ETUDE DE PROBLEMES D'OPTIMAL DESIGN

François MURAT et Jacques SIMON

Université Paris VI et CNRS
Laboratoire Associé 189
Analyse Numérique

Tour 55.65 - 5ème étage
4 place Jussieu - PARIS (5°)

On expose ici les grandes lignes d'un travail sur l'optimal design. On travaillera sous des hypothèses fortes pour simplifier l'exposition. Un article plus complet, à paraitre, donnera le détail des démonstrations (qui sont seulement esquissées ici) ainsi que d'autres résultats et la bibliographie.

Les données d'un problème d'optimal design sont sensiblement différentes de celles d'un problème de contrôle habituel, et l'objet de notre travail est de définir un cadre dans lequel on puisse étendre des méthodes et obtenir des résultats classiques en théorie du contrôle.

1. POSITION DU PROBLEME

On appelle "optimal design", ou "contrôle par un domaine géométrique", la recherche d'un domaine (ie ouvert connexe de R^N), réalisant le minimum d'une fonctionnelle donnée définie sur un ensemble de domaines.

On exposera ici les méthodes et les résultats obtenus en se référant à l'exemple suivant, relatif au problème de Neumann non homogène. Etant donnés

(1) $\quad f \in H^1(R^N)\ ,\ g \in H^2(R^N)\ ,\ z \in H^2(R^N)$

(2) $\quad D$ un domaine de R^N, variété à bord de classe $W^{2,\infty}$, et de frontière ∂D,

on définit l'état u_D du système de façon unique par :

$$(3) \qquad \begin{cases} -\Delta u_D + u_D = f & \text{dans } D \\ \dfrac{\partial u_D}{\partial n} = g & \text{sur } \partial D \\ u_D \in H^2(D) & \end{cases}$$

et la fonction coût J par

(4) $$J(D) = \int_{\partial D} |u_D - z|^2 \, ds$$

Etant donnée une famille $\mathcal{D}_{ad}$ de domaines vérifiant l'hypothèse (2) on cherche un élément D_o minimisant J sur $\mathcal{D}_{ad}$, i.e. tel que

(5) $$D_o \in \mathcal{D}_{ad} \quad \text{et} \quad J(D_o) \leqslant J(D) \qquad \forall D \in \mathcal{D}_{ad}$$

On se pose 3 problèmes classiques en optimisation :

i) Etablir l'existence d'un domaine optimal D_o

ii) Donner des propriétés des éléments optimaux

iii) Trouver une méthode de calcul d'un élément optimal.

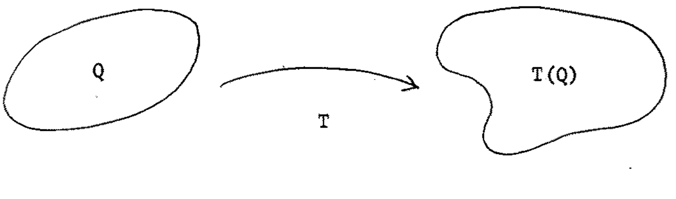
Dans les problèmes de contrôle habituels, le contrôle appartient en général à un sous ensemble d'un espace vectoriel normé, et on utilise des résultats de compacité pour résoudre i), et des résultats de dérivabilité de la fonction coût par rapport au contrôle pour résoudre ii) et iii).

Une des difficultés des problèmes d'optimal design est qu'il n'existe pas dans l'arsenal classique de l'analyse d'ensemble de domaines muni d'une structure d'espace vectoriel normé. Nous allons définir des espaces de domaines munis d'une structure métrique complète et d'une "structure différentielle". Dans ces espaces nous obtiendrons l'existence d'un domaine optimal D_o (quand la famille $\mathcal{D}_{ad}$ est compacte), et la dérivabilité de la fonction $D \to J(D)$ (ce qui conduit aux conditions nécessaires d'optimalité du 1er ordre et à des méthodes de descente).

