



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) $f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\|$ for a given $\mathbf{a} \in V$, and $\mathbf{x} \neq \mathbf{a}$.
- (b) $f(\mathbf{x}) = \|2\mathbf{x} - \mathbf{a}\|^2$ for a given $\mathbf{a} \in V$.
- (c) $f(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|}$, where $\mathbf{x} \neq \mathbf{0}$.
- (d) $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{Ax} - \mathbf{b}\|^2 + \lambda\|\mathbf{Bx}\|^2$, where $\lambda > 0$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{p \times n}$.
- (e) $V = \mathbb{R}^n$ and $f(\mathbf{x}) = \sum_{i=1}^n \sqrt{x_i^2 + c}$ for some $c > 0$.
- (f) Let $f : \mathbb{R}^m \rightarrow \mathbb{R}$ be differentiable everywhere. Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be a matrix. Let $g : \mathbb{R}^m \rightarrow \mathbb{R}$ be defined by

$$g(\mathbf{x}) = f(\mathbf{Ax}), \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Express $\nabla g(\mathbf{x})$ in terms of \mathbf{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) $f(\mathbf{x}) = \sum_{i=1}^n \sqrt{|x_i|}$ on \mathbb{R}^n .

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

- (a) Let

$$f(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Prove that f is convex.

- (b) Let

$$g(\mathbf{x}) = f_1(\mathbf{x}) - f_2(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Give an example of f_1 and f_2 such that g is not convex.

- (c) Let

$$h(\mathbf{x}) = \max\{f_1(\mathbf{x}), f_2(\mathbf{x})\}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Prove that h is convex.

- (d) Let

$$k(\mathbf{x}) = \min\{f_1(\mathbf{x}), f_2(\mathbf{x})\}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

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

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

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

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

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

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

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

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

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

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

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

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

b). Fix $x_0 \in V$,

$$\begin{aligned} f(x) - f(x_0) &= (\|2x-a\|^2 - \|2x_0-a\|^2) \\ &= \|2x-2x_0 + 2x_0-a\|^2 - \|2x_0-a\|^2 \\ &= \|2x-2x_0\|^2 + 2\langle 2x-2x_0, 2x_0-a \rangle \\ &= 4\|x-x_0\|^2 + \langle x-x_0, 4(2x_0-a) \rangle \end{aligned}$$

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

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

or equivalently,

$$\lim_{x \rightarrow \infty} \frac{|f(x) - f(x_0) - \langle x-x_0, 4(2x_0-a) \rangle|}{\|x-x_0\|}$$

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

$$= 0$$

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

Since $x_0 \in V$ is arbitrary,

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

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

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

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

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

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

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

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

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

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

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

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

$$\begin{aligned}
 \|Ax - b\|^2 &= (Ax - b)^T (Ax - b) \\
 &= x^T A^T Ax - (Ax)^T b - b^T Ax + b^T b \\
 &= x^T A^T A x - 2b^T A x + b^T b \\
 &= x^T A^T A x - 2(A^T b)^T x + b^T b
 \end{aligned}$$

Let $\varphi(x) = C^T x$ where $C = A^T b$

$$\varphi(x) = \sum_{j=1}^n c_j x_j$$

$$\begin{aligned}
 \frac{\partial \varphi}{\partial x_k} &= \sum_{j=1}^n c_j \frac{\partial x_j}{\partial x_k} = \sum_{j=1}^n c_j \delta_{jk} = c_k \\
 \Rightarrow D\varphi(x) &= C = A^T b
 \end{aligned}$$

Let $\psi(x) = x^T B x$ where $B = A^T A$

$$\psi(x) = \sum_{i=1}^n \sum_{j=1}^n b_{ij} x_i x_j$$

$$\begin{aligned}
 \frac{\partial \psi}{\partial x_k} &= \sum_{i=1}^n \sum_{j=1}^n b_{ij} \frac{\partial (x_i x_j)}{\partial x_k} \\
 &= \sum_{i=1}^n \sum_{j=1}^n b_{ij} (\delta_{ik} x_j + x_i \delta_{jk})
 \end{aligned}$$

$$= \sum_{j=1}^n b_{kj} x_j + \sum_{i=1}^m b_{ik} x_i$$

$$= (Bx)_k + \sum_{i=1}^m (B^T)_{ki} x_i$$

$$= (Bx)_k + (B^T x)_k$$

$$\therefore \nabla \phi(x) = (B + B^T)x \quad -(*)$$

$$\therefore \nabla (\|Ax-b\|_2^2) = (A^T A + (A^T A)^T)x - 2A^T b$$

$$\text{and } (A^T A)^T = A^T (A^T)^T = A^T A$$

$$\Rightarrow \nabla (\|Ax-b\|_2^2) = 2A^T A x - 2A^T b$$

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

$$\|Bx\|^2 = (Bx)^T (Bx)$$

$$\|Bx\|^2 = x^T B^T B x = x^T (B^T B)^T x$$

From $(*)$, we know that $\nabla (\|Bx\|^2)$

$$= (B^T B + (B^T B)^T)x$$

$$= (B^T B + B^T B)x$$

$$\begin{aligned} &= (\beta^T B + \beta^T B)x \\ &= 2(\beta^T B)x \end{aligned}$$

