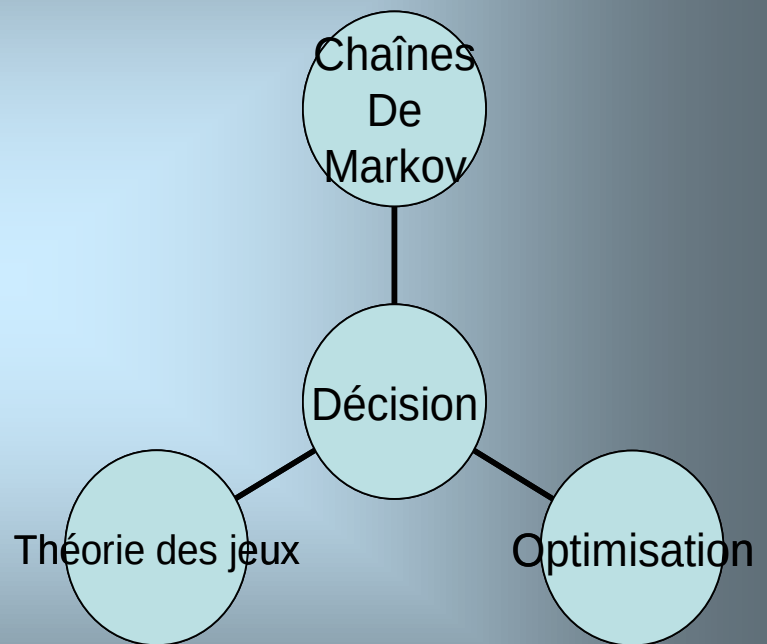


Université de Nice-Sophia Antipolis

Département de Mathématiques

Analyse de la décision



semaines de 1 à 5 :

Ch.1 : Les outils du décideur. L'importance de l'information.
Jeux à deux joueurs et à somme zéro.
Jeux ayant une valeur, les points selles, les stratégies optimales.
Les stratégies mixtes.
Stratégies dominantes, réduction d'une matrice de paiement.

Ch.2 : Les processus de Markov. Les matrices stochastiques, les matrices de transitions.
Marches aléatoires sur des graphes. Etats absorbants. Comportement à long terme.
Applications.

Semaine 6 ou 7 : test n°1

semaines de 7 à 13 :

Ch.3 : Loteries. Utilité espérée. Révision des programmes d'optimisation sans contraintes.
Programmes d'optimisation avec contraintes du type $g(x)=0$ et $h(x)\geq 0$.
CNPO, CSSO, fonctions convexes, fonctions concaves, formes quadratiques définies positives, définies négatives, semi-définies.

Ch.4 : La théorie de la décision. L'interprétation subjectiviste.
Valeur monétaire espérée (VME), Perte d'opportunité espérée (POE), calcul avec les probabilités corrigées, valeur de l'information parfaite (VIPE).

****** Vous avez la possibilité de poser vos questions relatives au cours du lundi par mail (zone@unice.fr) jusqu'au jeudi 12h, **seules les questions bien posée auront un droit de réponse !**
Les réponses seront soit adressées avant le lundi suivant, soit intégrées dans le poly si des commentaires se révèlent utiles.

******* Ni documents ni machines programmables autorisés pendant le test et l'examen final.
i.e. vous n'avez droit que à une calculette 4 opérations.

******* Durée du test : 2h. Durée de l'examen final : 3h.

N.B.)Des petits contrôles seront programmés en TD pour permettre une notation pondérée de contrôle continu, la présence est donc indispensable. Les salariés et dispensés de TD sont priés de communiquer leur nom au plus tard le lundi 24 février.

Ch1. Les jeux mathématiques.

Jeu à deux joueurs et à somme zéro.

§1.1 Jeux à deux joueurs et à somme zéro :

Si deux joueurs X et Y s'affrontent, dans un souci de simplification on notera X l'ensemble des stratégies possibles du joueur X et Y l'ensemble des stratégies possibles du joueur Y :

Déf.1 : On appellera alors jeu à deux joueurs et à somme zéro le triplet (X, Y, u) avec :

$X = \{x_i \in X, \text{ avec } x_i \text{ stratégie possible du joueur X et } i = 1, 2, \dots, n\}$; si le joueur X dispose de n stratégies possibles}

$Y = \{y_i \in Y, \text{ avec } y_i \text{ stratégie possible du joueur Y et } i = 1, 2, \dots, m\}$; si le joueur Y dispose de m stratégies possibles}

$U : X \times Y \rightarrow \mathbb{R}$ une fonction continue, qui à toute paire de stratégies $(x, y) \in X \times Y$ associe le paiement $r \in \mathbb{R}$
 $(x, y) \rightarrow u(x, y) = r \in \mathbb{R}$

Par convention on assumera ici que r est la somme que Y doit payer à X.

(Il va de soi que on aurait pu assumer que r est la somme que X doit payer à Y, l'important étant de choisir des le départ et de s'y tenir !!)

§1.2 Représentation d'un jeu sous forme normale

Si on représente les n stratégies du joueur X en ligne et les m stratégies du joueur Y en colonne, la simple lecture de la matrice $n \times m$ ayant comme entrées les valeurs de la fonction de paiement nous permettra une représentation très commode du jeu.

Exemple (1) :

	y_1	y_2	y_3
X_1	-2	+1	-1
X_2	-2	-4	+1
X_3	-3	+2	-4

On lira donc : « si X choisit la stratégie x_1 et Y choisit la stratégie y_1 , alors le joueur Y doit payer -2 au joueur X, i.e. X doit payer 2 à Y »

De même on lira : « si X choisit x_3 et Y choisit y_2 alors Y doit payer 2 à X, si X choisit x_3 et Y choisit y_3 alors X payera 4 à Y, etc »

Avec la convention relative au paiement il est évident que X doit choisir une stratégie qui maximise $u(x, y)$, alors que Y doit raisonnablement en choisir une qui le minimise.

Les deux joueurs sont supposés être également intelligents, rationnels et également informés.

Chacun est donc parfaitement capable d'effectuer le raisonnement de l'autre et d'agir de conséquence.

Si on reprend l'exemple (1) et on regarde la matrice des paiements successivement du point de vue de Y et de X, on peut donc observer que :

Point de vue de Y « si X choisit x_1 alors je dois choisir y_1 (paiement minimum vu le choix de X), car ainsi je minimiserai ma perte (je gagnerai 2), si X choisit x_2 je choisirai et si X choisit x_3 je choisirai y_3 ».

Point de vue de X « si Y choisit y_1 , je choisirai x_1 ou x_2 (paiement maximum vu le choix de Y), si Y choisit y_2 je choisirai x_3 »

Nous allons noter les choix de X et de Y respectivement en ligne et en colonne :

	y_1	y_2	y_3	
X_1	-2	+1	-1	2
X_2	-2	-4	+1	4
X_3	-3	+2	-4	4
	-2	+2	+1	

Maxmin = -2, Max (-2, -4, -4) = choix de X
minMax = -2, min (-2, 2, 1) = choix de Y

Parmi les paiements minimaux proposés par Y il va de soi que X choisira le plus grand et parmi les paiements maximaux proposés par X le joueur Y ne pourra que choisir le plus petit ! C'est la méthode du minMax et Maxmin. Si $\text{Maxmin} = \text{minMax}$ on dit que le jeu admet une valeur en stratégie pure et on notera $\text{val}(u) = \text{Maxmin} = \text{minMax}$. Ici $\text{val}(u) = -2$

Exemple (2) :

	y ₁	y ₂	y ₃	
X ₁	-1	+1	-1	-1
X ₂	-2	-4	+1	-4
X ₃	-3	+2	-4	-4
	-1	+2	+1	

$\text{Maxmin} = -1$, $\text{Max}(-1, -4, -4) = \text{choix de X}$
 $\text{Minmax} = -1$, $\text{min}(-2, 2, 1) = \text{choix de Y}$

$\text{Maxmin} = \text{minmax}$, donc ce jeu admet une valeur en stratégie pure.

Exercice (1) : Soit le jeu à deux joueurs et à somme zéro représenté par la matrice de paiement suivante : Déterminer, si elle existe, la valeur de ce jeu en stratégie pure.

		Stratégies de Y			
Stratégies De X		<u>-1</u>	+1	+1	<u>-1</u>
	-2	-3	+2	+2	-3
	-2	+1	-2	+1	-2
	<u>-1</u>	-3	+4	+1	-3
	-1	+1	+4	+2	

$-1 = \text{Maxmin}$
 $-1 = \text{minMax}$

Solution :

$\text{Maxmin} = -1 = \text{minmax}$, donc $\text{val}u = -1$

Il semble légitime à ce stade de se poser la question « Les points $P(x_1, y_1)$, $Q(x_1, y_4)$, $T(x_4, y_1)$ ayant tous pour image dans $\mathbf{R}$ la valeur $-1 = \text{val}(u)$, correspondent-ils à des paires de stratégies optimales pour les joueurs ? »

A savoir : si le joueur X choisit la stratégie x_1 et le joueur Y choisit y_1 , ou si X choisit x_1 et Y choisit y_4 , ou si X choisit x_4 et Y choisit y_1 , les deux joueurs sont-ils confrontés à des choix équivalents ?

Pour répondre à cette question il est indispensable de connaître et bien comprendre le paragraphe suivant :

§1.3 Les points selles

Un point $S(x^*, y^*)$ est un point selle pour un jeu à deux joueurs et à somme zéro ssi :

(déf.2) $u(x, y^*) \leq u(x^*, y^*) \leq u(x^*, y)$

A savoir "si X déplace son choix de stratégie de manière unilatérale de x^* à une autre stratégie $x \neq x^*$, alors que Y conserve son choix y^* ,

le paiement diminue (donc X s'auto-pénalise, car $u(x, y^*) \leq u(x^*, y^*)$) et de même si Y modifie

son choix de y^* à une autre stratégie $y \neq y^*$ alors que X conserve le choix x^* le paiement augmente (donc Y sera aussi pénalisé, car $u(x^*, y) \leq u(x^*, y^*)$)

Le point $S(x^*, y^*)$ constitue ainsi le meilleur compromis possible entre les deux joueurs : on parle alors de choix prudent.

Il ne faut jamais oublier que chaque joueur est parfaitement capable et dispose de toute l'information nécessaire pour simuler avec exactitude le raisonnement de l'autre !

On peut désormais utiliser la (déf.2) pour répondre à la question posée à la fin de l'exercice (1) :

Si $P(x_1, y_1)$ est un point selle alors le choix de stratégies (x_1, y_1) doit satisfaire: $u(x, y_1) \leq u(x_1, y_1) \leq u(x_1, y)$

Or $u(x, y_1) = -1$ ou $-2 \leq u(x_1, y_1) = -1 \leq u(x_1, y) = -1$ ou $+1$ donc P est un point selle.

Si $Q(x_1, y_4)$ est un point selle alors le choix de stratégies (x_1, y_4) doit satisfaire: $u(x, y_4) \leq u(x_1, y_4) \leq u(x_1, y)$ mais $u(x, y_4) = -1$ ou $+1$ ou $+2$ n'est pas toujours $\leq u(x_1, y_4) = -1$, la double inégalité de la (déf.2) n'est donc pas vérifiée : Q ne peut pas être un point selle pour le jeu donné.

Vous pouvez maintenant facilement démontrer que $T(x_4, y_1)$ ne sera pas un point selle !

Le jeu admet donc un seul point selle : le point $P(x_1, y_1)$.

X a donc une seule stratégie optimale : la stratégie x_1 ; Y a aussi une seule stratégie optimale : la stratégie y_1 .

Théorème 1 : Si l'ensemble $\mathbf{S}$ de tous les points selles d'un jeu est non vide ($\mathbf{S} \neq \emptyset$) alors le paiement est le même dans tous les points selles et ces derniers sont échangeables,

à savoir $\forall (x_1, x_2), (y_1, y_2) \in \mathbf{S} \times \mathbf{S} \Rightarrow (x_1, y_2) \in \mathbf{S}$ et $(x_2, y_1) \in \mathbf{S}$.

