abs 21-09-2017

(16p)

Optimisation

Ch 1 : Rappels mathématiques

1) Définition d'un problème d'optimisation :

D'un point de vue mathématique, l'optimisation consiste à rechercher le minimum ou le maximum d'une fonction avec ou sans contraintes.

Un problème d'optimisation avec contrainte est défini comme suit :

a) Problème de minimisation avec contrainte : recherche du minimum d'une fonction

$$PC_1 \begin{cases} \text{Min } f(x) \quad , x \in \mathbb{R}^n \\ \text{sous contraintes} \\ g_i(x) = 0 \; ; i = 1, \dots, m \quad \text{(contrainte d'égalité)} \\ h_j(x) \geq 0 \; ; j = 1, \dots p \quad \text{(contrainte inégalité)} \end{cases}$$

Min $f(x)$: signifie le minimum de $f(x)$, on dit que l'on cherche à minimiser $f(x)$.

b) Problème de maximisation avec contraintes : recherche du maximum d'une fonction.

$$PC_2 \begin{cases} \text{Max } f(x) \; ; x \in \mathbb{R}^n \\ \text{sous les contraintes} \\ g_i(x) = 0 \quad ; i = 1, \dots, m \\ h_j(x) \geq 0 \; ; j = 1, \dots, p \end{cases}$$

Max $f(x)$: signifie le maximum de $f(x)$, on dit que l'on cherche à maximiser $f(x)$.

Dans les problèmes d'optimisation PC_1 et PC_2 la fonction $f(x)$ définit de D dans $\mathbb{R}$ (i.e : $f : D \to \mathbb{R}$) porte divers noms : fonction de coût, fonction objectif ou critère

$x = (x_1, x_2, \dots, x_n) \in D \subset \mathbb{R}^n$

les variables $x_1, \dots, x_n$ sont appelées variables de décision

D est appelé ensemble ou domaine admissible défini par l'ensemble des contraintes comme suit :

$D = \{x \in \mathbb{R}^n / g_i(x) = 0,\ i = 1, \dots, m$
et $h_j(x) \geq 0;\ j = 1, \dots p\}$

Remarque 1 : les Problèmes PC1 et PC2, peuvent être écrits sous les formes simplifiées :

$$PC1 \begin{cases} \text{Min } f(x) \\ x \in D \subset \mathbb{R}^n \end{cases} \qquad PC2 \begin{cases} \text{Max } f(x) \\ x \in D \subset \mathbb{R}^n \end{cases}$$

Remarque 2 : en absence de contraintes $g_i(x) = 0$, $i=1, \dots m$ et $h_j(x) \geq 0$; $j = 1 \dots, p$, les Problèmes PC1 et PC2 deviennent :

$$P1 \begin{cases} \text{Min } f(x) \\ x \in \mathbb{R}^n \end{cases} \qquad P2 \begin{cases} \text{Max } f(x) \\ x \in \mathbb{R}^n \end{cases}$$

Dans ce cas P_1 et P_2 sont resp. des problèmes de minimisation et de maximisation sans contraintes.

Remarque 3 : $\underset{x}{\text{Min}}\,(f(x)) = -\underset{x}{\text{Max}}\,(-f(x))$

$\underset{x}{\text{Max}}\,(f(x)) = -\underset{x}{\text{Min}}\,(-f(x))$

Ainsi la recherche du maximum peut se ramener à la recherche du minimum et vice-versa.

2)- Convexité :

a)- ensemble convexe :

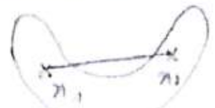
un ensemble $D \subset \mathbb{R}^n$ est dit convexe si pour tout couple $(x_1, x_2) \in D$ et $\forall\ \alpha \in [0,1]$:

on a : $\alpha x_1 + (1-\alpha) x_2 \in D$

Cette définition peut être interprétée en disant que le segment reliant x_1 et x_2 doit être dans D.

exp : ensembles convexes :

ensembles non-convexes ;

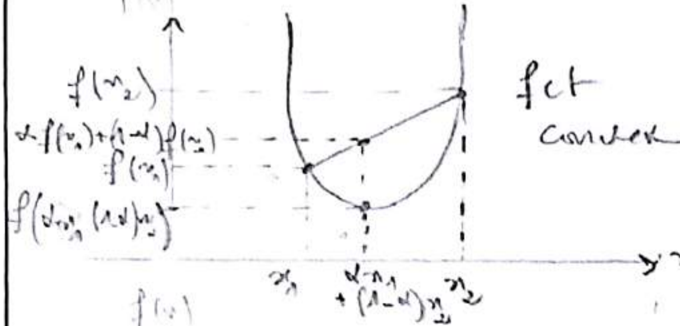


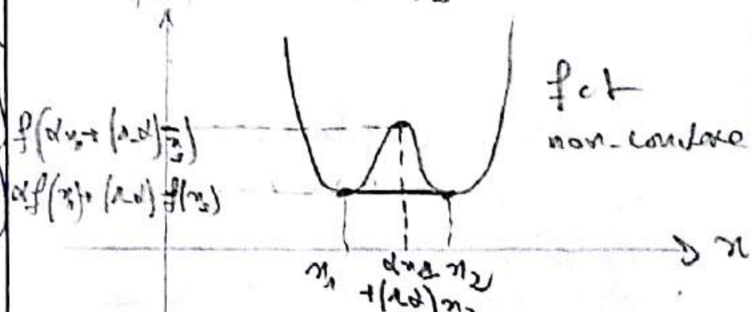
b)- Fonction convexe : une fct :

$f : D \to \mathbb{R}^n$ est convexe ssi :

$\forall x_1, x_2 \in D, \forall \alpha \in [0,1]$.

$f(\alpha x_1 + (1-\alpha) x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$





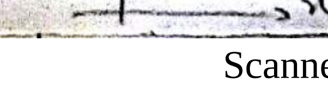
la définition de fonction convexe s'interprète géométriquement comme suit : le graphe de la fonction est toujours en dessous du segment reliant les pts $(x_1, f(x_1))$ et $(x_2, f(x_2))$

On dira que f est strictement convexe dans D ssi :

$\forall x_1, x_2 \in D, \forall \alpha\ [0,1]$

$f(\alpha x_1 + (1-\alpha) x_2) < \alpha f(x_1) + (1-\alpha) f(x_2)$

- Une fonct° concave si $-f$ est convexe
- La définit° de fct et ensemble se fait en inversant les inégalités dans les définitions précédentes



fct concave

28/09/2017

Definition :

Soit :

$$(P) \begin{cases} \text{Min ou Max } f(x) \\ \quad x \in \mathbb{R}^n \\ \text{s.c :} \\ g_i(x) = 0 \quad ; i = 1, \dots, m \\ h_j(x) \geq 0 \quad ; j = 1, \dots, p \end{cases}$$

Définition 1 :

tout vecteur x verifiant l'ensemble des contraintes (i.e: $x \in D$) est appelé solution admissible ou realisable du problème d'optimisation (P).

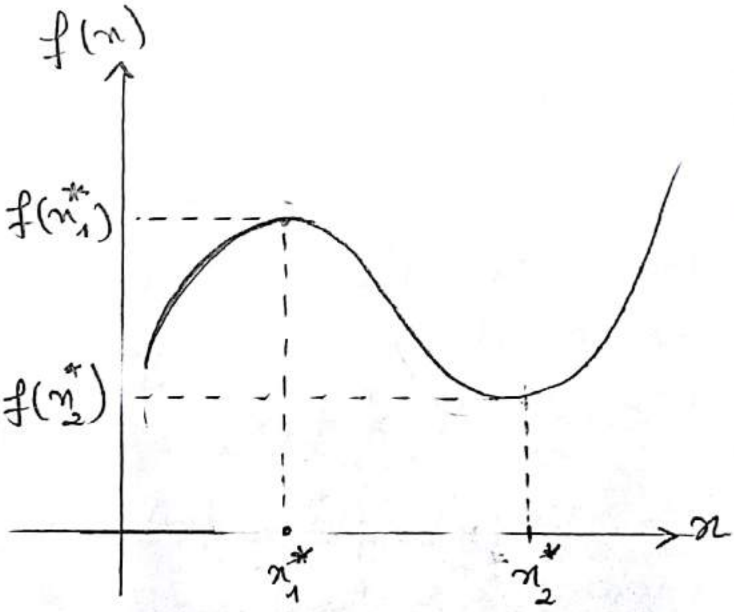
Définition 2 :

un optimum ou extremum d'une fonction f est soit un maximum soit un minimum ie : la valeur la plus haute ou la plus faible que prend la fonction sur son ensemble de définition D.

Cet optimum est donné par $f(x^*)$.

