MASTER Mathématiques 2ème année, 2011-2012, Marseille.
Spécialité EDP et Calcul Scientifique.
Notes de cours autorisées - Durée : 4h

ANALYSE NUMÉRIQUE DES EDP

Etude d'un problème elliptique non-linéaire

Partie I- Résultat préliminaire sur les opérateurs non-linéaires monotones :

On se donne un espace de Hilbert V (son produit scalaire est noté $(.,.)_V$) et $J : V \to V$ un opérateur *a priori* non linéaire. On suppose que J vérifie

$$\exists C_1 > 0,\ \ \|J(u_1) - J(u_2)\|_V \leq C_1 \|u_1 - u_2\|_V,\ \ \forall u_1, u_2 \in V, \tag{1}$$

$$\exists \alpha > 0,\ \ (J(u_1) - J(u_2), u_1 - u_2)_V \geq \alpha \|u_1 - u_2\|_V^2,\ \ \forall u_1, u_2 \in V. \tag{2}$$

Le but de cette partie est de montrer que J est bijectif et que son inverse est Lipschitzien.

1. Soit $v \in V$. Montrer qu'il existe $\rho > 0$ tel que l'application
$$\psi_v : u \in V \mapsto u - \rho(J(u) - v) \in V,$$
soit contractante dans V.
2. En déduire que J est bijectif.
3. Montrer que J^{-1} vérifie
$$\|J^{-1}(u_1) - J^{-1}(u_2)\|_V \leq \frac{1}{\alpha}\|u_1 - u_2\|_V.$$

Partie II- Existence et unicité de la solution d'un problème elliptique non-linéaire :

On notera $a \cdot b$ le produit scalaire de deux vecteurs $a, b \in \mathbb{R}^d$ et $|a| = \sqrt{a \cdot a}$ la norme euclidienne associée. On se donne :

- Une application $\varphi : \mathbb{R}^d \to \mathbb{R}^d$ vérifiant les propriétés
$$|\varphi(\xi_1) - \varphi(\xi_2)| \leq \mathrm{Lip}(\varphi)|\xi_1 - \xi_2|,\ \ \forall \xi_1, \xi_2 \in \mathbb{R}^d,$$
$$(\varphi(\xi_1) - \varphi(\xi_2)) \cdot (\xi_1 - \xi_2) \geq \alpha |\xi_1 - \xi_2|^2,\ \ \forall \xi_1, \xi_2 \in \mathbb{R}^d.$$
- Une application $\varphi_0 : \mathbb{R} \to \mathbb{R}$ vérifiant les propriétés
$$|\varphi_0(s) - \varphi_0(t)| \leq \mathrm{Lip}(\varphi_0)|s - t|,\ \ \forall s, t \in \mathbb{R},$$
φ_0 est croissante au sens large.

Soit Ω un ouvert borné, régulier, connexe de $\mathbb{R}^d$ ($d = 2$ ou $d = 3$). Pour une fonction $f \in L^2(\Omega)$ donnée, on s'intéresse au problème suivant : trouver une fonction $u : \Omega \to \mathbb{R}$ vérifiant

$$\begin{cases} -\mathrm{div}\,(\varphi(\nabla u)) + \varphi_0(u) = f, & \text{dans } \Omega \\ u = 0, & \text{sur } \partial\Omega. \end{cases} \tag{$\mathcal{P}$}$$

On se place dans l'espace fonctionnel $V = H_0^1(\Omega)$. On note $(.,.)_V$ (resp. $\|.\|_V$) le produit scalaire (resp. la norme) sur cet espace dont on rappelle la définition

$$(u, v)_V = \int_\Omega \nabla u \cdot \nabla v\, dx,\ \ \|u\|_V = \sqrt{(u,u)_V} = \|\nabla u\|_{L^2}. \tag{3}$$

On notera $C_P > 0$ la constante de Poincaré, c'est-à-dire la plus petite constante pour laquelle l'inégalité suivante est vraie

$$\|u\|_{L^2} \leq C_P \|u\|_V, \ \ \forall u \in V.$$

On introduit maintenant

$$a(u,v) = \int_\Omega \varphi(\nabla u) \cdot \nabla v \, dx + \int_\Omega \varphi_0(u) v \, dx, \ \ \forall u, v \in V,$$

$$L(v) = \int_\Omega f v \, dx, \ \ \forall v \in V.$$

On prendra garde au fait que a **n'est pas linéaire par rapport à sa première variable !**.

