

# ANALYSE NUMÉRIQUE

Prof. HAMRI Nasr-eddine  
Département de Mathématiques  
Université Abdelhafid Boussouf - Mila





# TABLE DES MATIÈRES

<b>1</b>	<b>NOTIONS D'ERREURS</b>	<b>7</b>
1	PRÉLIMINAIRES . . . . .	8
1.1	Exemples . . . . .	9
2	Erreurs absolues et Erreurs relatives . . . . .	9
2.1	Exemple . . . . .	10
2.2	Exemples . . . . .	10
3	PRINCIPALES SOURCES D'ERREURS . . . . .	11
4	PRECISION, CHIFFRES SIGNIFICATIFS . . . . .	11
4.1	Chiffres significatifs . . . . .	11
5	Cumulation des erreurs d'arrondi . . . . .	12
5.1	Erreurs d'arrondi sur une somme . . . . .	12
5.2	Erreurs d'arrondi sur un produit . . . . .	13
6	Représentation approchée des nombres réels . . . . .	13
6.1	Nombres en virgule flottante . . . . .	14
6.2	Non-associativité des opérations arithmétiques. . . . .	14
6.3	Phénomènes de compensation. . . . .	15
7	SERIE D'EXERCICES . . . . .	15
8	EXERCICES . . . . .	16
<b>2</b>	<b>APPROXIMATION</b>	<b>19</b>
1	GÉNÉRALITÉS . . . . .	19
2	APPROXIMATION . . . . .	19
2.1	Meilleure approximation . . . . .	19
3	APPROXIMATION AU SENS DES MOINDRES CARRÉS . . . . .	21
4	CARACTÉRISATION . . . . .	22
4.1	Norme . . . . .	22
5	SERIE D'EXERCICES . . . . .	24
<b>2</b>	<b>APPROXIMATION</b>	<b>19</b>
1	GÉNÉRALITÉS . . . . .	20
2	APPROXIMATION . . . . .	20
2.1	Meilleure approximation . . . . .	20
3	APPROXIMATION AU SENS DES MOINDRES CARRÉS . . . . .	21
4	CARACTÉRISATION . . . . .	22
4.1	Norme . . . . .	23

<b>3</b>	<b>INTERPOLATION POLYNOMIALE</b>	<b>25</b>
1	GÉNÉRALITÉS . . . . .	26
2	POLYNOME DE LAGRANGE . . . . .	27
2.1	Cas où les points sont equidistants . . . . .	29
3	Estimation de l'erreur dans l'interpolation de Lagrange . . . . .	30
4	POLYNOME DE NEWTON . . . . .	32
4.1	Différences finies . . . . .	32
4.2	Différences divisées . . . . .	33
5	Polynôme d'interpolation de Newton : . . . . .	35
5.1	Erreur d'interpolation . . . . .	36
5.2	Autre écriture du polynôme d'interpolation de Newton . . . . .	36
6	INTERPOLATION CUBIQUE DE HERMITE . . . . .	37
<b>4</b>	<b>INTEGRATION ET DÉRIVATION NUMÉRIQUE</b>	<b>39</b>
1	INTÉGRATION NUMÉRIQUE . . . . .	40
1.1	Méthode Générale . . . . .	40
1.2	Approximation d'une intégrale . . . . .	40
1.3	Utilisation de l'interpolation polynomiale . . . . .	41
1.4	Etude de l'erreur d'intégration . . . . .	42
1.5	Convergence des méthodes d'intégration . . . . .	42
1.6	Formules de Newton Cotes . . . . .	44
1.7	Formule de type fermé : des trapèzes et de Simpson . . . . .	44
1.8	Formule de type ouvert : . . . . .	45
1.9	Intégration par la méthode de Gauss . . . . .	45
1.10	Calcul de $\int_a^b f(x)dx$ . . . . .	47
1.11	Erreur de l'intégration par la méthode de Gauss . . . . .	47
2	DÉRIVATION NUMÉRIQUE . . . . .	48
2.1	Généralités : . . . . .	48
2.2	Utilisation de l'interpolation polynomiale . . . . .	50
2.3	Erreur de dérivation . . . . .	50
2.4	Algorithmes de dérivation . . . . .	54
2.5	Formules centrales de dérivation . . . . .	56
2.6	Formules non centrales de dérivation . . . . .	56
<b>5</b>	<b>RESOLUTION D'UN SYSTEME LINEAIRE</b>	<b>57</b>
1	METHODES DIRECTES . . . . .	57
1.1	Rappel . . . . .	57
1.2	Systèmes linéaires . . . . .	57
1.3	Résolution d'un système triangulaire supérieur . . . . .	58
2	Méthode de Gauss . . . . .	59
2.1	Interprétation matricielle de la méthode de Gauss . . . . .	60
3	Méthodes LU . . . . .	61
3.1	Décomposition LU . . . . .	61
4	Méthode de Cholesky . . . . .	62
4.1	Factorisation de Cholesky . . . . .	63
4.2	Algorithme de décomposition de Cholesky . . . . .	64
5	METHODES INDIRECTES . . . . .	66
5.1	Les méthodes itératives . . . . .	66