Définition 3 :

le point x^* où la fonction f possede un minimum (resp. Maximum) est dit un point de minimum ou un minimiseur.
(resp . un point de maximum ou maximiseur). x^* est aussi appelé solution optimale du problème d'optimisat°.



x_1^* : est un point de maximum

$f(x_1^*)$: maximum

x_2^* : est un point de minimum

$f(x_2^*)$: minimum

$f(x^*) = \min f(x)$ dans le cas d'un problème de minimisation

$\Rightarrow f(\vec{x})$, minimum

$f(x^*) = \max f(x)$ dans le cas d'un problème de maximisation $\Rightarrow f(x^*)$: maximum.

4)- Minima et Maxima d'une fonction :

Soit f est une fct de $D \subset \mathbb{R}^n$ dans $\mathbb{R}$ ($f : D \to \mathbb{R}$).

a/- minimum ou maximum local : f admet un minimum (resp. maximum) local (ou relatif) en $x^* \in D$, si et seulement s'il existe un voisinage $V_\varepsilon(x^*)$ de x^* tq :

$\forall x \in V_\varepsilon(x^*)$; $f(x^*) \leq f(x)$
(resp $f(x^*) \geq f(x)$).

On dit alors que $f(x^*)$ est un minimum local de la fct f sur D. x^* est un point de minimum local.

b/- minimum ou maximum local strict :

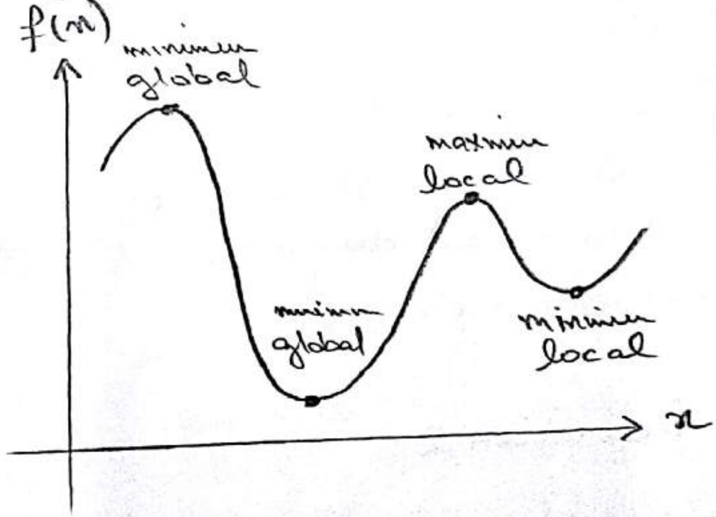
f admet un minimum (resp. maximum) local au sens strict en $x^* \in D$, ssi :
$\exists V_\varepsilon(x^*)$ de x^* tq :

$\forall x \in V_\varepsilon(x^*)$, $x \neq x^*$;
$f(x^*) < f(x)$ (resp. $f(x^*) > f(x)$.

c/. Minimum ou Maximum global : f admet un minimum (resp. maximum) global (ou absolu) en $x^* \in D$ si :

$\forall x \in D$, $f(x^*) \leq f(x)$
(resp $f(x^*) \geq f(x)$)

si les inégalités sont strictes, on obtient la définition d'un minimum (resp. maximum) global strict.



5) Minima des fonctions convexes :

soit f une fonction : $f : D \to \mathbb{R}$ $D \subset \mathbb{R}^n$ est convexe.

f est convexe sur D alors tout minimum local est un minimum global.

- Si f est strictement convexe, alors tout minimum local devient non seulement global mais aussi unique

6/- Gradient et Hessien d'une fonction :

- Gradient :

Soit : $f : \mathbb{R}^n \to \mathbb{R}$, le gradient de f noté $\nabla f(x)$ est donné par :

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

- Hessien ou Matrice Hessienne :

d'Hessien de f noté $\nabla^2 f(x)$ est donné par :

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x_n x_1} & \frac{\partial^2 f}{\partial x_n x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Remarque :

la matrice hessienne est la Jacobienne du gradient.

la matrice Hessienne est symétrique.

- point stationnaire ou critique d'une fonction f :

un point x^* vérifiant $\nabla f(x^*) = 0$ est appelé point stationnaire ou point critique de f.

05/10/2017

7/- Rappel sur les matrices définies semi-définies positives ou négatives.

1). Déf : soit A une matrice symétrique $n \times n$.

On dit que A est semi-définie positive (resp. semi-définie négative) et on note $A \geq 0$. (resp. $A \leq 0$) quand : $x^T A x \geq 0$ (resp. $x^T A x \leq 0$) $\forall x \in \mathbb{R}^n$,

* On dit que A est définie positive (resp. déf. négative) et on note $A > 0$ (resp. $A < 0$) quand :

$x^T A x > 0$ (resp. $x^T A x < 0$) $\forall x \in \mathbb{R}^n, x \neq 0$

* On dit que A est indéfinie s'il $\exists\ x, y \in \mathbb{R}^n$ tq: $x^TAx > 0$ et $y^TAy < 0$

2). Critère pour connaitre les matrices définies, semi-définies positive ou négative :

a). les valeurs propres :

* A est semi définie positive (resp. négative) ssi toutes ses valeurs propres sont ≥ 0 (resp. ≤ 0).

* A est définie positive (resp. négative) ssi toutes ses valeurs propres > 0 (resp < 0)

* A est indéfinie ssi elle admet une valeur propre > 0 et une valeur propre < 0

Rq: pour les calculs des valeurs propres, il faut trouver les racines de l'eqt : $\det(\lambda I - A) = 0$.

b). les mineurs principaux : les determinants des sous-matrices.

Si $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a & a_{22} & \cdots & a_{2n} \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix}$

alors : $D_1 = a_{11}$

$D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} \ldots D_n = \det(A)$

sont les mineurs principaux de A

* la matrice A est définie positive ssi ... $D_k > 0, k = 1 \ldots n$

* A est définie négative ssi : $(-1)^k D_k > 0 \quad k = 1 \ldots n$

* Si $D_k > 0$ $k = 1 \ldots n-1$ et $D_n = 0$ alors A est semi-définie positive

* si $(-1)^k D_k > 0$, $k = 1 \ldots n-1$ et $D_n = 0$ alors A est semi-définie négative.

* Si $D_n < 0$ et si "n" est paire alors A est indéfinie.

Rq: une matrice symetrique qui n'est ni définie positive ni semi positive ; définie ou semi-définie negative est indéfine. $E = \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix}$ $D_1 = 1 > 0$, $D_2 = 3 > 0$ $\Rightarrow$ A déf positive

Contour d'une fonction :

Soit : $f : \mathbb{R}^n \longrightarrow \mathbb{R}$

$y = f(x)$ définit une surface dans $\mathbb{R}^{n+1}$

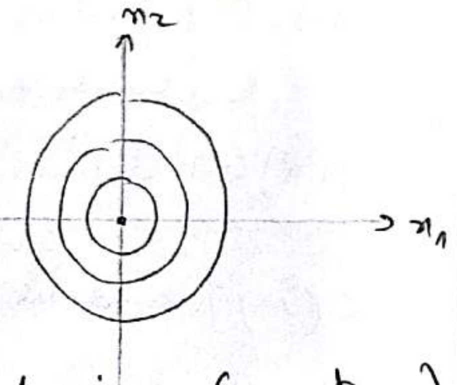
* $f(x) = c$ avec c constant définit des courbes (lignes) de niveaux ou des contours de cette surface

. Un contour de niveau c est l'ensemble de points :

$$S(c) = \{ x \,/\, f(x) = c \}$$

exp : $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$

$$f(x) = x_1^2 + x_2^2 = c$$



- courbe de niveau (ou contour).

9/- Classification des problèmes d'optimisat° :

Un problème d'optimisat° peut être classé suivant les propriétés de la fonction objectif et des contraintes par exp1, on parle :

* Programmation linéaire : lorsque la fct objectif et les contraintes sont linéaires.

* Programmation quadratique : lorsque la fct objectif est quadratique et les contraintes sont linéaires.

* Programmation non-linéaire : lorsque la fct obj et/ou les contraintes sont non-linéaires

* Programmat° convexe : lorsque la fct obje et les contraintes sont convexes.

12/10/2017

10/- Modélisation d'un problème d'optimisation :

Il consiste à représenter un problème d'optimisation mathématiquement, pour cela on distingue 3 étapes :

1- identification des variables de décision (vecteur x)

2- Définition d'une fct objectif : permettant d'évaluer l'état du système (gain, perte ...).

3- Description des contraintes imposées aux variable de décision (le domaine).

La resolution d'un problème d'optimisation consiste à determiner les variables de décision conduisant aux meilleures conditions de fonctionnement du système (ce qui revient à maximiser ou à minimiser la fct objé.) tout en respectant les contraintes

Ch 2: Optimisation sans Contraintes. Methodes locales.

1- Introduction : dans ce chapitre nous allons etudier les methodes de resolution d'un problème d'optimisation sans contraintes. Les methodes sont qualifiées de methodes locales, Pour souligner le fait que le minimum (ou le maximum) trouvé est local.