1. Démontrer que $a(u,v)$ est bien définie pour tout $u, v \in V$ et que pour tout $u \in V$, l'application

   $$v \in V \mapsto a(u,v)$$

   est continue.
2. Montrer qu'une fonction $u \in H_0^1(\Omega)$ est solution du problème $(\mathcal{P})$ (au sens des distributions) si et seulement on a

   $$a(u,v) = L(v), \ \ \forall v \in V. \tag{4}$$

   On va donc chercher à montrer qu'il existe une unique solution u au problème (4).
3. Montrer que pour tout $u \in V$, il existe un unique $J(u) \in V$ tel que

   $$a(u,v) = (J(u), v)_V, \ \ \forall v \in V.$$

4. Montrer que l'application J ainsi construite vérifie (1) et (2) avec la constante $C_1 = \mathrm{Lip}(\varphi) + \mathrm{Lip}(\varphi_0) C_P^2$.
5. Conclure.

Partie III- Approximation de Galerkin du problème :

Soit V_h un sous-espace de dimension finie de V. On considère le problème approché suivant : trouver $u_h \in V_h$ tel que

$$a(u_h, v_h) = L(v_h), \ \ \forall v_h \in V_h. \tag{5}$$

1. Expliquer, sans nécessairement refaire tous les calculs, comment adapter ce qui a été fait dans la partie précédente pour montrer que le problème (5) admet une unique solution dans V_h.
2. Démontrer que pour tout $v_h \in V_h$, on a

   $$a(u, u-u_h) - a(u_h, u-u_h) = a(u, u-v_h) - a(u_h, u-v_h).$$

3. En déduire que l'on a

   $$\|u-u_h\|_V \leq \frac{\mathrm{Lip}(\varphi) + \mathrm{Lip}(\varphi_0) C_P^2}{\alpha} \|u-v_h\|_V.$$

   Indication : on pourra récrire la formule de la question précédente à l'aide de l'opérateur J défini dans la partie 2.
4. En déduire que, si $d(u, V_h) \to 0$ quand $h \to 0$, alors on a

   $$u_h \xrightarrow[h\to 0]{} u, \ \text{ dans } V.$$

5. Soit $(\psi_i)_{1\leq i \leq N_h}$ une base de V_h. On cherche la solution u_h sous la forme $u_h = \sum_{i=1}^{N_h} u_i \psi_i$. Ecrire les équations satisfaites par les u_i.

   Quelle est la difficulté nouvelle qui se présente ici par rapport aux problèmes étudiés en cours ? Quelle méthode classique pourriez-vous envisager pour la résoudre ? On ne demande pas une réponse très détaillée.

   La dernière partie du problème sera consacrée à l'étude d'une méthode itérative adaptée à la résolution de ce problème.

Partie IV- Approximation $\mathbb{P}^1$ en dimension 2. Estimations d'erreur *a posteriori* :

On suppose ici que Ω est un domaine polygonal de $\mathbb{R}^2$ et qu'on dispose d'une triangulation géométriquement conforme $\mathcal{T} = (K)_{K\in\mathcal{T}}$. On considère l'espace d'approximation de Galerkin suivant

$$V_h = \{v \in H_0^1(\Omega),\ \ \forall K \in \mathcal{T}, v_{|K} \in \mathbb{P}^1\}.$$

On introduit également l'espace

$$G_h = \{g \in (L^2(\Omega))^d,\ \ \forall K \in \mathcal{T}, g_{|K} \in (\mathbb{P}^0)^d\},$$

dont on remarquera qu'il vérifie :

