Polynômes orthogonaux

1) Introduction: équations différentielles

Soient a et b deux fonctions définies sur un intervalle $I =]\alpha,\beta[$ de $\mathbf{R}$, borné ou non, avec $a > 0$ sur I. On se propose d'étudier les valeurs propres de l'opérateur différentiel:

$$T(y) = ay'' + by' .$$

On introduit pour cela une fonction résolvante $w > 0$ qui permet d'écrire T sous la forme $T(y) = \dfrac{1}{w}(awy')'$ (on verra bientôt l'utilité de cette écriture). L'égalité:

$$T(y) = ay'' + a'y' + \frac{aw'}{w}y'$$

montre que w doit être solution de l'équation différentielle linéaire du premier ordre:

$$aw' + (a' - b)w = 0 ,$$

donc de la forme $w = e^A$, où A est une primitive de $\dfrac{b - a'}{a}$. On voit alors que pour le produit scalaire $< f,g >= \int_I f(x)g(x)w(x)dx$, on a:

$$< T(f),g >= \int_I (awf')'(x)g(x)dx = [awf'g]_\alpha^\beta - \int_I a(x)w(x)f'(x)g'(x)dx .$$

Si de plus aw s'annule aux bornes de I (ou tend vers 0 en ces bornes si I est infini), on a:

$$< T(f),g >= -\int_I a(x)w(x)f'(x)g'(x)dx =< f,T(g) > \ ,$$

autrement dit l'opérateur T est symétrique.
Bien sûr, il faudrait préciser un peu les hypothèses sur a et b pour que tout cela ait un sens, et en particulier préciser sur quel domaine est défini le produit scalaire précédent. Nous nous limiterons ici au cas où a et b sont des fonctions polynomiales, a de degré ≤ 2 et b de degré ≤ 1: $a(x) = a_2x^2 + a_1x + a_0$, $b(x) = b_1(x) + b_0$. Dans ce cas, T est, pour tout entier $n \in \mathbf{N}$, une application linéaire de l'espace $\mathcal{P}_n$ des fonctions polynomiales de degré $\leq n$ dans lui-même et le produit scalaire $< \cdot,\cdot >$ est défini sur cet espace si $\int_I |x|^k w(x)dx < +\infty$ pour tout $k \leq n$. Sous cette hypothèse, $\mathcal{P}_n$ muni de ce produit scalaire est un espace vectoriel euclidien de dimension finie $n+1$, et T un endomorphisme symétrique de cet espace. Il existe donc une base orthonormée de $\mathcal{P}_n$ constituée de vecteurs propres de T. En particulier, il existe au moins un vecteur propre P_n de degré n, qu'on peut choisir unitaire, et qui vérifie donc:

$$aP_n'' + bP_n' = \lambda_n P_n .$$

En considérant le terme de degré n dans cette égalité, on obtient:

$$\lambda_n = n(a_2(n-1) + b_1)\,.$$

Deux sous-espaces propres d'un endomorphisme symétrique associés à des valeurs propres distinctes sont orthogonaux. Il en résulte que si les valeurs propres λ_n sont toutes distinctes, les polynômes P_n seront deux à deux orthogonaux: $\int_I P_n(x)P_m(x)w(x)dx = 0$ pour $m \neq n$. Ce sera le cas si l'équation :

$$a_2 n(n-1) + b_1 n = a_2 m(m-1) + b_1 m\,,$$

i.e.

$$(n-m)(a_2(n+m-1) + b_1) = 0 \tag{1}$$

n'admet pas de solutions entières positives $(n{,}m)$ avec $n \neq m$.

Remarque: on peut bien sûr ajouter à l'opérateur T un terme cy, avec c constant, sans changer la symétrie de T, ni les vecteurs propres; on ne fait ainsi que translater les valeurs propres.

2) Exemples

Considérons tout de suite quelques exemples, qui nous permettront d'introduire les familles classiques de polynômes orthogonaux:

Exemple 1: les polynômes de Legendre
C'est le cas $I =]-1{,}1[$, $T(y) = (1-x^2)y'' - 2xy'$, soit $a(x) = 1 - x^2$, $b(x) = -2x$. Dans ce cas $w = 1$, puisque T s'écrit déjà $T(y) = ((1-x^2)y')'$ et $a(x)w(x) = 1 - x^2$ s'annule bien en ± 1. L'équation (1) s'écrit $n + m + 1 = 0$ et n'a donc pas de solution entière. Les valeurs propres sont $\lambda_n = -n(n+1)$ et l'équation différentielle vérifiée par P_n est donc $(1-x^2)y'' - 2xy' + n(n+1)y = 0$.

