Cours de Mathémathiques — Ing1, Promo 2007

Mankalas — bouche_v

14 février 2005

Table des matières

1 Probabilités

1.1 Expérience aléatoire

Définition : Une expérience est qualifiée d'aléatoire si on ne peut prévoir par avance son résultat et si, répétée dans des conditions identiques, elle peut donner lieu à des résultats différents. On note ω le résultat de cette expérience.

Définition : On note Ω l'ensemble de tous les résultats possibles (on l'appelle Univers). Un événement est une proposition logique.

▶*Exemple :*
On lance deux dés et on fait la somme des points. Soit ω "la somme des points > 10". On dira qu'un événement est réalisé ou non suivant que la proposition est vraie ou fausse.

1.2 Algèbre des éléments

Définition : Soit $\mathscr{C}$ l'ensemble des événements. A tout événement A, on associe son contraire noté $\overline{A}$ tel que si A est réalisé, alors $\overline{A}$ ne l'est pas.

Définition : La classe $\mathscr{C}$ est définie par les trois axiomes

1. $\forall A \in \mathscr{C}, \overline{A} \in \mathscr{C}$.
2. Pour tout ensemble fini ou dénombrable de $A_i \in \mathscr{C}$, alors $\bigcup\limits_i A_i \in \mathscr{C}$.
3. $\Omega \in \mathscr{C}$.

▶*Remarque :*
Les trois axiomes impliquent que
– $\emptyset \in \mathscr{C}$ car $\overline{\Omega} = \emptyset \in \mathscr{C}$.
– $\bigcap\limits_i A_i \in \mathscr{C}$ car $A_i \in \mathscr{C} \Rightarrow \overline{A} \in \mathscr{C} \Rightarrow \bigcup\limits_i \overline{A_i} \in \mathscr{C} \Rightarrow \overline{\bigcup\limits_i \overline{A_i}} \Rightarrow \bigcap\limits_i \overline{\overline{A_i}} = \bigcap\limits_i A_i$
Ces propriétés définissent ce que l'on appelle une algèbre de Boole ou une tribu.

Définition : On appelle espace probabilisable le couple $(\Omega, \mathscr{C})$.

Définition : On appelle probabilité sur $(\Omega, \mathscr{C})$ (ou loi de probabilité) une application $P : \begin{array}{ccc} \mathscr{C} & \rightarrow & [0,1] \\ A & \rightarrow & P(A) \end{array}$
telle que

1. $P(\Omega) = 1$
2. Pour tout ensemble dénombrable d'événements incompatibles $(A_i \cap A_j \neq \emptyset, i \neq j)$,
$P\left(\bigcup\limits_i A_i\right) = \sum\limits_i P(A_i)$

Ce sont les axiomes de Kolmogorov
Propriétés élémentaires :

1. $P(\emptyset) = 0$
2. $P(\overline{A}) = 1 - P(A)$
3. $P(A) \leq P(B)$ si $A \subset B$
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

1.3 Lois de probabilités conditionnelles

Définition : Soit B un événement tel que $P(B) \neq 0$. On appelle probabilité conditionnelle de A sachant B le rapport

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Vérifions que les axiomes de Kolmogorov sont vérifiés :

1. $P(\Omega/B) = \dfrac{P(\Omega \cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$

2. Soit A_i tel que $A_i \cap A_j \neq \emptyset, i \neq j$, on a $P\left(\bigcup_i A_i/B\right) = \dfrac{P\left(\left(\bigcup_i A_i\right) \cap b\right)}{P(B)} = \dfrac{P\left(\bigcup_i (A_i \cap B)\right)}{P(B)}$. Or $A_i \cap B \cap A_j \cap B = A_i \cap A_j \cap B = \emptyset \cap B = \emptyset$. Donc $\dfrac{P\left(\bigcup_i (A_i \cap B)\right)}{P(B)} = \sum_i \dfrac{P(A_i \cap B)}{P(B)} = \sum_i P(A_i/B)$

Définition : A est indépendant de B si $P(A/B) = P(A)$. Donc

$$P(A/B) = P(A) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = P(A)P(B)$$

Formules de Bayes :

1. $P(B/A) = \dfrac{P(A/B)P(B)}{P(A)}$

2. $P(B_i/A) = \dfrac{P(A/B_i)P(B_i)}{\sum_j P(A/B_j)P(B_j)}$

Preuve :

1. $P(A/B) = \dfrac{P(A \cap B)}{P(B)}$ et $P(B/A) = \dfrac{P(A \cap B)}{P(A)} \Rightarrow P(B/A) = \dfrac{P(A/B)P(B)}{P(A)}$

