שיטות חישוביות באופטימיזציה 046197
תרגיל בית מספר 2

Alexander Shender: 328626114

Eliran Cohen: 204187801

תרגיל מספר 1 :

משפט : f קמורה אם ורק אם $\phi(t) = f(x + t\mathrm{d})$ קמורה לכל $x, \mathrm{d} \in R^n$

נוכיח את המשפט בשני הכיוונים

א. $\phi(t)$ קמורה

נרצה להראות שפונקציה f קמורה לפי הגדרה, לכן עבור שתי נקודות $x, y \in \mathbb{R}^n$ ו - $\lambda \in [0,1]$

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

נסמן : $\mathrm{d} = y - x$ ונקבל פונקציה מהצורה: $\phi(t) = f(x + t\mathrm{d})$ כלומר,

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= f\big(x + (1-\lambda)(y-x)\big) = f(x + (1-\lambda)\mathrm{d}) = \phi(1-\lambda) = \phi\big((1-\lambda)\cdot 1 + \lambda \cdot 0\big) \\ &\le \lambda\, \phi(0) + (1-\lambda)\, \phi(1) = \lambda f(x + 0\cdot \mathrm{d}) + (1-\lambda) f(x + 1 \cdot \mathrm{d}) \\ &= \lambda f(x) + (1-\lambda) f(x + y - x) = \lambda f(x) + (1-\lambda) f(y) \end{aligned}$$

כלומר הוכחנו שאם $\phi(t)$ קמורה פונקציה f קמורה לפי הגדרה

ב. f קמורה

נרצה להראות שפונקציה $\phi(t)$ קמורה לפי הגדרה, לכן עבור שתי נקודות $a, \mathrm{b} \in \mathbb{R}$ ו - $\lambda \in [0,1]$

$$\begin{aligned} \phi(\lambda a + (1-\lambda)b) &= f(x + (\lambda a + (1-\lambda)b)\mathrm{d}) = f(x + \lambda a\mathrm{d} + b\mathrm{d} - \lambda b\mathrm{d}) \\ &= f(x + \lambda x - \lambda x + \lambda a\mathrm{d} + b\mathrm{d} - \lambda b\mathrm{d}) = f\big(\lambda(x + a\mathrm{d}) + (1-\lambda)(x + b\mathrm{d})\big) \\ &\le \lambda f(x + a\mathrm{d}) + (1-\lambda) f(x + b\mathrm{d}) = \lambda\, \phi(a) + (1-\lambda)\, \phi(b) \end{aligned}$$

כלומר הוכחנו שאם f קמורה פונקציה $\phi(t)$ קמורה לפי הגדרה

תרגיל מספר 2 :

א. חישוב הגרדיאנט וההיסאן של $g(x)$

$$g(x) = f(x,x) = \left(\frac{1}{2}x^T Qx\right)\left(\frac{1}{2}x^T Rx\right)$$

$$\nabla g(x) = \frac{\partial}{\partial x}\left[\left(\frac{1}{2}x^T Qx\right)\left(\frac{1}{2}x^T Rx\right)\right] = \left(\frac{1}{2}x^T Rx\right)\frac{\partial}{\partial x}\left(\frac{1}{2}x^T Qx\right) + \left(\frac{1}{2}x^T Qx\right)\frac{\partial}{\partial x}\left(\frac{1}{2}x^T Rx\right)$$
$$= \frac{1}{2}[(x^T Rx)Qx + (x^T Qx)Rx]$$

$$\nabla^2 g(x) = \frac{\partial}{\partial x}\left[\frac{1}{2}[(x^T Rx)Qx + (x^T Qx)Rx]\right]$$
$$= \frac{1}{2}\left[(x^T Rx)\frac{\partial}{\partial x}[Qx] + \frac{\partial}{\partial x}[(x^T Rx)](Qx)^T + (x^T Qx)\frac{\partial}{\partial x}[Rx] + \frac{\partial}{\partial x}[(x^T Rx)](Rx)^T\right]$$
$$= \frac{1}{2}[x^T RxQ + 2Rxx^T Q + x^T QxR + 2Qxx^T R] = \frac{1}{2}x^T RxQ + \frac{1}{2}x^T QxR + Rxx^T Q + Qxx^T R$$

ב. האם $f(x,y)$ קמורה בהינתן ש - $g(x)$ אינה קמורה

אם $g(x)$ אינה קמורה אז מתקיים :

$$g(\lambda x + (1-\lambda)y) > \lambda g(x) + (1-\lambda)g(y)$$

נראה איך זה משפיע על $f(x,y)$:

$$f(\lambda x + (1-\lambda)y, \lambda x + (1-\lambda)y) = g(\lambda x + (1-\lambda)y) > \lambda g(x) + (1-\lambda)g(y) = \lambda f(x,x) + (1-\lambda)f(y,y)$$

כלומר קיימים x,y בהם הפונקציה לא קמורה ולכן f אינה קמורה.

