

On a le problème d'optimisation (minimisation ou maximisation) avec la fonction objective $f(x_1, x_2, \dots, x_n)$ de n variables x_i où $i = 1, \dots, n$ avec m contrainte(s) $g_j(x_1, x_2, \dots, x_n) = 0$ où $j = 1, \dots, m$ suivant :

$\underset{x \in \mathbb{R}^n}{\text{opt}} f(x_1, x_2, \dots, x_n)$ ici opt peut être remplacé par max ou min

sous les m contraintes

$$g_1(x_1, x_2, \dots, x_n) = 0$$

$$g_2(x_1, x_2, \dots, x_n) = 0$$

⋮

$$g_m(x_1, x_2, \dots, x_n) = 0$$

Par cohérence on écrit les contraintes de la façon suivante (comme dans Sampson):

$$g_j(x_1, x_2, \dots, x_n) = a_j - h_j(x_1, x_2, \dots, x_n) = 0$$

Ex. si la contrainte est linéaire on aura

$$g_j(x_1, x_2, \dots, x_n) = a_j - b_1 x_1 - b_2 x_2 \dots - b_n x_n = 0$$

On peut transformer le problème avec contraintes en un problème non-contraint en utilisant le lagrangien que l'on écrit de la manière suivante en introduisant les contraintes dans le lagrangien.

$$\mathcal{L}(\lambda_1, \lambda_2, \dots, \lambda_m, x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n) + \sum_{j=1}^m \lambda_j g_j(x_1, x_2, \dots, x_n)$$

Le lagrangien est une fonction de n variables x_i où $i = 1, \dots, n$ et de m contraintes $g_j(x_1, x_2, \dots, x_n) = 0$ où $j = 1, \dots, m$.

Le problème d'optimisation libre se définit alors de la manière suivante:

$$\underset{x \in \mathbb{R}^n, \lambda \in \mathbb{R}^m}{\text{opt}} \mathcal{L}(\lambda_1, \lambda_2, \dots, \lambda_m, x_1, x_2, \dots, x_n) \Leftrightarrow \underset{x \in \mathbb{R}^n, \lambda \in \mathbb{R}^m}{\text{opt}} f(x_1, x_2, \dots, x_n) + \sum_{j=1}^m \lambda_j g_j(x_1, x_2, \dots, x_n)$$

Rappel :

Si une fonction est (strictement) concave, elle est nécessairement (strictement) quasi-concave.

Si une fonction est (strictement) convexe, elle est nécessairement (strictement) quasi-convexe.

Pour déterminer si on a un maximum ou un minimum local on doit utiliser les règles basées sur les mineurs principaux primaires \bar{m}_i de la matrice Hessienne Bordée suivante en se basant sur les notions de quasi-concavité ou de quasi-convexité.

$$\bar{H} = \bar{H}\mathcal{L}(\lambda_1, \lambda_2, \dots, \lambda_m, x_1, x_2, \dots, x_n) = \left[\frac{\partial \mathcal{L}(\lambda_1, \lambda_2, \dots, \lambda_m, x_1, x_2, \dots, x_n)}{\partial \mathbf{z} \partial \mathbf{z}'} \right] \text{ où } \mathbf{z}' = [\lambda_1 \quad \dots \quad \lambda_m \quad x_1 \quad \dots \quad x_n]$$

$$\bar{H} = \begin{bmatrix} 0 & B \\ B' & H \end{bmatrix} = \begin{bmatrix} 0 & \dots & 0 & \frac{\partial^2 \mathcal{L}(\cdot)}{\partial \lambda_1 \partial x_1} & \dots & \frac{\partial^2 \mathcal{L}(\cdot)}{\partial \lambda_1 \partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \frac{\partial^2 \mathcal{L}(\cdot)}{\partial \lambda_m \partial x_1} & \dots & \frac{\partial^2 \mathcal{L}(\cdot)}{\partial \lambda_m \partial x_n} \\ \frac{\partial^2 \mathcal{L}(\cdot)}{\partial x_1 \partial \lambda_1} & \dots & \frac{\partial^2 \mathcal{L}(\cdot)}{\partial x_1 \partial \lambda_m} & \frac{\partial^2 \mathcal{L}(\cdot)}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 \mathcal{L}(\cdot)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \mathcal{L}(\cdot)}{\partial x_n \partial \lambda_1} & \dots & \frac{\partial^2 \mathcal{L}(\cdot)}{\partial x_n \partial \lambda_m} & \frac{\partial^2 \mathcal{L}(\cdot)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 \mathcal{L}(\cdot)}{\partial x_n \partial x_n} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \dots & 0 & \frac{\partial g_1(x_1, x_2, \dots, x_n)}{\partial x_1} & \dots & \frac{\partial g_1(x_1, x_2, \dots, x_n)}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \frac{\partial g_m(x_1, x_2, \dots, x_n)}{\partial x_1} & \dots & \frac{\partial g_m(x_1, x_2, \dots, x_n)}{\partial x_n} \\ \frac{\partial g_1(x_1, x_2, \dots, x_n)}{\partial x_1} & \dots & \frac{\partial g_m(x_1, x_2, \dots, x_n)}{\partial x_1} & \frac{\partial f(\cdot)}{\partial x_1 \partial x_1} + \sum_{j=1}^m \lambda_j \frac{\partial g_j(\cdot)}{\partial x_1 \partial x_1} & \dots & \frac{\partial f(\cdot)}{\partial x_1 \partial x_n} + \sum_{j=1}^m \lambda_j \frac{\partial g_j(\cdot)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1(x_1, x_2, \dots, x_n)}{\partial x_n} & \dots & \frac{\partial g_m(x_1, x_2, \dots, x_n)}{\partial x_n} & \frac{\partial f(\cdot)}{\partial x_n \partial x_1} + \sum_{j=1}^m \lambda_j \frac{\partial g_j(\cdot)}{\partial x_n \partial x_1} & \dots & \frac{\partial f(\cdot)}{\partial x_n \partial x_n} + \sum_{j=1}^m \lambda_j \frac{\partial g_j(\cdot)}{\partial x_n \partial x_n} \end{bmatrix}$$

