Intégration de polynômes
Points de Gauss

Commençons par un exercice classique de premier cycle.

Problème 1 *Trouver trois réels α, β et γ tels que, pour tout polynôme* P *de degré au plus* 2*, on ait :*

$$\int_0^1 \mathrm{P}(x)\, dx = \alpha \cdot \mathrm{P}(0) + \beta \cdot \mathrm{P}\left(\frac{1}{2}\right) + \gamma \cdot \mathrm{P}(1)$$

Montrer que cette relation reste valable pour les polynômes de degré 3*.*

Solution : En remarquant que les deux membres de l'égalité sont (quels que soient α, β et γ) des formes linéaires en P, il suffit que l'égalité soit vraie lorsque le polynôme P décrit une base de l'espace vectoriel $\mathbb{R}_2[\mathrm{X}]$ des polynômes de degré au plus 2. Par exemple, la base $\{1, \mathrm{X}, \mathrm{X}^2\}$.

On obtient donc la condition nécessaire et suffisante :

$$\left\{\begin{aligned} \int_0^1 1 \,\mathrm{d}x &= \alpha \cdot 1 + \beta \cdot 1 + \gamma \cdot 1 \\ \int_0^1 x \,\mathrm{d}x &= \alpha \cdot 0 + \beta \cdot \frac{1}{2} + \gamma \cdot 1 \\ \int_0^1 x^2 \,\mathrm{d}x &= \alpha \cdot 0^2 + \beta \cdot \left(\frac{1}{2}\right)^2 + \gamma \cdot 1^2 \end{aligned}\right.$$

soit

$$\left\{\begin{aligned} \alpha + \beta + \gamma &= 1 \\ \frac{\beta}{2} + \gamma &= \frac{1}{2} \\ \frac{\beta}{4} + \gamma &= \frac{1}{3} \end{aligned}\right.$$

On trouve $$\alpha = \gamma = \frac{1}{6} \text{ et } \beta = \frac{2}{3}\ .$$

c'est-à-dire, pour tout polynôme de degré au plus 2 :

$$\int_0^1 \mathrm{P}(x)\,\mathrm{d}x = \frac{1}{6}\left(\mathrm{P}(0) + 4\mathrm{P}\left(\frac{1}{2}\right) + \mathrm{P}(1)\right)\ .$$

Une simple vérification nous montre que cette formule reste valable pour les polynômes de degré 3. Il suffit là aussi d'examiner le cas du polynôme X^3.

Or $$\int_0^1 x^3\,\mathrm{d}x = 1/4 = 1/6\left(0 + 4 \cdot (1/2)^3 + 1\right)\ . \qquad \square$$

N.B. On vérifie en revanche que cette formule est fausse pour les polynômes de degré 4…

1 Intégration des polynômes et formes linéaires

1.1 Structure de $\mathbb{R}_n[X]$

Définition-Propriété : On note $\mathbb{R}_n[X]$ l'ensemble des polynômes à coefficients réels, et de degré au plus n. $\mathbb{R}_n[X]$ possède une structure d'espace vectoriel sur $\mathbb{R}$, de dimension finie.

La structure d'espace vectoriel découle de la définition des polynômes. Il suffit d'observer que la somme de polynômes de degré au plus n est de degré au plus n... D'autre part, la famille $\left\{1, X, X^2, \ldots, X^n\right\}$ constitue bien sûr une base de $\mathbb{R}_n[X]$, qui est donc de dimension $n+1$. Plus généralement :

Propriété 1.1 *Si $\{P_0, P_1, \ldots, P_n\}$ est une famille de polynômes telle que pour tout indice i, P_i est de degré exactement i, alors la famille $\{P_0, P_1, \ldots, P_n\}$ est une base de $\mathbb{R}_n[X]$. Cette famille est dite « échelonnée en degré ».*

Démonstration : La propriété se démontre aisément par récurrence : le résultat est évident pour $n = 0$, en effet le polynôme P_0 est alors une constante λ non nulle, et $\mathbb{R}_0[X]$ est l'ensemble des polynômes constants, donc des multiples de P_0.

Supposons le résultat établi jusqu'au rang n, et soit $\{P_0, P_1, \ldots, P_{n+1}\}$ une famille échelonnée en degré. Par hypothèse de récurrence, $\{P_0, P_1, \ldots, P_n\}$ est une base de $\mathbb{R}_n[X]$. Nous allons montrer que $\{P_0, P_1, \ldots, P_{n+1}\}$ est libre.