Remarque : Le paiement commun à tous les points selles est la valeur du jeu, on note $\text{val}(u) : u(\mathbf{S}) = \{\text{val}(u)\}$

Démonstration :

(1) Si $(x_1, y_1) \in \mathbf{S}$ alors la Déf.2 $\Rightarrow$ (a) $u(x, y_1) \leq u(x_1, y_1) \leq u(x_1, y), \forall x \neq x_1 \forall y \neq y_1$

si $(x_2, y_2) \in \mathbf{S}$ alors la Déf.2 $\Rightarrow$ (b) $u(x, y_2) \leq u(x_2, y_2) \leq u(x_2, y), \forall x \neq x_2 \forall y \neq y_2$

donc en particulier

si dans (a) on choisit $x=x_2$ et $y=y_2$ on a : $u(x_2, y_1) \leq u(x_1, y_1) \leq u(x_1, y_2)$

si dans (b) on choisit $x=x_1$ et $y=y_1$ on a : $u(x_1, y_2) \leq u(x_2, y_2) \leq u(x_2, y_1)$

on peut alors observer le troisième terme de la première inégalité est égal au premier terme de la deuxième inégalité et ceci $\Rightarrow$ l'égalité des six paiements, avec en particulier $u(x_1, y_1) = u(x_2, y_2) = \text{val}(u)$.

Nous venons donc de démontrer que le paiement est bien le même pour les différents points selles du jeu, car le raisonnement se généralise très facilement pour un nombre de points selles supérieur à deux.

(2) Nous allons maintenant démontrer l'échangeabilité des points selles :

Pour ce faire on observe d'abord que la Déf.2 peut s'écrire autrement car:

$$u(x, y^*) \leq u(x^*, y^*) \leq u(x^*, y) \quad \text{équivalent à} \quad u(x^*, y^*) = \sup_{x \in X} u(x, y^*) \quad \text{et} \quad u(x^*, y^*) = \inf_{y \in Y} u(x^*, y)$$

Au passage il est intéressant d'observer que un point (x^*, y^*) sera donc un point selle pour un jeu à deux joueurs

et à somme zéro ssi Déf.2' : $u(x^*, y^*) = \sup_{x \in X} u(x, y^*) \quad \text{et} \quad u(x^*, y^*) = \inf_{y \in Y} u(x^*, y)$

Ici donc $u(x_1, y_2) = u(x_2, y_2) = \sup_{x \in X} u(x, y_2)$ et $u(x_1, y_2) = u(x_1, y_1) = \inf_{y \in Y} u(x_1, y)$

Ce qui permet d'affirmer que (x_1, y_2) est aussi un point selle.

De même $u(x_2, y_1) = u(x_1, y_1) = \dots$, ce qui prouve que (x_2, y_1) est aussi un point selle.

Les points selles sont ainsi échangeables au sens précisé dans l'énoncé du Théorème 1 Q.E.D.

§1.4 Les stratégies optimales et l'existence de points selles pour un jeu.

A ce stade nous nous sommes limité à affirmer qu'une stratégie était optimale pour un joueur si elle était une coordonnée d'un point selle (abscisse pour X , ordonnée pour Y).

Le moment est venu d'être plus précis.

Etant donné que tous les points selles d'un jeu ont même paiement et sont échangeables il existe un sous-ensemble de X et un sous-ensemble de Y , que nous noterons $\mathbf{O}_X$ (et que nous appellerons l'ensemble de stratégies optimales du joueur X) et $\mathbf{O}_Y$ (que nous appellerons l'ensemble de stratégies optimales du joueur Y) t.q. $\mathbf{O}_X \times \mathbf{O}_Y = \mathbf{S}$

Si X choisit une stratégie de $\mathbf{O}_X$, alors contre toute défense de Y il sera certain d'obtenir un paiement au moins égal à la valeur du jeu.

Si Y choisit une stratégie de $\mathbf{O}_Y$, alors quelque soit le choix de X il sera certain de ne devoir payer plus que la valeur du jeu.

X tente donc d'imposer un paiement compris entre $\text{val}(u)$ et $+\infty$, alors que Y fera tout pour obtenir un paiement compris entre $-\infty$ et $\text{val}(u)$.

Si un des deux joueurs choisit une stratégie ailleurs que dans son propre ensemble de stratégies optimales, son adversaire peut exploiter la faute et gagner plus ! (ou perdre moins !)

Des joueurs rationnels et "prudents" choisiront toujours l'issue du jeu qui leur attribuera la valeur du jeu !!

On ne peut donc parler de stratégies optimales (pures) que pour des jeux qui ont des points selles.

Le théorème suivant nous permet de caractériser l'existence de points selles (admis sans démonstration).

Rappel (1) :

Un ensemble A est compact si il est à la fois fermé et borné. (cf. rappels de mathématiques , ci joint)

Rappel (2) : Pour toute fonction $f : X \times Y \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ il est toujours vrai que :

$$\sup_{x \in X} \inf_{y \in Y} f(x,y) \leq \inf_{y \in Y} \sup_{x \in X} f(x,y)$$

Théorème 2 : Si les ensembles X et Y sont compacts,

$$\begin{aligned} \sup_{x \in X} \inf_{y \in Y} u(x,y) &= \inf_{y \in Y} \sup_{x \in X} u(x,y) \Rightarrow \mathbf{S} \neq \emptyset \text{ et la val}(u) = \sup_{x \in X} \\ \inf_{y \in Y} &= \inf_{y \in Y} \sup_{x \in X} \end{aligned}$$

et réciproquement

$$\text{Si } \mathbf{S} \neq \emptyset \Rightarrow \sup_{x \in X} \inf_{y \in Y} u(x,y) = \inf_{y \in Y} \sup_{x \in X} u(x,y) =$$

val(u)

**Pour la démonstration voir H.Moulin, p.15 « Théorie des jeux » Ed.Hermann

Théorème 3 : Si $\mathbf{S} \neq \emptyset$ une stratégie x^* sera optimale pour le joueur X ssi $\inf_{y \in Y} u(x^*,y) = \text{val}(u)$

et une stratégie y^* sera optimale pour Y ssi $\sup_{x \in X} u(x,y^*) = \text{val}(u)$

Nous avons vu que le choix d'une stratégie optimale garantit au décideur un gain au moins égal à la valeur du jeu. Une telle stratégie étant abscisse d'un point selle pour X et ordonnée pour Y, il est clair que un point selle constitue une issue stable du jeu, car aucun joueur n'a intérêt à changer d'avis.

Représentation dans $\mathbb{R}^3$

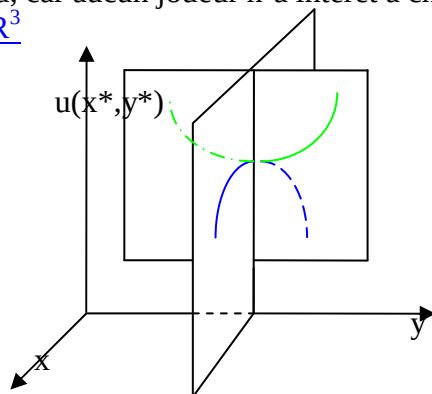


fig-1-

Réfléchir, puis écrire une définition pertinente de stratégie non optimale.

Si certaines hypothèses ne sont pas vérifiées le jeu tout en admettant une valeur

sans avoir de points selles ! On aura alors toujours $\sup_{x \in X} \inf_{y \in Y} u(x,y) = \inf_{y \in Y} \sup_{x \in X} u(x,y)$

mais vous pourrez vérifier que $\inf_{y \in Y} u(x,y)$ ne peut atteindre son maximum sur X ou

Pour éviter toute « confusion » l'écriture $\text{Maxmin}=\text{minMax}$, utilisée dans les exemples et les exercices traités dans les pages précédentes, sera maintenant systématiquement remplacée par $\text{supinf}=\text{infsup}$. Je suppose que les raisons de cette nécessité sont maintenant claires pour tous ! Si non à vos e-mails !

Exercices :



******au travail !!!**

vos propositions de démonstration complète avant vendredi 12h !

1) (H.Moulin, p21)

Soient $X= Y [0,1]$ les ensembles des stratégies possibles des joueurs X et Y, et soit $u(x,y)= 1 -(x-y)^2$ la fonction de paiement d'un jeu à deux joueurs et à somme zéro.

Démontrer que ce jeu n'admet pas de points selles.

Solution :

$$\text{supinf } u = \sup \inf = \frac{3}{4}$$

$$\text{infsup } u = \inf \sup = 1$$

donc pas de valeur en stratégie pure.

Le jeu du menteur



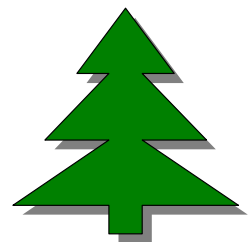
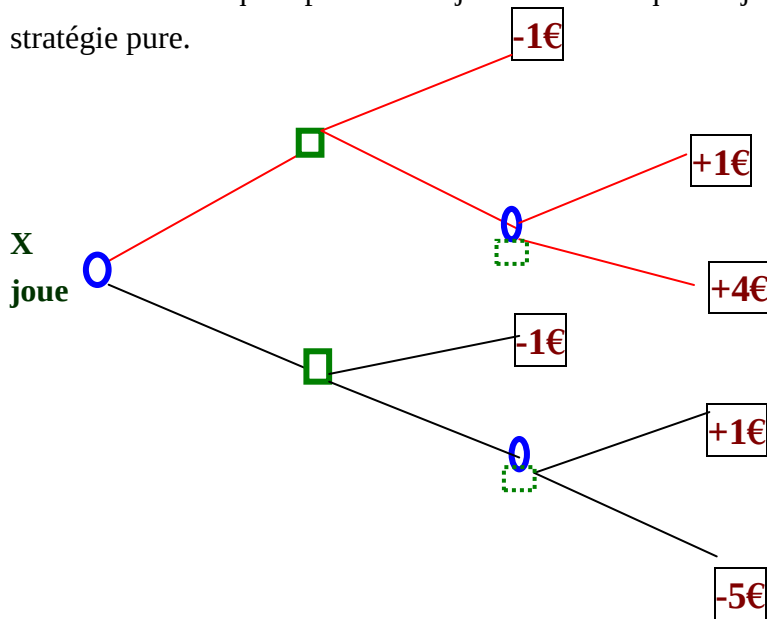
Mr.X tire une carte dans un jeu et ne la montre pas.

Il peut choisir de donner un euro ou de demander un euro à son adversaire Mr.Y.

Ce dernier peut soit accepter de donner l'euro demandé soit de dire « menteur ».

A ce stade du jeu X doit montrer sa carte : si elle est rouge il reçoit 4 euros de Y, si elle est noire il doit payer 5 euros à Y.

Ecrire la matrice qui représente ce jeu et montrer que les joueurs n'ont pas de stratégies optimales en stratégie pure.



Une observation attentive de l'arbre de décision est suffisante pour réaliser que toutes les informations nécessaires sont contenues dans le graphe.

1) Le joueur Y ne dispose que de deux stratégies : **(D)** donner 1€ ou **(M)** dire « menteur ! ».

Le joueur X dispose de quatre stratégies : il peut donner 1€ si il a tiré une carte rouge et donner 1€ si il a tiré une carte noire i.e. $x_1 = (\text{Donner, Donner})$

il peut demander 1€ si il a tiré une carte rouge donner 1€ si il a tiré une carte noire i.e. $x_2 = (\text{Demander, Donner})$

il peut donner 1€ si il a tiré une carte rouge et demander 1€ si il a tiré une carte noire i.e. $x_3 = (\text{Donner, Demander})$

il peut demander 1€ si il a tiré une carte rouge et demander 1€ si il a tiré une carte noire i.e. $x_4 = (\text{Demander, Demander})$

La matrice aura donc 4 lignes et 2 colonnes.