2. DEFINITION ET PROPRIETES METRIQUES DES ESPACES DE DOMAINES

Etant donné un ouvert connexe Q de R^N (on ne suppose pas de régularité sur Q) et k un entier, $k \geqslant 1$, on définit un espace de domaines homéomorphes à Q par :

$$\mathcal{D}_Q^{k,\infty} = \{ D \mid D = T(Q) \quad , \quad T \in \mathcal{C}^{k,\infty} \}$$



où $\mathcal{C}^{k,\infty}$ est l'espace de bijections régulières de R^N défini ainsi :

$$\mathcal{C}^{k,\infty} = \{ T \mid T \quad \text{bijection de } R^N \text{ sur lui-même} \quad T-I \in W^{k,\infty}(R^N,R^N) \quad \text{et} \quad T^{-1}-I \in W^{k,\infty}(R^N,R^N) \}$$

L'espace $\mathcal{C}^{k,\infty}$ est donc un ensemble de perturbations bornées, à dérivées bornées, de l'identité I de R^N. On munit par exemple $W^{k,\infty}(R^N,R^N)$ de la norme :

$$\|\varphi\|_{W^{k,\infty}(R^N,R^N)} = \underset{x\in R^N}{\text{Sup.ess.}} \{ \sum_{0\leq|\alpha|\leq k} |D^\alpha \varphi(x)|^2_{R^N} \}^{1/2}$$

et on définit sur $\mathcal{D}^{k,\infty}_Q \times \mathcal{D}^{k,\infty}_Q$ l'application :

$$d_{k,\infty}(D_1,D_2) = \inf_{\substack{T(D_1)=D_2 \\ T\in\mathcal{C}^{k,\infty}}} \{ \|T-I\|_{W^{k,\infty}(R^N,R^N)} + \|T^{-1}-I\|_{W^{k,\infty}(R^N,R^N)} \}$$

L'application $d_{k,\infty}$ est presque une distance, mais elle ne vérifie pas l'inégalité triangulaire. On a cependant les résultats suivants :

THEOREME 1

i) Il existe une constante positive η_k telle que l'application
$\hat{d}_{k,\infty} = \inf \{ \sqrt{d_{k,\infty}} , \eta_k \}$ soit une distance sur $\mathcal{D}^{k,\infty}_Q$.

ii) L'espace $\mathcal{D}^{k,\infty}_Q$ muni de cette métrique est complet .

iii) Si $k \geq 2$ et si Q est borné, l'injection de $\mathcal{D}^{k,\infty}_Q$ dans $\mathcal{D}^{k-1,\infty}_Q$ est compacte : Plus précisemment de toute suite $\{D_n\}$ de $\mathcal{D}^{k,\infty}_Q$ telle que $d_{k,\infty}(D_1,D_n) \leq$ cste on peut extraire une sous suite qui converge dans $\mathcal{D}^{k-1,\infty}_Q$ vers un élément $D_o \in \mathcal{D}^{k,\infty}_Q$. ■

3. UN RESULTAT D'EXISTENCE D'UN DOMAINE OPTIMAL

Utilisant les espaces de domaines munis de leur structure métrique introduits au paragraphe 2, on obtient pour l'exemple que nous avons choisi de traiter, le résultat d'existence :

THEOREME 2

Soit Q un ouvert de R^N borné, variété à bord de classe $W^{2,\infty}$. Etant donné $\mathcal{D}_{ad}$ sous ensemble borné de $\mathcal{D}_Q^{2,\infty}$, fermé pour la topologie induite par $d_{1,\infty}$, il existe un domaine D_o réalisant le minimum de J sur $\mathcal{D}_{ad}$, i.e. vérifiant (5). ∎

Principe de la démonstration

On utilise le théorème 1 iii) et on se ramène au domaine fixe Q par changement de variable. ∎

4. DERIVATION DANS LES ESPACES DE DOMAINES $\mathcal{D}_Q^{k,\infty}$

La notion de dérivée Fréchet s'introduit généralement dans un espace affine normé A d'espace vectoriel sous jacent E, et utilise l'addition, application de A × E dans A. Nous allons étendre la notion de dérivée Fréchet en faisant jouer à $W^{k,\infty}(R^N,R^N)$ le rôle d'espace vectoriel sous jacent à $\mathcal{D}_Q^{k,\infty}$, et à l'application définie par

(6) $\qquad (D,\Theta) \to (I+\Theta)(D)$

le rôle de l'addition. Cela est loisible car si Θ est assez petit dans $W^{k,\infty}(R^N,R^N)$, $(I+\Theta)$ appartient à $\mathcal{C}^{k,\infty}$ et $(I+\Theta)(D)$ est donc un élément de $\mathcal{D}_Q^{k,\infty}$ si $D \in \mathcal{D}_Q^{k,\infty}$. De plus l'application $\Theta \to (I+\Theta)(D)$ transforme toute base de voisinages de O dans $W^{k,\infty}(R^N,R^N)$ en une base de voisinages de D dans $\mathcal{D}_Q^{k,\infty}$. Cela nous conduit à poser les définitions :