Soit $\lambda_0, \ldots, \lambda_{n+1}$ des réels tels que $\sum_{k=0}^{n+1} \lambda_k P_k = 0$. Si λ_{n+1} est non nul, alors on peut écrire :

$$P_{n+1} = -\frac{1}{\lambda_{n+1}} \sum_{k=0}^{n} \lambda_k P_k$$

ce qui est absurde puisqu'un polynôme de degré $n+1$ ne peut s'écrire comme une combinaison linéaire de polynômes de degrés au plus n. Donc $\lambda_{n+1} = 0$, et l'on obtient :

$$\sum_{k=0}^{n} \lambda_k P_k = 0 \ .$$

Comme la famille $\{P_0, P_1, \ldots, P_n\}$ est libre (par hypothèse de récurrence), on obtient alors $\lambda_0 = \lambda_1 = \cdots = \lambda_n = 0$. On a donc montré que la famille $\{P_0, P_1, \ldots, P_{n+1}\}$ est libre. Comme elle possède exactement $n+2$ éléments, c'est une base de $\mathbb{R}_{n+1}[X]$. Le résultat est établi par récurrence.

□

1.2 Formes linéaires

Définition Soit E un espace vectoriel sur un corps $\mathbb{K}$. On appelle *forme linéaire* sur E une application linéaire de E dans $\mathbb{K}$.

EXEMPLE : Si a est un réel, l'application $P \longmapsto P(a)$ est une forme linéaire sur l'espace vectoriel $\mathbb{R}[X]$ des polynômes à coefficients réels. Cette application est appelée *évaluation au point a.*

N.B. La donnée de la valeur en chaque élément d'une base de E définit de manière unique une forme linéaire. Autrement dit, si $(e_i)_{i\in \mathrm{I}}$ désigne une base de E, si $(\alpha_i)_{i\in \mathrm{I}}$ désigne une famille d'éléments du corps $\mathbb{K}$, alors il existe une unique forme linéaire f sur E telle que $f(e_i) = \alpha_i$ pour tout i.

L'unicité est immédiate : si deux formes linéaires f et g coïncident sur une base $(e_i)_{i\in \mathrm{I}}$, alors la différence $f - g$ est nulle sur ladite base. Comme tout élément x de E se décompose sur cette base, comme une combinaison linéaire finie $\sum_{k=1}^{n} \lambda_k e_k$, on obtient par linéarité de $f - g$:

$$(f-g)(x) = \sum_{k=1}^{n} \lambda_k (f-g)(e_k) = \sum_{k=0}^{n} \lambda_k \cdot 0 = 0$$

et donc $f - g$ est la forme linéaire nulle, c'est-à-dire que f et g sont égales. Pour l'existence, on définit une fonction f de E dans $\mathbb{K}$ par :

$$x = \sum_{k=1}^{n} \lambda_k e_k \longmapsto \sum_{k=1}^{n} \lambda_k \alpha_k$$

et l'on vérifie aisément que cette application est bien linéaire sur E.

Propriété 1.2 *Soit* E *un espace vectoriel sur un corps* $\mathbb{K}$*. L'ensemble* E′ *des formes linéaires sur* E *possède une structure d'espace vectoriel sur* $\mathbb{K}$*. Si* E *est de dimension finie, alors* E′ *est aussi de dimension finie, et sa dimension est égale à celle de* E.

Démonstration : On vérifie sans problème que la somme de deux formes linéaires est encore une forme linéaire, tout comme le produit d'une forme linéaire par un élément du corps $\mathbb{K}$. E′ possède donc une structure d'espace vectoriel sur $\mathbb{K}$. Supposons désormais que E est de dimension finie n.

Soit $\{b_1, \ldots, b_n\}$ une base de E. Pour $1 \leq i \leq n$, soit b_i^* la forme linéaire définie par $b_i^*(b_i) = 1$ et $b_i^*(b_j) = 0$ pour tout $j \neq i$. Alors la famille $\{b_1^*, \ldots, b_n^*\}$ constitue une base de E′. En effet, si f est une forme linéaire donnée, alors on peut écrire f sous la forme :

$$f = \sum_{k=1}^{n} f(b_i) b_i^* \ .$$

Ces deux formes linéaires f et $\sum_{k=1}^{n} f(b_i) b_i^*$ coïncident sur chacun des vecteurs de base b_i $(1 \leq i \leq n)$, donc elles sont égales ! □

N.B. Inversement, si E′ est de dimension finie, on peut associer à toute base $\{f_1, \ldots, f_n\}$ de E′ une base $\{e_1, \ldots, e_n\}$ de E telle que pour tous i, j, on ait $f_i(e_j) = 1$ si $i = j$, 0 sinon.

1.3 Application aux polynômes

Si $a_0, a_1, \ldots, a_n$ sont des réels distincts, alors les évaluations des polynômes de degré au plus n en ces points forment un famille libre de $\mathbb{R}_n[\mathrm{X}]'$. En effet, soient $a_0, \ldots, a_n$ des réels distincts, et soit $f_i \in \mathbb{R}_n[\mathrm{X}]'$ l'évaluation au point a_i, c'est-à-dire la forme linéaire :

$$f_i \left\{ \begin{array}{rcl} \mathbb{R}_n[\mathrm{X}] & \to & \mathbb{R} \\ \mathrm{P} & \longmapsto & \mathrm{P}(a_i) \end{array} \right.$$

Soient $\lambda_0, \ldots, \lambda_n$ des réels tels que $\sum_{k=0}^{n} \lambda_k f_k$ soit la forme linéaire nulle sur $\mathbb{R}_n[X]$.

