Notes de cours d'Optimisation

Chapitre 1 : Généralités

1.1 Introduction, terminologie

On s'intéresse aux solutions éventuelles du programme

$$(P)\begin{cases} Max\, f(x) \\ x \in K \end{cases}$$

$f : \mathbb{R}^n \to \mathbb{R}$ est la *fonction objectif* du programme (P)

$K \subset \mathbb{R}^n$ est le *domaine* du programme (P)

C'est la forme générale d'un problème d'optimisation en dimension finie, l'ensemble K est quelconque donné à priori, de même que la fonction f .

L'étude d'un programme de minimisation se ramène à celle d'un programme de maximisation en remarquant que les minima de f sur K coïncident avec les maxima de la fonction -f sur K .

Définition

$x^* \in \mathbb{R}^n$ est une *solution optimale* de (P), ou un *maximum* de f sur K si

$$x^* \in K \text{ et } \forall x \in K,\ f(x) \le f(x^*)$$

L'ensemble (éventuellement vide) des solutions optimales de (P) est noté SO(P) .

Définition

On appelle *courbe de niveau* de f , tout ensemble de la forme

$$C_\alpha = \{x \in \mathbb{R}^n / f(x) = \alpha\}$$

où $\alpha \in \mathbb{R}$ est une valeur quelconque donnée (le niveau de f sur C_α)

Géométriquement, la résolution du programme (P) revient à déterminer la courbe de niveau maximale de la fonction objectif f qui rencontre l'ensemble K .

On appellera *extremum* de f sur K tout point de K qui est un maximum ou un minimum de f sur K .

Les questions qui nous intéressent sont les suivantes :

1) Existence des solutions
2) Unicité de la solution
3) Caractérisation des solutions, calcul explicite

4) Dépendance des solutions par rapport aux paramètres : statique comparative

On utilise la plupart du temps des outils de nature différente pour répondre à chacune de ces questions :

- le problème de l'existence se résout en général en exploitant les propriétés *topologiques* de l'ensemble K et de la fonction f.
- le problème de l'unicité n'a pas de solution générale satisfaisante sauf dans le cas particulier important de l'optimisation convexe .
- la caractérisation des solutions fait appel à une notion nouvelle : le cône normal, et au théorème de Kühn et Tücker.
- enfin l'étude de la dépendance des solutions par rapport aux paramètres utilise le théorème de l'enveloppe.

En ce qui concerne le calcul des solutions, on est amené à établir un certain nombre de conditions nécessaires d'optimalité :

Définition

(C) est une ***condition nécessaire*** *d'optimalité si toute solution du programme (P) vérifie la condition (C).*

L'application d'une CN fournira donc en général <u>trop de points</u>.

Parmi tous ces points il faudra déterminer la ou les solutions globales de (P), usuellement en comparant la valeur de la fonction objectif en ces points.

Exemples :

- $K =]a , b[\subset R$, f dérivable, alors la CN1 s'écrit

$$f'(x^*) = 0 .$$

- $K = [a , b] \subset R$, f dérivable, alors la CN1 s'écrit

$$\begin{array}{ll} f'(x^*) = 0 & \text{si } x^* \in]a , b[\\ f'(a) \leq 0 & \text{si } x^* = a \\ f'(b) \geq 0 & \text{si } x^* = b . \end{array}$$

Démonstration géométrique.

1.2. Solutions intérieures

Définitions

$x \in K$ est un point *intérieur* de K si

$$\exists \eta : B(x,\eta) \subset K$$

Une partie O de $\mathbb{R}^n$ est un *ouvert* si tout point de O est intérieur à O .

L'ensemble des points intérieurs de K, noté $int(K)$, est appelé intérieur de K. C'est le plus grand ouvert inclus dans K.

Exemples : l'intérieur d'un intervalle fermé $[a,b]$ de $\mathbb{R}$ est l'intervalle ouvert $]a,b[$, l'intérieur d'une droite dans $\mathbb{R}^2$ est vide

Définition

$x^* \in K$ est une *solution intérieure* de (P) si

$$x^* \in SO(P) \text{ et } x^* \in int(K)$$

On suppose maintenant f différentiable.