$$\nabla v_h \in G_h,\ \ \forall v_h \in V_h.$$

Les résultats de la partie précédente donnent en particulier une borne de l'erreur d'approximation $\|u - u_h\|_V$ en fonction de constantes indépendantes de h et de la quantité $d(u, V_h)$. Cette dernière quantité dépend de la solution u que l'on ne connaît évidemment pas. On a vu en cours que si on suppose une certaine régularité pour u et pour la famille de maillages considérés, alors on peut estimer $d(u, V_h)$ en fonction de puissances de h et de normes adéquates de u.
Dans cette partie du problème on va essayer de trouver une borne pour l'erreur qui ne dépende que de la solution approchée u_h (et non pas de u). De sorte que, aux constantes universelles près, cette borne sera effectivement calculable dans un code de calcul et donnera une idée *a posteriori* de l'erreur commise.
On notera $\mathcal{E} = (\sigma)_{\sigma\in\mathcal{E}}$ l'ensemble de toutes les arêtes du maillage, $\mathcal{E}_{int}$ l'ensemble des arêtes intérieures (i.e. qui ne sont pas sur le bord de Ω). On notera également $\mathcal{E}_K$ l'ensemble des arêtes d'un élément K. On notera h_σ la longueur d'une arête $\sigma \in \mathcal{E}$ et h_K le diamètre de tout élément $K \in \mathcal{T}$.
Pour toute arête intérieure $\sigma \in \mathcal{E}_{int}$ qui sépare deux éléments K et L, on note $n_{K,\sigma}$ (resp. $n_{L,\sigma}$) la normale sortante à K (resp. à L) sur σ, de sorte que l'on a

$$n_{K,\sigma} = -n_{L,\sigma}. \tag{6}$$

On rappelle la formule de Stokes suivante

$$\int_K (\nabla f)\, dx = \sum_{\sigma\in\mathcal{E}_K} \int_\sigma f n_{K,\sigma}\, dx,$$

pour toute fonction $f : K \to \mathbb{R}$ régulière.
Pour toute fonction $g_h \in G_h$, on introduit le saut dans la direction normale défini par

$$[\![g_h]\!]_\sigma = (g_h)_{|K} \cdot n_{K,\sigma} + (g_h)_{|L} \cdot n_{L,\sigma} \in \mathbb{R}.$$

Il s'agit bien d'un terme de saut d'après (6). On rappelle que g_h est constant sur chaque élément du maillage, ce qui justifie la définition ci-dessus.
On utilisera dans la suite l'existence d'un opérateur d'interpolation $I_h : V \to V_h$ (par exemple construit selon la méthode de Scott-Zhang vue en cours) vérifiant les propriétés suivantes ($C > 0$ indépendante de h)

$$I_h(v_h) = v_h,\ \ \forall v_h \in V_h,$$

$$\|I_h v\|_V \leq C\|v\|_V,\ \ \forall v \in V,$$

$$\sum_{K\in\mathcal{T}} \frac{1}{h_K^2}\|v - I_h v\|^2_{L^2(K)} \leq C\|v\|^2_V.$$

$$\sum_{\sigma\in\mathcal{E}} \frac{1}{h_\sigma}\|v - I_h v\|^2_{L^2(\sigma)} \leq C\|v\|^2_V.$$

On introduit l'opérateur d'interpolation $\mathbb{P}^0$ défini, pour toute fonction $v \in L^2(\Omega)$ par

$$(I_h^0 v)_{|K} = \frac{1}{|K|}\int_K v(x)\, dx,\ \ \forall K \in \mathcal{T},$$

et dont les propriétés ont été vues en cours.
Pour terminer l'introduction des notations, on notera

$$\text{reg}(\mathcal{T}) = \sup_{K\in\mathcal{T}} \frac{h_K}{\rho_K},$$

la constante de régularité du maillage $\mathcal{T}$ (dans le cours celle-ci était notée $\sigma_\mathcal{T}$ mais, ici, cette notation pourrait engendrer des confusions avec celle utilisée ici pour désigner les arêtes du maillage).