5.2	Différentes décomposition de A . . . . .	67
5.3	Méthode de Jacobi . . . . .	67
5.4	Méthode de Gauss-Seidel . . . . .	67
5.5	Méthode de relaxation . . . . .	68
6	Convergence des méthodes itératives . . . . .	68
6.1	Cas général . . . . .	68
<b>6</b>	<b>CALCUL DES VALEURS PROPRES ET VECTEURS PROPRES</b>	<b>71</b>
1	Introduction . . . . .	72
2	RAPPELS . . . . .	72
3	Calcul direct de $\det(A - \lambda I)$ . . . . .	72
4	Méthode de Krylov . . . . .	72
5	MÉTHODE DE LEVERRIER . . . . .	74
6	Valeurs et Vecteurs Propres . . . . .	75
7	La condition du calcul des valeurs propres . . . . .	75
7.1	Condition du calcul des vecteurs propres . . . . .	77
8	La méthode de la puissance . . . . .	78
9	Méthode de la puissance inverse de Wielandt . . . . .	79
10	VALEURS PROPRES ET VECTEURS PROPRES . . . . .	80
11	LA CONDITION DU CALCUL DES VALEURS PROPRES . . . . .	81
11.1	Condition du calcul des vecteurs propres . . . . .	83
12	LA METHODE DE LA PUISSANCE . . . . .	84
13	METHODE DE LA PUISSANCE INVERSE DE WIELANDT . . . . .	85
14	Transformation sous forme tridiagonale (ou de Hessenberg) . . . . .	87
14.1	a) A l'aide des transformations élémentaires . . . . .	87
14.2	b) A l'aide des transformations orthogonales . . . . .	88
14.3	Méthode de bisection pour des matrices tridiagonales . . . . .	88
14.4	Méthode de bisection. . . . .	90
15	L'itération orthogonale . . . . .	90
15.1	Généralisation de la méthode de la puissance (pour calculer les deux valeurs propres dominantes). . . . .	91
15.2	Méthode de la puissance (pour le calcul de toutes les valeurs propres) . . . . .	92
15.3	L' algorithme QR . . . . .	93
15.4	Accélération de la convergence . . . . .	94
15.5	Critère pour arrêter l'itération. . . . .	94
15.6	Le "double shift" de Francis . . . . .	95
15.7	Etude de la convergence . . . . .	96
16	Exercices . . . . .	96
17	TRANSFORMATION SOUS FORME TRIDIAGONALE (ou de HESSENBERG) . . . . .	99
17.1	a) A l'aide des transformations élémentaires . . . . .	100
17.2	b) A l'aide des transformations orthogonales . . . . .	100
17.3	Méthode de bisection pour des matrices tridiagonales . . . . .	101
17.4	Méthode de bisection. . . . .	102
18	L'ITERATION ORTHOGONALE . . . . .	103
18.1	Généralisation de la méthode de la puissance (pour calculer les deux valeurs propres dominantes). . . . .	103
18.2	Méthode de la puissance (pour le calcul de toutes les valeurs propres) . . . . .	105
18.3	L'algorithme QR . . . . .	105
18.4	Accélération de la convergence . . . . .	106
18.5	Critère pour arrêter l'itération. . . . .	107
18.6	Le "double shift" de Francis . . . . .	108
18.7	Etude de la convergence . . . . .	109
19	Exercices . . . . .	109





## 1 GÉNÉRALITÉS

Soit  $X = \{x_i, i = 1, 2, \dots, n\}$  un ensemble de  $n$  points appartenant à l'intervalle  $[a, b] \subset \mathbb{R}$ .