Les paiements (entrées de la matrice) se calculent simplement en effectuant le « bon parcours » dans l'arbre

si X choisit x_1 , alors le paiement espéré de X est : $\frac{1}{2}(1) + \frac{1}{2}(1) = \boxed{+1}$ si Y choisit la stratégie **(D)**

et $\frac{1}{2}(-5) + \frac{1}{2}(+4) = \boxed{-\frac{1}{2}}$ si Y choisit la stratégie **(M)**

et ainsi de suite : si X choisit x_3 , alors le paiement espéré de X est : $\frac{1}{2}(-1) + \frac{1}{2}(1) = \boxed{0}$ si Y choisit la stratégie (D) et $\frac{1}{2}(-1) + \frac{1}{2}(-5) = \boxed{-3}$ si Y choisit la stratégie (M)

si X choisit x_1 alors le paiement espéré de X est : $\frac{1}{2}(-1) + \frac{1}{2}(-1) = \boxed{-1}$ si Y choisit la stratégie (D) et $\frac{1}{2}(-1) + \frac{1}{2}(-1) = \boxed{-1}$ si Y choisit la st

et si X choisit x_2

Il est maintenant possible de fournir sans hésitation la matrice qui représente le jeu du menteur

-1	-1
0	$+\frac{3}{2}$
0	$-\frac{3}{2}$
+1	$-\frac{1}{2}$

avec les règles énoncées :

Si on observe de plus près on voit que la stratégie x_1 est dominée par x_2 et x_3 est aussi dominée

par x_2 , donc la matrice se réduit à :

0	$+\frac{3}{2}$
+1	$-\frac{1}{2}$

Après élimination des stratégies dominées X choisira donc ou x_2 ou x_4 , mais aucun des deux joueurs a des stratégies optimales en stratégie pure car :

$$\sup_x \inf_y = -\frac{1}{2} \neq \inf_y \sup_x = \frac{3}{2}$$

§1.5 Les stratégies mixtes.

Pour les jeux n'ayant pas de valeur en stratégie pure nous pouvons élargir l'ensemble des stratégies de chaque joueur en introduisant la notion de hazard.

On appelle alors stratégie mixte du joueur X une variable aléatoire qui peut prendre les valeurs x_i avec une probabilité p_i , X doit donc choisir une loi de probabilité sur l'ensemble des valeurs possibles de la variable aléatoire.

Le joueur Y doit alors utiliser aussi une stratégie mixte et les variables aléatoires choisies par X et par Y sont indépendantes.

(c'est comme si chaque joueur tirait au sort une stratégie sans connaître la stratégie tirée au sort par l'autre !)

Dans l'exemple du jeu du menteur X choisira x_2 avec une probabilité p et donc x_1 avec probabilité $(1-p)$,

alors que Y choisira (D) ou (M) avec probabilités respectives q et $(1-q)$

On a alors comme paiement minimal garanti pour X :

$$\sup_{0 \leq p \leq 1} \inf \{ p \times 0 + (1-p) \times (+1), p \times (+\frac{3}{2}) + (1-p) \times (-\frac{1}{2}) \} = \sup_{0 \leq p \leq 1} \inf \{ 1-p, 2p - \frac{1}{2} \} = \frac{1}{2}$$

$$\text{car } 1-p = 2p - \frac{1}{2} \text{ ssi } p = \frac{1}{2}$$

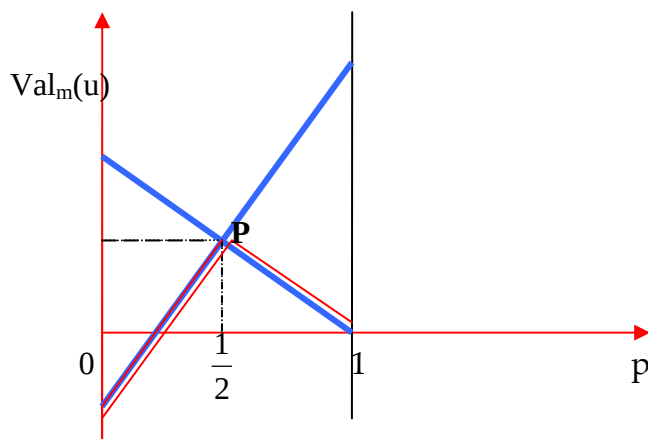
De même la perte minimale garantie pour Y sera :.....

à vous de jouer !!!..... = ?????

La valeur mixte de ce jeu est donc $\frac{1}{2}$ et les stratégies mixtes pour X sont $(\frac{1}{2}, \frac{1}{2})$

et pour Y sont $(\frac{2}{3}, \frac{1}{3})$.

Pour calculer la valeur de la stratégie mixte on peut utiliser une méthode graphique très suggestive :



(X choisira de parcourir le sup des inf autorisés par Y, **chemin rouge**)

Visiblement le point P est le $\sup_{X} \inf_{Y}$! donc il suffit de déterminer l'intersection entre les segments de droite

qui représentent les paiements pour X, on a : $1-p=2p-\frac{1}{2}$ ssi $p^*=\frac{1}{2}$ = abscisse du point P

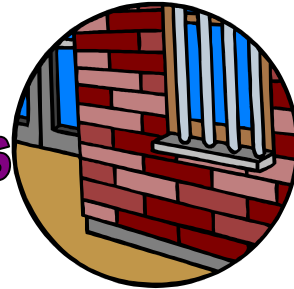
et la valeur mixte du jeu sera $\text{val}_m(u)=\frac{1}{2}$ = ordonnée de P.

Faire le calcul pour Y et vérifier que $\text{val}_m(u)=\frac{1}{2}$

*** Pour réfléchir sur les jeux en général :

Comment devraient être les entrées de la matrice qui fournit la forme normale du « baccara des bagnards » pour que un jouer puisse au moins s'assurer match nul ?

Le "baccara"...des bagnards



			
	0	−	+
	+	0	−
	−	+	0

Vous êtes M. X
à vous de jouer !

avez-vous des stratégies optimales ?

Ch2.Les processus de Markov

§1.1 Avant d'introduire la notion de chaîne de Markov nous allons rappeler quelques définitions indispensables.

- (a) Si v est un vecteur à n composantes et M est une matrice carrée à n lignes, on dit que le vecteur v est un **vecteur fixe à gauche** de M si $vM=v$
- Il est évident que si v est un vecteur fixe de M alors kv le sera aussi, $\forall k \neq 0$
- Une définition analogue est à fournir pour un vecteur fixe à droite de M
- (b) Un vecteur $s(s_1, s_2, \dots, s_n)$ est appelé vecteur probabilité si $s_i \geq 0$ et $\sum s_i = 1$
- (c) Une matrice $P(p_{ij})$ est appelée stochastique si chacune de ses lignes est un vecteur probabilité.
- (d) Une matrice stochastique P est appelée **régulière** si $\exists m$ t.q. P^m a des entrées toutes positives.

Remarques (1) : Une matrice stochastique $m \times m$, avec $m \geq 2$, dont toutes les entrées situées sur la diagonale principale sont des 1 ne peut être régulière. **Pourquoi ?**

(2) : Une matrice stochastique $m \times m$, avec $m \geq 2$, dont une entrée située sur la diagonale principale est 1 ne peut être régulière. **Pourquoi ?**

EXEMPLES :

(1) Soit $M = \begin{pmatrix} 0 & 1 \\ 1 & 2 \\ 3 & 3 \end{pmatrix}$, cherchons le vecteur probabilité fixe de cette matrice : on pose $(x, 1-x)$

x) $\begin{pmatrix} 0 & 1 \\ 1 & 2 \\ 3 & 3 \end{pmatrix} = (x, 1-x)$, on effectue donc le produit (ligne par colonne !) et on a :

$$\left(\frac{1}{3}(1-x), x + \frac{2}{3}(1-x)\right) \Rightarrow (x, 1-x) \Rightarrow \frac{1}{3} - \frac{1}{3}x = x \quad \text{et} \quad \frac{1}{3}x + \frac{2}{3} = 1-x \Rightarrow x = \frac{1}{4} \quad \text{et} \quad 1-x = \frac{3}{4} \quad \text{i.e. le vecteur}$$

$v\left(\frac{1}{4}, \frac{3}{4}\right)$ est un vecteur fixe à gauche de la matrice M .

Vérification :

$$\left(\frac{1}{4}, \frac{3}{4}\right) \begin{pmatrix} 0 & 1 \\ 1 & 2 \\ 3 & 3 \end{pmatrix} = \left(\frac{1}{4} \cdot 0 + \frac{3}{4} \cdot \frac{1}{3}, \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot \frac{2}{3}\right) = \left(\frac{1}{4}, \frac{3}{4}\right) \text{ q.e.d.}$$

(2) Soit $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$, cette matrice est-elle régulière ? Si on calcule P^5 on a $P^5 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \times$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{8} & \frac{3}{8} & \frac{1}{2} \end{bmatrix}$$

N.B. Vérifier le calcul de P^5 en utilisant par exemple Excel (fonction « produitmat »), pour mettre à profit ce que vous avez appris en DEUG1 !!

Il existe donc une puissance t.q. aucune entrée de la matrice est ≤ 0 , donc la matrice est bien régulière.

(3) Soit $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ prouver que cette matrice stochastique ne peut être régulière.

I étant la matrice unité, $\forall M_{2 \times 2}$ on a $MI=IM=M$ et en particulier pour $M=I$ on a $I=I \times I = I^2 = I^2 \times I = \dots = I^n, \forall n$ i.e. elle ne peut être régulière.

(4) Soit $M = \begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \\ 0.25 & 0.5 & 0.25 \end{bmatrix}$ pouvez vous affirmer que cette matrice est régulière ?

Il est facile de démontrer qu'une telle affirmation sera fausse, en effet la définition de produit de matrices nous dit que :

$$\begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \times \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \text{ avec } = \sum_{h=1}^3 \text{ donc ici } = a_{21}$$

$$a_{11} + a_{22}a_{21} + a_{23}a_{31} = 0.(0.5) + 1.(0) + 0.(0.25) = 0, \quad = 1 \text{ et } = 0$$

$$\begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \\ 0.25 & 0.5 & 0.25 \end{bmatrix} \times \begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \\ 0.25 & 0.5 & 0.25 \end{bmatrix} = \begin{bmatrix} " & " & " \\ 0 & 1 & 0 \\ " & " & " \end{bmatrix} \text{ et } \begin{bmatrix} " & " & " \\ 0 & 1 & 0 \\ " & " & " \end{bmatrix} \times \begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \\ 0.25 & 0.5 & 0.25 \end{bmatrix} = \begin{bmatrix} " & " & " \\ 0 & 1 & 0 \\ " & " & " \end{bmatrix} \dots$$

*****Utilisez cet exemple pour démontrer la remarque (2). Quelle méthode de démonstration allez vous choisir ?

Théorème 1 : Si P est une matrice stochastique régulière alors elle admet un vecteur fixe unique v à composantes toutes positives

et si p est un vecteur probabilité alors la suite $pP, pP^2, pP^3, \dots$ tend vers le vecteur fixe v ,

alors que la suite $P, P^2, \dots, P^n, \dots$ tend vers une matrice V dont les lignes sont toutes identiques et égales à v .

(admis sans démonstration : ce qui ne vous empêche pas d'en proposer !!)

EXEMPLE:

$$P = \begin{pmatrix} 1 & 3 \\ 4 & 4 \\ 1 & 0 \end{pmatrix}$$

Après avoir vérifié que P est bien stochastique et régulière, le vecteur probabilité

fixe $v(x, 1-x)$ de P se détermine en posant : $(x, 1-x) \begin{pmatrix} 1 & 3 \\ 4 & 4 \\ 1 & 0 \end{pmatrix} = (x, 1-x)$

Nous avons donc $x = \frac{4}{7}$ d'où $v = (\frac{4}{7}, \frac{3}{7})$ et le théorème 1 nous permet d'affirmer que la suite $P,$

$P^2, P^3, \dots, P^n, \dots \rightarrow V$ avec $V = \begin{pmatrix} \frac{4}{7} & \frac{3}{7} \\ \frac{4}{7} & \frac{3}{7} \\ \frac{4}{7} & \frac{3}{7} \end{pmatrix}$

§1.2 Un processus stochastique est dit de Markov si dans ce processus toute suite d'essais a des issues $X_1, X_2, \dots$ t.q. toute issue X_i appartient à un ensemble fini $\{e_1, e_2, \dots, e_n\}$, appelé l'ensemble des états du système, et n'est dépendante que de l'issue précédente.

On dira alors que le système se trouve dans l'état e_i à l' n -ième étape si l'issue de l' n -ième essai est e_i .

On appellera probabilité de transition p_{ij} la probabilité pour que le système passe de l'état e_i à l'état e_j en une seule étape

et matrice de transition la matrice $P = (p_{ij}) =$

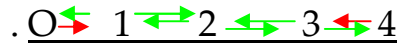
$$\begin{pmatrix} p_{11} & p_{12} & \dots & \dots & p_{1n} \\ p_{21} & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ p_{n1} & \dots & \dots & \dots & p_{nn} \end{pmatrix}$$

Il est évident que la matrice de transition d'un processus de Markov est une matrice stochastique.