Pour i un indice (désormais fixé), on appelle P_i le polynôme $\prod_{\substack{j=0 \\ j\neq i}}^{n} (X - a_j)$. Le polynôme P_i est de degré exactement n, en particulier on doit avoir :

$$\sum_{k=0}^{n} \lambda_k f_k(P_i) = 0 \ .$$

Mais, pour $k \neq i$, on a $f_k(P_i) = \prod_{\substack{j=0 \\ j\neq i}}^{n} (a_k - a_j) = 0$ (lorsque l'indice j est égal à k, on obtient un terme du produit de la forme $(a_k - a_k) = 0$). Il vient $\lambda_i f_i(P_i) = 0$.

Soit
$$\lambda_i \prod_{\substack{j=0 \\ j\neq i}}^{n} (a_i - a_j) = 0 \ .$$

Comme par hypothèse, les réels $a_0, \ldots, a_n$ sont distincts deux à deux, tous les termes $a_i - a_j$ du produit sont non nuls. On en déduit que $\lambda_i = 0$, et ce pour tout indice i, ce qui prouve bien que la famille $\{f_0, \ldots f_n\}$ est libre. C'est donc une base de l'espace vectoriel des formes linéaires sur $\mathbb{R}_n[X]$ (en tant que famille libre à $n+1$ éléments d'un espace vectoriel de dimension $n+1$).

En particulier, toute forme linéaire sur $\mathbb{R}_n[X]$ peut s'exprimer comme combinaison linéaire des formes $f_0, \ldots f_n$.

Or l'application de $\mathbb{R}_n[X]$ dans $\mathbb{R}$ qui à un polynôme P associe la valeur $\int_0^1 P(x)\, dx$ est une forme linéaire sur $\mathbb{R}_n[X]$ (par linéarité de l'intégrale !). Et donc, quels que soient les points $a_0, \ldots, a_n$ deux à deux distincts, on peut trouver des coefficients $\alpha_0, \ldots, \alpha_n$ tels que, pour tout polynôme P de degré au plus n, on ait :

$$\int_0^1 P(x)\, dx = \sum_{k=0}^{n} \alpha_k P(a_k) \ .$$

N.B. Ceci justifie, entre autre, le fait que nous ayons effectivement trouvé trois nombres α, β et γ qui répondent à notre problème, en préambule.

2 Produit scalaire et orthogonalité

Définition Soit E un espace vectoriel sur $\mathbb{R}$. Une application $(u,v) \longmapsto \langle u, v\rangle$ de E^2 dans $\mathbb{R}$ est un *produit scalaire* si et seulement si elle est :

– bilinéaire, c'est-à-dire linéaire par rapport à chaque variable :

$$\begin{array}{ll} \forall u, u', v \in E, \forall \lambda, \lambda' \in \mathbb{K}, & \langle \lambda u + \lambda' u', v\rangle = \lambda \langle u, v\rangle + \lambda' \langle u', v\rangle \\ \forall u, v, v' \in E, \forall \mu, \mu' \in \mathbb{K}, & \langle u, \mu v + \mu' v'\rangle = \mu \langle u, v\rangle + \mu' \langle u, v'\rangle \end{array}$$

– *symétrique* : $\forall u, v \in E, \quad \langle u, v\rangle = \langle v, u\rangle$

– *définie* : $\forall u \in E, \quad \langle u, u\rangle = 0 \Longrightarrow u = 0$

– *positive* : $\forall u \in E, \quad u \neq 0 \Longrightarrow \langle u, u\rangle > 0$

Exemple : Sur $\mathbb{R}^2$, et sur $\mathbb{R}^n$ de manière générale, le produit scalaire le plus classique, c'est l'application définie par :

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = x_1 y_1 + \cdots + x_n y_n \ .$$

Définition La norme d'un vecteur u de l'espace vectoriel E (muni du produit scalaire $\langle ., . \rangle$) est le réel :

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

2.1 Bases orthogonales, procédé de Gram-Schmidt

Lorsque l'on dispose d'un espace vectoriel (de dimension finie) sur $\mathbb{R}$, muni d'un produit scalaire, des bases particulièrement intéressantes sont les bases *orthonormales* ou *orthonormées.*

Définition Une famille $\{e_1, \dots, e_n\}$ de vecteurs non nuls (resp. une base) de l'espace vectoriel E, muni du produit scalaire $\langle ., . \rangle$, est dite *orthogonale* si et seulement si, pour tous indices $i \neq j$, on a $\langle e_i, e_j \rangle = 0$.

N.B. Une famille orthogonale est nécessairement libre : si $\lambda_1, \dots, \lambda_n$ sont tels que $\sum\limits_{k=1}^{n} \lambda_k e_k = 0$, alors en effectuant le produit scalaire de cette somme par e_i $(1 \leq i \leq n)$, on trouve $\langle e_i, \lambda_i e_i \rangle = 0$ soit $\lambda_i \|e_i\|^2 = 0$, et donc $\lambda_i = 0$ puisque les vecteurs sont non nuls.