Supposons qu'à chaque  $x_i$  de  $X$ , on sache associer un  $y_i \in \mathbb{R}$ , résultat d'une expérience ou valeur donnée par une table, mais que cette opération soit impossible avec  $x \in [a, b] \setminus X$ .

Notons

$$\mathcal{D} = \{(x_i, y_i), \text{ pour } x_i \in X\}.$$

On veut déterminer une fonction, *facilement calculable*, qui permette d'obtenir une estimation "raisonnable" de la réponse du phénomène pour la valeur  $x$ .

Notons  $g$  cette fonction et  $\mathcal{G} = \{(x_i, g(x_i)), \text{ pour } x_i \in X\}$ . Si on définit une "distance" entre  $\mathcal{D}$  et  $\mathcal{G}$ , nous pouvons chercher une fonction  $g$ , d'un type préalablement choisi (un polynôme par exemple), telle que cette distance soit minimale.

C'est ce qu'on appelle une **approximation**.

Un procédé général consiste à limiter la représentation de  $g$  aux combinaisons linéaires des  $n + 1$  fonctions d'une certaine base, choisies a priori.

Si on désigne par

$$u_1(x), u_2(x), u_3(x), \dots, u_{n+1}(x)$$

les fonctions de la base, les représentations utilisées seront les combinaisons

$$a_1 u_1(x) + a_2 u_2(x) + a_3 u_3(x) + \dots + a_{n+1} u_{n+1}(x) \quad (1.1)$$

qui dépendent des  $n + 1$  coefficients  $a_i$  que nous devrions calculer pour définir chaque approximation.

*Remarque 36.* On utilise un procédé analogue lorsque on étudie la représentation de fonctions par des séries de fonctions.

On prend une base infinie  $u_1(x), u_2(x), \dots, u_{n+1}(x), \dots$  et on considère les séries

$$a_1 u_1(x) + a_2 u_2(x) + a_3 u_3(x) + \dots + a_{n+1} u_{n+1}(x) + \dots \quad (1.2)$$

L'équation (1.1) apparaît alors comme la somme partielle des  $n + 1$  premiers termes de (2.1).

## 2 APPROXIMATION

### 2.1 Meilleure approximation

Soit  $(E, d)$  un espace métrique et  $F \subset E$ . Chercher à approcher un élément  $f$  de  $E$  par un élément de  $F$ , c'est donc déterminer  $g$  de  $F$  tel que :

$$d(f, g) = \min_{h \in F} d(f, h) \quad (2.1)$$

S'il existe, cet élément sera appelé *meilleure approximation* de  $f$  dans  $F$  au sens de la distance  $d$ .

*Théorème 37* (d'existence). Si  $F$  est une partie compacte de  $E$ , alors il existe au moins un élément  $g$  de  $F$  tel que :

$$d(f, g) = \min_{h \in F} d(f, h).$$

*Corollaire 38*. Soit  $E$  un espace normé. Si  $F$  est un sous espace vectoriel de  $E$  de dimension finie, alors il existe au moins un élément  $g$  de  $F$  tel que :

$$\|f - g\| = \min_{h \in F} \|f - h\|.$$

*Remarque 39*. 1. Nous supposons dorénavant que  $E$  est l'ensemble des fonctions continues sur un intervalle  $[a, b]$  de  $\mathbb{R}$ .