EXEMPLE: (Une marche aléatoire)

Prenons un point mobile P situé dans le point d'abscisse 3 sur l'axe des x et supposons qu'il puisse se déplacer d'une marche unitaire à droite avec une probabilité p sauf quand il arrive dans l'origine O

Où il est obligé de bouger vers la droite ou dans le point d'abscisse 4 où il est obligé de bouger à gauche.



Si on note X_n la position du mobile après n étapes, on peut affirmer que l'ensemble des états est $\{e_0, e_1, \dots, e_4\}$, où e_i signifie que le point P se trouve dans le point d'abscisse i .

La matrice de transition de ce processus de Markov est :

	e_0	e_1	e_2	e_3	e_4
e_0	0	1	0	0	0
e_1	$1-p$	0	p	0	0
e_2	0	$1-p$	0	p	0
e_3	0	0	$1-p$	0	p
e_4	0	0	0	1	0

Si le mobile s'arrête quand il arrive en 0 ou en 4, que se passe-t-il ?

Les entrées p_{ij} d'une matrice de transition sont données par la probabilité pour que le système passe de l'état e_i à l'état e_j en une seule étape, on veut maintenant déterminer

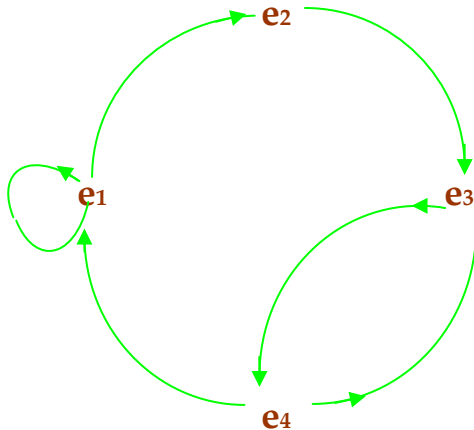
la probabilité $p_{ij}^{(n)}$ pour que le système passe de l'état e_i à l'état e_j en n étapes.

Ces probabilités constitueront les entrées de la matrice de transition de l' n -ième étape, matrice que l'on notera $P^{(n)}$.

N.B.

La matrice de transition peut être remplacée par un diagramme (le diagramme de transition):

Si, par exemple, on a :



La matrice de transition est alors :

$\frac{1}{2}$	$\frac{1}{2}$	0	0
0	0	1	0
0	0	0	1
$\frac{1}{2}$	0	$\frac{1}{2}$	0

Pour toute matrice de transition des exemples précédents écrire le diagramme de transition correspondant! Avant le 18 avril !

Si $\mathbf{p}$ est une distribution de probabilité et P une matrice de transition sur un ensemble d'états E , on appelle chaîne de Markov homogène de loi initiale $\mathbf{p}$ un processus aléatoire à valeurs dans E .
notion de tribu ??

Théorème 2 : Si P est la matrice de transition d'un processus de Markov et $p=(p_i)$ est la distribution de probabilité au temps t , alors pP sera la distribution de probabilité à l'instant $(t+1)$ et pP^n est la distribution de probabilité après n étapes i.e. au temps $(t+n)$

Démonstration :

Soit $\{e_0, e_1, \dots, e_m\}$, l'ensemble des états.

La probabilité pour que le système soit dans l'état e_j au temps t et dans l'état e_i au temps $(t+1)$ est : « probabilité d'être en e_j en t » $\times$ « probabilité d'être en e_i en $(t+1)$ » = $p_j \cdot p_{ji}$

Donc la probabilité pour que le système soit dans l'état e_i au temps $(t+1)$ sera : $\sum_{j=1}^m p_j p_{ji}$ car le système au temps t peut être ou en e_1 , ou en e_2 , ou en e_j , ou en e_m

Ceci nous permet d'affirmer que la distribution de probabilité au temps $(t+1)$ est :

$$p^* = \left(\sum_{j=1}^m p_j p_{j1}, \sum_{j=1}^m p_j p_{j2}, \dots, \sum_{j=1}^m p_j p_{jm} \right) \text{ et il suffit maintenant}$$

d'observer que l'expression de p^* n'est rien d'autre que pP !

Donc si la distribution initiale de probabilité est $p^{(0)}$ alors après une étape elle sera $p^{(1)} = p^{(0)}P$, après deux étapes elle sera $p^{(2)} = p^{(1)}P = p^{(0)}P^2$ et après n étapes on aura une distribution de probabilité égale à pP^n : **q.e.d.**

Théorème 3 : Soit P est la matrice de transition d'un processus de Markov au temps t , alors si on note $P^{(n)}$ la matrice de transition au temps $(t+n)$ et

on peut montrer que : $P^{(n)} \equiv P^n$

Démonstration :

Si le système est dans l'état e_i au temps t on note $p_{ij}^{(n)}$ la probabilité pour que le système soit dans l'état e_j au temps $(t+n)$.

La distribution de probabilité au temps t est $(0, 0, 0, \dots, 1, 0, 0)$, avec le 1 comme i -ième composante car le système est en e_i par hypothèse !!

Si on applique le Théorème 2 la distribution de probabilité en $(t+n)$ sera : $(0, 0, 0, \dots, 1, 0, 0)P^n$ à savoir la i -ième ligne de la matrice P^n (il suffit de se souvenir de la définition de produit de matrices.....)

Il est alors évident que $p_{ij}^{(n)}$ est la j -ième entrée de la i -ième ligne de la matrice P^n ce qui nous

permet de conclure (juste en visualisant les entrées de $P^{(n)}$)

que $P^{(n)} \equiv P^n$ **q.e.d.**

§1.4 Les états absorbants

Un état e_i d'un processus de Markov est « absorbant » si quand le système rentre dans l'état e_i il ne peut plus passer dans aucun autre état.

Peut-on dire que pour présenter des états absorbant un processus de Markov doit avoir une matrice de transition non régulière ?

Expliquer votre raisonnement.

EXEMPLE:

Soit M la matrice de transition d'un processus de Markov :

p_{11}	"	"	"	p_{15}
"	"	"	"	"
0	0	1	0	0
"	"	"	"	"
p_{51}	"	"	"	p_{55}

si $j \neq 3$ alors $p^{(n)}_{3j} = 0$ pour tout n , donc aucune puissance de M peut avoir des entrées toutes positives : M ne serait être régulière.

Généraliser cette constatation à toute matrice $n \times n$ ayant une ou plusieurs entrées unitaires sur la diagonale principale. !!!

Dans cet exemple e_3 est un état absorbent.

EXERCICES

1) Des électeurs sont appelés à déposer leur bulletin dans une urne les uns après les autres.

Sur un total de 500 votants 300 sont des hommes.

Quelle est la probabilité pour que le cinquième votant soit une femme ?

2) En observant le comportement d'un groupe de singes laissés dans une cage avec la porte ouverte, on a constaté que 70% des singes qui avaient quitté la cage la veille sortaient aussi le lendemain, alors que 40% de ceux qui n'étaient pas sortis la veille sortaient le lendemain.

Si le premier jour de l'observation seulement 50% des singes avaient quitté la cage que pouvez vous dire sur l'observation du troisième jour ? et du 365^{ème} jour ?

hint ! Ensemble des états = {sortir, ne pas sortir} distribution initiale de probabilité : (0.5 , 0.5)

matrice de transition :

.7	.3
.4	.6

Que faut-il faire avant d'utiliser l'aide proposée ??

(1) (2 points): Donner la définition d'état absorbant et de graphe transitif.

(2) (2 points): Trouver un multiple des vecteurs suivants, qui soit un vecteur de probabilité.

$$u = (1/4, 0, 1/2, 2/5) \quad v = (1, 2, 0, 4, 5)$$

(3) (2 points) : Parmi les matrices suivantes, quelles sont celles qui sont des matrices stochastiques :

(a)

1/2	1/3
1/4	3/4

(b)

2/3	0	1/3
0	1/2	1/2

(c)

2/3	-1/3	1/3
1/2	0	1/2
0	0	1

(d)

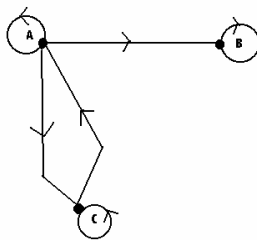
1/3	2/3
0	1

(e)

3/4	0	1/4
2/3	1/3	1/3
0	1	0

(4) (4 points): Lors de la coupe du monde de patinage artistique, une juge française a déterminé le graphique de transition de Markov mettant en scène les 3 plus grands compétiteurs de ces 5 dernières années.

Ce graphique expose la probabilité de victoire de chacun des compétiteurs d'une année sur l'autre.



a) Déterminer la matrice de transition P.

b) On suppose que C a gagné la 1ère année du recensement. Quelle est la probabilité pour que la juge française désigne B vainqueur à la fin de la 4^{ème} année.

c) Existe-t-il des états absorbants ; le diagramme de transition fait-il apparaître une périodicité, sinon vers quelle matrice T tend $P^{(n)}$?

d) Avec quelle fréquence C peut-il espérer triompher jusqu'à la fin de sa carrière en gardant la même juge ?

(5) (3 points):

Soit le jeu suivant :

1	2	3
0	-1	2
-1	1	3

Quel est la valeur du jeu, existe-t-il un ou plusieurs points selles, sont-ils échangeables ? Justifiez.

(6) (3 points) :

1	-2
0	1
2	0

Le jeu représenté ci-dessus sous forme normale a-t-il une valeur en stratégie pure ? En stratégie mixte ? Déterminer et donner la nature des stratégies optimales.

(7) (3 points) :

2	-3	-1	1	-1	1
2	a	-1	2	0	0
2	-2	1	0	1	2
1	-2	0	-1	0	0

On suppose, tout d'abord, le paramètre « a » positif. Simplifier la matrice.

Pour quelles valeurs de « a », n'a-t-on pas de valeur en stratégie pure ? Si $a = 4$ quelle est la valeur du jeu ?

(8) (3 points): Ecrire la forme normale d'un jeu (4,4) ayant au moins deux points selles, ne se trouvant pas sur la même colonne.

EXERCICES ++++

il suffit d'en demander.....

Ne pas utiliser pour le moment !!

****Rappels de mathématiques :**

1)Ensembles : Un ensemble E est bien défini quand on peut affirmer sans incertitude si un élément x appartient ou n'appartient pas à l'ensemble E. Un ensemble E est donc une collection d'objets appelés les éléments de E.

E est fini quand le nombre de ses éléments est fini i.e. quand $\text{card } E = n$

Notre intérêt principal sera d'apprendre à reconnaître aisément parmi tous les ensembles ceux qui seront ouverts, fermés, bornés, convexes, connexes.

Pour le moment on se limitera à des définitions plus qu'informelles qui doivent être complétées avec les observations effectuées en cours !!! A savoir :

Un ensemble E est ouvert si tous ses points sont intérieurs à E, i.e. $\forall x \in E \exists V(x) = \text{ensemble contenant } x \text{ et contenu dans } E$.



N.B. La différence entre les deux ensembles est évidente : le bord ne faisant pas partie de E_1 même si on prend le point x très proche du bord on trouvera toujours un ensemble qui

le contient et qui est contenu dans E_1 , alors que si un point est sur le bord de E_2 même si on choisit un ensemble très petit pour pouvoir contenir x il contiendra nécessairement des points qui n'appartiendront pas à E_2 !

Un ensemble E est convexe si $\forall x, y \in E$ le segment d'extrémités x et y est entièrement contenu dans E.

