IFT1575 Modèles de recherche opérationnelle (RO)

7. Programmation non linéaire

Fonctions convexes et concaves

- Soit x' et x'' deux points dans R^n
- Le **segment de droite** joignant ces deux points est l'ensemble des points

  $x = x' + \lambda(x''-x') = \lambda x'' + (1- \lambda)x'$, où $\lambda \in [0,1]$
- Une fonction f est **convexe** si tout point sur le segment de droite joignant $f(x')$ et $f(x'')$ se trouve au dessus du graphe de f :

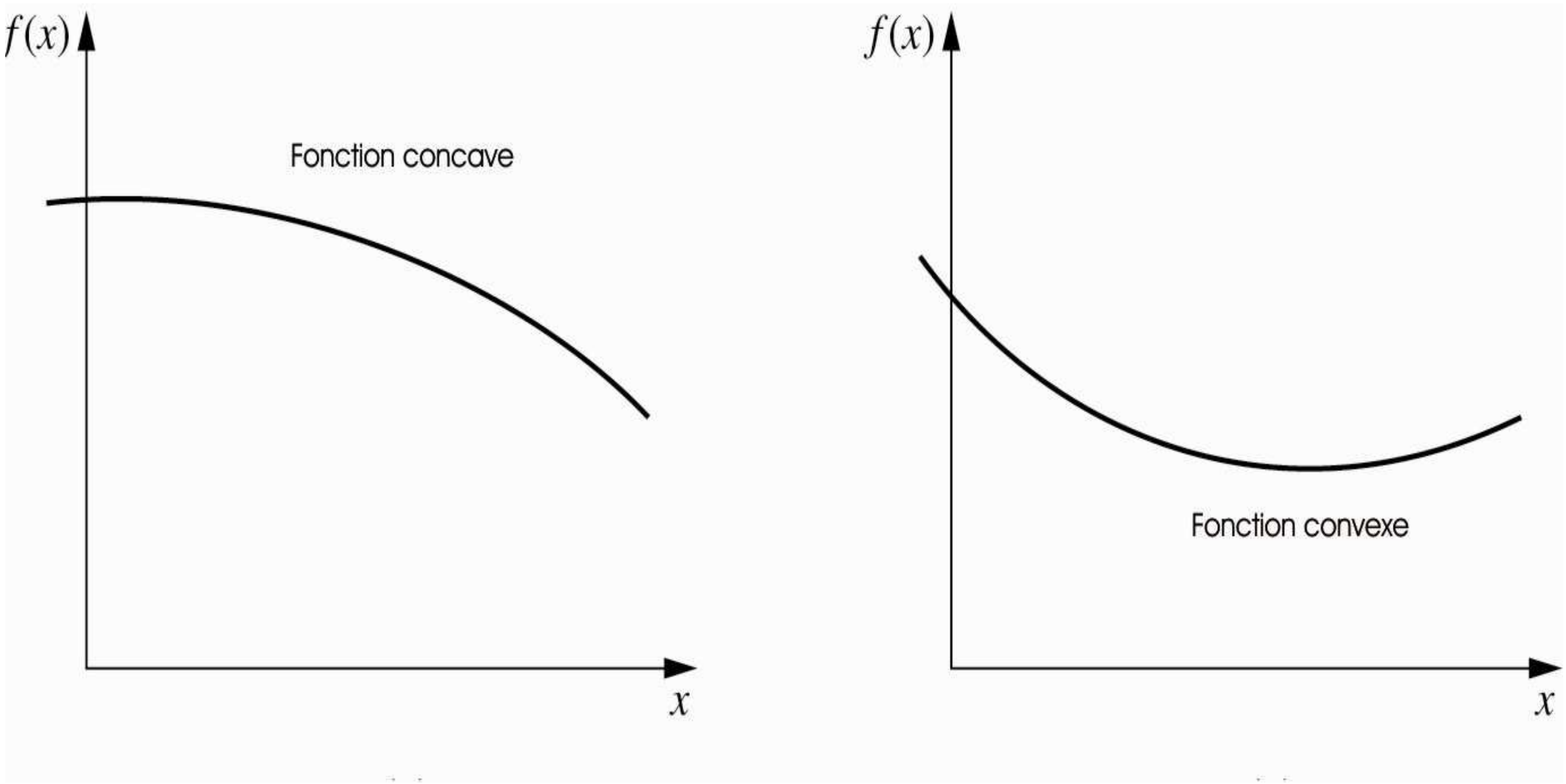
  $f(\lambda x'' + (1- \lambda)x') \leq \lambda f(x'') + (1- \lambda)f(x')$, pour toute paire de points x', x''

Fonctions convexes et concaves

- Si cette inégalité est stricte, la fonction f est dite **strictement convexe**
- Si on remplace $\leq$ par $\geq$ dans cette inégalité, la fonction f est dite **concave** :

  $f(\lambda x'' + (1- \lambda)x') \geq \lambda f(x'') + (1- \lambda)f(x')$, pour toute paire de points x', x''
- Si cette dernière inégalité est stricte, la fonction f est dite **strictement concave**