DEFINITION 1

Une application J de $\mathcal{D}_Q^{k,\infty}$ dans R est dérivable au point $D_o \in \mathcal{D}_Q^{k,\infty}$ si l'application $\Theta \to J^\star(\Theta) = J((I+\Theta)(D_o))$ est Fréchet dérivable (au sens usuel) de $W^{k,\infty}(R^N,R^N)$ dans R au point O. On dira que la dérivée de J au point D_o est

$$\frac{\partial J}{\partial D}(D_o) = \frac{\partial J^\star}{\partial \Theta}(0) \in \mathcal{L}_c(W^{k,\infty}(R^N,R^N);R)$$ ∎

DEFINITION 2

Une application G de R dans $\mathcal{D}_Q^{k,\infty}$ est dérivable au point r_o s'il existe une application ${}^\star G$ de R dans $W^{k,\infty}(R^N,R^N)$ telle que

$G(r) = (I+{}^\star G(r))(G(r_o))$ pour tout r dans un voisinage de r_o

${}^\star G(r_o) = 0$

${}^\star G$ est dérivable (au sens usuel) en r_o.

On dira alors que G admet comme dérivée au point r_o

$$\frac{\partial^{\star} G}{\partial r}(r_o) \in \mathcal{L}_c(R;W^{k,\infty}(R^N,R^N)) \equiv W^{k,\infty}(R^N,R^N) \quad \blacksquare$$

Ces définitions entrainent un certain nombre de propriétés analogues à celles des fonctions dérivables au sens usuel ; Ainsi une application dérivable est localement lipschitzienne, et la composée de 2 applications dérivables est dérivable. Par contre "l'addition" définie par (6) n'est pas injective, et la "soustraction" correspondante est multivoque, ce qui entraine les propriétés spécifiques suivantes :

PROPRIETE 1

Soit J une application de $\mathcal{D}_Q^{k,\infty}$ dans R, dérivable en un point $D_o \in \mathcal{D}_Q^{k,\infty}$ que l'on suppose ouvert borné de R^N, variété à bord de classe C^1 ; Si Θ_1 et Θ_2 sont des éléments de $W^{k,\infty}(R^N,R^N)$ tels que

$$\Theta_1-\Theta_2 \in C^k(R^N,R^N)$$

$$\langle n,\Theta_1\rangle = \langle n,\Theta_2\rangle \quad \text{sur } \partial D_o$$

(n désignant la normale extérieure et $\langle,\rangle$ le produit scalaire dans R^N, de sorte que $\langle n,\Theta\rangle = \sum_{i=1}^{N} n_i\, \Theta_i$), alors

$$\frac{\partial J}{\partial D}(D_o)\cdot\Theta_1 = \frac{\partial J}{\partial D}(D_o)\cdot\Theta_2$$

Cette première propriété peut se résumer simplement en disant que la dérivée $\frac{\partial J}{\partial D}(D_o)\cdot\Theta$ ne dépend de Θ presque que par la valeur de sa trace normale $\langle n,\Theta\rangle$ sur ∂D_o.

De la définition 2, il résulte qu'une application G dérivable en r_o peut admettre plusieurs dérivées en ce point. Si l'on note $\frac{\partial G}{\partial r}(e_o) \subset W^{k,\infty}(R^N,R^N)$ l'ensemble de ces dérivées, on peut presque caractériser cet ensemble. En effet :

PROPRIETE 2

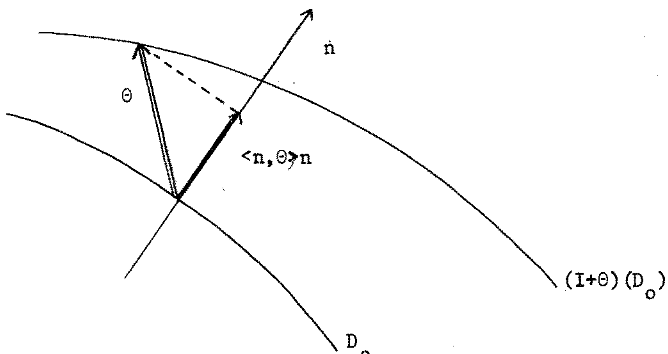
Soit G une application de $\mathcal{D}_Q^{k,\infty}$ dans R, dérivable en un point r_o, et soit $\gamma \in \frac{\partial G}{\partial r}(r_o)$; on suppose $D_o = G(r_o)$ ouvert borné de R^N, variété à bord de classe C^1. Alors

$$\frac{\partial G}{\partial r}(r_o) \subset \{\gamma+\eta \mid \eta \in W^{k,\infty}(R^N,R^N) \ , \ \langle n,\eta\rangle = 0 \text{ sur } \partial D_o\}$$