Les methodes d'optimisation qui seront presentées dans ce chapitre considerent le problème de minimisation sans contraintes suivantes

$$P : \begin{cases} \min f(x) \\ x \in \mathbb{R}^n \end{cases}$$

où : $f : \mathbb{R}^n \to \mathbb{R}$ est une fct donnée quant à la recherche d'un maximum d'une fct $f(x)$ il revient biensur à chercher le minimum de $-f(x)$

- conditions suffisantes pour l'existence d'un minimum ou maximum:

soit $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^2, alors :

x^* est un point de minimum local si :

$$\begin{cases} \nabla f(x^*) = 0 \quad \text{(point critique)} \\ \quad \text{(condition de 1}^{er}\text{ ordre)} \\ \nabla^2 f(x^*) \text{ est définie positive } (>0). \\ \quad \text{(condition du 2}^{eme}\text{ ordre)} \end{cases}$$

x^* : est un point de maximum local si :

$$\begin{cases} \nabla f(x^*) = 0 \quad (x^* \text{ point critique}) \\ \quad \text{(c.d : 1}^{er}\text{ ordre)} \\ \nabla^2 f(x^*) \text{ est définie negative } (<0) \\ \quad \text{(c.d : 2}^{eme}\text{ ordre)} \end{cases}$$

form $\nabla f(x^*) = 0$ et si $f(x)$ est convexe (resp. strict. convexe) alors x^* est un pt de minimum global (resp. global et unique).

* Si $\nabla f(x^*) = 0$ et si $f(x)$ est concave (resp stri. concave) alors x^* est un pt de maximum global (resp. global et unique)

* Si $\nabla f(x^*) = 0$ et $\nabla^2 f(x^*)$ est semi-défini positive (≥ 0) ou semi-défini negative (≤ 0) on ne peut rien conclure.

* Si $\nabla f(x^*) = 0$ et si $\nabla^2 f(x^*)$ est indéfinie, alors x^* est un pt selle ou pt col.

Rq:

Dans le cas d'une fct d'une seule variable $f(x)$, les conditions ci-dessous deviennent:

+ Si $f'(x^*) = 0$ et $f''(x^*) > 0$ $\Rightarrow$ x^* est un pt de minimum local.

+ Si $f'(x^*) = 0$ et $f''(x^*) < 0$ $\Rightarrow$ x^* un point de maximum local.

+ Si $f'(x^*) = 0$ et $f''(x^*) = 0$, x^* peut être un pt d'inflexion

2- Methodes d'optimisation:

les problèmes d'optimisation peuvent être resolus analytiquement ou avec des methodes numeriques iteratives.

1/- methode analytique:

cette methode est basée sur le calcul des pts critiques et l'analyse de la matrice hessienne avec pts critiques. les etapes de cett methodes sont:

1/- recherche des pts critiques

2/- pour chaque pt critique x^*:

* Si $\nabla^2 f(x^*)$ est défini positive (resp. négative): x^* est un pt minimum (resp. pt de maximum local).

* Si $\nabla^2 f(x^*)$ est indefinie, alors x^* est 1 pt selle ou pt col (x^* n'est pas 1 pt d'extremum)

* Si $\nabla^2 f(x^*)$ est semi-définie positive (resp. negative) alors on ne peut rien conclure.

exemple:

Determiner les extremums de les fct suivantes:

$f(x_1,x_2) = x_1^2 + x_2^3 - 2x_1x_2 - x_2$

- calcul des pts critiques :

$$\nabla f(x_1,x_2) = \begin{bmatrix} 2x_1 - 2x_2 \\ 3x_2^2 - 2x_1 - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

on trouve 2 pts critiques :

$(-1/3, -1/3)$ et $(1,1)$

- Calcul de la matrice hessienne :

$$H(x_1,x_2) = \nabla^2 f(x_1,x_2) = \begin{bmatrix} 2 & -2 \\ -2 & 6x_2 \end{bmatrix}$$

$$H(-1/3, -1/3) = \begin{bmatrix} 2 & -2 \\ -2 & -2 \end{bmatrix}$$

$\Delta_1 = 2 > 0$

$\Delta_2 = \det\left(H(-1/3, -1/3)\right) = -8 < 0$

$\Rightarrow H(-1/3, -1/3)$ est indéfinie.

$\Rightarrow (-1/3, -1/3)$ est un pt selle.

$$H(1,1) = \begin{bmatrix} 2 & -2 \\ -2 & 6 \end{bmatrix}$$

$D_1 = 2 >$

$\Delta_2 = \det\left(H(1,1)\right) = 8 > 0$

$\Rightarrow H(1,1)$ est définie positive

$\Rightarrow (1,1)$ est pt de minimum local

$A \in \mathbb{R}^{m \times n}$

le minimum de f est point f_0.

19/11/2.

2) - Methodes de descente (methodes numériques iteratives).

- Notion de direction de descente.

Soit $f : \mathbb{R}^n \to \mathbb{R}$ et $x \in \mathbb{R}^n$, le vecteur $d \in \mathbb{R}^n$ est une direction de descente pour f en x, s'il existe $\delta > 0$ tq :

$$f(x + \alpha d) < f(x), \forall \alpha \in]0, \delta].$$

Les methodes de descente sont des methodes numériques (algorithme) permettant de trouver un minimum local d'une fct. Partant d'un point $x_0 \in \mathbb{R}^n$ arbitrairement choisi, les methodes de descente génèrent une suite de points

$x^0, x^1, x^2, \dots x^k$ tq : $\forall k \in \mathbb{N}$

$f(x^{k+1}) \leq f(x^k)$

avec :

$$x^{k+1} = x^k + \alpha^k d^k \quad \alpha^k > 0$$

Le vecteur d^k est la direction de descente de f en x^k, i.e : la direction dans laquelle la valeur de f decroit.

$x_B = (x_1, \dots x_m, 0 \dots 0), 8$

Scanned by CamScanner

- de descente.

Proposition :

Soit $d \in \mathbb{R}^n$ vérifiant

$$\nabla f(x)^T d < 0$$

alors d est une direction de descente en x.

Démonstration :

on écrit le développement de Taylor de : $f(x+\alpha d)$, $\alpha > 0$ (1^{er} ordre).

$$f(x+\alpha d) = f(x) + \alpha \nabla f(x)^T d + \alpha \varepsilon(\alpha)$$

avec : $\varepsilon(\alpha) \underset{\alpha \to 0}{\longrightarrow} 0$

Pour α suffisament petit, on aura :

$f(x+\alpha d) - f(x) < 0$ ssi : $\nabla f(x)^T d < 0$

$\Rightarrow$ d est une direction de descente

- Méthode de descente :

a). méthode de gradient à pas fixe (ou à pas cte) :

à partir d'un $x^0 \in \mathbb{R}^n$, on construit la suite des points suivante :

$k = 0, 1, \dots,$

$$x^{k+1} = x^k - \alpha \nabla f(x^k)$$

... la direction de descente est

$$d^k = -\nabla f(x^k)$$

$\alpha > 0$: le pas de descente fixe a priori

b) - méthode de gradient à pas optimal :

Partant d'un pt x^0, l'algorithme de gradient à pas optimal est :

$k = 0, 1, \dots$

1 - Calculer un pas optimal α^k de $\underset{\alpha}{\operatorname{Min}} f(x^k - \alpha^k \nabla f(x^k))$

2 - $x^{k+1} = x^k - \alpha^k \nabla f(x^k)$.

critère d'arrêt :

un test d'arrêt devera être choisi pour garantir que l'algorithme s'arrête toujours après un nombre fini d'itérations et que le dernier point calculé soit suffisament proche de point de minimum local de f.

Soit $\varepsilon > 0$, la précision demandée plusieurs critères sont à notre disposition, on en distingue :

1 - $\|\nabla f(x^k)\| < \varepsilon$ (ce critère vient des conditions nécessaire d'optimialité du 1^{er} ordre)

2- $\|x^{k+1} - x^k\| < \varepsilon$ (stagnation de la solution).

donc l'un des critères ci-dessus peut être utilisé.

$\|\cdot\|$: norme euclidienne ($\|\cdot\|_2$)

$$\|x\| = \|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i} \;,\; x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Rq: 1

La methode du gradient à pas optimal est également appelée methode de la plus grande descente (ou pente)

Rq: 2

Dans cette methode on a :

$(x^{k+1} - x^k) \perp (x^k - x^{k-1})$

e) - methode du gradient conjugué :

L'algorithme prend la forme suivante :

on se donne x^0 et on pose :

$d^0 = -\nabla f(x^0)$

$k = 0, 1, \dots$

1- Calculer le pas α^k, la solution de $\min\limits_{\alpha \geq 0} f(x^k + \alpha^k d^k)$

2- $x^{k+1} = x^k + \alpha^k d^k$

3- $$B^k = \frac{\|\nabla f(x^{k+1})\|^2}{\|\nabla f(x^k)\|^2}$$

4- $d^{k+1} = -\nabla f(x^{k+1}) + B^k d^k$

Les critères d'arrêt sont les même que precedement.

cette methode converge rapidement que la methode du gradient.