2.  $g$  est le plus souvent cherchée sous la forme suivante :

$$g(x) = a_1 u_1(x) + a_2 u_2(x) + a_3 u_3(x) + \dots + a_{n+1} u_{n+1}(x)$$

où les  $u_i(x)$  sont des fonctions choisies dans :

— la classe des monômes :

$$1, x, x^2, \dots, x^n.$$

— la classe des fonctions trigonométriques :

$$1, \sin x, \sin 2x, \dots, \sin nx, \cos x, \cos 2x, \dots, \cos nx.$$

— la classe des fonctions exponentielles :

$$1, \exp x, \exp 2x, \dots, \exp nx$$

C'est-à-dire que :

—

$$g(x) = a_1 x^n + a_2 x^{n-1} + a_3 x^{n-2} + \dots + a_n x + a_{n+1},$$

ou ;

$$g(x) = c + a_1 \cos x + a_2 \cos 2x + a_3 \cos 3x + \dots + a_n \cos nx + \\ + b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + \dots + b_n \sin nx,$$

ou ;

$$g(x) = a_1 \exp b_1 x + a_2 \exp b_2 x + a_3 \exp b_3 x + \dots + a_n \exp b_n x,$$

ou aussi ;

$$g(x) = \frac{a_1 x^n + a_2 x^{n-1} + a_3 x^{n-2} + \dots + a_n x + a_{n+1}}{b_1 x^n + b_2 x^{n-1} + b_3 x^{n-2} + \dots + b_n x + b_{n+1}},$$

3. L'approximation de  $f$  par des polynômes, dite **approximation polynomiale**, est la plus fréquente.

*Théorème 40*. Si  $f$  est une fonction continue sur  $[a, b]$ , pour tout  $\varepsilon > 0$ , il existe un polynôme  $P_n$  de degré inférieur ou égal à  $n$  tel que :

$$\max_{x \in [a, b]} |f(x) - P_n(x)| < \varepsilon.$$

### 3 APPROXIMATION AU SENS DES MOINDRES CARRÉS

*Définition 41.* Soit  $\mathbf{E}$  un espace vectoriel. Pour tout couple  $(f, g)$  de  $\mathbf{E} \times \mathbf{E}$ , on définit le produit scalaire noté  $\langle f, g \rangle$  et la norme associée par :

$$\|f\| = \sqrt{\langle f, f \rangle}.$$

Pour tout  $f \in \mathbf{E}$ , il existe  $g \in \mathbf{F}$  ( $\mathbf{F}$  est un sous espace vectoriel de  $\mathbf{E}$ , de dimension finie) tel que :

$$\|f - g\| = \min_{h \in \mathbf{F}} \|f - h\|.$$

*Théorème 42.* Une condition nécessaire et suffisante pour que  $g$  soit une meilleure approximation de  $f$  est que :

$$\langle f - g, h \rangle = 0 \quad \text{pour tout } h \in \mathbf{F} \quad (3.1)$$

$g$  est appelée **meilleure approximation de  $f$  au sens des moindres carrés**.

*Démonstration.* La Condition est nécessaire : Soit  $g$  la meilleure approximation de  $f$ . Faisons un raisonnement par l'absurde. Supposons qu'il existe  $h_1 \in \mathbf{F}$  tel que :

$$\langle f - g, h_1 \rangle = c \neq 0.$$

Soit  $h_2 \in \mathbf{F}$  défini par :

$$h_2 = g + \frac{c}{\|h_1\|^2} h_1.$$

on a

$$\begin{aligned} \|f - h_2\| &= \sqrt{\langle f - g - \frac{c}{\|h_1\|^2} h_1, f - g - \frac{c}{\|h_1\|^2} h_1 \rangle} \\ &= \sqrt{\|f - g\|^2 - \frac{c^2}{\|h_1\|^2}} < \|f - g\| \end{aligned}$$

Ce qui est absurde.

La Condition est suffisante : Soit  $g_1 \in \mathbf{F}$  tel que :  $\langle f - g_1, h \rangle = 0$  pour tout  $h \in \mathbf{F}$  on a alors :

$$\begin{aligned} \|f - h\| &= \sqrt{\langle f - h, f - h \rangle} \\ &= \sqrt{\langle f - g_1 - h + g_1, f - g_1 - h + g_1 \rangle} \\ &= \sqrt{\|f - g_1\|^2 + \|h - g_1\|^2}. \end{aligned}$$

D'où  $\|f - g_1\| < \|f - h\|$  pour tout  $h \in \mathbf{F}$  et donc  $g_1 = g$ .

cqfd

*Théorème 43.* La meilleure approximation au sens des moindres carrés est unique.