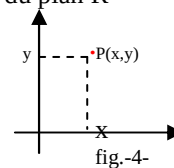


Pour E_1 $\lambda x + (1 - \lambda)y$, avec $0 < \lambda < 1$, $\in E_1$ alors que pour $E_2 \exists x$ et y dans E_2 t.q. $\lambda x + (1 - \lambda)y$, avec $0 < \lambda < 1$, $\notin E_2$

2)Produit cartésien : Si A et B sont deux ensembles le produit cartésien $A \times B$ est l'ensemble défini par $\{(a, b) : a \in A \text{ et } b \in B\}$

exemple :

$A = \mathbf{R}, B = \mathbf{R} \Rightarrow A \times B = \mathbf{R}^2 = \{P(x, y) : x \in \mathbf{R} \text{ et } y \in \mathbf{R}\} = \text{ensemble de tous les points P du plan } \mathbf{R}^2$



Ceci se généralise très facilement au produit cartésien de n ensembles :

$A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n) \text{ avec } a_1 \in A_1, \dots, a_n \in A_n\}$, $(a_1, a_2, \dots, a_n) = n\text{-uplet ordonné}$

3) Les Applications :

Si A et B sont deux ensembles, on appelle application de A dans B toute correspondance **f** qui à tout élément de A (domaine de définition) associe un et un seul élément de B (co-domaine) $f(A) \subset B$

$$f : A \rightarrow B$$

$$a \mapsto f(a) = b, \text{ avec } a \in A \text{ et } b \in B$$

Une application est donc une relation binaire $R \subset A \times B$, i.e. $a R b$ ssi $b = f(a)$

Si **f** et **g** sont deux applications t.q. $f : A \rightarrow B$ et $g : f(A) \rightarrow C$ alors on peut définir l'application composée de **f** par **g**, notée **g o f** $= g(f(a)) \in C$ (!! il est évident que $g \circ f \rightarrow f \circ g$)

Le **graphe** d'une application $f : A \rightarrow B$ sera donc l'ensemble $G(f) = \{(x, f(x)) \in A \times B, \text{ avec } x \in A\}$

Exemples d'applications :

- | | | |
|----|---|--|
| 1) | $f : A \rightarrow B$, avec $B = \mathbf{R}$ | est une fonction numérique |
| 2) | $f : A \rightarrow B$, avec $A \subset \mathbf{R}$ et $B = \mathbf{R}$ | est une fonction réelle d'une variable réelle |
| 3) | $f : A \rightarrow B$, avec $A \subset \mathbf{R}^n$ et $B = \mathbf{R}$ | est une fonction réelle de n variables |
| 4) | $f : A \rightarrow B$, avec $A = \mathbf{N}$ | est une suite à éléments dans B |
| 5) | $f : A \rightarrow B$, avec $A = \mathbf{N}$ et $B \subset \mathbf{R}$ | est une suite réelle ($f : n \mapsto u_n \in \mathbf{R}$) |
| 6) | $f : \mathbf{R} \rightarrow \mathbf{R}^+$ t.q. $x \mapsto x$ pour $x \geq 0$ et $x \mapsto -x$ pour $x < 0$ | est la valeur absolue , on note : $x \mapsto x $ |

6)Les fonctions réelles :

On appelle fonction réelle de n variables une application $f: E \subset \mathbb{R}^n \rightarrow \mathbb{R}$ t.q. $M \in E \rightarrow f(M) \in \mathbb{R}$
i.e. $(x_1, x_2, \dots, x_n) \in E \rightarrow y = f(x_1, x_2, \dots, x_n)$

(y est donc une variable dépendante du choix de M dans E)

$y = f(x_1, x_2, \dots, x_n)$ est l'image de $M(x_1, x_2, \dots, x_n)$ par la fonction f , c'est par abus de langage que souvent on appelle $y = f(x_1, x_2, \dots, x_n)$ fonction ! (cet abus est autorisé seulement si le contexte est clair, cf cours de math.)

Exemples :

(1) Soit $(x_1, x_2, \dots, x_n) \in X$ une dotation possible d'un consommateur, on notera $u = U(x_1, x_2, \dots, x_n)$ la valeur de l'utilité atteinte par le consommateur et on appellera fonction d'utilité l'application qui à toute dotation de X fait correspondre une seule valeur réelle u .

(2) Soit f de $E \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par : $f(x, y) = \frac{\sqrt{\frac{x}{y}}}{x+y-4}$,

alors $E = \{P(x, y) \in \mathbb{R}^2 : x \in \mathbb{R} \text{ et } y \in \mathbb{R} \text{ t.q. } x \geq 0, y > 0 \text{ et } x-y \rightarrow 4 \text{ ou } x=0, y < 0 \text{ et } x+y \rightarrow 4\}$

(3) Soit f de $E \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par :

$f(x, y) = \frac{xy}{x^2 + y^2 - 4}$, alors $E = \{P(x, y) \in \mathbb{R}^2 : \text{t.q. } x^2 + y^2 \rightarrow 4\}$

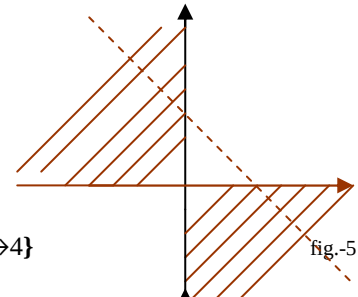


fig.-5

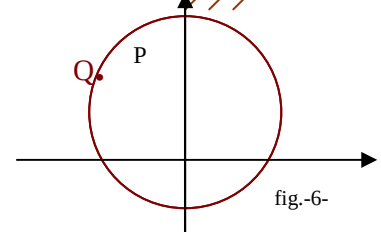


fig.-6-

$Q \notin E, \quad P \in E \text{ et } P' \in E$

4) Les dérivées partielles : Soit Ω un ensemble ouvert de $\mathbb{R}^n$ et $X_0(x_1^0, x_2^0, \dots, x_n^0)$ un point de Ω et soit $f: \Omega \rightarrow \mathbb{R}$ une fonction réelle de n variables réelles t.q.

$(x_1, x_2, \dots, x_n) \rightarrow f(x_1, x_2, \dots, x_n) = r$

Si on suppose les $(n-1)$ variables fixées en $x_1^0, x_2^0, \dots, x_{i-1}^0, x_{i+1}^0, \dots, x_n^0$, on obtient une fonction $f(x_1^0, x_2^0, \dots, x_{i-1}^0, x_i, x_{i+1}^0, \dots, x_n^0)$ de la seule variable x_i définie au voisinage de x_i^0

On dit alors que f admet une dérivée partielle première par rapport à la i -ème variable x_i au point M_0 si

$\lim_{h \rightarrow 0} \frac{f(x_1^0, x_2^0, \dots, x_{i-1}^0, x_i + h, x_{i+1}^0, \dots, x_n^0) - f(x_1^0, x_2^0, \dots, x_i^0, \dots, x_n^0)}{h}$ existe dans $\mathbb{R}$

On note : $f'_{xi}(X_0)$ ou $\frac{\partial f(x_0)}{\partial x_i}$

On appelle gradient de f le vecteur dont les composantes sont les n dérivées partielles de f i.e.

$\vec{\text{grad}} f = (f'_1(X_0), f'_2(X_0), \dots, f'_n(X_0))$ on note: $\nabla f = \vec{0}$

Si on considère l'application linéaire df_x on sait déterminer la matrice associée dans une base donnée, si on choisit la base canonique

$\{(1, 0, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, 0, \dots, 1)\}$ pour $\mathbb{R}^n$ on a :

$$\left\{ \begin{aligned} df_x(1, 0, \dots, 0) &= \frac{\partial f(X)}{\partial x_1} \\ &\dots\dots\dots \\ df_x(0, 0, \dots, 1) &= \frac{\partial f(X)}{\partial x_n} \end{aligned} \right\}$$



5) Les fonctions homogènes :

Soit $f : \Omega \rightarrow \mathbb{R}$ une fonction réelle de n variables réelles définie par $y=f(x)$,

on dit que f est homogène de degré g si $f(tx)=t^g f(x)$ (le domaine de définition de f doit être un cône de $\mathbb{R}^n$, à savoir si $x \in \Omega$, alors $\lambda x \in \Omega, \lambda > 0$)

Exemples :

(I) Pour $Y=x^2$ on a : $f(tx)=(tx)^2 = t^2 x^2 = t^2 f(x)$ donc il s'agit d'une fonction homogène de degré 2.

(II) Pour $Y=x^2 + 1$ on a : $f(tx)= t^2 x^2 + 1 \neq t^2 (x^2 + 1)$ donc cette fonction n'est pas homogène.

(III) Pour $Z=\sqrt{xy}$ on a : $f(tx)= \sqrt{txty} = t\sqrt{xy} = t^1 f(x,y)$ donc cette fonction est homogène de degré 1.

6) L'identité d'Euler :

Si $y=f(x)$ est une fonction de n variables réelles, homogène de degré g et ayant les dérivées partielles premières continues sur Ω , la définition nous permet d'écrire :

$f(tx)=t^g f(x), \forall X \in \Omega$ et si on dérive les deux membres de cette égalité par rapport à t , on obtient :

$\sum f_i'(tx)x = g t^{g-1} f(x)$, cette égalité étant vérifiée pour toute valeur de t , elle le sera en particulier pour $t=1$

i.e. $\sum_{i=1}^n f_i'(x).x = g.f(x)$ **Identité d'Euler** pour les fonctions homogènes.

Pour $n=1$ (fonction réelle d'une seule variable réelle) on a $f'(x)=g.f(x)$

i.e. la dérivée d'une fonction homogène d'une variable est égale au degré d'homogénéité multiplié par la fonction elle-même.

Pour $n=2$ (fonction réelle de deux variables réelles) on a : $f(tx,ty)=t^g f(x,y)$ et en dérivant les deux membres de l'égalité par rapport à t , on obtient $f'_x.x + f'_y.y = g.f(x,y)$



Fonctions équivalentes* :

(cf. cours de math. « tangente à une courbe en un point ».....)

$e^x - 1 \sim x$ dans un voisinage de $x=0$

$\ln(1+x) \sim x$ dans un voisinage de $x=0$,

Application :

$\lim_{x \rightarrow +\infty} \left(1 + \frac{1}{x}\right)^x = e$ en effet si on pose $\frac{1}{x} = t$ on voit que $t \rightarrow 0$ pour $x \rightarrow +\infty$

or $\ln\left(1 + \frac{1}{x}\right)^x = x \ln\left(1 + \frac{1}{x}\right) = \frac{1}{t} \ln(1+t)$ donc $\lim_{x \rightarrow +\infty} \ln\left(1 + \frac{1}{x}\right)^x = \lim_{t \rightarrow 0} \frac{1}{t} \ln(1+t) = 1$, (!! $\lim_{t \rightarrow 0} \left(\frac{1}{t}\right) = 1$)

de plus par définition de logarithme $v=e^{\ln v}$

donc $\lim_{x \rightarrow +\infty} \left(1 + \frac{1}{x}\right)^x = e^1 = e$

Si on garde ce résultat en mémoire, on peut alors obtenir facilement la solution au problème suivant :

Un agent place à la banque une somme S à un taux d'intérêt annuel r , après un an, si il n'effectue aucun prélèvement,

il pourra disposer de la somme $S' = S + rS = S(1+r)$,

après deux ans de la somme $S'' = S' + rS' = S(1+r) + S(1+r)r = S(1+r)(1+r) = S(1+r)^2$

..... après n années de la somme $S^* = S(1+r)^n$

Si on veut calculer S' quand le taux d'intérêt r est continu, on peut observer que sur une durée de temps $dt = \frac{1}{N}$ d'année, avec N très grand, l'intérêt accumulé sera $r dt = r \frac{1}{N}$

donc après un an $S' = S(1 + r \frac{1}{N})^N$ et $\lim_{N \rightarrow +\infty} S' = e^r$ $\Rightarrow$ après n années $S^* = S e^{r n}$

Faire au mois un exercice de calcul complet ! Prendre une somme $S=1000$ euros et calculer S^* pour un taux d'intérêt annuel de $r = 5\%$, puis avec un taux d'intérêt continu de $r = 4\%$, pour un placement de 10 ans.

Ceci vous servira pour pouvoir décider entre un taux annuel et un taux mensuel..... ?????

Fonctions Convexes :

~~Fonctions Convexes~~

Déjà vu en cours de math. ! à savoir parfaitement !!

Ensembles convexes :

Un ensemble $E \subset \mathbf{R}^n$ est convexe si $\forall x, y \in E$ le segment d'extrémités x et y est entièrement contenu dans E (on note : E convexe ssi $\forall x \in E, \forall y \in E \Rightarrow [x, y] \subset E$)



Pour E_1 $\lambda x + (1 - \lambda)y$, avec $0 < \lambda < 1$, $\in E_1$ alors que pour E_2 , $\exists x$ et y dans E_2 t.q. $\lambda x + (1 - \lambda)y$, avec $0 < \lambda < 1$, $\notin E_2$

Exemple :

Un hyperplan π de $\mathbf{R}^n$ est un ensemble convexe (**à vous de prouver qu'il est aussi concave !**)

Soit $f(x) = b$, avec $f(x) = a_1 x_1 + \dots + a_n x_n$ l'équation de l'hyperplan π .