ג. הוכחה ש - $g(x)$ קמורה אם $R = \alpha Q\ , \alpha > 0$

נבדוק אם מטריצת ההיסאן היא אי שלילית מוגדרת לכל x :

$$\nabla^2 g(x) = \tfrac{1}{2}x^T RxQ + \tfrac{1}{2}x^T QxR + Rxx^T Q + Qxx^T R = \tfrac{1}{2}\alpha x^T QxQ + \tfrac{1}{2}\alpha x^T QxQ + \alpha Qxx^T Q + \alpha Qxx^T Q =$$
$$= \alpha(x^T Qx)Q + 2\alpha Qx(Qx)^T$$

נראה את השפעת כלל הגורמים על ההיסאן :

1. $x^T Qx$ הוא סקלר חיובי לכל x מכיוון ש - $Q \succ 0$
2. $\alpha(x^T Qx)Q$ היא חיובית מוגדרת מכיוון שזו תוצאה של מכפלת סקלר חיובי במטריצה חיובית מוגדרת
3. $Qx(Qx)^T$ היא מטריצה חיובית מוגדרת מכיוון שהיא מכילה את הערכים הריבועיים של Qx
4. $2\alpha Qx(Qx)^T$ היא חיובית מוגדרת מכיוון שזו תוצאה של מכפלת סקלר חיובי במטריצה חיובית מוגדרת
5. $\alpha(x^T Qx)Q + 2\alpha Qx(Qx)^T$ היא חיובית מוגדרת מכיוון שהיא סכום של שתי מטריצות חיוביות מוגדרות

כלומר $\nabla^2 g(x) = \alpha(x^T Qx)Q + 2\alpha Qx(Qx)^T \succ 0$ לכל x ולכן $g(x)$ קמורה

ד. בדיקה האם $g(x)$ קמורה עבור מטריצות Q ו – R הנתונות

תחילה נראה כי $R, Q \succ 0$ על ידי מציאת הערכים העצמיים של המטריצות :

$$Q = \begin{pmatrix} 3 & -2 \\ -2 & 3 \end{pmatrix} \to \det[Q - \lambda I] = \begin{vmatrix} 3-\lambda & -2 \\ -2 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 - 4 \to \lambda_1 = 5\ , \lambda_2 = 1 \to \lambda_1, \lambda_2 > 0 \to Q \succ 0$$

$$R = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} \to \det[Q - \lambda I] = \begin{vmatrix} 4-\lambda & 3 \\ 3 & 4-\lambda \end{vmatrix} = (4-\lambda)^2 - 9 \to \lambda_1 = 7\ , \lambda_2 = 1 \to \lambda_1, \lambda_2 > 0 \to R \succ 0$$

נבדוק קמירות על ידי ההיסאן :

$$\nabla^2 g(x) = \frac{1}{2}x^T RxQ + \frac{1}{2}x^T QxR + Rxx^T Q + Qxx^T R$$

נבדוק גורם גורם בסכום על ווקטור $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$:

$$\frac{1}{2}x^T RxQ = \frac{1}{2}(x_1 \quad x_2)\begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\begin{pmatrix} 3 & -2 \\ -2 & 3 \end{pmatrix} = \frac{1}{2}(4x_1 + 3x_2 \quad 3x_1 + 4x_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\begin{pmatrix} 3 & -2 \\ -2 & 3 \end{pmatrix}$$
$$= (2x_1^2 + 3x_1x_2 + 2x_2^2)\begin{pmatrix} 3 & -2 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 6x_1^2 + 9x_1x_2 + 6x_2^2 & -4x_1^2 - 6x_1x_2 - 2x_2^2 \\ -4x_1^2 - 6x_1x_2 - 2x_2^2 & 6x_1^2 + 9x_1x_2 + 6x_2^2 \end{pmatrix}$$

$$\frac{1}{2}x^T QxR = \frac{1}{2}(x_1 \quad x_2)\begin{pmatrix} 3 & -2 \\ -2 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} = \frac{1}{2}(3x_1 - 2x_2 \quad -2x_1 + 3x_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}$$
$$= \frac{1}{2}(3x_1^2 - 4x_1x_2 + 3x_2^2)\begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 12x_1^2 - 16x_1x_2 + 12x_2^2 & 9x_1^2 - 12x_1x_2 + 9x_2^2 \\ 9x_1^2 - 12x_1x_2 + 9x_2^2 & 12x_1^2 - 16x_1x_2 + 12x_2^2 \end{pmatrix}$$

