Devoir maison d'analyse

Le but de ce petit problème est d'étudier les fonctions convexes. À partir de la définition géométrique, on démontrera les propriétés de base (inégalités des "pentes croissantes", résultats sur la dérivabilité et la continuité). On démontrera ensuite une inégalité plus générale, dont une conséquence sera une comparaison entre la moyenne arithmétique et la moyenne géométrique.

Dans tout le problème, I désigne un intervalle de $\mathbb{R}$ avec au moins deux points.

1 Définition et premiers exemples

Définition : Soit f une application de I dans $\mathbb{R}$.
On dit que f est une fonction convexe sur I si

$$\forall (x,y) \in I^2, \forall t \in [0,1], f(tx+(1-t)y) \leq tf(x)+(1-t)f(y). \qquad (*)$$

Définition : Pour $x < y$ dans I, on appelle *arc du graphe de* f entre x et y l'ensemble des points de coordonnées $(z, f(z))$ où z décrit l'intervalle $[x,y]$. Pour $x < y$ dans I, on appelle *corde du graphe de* f entre x et y le segment d'extrémités $(x, f(x))$ et $(y, f(y))$.

Interprétation géométrique : Une fonction est convexe sur I si et seulement si tout arc du graphe de f est en-dessous de sa corde.

Définition : Soit I un intervalle de $\mathbb{R}$ et f une application de I dans $\mathbb{R}$.
On dit que f est une fonction concave sur I si

$$\forall (x,y) \in I^2, \forall t \in [0,1], f(tx+(1-t)y) \geq tf(x)+(1-t)f(y).$$

Question 1.1 Déterminer géométriquement lesquelles des fonctions suivantes sont convexes :
a) $f : \mathbb{R} \to \mathbb{R}, f(x) = e^x$ si $x \in \mathbb{R}$.
b) $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$ si $x \in \mathbb{R}$.
c) $f : \mathbb{R} \to \mathbb{R}, f(x) = x^3$ si $x \in \mathbb{R}$.
d) $f : \mathbb{R}^+ \to \mathbb{R}, f(x) = x^3$ si $x \in \mathbb{R}^+$.
e) $f : [-1,1] \to \mathbb{R}, f(-1) = 2, f(1) = 2$ et $f(x) = x^2$ si $x \in]-1,1[$.
f) $f : [-1,1] \to \mathbb{R}, f(-1) = 0, f(1) = 0$ et $f(x) = -x^2$ si $x \in]-1,1[$.
g) $f : [-1,1] \to \mathbb{R}, f(x) = 0$ si $x < 0$ et $f(x) = x$ si $x \geq 0$.
(On tracera l'allure des graphes et si la fonction n'est pas convexe, on tracera une corde le prouvant.)

Question 1.2 Soit f une application de I dans $\mathbb{R}$. Montrer que f est convexe si et seulement si $-f$ est concave.

Question 1.3 Donner l'interprétation géométrique de la concavité. En déduire les applications qui sont à la fois convexes et concaves.

2 Propriétés de continuité et dérivabilité des fonctions convexes

2.1 Inégalité des pentes croissantes

Soit $f : I \to \mathbb{R}$. On veut montrer la caractérisation suivante de la convexité :

$$f \text{ convexe sur } I \Leftrightarrow \left(\forall (x,y,z) \in I^3, x < y < z \Rightarrow \frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x} \leq \frac{f(z)-f(y)}{z-y}\right)$$

On appelle la double inégalité du terme de droite *l'inégalité des pentes croissantes.*

Pour démontrer cette caractérisation, on va raisonner géométriquement. Pour $x < y < z$ dans I, on introduit les points suivants :
A de coordonnées $(x, f(x))$, B de coordonnées $(y, f(y))$, C de coordonnées $(z, f(z))$
et P le point d'abscisse y sur le segment $[AC]$.

Question 2.1.a (Représentation graphique d'un cas particulier)
Tracer l'allure de la courbe de $f = exp$ sur $[-1, 1]$, et placer les points A, B, C et P dans le cas $x = -1$, $y = 0$, $z = 1$. Interpréter géométriquement les trois valeurs dans l'inégalité des pentes croissantes.

On prouve maintenant la caractérisation dans le cas général.

Question 2.1.b Exprimer l'ordonnée de P (que l'on note y_P) en fonction de $f(x)$, $y - x$ et de la pente de la droite (AC). Exprimer y_P d'autre part en fonction de $f(z), y - z$ et de la pente de la droite (AC).

Question 2.1.c Déduire de la question précédente que si f est convexe sur I, alors l'inégalité des pentes croissantes est vérifiée.

Question 2.1.d Réciproquement, si on suppose que l'inégalité des pentes croissantes est vérifiée pour f, montrer que $f(y) \leq y_P$ et en déduire que f est convexe sur I.