Théorème

Si x^* est une solution <u>intérieure</u> de (P) alors

$$\text{(CN1)} \quad \nabla f(x^*) = 0 .$$

Preuve : Soit η tel que $B(x^*,\eta) \subset K$.

$\forall\, i = 1 , \dots , n$, on considère l'application partielle

$$\varphi_i : t \to f(x_i^* + t , x_{-i}^*) ,$$

on a $\forall\, t \in]-\eta , \eta[$, $\varphi_i(t) \leq \varphi_i(0)$, d'où $\varphi'_i(0) = 0$ //

Remarques :

- En changeant f in $-f$ et/ou en restreignant le domaine de (P), on voit que la condition ci-dessus est en fait une condition nécessaire d'extremalité <u>locale</u> en un point intérieur. L'ensemble des points vérifiant cette condition est appelé ensemble des <u>points critiques</u> de f. La condition d'annulation du gradient n'est pas une condition suffisante d'optimalité sauf dans le cas d'un programme convexe.
- La CN1 d'annulation du gradient n'est plus valable en un point situé sur la **frontière** de K comme on l'a déjà vu dans l'exemple $K = [a , b]$. La dérivation d'une CN1 générale d'optimalité (locale) valable en un point intérieur ou non intérieur forme l'essentiel de ce cours. Noter cependant qu'on peut parfois ramener l'étude d'un programme sous contraintes à celle d'un programme sans contraintes.

1.3 Théorèmes d'existence

Rappels de topologie :

Rappelons qu'une partie de R^n est *compacte* si et seulement si elle est *fermée* et *bornée* et qu'un ensemble est dit fermé si et seulement si il contient tous ses points adhérents (c'est à dire les limites de ses suites convergentes).

Rappelons aussi qu'un compact de R^n est caractérisé par la propriété que toute suite de points de l'ensemble admet une valeur d'adhérence dans l'ensemble. Intuitivement on comprend qu'un fermé borné vérifie une telle propriété : une suite de points de l'ensemble ne peut pas partir "à l'infini" puisque l'ensemble est borné, elle "doit donc"[1] s'accumuler quelque part dans R^n et une valeur d'adhérence de la suite est nécessairement dans l'ensemble puisque l'ensemble est fermé.

Théorème de Weierstrass et ses extensions :

Le prototype du théorème d'existence est le théorème de Weierstrass, il énonce une condition suffisante d'existence de solutions au programme (p).

Théorème

Si l'ensemble K est compact (non vide) et si f est continue sur K, alors le programme (P) admet (au moins) une solution.

Preuve : Soit α le Sup de la fonction f sur l'ensemble K (éventuellement égal à $+\infty$), par définition du Sup il existe une suite (x_n) de points de K telle que $(f(x_n))$ converge vers α . K étant compact, la suite (x_n) admet un point d'accumulation x^* dans K , il en résulte par continuité que $f(x^*)$ est égal à α , c'est dire que $\alpha < +\infty$ et x^* est une solution optimale du programme (P) //

Lorsque le domaine K n'est pas compact on cherche à se ramener à un domaine compact en montrant que f n'approche sa borne supérieure ni à l'infini, ni sur les bords de l'ensemble K .

Par exemple quand K est un fermé non vide, non nécessairement borné (comme R^n tout entier) on a le théorème d'existence suivant :

Théorème

1[1] Cette intuition n'est valable qu'en dimension finie : le nombre de sommets d'un cube de R^n augmente indéfiniment quand n augmente, ainsi une suite bornée peut ne pas posséder de point d'accumulation en dimension infinie.

Si K est <u>fermé</u> non vide, f continue sur K , et si de plus on a

$$\lim_{\|x\|\to+\infty} f(x) = -\infty$$

alors le programme (P) admet une solution

Preuve : Soit x_0 un point quelconque de K . Puisque f tend vers - ∞ à l'infini, il existe un réel A positif tel que $\|x\|>A \Rightarrow f(x)<f(x_0)$. Les solutions éventuelles de f sur K sont donc à chercher parmi les points de K de norme inférieure ou égale à A , c'est à dire les points de K∩B'(0, A) . Mais K∩B'(0, A) est compact et f est continue, donc (P) admet des solutions d'après (2.2.1) //

On peut multiplier les théorèmes de ce type, il découle de la démonstration ci-dessus qu'on a par exemple :

Théorème
Si K est <u>fermé</u>, f continue sur K , et si il existe $x_0 \in K$ et R > 0 tels que

$$\|x\|>R \Rightarrow f(x)<f(x_0) ,$$

alors le programme (P) admet une solution.