Exemple 2: les polynômes de Tchebychev
C'est le cas $I =]-1{,}1[$, $T(y) = (1-x^2)y'' - xy'$, soit $a(x) = 1 - x^2$, $b(x) = -x$. Dans ce cas $w(x) = \dfrac{1}{\sqrt{1-x^2}}$, $a(x)w(x) = \sqrt{1-x^2}$ s'annule en ± 1 et $P(x)w(x)$ est bien intégrable sur I pour tout polynôme P. L'équation (1) s'écrit $n + m = 0$, les valeurs propres sont $\lambda_n = -n^2$, et P_n vérifie l'équation différentielle $(1-x^2)y'' - xy' + n^2 y = 0$.

Exemple 3: les polynômes de Hermite
C'est le cas $I = \mathbf{R}$, $T(y) = y'' - 2xy'$, soit $a(x) = 1$, $b(x) = -2x$. Dans ce cas $w(x) = e^{-x^2}$, $a(x)w(x) = e^{-x^2}$ a pour limite 0 en $\pm\infty$ et $P(x)w(x)$ est bien intégrable sur I pour tout polynôme P. Les valeurs propres sont $\lambda_n = -2n$, et P_n vérifie l'équation différentielle $y'' - 2xy' + 2ny = 0$.

Exemple 4: les polynômes de Laguerre
C'est le cas $I =]0, +\infty[$, $T(y) = xy'' + (1-x)y'$, soit $a(x) = x$, $b(x) = 1 - x$. Dans ce cas $w(x) = e^{-x}$, $a(x)w(x) = xe^{-x}$ s'annule en 0 et a pour limite 0 en $+\infty$ et $P(x)w(x)$ est bien intégrable sur I pour tout polynôme P. Les valeurs propres sont $\lambda_n = -n$, et P_n vérifie l'équation différentielle $xy'' + (1-x)y' + ny = 0$.

3) Les théorèmes fondamentaux: relation de récurrence et racines

Nous considérerons dans toute la suite un intervalle I de $\mathbf{R}$, borné ou non, et une fonction $w > 0$ sur I, appelée poids, qui vérifie $\int_I |x|^n w(x)dx < +\infty$ pour tout entier $n \in \mathbf{N}$. Tout polynôme est donc intégrable pour la mesure de densité w par rapport à la mesure de Lebesgue sur I. Le procédé d'orthogonalisation de Schmidt permet alors de construire à partir de la base naturelle $1,X,\ldots,X^n$ de $\mathcal{P}_n$ une famille orthogonale de polynômes P_n, qu'on peut supposer unitaires. Si on veut une base orthonormée, il suffit de diviser les P_n par leur norme. On obtient ainsi une famille $\tilde{P}_n = \dfrac{P_n}{\|P_n\|}$, où $\|\ \|$ est la norme associée au produit scalaire. Ces familles vérifient alors des relations de récurrence qui s'écrivent:

$$P_n(X) = (X - \lambda_n)P_{n-1}(X) - \mu_n P_{n-2}(X) \quad (n \geq 1)$$

où:

$$\lambda_n = \frac{< XP_{n-1},P_{n-1} >}{\|P_{n-1}\|^2}, \quad \mu_n = \frac{\|P_{n-1}\|^2}{\|P_{n-2}\|^2}$$

(en particulier $\mu_n > 0$ pour tout $n \geq 2$) dans le cas des polynômes unitaires $P_n(X) = X^n + \ldots$ (en posant $P_{-1} = 0$). Plus généralement, on montre (ou on déduit de la relation précédente) que toute famille orthonormée P_n de polynômes orthogonaux de degré n vérifie une relation de récurrence du même type:

$$xP_{n-1} = \alpha_n P_n + \beta_n P_{n-1} + \gamma_n P_{n-2}$$

(cette remarque est utile, car on est parfois amené à utiliser d'autres familles orthogonales que celles normalisées par leur norme ou leur terme de plus haut degré; c'est en particulier le cas pour les polynômes de Tchebychev T_n, qui interviennent naturellement sous la forme $T_n(x) = \cos(n \arccos(x))$). Les P_n sont normalisés par leur coefficient de plus haut degré si $\alpha_n = 1$ pour tout n (à condition d'avoir pris $P_0 = 1$) et par leur norme euclidienne si $\gamma_n = \alpha_{n-1}$ pour tout n.
De plus, les racines du polynôme P_n sont, pour tout $n \geq 1$, réelles, intérieures à l'intervalle I et distinctes.
Les démonstrations détaillées se trouvent par exemple dans [Demailly], pp. 52-54.