Définition Une famille de vecteurs (resp. une base) est dite *orthonormée* si elle est orthogonale, et constituée de vecteurs de norme 1.

Exemple : Pour le produit scalaire classique sur $\mathbb{R}^n$, la base canonique est orthonormée. En effet, si e_i désigne le vecteur $(0, \dots, 0, 1, 0, \dots, 0)$ dont la seule composante non nulle est à la i-ème position, alors ces vecteurs sont clairement orthogonaux et de norme 1.

Théorème 2.1 (Orthonormalisation de Gram-Schmidt) *Quelle que soit la famille libre $\{u_1, \dots, u_p\}$ d'un espace vectoriel* E *muni d'un produit scalaire $\langle ., . \rangle$, il existe une unique famille orthonormée $\{e_1, \dots, e_p\}$ telle que :*

$$\forall k \leq p, \quad \begin{cases} \operatorname{Vect}\{e_1, \dots, e_k\} = \operatorname{Vect}\{u_1, \dots, u_k\} \\ \langle u_k, e_k \rangle > 0 \end{cases}$$

Démonstration : On procède par récurrence sur p. Pour $p = 1$, la condition $\operatorname{Vect}\{e_1\} = \operatorname{Vect}\{u_1\}$ impose que e_1 soit de la forme λu_1, avec λ réel. On veut que e_1 soit de norme 1, c'est-à-dire $\|\lambda u_1\| = |\lambda| \, \|u_1\| = 1$. Ceci nous donne deux solutions possibles pour λ. Comme on doit aussi avoir $\langle u_1, e_1 \rangle = \lambda \|u_1\|^2 > 0$, on obtient $\lambda > 0$, et donc e_1 est déterminé de manière unique.

Soit maintenant $p \geq 2$, et $\{u_1, \dots, u_p\}$ une famille libre d'un espace vectoriel E muni d'un produit scalaire. Soit $\{e_1, \dots, e_{p-1}\}$ l'unique famille orthonormée construite (grâce à l'hypothèse de récurrence), de façon à ce que l'on ait :

$$\forall k \leq p-1, \quad \begin{cases} \operatorname{Vect}\{e_1, \dots, e_k\} = \operatorname{Vect}\{u_1, \dots, u_k\} \\ \langle u_k, e_k \rangle > 0 \end{cases}$$

La condition $\text{Vect}\{e_1, \ldots, e_p\} = \text{Vect}\{u_1, \ldots, u_p\}$ nous impose une relation du type :

$$e_p = \sum_{k=1}^{p} \alpha_k u_k = \alpha_p u_p + \sum_{k=1}^{p-1} \lambda_k e_k \ .$$

(En effet, la somme $\sum_{k=1}^{p-1} \alpha_k u_k$ est dans $\text{Vect}\{e_1, \ldots, e_{p-1}\} = \text{Vect}\{u_1, \ldots, u_{p-1}\}$, donc peut s'écrire comme combinaison linéaire des vecteurs $e_1, \ldots, e_{p-1}$.)

Les conditions d'orthogonalité $\langle e_p, e_i \rangle = 0$ pour $i < p$ nous donnent alors :

$$\forall j \leq p-1, \quad \lambda_j + \alpha_p \langle u_p, e_j \rangle = 0 \ .$$

Donc e_p s'écrit $$e_p = \alpha_p \left(u_p - \sum_{k=1}^{p-1} \langle u_p, e_k \rangle e_k \right) \ .$$

Comme la famille $\{u_1, \ldots, u_p\}$ est libre, le vecteur u_p n'est pas dans dans le sous-espace vectoriel $\text{Vect}\{u_1, \ldots, u_{p-1}\} = \text{Vect}\{e_1, \ldots, e_{p-1}\}$, c'est-à-dire que le terme $u_p - \sum_{k=1}^{p-1} \langle u_p, e_k \rangle e_k$ est forcément non nul. On détermine alors (au signe près) le coefficient α_p par la condition $\|e_p\| = 1$. Enfin, on doit avoir :

$$0 < \langle u_p, e_p \rangle = \langle \frac{1}{\alpha_p} e_p + \sum_{k=1}^{p-1} \langle u_p, e_k \rangle e_k, e_p \rangle = \frac{1}{\alpha_p}$$

ce qui détermine le signe de α_p, donc α_p.

□

2.2 Des formules "économiques" d'intégration

Revenons à notre espace vectoriel $\mathbb{R}_n[\mathrm{X}]$ des polynômes de degré au plus n.

Propriété 2.2 *L'application* $(\mathrm{P}, \mathrm{Q}) \longmapsto \langle \mathrm{P}, \mathrm{Q} \rangle = \int_0^1 \mathrm{P}(x)\mathrm{Q}(x)\, dx$ *définit un produit scalaire sur* $\mathbb{R}_n[\mathrm{X}]$.