*Démonstration.* Soient  $g_1$  et  $g_2$  deux meilleures approximations de  $f$  on a donc :

$$\langle f - g_1, h \rangle = 0 = \langle f - g_2, h \rangle \quad \text{pour tout } h \in \mathbf{F}$$

en particulier pour  $h = g_1 - g_2$ , d'où

$$\begin{aligned} \|g_1 - g_2\| &= \sqrt{\langle g_1 - g_2 + f - f, g_1 - g_2 \rangle} = \\ &= \sqrt{\langle f - g_1, g_1 - g_2 \rangle + \langle f - g_2, g_1 - g_2 \rangle} = 0 \end{aligned}$$

donc  $g_1 = g_2$ .

cqfd

## 4 CARACTÉRISATION

Il s'agit maintenant de construire le polynôme d'approximation qu'on note  $Q_n(x)$  défini par :

$$Q_n(x) = a_1 + a_2x + a_3x^2 + \dots + a_{n+1}x^n.$$

tel que la différence

$$\varepsilon(x) = f(x) - Q_n(x)$$

ait une norme aussi petite que possible.

Soit  $\mathbf{P}_n$  l'ensemble des polynômes de degré inférieur ou égal à  $n$ . Et soit  $f$  une fonction continue sur un intervalle  $[a, b]$ . La meilleure approximation de  $f$  par un polynôme  $Q_n$  de  $\mathbf{P}_n$  s'écrit :

$$Q_n(x) = b_1u_n(x) + b_2u_{n-1}(x) + \dots + b_{n+1}u_0(x) = \sum_{i=1}^{n+1} b_i u_{n+1-i}(x).$$

(Où les  $u_i$  forment une base de  $\mathbf{P}_n$ , c'est-à-dire :  $u_i(x) = x^i$   $i = 0, 1, \dots, n$ ), et un polynôme quelconque  $P_n$  de  $\mathbf{P}_n$  s'écrit :

$$P_n(x) = a_1u_n(x) + a_2u_{n-1}(x) + \dots + a_{n+1}u_0(x) = \sum_{i=1}^{n+1} a_i u_{n+1-i}(x).$$

La condition nécessaire et suffisante du théorème s'écrit :

$$\left\langle f - \sum_{i=1}^{n+1} b_i u_{n+1-i}, \sum_{i=1}^{n+1} a_i u_{n+1-i} \right\rangle = 0$$

pour tout élément  $(a_1, a_2, \dots, a_{n+1})$  de  $\mathbb{R}^{n+1}$ . Ce qui donne le système d'équations :

$$\sum_{i=1}^{n+1} b_i \langle u_{n+1-i}, u_{n+1-j} \rangle = \langle f, u_{n+1-j} \rangle \quad j = 1, 2, \dots, n+1 \quad (4.1)$$

qui admet une solution unique.

### 4.1 Norme

*Remarque 44.* Sur l'ensemble des fonctions continues  $\mathcal{C}([a, b])$ , on utilise le produit scalaire suivant : Soit  $\omega$  une fonction positive n'ayant qu'un nombre fini de racines sur  $[a, b]$  et telle que :  $\int_a^b \omega(x)h(x)dx$  existe pour tout  $h \in \mathcal{C}([a, b])$ . On suppose  $\omega$  continue par morceaux et on pose :

$$\langle f, g \rangle = \int_a^b \omega(x)f(x)g(x)dx.$$

*Remarque 45.* Le plus souvent dans le cas de l'approximation polynomiale on prend  $\omega(x) = 1$ .

La meilleure approximation de  $f \in \mathcal{C}([a, b])$  par un polynôme de  $\mathbf{P}_n$  est donc le polynôme  $Q_n$  défini par :

$$\int_a^b \omega(x)[f(x) - Q_n(x)]^2 dx = \min_{P_n \in \mathbf{P}_n} \int_a^b \omega(x)[f(x) - P_n(x)]^2 dx.$$

Ce polynôme existe et vérifie :

$$\int_a^b \omega(x)[f(x) - Q_n(x)]P_n(x)dx = 0 \quad \text{pour tout } P_n \in \mathbf{P}_n$$

Ses coefficients  $b_i$  sont donnés par le système d'équations linéaires :

$$\sum_{i=1}^{n+1} b_i \int_a^b \omega(x) x^{2n+2-i-j} dx = \int_a^b \omega(x) f(x) x^{n+1-j} dx \quad j = 1, \dots, n+1 \quad (4.2)$$