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

$$\begin{aligned} &= \frac{1}{2}(2A^T A x - 2A^T b) + 2\lambda(\beta^T B)x \\ &= A^T A x - A^T b + 2\lambda(\beta^T B)x \end{aligned}$$

c)

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

$$\begin{aligned} f(x) &= \sum_{i=1}^n \sqrt{x_i^2 + c} = \sum_{i=1}^n \left(\sqrt{x_i^2 + c} - \sqrt{x_{0,i}^2 + c} \right) + \\ &\quad \sum_{i=1}^n \sqrt{x_{0,i}^2 + c} \end{aligned}$$

for some $x_0 = (x_{0,1}, x_{0,2}, \dots, x_{0,n}) \in V$

$$\begin{aligned} f(x) &= \sum_{i=1}^n \frac{(x_i^2 + c) - (x_{0,i}^2 + c)}{\sqrt{x_i^2 + c} + \sqrt{x_{0,i}^2 + c}} + f(x_0) \\ &= \sum_{i=1}^n (x_i - x_{0,i}) \frac{(x_i + x_{0,i})}{\sqrt{x_i^2 + c} + \sqrt{x_{0,i}^2 + c}} + f(x_0) \end{aligned}$$

$$\sum_{i=1}^n (x_i - x_{0,i}) \frac{x_{0,i}}{\sqrt{x_i^2 + c} + \sqrt{x_{0,i}^2 + c}} + f(x_0)$$

$$= \sum_{i=1}^n (x_i - x_{0,i}) \frac{x_{0,i}}{\sqrt{x_{0,i}^2 + c}} + \sum_{i=1}^n (x_i - x_{0,i})$$

$$\left[\frac{x_{0,i}}{\sqrt{x_{0,i}^2 + c}} - \frac{(x_i + x_{0,i})}{\sqrt{x_i^2 + c} + \sqrt{x_{0,i}^2 + c}} \right] + f(x_0)$$

For pink term, let $K = \left(\frac{x_{0,1}}{\sqrt{x_{0,1}^2 + c}}, \frac{x_{0,2}}{\sqrt{x_{0,2}^2 + c}}, \dots, \frac{x_{0,n}}{\sqrt{x_{0,n}^2 + c}} \right)^T$

$$\therefore \sum_{i=1}^n (x_i - x_{0,i}) \frac{x_{0,i}}{\sqrt{x_{0,i}^2 + c}} = \langle x - x_0, K \rangle$$

For yellow term, $x_0 \rightarrow x_{0,i}$ as $x \rightarrow x_0$,

$$\frac{x_{0,i}}{\sqrt{x_{0,i}^2 + c}} - \frac{(x_i + x_{0,i})}{\sqrt{x_i^2 + c} + \sqrt{x_{0,i}^2 + c}} \rightarrow 0$$

$$(x_i - x_{0,i}) \left[\frac{x_{0,i}}{\sqrt{x_{0,i}^2 + c}} - \frac{x_i + x_{0,i}}{\sqrt{x_i^2 + c} + \sqrt{x_{0,i}^2 + c}} \right] = O(x_i - x_{0,i})$$

$$\sum_{i=1}^n (x_i - x_{0,i}) \left[\frac{x_{0,i}}{\sqrt{x_{0,i}^2 + c}} - \frac{x_i + x_{0,i}}{\sqrt{x_i^2 + c} + \sqrt{x_{0,i}^2 + c}} \right] = O(\|x - x_0\|)$$

$$\therefore f(x) - f(x_0) - \langle x - x_0, K \rangle = O(\|x - x_0\|)$$

$$\Rightarrow \nabla(f(x_0)) = K = \left(\frac{x_{0,1}}{\sqrt{x_{0,1}^2 + c}}, \frac{x_{0,2}}{\sqrt{x_{0,2}^2 + c}}, \dots, \frac{x_{0,n}}{\sqrt{x_{0,n}^2 + c}} \right)^T$$

$$\nabla f(x) = \left(\frac{x_1}{\sqrt{x_1^2 + c}}, \frac{x_2}{\sqrt{x_2^2 + c}}, \dots, \frac{x_n}{\sqrt{x_n^2 + c}} \right)^T$$

f).