Si x et x' sont deux points de π alors $f(x) = f(x') = b$

Mais $f(\lambda x + \lambda' x') = \lambda f(x) + (1 - \lambda) f(x') = \lambda b + (1 - \lambda) b = b$ q.e.d.

Tout segment $[x, x']$ appartient à π , $\forall x \in \pi, \forall x' \in \pi$ donc l'hyperplan est convexe.

Vous pouvez prouver que :

- (1) Si f est une application linéaire de $C \subset \mathbf{R}^n$ (convexe) dans $\mathbf{R}^p$ alors $f(C)$ sera un ensemble convexe de $\mathbf{R}^p$.
- (2) Si C et C' sont deux ensembles convexes de $\mathbf{R}^n$ alors $C \cup C'$ est aussi un convexe de $\mathbf{R}^n$.

On appellera enveloppe convexe d'un ensemble $S \subset \mathbf{R}^n$ l'intersection de tous les sous-ensembles convexes qui contiennent S , on notera $[S]$ l'enveloppe convexe de S .

On appellera point extrémal d'un ensemble convexe S tout point $x \in S$ pour lequel il est impossible de trouver deux autres points x^1 et x^2 de S t.q. $x = \lambda x^1 + (1 - \lambda) x^2$ avec $0 < \lambda < 1$.

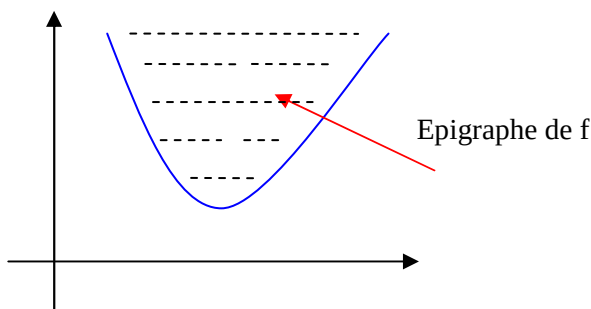
Exemple :

Tous les points de la circonférence d'un disque de $\mathbf{R}^2$ sont extrémaux.

Fonctions convexes :

On appelle épigraphe d'une fonction $f : D_f \subset \mathbf{R}^n \rightarrow \mathbf{R}$ l'ensemble des points situés « au dessus » de son graphe.

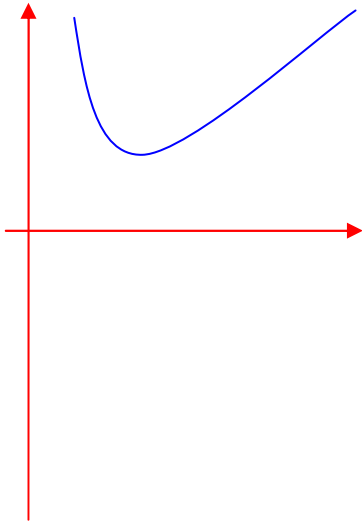
On note $E(f)$ et vous pouvez en préciser la définition..... $E(f) = \{(x_1, \dots, x_n, x_{n+1}) \dots\}$



On peut maintenant énoncer qu'on appelle convexe toute fonction dont l'épigraphe est convexe.

Une fonction g sera concave si $-g$ est convexe .

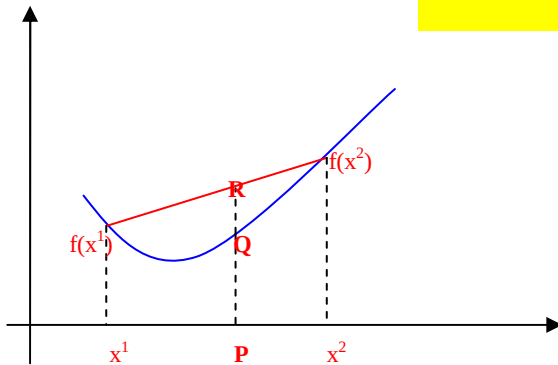
Dessinez sur un même repère une fonction g et son opposée $-g$



puis constatez !

Caractérisation des fonctions convexes :

Une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ est convexe ssi $f(\lambda_1 x^1 + \lambda_2 x^2) \leq \lambda_1 f(x^1) + \lambda_2 f(x^2)$, avec $0 < \lambda_i < 1$ et $\sum_i \lambda_i = 1$



Tout point P du segment $[x^1, x^2]$ s'écrit $P = \lambda_1 x^1 + \lambda_2 x^2$

De même $Q = f(P) = f(\lambda_1 x^1 + \lambda_2 x^2)$ et $R = \lambda_1 f(x^1) + \lambda_2 f(x^2)$ donc ordonnée de $R >$ ordonnée de Q

Dès que vous avez complété la caractérisation des fonctions convexes dans $\mathbb{R}^2$, passez au cas général :

Proposition 1 :

Soit $f : C \subset \mathbb{R}^n \rightarrow \mathbb{R}$ alors f est convexe ssi $f(\sum_{i=1}^p \lambda_i x^i) \leq \sum_{i=1}^p \lambda_i f(x_i)$ $\forall p \in \mathbb{N}, \forall x^1, \dots, x^i \in C, \forall \lambda_i > 0$ t.q. $\sum \lambda_i = 1$

A démontrer.....(par récurrence ?)

Proposition 2 :

Toute combinaison linéaire à coefficients positifs de fonctions convexes est une fonction convexe.

Proposition 3 :

Si f est une fonction de $I \subset \mathbb{R}$ dans $\mathbb{R}$, alors f est convexe ssi f' est une fonction croissante (i.e. $f'' \geq 0$).

Proposition 4 :

Si f est une fonction de $I \subset \mathbb{R}$ dans $\mathbb{R}$, alors f est convexe ssi $f(x) \geq f(x_0) + (x - x_0)f'(x_0) \quad \forall x_0 \in I$,
i.e. f convexe ssi la courbe représentative (graphe) de f est toujours au dessus de la tangente en tout point.

Proposition 5 :

Soit $f : C \subset \mathbb{R}^n \rightarrow \mathbb{R}$ alors f est convexe ssi son graphe est situé au dessus de tout hyperplan tangent,

à savoir ssi $f(x) \geq f(x^0) + \sum_{i=1}^n (x_i - x_{i^0}^0) \frac{\partial f(x^0)}{\partial x_i}, \forall x^0 \in C.$

Considérons maintenant une fonction $f(x,y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F$ (cf ; formes quadratiques....)
Nous avons :

Ch3. OPTIMISATION

§3.0 Généralités



****** Vous avez déjà appris que :

- 1) La matrice des dérivées partielles premières d'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ différentiable dans $x \in \Omega$, ensemble ouvert, s'appelle la matrice Jacobienne, notée J

$$J = \begin{pmatrix} f_1^1 & \dots & \dots & f_n^1 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ f_1^m & \dots & \dots & f_n^m \end{pmatrix}$$

2) Pour $m=1$, la matrice se réduit au vecteur ligne $(f'_1(x_0), f'_2(x_0), \dots, f'_n(x_0))$ appelé le gradient de la fonction f , calculé au point x_0 de Ω , on note :

$$\nabla f = \overrightarrow{\text{grad } f} = (f'_1(x_0), f'_2(x_0), \dots, f'_n(x_0))$$

3) La matrice des dérivées partielles secondes s'appelle la matrice Hessienne, notée H

$$H = \begin{pmatrix} f''_{11} & \dots & \dots & f''_{1n} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ f''_{n1} & \dots & \dots & f''_{nn} \end{pmatrix}$$

****Attention !**

Il n'est pas question ici de vous fournir un recueil de « définitions », mais de préciser notations et vocabulaire utilisés par la suite afin d'éviter toute mésentente

A vous de pousser votre zèle jusqu'à réviser sérieusement votre cours de mathématiques et (mieux encore !) d'en profiter pour vous poser quelques questions.....

Par exemple :

Quand $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est différentiable ?

L'existence et la continuité des dérivées partielles en un point pour une telle fonction suffisent-elles à garantir que $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sera différentiable dans ce point ? etc.....

Ne pas oublier que :

On appelle **conique** toute courbe algébrique représentée par :

$$(1) f(x,y) = ax^2 + 2bxy + cy^2 + 2dx + 2ey + f$$

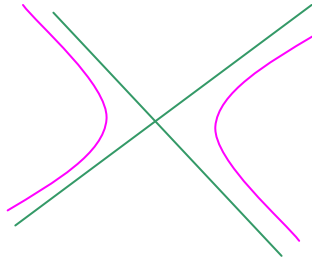
Si $f(x, y)$ peut s'écrire sous la forme d'un produit de deux polynômes de premier degré,

La conique est dégénérée et se décompose en deux droites. Si elle n'est pas dégénérée alors :

(1') $f(x,y) = ax^2 + 2bxy + cy^2$ est l'équation de l'ensemble des droites de direction asymptotique passant par l'origine.

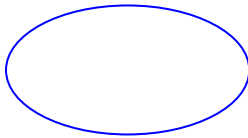
L'étude des directions asymptotiques donne les résultats suivants :

- (i) $b^2 - 4ac > 0$ ssi $\exists$ deux directions asymptotiques réelles et distinctes :



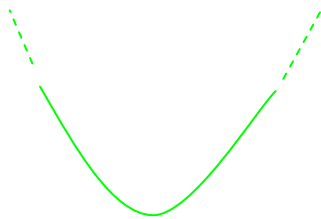
la courbe est une hyperbole

- (ii) $b^2 - 4ac < 0$ ssi il n'y a pas de direction asymptotique réelle :



la courbe est une ellipse

- (iii) $b^2 - 4ac = 0$ ssi $\exists$ une direction asymptotique double :



la courbe est une parabole

N.B. La nature d'une conique est indépendante du système d'axes choisi, si on choisit les axes coordonnés on a :

- (i) $\frac{x^2}{A^2} - \frac{y^2}{B^2} = 1$ pour l'hyperbole

- (ii) $\frac{x^2}{A^2} + \frac{y^2}{B^2} = 1$ pour l'ellipse (si $A=B$ on retrouve l'équation du cercle de rayon A)

- (iii) $y^2 - 2px = 0$ pour la parabole

Applications :

(1) Considérons le programme :

$$\text{Max } (6x + 2xy - 2x^2 - 2y^2) \quad \text{t.q. } 3x + 4y - 6 \leq 0$$

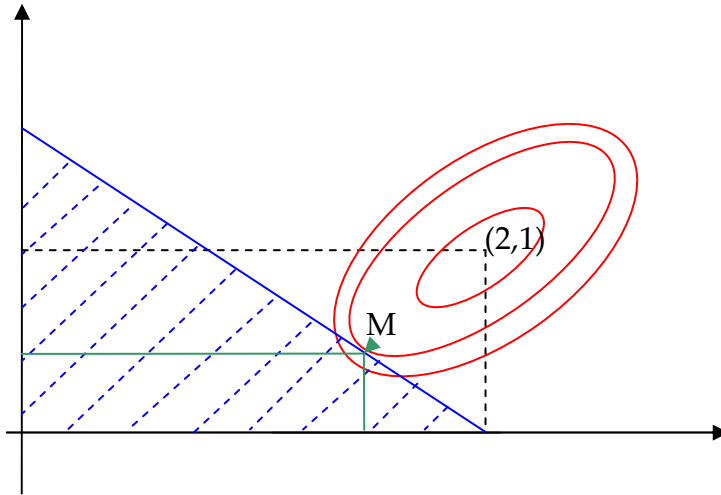
On peut d'abord étudier l'allure de la fonction objectif i.e. $\text{Max } (6x + 2xy - 2x^2 - 2y^2)$

$$\text{CNPO : } \begin{cases} 6 - 4x + 2y = 0 \\ -4y + 2x = 0 \end{cases} \Rightarrow \begin{cases} x = 2 \\ y = 1 \end{cases}$$

CSSO : $\det \begin{pmatrix} -4 & 2 \\ 2 & -4 \end{pmatrix} > 0$ et $\det \begin{pmatrix} -4 \end{pmatrix} < 0$ la forme quadratique des dérivées secondes de la fonction

objectif est définie négative, donc la fonction objectif admet un maximum en (2,1)

Dans le plan (2,1) est donc le centre d'une famille d'ellipse

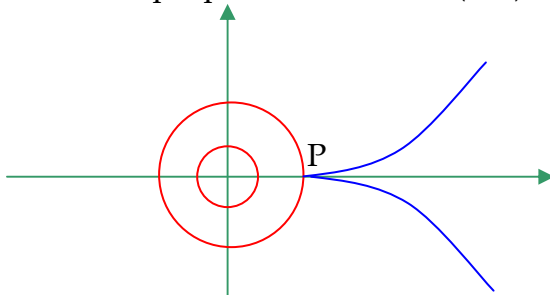


la contrainte étant représentée par la partie hachurée en figure le maximum de f sous la contrainte donnée sera donc $f(M)$ avec M point de tangence entre la droite $3x+4y-6=0$ et une des ellipses.