D'autre part si $\gamma_1 \in \frac{\partial G}{\partial r}(r_o)$ appartient à $C^k(R^N,R^N)$, on a

$$\{ \gamma_1+\eta_1 \mid \eta_1 \in W^{k,\infty}(R^N,R^N) \cap C^k(R^N,R^N) \; \langle n,\eta_1\rangle = 0 \text{ sur } \partial D_o\} \subset \frac{\partial G}{\partial r}(r_o). \quad \blacksquare$$

Ces propriétés s'interprètent géométriquement en observant que, quand Θ_1 et Θ_2 sont voisins de 0 dans $W^{k,\infty}(R^N,R^N)$, $(I+\Theta_1)(D_o)$ et $(I+\Theta_2)(D_o)$ définissent le même domaine au second ordre près, "si et seulement si" $\langle n,\theta_1\rangle = \langle n,\Theta_2\rangle$ sur ∂D_o.



5. CONDITIONS NECESSAIRES D'OPTIMALITE

Les espaces de domaines $\mathcal{D}_Q^{k,\infty}$ étant munis de la "structure différentielle" définie au paragraphe 4, on peut donner des conditions nécessaires d'optimalité générales :

THEOREME 3

Soit $\mathcal{D}_{ad}$ un sous ensemble de $\mathcal{D}_Q^{k,\infty}$ et J une fonction de $\mathcal{D}_{ad}$ dans R. Si D_o réalise le minimum de J sur $\mathcal{D}_{ad}$ et si J est dérivable au point D_o, alors pour tout Θ appartenant à l'espace tangent à $\mathcal{D}_{ad}$ en D_o, on a

(7) $$\frac{\partial J}{\partial D}(D_o) \cdot \theta \geq 0$$

où l'espace tangent à $\mathcal{D}_{ad}$ en D_o est l'ensemble des $\Theta \in W^{k,\infty}(R^N,R^N)$, tels qu'il existe une application G de $[0,1]$ dans $\mathcal{D}_Q^{k,\infty}$, dérivable en zéro, telle que $G(0)=D_o$, $G(h) \in \mathcal{D}_{ad}$ et $\theta \in \frac{\partial G}{\partial h}(0)$. ■

L'espace tangent à $\mathcal{D}_{ad}$ en D_o contient en particulier l'ensemble des $\Theta \in W^{k,\infty}(R^N,R^N)$ tels que $(I+h\Theta)(D_o)$ soit dans $\mathcal{D}_{ad}$ pour h assez petit.

6. DERIVABILITE D'UNE INTEGRALE SUPERFICIELLE PAR RAPPORT A SON DOMAINE D'INTEGRATION

Avant d'énoncer le théorème de dérivabilité de la fonctionnelle associée au problème de Neumann par (4), donnons un théorème de dérivabilité d'une fonctionnelle plus simple définie par une intégrale superficielle, et qui éclaire le résultat et le principe de la démonstration.

On suppose Q ouvert de R^N borné, variété à bord de classe $W^{2,\infty}$. Alors tout $D \in \mathcal{D}_Q^{2,\infty}$ présente les mêmes propriétés. Soit $\tilde{f} \in W^{2,1}(R^N,R)$; on définit la fonctionnelle $\tilde{J}$ par

$$(8) \qquad \tilde{J}(D) = \int_{\partial D} \tilde{f}(s)\cdot ds \qquad \forall\, D \in \mathcal{D}_Q^{2,\infty}$$

THEOREME 4

L'application $\tilde{J}$ définie par (8) est dérivable de $\mathcal{D}_Q^{2,\infty}$ dans R en tout point $D_o \in \mathcal{D}_Q^{2,\infty}$ et l'on a

$$(9) \qquad \forall\, \tau \in W^{2,\infty}(R^N,R^N) \qquad \frac{\partial \tilde{J}}{\partial D}(D_o)\cdot\tau = \int_{\partial D_o} \langle n,\tau\rangle \left(\frac{\partial \tilde{f}}{\partial n} + Hf\right)$$

où n désigne la normale extérieure, $\frac{\partial \tilde{f}}{\partial n}$ la dérivée normale de $\tilde{f}$ et H la courbure moyenne de la variété ∂D_o. ∎