(f) Let $f : \mathbb{R}^m \rightarrow \mathbb{R}$ be differentiable everywhere. Let $A \in \mathbb{R}^{m \times n}$ be a matrix. Let $g : \mathbb{R}^m \rightarrow \mathbb{R}$ be defined by

$$g(x) = f(Ax), \quad \forall x \in \mathbb{R}^n.$$

Express $\nabla g(x)$ in terms of A and gradient of f .

$g(x) = f(Ax)$, by chain rule,

$$g'(x) = f'(Ax) A$$

since gradient is a column vector,

$$\nabla g(x) = g'(x)^T = A^T \nabla f(Ax)$$

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

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

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

2a). $f(x) = e^x - 1$

$\forall x, y \in \mathbb{R}, \alpha \in [0, 1]$,

if $x = y$,

$$f(\alpha x + (1-\alpha)y) = f(\alpha x + (1-\alpha)x) = f(x) =$$

$$\alpha f(x) + (1-\alpha)f(x) = \alpha f(x) + (1-\alpha)f(y)$$

If $x \neq y$,

$$f(\alpha x + (1-\alpha)y) - \alpha f(x) - (1-\alpha)f(y)$$

$$= e^{\alpha x + (1-\alpha)y} - 1 - \alpha(e^x - 1) - (1-\alpha)(e^y - 1)$$

$$= e^{\alpha x + (1-\alpha)y} - \alpha e^x - (1-\alpha)e^y$$

$$= e^{\alpha(x-y)+y} - \alpha(e^x - e^y) - e^y$$

let $g(\alpha) = e^{\alpha(x-y)+y} - \alpha(e^x - e^y) - e^y$

$$g'(\alpha) = (x-y) e^{\alpha(x-y)+y} - (e^x - e^y)$$

$$g''(\alpha) = (x-y)^2 e^{\alpha(x-y)+y} > 0 \quad \forall \alpha \in [0, 1]$$

$$g''(\alpha) = (x-y)^2 e^{\alpha(x-y)+y} > 0 \quad \forall \alpha \in [0,1]$$

$\therefore g$ is minimum at boundary, 0 or 1.

$$g(0) = 0$$

$$g(1) = 0$$

$$\therefore g(\alpha) \leq 0 \quad \forall \alpha \in [0,1]$$

$$\Rightarrow f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)y$$

$\Rightarrow f(x)$ is convex

$$b). \quad f(x) = \sum_{i=1}^n \sqrt{|x_i|}$$

$$\text{let } x = (1, 0, 0, \dots, 0)^T, y = (16, 0, 0, \dots, 0)^T \in \mathbb{R}^n,$$

$$\alpha = \frac{7}{15} \in (0, 1)$$

$$\alpha x + (1-\alpha)y = \frac{7}{15}(1, 0, 0, \dots, 0)^T + \frac{8}{15}(16, 0, 0, \dots, 0)^T$$

$$= (9, 0, 0, \dots, 0)^T$$

$$f(\alpha x + (1-\alpha)y) = f((9, 0, 0, \dots, 0)^T)$$

$$= \sqrt{9} + \sqrt{0} + \sqrt{0} + \dots + \sqrt{0} = 3$$

$$f(x) = f((1, 0, \dots, 0)^T) \quad f(y) = f((16, 0, 0, \dots, 0)^T)$$

$$= \sqrt{16} + \sqrt{0} + \sqrt{0} + \dots + \sqrt{0}$$

$$= 4$$

$$\alpha f(x) + (1-\alpha)f(y) = \frac{7}{15}(1) + \frac{8}{15}(4) = \frac{39}{15} < 3$$

$$\therefore f(\alpha x + (1-\alpha)y) > \alpha f(x) + (1-\alpha)f(y)$$

when $x = (1, 0, \dots, 0)^T$, $y = (16, 0, 0, \dots, 0)^T$, $\alpha = \frac{7}{15}$.

$\Rightarrow f(x)$ is not convex.

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

(a) Let

$$f(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Prove that f is convex.

(b) Let

$$g(\mathbf{x}) = f_1(\mathbf{x}) - f_2(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Give an example of f_1 and f_2 such that g is not convex.

(c) Let

$$h(\mathbf{x}) = \max\{f_1(\mathbf{x}), f_2(\mathbf{x})\}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Prove that h is convex.