Les **mineurs principaux primaires de la hessienne bordée** notés \bar{m}_i que nous allons regarder sont les suivants :

$$\bar{m}_{2m+1} = \det(\bar{H}_{2m+1})$$

$$\bar{m}_{2m+2} = \det(\bar{H}_{2m+2})$$

⋮

$$\bar{m}_{m+n} = \det(\bar{H}_{m+n})$$

On ne regarde pas les mineurs principaux primaires de la hessienne bordée d'ordre plus petit que $2m+1$.

Pour un k donné, \bar{H}_{m+k} est la sous-matrice de \bar{H} prenant les $m+k$ -ièmes premières lignes et $m+k$ -ièmes colonnes de la matrice hessienne bordée \bar{H} , donc \bar{H}_{m+k} comporte $(m+k) \times (m+k)$ éléments.

Ainsi $\bar{m}_{m+k} = \det(\bar{H}_{m+k})$ est le $m+k$ -ième **mineur principal primaire**, c'est donc le déterminant de la matrice \bar{H}_{m+k} .

Les règles pour déterminer si on a un maximum ou un minimum sont les suivantes. Elles sont basées sur la quasi-concavité et la quasi-convexité (une généralisation moins contraignante des notions de concavité et de convexité).

Pour n variables et m contrainte(s) ($m \geq 1$)

Valide pour $n > m$, donc pour $n \geq 2$

| Maximum local (basé sur la quasi-concavité) | Minimum local (basé sur la quasi-convexité) |
|--|---|
| <p>Règles modifiées et simplifiées</p> <p>On a uniquement besoin de regarder les $n - m$ derniers mineurs principaux primaires</p> <p>Ici on doit avoir l'alternance des signes des mineurs principaux primaires $\bar{m}_p = \det(\bar{H}_p)$ pour les $n - m$ derniers mineurs principaux primaires jusqu'au dernier d'ordre $m + n$. C'est-à-dire qu'il faut regarder les $\bar{m}_p = \det(\bar{H}_p)$ pour $p = 2m + 1, \dots, m + n$. Ainsi le premier à regarder est $\bar{m}_{2m+1} = \det(\bar{H}_{2m+1})$.</p> <p>Notez aussi que $2m + 1$ est un nombre impair.</p> <p>Basé sur le nombre de variables</p> <p>Pour un maximum local, on doit avoir alternance des signes avec le dernier $\bar{m}_{m+n} = \det(\bar{H}_{m+n})$ prenant le signe de $\text{signe}((-1)^n)$</p> <p>Si le nombre de variables n est pair on doit avoir que le dernier mineur principal primaire \bar{m}_{m+n} soit de signe positif.</p> <p>Si le nombre de variables n est impair on doit avoir que le dernier mineur principal primaire \bar{m}_{m+n} soit de signe négatif.</p> <p>Les autres $\bar{m}_p = \det(\bar{H}_p)$ pour $p = 2m + 1, \dots, m + n$ doivent alterner de signe jusqu'au dernier suivant la règle ci-haut.</p> <p>Basé sur le nombre de contraintes</p> <p>Si le nombre de contrainte m est impair on doit avoir que le premier mineur principal primaire à regarder $\bar{m}_{2m+1} = \det(\bar{H}_{2m+1})$ soit de signe positif.</p> <p>Si le nombre de contrainte m est pair on doit avoir que le premier mineur principal primaire à regarder $\bar{m}_{2m+1} = \det(\bar{H}_{2m+1})$ soit de signe négatif.</p> <p>Bref on a toujours $\text{signe}(\bar{m}_{2m+1}) = \text{signe}((-1)^{m+1})$ pour le premier à regarder</p> <p>Les autres $\bar{m}_p = \det(\bar{H}_p)$ pour $p = 2m + 1, \dots, m + n$ doivent alterner de signe à partir de celui d'ordre $2m + 1$.</p> <p>Règle générale : On peut établir la règle générale suivante</p> $\text{signe}(\bar{m}_p) = \text{signe}((-1)^{(p+m+(-1)^n+1)})$ <p>où $p = 2m + 1, \dots, m + n$</p> | <p>Règles simplifiées</p> <p>On a uniquement besoin de regarder les $n - m$ derniers mineurs principaux primaires</p> <p>Ici on doit avoir le même signe pour tous les mineurs principaux primaires $\bar{m}_p = \det(\bar{H}_p)$ que l'on va regarder, c'est-à-dire les pour les $n - m$ derniers mineurs principaux primaires jusqu'à celui d'ordre $m + n$ en partant de celui d'ordre $2m + 1$.</p> <p>Le signe de tous les mineurs principaux primaires à regarder sera déterminé par $\text{signe}((-1)^m)$</p> <p>Donc on a la règle</p> $\text{signe}(\bar{m}_p) = \text{signe}((-1)^m)$ <p>Pour tous $p = 2m + 1, \dots, m + n$</p> <p>Si le nombre de contrainte est impair on doit avoir que des signes négatifs.</p> <p>Si le nombre de contrainte est pair on doit avoir que des signes positifs.</p> |