Fonctions convexes et concaves



Tests de convexité et de concavité

- Supposons que f est une fonction deux fois dérivable d'une seule variable
- f est convexe ssi $\frac{d^2 f(x)}{d^2 x} \geq 0, \forall x$
- f est strictement convexe ssi $\frac{d^2 f(x)}{d^2 x} > 0, \forall x$
- f est concave ssi $\frac{d^2 f(x)}{d^2 x} \leq 0, \forall x$
- f est strictement concave ssi $\frac{d^2 f(x)}{d^2 x} < 0, \forall x$

Tests de convexité et de concavité

- Supposons que f est une fonction deux fois dérivable de deux variables
- Dans ce cas, on utilise le tableau suivant :

	Convexe	Strict. convexe	Concave	Strict. concave
$\dfrac{\partial^2 f(x_1,x_2)}{\partial x_1^2}$	≥ 0	> 0	≤ 0	< 0
$\dfrac{\partial^2 f(x_1,x_2)}{\partial x_2^2}$	≥ 0	> 0	≤ 0	< 0
$\dfrac{\partial^2 f(x_1,x_2)}{\partial x_1^2} \cdot \dfrac{\partial^2 f(x_1,x_2)}{\partial x_2^2} - \left[\dfrac{\partial^2 f(x_1,x_2)}{\partial x_1 x_2}\right]^2$	≥ 0	> 0	≥ 0	> 0

Tests de convexité et de concavité

- Ces tests se généralisent à des fonctions de plus de deux variables en utilisant le Hessien (la matrice des dérivées secondes)
- Deux résultats importants pour tester la convexité et la concavité de fonctions quelconques :
  - Une fonction qui s'exprime comme la somme de fonctions convexes (concaves) est convexe (concave)
  - L'opposé d'une fonction concave est convexe et vice-versa
- Voir appendice 2, H&L, et TP 9 pour des exemples

Ensembles convexes

- Un ensemble est convexe si, pour toute paire de points de l'ensemble, le segment de droite joignant ces deux points est contenu dans l'ensemble
- On peut démontrer qu'un ensemble est convexe grâce aux propriétés suivantes :
  - Si f est convexe, alors $\{x \mid f(x) \leq b\}$ est convexe
  - Si f est concave, alors $\{x \mid f(x) \geq b\}$ est convexe
  - L'intersection d'ensembles convexes est convexe

Programmation non linéaire

- Un modèle de programmation non linéaire prend la forme suivante :

$$\max f(x_1, x_2, \ldots, x_n)$$

$$g_i(x_1, x_2, \ldots, x_n) \leq b_i, \quad i = 1, 2, \ldots, m$$

- On suppose qu'au moins une des fonctions *f* et g_i est non linéaire
- En général, on suppose aussi que toutes ces fonctions sont deux fois dérivables