Démonstration : La bilinéarité est immédiate. D'autre part, pour tout polynôme P à coefficients réels, l'application $x \longmapsto \mathrm{P}(x)^2$ est à valeur dans $\mathbb{R}_+$, et donc le nombre $\langle \mathrm{P}, \mathrm{Q} \rangle$ est positif ou nul.

Enfin, si un polynôme P est non nul, alors l'application $x \longmapsto \mathrm{P}(x)^2$, continue sur $[\,0\,;1\,]$, prend sur cet intervalle des valeurs strictement positives (sinon P aurait une infinité de racines). On en conclut que $\int_0^1 \mathrm{P}(x)^2 \, \mathrm{d}x$ est strictement positive (propriété de positivité de l'intégrale). □

Le procédé d'orthonormalisation de Gram-Schmidt, partant de la base canonique $\{1, \mathrm{X}, \mathrm{X}^2, \ldots, \mathrm{X}^n\}$, nous fournit alors une base orthonormée $\{\mathrm{P}_0, \ldots, \mathrm{P}_n\}$. Comme l'une des conditions du procédé de Gram-Schmidt est que pour tout i, on ait $\text{Vect}\{\mathrm{P}_0, \ldots, \mathrm{P}_i\} = \text{Vect}\{1, \mathrm{X}, \ldots, \mathrm{X}^i\} = \mathbb{R}_i[\mathrm{X}]$, la famille construite est nécessairement constituée de polynômes de degrés échelonnés (P_i est de degré i pour tout i). En outre, P_i est orthogonal à $\mathbb{R}_{i-1}[\mathrm{X}]$ (c'est-à-dire à tous les polynômes de degré au plus $i-1$), ce pour tout i.

Lemme 2.3 *Pour tout i, le polynôme P_i possède i racines réelles distinctes, toutes dans l'intervalle* $]\,0\,;1\,[$.

Démonstration : La propriété est vraie pour $i = 0$ (P_0 est un polynôme constant, sans racine). Soit $i \geq 1$ un indice désormais fixé. On a :

$$\int_0^1 \mathrm{P}_i(t)\mathrm{P}_0(t)\ \mathrm{d}t = \int_0^1 \mathrm{P}_i(t)\ \mathrm{d}t = 0$$

donc P_i change nécessairement de signe sur l'intervalle $]\,0\,;1\,[$, et possède donc au moins une racine sur cet intervalle.

Soient $x_1 < \cdots < x_k$ les racines (distinctes) de P_i dans l'intervalle $]\,0\,;1\,[$. Alors P_i se décompose sous la forme d'un produit :

$$\prod_{j=1}^{k} (\mathrm{X} - x_j)^{\alpha_j}\ \mathrm{Q}$$

où les nombres α_j sont des entiers naturels et Q est un polynôme sans racine dans l'intervalle $]\,0\,;1\,[$. Considérons maintenant la séquence $j_1 < j_2 < \cdots < j_r$ des indices j compris entre 1 et k et tels que l'entier α_j soit impair.
Alors le polynôme $\mathrm{P}_i \cdot \prod_{m=1}^{r} (\mathrm{X} - x_{j_m})$ s'écrit :

$$\prod_{j=1}^{k} (\mathrm{X} - x_j)^{\beta_j}\ \mathrm{Q}$$

où $\beta_j = \alpha_j$ si α_j est pair, et $\beta_j = \alpha_j + 1$ sinon. C'est-à-dire que les nombres β_j sont tous pairs, et donc le produit $\prod_{j=1}^{k} (\mathrm{X} - x_j)^{\beta_j}$ est de signe constant (positif) sur l'intervalle $[\,0\,;1\,]$. D'autre part, comme le polynôme Q n'a pas de racine sur $]\,0\,;1\,[$, il est lui aussi de signe constant sur $[\,0\,;1\,]$. On a donc :

$$\langle \mathrm{P}_i, \prod_{m=1}^{r} (\mathrm{X} - x_{j_m})\rangle = \int_0^1 \mathrm{P}_i(x) \prod_{m=1}^{r} (x - x_{j_m})\ \mathrm{d}x \neq 0\ .$$

Or, par construction, on sait que le polynôme P_i est orthogonal à $\mathbb{R}_{i-1}[\mathrm{X}]$. On en déduit donc que le polynôme $\prod_{m=1}^{r} (\mathrm{X} - x_{j_m})$ est de degré au moins i. Par définition de ce polynôme, ceci signifie que P_i a i racines distinctes dans $]\,0\,;1\,[$.
□

Soit P_n le polynôme de degré n obtenu par le procédé de Gram-Schmidt à partir de la base $1, \mathrm{X}, \ldots, \mathrm{X}^n$ de $\mathbb{R}_n[\mathrm{X}]$. Soient $a_1 < \cdots < a_n$ les racines (distinctes d'après le lemme qui précède) de P_n. Nous avons vu que les formes linéaires définies par l'évaluation des polynômes en $a_1, \ldots, a_n$ forme une base de $\mathbb{R}_{n-1}[\mathrm{X}]'$ (on a n points), et qu'en particulier il existe des coefficients $\alpha_1, \ldots, \alpha_n$ tels que pour tout polynôme P de degré au plus $n - 1$, on ait :

$$\int_0^1 \mathrm{P}(x)\ \mathrm{d}x = \sum_{k=1}^{n} \alpha_k \mathrm{P}(a_k)\ . \tag{1}$$

Théorème 2.4 *La formule précédente reste valable pour tout polynôme de degré au plus* $2n-1$.