1. Soit $v \in V$ quelconque. Montrer, en utilisant $I_h v$ comme fonction test dans le problème (5), que l'on a

$$a(u,v) - a(u_h,v) = \sum_{K\in\mathcal{T}} \left(\int_K (f - \varphi_0(u_h))(v - I_h v)\, dx \right) - \sum_{\sigma\in\mathcal{E}} \left(\int_\sigma [\![\varphi(\nabla u_h)]\!]_\sigma (v - I_h v)\, dx \right).$$

2. En déduire que

$$\|u - u_h\|_V^2 \leq C \left(\sum_{K\in\mathcal{T}} h_K^2 \|f - \varphi_0(u_h)\|_{L^2(K)}^2 + \sum_{\sigma\in\mathcal{E}} h_\sigma^2 [\![\varphi(\nabla u_h)]\!]_\sigma^2 \right). \tag{7}$$

   On a donc bien une majoration de l'erreur en fonction de quantités calculables à partir des données du problème (f, φ_0 et φ) et de la solution approchée u_h. Pour l'instant rien ne nous dit que cette majoration n'est pas trop grossière. On va montrer dans la suite que ça n'est pas le cas, au moins pour le premier des deux termes du membre de droite de (7). L'autre terme pourrait se traiter de manière analogue mais on ne le fera pas ici.

3. On fixe dans toute la suite une fonction $\hat{b}$ de classe $\mathcal{C}^1$, définie sur le triangle de référence $\hat{K}$, et qui vérifie

$$0 \leq \hat{b} \leq 1,$$
$$\int_{\hat{K}} \hat{b}\, d\hat{x} > 0,$$
$$\hat{b} = 0, \text{ sur } \partial\hat{K}.$$

   Pour tout élément $K \in \mathcal{T}$, on note $T_K : \hat{K} \to K$ la transformation affine qui envoie l'élément de référence sur K et on définit maintenant la fonction b_K par

$$b_K = \begin{cases} \hat{b} \circ (T_K)^{-1}, & \text{dans } K, \\ 0, & \text{en dehors de } K. \end{cases}$$

   (a) Montrer que b_K est identiquement nulle sur ∂K. En déduire que $b_K \in H_0^1(\Omega)$.

   (b) Montrer que
$$\int_K \nabla b_K\, dx = 0.$$

   (c) Montrer qu'il existe $C_1 > 0$ indépendant de K telle que
$$0 \leq b_K \leq 1, \text{ sur } K,$$
$$\int_K b_K(x)\, dx = C_1 |K|.$$

   Montrer qu'il existe C_2 dépendant seulement de $\mathrm{reg}(\mathcal{T})$ telle que
$$\|\nabla b_K\|_{L^2(K)} \leq \frac{C_2 |K|^{1/2}}{h_K}.$$

   (d) On note $R_h = f - \varphi_0(u_h) \in L^2(\Omega)$. On se fixe un $K \in \mathcal{T}$.

      i. Montrer que $\|R_h\|_{L^2(K)} \leq \|I_h^0 R_h\|_{L^2(K)} + \|R_h - I_h^0 R_h\|_{L^2(K)}$.

      ii. Montrer que
$$C_1 \|I_h^0 R_h\|_{L^2(K)}^2 \leq (R_h, b_K I_h^0 R_h) + \|R_h - I_h^0 R_h\|_{L^2(K)} \|I_h^0 R_h\|_{L^2(K)}.$$

      iii. En utilisant $b_K I_h^0 R_h \in V$ comme fonction-test dans (4) et la définition de R_h, montrer que
$$(R_h, b_K I_h^0 R_h) = \int_K (\varphi(\nabla u) - \varphi(\nabla u_h)) \cdot \nabla(b_K I_h^0 R_h)\, dx$$
$$+ \int_K (\varphi_0(u) - \varphi_0(u_h)) b_K I_h^0 R_h\, dx.$$

iv. En déduire que

$$(R_h, b_K I_h^0 R_h) \leq \frac{C}{h_K} \|\nabla(u-u_h)\|_{L^2(K)} \|I_h^0 R_h\|_{L^2(K)} + C\|u-u_h\|_{L^2(K)} \|I_h^0 R_h\|_{L^2(K)}.$$

v. Déduire enfin de tout cela que

$$\sum_{K\in\mathcal{T}} h_K^2 \|R_h\|_{L^2(K)}^2 \leq C\|u-u_h\|_V^2 + C \sum_{K\in\mathcal{T}} h_K^2 \|R_h - I_h^0 R_h\|_{L^2(K)}^2. \tag{8}$$