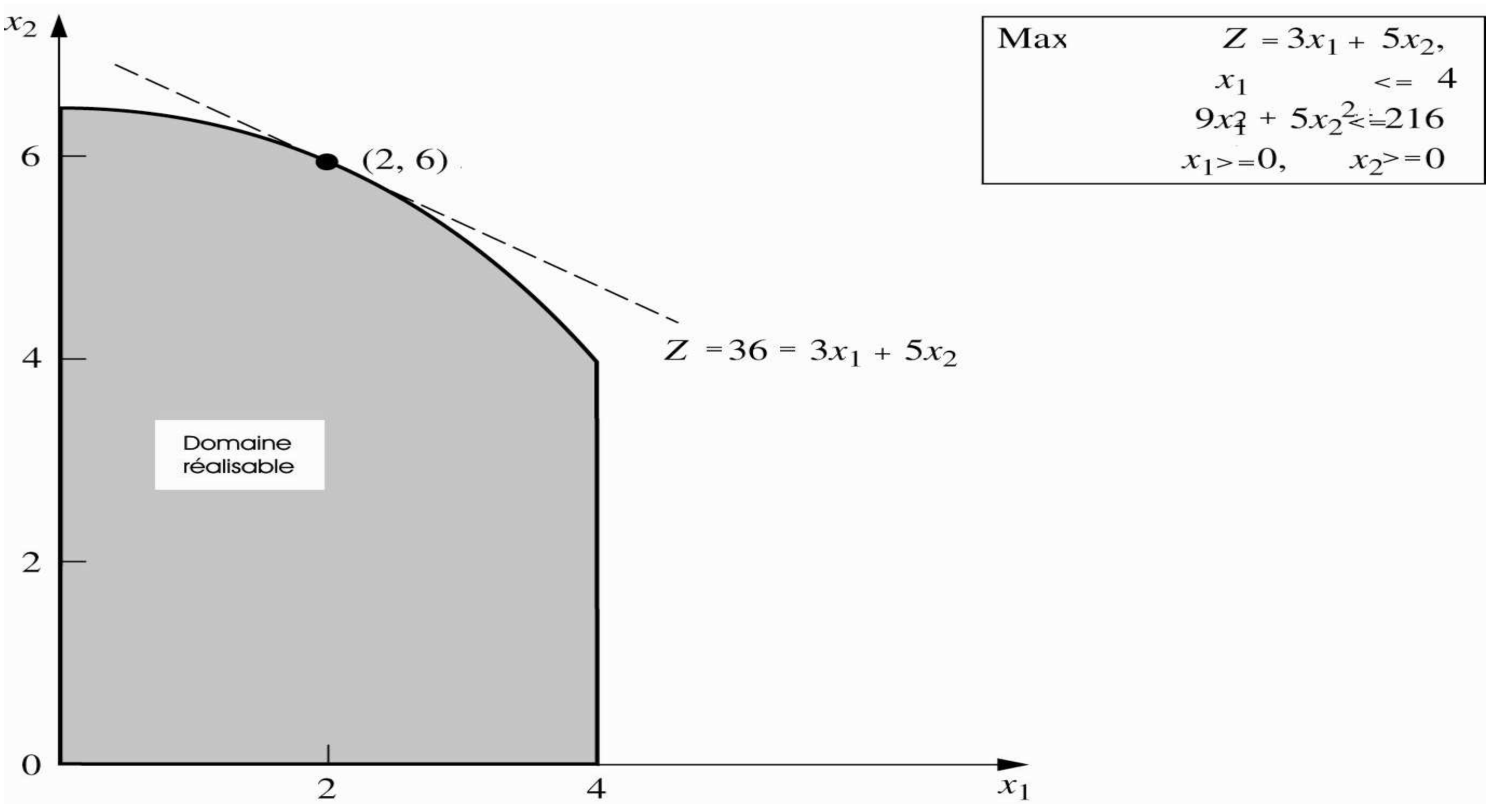
Exemple 1

- On modifie l'exemple *Wyndor Glass* (sect. 2) en remplaçant certaines contraintes par une contrainte non linéaire

$$\max Z = 3x_1 + 5x_2$$
$$x_1 \leq 4$$
$$9x_1^2 + 5x_2^2 \leq 216$$
$$x_1 \geq 0, x_2 \geq 0$$

- Voyons comment on peut représenter ce problème graphiquement

Exemple 1 (suite)



Exemple 1 (suite)

- On remarque que l’objectif est concave (car linéaire) et que le domaine réalisable est convexe : c’est un modèle de **programmation convexe**
- Ici, la solution optimale est sur la frontière du domaine réalisable, mais ne correspond pas à un coin (l’intersection de deux contraintes)
- Résolution par Excel : *Exemple1_NL.xls*

Exemple 2

- On modifie encore l'exemple *Wyndor Glass* (sect. 2), mais cette fois en remplaçant l'objectif linéaire par une fonction non linéaire

$$\max Z = 126x_1 - 9x_1^2 + 182x_2 - 13x_2^2$$

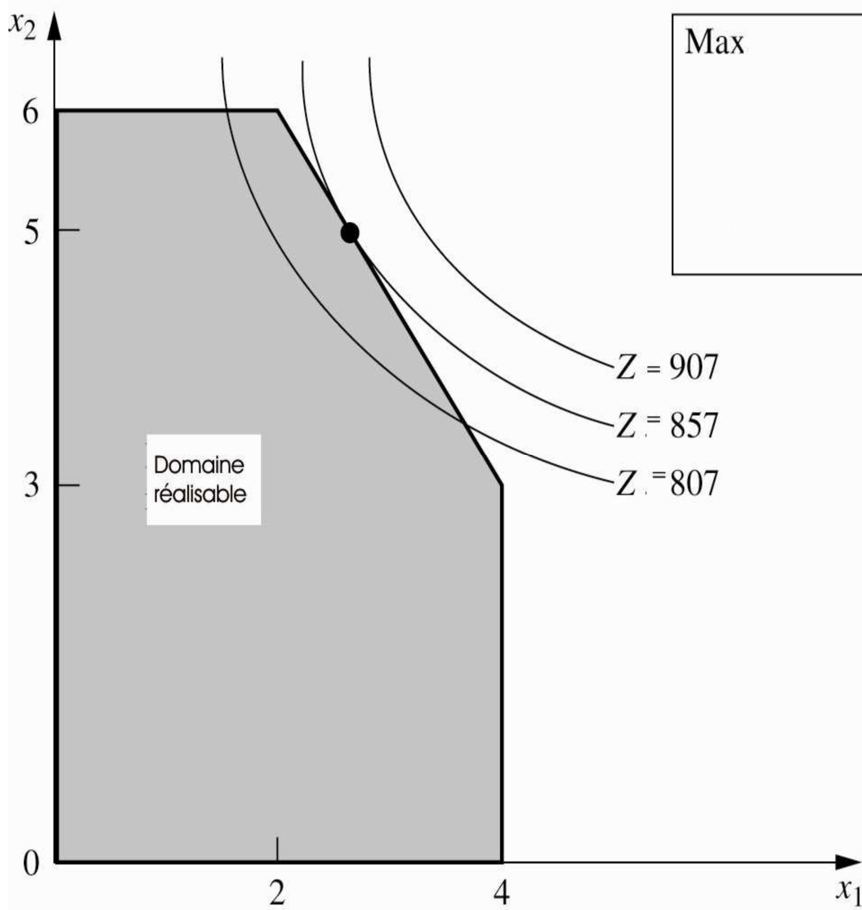
$$x_1 \le 4$$

$$2x_2 \le 12$$

$$3x_1 + 2x_2 \le 18$$

$$x_1 \ge 0, x_2 \ge 0$$

Exemple 2 (suite)