Rappelons enfin que la situation est complètement différente dans le cas particulier important de l'optimisation linéaire : il suffit alors que le domaine soit non vide et que la fonction objectif soit majorée sur le domaine, i.e. $\text{Sup}_K f < +\infty$.
Ceci est bien sûr très spécifique au cas de l'optimisation linéaire, le domaine est alors un polyèdre convexe fermé tandis que la courbe de niveau correspondant au Sup de f sur K est un hyperplan, on conçoit que ces deux ensembles ne peuvent être tangents en l'infini mais doivent se rencontrer.

Quand on sort du cadre de l'optimisation linéaire, ceci n'est bien sûr plus vrai, par exemple le programme
(P) : Max x_2 s.c. $x_2 \leq -1/x_1$ et $x_1 \geq 0$ n'a pas de solutions bien que son domaine soit fermé et son Sup fini.

Chapitre 2 : Condition géométrique d'optimalité

3.1 Courbes de niveau et gradients

Proposition
Une fonction augmente localement dans la direction de son gradient

Preuve : Par définition de la différentiabilité de f :

$$f(x+t\nabla f(x))=f(x)+<\nabla f(x), t\nabla f(x)>+\|t\nabla f(x)\|\epsilon(t\nabla f(x))$$

où ε est une fonction qui tend vers zéro en zéro, donc

$$\frac{f(x+t\nabla f(x))-f(x)}{t}=\|\nabla f(x)\|^2+\|\nabla f(x)\|\epsilon(t\nabla f(x))$$

Le terme de droite tend vers $\|\nabla f(x)\|^2$ quand t tend vers 0 ,
donc $f(x+t\nabla f(x))-f(x)$ est strictement positif si $\nabla f(x)$ est non nul et si t est strictement positif et assez petit. //

Proposition
Les gradients sont perpendiculaires aux courbes de niveau

Preuve : Par définition, un vecteur tangent u à une courbe de niveau C d'une fonction g en un point x de C est la dérivée d'un chemin tracé dans C au point x :
$u=\phi'(0)$ avec $\phi:\mathbb{R}\rightarrow C$ et $\phi(0)=x$ (comparer avec, plus loin, la définition du cône tangent)
Puisque g est constante sur C , la fonction $g o\phi$ est constante,
donc $(g o\phi)'(0)=0$. Or $(g o\phi)'(0)=<\nabla g(\phi(0)), \phi'(0)>=<\nabla g(x), u>$ //

3.2 Cône normal et CN1 d'optimalité

La notion de cône tangent généralise les notions usuelles de tangente à une courbe et de plan tangent à une surface :

Définition
Le *cône tangent* à K au point x est l'ensemble des dérivées des chemins rentrants dans K en x :

$$T_x K=\{\phi'(0)/\phi:\mathbb{R}_+\rightarrow K \text{ et } \phi(0)=x\}$$

Proposition
Soit $K\subset\mathbb{R}^n$ et $x\in\text{int}(K)$, alors $T_x K=\mathbb{R}^n$

Preuve :
Soit $u\in\mathbb{R}^n$ et soit $\eta>0$ tel que $B(x,\eta)\subset K$.
Considérons le chemin $\phi:\mathbb{R}_+\rightarrow\mathbb{R}^n$ défini par
$\forall t\in[0,\eta/2], \phi(t)=x+tu$ et $\forall t\geq\eta/2, \phi(t)=x$
Le chemin ϕ est bien à valeurs dans K et vaut x en 0 , donc $\phi'(0)=u \in T_x K$ //

La notion de cône normal généralise les notions de normale à une courbe ou à une surface :