$$Rxx^T Q = \begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}(x_1 \quad x_2)\begin{pmatrix} 3 & -2 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 4x_1 + 3x_2 \\ 3x_1 + 4x_2 \end{pmatrix}(3x_1 - 2x_2 \quad -2x_1 + 3x_2)$$
$$= \begin{pmatrix} 12x_1^2 + x_1x_2 - 6x_2^2 & -8x_1^2 + 6x_1x_2 + 9x_2^2 \\ 9x_1^2 + 6x_1x_2 - 8x_2^2 & -6x_1^2 + x_1x_2 + 12x_2^2 \end{pmatrix}$$

$$Qxx^T R = (Rxx^T Q)^T = \begin{pmatrix} 12x_1^2 + x_1x_2 - 6x_2^2 & -8x_1^2 + 6x_1x_2 + 9x_2^2 \\ 9x_1^2 + 6x_1x_2 - 8x_2^2 & -6x_1^2 + x_1x_2 + 12x_2^2 \end{pmatrix}^T =$$
$$= \begin{pmatrix} 12x_1^2 + x_1x_2 - 6x_2^2 & 9x_1^2 + 6x_1x_2 - 8x_2^2 \\ -8x_1^2 + 6x_1x_2 + 9x_2^2 & -6x_1^2 + x_1x_2 + 12x_2^2 \end{pmatrix}$$

$$Rxx^T Q + Qxx^T R = \begin{pmatrix} 24x_1^2 + 2x_1x_2 - 12x_2^2 & x_1^2 + 12x_1x_2 + x_2^2 \\ x_1^2 + 12x_1x_2 + x_2^2 & -12x_1^2 + 2x_1x_2 + 24x_2^2 \end{pmatrix}$$

נפריך בדוגמא נגדית עבור $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$:

$$\frac{1}{2}x^T RxQ = \begin{pmatrix} 6 & -4 \\ -4 & 6 \end{pmatrix}$$

$$\frac{1}{2}x^T QxR = \frac{1}{2}\begin{pmatrix} 12 & 9 \\ 9 & 12 \end{pmatrix}$$

$$Rxx^T Q + Qxx^T R = \begin{pmatrix} 24 & 1 \\ 1 & -12 \end{pmatrix}$$

$$\nabla^2 g(x) = \frac{1}{2}x^T RxQ + \frac{1}{2}x^T QxR + Rxx^T Q + Qxx^T R = \begin{pmatrix} 6 & -4 \\ -4 & 6 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 12 & 9 \\ 9 & 12 \end{pmatrix} + \begin{pmatrix} 24 & 1 \\ 1 & -12 \end{pmatrix}$$
$$= \begin{pmatrix} 36 & 1.5 \\ 1.5 & 0 \end{pmatrix}$$

$$\nabla^2 g(x) = \begin{pmatrix} 36 & 1.5 \\ 1.5 & 0 \end{pmatrix} \to \det[\nabla^2 g(x) - \lambda I] = \begin{vmatrix} 36 - \lambda & 1.5 \\ 1.5 & -\lambda \end{vmatrix} = \lambda(\lambda - 36) - \frac{9}{4} \to \lambda_1 = 36.06, \lambda_2 = -0.06$$
$$\to \lambda_2 < 0 \to Q \ not \ PSD$$

$\nabla^2 g(x)$ לא אי שלילית מוגדרת , כלומר $g(x)$ לא קמורה.

תרגיל מספר 3 :

א. הוכחה ש - $f(x)$ קמורה עבור $0 \le x < 1$

נגזור את הפונקציה פעמיים כדי לראות אם היא אי שלילית בתחום הנתון :

$$f(x) = \ln\left(\frac{1+x}{1-x}\right) = \ln(1+x) - \ln(1-x)$$

$$\frac{d}{dx}f(x) = \frac{d}{dx}[\ln(1+x) - \ln(1-x)] = \frac{1}{1+x} + \frac{1}{1-x} = \frac{2}{(1+x)(1-x)}$$

$$\frac{d^2}{dx^2}f(x) = \frac{d}{dx}\left[\frac{1}{1+x} + \frac{1}{1-x}\right] = -\frac{1}{(1+x)^2} + \frac{1}{(1-x)^2} = \frac{4x}{\left((1+x)(1-x)\right)^2}$$