Autrement dit, l'évaluation des polynômes en n points bien choisis suffit à calculer l'intégrale de polynômes de degré allant jusqu'à $2n-1$!

Démonstration : Supposons construit la famille de polynômes $P_0, \dots, P_n$, selon le procédé de Gram-Schmidt. Soient $\alpha_1, \dots, \alpha_n$ tels que pour tout polynôme P de degré au plus $n-1$, on ait :

$$\int_0^1 P(x)\, dx = \sum_{k=1}^{n} \alpha_k P(a_k) \ .$$

Considérons un polynôme P de degré au plus $2n-1$. La division euclidienne de P par P_n nous donne deux polynômes Q et R, avec R de degré strictement inférieur au degré de P_n (donc de degré au plus $n-1$), tels que $P = QP_n + R$. D'autre part, on a alors $\deg(Q) + \deg(P_n) \leq \deg(P) \leq 2n-1$, et donc le quotient Q est lui aussi de degré au plus $n-1$.

Mais alors, par linéarité de l'intégrale, on a :

$$\int_0^1 P(x)\, dx = \int_0^1 Q(x)P_n(x)\, dx + \int_0^1 R(x)\, dx \ .$$

On reconnaît dans le premier terme de la somme le produit scalaire $\langle Q, P_n \rangle$. Or, par construction, le polynôme P_n est orthogonal à $\mathbb{R}_{n-1}[X]$, c'est-à-dire que ce produit scalaire est nul. On obtient :

$$\int_0^1 P(x)\, dx = \int_0^1 R(x)\, dx = \sum_{k=1}^{n} \alpha_k R(a_k) \ .$$

En effet le polynôme R est de degré au plus $n-1$, donc la formule (1) s'applique. Enfin, évaluant l'égalité $P = QP_n + R$ en un point a_i, on obtient :

$$P(a_i) = Q(a_i)P_n(a_i) + R(a_i) = R(a_i)$$

puisque a_i est par construction racine de P_n.

D'où
$$\int_0^1 P(x)\, dx = \sum_{k=1}^{n} \alpha_k P(a_k) \ . \qquad \square$$

Définition Les points $a_1, \dots, a_n$, racines du n-ième polynôme de la famille orthogonale construite, s'appellent les *points de Gauss.*

N.B. On peut énoncer une réciproque au théorème 2.4 : si l'on a n points $a_1, \dots, a_n$, et n réels $\lambda_1, \dots, \lambda_n$ tels que, pour tout polynôme de degré au plus $2n-1$, on ait :

$$\int_0^1 P(x)\, dx = \sum_{k=1}^{n} \lambda_i P(a_i)$$

alors $a_1, \dots, a_n$ sont les points de Gauss d'ordre n, c'est-à-dire les racines du polynôme P_n.

En effet, soit P le polynôme $\prod_{i=1}^{n} (X - a_i)$. Pour tout polynôme Q de degré au plus $n-1$, le produit PQ est de degré au plus $2n-1$, et donc vérifie :

$$\int_0^1 P(x)Q(x)\, dx = \sum_{k=1}^{n} \lambda_i P(a_i)Q(a_i) = 0 \ .$$

En d'autres termes, $\langle \mathrm{P}, \mathrm{Q} \rangle = 0$, c'est-à-dire que P est orthogonal à $\mathbb{R}_{n-1}[\mathrm{X}]$. Le polynôme P est donc, à constante multiplicative près, le n-ième polynôme de la famille orthonormale précédemment construite.

2.3 Détermination des points de Gauss

En pratique, le procédé de Gram-Schmidt est souvent pénible à mettre en oeuvre. Ici, nous ne recherchons que les racines des polynômes orthogonaux en question. On peut donc alléger la méthode en ne recherchant les polynômes qu'à constante multiplicative près (on y gagne, en particulier, des calculs d'intégrales). Ainsi, on trouve successivement :