(d) Let

$$k(\mathbf{x}) = \min\{f_1(\mathbf{x}), f_2(\mathbf{x})\}, \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

Give an example of f_1 and f_2 such that k is not convex.

a). $f(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x})$

Given f_1 and f_2 are convex

\Rightarrow for any $\alpha, \gamma \in \mathbb{R}^n$, $\alpha \in [0, 1]$,

$$f_1(\alpha x + (1-\alpha)y) \leq \alpha f_1(x) + (1-\alpha)f_1(y)$$

$$f_2(\alpha x + (1-\alpha)y) \leq \alpha f_2(x) + (1-\alpha)f_2(y)$$

$$\begin{aligned} \therefore f(\alpha x + (1-\alpha)y) &\leq [\alpha f_1(x) + (1-\alpha)f_1(y)] + [\alpha f_2(x) + (1-\alpha)f_2(y)] \\ &= \alpha(f_1(x) + f_2(x)) + (1-\alpha)(f_1(y) + f_2(y)) \\ &= \alpha f(x) + (1-\alpha)f(y) \end{aligned}$$

$$= \alpha f(x) + (1-\alpha)f(y)$$

\Rightarrow f is convex function.

b).

I don't know ...

c). $\forall x, y \in \mathbb{R}^n$, $\alpha \in [0, 1]$,

$$h(\alpha x + (1-\alpha)y) = \max \left\{ f_1(\alpha x + (1-\alpha)y), f_2(\alpha x + (1-\alpha)y) \right\}$$

Given $f_1(x)$ and $f_2(x)$ are convex,

$$f_1(\alpha x + (1-\alpha)y) \leq \alpha f_1(x) + (1-\alpha)f_1(y)$$

$$f_2(\alpha x + (1-\alpha)y) \leq \alpha f_2(x) + (1-\alpha)f_2(y)$$

$$h(\alpha x + (1-\alpha)y) \leq \max \left\{ \alpha f_1(x) + (1-\alpha)f_1(y), \alpha f_2(x) + (1-\alpha)f_2(y) \right\}$$

$$\leq \max \left\{ \alpha f_1(x), \alpha f_2(x) \right\} +$$

$$\max \left\{ (1-\alpha)f_1(y), (1-\alpha)f_2(y) \right\}$$

$$\leq \alpha \max \left\{ f_1(x), f_2(x) \right\} + (1-\alpha)$$

$$\max \left\{ f_1(y), f_2(y) \right\}$$

$$= \alpha h(x) + (1-\alpha) h(y)$$

$\Rightarrow h$ is convex

$\Rightarrow h$ is convex

d). let $f_1(x) = x_1, f_2(x) = 0, \forall x \in \mathbb{R}^n$

$$f_1(\alpha x + (1-\alpha)y) = \alpha x_1 + (1-\alpha)y_1 = \alpha f_1(x) + (1-\alpha)f_1(y)$$

$$f_2(\alpha x + (1-\alpha)y) = 0 = \alpha f_2(x) + (1-\alpha)f_2(y)$$

$\Rightarrow f_1, f_2$ are convex.

Let $x = (-1, 0, 0, \dots, 0)^T, y = (1, 0, 0, \dots, 0)^T \in \mathbb{R}^n$,

$$\alpha = 0.5 \in [0, 1],$$

$$K(\alpha x + (1-\alpha)y) = K\left(\frac{1}{2}(-1, 0, 0, \dots, 0)^T + \frac{1}{2}(1, 0, 0, \dots, 0)^T\right)$$

$$= K((0, 0, \dots, 0)^T)$$

$$= \min \{f_1(0, 0, \dots, 0)^T, f_2(0, 0, \dots, 0)^T\}$$

$$= \min \{0, 0\} = 0$$

$$K(x) = K((-1, 0, 0, \dots, 0)^T)$$

$$= \min \{f_1(-1, 0, \dots, 0)^T, f_2(-1, 0, 0, \dots, 0)^T\}$$

$$= \min \{-1, 0\} = -1$$

$$K(y) = K((1, 0, 0, \dots, 0)^T)$$

$$\begin{aligned}
 K(y) &= K((1, 0, 0, \dots, 0)) \\
 &= \min \left\{ f_1(1, 0, 0, \dots, 0)^T, f_2(1, 0, 0, \dots, 0)^T \right\} \\
 &= \min \{1, 0\} = 0
 \end{aligned}$$

$$2K(x) + (1-\alpha)K(y) = \frac{1}{2}(-1) + \frac{1}{2}(0) = -\frac{1}{2} < 0$$

$$\therefore K(\alpha x + (1-\alpha)y) > \alpha K(x) + (1-\alpha)K(y) \Rightarrow$$

K is not convex.