2.2 Dérivabilité et continuité

Soit f une fonction convexe de I dans R. On veut montrer qu'en tout point intérieur à I, f est continue, dérivable à gauche et dérivable à droite.
Soit x_0 un point intérieur à I. On introduit la fonction $\phi : I \cap]-\infty, x_0[\to \mathbb{R}$ définie par :

$$\forall x \in I \cap]-\infty, x_0[,\ \phi(x) = \frac{f(x) - f(x_0)}{x - x_0}.$$

Question 2.2.a Soit x et y dans I tels que $x < y < x_0$. En utilisant l'inégalité des pentes croissantes, démontrer que $\phi(x) \leq \phi(y)$.

Question 2.2.b Soit λ dans I avec $\lambda > x_0$. Démontrer que ϕ est majorée par $\dfrac{f(\lambda) - f(x_0)}{\lambda - x_0}$.
(on pourra utiliser l'inégalité des pentes croissantes)

Question 2.2.c En déduire que f est dérivable à gauche en x_0 et est continue à gauche en x_0.

Question 2.2.d Comment pourrait-on procéder pour montrer que f est dérivable à droite en x_0 et est continue à droite en x_0 ?

Remarque à retenir : On a montré qu'une fonction convexe était dérivable à gauche et à droite (donc continue) en tout point intérieur à I. Cependant, elle peut très bien ne pas être dérivable en un point (cf. l'exemple g de la question 1.1), et peut ne pas être continue aux extrémités de l'intervalle (cf l'exemple e de la question 1.1).
Concernant les dérivées à gauche et à droite, on pourrait montrer de plus que $f'_g(x_0) \leq f'_d(x_0)$ pour x_0 intérieur à I (et même que l'ensemble des points de non-dérivabilité est au plus dénombrable).

2.3 Caractérisation des fonctions convexes dérivables

Dans toute cette partie 2.3, on suppose que f est une fonction dérivable de I dans $\mathbb{R}$.

Question 2.3.a Supposons f convexe. Soit $x < z$ fixés dans I, et y dans $]x, z[$.
En utilisant l'inégalité des pentes croissantes, démontrer que $f'(x) \leq \dfrac{f(z) - f(x)}{z - x} \leq f'(z)$.
(on pourra faire tendre y vers des valeurs bien choisies pour prouver chacune des inégalités)

Question 2.3.b En déduire que si f est convexe, alors f' est croissante.

On s'intéresse maintenant à l'implication réciproque. On suppose donc f' croissante dans les questions à venir de la partie 2.3, et on souhaite démontrer que f est convexe.
Soit $x < y$ fixés dans I, et $\phi : [0, 1] \to \mathbb{R}$ définie par

$$\forall t \in [0, 1],\ \phi(t) = tf(x) + (1 - t)f(y) - f(tx + (1 - t)y).$$

Question 2.3.c Démontrer que ϕ est dérivable et calculer sa dérivée. Calculer $\phi(0)$ et $\phi(1)$.

Question 2.3.d Ecrire l'égalité des accroissements finis pour f entre x et y. En déduire qu'il existe $\lambda \in]0, 1[$ tel que pour tout t de $[0, 1]$,

$$\phi'(t) = (x - y)\Big(f'(\lambda x + (1 - \lambda)y) - f'(tx + (1 - t)y)\Big).$$

Question 2.3.e Résoudre l'inégalité $\lambda x + (1 - \lambda)y \geq tx + (1 - t)y$ d'inconnue t et en déduire le signe de ϕ' sur $[0, \lambda]$ et sur $[\lambda, 1]$.

Question 2.3.f Etablir le tableau de variation de ϕ (on ne précisera pas $f(\lambda)$).

Question 2.3.g Conclure.

Remarque à retenir : Cette partie 2.3 prouve donc que pour une fonction f dérivable, il y a équivalence entre la convexité de f et la croissance de f'. On en déduit aisément que pour f deux fois dérivable, il y a équivalence entre la convexité de f et la positivité de f''.
De façon similaire, on peut montrer qu'une fonction dérivable est convexe si et seulement si son graphe est au-dessus de toutes ses tangentes.

3 Des utilisations de la convexité

Soit f une application de I dans $\mathbb{R}$.
Soit $n \in \mathbb{N}^*, (x_1, \ldots, x_n) \in I^n$ et $(\lambda_1, \ldots, \lambda_n) \in [0,1]^n$ vérifiant $\lambda_1 + \ldots + \lambda_n = 1$.

On admet pour les parties 3.1, 3.2 et 3.3 que l'inégalité suivante est vérifiée si f est convexe.