$$
\begin{array}{lrl}
\text{Max} & Z = 126x_1 - 9x_1^2 + 182x_2 - 13x_2^2, & \\
& x_1 & <= 4 \\
& 2x_2 & <= 12 \\
& 3x_1 + 2x_2 & <= 18 \\
& x_1 >= 0, \quad x_2 >= 0 &
\end{array}
$$

Exemple 2 (suite)

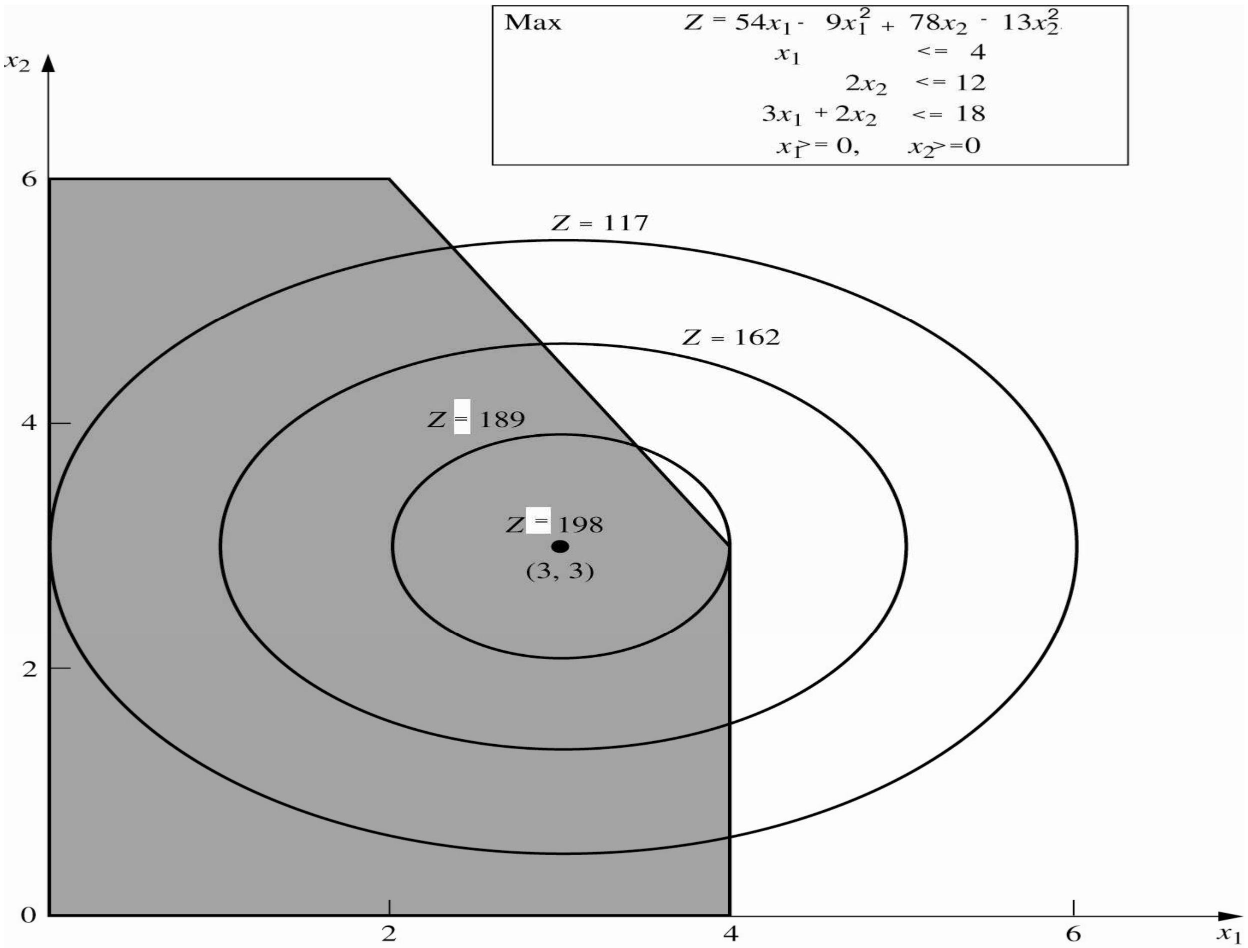
- On remarque que l'objectif est concave et que le domaine réalisable est convexe : c'est aussi un modèle de programmation convexe
- Ici, la solution optimale (8/3,5) est sur la frontière du domaine réalisable, mais ne correspond pas à un point extrême du domaine réalisable
- Résolution par Excel : *Exemple2_NL.xls*