26.10.2017

d) - Methode de Newton :

soit $f : \mathbb{R}^n \to \mathbb{R}$, $f \in C^2$

Le principe de la methode de Newton est de calculer une approximation $p(x)$ de $f(x)$ autour d'un pt x^k par son developpement de Taylor du second ordre.

$$f(x) \simeq p(x) = f(x^k) + (x - x^k)^T \nabla f(x^k) + \frac{1}{2}(x - x^k)^T H(x^k)(x - x^k)$$

$H(x^k)$: est la matrice hessienne au point x^k $H(x^k) = \nabla^2 f(x^k)$

on calcule alors le nouveau pt x^{k+1} qui minimise $p(x)$.

$\nabla f(x) = \nabla f(x^k) + H(x^k)(x - x^k) = 0$

si la matrice hessienne est inversible $\Rightarrow \boxed{x^{k+1} = x^k - H^{-1}(x^k)\nabla f(x^k)}$

e/- Methode de Newton Modifiée :

L'algorithme de cette methode est comme suit :

$$\boxed{x^{k+1} = x^k - \alpha^k H^{-1}(x^k)\nabla f(x^k)}$$

avec :

α^k est la solution de :

$$\boxed{\min_{\alpha} f\left(x^k - \alpha H^{-1}(x^k)\nabla f(x^k)\right)}$$

f)- Methode Quasi-Newton :

Pour des problèmes de grandes dimensions, le calcul de l'hessienne et de son inverse sont trop coûteux, on peut alors utiliser des algorithmes dit Quasi-Newton qui calculent donc une approximation B^k de $H^{-1}(x^k)$ en fonction de B^{k-1}, $\nabla f(x^k)$, $\nabla f(x^{k-1})$, x^k et x^{k-1}, on trouve notament 2 methodes de calcul :

+ la methode de Davidson - Fletcher - Powell (DFP)

+ La methode de Broyden - Fletcher - Goldfarb - Shanno (BFGS).

– Methodes de recherche unidimensionnelle :

On s'interesse dans cette partie à l'optimisation des fonctions monovariables (Fonction d'une seule variable). Plusieurs methodes ont été proposées, on en distingue :

- methode analytique :

Soit x^*, le pt critique ($f'(x^k)=0$)

$f''(x^k) > 0 \Rightarrow x^k$ pt de minimum local

$f''(x^k) < 0 \Rightarrow x^k$ pt de maximum local

$f''(x^k) = 0 \Rightarrow x^k$ pt d'inflexion

- methode de descente :

Les methodes de descente presentées précédement peuvent être utilisées dans le cas d'une fonction monovariable.

- methode de Dichotomie :

Soit f une fct unimodale.

La 1ère etape consiste en la recherche de a et b tq :

on ait les 2 relations :

$f'(a) < 0$ et $f'(b) > 0$

Après cette etape, on pose :

$x = \frac{1}{2}(a+b)$

si : $f'(x) > 0$, on remplace b par x, sinon on remplace a par x. on repète l'operation jusqu'à $b - a < \varepsilon$

Rq :

on dit qu'une fonction est unimodale, s'il $\exists$ un réel x^* pour lequel la fonction est strictement decroissante sur $]-\infty, x^*]$ et strictement croissante sur $[x^*, +\infty[$.

Ch3 : Programmation linéaire

1) La programmation linéaire a pour objet, l'etude de la resolution des programmes linéaires (P.L)

2) Un programme linéaire : est un problème d'optimisation dont la fonction objectif et les contraintes sont exprimées de manière li...

La forme canonique (ou standard) d'un P.L est :

$$P.L = \begin{cases} \text{Min } c^T x \quad (\text{ou Max } c^T x) \\ s.c : \\ A.x = b \quad (1) \quad (3,1) \\ x \geq 0 \end{cases}$$

avec : $c^T \in \mathbb{R}^n$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$

avec $b \geq 0$, $A \in \mathbb{R}^{m \times n}$

Rq :

Les contraintes (1) peuvent être de type inégalité. ie : $Ax \geq b$ et/ou $Ax \leq b$, et il est tjrs possible de revenir à la forme standard (3,1)

* Transformation sous forme standard :

considerons le P.L suivant

$$\begin{cases} \text{Min } x_1 + x_2 + x_3 \\ s.c : \qquad\qquad \underline{P.L1} \\ x_1 + x_2 - x_3 \geq 2 \\ 2x_1 + 3x_2 + 5x_3 \quad 4 \\ x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0 \end{cases}$$

en introduisant 2 variable d'ecart x_4 et x_5, le problème P.L1 ...

Après cette etape, on pose :

$x = \frac{1}{2}(a+b)$

si : $f'(x) > 0$, on remplace b par x, sinon on remplace a par x. on repète l'operation jusqu'à $b - a < \varepsilon$

Rq:

on dit qu'une fonction est unimodale, s'il $\exists$ un réel x^* pour lequel la fonction est strictement decroissante sur $]-\infty, x^*]$ et strictement croissante sur $[x^*, +\infty[$.

Ch3: Programmation linéaire

1) La programmation linéaire a pour objet, l'etude de la resolution des programmes linéaires (P.L)

2) Un programme linéaire : est un problème d'optimisation dont la fonction objectif et les contraintes sont exprimées de manière linéaire.

La forme canonique (ou standard) d'un P.L est :

$$P.L = \begin{cases} \text{Min } c^T x \quad (\text{ou Max } c^T x) \\ s.c : \\ A.x = b \quad ① \quad (3,1) \\ x \geq 0 \end{cases}$$

avec : $c^T \in \mathbb{R}^n$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$
avec $b \geq 0$, $A \in \mathbb{R}^{m \times n}$

Rq:

Les contraintes ① peuvent être de type inégalité. ie : $Ax \geq b$ et/ou $Ax \leq b$, et il est tjrs possible de revenir à la forme standard (3,1)

* Transformation sous forme standard :

considerons le P.L suivant

$$\begin{cases} \text{Min } x_1 + x_2 + x_3 \\ s.c : \qquad\qquad \underline{P.L1} \\ x_1 + x_2 - x_3 \geq 2 \\ 2x_1 + 3x_2 + 5x_3 \quad 4 \\ x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0 \end{cases}$$

en introduisant 2 variable d'ecart x_4 et x_5, le problème P.L1 peut être ecrit sous la forme canonique suivante.

Min $x_1+x_2+x_3+0x_4+0x_5$

s.c

$x_1+x_2-x_3-x_4=2$

$2x_1+3x_2+5x_3+x_5=4$

$x_1\geq 0 \quad x_2\geq 0 \quad x_3\geq 0 \quad x_4\geq 0 \quad x_5\geq 0$

$$\text{Min } \begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\begin{bmatrix}1 1 1 0 0\end{bmatrix}\underbrace{\begin{bmatrix}x_1\\ \vdots\\x_5\end{bmatrix}}_{x}$$

s.c :

$$\underbrace{\begin{bmatrix}1 & -1 & -1 & 0\\ 2 & 3 & 5 & 0\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}x_1\\ \vdots\\x_5\end{bmatrix}}_{x}=\underbrace{\begin{bmatrix}2\\4\end{bmatrix}}_{b} ; x\geq 0$$

Les variables x_1, x_2 et x_3 sont appelées variables principales ou ponctuelles (Variable de disicsion), les variables x_4 et x_5 sont appelées les variables d'ecart.

3) Définition :

considerons le programme linéaire suivant :

Max c^Tx

$Ax=b$ ①

$x\geq 0$ ②

$b\geq 0$, $x\in\mathbb{R}^n$, $b\in\mathbb{R}^m$

$A\in\mathbb{R}^{m\times n}$

Déf 1 :

Tout vecteur x vérifiant la contrainte ① et ② est appelé solution admissible ou realisable du problème posé

Déf 2 :

Un optimum ou extremum d'une fct f est soit un maximum ou un minimum i.e : la valeur la plus faible que prend la fct sur son ensemble de definition D. cet optimum est donné par $f(x^*)$

Déf 3 :

une solution realisable x^0 est optimal si : $c^Tx^0=\max(c^Tx)$

Déf 4 :

une solution réalisable x est de base si $n-m$ de ses composants sont nulles et aux autres correspondent m vecteur linéairement independants.

exp1 :

$A=[a_1 \quad a_2 \cdots \quad a_m \cdots a_n]$

$x=\begin{bmatrix}x_1\\ \vdots\\x_n\end{bmatrix}$

$Ax=a_1x_1+a_2x_2+\cdots+a_mx_m+\cdots+a_nx_n=b$

$x_B=(x_1,\cdots x_m,0\cdots0)$,

$AB = [a_1 \cdots a_m]$

avec : $det(AB) \neq 0$

$x_1 \cdots x_m$: variables de base

$x_{m+1} \cdots x_n$: variables hors base

02/11/2017

4)- Resolution geometrique d'un P.L à 2 variables :

La resolution geometrique d'un P.L consiste à representer le domaine des solutions admissibles définies par les contraintes ensuite evaluer la fct aux differentes sommets du domaine.