Définition

Le *cône normal* à K au point x est le polaire négatif du cône tangent :

$$N_x K = \{ p \in \mathbb{R}^n / \forall u \in T_x K , < p , u > \leq 0 \}$$

Proposition

Soit $K \subset \mathbb{R}^n$ et $x \in \text{int}(K)$, alors $N_x K = \{0\}$

Preuve : immédiat puisque $T_x K = \mathbb{R}^n$ //

Proposition

Soit $x^* \in K \subset \mathbb{R}^n$ et f , g deux fonctions $\mathbb{R}^n \to \mathbb{R}$ différentiables en x^*

a) Si $K = \{ x \in \mathbb{R}^n / f(x) = 0 \}$ et $\nabla f(x^*) \neq 0$, alors $N_{x^*} K = \mathbb{R} \nabla f(x^*)$

b) Si $K = \{ x \in \mathbb{R}^n / f(x) \leq 0 \}, f(x^*) = 0$ et $\nabla f(x^*) \neq 0$, alors $N_{x^*} K = \mathbb{R}_+ \nabla f(x^*)$

c) Si $K = \{ x \in \mathbb{R}^n / f(x) \leq 0$ et $g(x) \leq 0 \}, f(x^*) = g(x^*) = 0$ et si les vecteurs $\nabla f(x^*), \nabla g(x^*)$ sont libres, alors $N_{x^*} K = \mathbb{R}_+ \nabla f(x^*) + \mathbb{R}_+ \nabla g(x^*)$

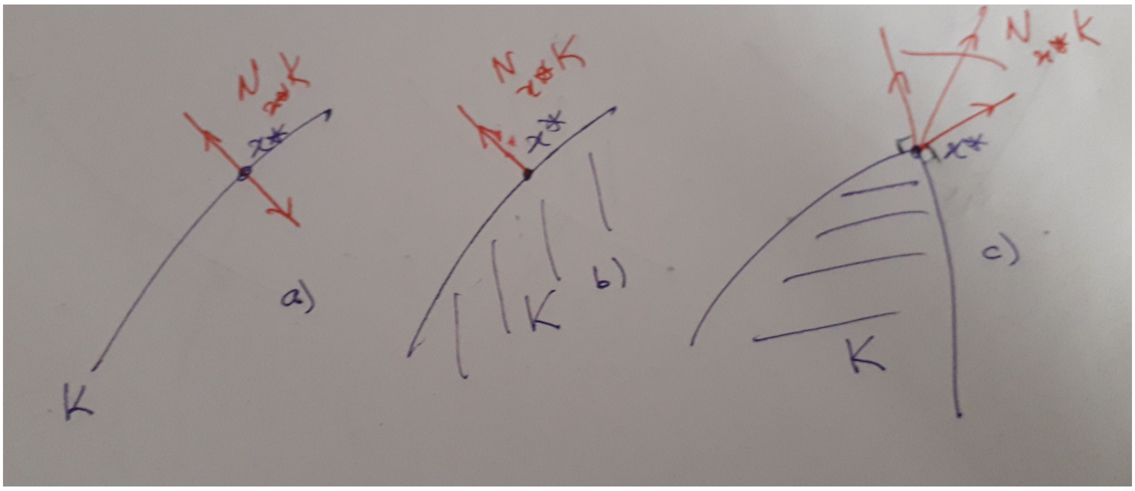
Preuve : La preuve fait appel au théorème des fonctions implicites et est omise //

Dans la pratique, lorsque n = 2 , on représente sur une figure le translaté en x* du cône normal. On obtient ainsi dans les différents cas envisagés ci-dessus :

a) la normale à la courbe K en x*

b) la normale sortante à K en x*

c) le cône de sommet x* engendré par les normales sortantes à K en x*



Théorème

Si x* est solution de $(P)\begin{cases} Max\, f(x) \\ x\in K \end{cases}$, alors on doit avoir $\nabla f(x^*)\in N_{x^*}K$

Preuve : Soit $u=\phi'(0)$, avec $\phi:\mathbb{R}_+\to K$ et $\phi(0)=x^*$, un élément de $T_{x^*}K$.
On doit montrer que $<\nabla f(x^*),u>\leq 0$.
Puisque f atteint son maximum en x* sur K , la fonction $f\,o\,\phi$ doit atteindre son maximum en t* = 0 sur $\mathbb{R}_+$, donc on doit avoir la CN1 $(f\,o\,\phi)'(0)\leq 0$.
Or $(f\,o\,\phi)'(0)=<\nabla f(\phi(0)),\phi'(0)>=<\nabla f(x^*),u>$ //

Remarques :
- Pour un programme de minimisation la CN1 est bien sûr que l'opposé du gradient appartient au cône normal.
- Il faut bien noter que la condition du premier ordre ne fournit qu'une condition <u>nécessaire</u> d'optimalité. Autrement dit, l'application de ce critère fournira en général trop de points parmi lesquels il faudra déterminer la ou les solutions du programme (P) , usuellement en comparant les valeurs de la fonction objectif f en ces points.