Exemple 3

- Un exemple semblable au précédent, avec une fonction objectif toujours concave, mais différente

$$\max Z = 54x_1 - 9x_1^2 + 78x_2 - 13x_2^2$$
$$x_1 \leq 4$$
$$2x_2 \leq 12$$
$$3x_1 + 2x_2 \leq 18$$
$$x_1 \geq 0, x_2 \geq 0$$

Exemple 3 (suite)



Exemple 3 (suite)

- Il s'agit à nouveau d'un modèle de programmation convexe
- Ici, la solution optimale (3,3) est à l'intérieur du domaine réalisable
- La fonction objectif est la somme de deux fonctions d'une seule variable
- Si on annule les dérivées de chacune de ces fonctions, on obtient la solution unique (3,3), qui se trouve à l'intérieur du domaine : c'est donc nécessairement la solution optimale
- Résolution par Excel : *Exemple3_NL.xls*

Exemple 4

- Un exemple semblable au premier, mais on introduit une contrainte non linéaire qui définit un domaine réalisable non convexe

$$\max Z = 3x_1 + 5x_2$$

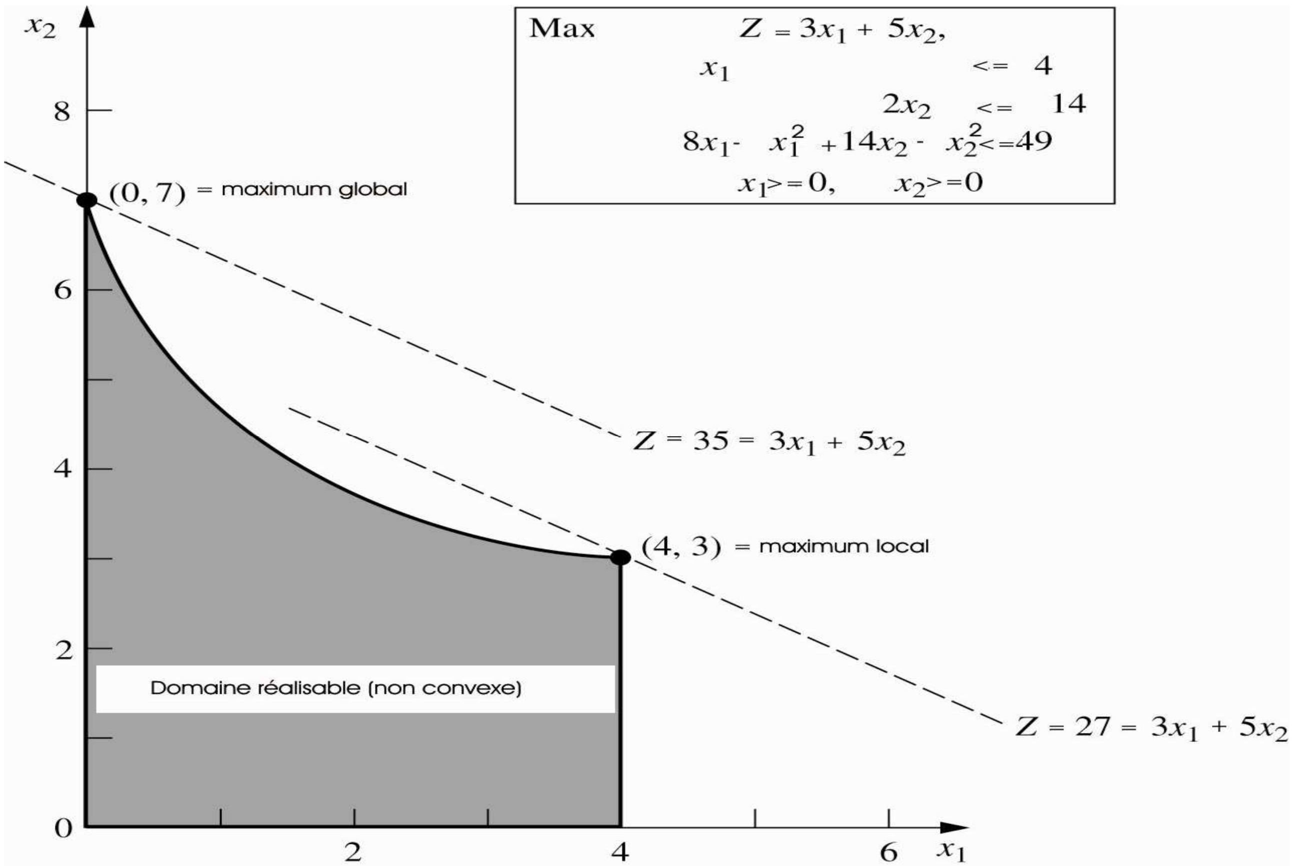
$$x_1 \leq 4$$

$$2x_2 \leq 14$$

$$8x_1 - x_1^2 + 14x_2 - x_2^2 \leq 49$$

$$x_1 \geq 0, x_2 \geq 0$$

Exemple 4 (suite)



Exemple 4 (suite)