4) Meilleure approximation hilbertienne

Les polynômes orthogonaux permettent d'expliciter la meilleure approximation au sens des moindres carrés, c'est-à-dire de la norme $\|\ \|$ définie par $\|f\|^2 = \int_I |f(x)|^2 w(x)dx = \int_I |f(x)|^2 d\mu(x)$ d'une fonction de $L^2(I,d\mu)$, où μ est la mesure de densité w par rapport à la mesure de Lebesgue sur I, par une fonction polynomiale de degré $\leq n$. Cet espace est en effet un espace de Hilbert et la meilleure approximation de f n'est autre que la projection orthogonale de f sur le sous-espace de dimension finie (donc fermé) $\mathcal{P}_n$.
Cette approximation convergera (au sens de la norme $\|\ \|$) vers f si l'espace des fonctions polynomiales est dense dans $L^2(I,d\mu)$. C'est le cas si I est borné: les fonctions polynomiales sont alors denses dans l'espace des fonctions continues sur I muni de la norme uniforme par le théorème de Stone-Weierstrass, et il suffit de comparer les deux normes pour établir le résultat (voir par exemple [Demailly]).
Par contre, si I n'est pas borné, le théorème de Stone-Weierstrass ne s'applique plus, et il faut alors examiner séparément chaque cas, ce qui nécessite des arguments d'analyse plus

poussés. Un oral d'algèbre n'est sans doute pas le lieu le plus indiqué pour développer en détail de tels arguments.
Remarque: au lieu de travailler avec des fonctions à valeurs réelles, on aurait aussi pu travailler avec des fonctions à valeurs complexes, en introduisant le produit scalaire hermitien:

$$< f,g >= \int_I \overline{f(x)} g(x) w(x) dx \,.$$

Cela n'aurait rien changé à la théorie, et en particulier au fait que les polynômes orthogonaux obtenus sont à coefficients réels.

5) Lien avec les matrices tridiagonales symétriques

Considérons la matrice tridiagonale symétrique réelle:

$$M_n = \begin{pmatrix} \lambda_1 & \sqrt{\mu_2} & 0 & \dots & 0 & 0 \\ \sqrt{\mu_2} & \lambda_2 & \sqrt{\mu_3} & \dots & 0 & 0 \\ 0 & \sqrt{\mu_3} & \lambda_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_{n-1} & \sqrt{\mu_n} \\ 0 & 0 & 0 & \dots & \sqrt{\mu_n} & \lambda_n \end{pmatrix}$$

et son polynôme caractéristique $\Delta_n(X) = \det(XI_n - M_n)$. On voit alors facilement, en développant le déterminant par rapport à la dernière ligne ou la dernière colonne, que la suite Δ_n vérifie la même relation de récurrence que la suite P_n, et lui est donc égale. Les racines de P_n apparaissent donc comme les valeurs propres d'une matrice symétrique réelle, et sont donc réelles. De plus, si on note $x_1, \dots, x_n$ les racines de P_n et $y_1, \dots, y_{n-1}$ celles de P_{n-1}, il résulte du principe du minimax (voir par exemple [Serre]) que l'on a $x_1 \le y_1 \le x_2 \le \dots \le y_{n-1} \le x_n$, puisque la matrice M_{n-1} s'obtient en barrant la dernière ligne et la dernière colonne de M_n. Nous retrouverons cette propriété d'entrelacement des racines d'une autre manière plus loin.
On peut également mettre en évidence les vecteurs propres correspondants: en effet on vérifie facilement que, pour toute racine x de P_n, le vecteur

$${}^t(\sqrt{\mu_2\mu_3 \dots \mu_n} P_0(x), \sqrt{\mu_3 \dots \mu_n} P_1(x), \dots, \sqrt{\mu_n} P_{n-2}(x), P_{n-1}(x))$$

est vecteur propre de M_n associé à la valeur propre x (seule la dernière égalité à vérifier fait intervenir le fait que x est racine de P_n; les autres sont vraies pour tout x).