הפונקציה מוגדרת בתחום $-1 \le x < 1$ והנגזרת השניה אי שלילית בתחום : $x \ge 0$

כלומר הפונקציה קמורה בתחום : $0 \le x < 1$

ב. נראה כי מתקיים $f(x) \ge 2x$ עבור $0 \le x < 1$

נתבונן בפונקציה $q(x)$ המתארת את ההפרש בין הפונקציות :

$$q(x) = \ln\left(\frac{1+x}{1-x}\right) - 2x$$

נראה כי הפונקציה אי שלילית בתחום : $0 \le x < 1$.

$$q'^{(x)} = \frac{d}{dx}\left[\ln\left(\frac{1+x}{1-x}\right) - 2x\right] = \frac{2}{(1+x)(1-x)} - 2 = \frac{2x^2}{(1+x)(1-x)}$$

הנגזרת חיובת בתחום $0 \le x < 1$ כלומר בתחום זה היא עולה

כמו כן נשים לב ש :

$$q(0) = 0$$

לכן ההפרש בין הפונקציות מתחיל ב-0 וההפרש גדל בתחום $0 \le x < 1$ כלומר מתקיים $f(x) \ge 2x$.

ג. הראו כי $h(x)$ קעורה עבור $0 \le x < 1$

$$h(x) = -x\ln(x) - (1-x)\ln(1-x)$$

נתבונן בפונקציה :

$$p(x) = -h(x) = x\,ln(x) + (1-x)\,ln(1-x)$$

הפונקציה $x\,ln(x)$ היא פונקציה קמורה ועולה בתחום הנתון

הפונקציה $1 - x$ היא פונקציה קמורה

ולכן הרכבת שתי הפונקציות $(1-x)\,ln(1-x)$ היא פונקציה קמורה בתחום הנתון.

לכן הפונקציה $p(x) = x\,ln(x) + (1-x)\,ln(1-x)$ היא פונקציה קמורה כסכום של שתי פונקציות קמורות עם מקדמים חיוביים.

לכן , הפוקנציה הנגדית : $h(x) = -p(x) = -x\,ln(x) - (1-x)\,ln(1-x)$ היא פונקציה קעורה בתחום הנתון.

ד. הוכחה כי $g(x)$ היא פונקציה קמורה בתחום $0 \le x < 1$

$$g(x) = h\left(\frac{1-\sqrt{1-x^2}}{2}\right)$$

נחשב את הנגזרת השניה של הפונקציה כדי להראות שהיא תמיד אי שלילית

$$g'(x) = h'\left(\frac{1-\sqrt{1-x^2}}{2}\right)\cdot\frac{2x}{4\sqrt{1-x^2}} = h'\left(\frac{1-\sqrt{1-x^2}}{2}\right)\frac{x}{2\sqrt{1-x^2}}$$

$$g''(x) = h''\left(\frac{1-\sqrt{1-x^2}}{2}\right)\cdot\frac{2x}{4\sqrt{1-x^2}}\cdot\frac{x}{2\sqrt{1-x^2}} + h'\left(\frac{1-\sqrt{1-x^2}}{2}\right)\cdot\left[\frac{2\sqrt{1-x^2}-x\frac{-2x}{\sqrt{1-x^2}}}{4(1-x^2)}\right]$$

$$= h''\left(\frac{1-\sqrt{1-x^2}}{2}\right)\frac{x^2}{4(1-x^2)} + h'\left(\frac{1-\sqrt{1-x^2}}{2}\right)\cdot\frac{2(1-x^2)+2x^2}{4(1-x^2)\sqrt{1-x^2}} =$$

$$= h''\left(\frac{1-\sqrt{1-x^2}}{2}\right)\frac{x^2}{4(1-x^2)} + h'\left(\frac{1-\sqrt{1-x^2}}{2}\right)\cdot\frac{1}{2(1-x^2)\sqrt{1-x^2}}$$

נחשב את הנגזרות של $h(x)$

$$h(x) = -x\ln(x) - (1-x)\ln(1-x)$$

$$h'(x) = -[\ln(x) + 1 - \ln(1-x) - 1] = \ln(1-x) - \ln(x)$$

$$h''(x) = \frac{-1}{1-x} - \frac{1}{x} = \frac{1}{x(x-1)}$$

$$h'\left(\frac{1-\sqrt{1-x^2}}{2}\right) = \ln\left(1-\frac{1-\sqrt{1-x^2}}{2}\right) - \ln\left(\frac{1-\sqrt{1-x^2}}{2}\right) = \ln\left(\frac{1+\sqrt{1-x^2}}{1-\sqrt{1-x^2}}\right)$$

$$h''\left(\frac{1-\sqrt{1-x^2}}{2}\right) = \frac{1}{\left(\frac{1-\sqrt{1-x^2}}{2}\right)\left(\frac{1-\sqrt{1-x^2}}{2}-1\right)} = \frac{-4}{(1-\sqrt{1-x^2})(1+\sqrt{1-x^2})} = \frac{-4}{x^2}$$