- Dans ce modèle de **programmation non convexe**, on remarque la présence de deux maxima locaux :
  - x est un **maximum local** si f(x) ≥ f(x’), pour tout x’ réalisable suffisamment près de x
  - (4,3) et (0,7) sont des maxima locaux, mais x*=(0,7) est le **maximum global** : f(x*) ≥ f(x’), pour tout x’ réalisable
- Les méthodes classiques de programmation non linéaire permettent d’identifier un maximum local, mais pas nécessairement le maximum global!
- Programmation convexe : maximum local = global
- Résolution par Excel : *Exemple4_NL.xls*

Optimisation sans contrainte

- On considère le cas d'un modèle de programmation non linéaire dans lequel il n'y a aucune contrainte :

$$\max f(x_1, x_2, \ldots, x_n)$$

- On peut alors montrer que :

  si $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ est un maximum local, alors

$$\frac{\partial f(x)}{\partial x_j} = 0,\ j = 1,2,\ldots,n,\ \text{en } x = x^*$$

- En d'autres mots, lorsque la dérivée s'annule en un point donné, ce point **peut être** un maximum local
- Il peut aussi être : minimum local ou point de selle

Optimisation sans contrainte (suite)

- Par contre, si *f* est concave, un point où la dérivée s'annule est nécessairement un maximum global
- De plus, si *f* est strictement concave, un tel maximum global est aussi unique
- Sinon, que faire pour trouver un maximum global ?
  - Identifier tous les maxima locaux
  - Identifier celui de plus grande valeur
  - Vérifier que la fonction est bornée supérieurement (sinon, il n'y a pas de maximum global!)
- Problème : identifier tous les maxima locaux peut être difficile!!!

Méthode de la bisection

- On considère d'abord le cas le plus simple : la fonction objectif comporte **une seule variable**
- On suppose que *f* est concave
- On peut montrer alors que, si x^* est une solution optimale, il existe a et b tels que $a \leq x^* \leq b$ et

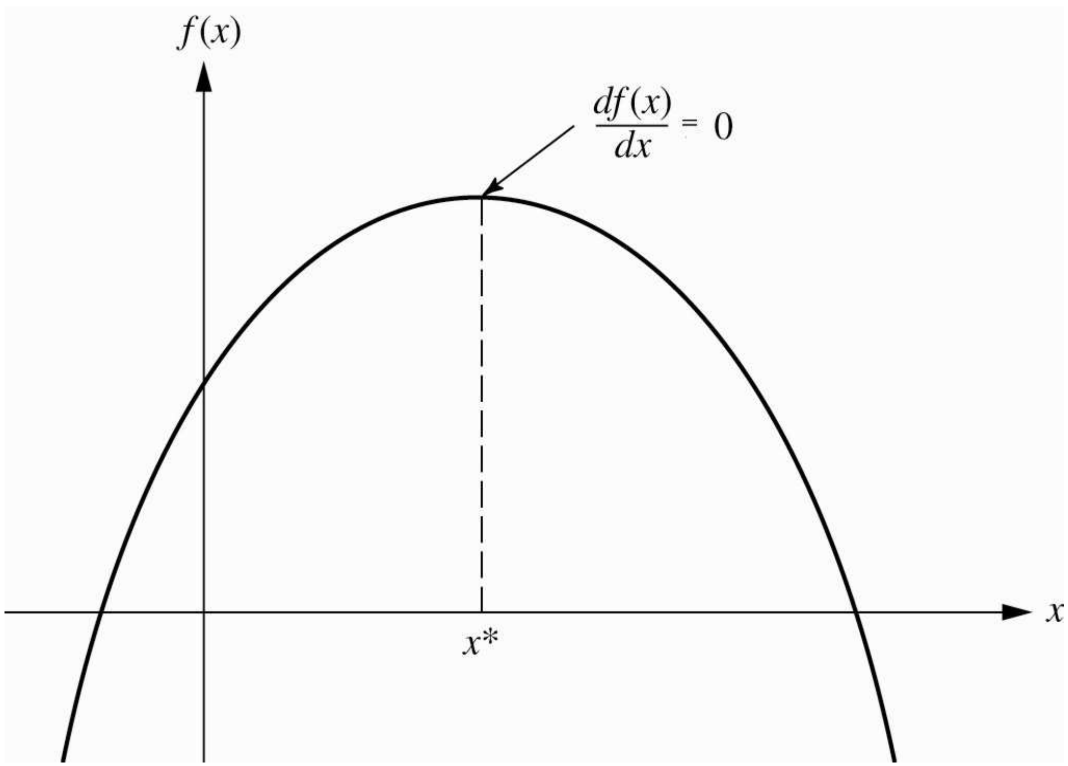
$$\frac{df(x)}{dx} > 0, \text{si } x < a$$

$$\frac{df(x)}{dx} = 0, \text{si } x = x^*$$

$$\frac{df(x)}{dx} < 0, \text{si } x > b$$

Méthode de la bisection (suite)

Si *f* est strictement concave : $a = x^* = b$



Méthode de la bisection (suite)