6) Noyau reproduisant et relation de Darboux-Christoffel

On sait que la décomposition d'un vecteur P d'un espace vectoriel euclidien dans une base orthonormée $(\tilde{P}_0, \dots, \tilde{P}_n)$ s'écrit:

$$P = \sum_{k=0}^{n} < \tilde{P}_k, P > \tilde{P}_k \,.$$

Cette décomposition s'écrit ici:

$$\begin{aligned} P(x) &= \sum_{k=0}^{n} \left(\int_I \tilde{P}_k(y) P(y) w(y) dy \right) \tilde{P}_k(x) \\ &= \int_I \left(\sum_{k=0}^{n} \tilde{P}_k(y) \tilde{P}_k(x) \right) P(y) w(y) dy \\ &= \int_I K_n(x,y) P(y) w(y) dy \\ &= < K_n(x,\cdot), P > \ , \end{aligned}$$

où l'on a posé:

$$K_n(x,y) = \sum_{k=0}^{n} \tilde{P}_k(y) \tilde{P}_k(x) = \sum_{k=0}^{n} \frac{P_k(y) P_k(x)}{\|P_k\|^2} \,.$$

Ce polynôme K_n, symétrique en les deux variables x et y, peut encore s'écrire:

$$K_n(x,y) = \frac{1}{\|P_n\|^2} \frac{P_{n+1}(x) P_n(y) - P_n(x) P_{n+1}(y)}{x - y}$$

pour $x \neq y$, et:

$$K_n(x,x) = \frac{1}{\|P_n\|^2} \left[P'_{n+1}(x) P_n(x) - P'_n(x) P_{n+1}(x) \right]$$

si $x = y$. C'est la *formule de Darboux-Christoffel*, qui découle immédiatement de la relation de récurrence vérifiée par les P_n. Posons en effet:

$$R_n(x,y) = \frac{1}{\|P_n\|^2} \frac{P_{n+1}(x) P_n(y) - P_n(x) P_{n+1}(y)}{x - y}$$

pour $x \neq y$. Il suffit alors de remplacer $P_{n+1}(x)$ par $(x - \lambda_{n+1}) P_n(x) - \mu_{n+1} P_{n-1}(x)$ (idem pour $P_{n+1}(y)$) et de se souvenir que $\mu_{n+1} = \dfrac{\|P_n\|^2}{\|P_{n-1}\|^2}$ pour obtenir:

$$R_n(x,y) = \frac{P_n(x) P_n(y)}{\|P_n\|^2} + R_{n-1}(x,y) \,.$$

La première expression en découle. La valeur pour $y = x$ s'obtient en écrivant:

$$P_{n+1}(x) P_n(y) - P_n(x) P_{n+1}(y) = P_{n+1}(x) \left(P_n(y) - P_n(x) \right) - P_n(x) \left(P_{n+1}(y) - P_{n+1}(x) \right)$$

et en faisant tendre y vers x.

Conséquence: entrelacement des racines

$K_n(x,x) = \displaystyle\sum_{k=0}^{n} \frac{P_k^2(x)}{\|P_k\|^2}$ est partout >0 (ce qui permet de retrouver la simplicité des racines de P_n), on a donc $P'_{n+1}(x) P_n(x) > 0$ pour toute racine x de P_{n+1}. Or si $x_1 < x_2 < \cdots < x_{n+1}$ sont les racines de P_{n+1}, les signes des $P'_{n+1}(x_i)$ sont alternés (faire un dessin ou un calcul). La même propriété est donc vraie pour les signes des $P_n(x_i)$, ce qui montre qu'entre deux racines consécutives de P_{n+1} il y a au moins une racine de P_n. Comme P_n ne peut avoir plus de n racines, il y en a exactement une dans chaque intervalle $]x_i, x_{i+1}[$. Les racines de P_n s'intercalent donc entre celles de P_{n+1}.

7) Application: méthodes de quadrature gaussiennes

Les méthodes de Gauss pour le calcul numérique d'intégrales constituent une application directe de la théorie des polynômes orthogonaux et sont bien décrites dans [Demailly], pp. 73-77. Les démonstrations sont très algébriques et rentrent donc bien dans le cadre de l'oral d'algèbre.
Le problème est le suivant: on voudrait calculer une valeur numérique approchée d'une intégrale $\int_\alpha^\beta f(x)w(x)dx$, où $I = [\alpha,\beta]$ est un intervalle compact de $\mathbf{R}$ et w un poids. On cherche pour cela une valeur approchée de la forme $\sum_{k=1}^n \lambda_k f(x_k)$, où les λ_k sont des réels (ne dépendant pas de la fonction f) et les x_k des points de I, indépendants eux aussi de f. On dit qu'une telle méthode est *d'ordre* N si la formule:

$$\int_\alpha^\beta f(x)w(x)dx = \sum_{k=1}^n \lambda_k f(x_k)$$

est exacte pour toute fonction f polynomiale de degré $\leq N$ et fausse pour au moins une fonction polynomiale f de degré $N+1$. On pourrait s'attendre à ce qu'une telle méthode soit d'ordre n, où n est le nombre de points intervenant dans la formule. En fait, si on choisit comme points x_k les racines du n-ième polynôme orthogonal associé au poids w, on montre que la méthode est exactement d'ordre $2n-1$, donc bien meilleure que ce qu'on pouvait attendre, à condition naturellement de choisir les coefficients λ_k de manière appropriée.
Tout cela est démontré dans [Demailly]. On peut seulement ajouter deux remarques:
a) les coefficients λ_k obtenus sont tous positifs, comme on le voit en appliquant la formule aux polynômes $\displaystyle\prod_{\substack{i=1\\ i\neq k}}^n (x-x_i)^2$;
b) si on veut simplement montrer que la méthode n'est pas d'ordre $> 2n-1$ (et non pas obtenir une majoration de l'erreur, comme cela découle de la démonstration du livre), il suffit de considérer la fonction polynomiale $\displaystyle\prod_{k=1}^n (x-x_k)^2$: elle est de degré $2n$, son intégrale sur I est strictement positive, alors que la formule approchée donne un résultat nul.

8) Autre application: les fonctions de Hermite

Posons, pour tout entier $n \in \mathbf{N}$, $\psi_n(x) = H_n(x)e^{-x^2/2}$, où H_n est le n-ième polynôme de Hermite (exemple 3). Ces fonctions, appelées fonctions de Hermite, constituent une famille orthogonale dans $L^2(\mathbf{R})$. Il suffit de les diviser par leur norme pour obtenir une base hibertienne de $L^2(\mathbf{R})$ (*attention: une base hilbertienne n'est pas une base au sens de l'algèbre*). Ceci répond à une question parfois posée à l'oral d'analyse concernant de telles bases. De plus, cette base est particulièrement intéressante car elle est constituée de fonctions propres de la transformation de Fourier: on peut en effet montrer (mais un oral d'algèbre n'est sans doute le lieu le plus approprié pour le faire) qu'avec la normalisation $\hat{f}(\xi) = \int_{\mathbf{R}} e^{i\xi x} f(x)dx$, pour $f \in L^1(\mathbf{R})$, on a $\hat{\psi}_n(\xi) = \sqrt{2\pi} i^n \psi_n(\xi)$ pour tout $n \in \mathbf{N}$. Un outil très efficace dans cette démonstration est l'existence d'une *fonction génératrice* très simple pour les polynômes de Hermite. On a en effet, en posant:

$$G(x,t) = \sum_{n=0}^\infty \frac{t^n}{n!} H_n(x)\,,$$

l'identité $G(x,t) = e^{2tx-t^2}$, qui permet à la fois d'expliciter les H_n et de calculer la transformée de Fourier des fonctions de Hermite.

Commentaires

Les parties 3 et 4 sont centrales et paraissent incontournables; elles bénéficient en outre d'une solide bibliographie.
L'application aux méthodes de quadrature peut faire l'objet d'un développement d'une longueur raisonnable, de nature très algébrique et est bien documentée.
Les autres parties paraissent moins indispensables. Il est toutefois indispensable de savoir que les polynômes orthogonaux sont solutions d'équations différentielles simples et interviennent naturellement dans de nombreux problèmes d'analyse. L'introduction permet d'avoir une vue d'ensemble de plusieurs cas particuliers développés par exemple dans [Rombaldi].
Plusieurs des thèmes développés dans diverses parties (noyau reproduisant, équations différentielles) sont traités sous forme d'exercices corrigés dans [Crouzeix-Mignot] et sont donc très abordables. Le lien avec les matrices tridiagonales provient de [Stoer-Bulirsch].
On peut naturellement choisir aussi d'étudier plus en détail l'une ou l'autre des familles classiques introduites dans la partie 2, ce qui est fait dans un grand nombre d'ouvrages (par exemple [Rombaldi]). Les polynômes de Tchebychev et leurs racines interviennent de ce point de vue dans de nombreux problèmes d'approximation, en particulier pour la norme $\| \|_\infty$ (voir [Demailly], ch. II).

Références

[Crouzeix-Mignot] M. Crouzeix, A.L. Mignot, *Exercices d'analyse numérique des équations différentielles,* Masson, 1986.

[Demailly] J.-P. Demailly, *Analyse numérique et équations différentielles,* Presses Universitaires de Grenoble, 1991.

[Rombaldi] J.-E. Rombaldi, *Thèmes pour l'agrégation de mathématiques*, EDP Sciences, 1999.

[Serre] D. Serre, *Les matrices: théorie et pratique*, Dunod, 2001.

[Stoer-Bulirsch] J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis,* Springer-Verlag, 1980.