*Exemple 46.* Le polynôme  $Q_2$  qui réalise la meilleure approximation au sens des moindres carrés de la fonction  $f(x) = x^3 - x^2 - \frac{1}{4}x + \frac{1}{4}$  sur l'intervalle  $[-1, 1]$  avec  $\omega(x) = 1$  est donné par la condition :

$$\sum_{i=1}^3 b_i \int_{a=-1}^{b=1} 1 \cdot x^{2 \cdot 2 + 2 - i - j} dx = \int_{a=-1}^{b=1} 1 \cdot (x^3 - x^2 - \frac{1}{4}x + \frac{1}{4}) x^{2+1-j} dx \quad j = 1, 2, 3$$

c'est-à-dire le système :

$$\begin{cases} \frac{2}{5}b_1 + \frac{2}{3}b_3 & = & -\frac{2}{5} + \frac{1}{6} & = & -\frac{7}{30} \\ \frac{2}{3}b_2 & = & \frac{2}{5} - \frac{1}{6} & = & \frac{7}{30} \\ \frac{2}{3}b_1 + 2b_3 & = & -\frac{2}{3} + \frac{1}{2} & = & -\frac{1}{6} \end{cases}$$

qui donne comme solution :

$$b_1 = -1; b_2 = \frac{7}{20}; b_3 = \frac{1}{4}.$$

le polynôme cherché est donc :

$$Q_2(x) = -x^2 + \frac{7}{20}x + \frac{1}{4}.$$

L'erreur commise pour cette approximation s'évalue comme suit :

$$\varepsilon_2 = f(x) - Q_2(x) = (x^3 - x^2 - \frac{1}{4}x + \frac{1}{4}) - (-x^2 + \frac{7}{20}x + \frac{1}{4}) = x^3 - \frac{3}{5}x.$$

*Exemple 47.* Le polynôme  $Q_2$  qui réalise la meilleure approximation au sens des moindres carrés de la fonction  $f(x) = |x|$  sur l'intervalle  $[-1, 1]$  avec  $\omega(x) = 1$  est donné par la condition :

$$\sum_{i=1}^3 b_i \int_{-1}^1 x^{6-i-j} dx = \int_{-1}^1 |x| x^{3-j} dx \quad j = 1, 2, 3$$

c'est-à-dire le système :

$$\begin{cases} \frac{2}{5}b_1 + \frac{2}{3}b_3 & = & \frac{1}{2} \\ \frac{2}{3}b_2 & = & 0 \\ \frac{2}{3}b_1 + 2b_3 & = & 1 \end{cases}$$

qui donne comme solution :

$$b_1 = \frac{15}{16}; b_2 = 0; b_3 = \frac{3}{16}.$$

le polynôme cherché est donc :

$$Q_2(x) = \frac{15}{16}x^2 + \frac{3}{16}.$$

L'erreur commise pour cette approximation s'évalue comme suit :

$$\varepsilon_2 = f(x) - Q_2(x) = |x| - \left( \frac{15}{16}x^2 + \frac{3}{16} \right) = \begin{cases} \frac{15}{16}x^2 - x - \frac{3}{16} & \text{si } x < 0 \\ -\frac{15}{16}x^2 + x - \frac{3}{16} & \text{si } x > 0 \end{cases}.$$

## 5 SERIE D'EXERCICES

*Exercice 48.* Trouver le polynôme  $P_2$  qui réalise la meilleure approximation au sens des moindres carrés de la fonction  $f(x) = |x - 1|$  sur l'intervalle  $[0, 2]$  en prenant  $\omega(x) = 1$ .

*Exercice 49.* Trouver le polynôme  $P_2$  qui réalise la meilleure approximation au sens des moindres carrés de la fonction  $f(x) = |3x - 5|$  sur l'intervalle  $[-1, 2]$  en prenant  $\omega(x) = 1$ .

*Exercice 50.* Trouver le polynôme  $P_2$  qui réalise la meilleure approximation au sens des moindres carrés de la fonction donnée par le tableau suivant :

$x$	-1	-0,5	0	0,5	1
$f(x)$	-0,75	0	0,25	0	0