- On fixe d'abord une borne inférieure x^i pour laquelle la dérivée en ce point est > 0
- On détermine également une borne supérieure x^s pour laquelle la dérivée en ce point est < 0
- Si les deux bornes ne sont pas suffisamment près l'une de l'autre, on prend le point milieu entre les deux bornes comme point candidat x^c
- Si la dérivée en x^c est ≥ 0, on pose $x^i = x^c$
- Si la dérivée en x^c est ≤ 0, on pose $x^u = x^c$
- On itère à nouveau tant que $x^u - x^i$ n'est pas suffisamment petit

Méthode de la bisection (suite)

1. Identifier des bornes inférieure et supérieure initiales x^i et x^u
2. Déterminer le point candidat : $x^c = (x^u + x^i)/2$
3. Si la dérivée en x^c est ≥ 0 : $x^i = x^c$
4. Si la dérivée en x^c est ≤ 0 : $x^u = x^c$
5. Si $x^u - x^i \leq 2\varepsilon$ (où ε est un nombre suffisamment petit), arrêter; sinon, retourner à l'étape 2

Exemple : $f(x) = 12x - 3x^4 - 2x^6$

Exécuter l'algorithme avec IOR Tutorial

Méthode du gradient

- Considérons une fonction concave *f* de **plusieurs variables**
- Le gradient de *f* au point *x′* est défini ainsi :

$$\nabla f(x') = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n}\right) \text{ en } x = x'$$

- On peut montrer que le gradient correspond à une direction d'augmentation de la valeur de *f*
- Dans la méthode du gradient, on se déplace dans la direction du gradient en tentant d'augmenter au maximum la valeur de l'objectif

Méthode du gradient (suite)

- À partir d'un point initial *x'*, on effectue un déplacement dans la direction du gradient vers un nouveau point *x* : $x = x' + t^* \nabla f(x')$
- Dans cette formule, t^* est la solution du problème de maximisation suivant :

$$\max_{t \geq 0} f(x' + t\nabla f(x'))$$

- C'est un problème de maximisation d'une fonction concave d'une seule variable : on peut donc le résoudre par la méthode de la bisection
- On pose $x' = x$ et on itère ainsi jusqu'à ce que le gradient s'annule (ou presque)

Méthode du gradient (suite)

1. Déterminer une solution initiale x'
2. Résoudre le problème suivant : $\max_{t\geq 0} f(x'+t\nabla f(x'))$
3. Soit t^* la solution optimale : $x' = x' + t^* \nabla f(x')$
4. Si $\left|\frac{\partial f(x')}{\partial x_j}\right| \leq \varepsilon, \ j = 1,2,\ldots,n$, où ε est un nombre

   suffisamment petit, arrêter; sinon, retourner en 2

Exemple : $f(x_1, x_2) = 2x_1x_2 + 2x_2 - x_1^2 - 2x_2^2$

Exécuter l'algorithme avec IOR Tutorial

Optimisation sous contraintes

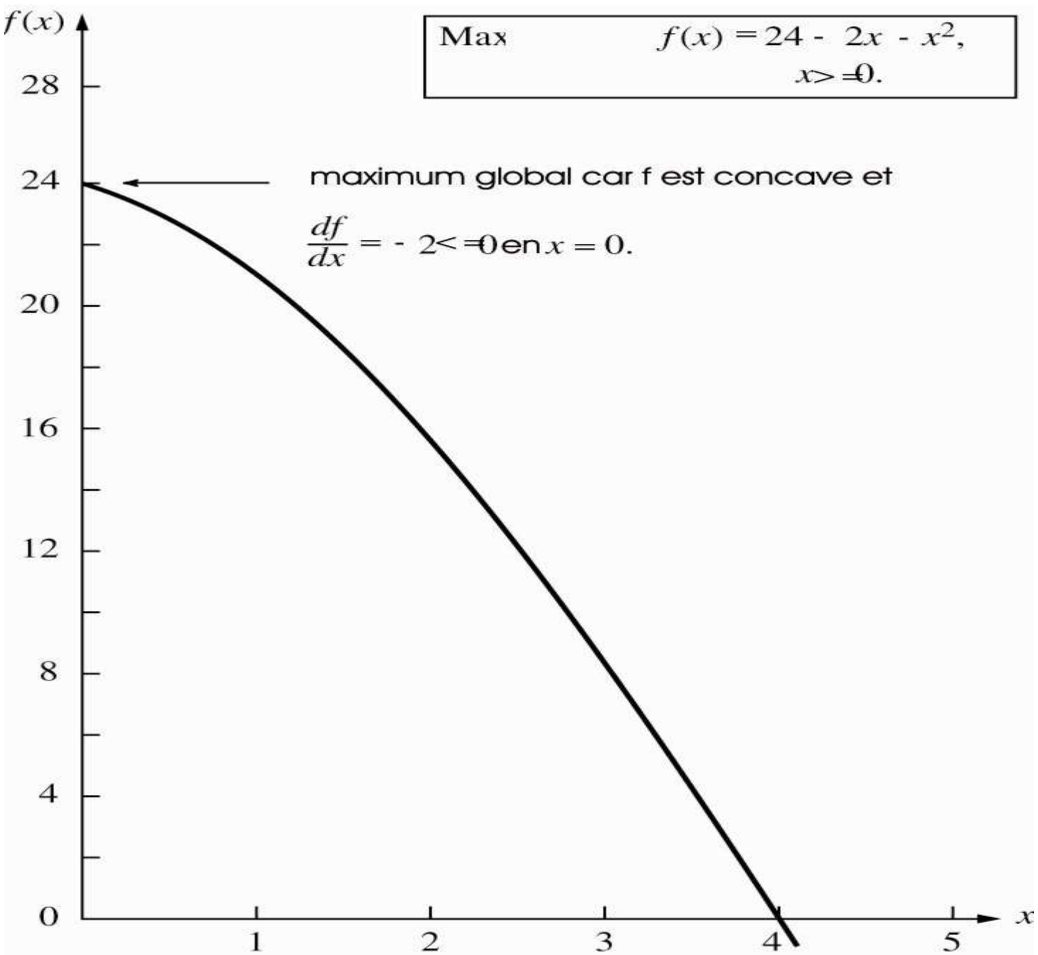
- Dans le cas d'un modèle de programmation non linéaire sans contrainte, nous avons vu la condition d'optimalité suivante :