Dans le cadre de l'optimisation convexe cependant, la condition du premier ordre est également suffisante. (voir Annexe)

Chapitre 3. Théorème de Kühn et Tücker.

On s'intéresse toujours au programme $(P)\begin{cases} Max\, f(x) \\ x\in K \end{cases}$

Où le domaine $K\subset\mathbb{R}^n$ est maintenant défini explicitement par un nombre fini d'égalités et d'inégalités larges et strictes (les contraintes du programme) :

$$x\in K \iff \begin{cases} \forall i\in I,\ f_i(x) = 0 \\ \forall j\in J,\ g_j(x) \leq 0 \\ \forall l\in L,\ h_l(x) < 0 \end{cases}$$

I , J , L sont des ensembles finis (éventuellement vides).
Notons p = Card(I) , q = Card(J) .

Définition
Le *Lagrangien* de (P) est la fonction

$$L:\ \mathbb{R}^n X \mathbb{R}^p X \mathbb{R}^q \to \mathbb{R}\ ,\ (x,\lambda,\mu) \to f(x) - \sum_{i=1}^{p}\lambda_i f_i(x) - \sum_{j=1}^{q}\mu_j g_j(x)$$

Notons que les contraintes d'inégalité stricte n'interviennent pas dans le Lagrangien.

Si $x \in K$, on note J(x) l'ensemble des contraintes d'inégalité large qui sont saturées au point x :

$$J(x) = \{ j \in J / g_j(x) = 0 \}$$

Le théorème de Kühn et Tücker sera valable sous certaines conditions, dites conditions de qualification.

Définition

On dit que le point x dans K est *qualifié* si l'une des deux conditions suivantes est satisfaite :

1. Toutes les contraintes actives au point x sont affines :

   les fonctions $((f_i)_{i \in I}, (g_j)_{j \in J(x)})$ sont affines

2. Les gradients des contraintes actives au point x forment une famille libre :

   $((\nabla f_i(x))_{i \in I}, (\nabla g_j(x)_{j \in J(x)}))$ est libre (condition de Lagrange)

Théorème (Kühn et Tücker)

Si x* est un point qualifié et est solution du programme $(P) \begin{cases} Max\, f(x) \\ x \in K \end{cases}$, alors on a les conditions de Kühn et Tücker :

$$\exists (\lambda^*, \mu^*) \in \mathbb{R}^p X \mathbb{R}^q / \quad (KT) \quad \begin{cases} \dfrac{\partial L}{\partial x}(x^*, \lambda^*, \mu^*) = 0 \\ \forall j \in J, \quad \mu_j^* g_j(x^*) = 0 \\ \forall j \in J, \quad \mu_j^* \geq 0 \\ x^* \in K \end{cases}$$

Des réels $\lambda_i^*, i \in I$, et $\mu_j^*, j \in J$, vérifiant les conditions ci-dessus sont appelés *mutiplicateurs de Kühn et Tücker* associés au point x* .

Les conditions $\mu_j^* g_j(x^*) = 0$ sont appelées *relations d'exclusion*.

Elles expriment le fait que les contraintes non saturées en x* ne jouent aucun rôle dans les conditions du premier ordre puisque leurs multiplicateurs associés sont nuls, d'un autre côté si un multiplicateur μ_j^* est strictement positif, la contrainte associée $g_j(x) \leq 0$ doit nécessairement être saturée (active) en x* .

La méthode de Kühn et Tücker conduit donc à rechercher n+p+q inconnues : les n composantes de x^* , les p multiplicateurs associés aux contraintes d'égalité et les q multiplicateurs associés aux contraintes d'inégalité large.
Le système (KT) comporte (en plus de conditions d'inégalité) n+p+q équations : les n dérivées partielles du Lagrangien par rapport aux coordonnées de x , les q relations d'exclusion et les p contraintes d'égalité du domaine.