(2) $\min x^2+y^2$ t.q. $(x-1)^3-y^2=0$ Att !! les hypothèses du théorème de la fonction implicite ne sont pas vérifiées ! donc la méthode de Lagrange ne peut pas être utilisée.

Le problème est du type $\min f(x,y)$ t.q. $g(x,y)=0$ avec $g'_1=0$ pour $x=1$ et $g'_2=0$ pour $x=0$, donc la contrainte n'est pas qualifiée en (1,0) ! que faire ?

Pourtant si on observe la fonction objectif, son intersection avec le plan $z=c$ est représentée en figure par la famille de cercles de centre (0,0) qui ne peut avoir une intersection possible avec la contrainte que pour $c \geq 1$, donc $P(1,0)$ t.q. $f(1,0)=1$ sera le minimum sous la contrainte donnée.



En effet il s'agissait de déterminer sur le parabolôide d'équation $z=x^2+y^2$ le point "le plus bas" autorisé par la contrainte.....

i.e. chercher le point d'ordonnée minimale appartenant à la courbe intersection des surfaces représentées par la fonction objectif et par la contrainte !

SI VOUS ETES TETUS :

Prenez $L(x,y,\lambda) = x^2+y^2 - \lambda((x-1)^3-y^2)$ et vérifiez l'absurdité de votre procédé !

Proposition d'un étudiant :

$$\text{CNPO} \Rightarrow \begin{cases} 2x - 3\lambda(x-1)^2 = 0 \\ 2y + 2\lambda y = 0 \end{cases} \Rightarrow \lambda = -1 \Rightarrow x = \frac{2 \pm \sqrt{-5}}{3} \text{ pas de solution réelle}$$

Votre avis ??

$$(3) \text{ opt } 2x^2 - 3xy + y - 3z + 2yz$$

$$\text{CNPO} : \nabla f = \overrightarrow{\text{grad } f} = \vec{0}$$

$$\begin{cases} 4x - 3y = 0 \\ -3x + 1 + 2z = 0 \\ -3 + 2y = 0 \end{cases} \Rightarrow \begin{cases} x = \frac{9}{8} \\ 2z = 3x - 1 \\ y = \frac{3}{2} \end{cases} \Rightarrow \begin{cases} x = \frac{9}{8} \\ z = \frac{19}{16} \\ y = \frac{3}{2} \end{cases} \quad \text{vérifier les calculs !}$$

CSSO : $f''_{xx} = 4$; $f''_{yy} = 0$; $f''_{zz} = 0$; $f''_{xy} = -3$; $f''_{xz} = 0$; $f''_{yz} = 2$ (les dérivées secondes étant continues l'ordre de dérivation.....compléter.....)

$$\text{Donc } H = \begin{bmatrix} 4 & -3 & 0 \\ -3 & 0 & 2 \\ 0 & 2 & 0 \end{bmatrix}$$

$$\text{Et } \det 4 > 0 ; \det \begin{bmatrix} 4 & -3 \\ -3 & 0 \end{bmatrix} < 0 ; \det H < 0$$

Que pouvez vous en déduire ????

Que devez vous faire ?

$\Delta = 9 > 0$ donc la fonction objectif n'est ni concave, ni convexe.....

"OPTIMISATION" Exercices

1) Montrer que si on veut déterminer un extremum relatif de la fonction

$f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x,y) = 3x^2 + 2xy + y^2 - x + 2$ et les CNPO (conditions nécessaires du premier ordre) sont vérifiées, il ne sera pas indispensable de vérifier les CSSO (conditions suffisantes du second ordre).

2) Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par $f(x,y) = \ln(x^2 + y^2) - xy$.

Prouver que $f(1,1)$ ne sera ni un maximum ni un minimum relatif.

3) Déterminer le profit maximum Π d'une entreprise qui produit deux biens X et Y , sachant que le prix de vente d'une unité x de X est $p = (6-x)$ et le prix de vente d'une unité y de Y est $q = (2-y)$.

L'entreprise est en situation de monopole et sa fonction coût est $c(x,y) = x^2 - xy + y^2$.

"Rappel" : $\Pi(x,y) = \text{Revenu} - \text{coût}$, ici $\Pi(x,y) = (6-x)x + (2-y)y - c(x,y)$.

4) Résoudre :

(i) $\text{opt } xy \text{ t.q. } x+y=6$.

(ii) $\text{Max } 3 \cdot x^{\frac{1}{2}} \cdot y^{\frac{1}{2}} \text{ t.q. } 2x+3y=5$.

(iii)* $\text{min } x^2 + y^2 \text{ t.q. } (x-1)^3 - y^2$.

(4) $\text{Max } xyz \text{ t.q. } x+y+z=1 \text{ et } x-y=0$.

(v)* $\text{opt } 2x^2 - 3xy + y - 3z + 2yz \text{ t.q. } 3x - 2y = 4 \text{ et } x+z=1$

(vi) $\text{opt } x^2 + y^2 \text{ t.q. } x^2 - 2xy + y^2 = \alpha^2$.

5) Résoudre le programme suivant :

$$\begin{array}{l} \text{Max } 4x^2 - 2y^2 + 8xy \\ \text{t.q. } x - 2y = 0 \end{array}$$

6)* Considérons le programme

$$\text{Max } f(x,y) - vx - wy, \text{ avec } f \text{ continûment différentiable}$$

et $\begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$ forme quadratique définie négative,

$\forall \mathbf{M}(\mathbf{x}, \mathbf{y}) \in \Omega$, domaine de définition de f .

a) Démontrer que la valeur optimale $\mathbf{x}^*$ peut s'exprimer en fonction de v et de w .

b) Démontrer que $\frac{\partial x}{\partial v} < 0$.

7) * Soient $\mathbf{y} = (y_1, y_2, \dots, y_n)$ et $\mathbf{x} = (x_1, x_2, \dots, x_n)$ deux points de $\mathbb{R}^n$.

Déterminer les valeurs de $\mathbf{a}$ et de $\mathbf{b}$ qui minimisent $\sum_{i=1}^n (y_i - (a + bx_i))^2$.

8) Un consommateur a une fonction d'utilité à variables séparables i.e.

$U(x_1, x_2, x_3) = f(x_1) + g(x_2) + h(x_3)$, avec $\mathbf{x} = (x_1, x_2, x_3) =$ vecteur de consommation.

Soit p_i , pour $i=1, 2, 3$ le prix du bien i et soit y le revenu de l'agent

(p_i et y sont indépendants de $\mathbf{x}$) :

(a) Ecrire le programme de maximisation de l'utilité qui décrit le comportement de l'agent, dans le cas où il dépense totalement son revenu.

(b) La contrainte est-elle qualifiée ?

(c) Ecrire les CNPO et les CSSO pour ce programme, puis montrer que si f', g', h' sont strictement positives et f'', g'', h'' sont strictement négatives $\forall \mathbf{x} \in \Omega$, domaine de définition de U , alors les CSSO sont toujours vérifiées.

(d) Si $\mathbf{x}^*$ est une solution régulière que pouvez vous dire des signes

de $\frac{\partial x_1^*}{\partial p_1}$ et $\frac{\partial x_2^*}{\partial p_2}$?

(e) Sachant qu'un bien est inférieur si $\frac{\partial x_i^*}{\partial y} < 0$, un de biens concernés peut-il être inférieur ? Expliquer clairement.

9) (Sujet d'examen – Université Paris 1- DEUG 2 Sciences Economiques)

(i) Déterminer les valeurs de x et de y qui permettent à la fonction définie

Par $f(x,y)= 10+2x+3y-x^2-2y^2$ d'atteindre son optimum z^* , dont on précisera la nature, quand les variables sont liées par la contrainte $x+3y= \frac{1}{2}$.

(ii) Résoudre et commenter ensuite le programme

$$\text{opt } x+3y- \frac{1}{2} \text{ t.q. } 10+2x+3y-x^2-2y^2= z^*$$

Solutions proposées par



"OPTIMISATION" l'essentiel

à développer !!!!!

§1- Les formes quadratiques

Les techniques de l'algèbre linéaire sont indispensables pour traiter les fonctions numériques de n variables de la forme :

$$F(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i y_j \quad \text{i.e. les formes quadratiques, qui ont un rôle très important dans la détermination}$$

des CN et des CS pour la détermination des extréma.

Vous savez déjà que $F(x)$ peut s'écrire sous forme matricielle : $F(x) = x'Ax$

(avec A $n \times n$ matrice symétrique et x vecteur $n \times 1$)

On dit que la matrice A , et donc la forme quadratique à la quelle elle est associée,

est : (1) **définie positive** ssi $x'Ax > 0, \forall x \neq 0$

définie négative ssi $x'Ax < 0, \forall x \neq 0$

Att !!

Une forme définie positive, ou négative, reste définie positive, ou négative, même si exprimée dans de nouvelles variables à condition que la transformation effectuée ne soit pas singulière.

On appelle forme diagonale une forme quadratique qui contient seulement les termes au carré. Il est possible de transformer une forme quadratique dans une forme diagonale en effectuant une transformation de variables orthogonale.

Prenons la transformation $x = Qy$ avec Q matrice dont les colonnes sont des vecteurs propres ortho normaux de A qui engendrent E^n

(*réviser le cour d'algèbre au plus vite !)

la forme diagonale résultante aura comme coefficients des y_i^2 les valeurs propres λ_i de A .



au travail !!! Que faut-il faire pour obtenir les valeurs propres ?..... tout ceci est essentiel, car si on connaît les valeurs propres de A on peut énoncer que:

A est : (1') **définie positive** ssi $\lambda_i > 0, \forall i$
définie négative ssi $\lambda_i < 0, \forall i$

Si n n'est pas trop grand il sera demandé, dans les exercices d'application, d'utiliser le «fait» suivant :

A est : (1'') **définie positive** ssi $\det A_i > 0, \forall i$
définie négative ssi $(-1)^i \det A_i > 0, \forall i$

avec A_i = sous matrices obtenues à partir de A en supprimant les $(n-i)$ dernières lignes et les $(n-i)$ colonnes correspondantes. Les A_i sont les mineurs principaux de la matrice A .



Pouvez vous démontrer tous seuls que (1') et (1'') sont deux énoncés équivalents ?

Théorème (C.S.S.O.): Soit f une fonction de n variables définie sur un ouvert Ω de $\mathbb{R}^n$.

Dans un point $M_0 \in \Omega$ t.q. $\nabla f(M_0)=0$ la fonction f atteindra un maximum local si la forme quadratique $dx' \left(\frac{\partial^2 f}{\partial x^2} \right) dx$ est définie négative (et un minimum locale si la même forme quadratique est définie positive).

- Pour $n=2$ vérifier que vous retrouvez les CSSO énoncées en DEUG1 en utilisant la formule de Taylor.... Vos questions ???

Petit rappel provisoire :

Donc, par exemple, pour résoudre un programme de maximisation sans contraintes avec une fonction objectif f définie sur un ouvert Ω de $\mathbb{R}^n$ et ayant des dérivées partielles continues (i.e. pour chercher un maximum relatif de f dans un point $M_0 \in \Omega$)

on doit :

- 1) . se souvenir (et avoir compris !) que : condition nécessaire (mais non suffisante !)

pour que (*) $\begin{cases} \text{Max} f(x) \\ x \in \Omega \end{cases}$ admette la solution M_0 est que $\nabla f(M_0)=0$

Att ! ceci revient à dire que si $f(M_0)$ est un maximum alors.....

la solution du système $\nabla f(M_0)=0$ ne vous fournira que les coordonnées du point susceptible de maximiser f , c'est tout !! vrai ou faux ??