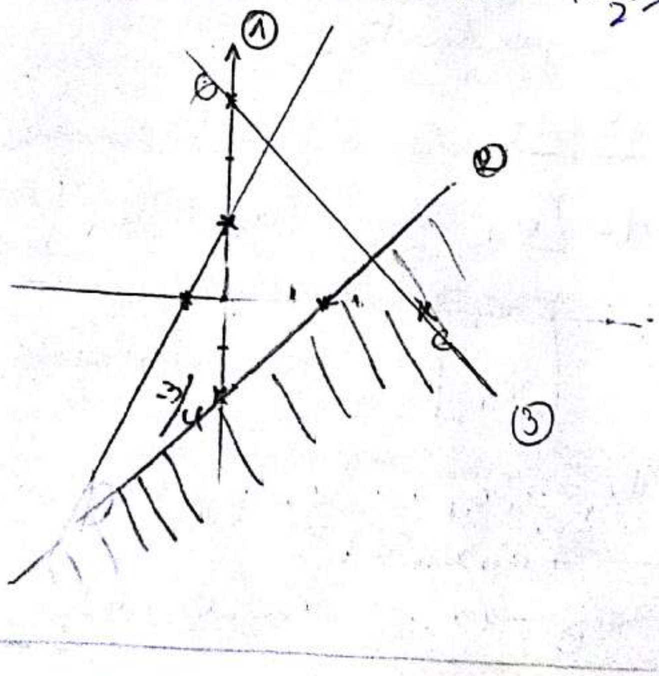
expl : Resoudre géometriquement le P.L suivant :

Min $2x_1 - x_2$

s.c : $-3x_1 + 2x_2 \leq 2$

$2x_1 - 4x_2 \leq 3$

$x_1 + x_2 \leq 6$; $x_1 \geq 0$; $x_2 \geq 0$



① $-3x_1 + 2x_2 = 2$ $(0,1)$ $(-2, ...)$

② $2x_1 - 4x_2 = 3$ $(0, -3/4)$ $(3/2, 0)$

③ $x_1 + x_2 = 6$ $(0,6)$ $(6,0)$

on a 5 sommets :

$D(0,0)$ $C(3/2, 0)$ $E(0,1)$

$A(2,4)$ $B(9/2, 3/2)$

resultat les cpt ① et ③ — l'cpt ③ et ②

$f(D) = 0$ $f(C) = 3$, $f(E) = -1$

$f(A) = 0$ $f(B) = \frac{15}{2}$

$\Rightarrow Min(f) = f(E) = -1$

donc la solution optimale (point de minimum) est $E(0,1)$

Rq :

si le nbr de variable de P.L est superieur à 2, on utilise la methode de simplexe.

5)- Methode de Simplexe :

soit le programme linéaire sur la forme canonique :

Max $c_1x_1 + c_2x_2 + \cdots + c_nx_n$

s.c :

$x_1 + a_{1m+1}x_{m+1} + \cdots + a_{1n}x_n = b_1$

$x_2 + a_{2m+1}x_{m+1} + \cdots + a_{2n}x_n = b_2$

$\vdots$

$x_m + a_{mm+1}x_{m+1} + \cdots + a_{mn}x_n = b_m$

$x_1 \geq 0$ $x_2 \geq 0$, $\cdots$, $x_n \geq 0$

Scanned by CamScanner

$x_j \,/\, j \in J = \{1 \dots n\}$

La methode du simplexe est une methode iterative, elle demarre avec une solution realisable de base.

etape 1 :

choix de la solution réalisable de base de depart.

on a : $J = \{1, \dots n\}$ $J_B = \begin{cases} x_1 \dots x_m \\ 1, 2 \dots m \end{cases}$

$JH = \begin{cases} x_{m+1} & x_n \\ m+1 & n \end{cases}$

$A_B = I_m = [A_1 A_2, \dots A_m]$

J_B choisi pour avoir :

A_B : une matrice identité

$x_1 \dots x_m$: variable de base

$x_{m+1} \dots x_n$: " hors base.

La solution realisable de base de depart $x = (b_1 \dots b_m, 0 \dots 0)$

etape 2 :

construction du Tableau de Simplexe et calcul des E_j

	C^T	C_1	C_2	...	C_m	C_{m+1}	...	C_n	
Valeurs de base	b	A_1	A_2	...	A_m	A_{m+1}	...	A_n	θ_j
A_1	b_1	1	0		0	a_1^{m+1}		a_1^n	θ_1
A_2	b_2	0	1		0	a_2^{m+1}		a_2^n	θ_2
...	...	.	0		.	.			
	b_m	0	0	...	1	a_m^{m+1}		a_m^n	θ_m
	E_j	$E_1 = 0$	$E_2 = 0$		$E_m = 0$	E_{m+1}		E_n	

$E_j = C_B^T A_B^{-1} A_j - C_j$; $C_B = \begin{bmatrix} C_1 \\ C_m \end{bmatrix}$

si $E_j \geq 0$; $j = \overline{1,n} \Rightarrow$ Stop,

$x = [b_1 \dots b_m, 0, 0]$

etape 3 :

a/- Determination du vecteur rentrant dans la base :

le vecteur A_{j_0} rentrant dans la base est celui qui correspond à E_{j_0}, avec : $E_j < 0$, $j \in J_H$

par la division de l'ancienne ligne par l'élément pivot α.

09/11/2017

e)- Détermination des autres lignes :

On cherche des combinaisons linéaire entre la nouvelle ligne et le reste des lignes de façon à rendre le vecteur entrant A_{j_0} un vecteur de base.

f)- Calcul des E_j :

si $E_j \geq 0$, $j = \overline{1,n}$ $\Rightarrow$ stop.
x est une solution optimale
si non on refait l'étape 3.

$L_6 - 2L_1$

Vecteur de base	c^T	4	-2	2	0	0	0	
	b	A_1	A_2	A_3	A_4	A_5	A_6	θ_j
L_4 A_4	6	1	-1	1	1	0	0	6/1
L_5 A_5	2	1	1	1	0	1	0	2/1 →
L_6 A_6	4	2	2	1	0	0	1	4/2
	E_j	-4 ↑	2	-2	0	0	0	
L_4' A_4	4	0	-2	0	1	-1	0	
L_1 A_1	2	1	1	1	0	1	0	$L_4' = L_4 - L_1$
L_6' A_6	0	0	0	-1	0	-2	1	
	E_j	0	6	2	0	4	0	

$E_j = c^T_B \; A_j - c_j$ $\quad [0\,0\,0]\begin{bmatrix}-1\\1\\2\end{bmatrix}+2$

$E_1 = \underbrace{[0\ 0\ 0]}_{c^T_B}\underbrace{\begin{bmatrix}1\\1\\2\end{bmatrix}}_{A_j} - \underbrace{4}_{c_j} = -4$

$L_1 = \frac{L_5}{\alpha} = L_5$ $\quad L_4' = L_4 - L_1$

$L_6' = L_6 - 2L_1$

$c^T_B = [c_4\, c_1\, c_6] = [0\ 4\ 0]$

$y \geq 0 \Rightarrow \overset{*}{x} = (2\ 0\ 0\ 4\ 00)$ est la solution optimale (pr de Max)

$Max(f) = f(x^*) = 8$

- Methode des 2 phases de simplexe :

considerons le probleme de programmation linéaire suivant :

$$(P) \begin{cases} Max\ c^T x \\ s.c. \\ Ax = b \\ x \geq 0 \quad x \in \mathbb{R}^n \quad b \in \mathbb{R}^m \quad c \in \mathbb{R}^n \quad b \geq 0 \end{cases}$$

Dans le cas ou la solution de base de depart n'est pas evidente ou n'est pas realisable on ajoute des variables artificielles.

Ainsi, la resolution du probleme (P) sera faite en 2 phase.

a)- Premiere phase :

La 1ere phase de resolution du probleme (P) consiste a resoudre le probleme auxilliaire suivant par la methode de simplexe :

$$(P_1) \begin{cases} Max\left(-\sum_{i=1}^{m} x_{n+i}\right) \\ s.c : \\ [Ax]_i + x_{n+i} = b_i \quad ; \ i = \overline{1,m} \\ x \geq 0 \quad x_{n+i} \geq 0 \end{cases}$$

où les x_{n+i} sont appelées variables artificielle, Ainsi le probleme (P_1) possède $n+m$ variables :

$X = (x_1 \ldots x_n\ x_{n+1} \ldots x_{n+m})$

Soit $X^* = (x_1^* \ldots x_n^*\ x_{n+1}^* \ldots x_{n+m}^*)$ la solution optimale de (P_1)

2 cas peuvent se produire :

1er cas :

$x_{n+i}^* \neq 0$ pour au moins un indice, dans ce cas le probleme (P) n'admet pas de solution.