<u>Attention</u> : si un multiplicateur associé à une contrainte d'inégalité est nul, la contrainte associée peut quand même être active en x^* !
Exemple : dans R , soit le programme :

$$\text{Max } (-x^2) \text{ sous la contrainte } g(x) = x \leq 0 ,$$

$x^* = 0$ est évidemment la solution de ce programme et en ce point la contrainte est qualifiée puisque $\nabla g(0) = g'(0) = 1 \neq 0$ et est donc un vecteur libre dans $\mathbb{R}$ (ou, plus simplement, la contrainte est affine donc qualifiée)
Le Lagrangien du programme est $L(x , \mu) = -x^2 - \mu x$ et l'application des conditions de Kühn et Tücker donne $\exists \ \mu^* \geq 0 \ / \ \frac{\partial L}{\partial x}(0 , \mu^*) = -\mu^* = 0$

Chapitre 4. Le théorème de l'enveloppe.

On s'intéresse à des programmes paramétrés :

$$(P(a))\begin{cases} Max_x \ f(x,a) \\ g(x,a) = 0 \end{cases}$$

Les fonctions f et g sont supposées de classe C^2 sur un ouvert $U \text{X} I$ de $\mathbb{R}^n \text{X} \mathbb{R}$

(le paramètre a est supposé unidimensionnel sans perte de généralité : on peut toujours fixer les autres paramètres éventuels, par contre g peut être à valeurs vectorielles)

Supposons que (P(a)) admette une solution optimale $x^*(a)$ pour tout a dans I et que la fonction ainsi définie $x^*: I \rightarrow \mathbb{R}$ soit caractérisée par les conditions de Kühn et Tücker et soit de classe C^1
La fonction valeur du programme $V: I \rightarrow \mathbb{R}, a \rightarrow V(a) = f(x^*(a), a)$ est alors dérivable et sa dérivée est donnée par le résultat suivant :

Théorème

$$V'(a)=\frac{\partial L}{\partial a}(x^*(a)\ ,\ \lambda^*(a)\ ,\ a)$$

Où $L(x\ ,\ \lambda\ ,\ a)=f(x,a)-\lambda g(x,a)$ est le Lagrangien du programme (P(a)) et $\lambda^*(a)$ est un multiplicateur de Kühn et Tücker associé à $x^*(a)$

Preuve :

On a immédiatement $V'(a)=<\frac{\partial f}{\partial x}(x^*(a),a),x^{*\prime}(a)> + \frac{\partial f}{\partial a}(x^*(a),a)$

Par ailleurs KT nous dit que

$$\frac{\partial L}{\partial x}(x^*(a),\lambda^*(a),a)=0=\frac{\partial f}{\partial x}(x^*(a),\lambda^*(a),a) - \lambda^*(a)\frac{\partial g}{\partial x}(x^*(a),\lambda^*(a),a)$$

et, en dérivant la contrainte $g(x^*(a),a) = 0$, on obtient

$$<\frac{\partial g}{\partial x}(x^*(a),a),x^{*\prime}(a)> + \frac{\partial g}{\partial a}(x^*(a),a) = 0$$

D'où, en reportant dans la première équation

$$V'(a) = -\lambda^*(a)\frac{\partial g}{\partial a}(x^*(a),a) + \frac{\partial f}{\partial a}(x^*(a),a) = \frac{\partial L}{\partial a}(x^*(a),\lambda^*(a),a) \quad //$$

Un cas particulier important se présente quand le domaine ne dépend pas du paramètre :

Proposition

Si la fonction valeur V du programme $(P(a))\begin{cases} Max_x\ f(x,a) \\ x \in K \end{cases}$ est dérivable, alors pour tout a , $V'(a)=\frac{\partial f}{\partial a}(x^*(a)\ ,\ a)$

Preuve :

La fonction $h: I \rightarrow \mathbb{R}, b \rightarrow f(x^*(a),b)-V(b)$ est négative ou nulle sur tout I et est nulle au point a , elle atteint donc son maximum au point a , d'où

$$h'(a)=\frac{\partial f}{\partial a}(x^*(a)\ ,\ a)-V'(a)=0 \quad //$$

Chapitre 5. Programmation dynamique

De nombreux programmes d'optimisation ont une structure dynamique. On s'intéresse dans ce chapitre à des programmes stationnaires en horizon infini du type

$$(P(k_{-1}))\begin{cases} Max \sum_{t=0}^{+\infty} \beta^t u(c_t) \\ \forall t \in \mathbb{N} \ , \ k_t = f(k_{t-1}, c_t) \end{cases}$$

avec k_{-1} donné

La fonction objectif est stationnaire de même que la contrainte à tout date, appelée *équation de transition*.
L'équation de transition $k_t = f(k_{t-1}, c_t)$ peut par exemple être $k_t = k_{t-1} - c_t$ pour un modèle simple de consommation intertemporelle où c_t et k_t sont respectivement la consommation et l'épargne à date t.