Exemples

| Contraintes m | Variables n | Maximum local (basé sur la quasi-concavité) Les $n - m$ derniers $\bar{m}_p \equiv \det(\bar{H}_p)$ alternent On regarde que les $n - m$ derniers \bar{m}_p pour $p = 2m + 1, \dots, m + n$ Regardez m pour le signe du premier \bar{m}_{2m+1} Regardez n pour le signe du dernier \bar{m}_{m+n} | Minimum local (basé sur la quasi-convexité) Les $n - m$ derniers \bar{m}_p prennent le signe de $\text{signe}((-1)^m)$ On regarde que les $n - m$ derniers \bar{m}_p pour $p = 2m + 1, \dots, m + n$ |
|--------------------|------------------|--|---|
| 1 | 2 | $\bar{m}_3 \equiv \det(\bar{H}_3) > 0$ (le dernier) m impair : Premier de signe positif n pair : Dernier de signe positif | $\text{signe}(\bar{m}_p) = \text{signe}((-1)^m) = \text{signe}((-1)^1)$ $\bar{m}_3 < 0$ signe négatif (le dernier) |
| 1 | 3 | $\bar{m}_3 > 0, \bar{m}_4 < 0$ (2 derniers) m impair : Premier de signe positif n impair : Dernier de signe négatif | $\text{signe}(\bar{m}_p) = \text{signe}((-1)^1)$ $\bar{m}_3 < 0, \bar{m}_4 < 0$ signes tous négatifs (2 derniers) |
| 1 | 4 | $\bar{m}_3 > 0, \bar{m}_4 < 0, \bar{m}_5 > 0$ (3 derniers) m impair : Premier de signe positif n pair : Dernier de signe positif | $\text{signe}(\bar{m}_p) = \text{signe}((-1)^1)$ $\bar{m}_3 < 0, \bar{m}_4 < 0, \bar{m}_5 < 0$ signes tous négatifs (3 derniers) |
| 1 | 5 | $\bar{m}_3 > 0, \bar{m}_4 < 0, \bar{m}_5 > 0, \bar{m}_6 < 0$ (4 derniers) m impair : Premier de signe positif n impair : Dernier de signe négatif | $\text{signe}(\bar{m}_p) = \text{signe}((-1)^1)$ $\bar{m}_3 < 0, \bar{m}_4 < 0, \bar{m}_5 < 0, \bar{m}_6 < 0$ signes tous négatifs (4 derniers) |
| 2 | 3 | $\bar{m}_5 < 0$ (le dernier) m pair : Premier de signe négatif n impair : Dernier de signe négatif | $\text{signe}(\bar{m}_p) = \text{signe}((-1)^m) = \text{signe}((-1)^2)$ $\bar{m}_5 > 0$ signe positif (le dernier) |
| 2 | 4 | $\bar{m}_5 < 0, \bar{m}_6 > 0$ (2 derniers) m pair : Premier de signe négatif n pair : Dernier de signe positif | $\text{signe}(\bar{m}_p) = \text{signe}((-1)^2)$ $\bar{m}_5 > 0, \bar{m}_6 > 0$ signes tous positifs (2 derniers) |
| 2 | 5 | $\bar{m}_5 < 0, \bar{m}_6 > 0, \bar{m}_7 < 0$ (3 derniers) m pair : Premier de signe négatif n impair : Dernier de signe négatif | $\text{signe}(\bar{m}_p) = \text{signe}((-1)^2)$ $\bar{m}_5 > 0, \bar{m}_6 > 0, \bar{m}_7 > 0$ signes tous positifs (3 derniers) |
| 3 | 4 | $\bar{m}_7 > 0$ (le dernier) m impair : Premier de signe positif n pair : Dernier de signe positif | $\text{signe}(\bar{m}_p) = \text{signe}((-1)^m) = \text{signe}((-1)^3)$ $\bar{m}_7 < 0$ signes négatifs (le dernier) |

Règle générale pour un **maximum** valide $\forall n > m$ (avec $n \geq 2$) : $\text{signe}(\bar{m}_p) = \text{signe}((-1)^{(p+m+(-1)^n+1)})$ et
 $\forall p = 2m + 1, \dots, m + n$