  si $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ est un maximum local, alors

  $$\frac{\partial f(x)}{\partial x_j} = 0,\ j = 1,2,\ldots,n, \text{ en } x = x^*$$

- Cette condition n'est plus valable lorsqu'il y a des contraintes, car une solution optimale peut se trouver sur la frontière du domaine réalisable

Exemple avec contrainte $x \geq 0$



Conditions d'optimalité

- Avec une contrainte de la forme x ≥ 0, les conditions d'optimalité s'énoncent ainsi :

  si $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ est un maximum local, alors

$$\frac{\partial f(x)}{\partial x_j} \begin{cases} = 0 \text{ en } x = x^* \text{ si } x_j^* > 0 \\ \leq 0 \text{ en } x = x^* \text{ si } x_j^* = 0 \end{cases}$$

- Mais quelles sont ces conditions dans le cas général, où les contraintes sont de la forme :

$$g_i(x_1, x_2, \ldots, x_n) \leq b_i, \quad i = 1,2,\ldots,m$$

- Elles s'expriment en fonction des **multiplicateurs de Lagrange** u_i associés à chaque contrainte

Conditions KKT(Karush-Kuhn-Tucker)

Si $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ est un maximum local, alors il existe *m* nombres $u_1, u_2, \ldots, u_m$ tels que

$$\frac{\partial f(x)}{\partial x_j} - \sum_{i=1}^{m} u_i \frac{\partial g_i(x)}{\partial x_j} \leq 0 \text{ en } x = x^*, j = 1,2,\ldots,n$$

$$x_j^* \left(\frac{\partial f(x)}{\partial x_j} - \sum_{i=1}^{m} u_i \frac{\partial g_i(x)}{\partial x_j}\right) = 0 \text{ en } x = x^*, j = 1,2,\ldots,n$$

$$g_i(x^*) - b_i \leq 0, i = 1,2,\ldots,m$$

$$u_i(g_i(x^*) - b_i) = 0, i = 1,2,\ldots,m$$

$$u_i \geq 0, i = 1,2,\ldots,m$$

Conditions KKT : exemple

$$\max f(x_1, x_2) = \ln(x_1 + 1) + x_2$$
$$2x_1 + x_2 \leq 3$$
$$x_1 \geq 0, x_2 \geq 0$$

- D'abord, il est facile de vérifier que les multiplicateurs de Lagrange associés aux contraintes de non négativité doivent être nuls
- On peut donc considérer qu'il n'y a qu'une seule contrainte avec $g_1(x_1, x_2) = 2x_1 + x_2$ et $b_1 = 3$
- On associe à cette contrainte un multiplicateur $u_1 \geq 0$
- Outre les contraintes de non négativité, on a alors les conditions suivantes

Conditions KKT : exemple (suite)

$$\frac{1}{x_1^* + 1} - 2u_1 \leq 0$$

$$1 - u_1 \leq 0$$

$$x_1^* \left(\frac{1}{x_1^* + 1} - 2u_1\right) = 0$$

$$x_2^* (1 - u_1) = 0$$

$$2x_1^* + x_2^* - 3 \leq 0$$

$$u_1 (2x_1^* + x_2^* - 3) = 0$$

- On obtient $u_1 \geq 1$
- Puisque $x_1^* \geq 0$, on en déduit $\frac{1}{x_1^* + 1} - 2u_1 < 0$
- Donc $x_1^* = 0$

Conditions KKT : exemple (suite)

- Puisque $u_1 \neq 0$, on a $2x_1^* + x_2^* - 3 = 0$
- D'où $x_2^* = 3$
- Puisque $x_2^* \neq 0$, on déduit $u_1 = 1$
- Les conditions KKT sont donc satisfaites en un seul point : (0,3)
- Il s'agit bien d'un maximum global, car la fonction objectif est concave et le domaine réalisable est convexe (modèle de programmation convexe)