Determine whether or not the following functions are convex.

   (a) $f(x) = e^x - 1$ on $\mathbb{R}$.

   (b) $f(\boldsymbol{x}) = \sum_{i=1}^n \sqrt{|x_i|}$ on $\mathbb{R}^n$.

3. Let f_1 and f_2 be two convex functions from $\mathbb{R}^n$ to $\mathbb{R}$.

   (a) Let
   $$f(\boldsymbol{x}) = f_1(\boldsymbol{x}) + f_2(\boldsymbol{x}), \quad \forall\, \boldsymbol{x} \in \mathbb{R}^n.$$
   Prove that f is convex.

   (b) Let
   $$g(\boldsymbol{x}) = f_1(x) - f_2(\boldsymbol{x}), \quad \forall\, \boldsymbol{x} \in \mathbb{R}^n.$$
   Give an example of f_1 and f_2 such that g is not convex.

   (c) Let
   $$h(\boldsymbol{x}) = \max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\}, \quad \forall\, \boldsymbol{x} \in \mathbb{R}^n.$$
   Prove that h is convex.

   (d) Let
   $$k(\boldsymbol{x}) = \min\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\}, \quad \forall\, \boldsymbol{x} \in \mathbb{R}^n.$$
   Give an example of f_1 and f_2 such that k is not convex.

1. Find the gradient of the following functions $f : V \mapsto \mathbb{R}$, where V is a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$.

   (a) $f(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{a}\|$ for a given $\boldsymbol{a} \in V$, and $\boldsymbol{x} \neq \boldsymbol{a}$.

Fix $x \in V - \{a\}$, we have

$$f(x) = \|x-a\| = (\|x-a\|^2)^{\frac{1}{2}}$$

Let $g(t) = \sqrt{t}$ and $h(x) = \|x-a\|^2$,

$$f(x) = g(h(x))$$

$$\nabla g(h(x)) = g'(h(x))\, \nabla h(x)$$

$$g'(t) = \frac{1}{2} t^{-\frac{1}{2}}$$

$$\nabla h(x) = \nabla \|x-a\|^2 = 2(x-a)$$

$$\therefore \quad \nabla f(x) = \frac{1}{2}(\|x-a\|^2)^{-\frac{1}{2}} \cdot 2(x-a)$$

$$= \frac{x-a}{\|x-a\|}$$

(b) $f(\boldsymbol{x}) = \|2\boldsymbol{x} - \boldsymbol{a}\|^2$ for a given $\boldsymbol{a} \in V$.

b). Fix $x_0 \in V$,

$$f(x) - f(x_0) = \|2x-a\|^2 - \|2x_0 - a\|^2$$

$$= \|2x - 2x_0 + 2x_0 - a\|^2 - \|2x_0 - a\|^2$$

$$= \|2x - 2x_0 + 2x_0 - a\|^2 - \|2x_0 - a\|^2$$

$$= \|2x - 2x_0\|^2 + 2\langle 2x - 2x_0, 2x_0 - a\rangle$$

$$= 4\|x - x_0\|^2 + \langle x - x_0, 4(2x_0 - a)\rangle$$

$$\therefore f(x) - f(x_0) - \langle x - x_0, 4(2x_0 - a)\rangle$$

$$= 4\|x - x_0\|^2$$

or equivalently,

$$\lim_{x \to \infty} \frac{|f(x) - f(x_0) - \langle x - x_0, 4(x_0 - a)\rangle|}{\|x - x_0\|}$$

$$= \lim_{x \to \infty} 4\|x - x_0\|$$

$$= 0$$

$$\therefore \nabla f(x_0) = 4(2x_0 - a)$$

Since $x_0 \in V$ is arbitrary,

$$\nabla f(x) = 4(2x - a) \quad \forall x \in V.$$

$$\nabla f(x) = 4(2x - a) \quad \forall\, x \in V.$$

(c) $f(\boldsymbol{x}) = \frac{1}{\|\boldsymbol{x}\|}$, where $\boldsymbol{x} \neq \mathbf{0}$.

c). Fix $x \in V - \{0\}$,

$$f(x) = \frac{1}{\|x\|} = (\|x\|^2)^{-\frac{1}{2}}$$

let $g(t) = \frac{1}{\sqrt{t}}$ and $h(x) = \|x\|^2$

$$f(x) = g(h(x))$$

$$\nabla g(h(x)) = g'(h(x))\, \nabla h(x)$$

$$g'(t) = -\frac{1}{2} t^{-\frac{3}{2}}$$

$$\nabla h(x) = \nabla \|x\|^2 = 2x$$

$$\therefore \nabla f(x) = -\frac{1}{2}(\|x\|^2)^{-\frac{3}{2}} \cdot 2x$$

$$= -\frac{x}{\|x\|^3}$$

(d) $f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2 + \lambda\|\boldsymbol{B}\boldsymbol{x}\|^2$, where $\lambda > 0$, $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{B} \in \mathbb{R}^{p\times n}$.

$$\|Ax-b\|^2 = \langle Ax-b, Ax-b\rangle$$

$$\nabla(\|Ax-b\|^2) = \nabla(\langle Ax-b, Ax-b\rangle)$$

$$= \nabla(Ax-b)\cdot(Ax-b) + \nabla(Ax-b)\cdot(Ax-b)$$

$$= 2\,\nabla(Ax-b)(Ax-b)$$

Now solving $\nabla(Ax-b)$,

$$Ax-b = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n - b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n - b_m \end{bmatrix}$$

$$\therefore \nabla(Ax-b) = \begin{bmatrix} \frac{\partial}{\partial x_1}(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - b_1) \\ \frac{\partial}{\partial x_2}(a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n - b_2) \\ \vdots \\ \frac{\partial}{\partial x_n}(a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n - b_m) \end{bmatrix}$$

$$\therefore \nabla(Ax-b) = \begin{bmatrix} \frac{\partial}{\partial x_1}(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - b_1) \\ \frac{\partial}{\partial x_2}(a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n - b_2) \\ \vdots \\ \frac{\partial}{\partial x_n}(a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n - b_m) \end{bmatrix}$$

$$= (a_{11}, a_{22}, \ldots, a_{mn})$$

$$\therefore \nabla(\|Ax-b\|^2) = 2(a_{11}, a_{22}, \cdots, a_{mn})(Ax-b)$$

Now solving $\nabla(\|Bx\|^2)$

$$\|Bx\|^2 = \langle Bx, Bx \rangle$$

$$\nabla\|Bx\|^2 = 2\nabla(Bx)Bx.$$

Solving $\nabla(Bx)$:

$$Bx = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & & & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\begin{pmatrix} \vdots & & & \\ b_{p1} & b_{p2} & \cdots & b_{pn} \end{pmatrix} \begin{pmatrix} \vdots \\ x_n \end{pmatrix}$$

$$Bx = \begin{bmatrix} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n \\ b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n \\ \vdots \\ b_{p1}x_1 + b_{p2}x_2 + \cdots + b_{pn}x_n \end{bmatrix}$$

$$\therefore \nabla(Bx) = (b_{11}, b_{22}, \ldots, b_{pn})$$

$$\therefore \nabla f(x)$$

$$= \frac{1}{2}\nabla(\|Ax-b\|^2) + \lambda \nabla(\|Bx\|^2)$$

$$= (a_{11}, a_{22}, \ldots, a_{mn})(Ax-b) + 2\lambda(b_{11}, b_{22}, \ldots, b_{pn}) \cdot Bx$$