Des contraintes éventuelles d'inégalité stricte portant sur les *variables de contrôle* c_t et les *variables d'état* k_t sont sous-entendues.

On peut s'attaquer à la résolution de ce programme selon trois approches différentes

Approche 1 : En utilisant les conditions de Kühn et Tücker
Sous des conditions de régularité à préciser, on peut utiliser les conditions de Kühn et Tücker :
Le Lagrangien est $L(c \ , \ k \ , \ \lambda) \ = \ \sum_{t=0}^{+\infty} \beta^t \{u(c_t) \ + \ \lambda_t(-k_t + f(k_{t-1}, c_t))\}$
(en ayant multiplié les contraintes par β^t pour simplifier l 'écriture)
et les conditions de Kühn et Tücker s'écrivent $\forall t \in \mathbb{N} \ , \ \frac{\partial L}{\partial c_t} \ = \ \frac{\partial L}{\partial k_t} \ = \ 0$

C'est à dire

$$\forall t \in \mathbb{N} \ , \ \begin{cases} u'(c_t) \ + \ \lambda_t \frac{\partial f}{\partial c}(k_{t-1}, c_t) & = \ 0 \\ -\lambda_t \ + \ \beta \lambda_{t+1} \frac{\partial f}{\partial k}(k_t, c_{t+1}) & = \ 0 \end{cases}$$

Approche 2 : En éliminant les variables de contrôle
Dans le cas courant où l'équation de transition $k_t = f(k_{t-1}, c_t)$ peut être résolue en c_t sous la forme $c_t = g(k_{t-1}, k_t)$, on peut éliminer les variables de contrôle pour aboutir au programme

$$(Q(k_{-1}))\left\{ \ Max \sum_{t=0}^{+\infty} \beta^t u(g(k_{t-1}, k_t)) \right.$$

Dans ce programme sans contraintes, autres que d'inégalité stricte, on peut écrire la condition du premier ordre en un point intérieur :

$$\forall t,\ u'(g(k_{t-1},k_t))\frac{\partial g}{\partial x_2}(g(k_{t-1},k_t)) \ + \ \beta u'(g(k_t,k_{t+1}))\frac{\partial g}{\partial x_1}(g(k_t,k_{t+1})) \ = \ 0$$

On obtient ainsi une équation aux différences du second ordre sur les variables d'état.

Approche 3 : via l'équation de Bellmann

On peut enfin, en exploitant la stationnarité de la fonction objectif et des équations de transition, adopter une approche plus dynamique :

Réécrivons le programme sous la forme

$$(P(k_{-1}))\begin{cases} Max \ \ u(c_0)+\beta\sum_{t=0}^{+\infty}\beta^t u(c_{t+1}) \\ k_0=f(k_{-1},c_0) \\ \forall t\in\mathbb{N}, k_{t+1}=f(k_t,c_{t+1}) \end{cases}$$

Si $(k_t^*,c_t^*)_{t\in\mathbb{N}}$ est une solution optimale de $(P(k_{-1}))$ il apparaît que $(k_{t+1}^*,c_{t+1}^*)_{t\in\mathbb{N}}$ est une solution optimale de $(P(k_0^*))$ avec $k_0^*=f(k_{-1},c_0^*)$

De plus, en notant V la fonction valeur du programme (P) , fonction de l'état initial, on voit que

