Résumé d'optimisation sous contraintes
Méthode de Lagrange

Tout comme pour l'optimisation libre, la démarche pour optimiser localement une fonction $f(\vec{x})$ de plusieurs variables sous contraintes $\vec{h}(\vec{x}) = \vec{0}$ consiste à

1. chercher les points stationnaires du problème sous contraintes
2. étudier la nature de chaque point stationnaire en étudiant le "signe" d'une hessienne bien choisie.

Cette hessienne ne fait plus intervenir seulement la fonction f mais aussi les contraintes $\vec{h}$ par l'intermédiaire du *Lagrangien*.

1 Rappels de certains points essentiels vus en cours

On considère une fonction de n variables $f : \vec{x} = (x_1, \ldots, x_n) \in \Omega \subset \mathbb{R}^n \mapsto f(x_1, \ldots, x_n) \in \mathbb{R}$ qu'on cherche à optimiser sous les contraintes $\vec{h}(\vec{x}) = \vec{0}$ i.e. sur la surface $S = \{\vec{x} \in \Omega : \vec{h}(\vec{x}) = \vec{0}\}$.

Ici, l'application $\vec{h} : \vec{x} = (x_1, \ldots, x_n) \in \Omega \subset \mathbb{R}^n \mapsto \vec{h}(\vec{x}) = \big(h_1(\vec{x}), \ldots, h_p(\vec{x})\big) \in \mathbb{R}^p$ avec $1 \leq p \leq n$ est une application $\mathcal{C}^2$.

Hypothèse importante: on suppose que $\vec{h}$ est **régulière**, i.e. que, pour tout $\vec{x} \in S$, la famille de vecteurs $\big(Dh_1(\vec{x}), \ldots, Dh_p(\vec{x})\big)$ est une famille libre de $\mathbb{R}^n$. Nous ne rappelerons plus cette hypothèse dans la suite.

On a vu en cours qu'on peut alors caractériser facilement l'espace tangent à S en n'importe quel point $\vec{x}$ de S par p équations linéaires :

Théorème 1

$$\boxed{T_{\vec{x}}S = Vect\,(Dh_1(\vec{x}), \ldots, Dh_p(\vec{x}))^{\perp}\,.} \tag{1}$$

L'espace tangent à S en $\vec{x}$ est un s.e.v. de $\mathbb{R}^n$ de dimension $n-p$ d'équation

$$T_{\vec{x}}S = \{\vec{v} \in \mathbb{R}^n : \forall i = 1, \ldots, p,\ \langle \vec{v}, Dh_i(\vec{x}) \rangle = 0\}.$$

Il faut bien comprendre que chacune des p relations $\langle \vec{v}, Dh_i(\vec{x}) \rangle = 0$ est une équation linéaire et que le théorème précédent donne bien l'équation du s.e.v. $T_{\vec{x}}S$.

On supposera toujours que $\vec{h}$ est régulière. S est alors dite surface régulière de dimension $n - p$.

Remarque 2 *La preuve du théorème 1 comprend une partie difficile qui sort du cadre de ce cours. Elle utilise un théorème profond à la base de la géométrie différentielle : le théorème des fonctions implicites. Nous ne ferons que mentionner cette difficulté.*

2 Extrema locaux sous contraintes

Définition 3 *La fonction $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ admet un **maximum local** sous la contrainte $\vec{h}(\vec{x}) = \vec{0}$ au point $\vec{x}_* \in S = \left\{\vec{x} \in \Omega : \vec{h}(\vec{x}) = \vec{0}\right\}$ si on peut trouver $r > 0$ tel que*

1. $B(\vec{x}_*, r) \subset \Omega$,
2. *Pour tout $\vec{y} \in B(\vec{x}_*, r) \cap S$, on a $f(\vec{y}) \leq f(\vec{x}_*)$.*

Cela signifie que les valeurs de f des points de la surface S autour de $\vec{x}_*$ sont inférieures à $f(\vec{x}_*)$.

On définit de même un minimum local sous contraintes en changeant juste le signe $\leq$ par $\geq$. Un extremum est un minimum ou un maximum.

3 Conditions nécessaires du 1er ordre : Lagrangien et points stationnaires

Définition 4 (Lagrangien) *Le Lagrangien du problème d'optimisation de f sous la contrainte*
$\vec{h}(\vec{x}) = \vec{0}$ est la fonction $\mathcal{L}(\vec{x}, \vec{\lambda})$ définie pour $(\vec{x}, \vec{\lambda}) = (x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_p) \in \Omega \times \mathbb{R}^p$ par

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) + \lambda_1 h_1(\vec{x}) + \cdots + \lambda_p h_p(\vec{x}) = f(\vec{x}) + \langle \vec{\lambda}, \vec{h}(\vec{x}) \rangle \in \mathbb{R}.$$