2eme cas :

$x_{n+i}^* = 0 \ , \forall i \Rightarrow$ on passe à la 2eme phase.

b)- 2eme Phase :

Dans cette phase, on resoud (P), en prenant $x^* = (x_1^* \ldots x_n^*)$ comme solution de base de depart

exepl : resoudre par Simplexe le P.L suivant :

$$(P) \begin{cases} Max\ (x_1 - x_2 + 2x_3) \\ s.c : \\ \frac{1}{2}x_1 + x_2 + x_3 = 10 \\ x_1 - x_2 = 20 \\ x_j \geq 0 \quad j = \overline{1,3} \end{cases}$$

16. 11. 2017

La variable de base n'est pas évidente dans la 2ème contrainte donc on ajoute une variable artificielle x_4 et on passe à la 1ère phase.

phase 1 : $Max(-x_4)$

$$\frac{1}{2}x_1 + x_2 + x_3 = 10$$
$$x_1 - x_2 + x_4 = 20$$
$$x_j \geq 0 \quad j = \overline{1.4}$$

$I_B = \{x_3, x_4\} = \{3, 4\}$

$I_H = \{x_1, x_2\} = \{1, 2\}$

$x^* = (0\ 0\ 10\ 20)$ est la solution de départ de base. × si tous 0 ?

⇒ tableau de simplexe :

Vecteur de base	c^T	0	0	0	-1	
	b	A_1	A_2	A_3	A_4	θ_j
A_3	10	1/2	1	1	0	20
A_4	20	1	-1	0	1	20 →
	E_j	-1	1	0	0	
A_3	0	0	3/2	1	-1/2	
A_1	20	1	-1	0	1	
	E_j	0	0	0	1	

$E_j \geq 0 \Rightarrow x^* = (20\ 0\ 0\ 0)$ la solution optimal.

$c^T_B = (0\ -1)$

$L'_3 = L_3 - \frac{1}{2}L_1$

$L_1 = \frac{L_4}{\alpha} = L_4$

$c^T_B = (0.0)$

Comme $x_4 = 0$ donc on passe à la 2ème phase avec $x^* = (20\ 0\ 0)$ est la solution de base de départ de (P). on utilise donc le dernier tableau de la phase 1, en éliminant la variable artificielle x_4

	c^T	1	-1	2
	b	A_1	A_2	A_3
A_3	0	0	3/2	1
A_1	20	1	-1	0
	E_j	0	3	0

$c^T_B = (2\ 1)$

$E_j \geq 0 \Rightarrow x^* = (20\ 0\ 0)$ est la solution optimal (pt de max) de (P). $Max(f) = f(x^*) = 20$

- Problème dual d'un problème de programmation linéaire :

considérons le problème de programmation linéaire suivant

$$\begin{cases} Max(c^T x) \\ s.c : Ax = b \\ x \geq 0 \end{cases} \quad (1)$$

= problème ① est dit problème primal, son problème dual est donné par :

$$\begin{cases} \text{Min}(b^T y) \\ A^T y \geq c \\ y \in \mathbb{R}^m \end{cases} \quad ②$$

Rq :
Le dual du dual est le primal

- Methode de construction du dual :

Primal (Dual)	Dual (Primal)
fct à maximiser	fct à minimiser
ième contrainte ≥	ième variable ≤ 0
" " ≤	" ≥ 0
" " =	" ∈ IR
jème variable ≥ 0	jème ctrainte ≥
" " ≤ 0	" " <
" " ∈ IR	" " =

Propretés de la dualité :
etant donné un problème primal ① et son dual ②

Théorème 1 : soitut x^* et y^* 2 solutions realisables du primal et du dual respect. si : $c^T x^* = b^T y^*$ alors x^* et y^* sont des solutions optimales respectivement du primal et du dual

Théorème des Ecarts complementaires)

2 solutions realisables x^* et y^* respectivement du primal et du dual sont optimales si et ssi :

$$\left(\sum_{i=1}^{m} a_{ij} y_i^* - c_j\right) x_j^* = 0$$

$$\left(\sum_{j=1} a_{ij} x_j^* - b_i\right) y_i^* = 0$$

$j = \overline{1,n} \quad i = \overline{1,m}$

Théorème 3 : soit $X = \{x \in \mathbb{R}^n \quad Ax = b,\ x \geq 0\}$

si $X \neq \emptyset$ et si $c^T x \to +\infty$ alors les contraintes du dual sont contradictoires. De même, si l'ensemble des solutions admissibles du dual est non valide et si $b^T y \to -\infty$ alors les contraintes du primal sont contradictoires

expl : ecrire la forme duale du problème suivant

$\text{Max}\ (2x_1 - x_2 + x_3)$

$3x_1 + 2x_2 - x_3 = 2$

$5x_1 - x_2 + x_3 = 3$

$x_1 \geq 0 \quad x_2 \geq 0 \quad x_3 \geq 0$

Dual $\text{Min}\,(b^T y) = \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

$= 2y_1 + 3y_2$

s.c : $3y_1 + 5y_2 \geq 2$

$2y_1 - y_2 \geq -1$

$-y_1 + y_2 \geq 1$

$-y_1 \in \mathbb{R} \quad y_2 \in \mathbb{R}$.

30. 11. 2017

Ch 4: Programmation non linéaire

1. Introduction :

La programmation non linéaire consiste à résoudre des problèmes d'optimisation dont la fct objectif et/ou les contraintes sont non linéaires.

2. Problème avec contraints d'égalité : méthode de multiplicateur de Lagrange

Soit le problème suivant :

Min $f(x)$

s.c :

$g_i(x) = 0 \quad i = 1, \dots, m$

$m \leq n \quad x \in \mathbb{R}^n$

Soit la fonction de Lagrange (ou le Lagrangien) du problème

$$\boxed{\mathcal{L}(x, \lambda) = f(x) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)}$$

où : λ_i sont appelés "multiplicateur de Lagrange".

Soit $g(x) = \begin{bmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{bmatrix}$

Le problème revient à trouver le min de $\mathcal{L}$

1/. Conditions necessaires et suffisantes pour un minimum ou un maximum local :

x^* est un pt de minimum local (resp maximum) de $f(x)$ sous les contraintes $g_i(x) = 0$; $i = 1, \dots m$ si :

1° $\begin{cases} \nabla_x \mathcal{L}(x^*, \lambda^*) = 0 \\ g_i(x) = 0 = \nabla_\lambda \mathcal{L}(x^*, \lambda^*) \end{cases}$

2. $y^T \nabla^2_x \mathcal{L}(x^*, \lambda^*) y > 0$ (resp < 0)

$\forall y \neq 0 \ / \ J_g(x^*) y = 0$

où : $\nabla_x \mathcal{L}(x,\lambda)$: gradient de $\mathcal{L}$ par rapport à x.

$\nabla^2_x \mathcal{L}(x,\lambda)$: hessien de $\mathcal{L}$ % x

$J_g(x^*)$: matrice Jacobienne $(m \times n)$ de $g(x)$.

$$J_g(x) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & & \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \nabla g_1(x)^T \\ \vdots \\ \nabla g_m(x) \end{bmatrix}$$

exple:

$\text{Min}(x_1^2 + x_2^2)$

s.c :

$g(x) = x_1^2 + 2x_2^2 - 1 = 0$

La fct de Lagrange :

$\mathcal{L}(x,\lambda) = f(x) + \lambda g(x)$

$\mathcal{L}(x,\lambda) = x_1^2 + x_2^2 + \lambda(x_1^2 + 2x_2^2 - 1)$

point critique :

1°) :

$$\begin{cases} \nabla_x \mathcal{L}(x,\lambda) = 0 \\ g(x) = 0 = \nabla_\lambda \mathcal{L}(x,\lambda) \end{cases}$$

$$\Rightarrow \begin{cases} 2x_1 + 2\lambda x_1 = 0 \\ 2x_2 + 4\lambda x_2 = 0 \\ x_1^2 + 2x_2^2 - 1 = 0 \end{cases}$$

$$\begin{cases} x_1(1+\lambda) = 0 & (1) \\ x_2(1+2\lambda) = 0 & (2) \\ x_1^2 + 2x_2^2 - 1 = 0 & (3) \end{cases}$$

(1) $\Rightarrow x_1 = 0$ ou $\lambda = -1$

(2) $\Rightarrow x_2 = 0$ ou $\lambda = -1/2$

$x_1 = 0$: (3) $\Rightarrow x_2 = \pm \frac{1}{\sqrt{2}}$

(2) $\Rightarrow \lambda = -1/2$

pt1 $(0, \frac{+1}{\sqrt{2}}, -\frac{1}{2})$ pt2 $(0, \frac{-1}{\sqrt{2}}, -\frac{1}{2})$

$\lambda = -1$: (2) $\Rightarrow x_2 = 0$; (3) $\Rightarrow x_1 = \pm 1$

pt3 $(1, 0, -1)$ pt4 $(-1 \ 0 \ -1)$

$x_2 = 0$, (3) $\Rightarrow x_1 = \pm 1$ (1) $\Rightarrow \lambda = -1$

$\lambda = -1/2$; (1) $\Rightarrow x_1 = 0$; $x_2 = \pm \frac{1}{\sqrt{2}}$

on a 4 pts critiques:

$x^0 = (0 \ \ \frac{+1}{\sqrt{2}})$ avec $\lambda = -1/2$

$x^1 = (0 \ \ \frac{-1}{\sqrt{2}})$ avec $\lambda = -1/2$

$x^2 = (1\ 0)$ avec $\lambda = -1$

$x^3 = (-1\ 0)$ avec $\lambda = -1$

2°):

$$\nabla_x^2 \mathcal{L}(x,\lambda) = \begin{bmatrix} 2+2\lambda & 0 \\ 0 & 2+4\lambda \end{bmatrix}$$

$$Jg(x) = \begin{bmatrix} \frac{\partial g}{\partial x_1} & \frac{\partial g}{\partial x_2} \end{bmatrix}$$

$$Jg(x) = \begin{bmatrix} 2x_1 & 4x_2 \end{bmatrix}$$

Nature du 1er pt x^0 avec $\lambda = -1/2$

$$Jg(x^0) = \begin{bmatrix} 0 & 2\sqrt{2} \end{bmatrix}$$

$$Jg(x^0)\,y = 0 \Rightarrow \begin{bmatrix} 0 & 2\sqrt{2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0$$

$$0y_1 + 2\sqrt{2}\,y_2 = 0 \Rightarrow y = \begin{bmatrix} y_1 \\ 0 \end{bmatrix}$$

$$\nabla_x^2 \mathcal{L}(x^0, -\tfrac{1}{2}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$y^T \nabla_x^2 \mathcal{L}(x^0, -\tfrac{1}{2})\,y = y_1^2 > 0$$

x^0 est point de min local.

$x^0 = (0 ; \frac{+1}{\sqrt{2}})$

Nature du 4ème pt:

$x^3 = (-1, 0)$ $\lambda = -1$

$$Jg(x) = \begin{bmatrix} -2 & 0 \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

07.12.

Nature du 3ème point:

$x^2 = (1\ 0)$ $\lambda = -1$

$$Jg(x) = \begin{bmatrix} 2 & 0 \end{bmatrix} \Rightarrow Jg(x^2)y = \begin{bmatrix} 2 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0$$

$$\Rightarrow y = \begin{bmatrix} 0 \\ y_2 \end{bmatrix}$$

$$\nabla_x^2 \mathcal{L}(x^2, -1) = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}$$

$$\Rightarrow y^T \nabla_x^2 \mathcal{L}(x^2, -1)\,y = -2y_2^2 < 0$$

$x^2 = (1\ 0)$ est un point de max.

Nature du 2ème point:

$x^1 = (0, -\frac{1}{\sqrt{2}})$ $\lambda = -1/2$

$$Jg(x^1) = \begin{bmatrix} 0 & -2\sqrt{2} \end{bmatrix}$$

$$Jg(x^1) = \begin{bmatrix} 0 & -2\sqrt{2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0$$

$$y = \begin{bmatrix} y_1 \\ 0 \end{bmatrix}$$

$$\nabla_x^2 \mathcal{L}(x^1, -1/2) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow y^T \nabla_x^2 \mathcal{L}(x^1, -1/2)\,y = y_1^2 > 0$$

$\Rightarrow x^1 = (0\ -1/\sqrt{2})$ est un pt de min local.

on a 2 pts de min:

$\operatorname{Min} f(x) = f(x^0)$

$\operatorname{Min} f(x) = f(x^1)$

Problème avec contraintes d'inégalité : conditions de Kuhn-Tucker

Soit le problème suivant :

$\text{Min } f(x)$

s.c : $h_j(x) \leq 0 \quad j = 1, \dots p$

Rq : les contraintes de type $\geq$ peuvent être ramenées aux contraintes $\leq$.

Définition : une contrainte d'inégalité $h_j(x) \leq 0$ est active (ou saturée) en x si $h_j(x) = 0$. Elle est inactive si $h_j(x) < 0$

1/ conditions nécessaires et suffisantes pour un minimum ou maximum local : condition de Kuhn Tucker :

soit la fct de Lagrange :

$$\boxed{\mathcal{L}(x, \mu) = f(x) + \sum_{j=1}^{p} \mu_j h_j(x)}$$

μ_j : multiplicateur de Lagrange

Soit $h(x) = \begin{bmatrix} h_1(x) \\ \vdots \\ h_p(x) \end{bmatrix}$; $\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix}$

$\overset{*}{x}$ est un point de min (resp. max) local de $f(x)$ sous les contraintes $h_j(x) \leq 0$; $j = 1 \dots p$ si :

1°)
$$\begin{cases} \nabla_x \mathcal{L}(\overset{*}{x}, \overset{*}{\mu}) = 0 \\ \nabla_{\mu_j} \mathcal{L}(\overset{*}{x}, \overset{*}{\mu}) = h_j(\overset{*}{x}) \leq 0 \quad j = 1, \dots p \\ \overset{*}{\mu}_j h_j(\overset{*}{x}) = 0 \ ; \ j = \overline{1, p} \\ \overset{*}{\mu}_j \geq 0 \ ; \ j = 1, \dots p \ (\text{resp } \overset{*}{\mu}_j \leq 0) \end{cases}$$

2°) $y^T \nabla^2_x \mathcal{L}(\overset{*}{x}, \overset{*}{\mu}) y > 0$ (resp < 0) $\forall y \neq 0 / y^T \nabla h_j(\overset{*}{x}) = 0$; $\forall j \in I(\overset{*}{x})$

$I(\overset{*}{x})$ est l'ensemble des indices des contraintes actives

exemple :

$$\text{Min} (\underbrace{4x_1^2 + 5x_2^2}_{f(x)})$$

s.c :

$h(x) = x_1 - 1 \leq 0$

La fonction de Lagrange (Lagrangien)

$\mathcal{L}(x, \mu) = f(x) + \mu h(x)$

$= 4x_1^2 + 5x_2^2 + \mu(x_1 - 1)$

* conditions de Kuhn Tucker :

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x_1} = 8x_1 + \mu = 0 \quad \text{--- (1)} \\ \frac{\partial \mathcal{L}}{\partial x_2} = 10x_2 = 0 \quad \text{--- (2)} \\ x_1 - 1 \leq 0 \quad \text{--- (3)} \\ \mu(x_1 - 1) = 0 \quad \text{--- (4)} \\ \mu \geq 0 \quad \text{--- (5)} \end{cases}$$

(4) $\Rightarrow \mu = 0$ ou $x_1 = 1$

1^er^ cas : $\mu = 0 \quad x_1 \neq 1 \; (h<0)$
$\Rightarrow h$ inactive

$$\begin{cases} 8x_1 = 0 \\ x_2 = 0 \\ x_1 - 1 < 0 \Rightarrow -1 < 0 \\ \mu = 0 \end{cases}$$

Le 1^er^ pt critique $(\overset{*}{x_1}, \overset{*}{x_2}, \mu) = (0 \quad 0 \quad 0)$

2^eme^ cas : $\mu \neq 0 \quad x_1 = 1$ (h est active $\Rightarrow h = 0$)

$8x_1 + \mu = 0 \Rightarrow \mu = -8x_1 = -8$

$\mu < 0 \Rightarrow$ solution rejectée

on a trouvé un seul pt critique $x^1 = (0 \; 0)$; $\overset{*}{\mu} = 0$

- Etude de la nature du pt critique : x^1

$y^T \nabla_x^2 \mathcal{L}(x^1, \mu^*) y > 0, \; \forall y \neq 0$ $\cancel{y^T \nabla h(x) = 0}$

$$\nabla_x^2 \mathcal{L}(x, \mu) = \begin{bmatrix} 8 & 0 \\ 0 & 10 \end{bmatrix}$$

$\Rightarrow \Delta_1 = 8 > 0$
$\quad \Delta_2 = 80 > 0$

$\Rightarrow \nabla_x^2 \mathcal{L}(x^1, \mu^*) > 0 \Rightarrow x^1$ est un pt de min

$\min(f) = f(x^1) = 0$

4- Probleme avec contraintes d'égalité et d'inégalité : condition de Kuhn Tucker

14-12-2017

Soit le problème suivant :

$\min f(x)$

s.c

$g_i(x) = 0 \quad i = 1, \dots n$

$h_i(x) \leq 0 \quad j = 1, \dots, p$

1)- condition nécessair suffisante pour un min ou un max local _ Kuhn Tucker

Soit la fct de Lagrange :

$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{p} \mu_j h_j(x)$

λ_i, μ_j multiplicateur de Lagrange.