(e) Si on fait un instant abstraction du dernier terme, l'inégalité précédente donne presque le résultat escompté, c'est-à-dire que le terme calculable $\sum_{K\in\mathcal{T}} h_K^2 \|R_h\|_{L^2(K)}^2$ ne peut pas être beaucoup plus gros que le carré de l'erreur $\|u-u_h\|_V^2$.
On suppose maintenant que $f \in H^1(\Omega)$ et que φ_0 est de classe $\mathcal{C}^1$. Montrer que

$$\sum_{K\in\mathcal{T}} h_K^2 \|R_h - I_h^0 R_h\|_{L^2(K)}^2 \leq C \sum_{K\in\mathcal{T}} h_K^4 (\|\nabla f\|_{L^2(K)}^2 + \|\nabla u_h\|_{L^2(K)}^2).$$

Expliquer pourquoi on peut raisonnablement dire que le dernier terme dans (8) est négligeable devant le terme $\|u-u_h\|_V^2$.

4. A partir du maillage $\mathcal{T}$ et de la solution approchée u_h calculée sur ce maillage. Comment construiriez-vous un nouveau maillage un peu plus fin sur lequel la nouvelle solution approchée serait probablement une meilleure approximation de la solution u ?
On ne demande pas une réponse très détaillée.

Partie V- Résolution par décomposition-coordination :

Important : Dans cette dernière partie on suppose, pour simplifier un peu, que $\varphi_0 = 0$.

On se donne un nombre $r > 0$. On notera $(u, v) = \int_\Omega uv\, dx$ le produit scalaire de $L^2(\Omega)$ (ou de $(L^2(\Omega))^d$ le cas échéant).

1. En guise de préliminaire, démontrer que l'application $g \in \mathbb{R}^d \mapsto \varphi(g) + rg \in \mathbb{R}^d$ est bijective.
2. Démontrer que $u_h \in V_h$ est solution de (5) si et seulement s'il existe $g_h, \lambda_h \in G_h$ tels que

$$r(u_h, v_h)_V = (f, v_h) + (\lambda_h + rg_h, \nabla v_h), \ \ \forall v_h \in V_h, \tag{9}$$

$$\varphi(g_h) + r(g_h - \nabla u_h) + \lambda_h = 0, \tag{10}$$

$$g_h - \nabla u_h = 0, \tag{11}$$

Montrer qu'il existe un unique triplet $(u_h, g_h, \lambda_h) \in V_h \times G_h \times G_h$ solution de ce nouveau système (9)-(10).

3. On propose une méthode itérative pour résoudre les équations (9)-(11). On se donne $g_h^0, \lambda_h^0 \in G_h$ (par exemple les fonctions nulles ...) et pour tout $n \geq 1$, on définit $u_h^n \in V_h$, $g_h^n, \lambda_h^n \in G_h$ par les équations

$$r(u_h^n, v_h)_V = (f, v_h) + (\lambda_h^{n-1} + rg_h^{n-1}, \nabla v_h), \ \ \forall v_h \in V_h, \tag{12}$$

$$\varphi(g_h^n) + r(g_h^n - \nabla u_h^n) + \lambda_h^{n-1} = 0, \tag{13}$$

$$\lambda_h^n = \lambda_h^{n-1} + r(g_h^n - \nabla u_h^n), \tag{14}$$

(a) Si u_h^{n-1}, g_h^{n-1} et λ_h^{n-1} sont connus. Montrer qu'il existe un unique $u_h^n \in V_h$ solution de (12). Quelle est la nature du problème discret que l'on doit résoudre dans cette étape. Commentez.

(b) La fonction u_h^n étant maintenant calculée, montrer qu'il existe un unique $g_h^n \in G_h$ vérifiant (13). On pourra raisonner élément par élément. Comment, en pratique, calculer g_h^n ?

(c) Montrer enfin que (14) définit bien un unique λ_h^n dans G_h.