Définition 5 (Point stationnaire et multiplicateurs) *Un point $\vec{x}_* \in \Omega$ est dit **point stationnaire** du problème d'optimisation*

$$\left\{ \begin{array}{l} \textit{Optimiser localement } f(\vec{x}) \\ \textit{sous la contrainte } \vec{h}(\vec{x}) = \vec{0} \end{array} \right.$$

s'il existe un vecteur $\vec{\lambda}_* = (\lambda_1^*, \ldots, \lambda_p^*) \in \mathbb{R}^p$ tel que

$$D\mathcal{L}(\vec{x}_*, \vec{\lambda}_*) = \vec{0}.$$

*Les nombres $\lambda_1^*, \ldots, \lambda_p^*$ sont alors appelés **multiplicateurs de Lagrange** du point stationnaire $\vec{x}_*$.*

Les points stationnaires du problème sous contraintes sont ainsi les points $\vec{x}_* \in \Omega$ tels qu'il existe $\vec{\lambda}_* \in \mathbb{R}^p$ pour lequel $(\vec{x}_*, \vec{\lambda}_*)$ est un point stationnaire de $\mathcal{L}$. On notera ici $\mathcal{S}(f|\vec{h}(\vec{x}) = \vec{0})$ leur ensemble.

Pour trouver les extrema d'une fonction sous contraintes, la première étape est encore de rechercher les points stationnaires. C'est ce qu'affirme le

Théorème 6 (Condition nécessaire du 1er ordre) *Soit $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ une fonction $\mathcal{C}^2$ et $\vec{h} : \Omega \to \mathbb{R}^p$ une application $\mathcal{C}^2$ régulière. Soit $S = \{\vec{x} \in \Omega : \vec{h}(\vec{x}) = \vec{0}\}$ la surface régulière associée. Si $\vec{x}_* \in S$ est un extremum local de f sous la contrainte $\vec{h}(\vec{x}) = \vec{0}$, $\vec{x}_*$ est un point stationnaire du problème d'optimisation sous contraintes* $\begin{cases} \text{Optimiser localement } f(\vec{x}) \\ \text{sous la contrainte } \vec{h}(\vec{x}) = \vec{0} \end{cases}$.

4 Conditions suffisantes du 2nd ordre

Soit $\vec{x}_* \in \mathcal{S}(f|\vec{h}(\vec{x}) = \vec{0})$ un point stationnaire et $\vec{\lambda}_* \in \mathbb{R}^p$ un vecteur de multiplicateurs associé. C'est-à-dire

$$D\mathcal{L}(\vec{x}_*, \vec{\lambda}_*) = \vec{0}.$$

Question: $\vec{x}_*$ est-il ou non un extremum local du problème sous contraintes ?

On peut répondre dans les cas "favorables" grâce à l'étude d'une matrice hessienne bien choisie.

On suppose que f et $\vec{h}$ sont $\mathcal{C}^2$ et que la contrainte $\vec{h}$ est régulière. On note

$$H = D^2_{\vec{x}\vec{x}}\mathcal{L}(\vec{x}_*, \vec{\lambda}_*) = \left(\frac{\partial^2 \mathcal{L}}{\partial x_i\, \partial x_j}(\vec{x}_*, \vec{\lambda}_*)\right)_{i,j=1,\dots,n} \in M_n(\mathbb{R})$$

la **matrice hessienne de $\mathcal{L}$ en la variable $\vec{x}$** au point $(\vec{x}_*, \vec{\lambda}_*)$. Soit $q_H(\vec{u}) = \vec{u}\, H\, {}^t\vec{u}$ sa forme quadratique associée. On note toujours $S = \{\vec{x} : \vec{h}(\vec{x}) = \vec{0}\}$ et $T_{\vec{x}_*}S$ l'espace tangent à S en $\vec{x}_*$.

Théorème 7 (Conditions suffisantes du 2nd ordre) *Sous les hypothèses ci-dessus,*

*i) Si la restriction de q_H à $T_{\vec{x}_*}S$ est **définie positive** , $\vec{x}_*$ est un **minimum local** de f sous la contrainte $\vec{h}(\vec{x}) = \vec{0}$.*

*ii) Si la restriction de q_H à $T_{\vec{x}_*}S$ est **définie négative**, $\vec{x}_*$ est un **maximum local** de f sous la contrainte $\vec{h}(\vec{x}) = \vec{0}$.*

*iii) Si $\vec{0}$ est un **point col** de la restriction de q_H à $T_{\vec{x}_*}S$, $\vec{x}_*$ est un point col de f et **n'est pas un extremum local** de f sous la contrainte $\vec{h}(\vec{x}) = \vec{0}$.*

iv) Dans tous les autres cas, on ne peut pas conclure (à ce stade).

Remarquez bien que la discussion est similaire à celle déjà vue en optimisation libre mais que

- La matrice hessienne H à considérer est maintenant **la matrice hessienne de $\mathcal{L}$ en la variable** $\vec{x}$ au point $(\vec{x}_*, \vec{\lambda}_*)$.
- La forme quadratique dont on étudie le "signe" est la **restriction** de q_H à l'espace tangent $T_{\vec{x}_*}S$.