נציב ב - $g''(x)$

$$g''(x) = \frac{-4}{x^2}\cdot\frac{x^2}{4(1-x^2)} + \ln\left(\frac{1+\sqrt{1-x^2}}{1-\sqrt{1-x^2}}\right)\cdot\frac{1}{2(1-x^2)\sqrt{1-x^2}}$$

$$= \frac{-1}{(1-x^2)} + \ln\left(\frac{1+\sqrt{1-x^2}}{1-\sqrt{1-x^2}}\right)\frac{1}{2(1-x^2)\sqrt{1-x^2}}$$

בסעיף ב׳ הוכחנו כי מתקיים בתחום $0 \le x < 1$:

$$\ln\left(\frac{1+x}{1-x}\right) \ge 2x$$

התחום הוא אותו תחום ולכן נציב ב – $g''(x)$ כחסם תחתון את :

$$\ln\left(\frac{1+\sqrt{1-x^2}}{1-\sqrt{1-x^2}}\right) \ge 2\sqrt{1-x^2}$$

$$g''(x) = \frac{-1}{(1-x^2)} + \ln\left(\frac{1+\sqrt{1-x^2}}{1-\sqrt{1-x^2}}\right)\frac{1}{2(1-x^2)\sqrt{1-x^2}} \geq \frac{-1}{(1-x^2)} + 2\sqrt{1-x^2}\,\frac{1}{2(1-x^2)\sqrt{1-x^2}}$$
$$= \frac{-1}{(1-x^2)} + \frac{1}{(1-x^2)} = 0$$

הוכחנו כי : $g''^{(x)} \geq 0$ בתחום $0 \leq x < 1$ ולכן בתחום זה הפונקציה קמורה.

Question 4.

Prove that the following function is convex for $x > 0 \,; x \in R$:

$$g(x) = -\prod_{i=1}^{n} x_i^{p_i}, \quad p_i \geq 0 \,; \sum_i p_i = 1$$

We will use the composition statement from the book from Stephen Boyd. (page 86), which says:

If $f(x)$ is concave and positive, then $1/f(x)$ is convex.

Indeed: Let's take $f(x)$ concave positive function ($f(x) > 0$). The composition function will be $h(x) = \frac{1}{x}$, which is convex ($h''^{(x)} > 0$), and decreasing ($h'^{(x)} < 0$) for $x \in R^+$. Then:

$$g(x) = h\big(f(x)\big) = \frac{1}{f(x)}$$

Showing that its second derivative is positive:

$$g''(x) = \underbrace{h''\big(f(x)\big)}_{>0}\underbrace{f'(x)^2}_{>0} + \underbrace{h'\big(f(x)\big)}_{<0}\underbrace{g''(x)}_{<0} > 0$$

In our case, proving that $-\prod_{i=1}^{n} x_i^{p_i}$ is convex is equal to proving that $\prod_{i=1}^{n} x_i^{p_i}$ is concave.

From the previous statement, since $\prod_{i=1}^{n} x_i^{p_i}$ is positive ($x > 0 \,;\ p_i \geq 0 \ \forall\, i$) and concave, it is equal to proving that $\frac{1}{\prod_{i=1}^{n} x_i^{p_i}}$ is convex. Rewriting. Need to prove:

$$g_2(x) = \left(\prod_{i=1}^{n} x_i^{p_i}\right)^{-1} = \prod_{i=1}^{n} x_i^{-p_i}$$

Is convex.

Using exp and log trick to turn multiplication into sum:

$$g_2(x) = e^{\log\left(\prod_{i=1}^{n} x_i^{-p_i}\right)} = e^{\sum_{i=1}^{n} \log\left(x_i^{-p_i}\right)} = e^{\sum_{i=1}^{n} -\log\left(x_i^{p_i}\right)}$$

Since log is a concave function for positive values, thus $-\log$ is convex. So,

$$y = \sum_{i=1}^{n} -\log\left(x_i^{p_i}\right)$$

Is the sum of convex functions, thus is also convex.