Soit $g(x) = \begin{bmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{bmatrix}$ $h(x) = \begin{bmatrix} h_1(x) \\ \vdots \\ h_p(x) \end{bmatrix}$

$\mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix}$ $\lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix}$

x^* est un pt de minimum (resp max) local de $f(x)$, sous les contraintes $g_i(x) = 0$ $i = 1, \ldots m$ et $h_j(x) \leq 0$ $j = 1, \ldots p$ si :

1) $\boxed{\Delta x \mathcal{L}(x^*, \lambda^*, \mu^*) = 0}$

$$\begin{cases} g_i(x^*) = 0 & i = 1, \ldots m \\ h_j(x^*) \leq 0 & j = 1, \ldots p \\ \mu_j^* h_j(x^*) = 0 & j = \overline{1, p} \\ \mu_j^* \geq 0 & j = 1 \ldots p \\ (\text{resp } \mu_j^* \leq 0) \end{cases}$$

2) $y^T \nabla (x^*, \lambda^*, \mu^*) y > 0$ (resp < 0)

$\forall y \neq 0$

$$\begin{cases} J_g(x^*) y = 0 \\ y^T \nabla h_j(x^*) = 0 & \forall j \in I(x^*) \end{cases}$$

(exp. serie 3 exo 3)

5- Methode de Lagrange-Newton pour des contraintes d'egalité :

Soit le probleme suivant :

Min $f(x)$

s.c

$g_i(x) = 0$; $i = 1, \ldots m$

Soit la fonction de Lagrange

$\mathcal{L}(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$

Soit : $\lambda = \begin{bmatrix} \lambda_i \\ \vdots \\ \lambda_m \end{bmatrix}$

Le probleme revient à chercher le min de $\mathcal{L}(\cdot)$ en utilisant la methode iterative de Newton en partant d'un pt initial $\begin{bmatrix} x^0 \\ \lambda^0 \end{bmatrix}$

algorithme de Newton : (exo 4 TD3 exp)

$\begin{bmatrix} x^0 \\ \lambda^0 \end{bmatrix}$ pt initial $k = 0, 1 \ldots$

$$\begin{bmatrix} x^{k+1} \\ \lambda^{k+1} \end{bmatrix} = \begin{bmatrix} x^k \\ \lambda^k \end{bmatrix} - \left[\nabla^2 \mathcal{L}(x^k, \lambda^k) \right]^{-1} \nabla \mathcal{L}(x^k, \lambda^k)$$

$\nabla^2 \mathcal{L}$: matrice hessienne de $\mathcal{L}$

6 - Autres methodes.

Il ∃ d'autres methodes destinées à la solution d'un programme non-linéaire, on en definit.

* methode de pénalisation
* " de gradient projeté
* " de Newton projeté
* " de dualité (methode d'UZAWA).

Ch 5: Optimisation sans contraintes

methode globales.

1 - Introduction:

contrairement à toutes les methodes d'optimisation vues jusqu'ici, les methodes d'optimisation globales permettent d'obtenir un optimum global d'une fonction. Parmi les methodes globales, nous citons les algorithmes genetiques (AG).

Un algorithme génétique est une methode d'optimisation iterative aléatoire (stakastip) basée sur les probabilités. Il utilise une population d'individus represent les solutions potentielles du problème d'optimisation à resoudre. Cette population va evoluer de generation en generation à l'aide de differents operations : selection, croisement, mutation. L'inconvenient majeur de AG réside dans le nbr important d'evaluation nécessaire et leur convergence.

2 - Les etapes de l'optimisation avec les AG:

Soit le problème d'optimisation suivant :

$$\text{Min } f(x)$$

$$x = (x_1 \dots x_n)$$

L'optimisation avec les AG se base sur les etapes suivantes

1- Generation d'une population initiale et codage :

chaque x_i, $i = 1, \ldots, n$ correspond à une valeur réelle qui peut être codée à l'aide d'un alphabet.

L'alphabet le plus simple composé de 2 caractères, est l'alphabet binaire (composé de 0 et 1)

Ce nombre une fois codé, est appelé chromosome.

chaque caractère du chromosome (0 ou 1) est appelé gène ou (allèle). Le point x est alors construit en ajoutant bout à bout tous les codages des x_i. A l'aide de ce codage, on génère aléatoirement une population initiale de N individus, Une fois les individus generés, les operations de selection, de croisement, de mutation peuvent être appliqués à cette population.

exemple de codage :

une fonction à 5 variable

$x = (x_1, \ldots x_5)$

$x_1 = 14$; $x_2 = 12$; $x_3 = 2$

$x_4 = 7$ $x_5 = 9$ $= x$

$$\underbrace{\overset{14}{1110}}_{x_1}\ \underbrace{\overset{12}{1100}}_{x_2}\ \underbrace{\overset{2}{0010}}_{x_3}\ \underbrace{\overset{7}{0111}}_{x_4}\ \underbrace{\overset{9}{1001}}_{x_5}$$

taille de chromosome = 20 bits

<u>21-12-2017</u>

b/- Evaluation des individus de la population :

Dans cette etape, les individus sont evalués par les fcts d'adaptation (fct $f(x)$ à optimiser (fitness))

c/- Selection :

cette etape permet de choisir les individus (Parents), sur sur lesquels vont s'appliquer les operations de reproduction (croisement et mutation) pour la creation de la future generation.

La selection proportionnelle à la methode de selection la plus utilisée. Cette methode consiste à choisir un nombre $d \in [0\ 1]$ puis pour chaque individu x^i de la population on calcule la quantité suivante:

$$\boxed{n_i = \frac{f(x^j)}{\sum_{j=1}^{N_p} f(x)}}$$

9 → 1001
11 → 1011
B → 1011
C → 1100
D → 1101
E → 1110
F → 1111

N_p: nbr d'individus fait la population

$f(x)$: valeur d'adaptation de l'individu j.

n_i: le nbr de sélection operée de l'individu

si $n_i > x$: selectionner x^i à partir des elements selectionnés on forme une nouvelle population dont la taille est la mm que celle de la population initiale (N_p).

d/- Croisement :

L'idée du croisement est de créer des enfants. On associe donc aléatoirement des individus de la population 2 à 2 afin de former des couples, chaque couple va donner naissance à 2 enfants comportant chacun des sequences chromosomiques de chaque parent.

exemple de croisement :
(fct à 5 variables)

$x = (x_1 \text{ --- } x_5)$

population initiale:

parent ①:

$$\underbrace{\overset{14}{1110}}_{x_1}\ \underbrace{\overset{12}{1100}}_{x_2}\ \underbrace{\overset{2}{0010}}_{x_3}\ \underbrace{\overset{7}{0111}}_{x_4}\ \underbrace{\overset{9}{1001}}_{x_5}$$

parent ②:

$$\underbrace{\overset{5}{0101}}\ \overset{15}{1111}\ \overset{11}{1011}\ \overset{0}{0000}\ \overset{10}{1010}$$

Nouvelle population:

enfant ①:

$$\overset{14}{1110}\ \overset{13}{1101}\ \overset{11}{1011}\ \overset{00}{0000}\ \overset{10}{1010}$$

enfant ②:

0101 1110 0010 0111 1001

e/- Mutation:

cette operation consiste à modifier aléatoirement certains gènes d'un individu exemple de mutation :

ancien individu :

1110 1100 0010 0111 1001

nouveau individu

1110 1110 0010 0111 1001

exemple :

considérons un problème de maximisation suivant :

$$\text{Max } x \longrightarrow f(x)$$

s.c :

$$0 \leq x \leq 31$$

où x est un entier.

1/- codage et choix de la population initiale :

nous utiliserons un codage binaire de x. Les séquences (chromosomes) contenant au max 5 bits, car 31 se note

$$(31)_{10} \rightarrow (11111)_2$$

nous fixons la taille de la population initiale à $N_p = 4$ (4 individus). Supposons que nous avons choisi de façon aléatoire les individus suivants :

N°	chromosome	x
1	01101	13
2	11000	24
3	01000	8
4	10011	19

2)- évaluation des individus

N°	Chromosome	x	$f(x)=x^2$ adaptation	r_i
1	01101	13	169	0,144
2	11000	24	576	0,493
3	01000	8	64	0,0547
4	10011	19	361	0,3085

$$r_i = \frac{f(13)}{f(13)+f(24)+f(19)+f(8)}$$

3/ sélection des parents :

on va tirer 4 individus parmi la population en tenant compte de leur adaptation respective par exp :

* 1 copie de l'individu ①
* 〃 〃 ④
* 2 〃 〃 ②
* aucune copie 〃 〃 ③

N°	chromosome	x
1	01101	13
2	11000	24
3	01000	8
4	10011	19

24

4/. Croisement:

Les parents, sont selectionnés au hasard, nous supposons que l'individu 3 s'accouple avec 4, les chromosomes seront coupler au hasard

N°	Parent	Enfant
1	01101	01100
2	11000	11001
3	11000	11011
4	10011	10000

la nouvelle génération :

N°	Chromosome	x
1	01100	12
2	11001	25
3	11011	27
4	10000	16

Maintenant que la nouvelle generation (population) est entierement crée nous pouvons de nouveau l'evaluer aléatoirement les individus suivants

N°	chromosome	x
1	01101	13
2	11000	24
3	01000	8
4	10011	19

Scanned by CamScanner