4. On veut montrer que les suites $(g_h^n)_n$, $(\lambda_h^n)_n$ et $(u_h^n)_n$ convergent respectivement vers g_h, λ_h et u_h. Pour cela on pose

$$v_h^n = u_h^n - u_h,$$

$$\mu_h^n = \lambda_h^n - \lambda_h,$$

$$\delta_h^n = g_h^n - g_h.$$

(a) Justifier la formule

$$\frac{1}{2}\|a\|_{L^2}^2 - \frac{1}{2}\|b\|_{L^2}^2 - \frac{1}{2}\|a-b\|_{L^2}^2 = (b, a-b),\ \forall a, b \in L^2(\Omega), \tag{15}$$

puis montrer que

$$\frac{1}{2r}\|\mu_h^n\|_{L^2}^2 - \frac{1}{2r}\|\mu_h^{n-1}\|_{L^2}^2 - (\mu_h^{n-1}, \delta_h^n - \nabla v_h^n) - \frac{r}{2}\|\delta_h^n - \nabla v_h^n\|_{L^2}^2 = 0.$$

(b) Montrer que

$$\varphi(g_h^n) - \varphi(g_h) + r(\delta_h^n - \nabla v_h^n) + \mu_h^{n-1} = 0,$$

et que

$$r\|\nabla v_h^n\|_{L^2}^2 = (\mu_h^{n-1} + r\delta_h^{n-1}, \nabla v_h^n).$$

En déduire que

$$(\varphi(g_h^n) - \varphi(g_h), g_h^n - g_h) + r\|\delta_h^n - \nabla v_h^n\|_{L^2}^2 + (\mu_h^{n-1}, \delta_h^n - \nabla v_h^n) + r(\delta_h^n - \delta_h^{n-1}, \nabla v_h^n) = 0.$$

(c) Montrer que

$$\varphi(g_h^n) - \varphi(g_h^{n-1}) + r(\delta_h^n - \delta_h^{n-1}) + r\delta_h^{n-1} - r\nabla v_h^n = 0.$$

En déduire que

$$(\varphi(g_h^n) - \varphi(g_h^{n-1}), g_h^n - g_h^{n-1}) + r\|\delta_h^n - \delta_h^{n-1}\|_{L^2}^2 + r(\delta_h^{n-1}, \delta_h^n - \delta_h^{n-1}) - r(\nabla v_h^n, \delta_h^n - \delta_h^{n-1}) = 0.$$

(d) En utilisant à nouveau (15) montrer que

$$\frac{r}{2}\|\delta_h^n\|_{L^2}^2 - \frac{r}{2}\|\delta_h^{n-1}\|_{L^2}^2 + \frac{r}{2}\|\delta_h^n - \delta_h^{n-1}\|_{L^2}^2 + (\varphi(g_h^n) - \varphi(g_h^{n-1}), g_h^n - g_h^{n-1}) - r(\nabla v_h^n, \delta_h^n - \delta_h^{n-1}) = 0.$$

(e) Déduire de tout ce qui précède que

$$\begin{aligned}\frac{1}{2}\left(r\|\delta_h^n\|_{L^2}^2 + \frac{1}{r}\|\mu_h^n\|_{L^2}^2\right) - \frac{1}{2}\left(r\|\delta_h^{n-1}\|_{L^2}^2 + \frac{1}{r}\|\mu_h^{n-1}\|_{L^2}^2\right)\\ + (\varphi(g_h^n) - \varphi(g_h^{n-1}), g_h^n - g_h^{n-1}) + (\varphi(g_h^n) - \varphi(g_h), g_h^n - g_h)\\ + \frac{r}{2}\|\delta_h^n - \delta_h^{n-1}\|_{L^2}^2 + \frac{r}{2}\|\delta_h^n - \nabla v_h^n\|_{L^2}^2 = 0.\end{aligned}$$

(f) Montrer maintenant que la série numérique

$$\sum_{n\geq 1}(\varphi(g_h^n) - \varphi(g_h), g_h^n - g_h)$$

est convergente. En déduire que la suite $(g_h^n)_n$ converge vers g_h dans G_h.

(g) Montrer que la série numérique

$$\sum_{n\geq 1}\|\delta_h^n - \nabla v_h^n\|_{L^2}^2$$

est convergente. En déduire que la suite $(u_h^n)_n$ converge vers u_h dans $H_0^1(\Omega)$.

(h) Conclure enfin que $(\lambda_h^n)_n$ converge vers λ_h.

5. Quels sont les avantages et inconvénients de cette méthode itérative par rapport aux méthodes de résolution du système (5) que vous avez éventuellement envisagées à la fin de la partie III ?

FIN DU PROBLÈME