Théorème (équation de Bellmann)

$$V(k)=Max_c\{u(c)+\beta V(f(k,c))\}$$

avec, pour tout t, $c=c_t^*$ solution de ce programme pour $k=k_{t-1}^*$

Montrons que la CN1 et le théorème de l'enveloppe appliqués à l'équation de Bellmann donnent les mêmes équations que les conditions de Kühn et Tücker du programme initial $(P(k_{-1}))$:

$$\forall t\in\mathbb{N}\ ,\ \begin{cases} u'(c_t) \ +\beta V'(k_t)\frac{\partial f}{\partial c}(k_{t-1},c_t) = 0 \\ V'(k_{t-1}) = \beta V'(k_t)\frac{\partial f}{\partial k}(k_{t-1},c_t) \end{cases}$$

Et on retrouve le système vu précédemment en posant $\beta V'(k_t)=\lambda_t$

Annexe : Complément sur les programmes convexes

A1. Ensembles convexes

Définition

$K\subset\mathbb{R}^n$ est convexe si $\forall\ x,y\ \in\ K, \forall\ \lambda\ \in\ [0,1],\ \lambda x+(1-\lambda)y\ \in\ K$

Autrement dit, pour tous points x , y de K , le segment [x , y] est tout entier dans K
Exemples : boules, sous-espaces affines

Propriété
L'intersection de convexes est convexe

On verra plus loin (voir Proposition 2.5) un moyen important pour reconnaître des ensembles convexes.

A2. Fonctions concaves
Définition
Soit C un convexe de $\mathbb{R}^n$. $f: C \to \mathbb{R}$ est concave si
$\forall\ x, y\ \in C, \forall\ \lambda \in [0,1],\ \ f(\lambda x + (1-\lambda)y)\ \geq\ \lambda f(x) + (1-\lambda) f(y)$

Autrement dit : "l'image de la moyenne est plus grande que la moyenne des images".
Géométriquement : le graphe de f est au dessus de ses cordes (les segments qui lient les points du graphe)
Exemples : toute fonction affine, $t \to -t^2$, $t \to \ln(t)$, $t \to -e^t$

Propriétés
1) Soit $A: \mathbb{R}^p \to \mathbb{R}^n$ une application affine et $f: C \to \mathbb{R}$ une fonction concave, alors foA est concave.
2) Soit $f_1, \ldots, f_p$ des fonctions concaves et $\alpha_1, \ldots, \alpha_p$ des réels positifs ou nuls, alors $\sum_{i=1}^{p} \alpha_i f_i$ est concave.

Exemple : $(x_1, x_2, x_3) \to 2\ln(x_1 + x_2) - 3(x_2 - x_3)^2$ est concave

Pour les fonctions d'une seule variable on a le résultat bien connu suivant :
Proposition
Une fonction définie sur un intervalle de $\mathbb{R}$ et deux fois dérivable est concave ssi sa dérivée seconde est négative ou nulle en tout point

Pour les fonctions de plusieurs variables et deux fois différentiables, c'est plus compliqué :
Proposition
Soit C un convexe de $\mathbb{R}^n$. $f: C \to \mathbb{R}$ est concave ssi sa matrice Hessienne

$$D^2 f(x) = \left(\frac{\partial^2 f_i}{\partial x_j^2}(x)\right)_{1 \le i \le n, 1 \le j \le n}$$

a toutes ses valeurs propres négatives ou nulles en tout point x de C

Proposition

Soit C un convexe de $\mathbb{R}^n$ et $g_1, \ldots, g_p$ des fonctions concaves $C \to \mathbb{R}$
alors l'ensemble $K = \{x \in C \;/\; \forall\, j,\; g_j(x) \ge 0\}$ est convexe

Remarque : certaines inégalités peuvent être strictes , le résultat reste vrai.

A3. Programmes convexes

Définition

On dit que le programme

$$(P) \begin{cases} Max\, f(x) \\ x \in K \end{cases}$$

est *convexe* si le domaine K est convexe et la fonction objectif f est concave.

Remarque : noter bien qu'il s'agit d'un programme de maximisation

Proposition

L'ensemble des solutions d'un programme convexe est convexe.

Proposition

Pour un programme convexe, la CN1 géométrique est une condition nécessaire et suffisante d'optimalité :

$$x^* \in SO(P) \iff \nabla f(x^*) \in N_{x^*} K$$

Proposition

Pour un programme convexe et qualifié, le système des conditions de Kühn et Tücker caractérise l'ensemble des solutions optimales

Remarque : ne pas oublier dans le système de K.T. la condition de réalisabilité
$x^* \in K$