- $\mathrm{P}_0 = 1$.
- P_1 doit être de degré 1, et orthogonal à P_0. En posant P_1 sous la forme $\mathrm{X} + a$ (comme on cherche P_1 à constante multiplicative près, on peut le supposer unitaire), on trouve la condition :
$$\int_0^1 (x+a)\,\mathrm{d}x = 0 = 1/2 + a\ .$$
Donc $\mathrm{P}_1 = \mathrm{X} - 1/2$.
- On pose $\mathrm{P}_2 = \mathrm{X}^2 + a\mathrm{X} + b$. P_2 doit vérifier $\langle \mathrm{P}_2, \mathrm{P}_0 \rangle = \langle \mathrm{P}_2, \mathrm{X} \rangle = 0$, c'est-à-dire :
$$\int_0^1 (x^2 + ax + b)\,\mathrm{d}x = \int_0^1 (x^3 + ax^2 + bx)\,\mathrm{d}x = 0\ .$$
On trouve $\qquad 1/3 + a/2 + b = 1/4 + a/3 + b/2 = 0$

donc $\mathrm{P}_2 = \mathrm{X}^2 - \mathrm{X} + 1/6 = \left(\mathrm{X} - 1/2 - 1/2\sqrt{3}\right)\left(\mathrm{X} - 1/2 + 1/2\sqrt{3}\right)$.

Les calculs deviennent, on le voit, assez vite compliqués (on a des systèmes $n \times n$ à résoudre pour déterminer P_n). Dans le cas particulier de notre produit scalaire $\langle \mathrm{P}, \mathrm{Q} \rangle = \displaystyle\int_0^1 \mathrm{P}(x)\mathrm{Q}(x)\,\mathrm{d}x$, une famille orthogonale peut être obtenue de la façon suivante :

Théorème 2.5 *Les polynômes* $\mathrm{P}_n = \left((\mathrm{X} - \mathrm{X}^2)^n\right)^{(n)}$ *(dérivée n-ième du polynôme* $(\mathrm{X} - \mathrm{X}^2)^n$*) forment une famille échelonnée orthogonale.*

Démonstration : En effet, pour tout n, le polynôme P_n est de degré $2n - n = n$ (on dérive n fois un polynôme de degré $2n$). Pour montrer que cette famille de polynômes est orthogonale, nous allons montrer que pour tout n, P_n est orthogonal à $\mathbb{R}_{n-1}[\mathrm{X}]$.

Soit donc n un entier, soit R le polynôme $(\mathrm{X} - \mathrm{X}^2)^n$ (de sorte que $\mathrm{P}_n = \mathrm{R}^{(n)}$), et Q est un polynôme de degré au plus $n - 1$. Alors, n intégrations par parties successives nous donnent :

$$\begin{aligned}\langle \mathrm{P}_n, \mathrm{Q} \rangle &= \int_0^1 \mathrm{R}^{(n)}(x)\mathrm{Q}(x)\,\mathrm{d}x \\ &= \left[\sum_{k=1}^{n} (-1)^{k+1}\mathrm{R}^{(n-k)}(x)\mathrm{Q}^{(k-1)}(x)\right]_0^1 \\ &\quad + (-1)^n \int_0^1 (x - x^2)^n \mathrm{Q}^{(n)}(x)\,\mathrm{d}x\ .\end{aligned}$$

D'autre part, comme le polynôme $\mathrm{R} = (\mathrm{X} - \mathrm{X}^2)^n$ ainsi que ses dérivées successives jusqu'à l'ordre $n-1$ s'annulent en 0 et en 1, les termes d'intégration par partie sont tous nuls... On trouve donc :

$$\langle \mathrm{P}_n, \mathrm{Q} \rangle = (-1)^n \int_0^1 (x - x^2)^n \mathrm{Q}^{(n)}(x)\,\mathrm{d}x \;.$$

Enfin, comme le polynôme Q est de degré au plus $n-1$, sa dérivée n-ième est nulle, et l'on trouve $\langle \mathrm{P}_n, \mathrm{Q} \rangle = 0$, c'est-à-dire que le polynôme P_n est orthogonal à $\mathbb{R}_{n-1}[\mathrm{X}]$.

□

Exemple : On obtient ainsi sans beaucoup d'effort le polynôme P_3 comme la dérivée d'ordre 3 du polynôme $\left(\mathrm{X} - \mathrm{X}^2\right)^3 = \mathrm{X}^3 - 3\mathrm{X}^4 + 3\mathrm{X}^5 - \mathrm{X}^6$. C'est-à-dire que $\mathrm{P}_3 = -120\mathrm{X}^3 + 180\mathrm{X}^2 - 72\mathrm{X} + 6$. On remarque que $1/2$ est racine de P_3, ce qui permet de factoriser le polynôme. On trouve :

$$\mathrm{P}_3 = -120\left(\mathrm{X} - \frac{1}{2}\right)\left(\mathrm{X} - \frac{1}{2} - \frac{1}{2}\sqrt{\frac{3}{5}}\right)\left(\mathrm{X} - \frac{1}{2} + \frac{1}{2}\sqrt{\frac{3}{5}}\right) \;.$$