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i) \tag{**}$$

On prouvera dans la partie 3.4 que l'inégalité $(*)$ implique bien l'inégalité $(**)$.

3.1 Moyennes arithmétique et géométrique

Question 3.1.a Grâce aux résultats de la partie 2.3, démontrer que la fonction logarithme népérien est concave.

Question 3.1.b On considère maintenant $n \in \mathbb{N}^*, (x_1, \ldots, x_n) \in (\mathbb{R}_*^+)^n$. Démontrer que

$$\ln\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \geq \ln\left(\left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}\right).$$

Question 3.1.c En déduire que la moyenne arithmétique des nombres $x_1, \ldots, x_n$ est supérieure à leur moyenne géométrique.

3.2 Entropie d'une probabilité

On considère l'univers $\Omega = \{1, ..., n\}$ et $\mathbb{P}$ une probabilité sur Ω. On rappelle que $\mathbb{P}$ est caractérisée par les réels $p_k = \mathbb{P}(\{k\})$ pour $1 \leq k \leq n$. L'entropie de la probabilité $\mathbb{P}$ est définie par

$$H(\mathbb{P}) = -\sum_{k=1}^{n} p_k \ln(p_k).$$

Avec pour convention $0 \ln(0) = 0$.

Question 3.2.a Montrer que H est à valeurs dans $\mathbb{R}_+$ et trouver une $\mathbb{P}$ telle que $H(\mathbb{P}) = 0$.

Question 3.2.b Calculer $H(\mathbb{P})$ lorsque $\mathbb{P}$ est la loi uniforme sur $\{1, ..., n\}$.

Question 3.2.c Soit $(q_k, k \in [|1,n|])$ et $(p_k, k \in [|1,n|])$ deux probabilités sur $\{1, \ldots, n\}$, telles qu'aucun p_k ou q_k ne s'annule. Démontrer à l'aide de l'inégalité $(**)$ appliquée à la fonction $-\ln$, aux points $x_k = p_k/q_k$ et à $\lambda_k = q_k$ que

$$\sum_{k=1}^{n} q_k \ln(p_k) \leq \sum_{k=1}^{n} q_k \ln(q_k).$$

En choisissant $p_k = 1/n$ pour tout $k \in \{1, \ldots, n\}$, en déduire que la probabilité uniforme réalise le maximum de H. [1]

[1]De façon heuristique, l'entropie mesure le manque d'information d'une loi : on voit ainsi que la loi uniforme, qui ne favorise aucun des possibles, est la loi qui apporte le moins d'information. Inversement, la loi de probabilité où un des possibles a probabilité 1 de se réaliser réalise le minimum de l'entropie sur $\Omega = \{1, ..., n\}$.

3.3 Une autre application

Soient a, b, c trois réels strictement positifs. Démontrer que

$$\frac{a}{b+c} + \frac{b}{a+c} + \frac{c}{a+b} \geq \frac{3}{2}.$$

Indication : On pourra commencer par supposer que $a + b + c = 1$.

3.4 Inégalité généralisée de convexité

On veut ici démontrer l'inégalité $(**)$ (souvent appelée inégalité de Jensen) pour une fonction convexe f. On procède par récurrence sur n.
Question 3.3.a Démontrer que $(**)$ est vérifiée pour $n = 1$ et pour $n = 2$.

Pour toute la suite, soit n fixé tel que $(**)$ soit vraie. On veut démontrer que $(**)$ est vraie au rang suivant.
Fixons donc $(x_1, \ldots, x_{n+1}) \in I^{n+1}$ et $(\lambda_1, \ldots, \lambda_{n+1}) \in [0,1]^{n+1}$ vérifiant $\lambda_1 + \ldots + \lambda_{n+1} = 1$.

Question 3.3.b Montrer que si $\lambda_{n+1} = 1$, alors $(**)$ est vérifiée.

Question 3.3.c On s'intéresse maintenant au cas $\lambda_{n+1} < 1$.
On a donc $\sum_{i=1}^{n} \lambda_i = 1 - \lambda_{n+1} > 0$.
Posons $y = \dfrac{1}{\sum_{i=1}^{n} \lambda_i} \displaystyle\sum_{i=1}^{n} \lambda_i x_i$. On admet que y est bien un point de I.
Montrer $f\left(\displaystyle\sum_{i=1}^{n+1} \lambda_i x_i\right) \leq (1 - \lambda_{n+1}) f(y) + \lambda_{n+1} f(x_{n+1})$.
D'autre part, montrer via l'hypothèse de récurrence que $f(y) \leq \dfrac{1}{1 - \lambda_{n+1}} \displaystyle\sum_{i=1}^{n} \lambda_i f(x_i)$.
Combiner les deux inégalités précédentes pour obtenir l'inégalité $(**)$ au rang $n + 1$.