Further, e^y is also convex, since it is a composition of 2 convex functions, where y is convex, and e is convex and non-decreasing. Thus, we prove that $g_2(x)$ is convex, thus $g(x)$ is also convex.

Another way to prove it is to use the Hessian matrix of the original $g(x)$ function, and prove that it's PSD.

Question 5.

Need to prove:

$$f(p_1x_1 + p_2x_2) \geq p_1f(x_1) + p_2f(x_2)\,; \quad p_1 + p_2 = 1\,; \;\; p_1 > 0 \;\; ; \;\; p_2 < 0$$

Let's assume there's a point $x_1 \in U$. Extending:

$$x_1 = \frac{p_1x_1 + p_2x_2 - p_2x_2}{p_1}$$

Thus,

$$f(x_1) = f\left(\frac{p_1x_1 + p_2x_2 - p_2x_2}{p_1}\right)$$

Since the function f is convex, the following inequality holds:

$$f\left(\frac{p_1x_1 + p_2x_2 - p_2x_2}{p_1}\right) \leq \underbrace{\left(\frac{1}{p_1}\right)}_{t} f(p_1x_1 + p_2x_2) + \underbrace{\left(\frac{-p_2}{p_1}\right)}_{1-t} f(x_2)$$

Only if the conditions for t is met $(0 < t < 1)$. And that those coefficients sum to 1.

And indeed:

$$t + (1 - t) = \left(\frac{1}{p_1}\right) + \left(\frac{-p_2}{p_1}\right) = \frac{1 - p_2}{p_1} = \frac{p_1}{p_1} = 1$$

And:

$$t = \frac{1}{p_1} \;\; ; \;\; 0 < \frac{1}{p_1} < 1$$

So we get:

$$f(x_1) \leq \left(\frac{1}{p_1}\right) f(p_1x_1 + p_2x_2) + \left(\frac{-p_2}{p_1}\right) f(x_2)$$

Multiply by p_1:

$$p_1f(x_1) \leq f(p_1x_1 + p_2x_2) - p_2f(x_2)$$

$$p_1f(x_1) + p_2f(x_2) \leq f(p_1x_1 + p_2x_2)$$

Which is what is required.

Question 6.

Given function in $I \in R.\ x \neq y.$

$$\Delta_f(x,y) = \frac{f(x) - f(y)}{x - y}$$

a.

$$\Delta_f(x,y) = \frac{f(x) - f(y)}{x - y} = \frac{-\big(f(y) - f(x)\big)}{-(y - x)} = \frac{f(y) - f(x)}{y - x} = \Delta_f(y,x)$$

---

b.

First direction.

Expressing x_2 with the convex combination of x_1 ,x_3:

$$x_2 = tx_1 + (1 - t)x_3$$

We can then express t and $(1 - t)$:

$$x_2 = tx_1 + x_3 - tx_3$$

$$x_2 - x_3 = t(x_1 - x_3)$$

$$t = \frac{x_2 - x_3}{x_1 - x_3}$$

$$1 - t = \frac{x_1 - x_2}{x_1 - x_3}$$

The function at point x_2:

$$f(x_2) = f(tx_1 + (1 - t)x_3)$$

It is given that the function $f(x)$ is convex, thus the following holds:

$$f(tx_1 + (1 - t)x_3) \leq tf(x_1) + (1 - t)f(x_3)$$

Inserting t and $f(x_2)$ back:

$$f(x_2) \geq \left(\frac{x_2 - x_3}{x_1 - x_3}\right) f(x_1) + \left(\frac{x_1 - x_2}{x_1 - x_3}\right) f(x_3)$$

We multiply by $x_1 - x_3$. Since $x_1 < x_3$, it is a negative number, so we change the direction of inequality.

$$f(x_2)(x_1 - x_3) \geq (x_2 - x_3)f(x_1) + (x_1 - x_2)f(x_3)$$

$$(x_3 - x_2)f(x_1) + f(x_2)(x_1 - x_3) + f(x_3)(x_2 - x_1) \geq 0$$

Which proves the statement.

Second direction

Given some 3 points x_1, x_2, x_3, which satisfy $x_1 < x_2 < x_3$ and satisfy:

$$(x_3 - x_2)f(x_1) + f(x_2)(x_1 - x_3) + f(x_3)(x_2 - x_1) \geq 0$$

Prove that $f(x)$ is convex.