Les points de Gauss correspondant sont donc $1/2$, $1/2 - 1/2\sqrt{3/5}$ et $1/2 + 1/2\sqrt{3/5}$. Le même type de calculs que ceux effectués au tout début de ce texte nous fournit les coefficients correspondants (il suffit là aussi d'examiner les polynômes $1, \mathrm{X}$ et X^2). On obtient, pour tout polynôme de degré inférieur ou égal à $2 \cdot 3 - 1 = 5$:

$$\int_0^1 \mathrm{P}(x)\,\mathrm{d}x = \frac{5}{18} \cdot \mathrm{P}\left(\frac{1}{2} - \frac{1}{2}\sqrt{\frac{3}{5}}\right) + \frac{4}{9} \cdot \mathrm{P}\left(\frac{1}{2}\right) + \frac{5}{18} \cdot \mathrm{P}\left(\frac{1}{2} - \frac{1}{2}\sqrt{\frac{3}{5}}\right) \;.$$

N.B. À translation près (c'est-à-dire si l'on considère dans tout ce qui précède l'intégrale des polynômes sur l'intervalle $[-1\,;1]$ et non plus sur $[0\,;1]$), ces polynômes s'appellent les *polynômes de Legendre*. Le n-ième polynôme de Legendre est ainsi la dérivée n-ième de $\left(1 - \mathrm{X}^2\right)^n$.

2.4 Coefficients

Trouver les points de Gauss revient, on l'a vu, à factoriser un polynôme de degré n. C'est a priori un problème compliqué, il devient d'ailleurs très vite impossible de déterminer les racines de manière exacte. En revanche, comme on sait que toutes ces racines sont dans l'intervalle $]\,0\,;1\,[$, de nombreuses méthodes numériques nous permettent de trouver des valeurs approchées de ces racines, connaissant le polynôme (ce qui est assez simple, en utilisant le théorème 2.5).

La détermination des coefficients ne pose en revanche pas de gros problèmes, une fois connus les points de Gauss $a_1, \ldots, a_n$. Il suffit d'écrire les relations $\displaystyle\int_0^1 \mathrm{P}(x)\,\mathrm{d}x = \sum_{i=1}^n \lambda_i \mathrm{P}(a_i)$ pour les polynômes $1, \mathrm{X}, \ldots, \mathrm{X}^{n-1}$. On obtient ainsi un système linéaire de n équations à n inconnues $\lambda_1, \ldots, \lambda_n$, qui se résout sans trop de problème.

Une autre solution consiste à utiliser des polynômes interpolateurs de Lagrange. On obtient au passage une propriété supplémentaire pour nos coefficients : ils sont tous positifs.

En effet, si l'on considère le polynôme $\mathrm{L}_i = \prod_{\substack{k=1 \\ k \neq i}}^{n} (\mathrm{X} - a_k)^2$, de degré $2n-2$, on trouve :

$$\int_0^1 \mathrm{L}_i(x)\,\mathrm{d}x = \sum_{k=1}^{n} \lambda_k \mathrm{L}_i^2(a_k) = \lambda_i \mathrm{L}_i(a_i)\ .$$

Comme le polynôme L_i est positif, il en va de même pour son intégrale et pour sa valeur en a_i, et donc λ_i est nécessairement positif.

Conclusion : Ce procédé d'intégration des polynômes par évaluation en des points particuliers s'inscrit dans un cadre bien plus large. Tout ce qui fait marcher cette méthode, c'est le fait que la fonction :

$$\varphi : \mathrm{P} \longmapsto \int_0^1 \mathrm{P}(t)\,\mathrm{d}t$$

est une forme linéaire sur $\mathbb{R}[\mathrm{X}]$, et que $\langle \mathrm{P}, \mathrm{Q} \rangle = \varphi(\mathrm{PQ})$ définit bien un produit scalaire.

On pourrait bien sûr changer l'intervalle d'intégration, ou encore considérer une fonction du type :

$$\mathrm{P} \longmapsto \int_{\mathrm{I}} \mathrm{P}(t)\omega(t)\,\mathrm{d}t$$

où ω est une fonction (encore appelée *poids*) choisie de façon à ce que l'intégrale sur I de la fonction $\mathrm{P}\omega$ soit bien définie quel que soit le polynôme P.

En fonction de l'intervalle I choisi et du poids ω, on construit alors diverses familles de polynômes, orthogonales pour les produits scalaires correspondant. Et, à chaque fois, le polynôme de degré n de la famille possède n racines distinctes dans I, qui sont les points de Gauss correspondant au produit scalaire en question…

Ainsi, si l'on prend $\mathrm{I} = \left]0\,;+\infty\right[$ et ω de la forme $t \longmapsto t^{\alpha}e^{-t}$, la famille de polynômes correspondant s'appelle les *polynômes de Laguerre*. Si l'on prend $\mathrm{I} = \mathbb{R}$ et $\omega : t \longmapsto e^{-t^2}$, ce sont les *polynômes d'Hermitte*, etc.

On pourra trouver de plus amples informations sur ces familles de polynômes, et bien d'autres choses, dans un texte de Jean-Etienne Rombaldi intitulé `Approximation et Interpolation` disponible sur le site « Maths en Prépa », à l'adresse `http://www.mathprepa.com/`.