- 2) Si on ne sait pas étudier la convexité de f ,
si on ne sait pas déduire la « forme » de la surface à partir des courbes de niveau,
il faut appliquer le théorème qui nous fournit les CSSO en choisissant entre (1') et (1'')
ou encore vous amuser à faire un étude local.

Nous pouvons commencer « à rédiger «

§2.1' optimisation sans contraintes

Le théorème de Weierstrass (cf cours de Math.) nous assure que

le programme $\begin{cases} \text{Max} f(x) \\ x \in S \end{cases}$ admet une solution pour f continue et S compact.

Nous pouvons donc énoncer :

Théor-CNPO : Si $x^* \in \text{Int } S$ est t.q. $f(x^*)$ est un maximum relatif pour la fonction f et si $\nabla f(x)$ existe dans un voisinage de x^* , alors nécessairement $\nabla f(x^*)=0$

Démonstration :

Si f admet des dérivées partielles continues la démonstration n'est qu'une conséquence du Théorème de la valeur intermédiaire. **Pourquoi ? Expliquer !!!!!**

Aussi :

Soit $g(x_i) \equiv f(x_1^*, \dots, x_{i-1}^*, x_i, x_{i+1}^*, \dots, x_n)$, avec $x^*=(x_1^*, \dots, x_n^*)$ une fonction de la seule variable x_i ayant un maximum en x_i^* donc t.q. $g'(x_i^*)=0$, mais alors $f'_i(x^*)=0$ q.e.d.

est suffisamment clair ?

A propos de la question posée à la page 35

Réservé (à ceux qui ont envie de le faire : de 30 à 90 minutes !) après jetez l'éponge !

Considérons la suite de matrices $C_n \equiv A - \frac{1}{n} I$

Si A est définie négative C_n le sera aussi, car $x' C_n x \equiv x' A x - \frac{1}{n} x' x$

le déterminant : $\det(C_n)_i$ est une fonction continue des éléments de C_n
puis....



En cours on vous a demandé d'écrire les CNPO et les CSSO pour le programme :

$$\text{Max } f(x, y) - \omega_1 x - \omega_2 y$$

On précise maintenant que f est continûment différentiable et que la matrice Hessienne des dérivées seconde est définie négative $\forall x, \forall y$

(a) Démontrer que x^* , valeur optimale de x, peut s'exprimer en fonction de ω_1, ω_2 et que $\frac{\partial x}{\partial \omega_1} < 0$

Solution 1 :

$$\text{CNPO (1)} \begin{cases} f'_1(x^*, y^*) - \omega_1 = 0 \\ f'_2(x^*, y^*) - \omega_2 = 0 \end{cases} \quad \text{i.e.} \quad \begin{cases} F(x, y, \omega_1) = 0 \\ G(x, y, \omega_2) = 0 \end{cases}$$

Observer avec attention F et G, comment est-elle la matrice Jacobienne ??

Je prétends qu'elle est définie négative, donc que je peux utiliser le théorème de la fonction implicite, Vrai ou faux ? Expliquez synthétiquement !

On continue....

Si vrai, il existe ϕ et ψ t.q. $x = \phi(\omega_1, \omega_2)$ et $y = \psi(\omega_1, \omega_2)$

Pourquoi ?

Si on calcule la différentielle des CNPO par rapport à ω_1 on a :

$$(2) \begin{cases} f_{11} \frac{\partial x}{\partial \omega_1} + \frac{\partial y}{\partial \omega_2} = 1 \\ f_{21} \frac{\partial x}{\partial \omega_1} + f_{22} \frac{\partial y}{\partial \omega_2} = 0 \end{cases}$$

Deux équations, deux variables : on peut résoudre par rapport à $\frac{\partial x}{\partial \omega_1}$

A savoir : $\frac{\partial x}{\partial \omega_1} = \frac{f_{22}}{f_{22} \cdot f_{11} - (f_{21})^2} < 0$ pourquoi ? Utilisez les informations sur la matrice hessienne, puis

la règle des signes pour justifier formellement.

Comment calculez vous les solutions du système (2) ?

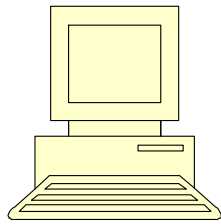
Est-il un système de Cramer ?

Les réponses aux questions posées et les commentaires demandés sur les propositions et les suggestions (hints) sur le ch. "optimisation" doivent être formulés avant le lundi 28 avril.

Seuls 4 étudiants ont participé cette semaine ! Vacances obligent ?

N'oubliez pas les révisions impératives pour lundi.

A vos ordi !!



$f_{11} < 0$ et $f_{11} \cdot f_{22} - (f_{12})^2 > 0$, car H définie négative par hypothèse, donc les déterminants des mineurs principaux doivent vérifier $(-1)^i \det(H_i) > 0 \quad \forall i$, donc le produit $f_{11} \cdot f_{22} > 0$ ce qui $\Rightarrow f_{22} > 0$

$\Rightarrow \frac{\partial x}{\partial \omega_1} < 0$ q.e.d.

lundi 28 avril.....

Optimisation avec contraintes :

On considère le problème (1)
$$\underset{x}{Max} f(x) \text{ t.q. } g(x)=a$$

Avec $f : \mathbb{R}^n \rightarrow \mathbb{R}$ et $g : \mathbb{R}^n \rightarrow \mathbb{R}^k$ fonctions continûment différentiables dans un voisinage d'une solution locale x^* .

Théor.1 : Si le rang de $\left[\frac{\partial g^i}{\partial x_j} \right]_{x^*} = k$ i.e. si les contraintes sont qualifiées, alors $\exists \lambda$ t.q. $\nabla f(x^*) = \lambda \nabla g(x^*)$

(application immédiate du théorème de la fonction implicite)

Remarques : $\nabla f(x^*)$ vecteur $1 \times n$, $\lambda = 1 \times k$ et $\nabla g(x^*)$ matrice $k \times n$
avec nombre de variables > nombre de contraintes

En présence d'une seule contrainte les deux gradients auront même direction, donc une des isoquantes (courbe de niveau) sera tangente à $g(x)=a$

Le théor.1 signifie que le gradient de la fonction objectif peut s'exprimer comme combinaison linéaire des gradients des contraintes.

Ce qui revient à dire que les mêmes conditions peuvent être obtenues en introduisant :

$L(x, \lambda, a) \equiv f(x) - \lambda \cdot (g(x) - a)$ et en écrivant le CNPO pour la recherche d'un point stationnaire de L .

Théor.2 : Soit B une matrice réelle $k \times n$ avec $\text{rang}(B) = k < n$ et A une matrice symétrique $n \times n$.

On construit H matrice symétrique constituée par $H \equiv \begin{pmatrix} 0 & B \\ B & A \end{pmatrix}$

Et on appelle H_i les mineurs de H obtenus en supprimant tout sauf les i premières lignes et les i premières colonnes.

Alors on peut démontrer que :

$$(1) \quad x'Ax > 0 \quad \forall x \neq 0 \text{ avec } Bx=0 \quad \text{ssi} \quad (-1)^k \det(H_{2k+i}) > 0 \text{ pour } i=1, \dots, (n-k)$$

$$(2) \quad x'Ax < 0 \quad \forall x \neq 0 \text{ avec } Bx=0 \quad \text{ssi} \quad (-1)^{k+i} \det(H_{2k+i}) > 0 \text{ pour } i=1, \dots, (n-k)$$

Théor.3 :

Avant tout commentaire et toute explication un exercice avec obligation d'utiliser les théorèmes 1&2:

Et maintenant lisez avec attention la solution de l'exercice suivant puis communiquez vos remarques !



$$\min (xy+xz+yz) \quad \text{t.q. } xyz=125$$

Solution A :

D'abord on utilise le (Théor1)

Ici $f(x,y,z)=xyz$ $g(x,y,z)=a$ c'est $xyz=125$ ce qui nous assure que $\Rightarrow x \neq 0, y \neq 0, z \neq 0$
(car une seule des variables égale à zéro impliquerait un produit nul)

Nous avons ici une seule contrainte : $k=1$ (une seule contrainte) et $g'_x=yz, g'_y=xz, g'_z=xy$

nous assure que la contrainte est qualifiée, car $\nabla(g) \neq (0,0,0)$ donc $\text{rang} \left[\frac{\partial g}{\partial x_j} \right] = k=1$

il est donc possible d'utiliser le Lagrangien :

$$L(x,y,z,\lambda)=xy+xz+yz - \lambda(xyz-125)$$

$$\text{CNPO : } \begin{cases} y + z = \lambda yz \\ x + z = \lambda xz \\ x + y = \lambda xy \\ xyz = 125 \end{cases} \Rightarrow \text{par simple quotient terme à terme} \quad \frac{y+z}{x+z} = \frac{y}{x} \quad \text{d'où } x^*=y^*=z^*$$

(on peut continuer le quotient t.à t. ou observer la symétrie....)

et $x=y=z \Rightarrow x^3=125 \Rightarrow x^*=5$

donc $P(5,5,5)$ est le point susceptible de minimiser la fonction objectif sous la contrainte donnée.

Et maintenant les (Théors. 3 & 2)

$$\text{CSSO : } A = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} \quad \text{car } \frac{\partial^2 L}{\partial x^2} = 0 \text{ et } \frac{\partial^2 L}{\partial x \partial y} = -1 \quad (\text{d'où les autres par symétrie})$$

D'autre part $\frac{\partial^2 L}{\partial \lambda^2} = 0$ et $\frac{\partial^2 L(5,5,5)}{\partial \lambda \partial x} = 25$ nous avons tous les éléments pour écrire :

$$H = \begin{bmatrix} 0 & 25 & 25 & 25 \\ 25 & 0 & -1 & -1 \\ 25 & -1 & 0 & -1 \\ 25 & -1 & -1 & 0 \end{bmatrix} \quad n=3, k=1 \text{ donc } (n-k) = 2 \text{ d'où } i=1,2$$

On détermine les deux déterminants de H_{2k+1} et de H_{2k+2} :

$$\det(H_{2k+1}) = \det(H_3) = \det \begin{pmatrix} 0 & 25 & 25 \\ 25 & 0 & -1 \\ 25 & -1 & 0 \end{pmatrix} < 0 \quad \text{et} \quad \det(H_4) = \det(H) < 0$$

théor. 2:

pour $i=1$ on a $(-1)^1 \det(H_{2k+1}) = (-1) \cdot (-) > 0$ et pour $i=2$ $(-1)^1 \det(H_{2k+2}) > 0$
donc A est définie positive et f atteint un minimum en (5,5,5) sous la contrainte donnée. (minf= 75)

Aussi

Solution B :

$xyz=125$ avec $x=y=z \neq 0$ (car... cf. solution précédente) d'où $z = \frac{125}{xy}$ et si on remplace dans la fonction objectif il nous faudra résoudre le programme sans contrainte suivant :

$$\min h(x,y) = \min \left(xy + x \cdot \frac{125}{xy} + y \cdot \frac{125}{xy} \right) = \min \left(xy + \frac{125}{y} + \frac{125}{x} \right)$$

$$\text{CNPO : } h'_x = 0 \text{ et } h'_y = 0 \text{ i.e. } y = -\frac{125}{x^2} \text{ et } x = -\frac{125}{y^2} \text{ donc } x^* = y^* = 5$$

$$\text{CSSO : } \frac{\partial^2 h}{\partial x^2} = \frac{\partial^2 h}{\partial y^2} = 2 \text{ et } \frac{\partial^2 h}{\partial x \partial y} = 1 \text{ d'où } H = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \text{ les déterminants des mineurs principaux}$$

Sont : $\det(2) = 2 > 0$ et $\det(H) = 3 > 0$, donc la forme quadratique des dérivées secondes est définie positive et $h(5,5) = 25 + \frac{125}{5} + \frac{125}{5} = 75$ est le minimum cherché.

A vous de décider !

Solution A ???



Solution B ???



Solution A et B ???



Autre solution ???