Solution:

We will try to use the same characteristic of the convex function. Since $x_1 < x_2 < x_3$, we can express x_2 as a combination of x_1, x_3.:

$$x_2 = tx_1 + (1 - t)x_3$$

We will use the same derivations that we got from previous direction proof:

$$t = \frac{x_2 - x_3}{x_1 - x_3}$$

$$1 - t = \frac{x_1 - x_2}{x_1 - x_3}$$

From this expression:

$$(x_3 - x_2)f(x_1) + f(x_2)(x_1 - x_3) + f(x_3)(x_2 - x_1) \geq 0$$

Divide both by $(x_1 - x_3)$:

$$f(x_2) + \left(\frac{x_3 - x_2}{x_1 - x_3}\right) f(x_1) + \left(\frac{x_2 - x_1}{x_1 - x_3}\right) f(x_3) \geq 0$$

$$f(x_2) \geq -\left(\frac{x_3 - x_2}{x_1 - x_3}\right) f(x_1) - \left(\frac{x_2 - x_1}{x_1 - x_3}\right) f(x_3)$$

$$f(x_2) \geq \left(\frac{x_2 - x_3}{x_1 - x_3}\right) f(x_1) + \left(\frac{x_1 - x_2}{x_1 - x_3}\right) f(x_3)$$

$$f(x_2) \geq tf(x_1) + (1 - t)f(x_3)$$

Putting x_2:

$$f(tx_1 + (1 - t)x_3) \geq tf(x_1) + (1 - t)f(x_3)$$

And since $x_1 < x_3$, it proves that the function $f(x)$ is convex.

c.

If $f(x)$ is convex and $x_1 < x_2 < x_3$, then this holds:

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(x_3)-f(x_1)}{x_3-x_1} \leq \frac{f(x_3)-f(x_2)}{x_3-x_2}$$

We can use the expression from previous paragraph, reorganize it, and get:

$$(x_3-x_2)f(x_1)+f(x_2)(x_1-x_3)+f(x_3)(x_2-x_1) \geq 0$$

$$x_3f(x_1)-x_2f(x_1)+f(x_2)(x_1-x_3)+f(x_3)(x_2-x_1) \geq 0$$

Add and subtract $f(x_1)x_1$:

$$x_3f(x_1)-f(x_1)x_1+f(x_1)x_1-x_2f(x_1)+f(x_2)(x_1-x_3)+f(x_3)(x_2-x_1) \geq 0$$

$$f(x_1)(x_3-x_1)+f(x_1)(x_1-x_2)+f(x_2)(x_1-x_3)+f(x_3)(x_2-x_1) \geq 0$$

$$(f(x_2)-f(x_1))(x_1-x_3)+(f(x_3)-\ f(x_1))(x_2-x_1) \geq 0$$

Divide by $(x_1-x_3)(x_2-x_1)$ (which is negative, changing sign):

$$\frac{f(x_2)-f(x_1)}{(x_2-x_1)}+\frac{f(x_3)-\ f(x_1)}{(x_1-x_3)} \leq 0$$

$$\frac{f(x_2)-f(x_1)}{(x_2-x_1)}-\frac{f(x_3)-\ f(x_1)}{(x_3-x_1)} \leq 0$$

$$\frac{f(x_2)-f(x_1)}{(x_2-x_1)} \leq \frac{f(x_3)-\ f(x_1)}{(x_3-x_1)}$$

Which proves the left part of the inequality.

To obtain the second part of the inequality, instead of adding and subtracting $f(x_1)x_1$, we add and subtract $f(x_3)x_3$. And follow same steps.

d.

Given: $x < y < z < w$

Function $f(x)$ is convex.

Prove: $a \leq b \leq c \leq d \leq e$

We use the results from paragraph 3. From the definition of the gradient (approximation):

$$\Delta_f(x, y) = \frac{f(x) - f(y)}{x - y}$$

So we can mark the gradients:

$$a = \frac{f(x) - f(y)}{x - y} = \frac{f(y) - f(x)}{y - x}$$

$$b = \frac{f(x) - f(z)}{x - z}$$

$$c = \frac{f(y) - f(z)}{y - z}$$

$$d = \frac{f(y) - f(w)}{y - w}$$

$$e = \frac{f(z) - f(w)}{z - w}$$

Since: $x < y < z$ and using the inequation from paragraph 3:

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(x) - f(z)}{x - z}$$

$$a \leq b$$

$$\frac{f(x) - f(z)}{x - z} \leq \frac{f(y) - f(z)}{y - z}$$

$$b \leq c$$

Since $y < z < w$ and using the inequation from paragraph 3:

$$\frac{f(y) - f(z)}{y - z} \leq \frac{f(y) - f(w)}{y - w}$$

$$c \leq d$$

$$\frac{f(y) - f(w)}{y - w} \leq \frac{f(z) - f(w)}{z - w}$$

$$d \leq e$$

Summing all together, we obtain:

$$a \leq b \leq c \leq d \leq e$$

Which is what was required.

e.