Principe de la démonstration

D'après la définition 1, il faut montrer que l'application $\tilde{J}^\star$ définie par

$$\tilde{J}^\star(\Theta) = \int_{\partial(I+\Theta)(D_o)} \tilde{f}(s)\cdot ds$$

est dérivable en zéro. Pour cela on effectue le changement de variable $(I+\Theta)$. On montre que :

$$\tilde{J}^\star(\Theta) = \int_{\partial D_o} \tilde{f}\circ(I+\Theta)\cdot |\det(I+\Theta)'|\cdot|{}^t(I+\Theta)'^{-1}n|\, ds$$

où $(I+\Theta)'$ designe la matrice dérivée de $(I+\Theta) \in C^1(R^N,R^N)$.

Dans cette expression, on montre la dérivabilité de chacun des trois termes du produit par rapport à Θ, ce qui démontre la dérivabilité de $\tilde{J}$ et donne la valeur de la dérivée :

$$\begin{cases} \forall\ \tau \in W^{2,\infty}(R^N,R^N) \\ \dfrac{\partial \tilde{J}}{\partial D}(D_o)\cdot\tau = \displaystyle\int_{\partial D_o} \{ <\text{grad}\ \tilde{f},\tau> + \tilde{f}(\text{div}\ \tau - <(\tau)'n,n> \}ds \end{cases}$$

Une intégration par parties sur ∂D_o, analogue à la formule de Stokes sur un ouvert, conduit alors à (9). ∎

7. RESULTAT DE DERIVABILITE POUR LE PROBLEME DE NEUMANN

THEOREME 5.

On suppose Q ouvert borné de R^N, variété à bord de classe $W^{3,\infty}$. Alors la fonction coût définie par (4) est dérivable de $\mathcal{D}_Q^{3,\infty}$ dans R en tout point $D_o \in \mathcal{D}_Q^{3,\infty}$ et sa dérivée est donnée par :

$$(10) \qquad \forall\ \Theta \in W^{3,\infty}(R^N,R^N) \qquad \frac{\partial J}{\partial D}(D_o)\cdot\Theta = \int_{\partial D_o} <n,\Theta> \cdot F(u_{D_o},p_{D_o}) \cdot ds$$

où u_{D_o} est l'état du système défini par (3), où p_{D_o} est l'état adjoint défini par :

$$(11) \qquad \begin{cases} -\Delta p_{D_o} + p_{D_o} = 0 & \text{dans } D_o \\ \dfrac{\partial p_{D_o}}{\partial n} = 2(u_{D_o}-z) & \text{sur } \partial D_o \\ p_{D_o} \in H^2(D_o) \end{cases}$$

et où F est la fonctionnelle définie par

$$(12) \qquad F(u,p) = -<\text{grad}\ u,\text{grad}\ p> + p(H.g + \frac{\partial g}{\partial n} + f-u) + (u-z).(H.(u-z)+4g-2\ \frac{\partial z}{\partial n})$$

(H est la courbure moyenne de la variété ∂D_o). ∎

Remarquons que, comme l'indique la propriété 1, la dérivée ne dépend de Θ que par sa trace normale $<n,\Theta>$.

Ce résultat de dérivabilité permet d'obtenir immédiatement les conditions nécessaires d'optimalité du premier ordre, grâce au théorème 3. Il permet également d'envisager l'emploi de méthodes de descente pour le calcul d'un domaine optimal.

<u>Principe de la démonstration du théorème 5.</u>

Comme dans la démonstration du théorème 4, on effectue le changement de variable $(I+\Theta)$ dans l'intégrale superficielle $J^{*}(\Theta) = J((I+\Theta)(D_o))$.

On est alors conduit à démontrer la dérivabilité par rapport à Θ d'un certain nombre de termes, et notamment la dérivabilité en zéro de l'application :

$$\Theta \to (u_{(I+\Theta)(D_o)}) \circ (I+\Theta) \qquad \text{de} \quad W^{3,\infty}(R^N,R^N) \text{ dans } H^2(D_o)$$

ce qui constitue un résultat de dérivabilité de l'état du système par rapport au domaine. Pour cela, on effectue le changement de variable $(I+\Theta)$ dans l'équation (3) relative à $(I+\Theta)(D_o)$ et on utilise le théorème de dérivabilité de la fonction implicite.

On termine la démonstration par des intégrations par parties sur D_o et ∂D_o, ce qui conduit à (10). ■