Prove that if $f(x)$ convex, if and only if for all $x \neq y \in I$; $g_y(x) = \Delta_f(x, y)$ is not decreasing.

Solution:

Direction 1

$f(x)$ is convex. We will use the previous statement from paragraph 3 again.

Case 1.

We assume $y < x_2 < x_1;\ x_2, x_1 \in I$.

Then:

$$\frac{f(x_2) - f(y)}{x_2 - y} \leq \frac{f(x_1) - f(y)}{x_1 - y} \leq \frac{f(x_1) - f(x_2)}{x_1 - x_2}$$

$$\Delta_f(x_2, y) \leq \Delta_f(x_1, y) \leq \Delta_f(x_1, x_2)$$

$$g_y(x_2) \leq g_y(x_1) \leq g_{x_2}(x_1)$$

Thus: $g_y(x_2) \leq g_y(x_1)$

Which means that the gradient is not decreasing.

Case 2.

We assume $x_2 < y < x_1;\ x_2, x_1 \in I$.

Then:

$$\frac{f(y) - f(x_2)}{y - x_2} \leq \frac{f(x_1) - f(x_2)}{x_1 - x_2} \leq \frac{f(x_1) - f(y)}{x_1 - y}$$

$$\Delta_f(x_2, y) \leq \Delta_f(x_1, x_2) \leq \Delta_f(x_1, y)$$

$$g_y(x_2) \leq g_{x_1}(x_2) \leq g_y(x_1)$$

Thus, $g_y(x_2) \leq g_y(x_1)$

Which means the gradient is not decreasing.

Case 3.

We assume $x_2 < x_1 < y\ ;\ x_2, x_1 \in I$.

Then:

$$\frac{f(x_1)-f(x_2)}{x_1-x_2} \le \frac{f(y)-f(x_2)}{y-x_2} \le \frac{f(y)-f(x_1)}{y-x_1}$$

$$\Delta_f(x_2,x_1) \le \Delta_f(x_2,y) \le \Delta_f(x_1,y)$$

$$g_{x_1}(x_2) \le g_y(x_2) \le g_y(x_1)$$

Thus, $g_y(x_2) \le g_y(x_1)$

Which means the gradient is not decreasing.

Direction 2

$g_y(x) = \Delta_f(x,y)$ is not decreasing.

Let's take 3 points, s.t. $x_1 < y < x_3$

From the property of not decreasing gradient, we get:

$$g_y(x_1) \le g_{x_1}(x_3) \le g_y(x_3)$$

Which means:

$$\frac{f(y)-f(x_1)}{y-x_1} \le g_{x_1}(x_3) \le \frac{f(y)-f(x_3)}{y-x_3}$$

$$\frac{f(y)-f(x_1)}{y-x_1} \le \frac{f(y)-f(x_3)}{y-x_3}$$

We multiply by $(y-x_1)(y-x_3)$ which is negative. Change sign.

$$\big(f(y)-f(x_1)\big)(y-x_3) \ge \big(f(y)-f(x_3)\big)(y-x_1)$$

We can express y as a convex combination of x_1 and x_3:

$$y = tx_1 + (1-t)x_3$$

Then:

$$(y-x_3) = \ (tx_1 + (1-t)x_3 - x_3) = (tx_1 - tx_3) = t(x_1 - x_3)$$

$$(y-x_1) = (tx_1 + (1-t)x_3 - x_1) = \big((1-t)x_3 - x_1(1-t)\big) = (t-1)(x_1-x_3)$$

Pasting:

$$\big(f(y)-f(x_1)\big)t(x_1-x_3) \ \ge \big(f(y)-f(x_3)\big)(t-1)(x_1-x_3)$$

$$f(y)\big(t(x_1-x_3) - (t-1)(x_1-x_3)\big) \ge f(x_1)t(x_1-x_3) + f(x_3)(1-t)(x_1-x_3)$$

$$f(y)(x_1-x_3) \ge f(x_1)t(x_1-x_3) + f(x_3)(1-t)(x_1-x_3)$$

We can divide both parts by (x_1-x_3). Since its negative, we change the sign:

$$f(y) \le f(x_1)t + f(x_3)(1-t)$$

And replace the y with the declaration:

$$f(tx_1 + (1-t)x_3) \leq tf(x_1) + (1-t)f(x_3)$$

Which is the basic characteristic of the convex function.

Thus, $f(x)$ is convex.