2. Soit B_i un système complet d'événements $\begin{cases} B_i \cap B_j = \emptyset & i \neq j \\ \bigcup_i B_i = \Omega & \text{(partition de } \Omega) \end{cases}$

   On a $P(A \cap B_i) = P(A/B_i)P(B_i)$ or $A = \bigcup_i (A \cap B_i) = \bigcup_i B_i \cap A = \Omega \cap A = A$, donc

$$P(A) = P\left(\bigcup_i A \cap B_i\right) = \sum_i P(A \cap B_i) = \sum_i P(A/B_i)P(B_i)$$

►*Exemple :*
Dans une usine, trois machines M_1, M_2 et M_3 fabriquent des boulons de même type. M_1 sort en moyenne 0,3% de boulons défectueux, M_2 en sort 0,8% et M_3 1%. On mélange 1000 boulons dans une caisse, 500 provenant de M_1, 350 de M_2 et 150 de M_3. On tire au hasard un boulon de la caisse et il est défectueux. Quelle est la probabilité qu'il ait été fabriqué par M_1 (resp. M_2 et M_3) ?
Evénements : D : "boulons défectueux", M_i : "boulon fabriqué par M_i".

On cherche $P(M_1/D)$. On connait $P(D/M_1) = 0,3\%$, $P(D/M_2) = 0,8\%$, $P(D/M_3) = 1\%$, $P(M_1) = 0,5$, $P(M_2) = 0,35$ et $P(M_3) = 0,15$. Donc

$$P(M_1/D) = \frac{P(D/M_1)P(M_1)}{P(D)} = \frac{P(D/M_1)P(M_1)}{P(D/M_1)P(M_1) + P(D/M_2)P(M_2) + P(D/M_3)P(M_3)} = 0,26\%$$

$$P(M_2/D) = 0,48\%, P(M_3/D) = 0,26\%$$

1.4 Variables aléatoires

►*Exemple :*
Considérons le lancer de deux dés. On s'intéresse à la somme des points marqués.
Soit $S : \begin{array}{ccc} \Omega & \to & E = \{2, \ldots, 12\} \\ \omega & \mapsto & S(\omega) \end{array}$. Pour obtenir la probabilité d'une valeur quelconque de S, il suffit de dénombrer les ω qui réalisent cette valeur $P(S = 5) = P(\{(1,4),(2,3),(3,2),(4,1)\}) = \frac{4}{36} = \frac{1}{9}$.
Généralement, $P(S = s) = P(\{S^{-1}(s)\})$.
Si X est une application de $(\Omega, \mathscr{C}, P)$ dans E, il faut que E soit probabilisable, c'est-à-dire muni d'une tribu $\mathscr{F}$ et que l'image réciproque de tout élément de $\mathscr{F}$ est un élément de $\mathscr{C}$.

Définition : Une variable aléatoire X est une application mesurable de $(\Omega, \mathscr{C}, P) \to (E, \mathscr{F})$. Lorsque $E = \mathbb{R}$, on utilise comme tribu la tribu borélienne $\mathscr{B}$ engendrée par les intervalles de $\mathbb{R}$.

$$\forall B \in \mathscr{B}, P_X(B) = P(\{\omega, X(\omega) \in B\}) = P(\{X^{-1}(B)\})$$

Définition : La fonction de répartition d'une variable aléatoire X est l'application $F : \mathbb{R} \to [0,1]$ vérifiant
- $F(x) = P[X < x]$
- $F(-\infty) = 0$ et $F(+\infty) = 1$

Définition : Une loi de probabilité P_X (X une variable aléatoire) admet une densité f si pour tout intervalle I de $\mathbb{R}$ on a $P_X(I) = \displaystyle\int_I f(x)dx$.

Définition : On appelle espérance de X

$$E(X) = \sum_k x_k P[X = x_k] \quad \text{si } X \text{ est discrète}$$

$$E(X) = \int_{\mathbb{R}} x f(x)dx \quad \text{si } X \text{ est continue}$$

C'est une moyenne pondérée des valeurs de la variable aléatoire.

Définition : La variance de X

$$V(X) = E((X - E(X))^2) \quad \text{et} \quad \sigma(X) = \sqrt{V(X)} = \|X - E(X)\|$$

L'écart-type est la distance entre X et la valeur centrale.

Définition :

$$E[\varphi(X)] = \sum_x \varphi(x)P[X = x] \quad \text{ou} \quad E[\varphi(X)] = \int_{\mathbb{R}} \varphi(x)f(x)dx$$

Formules de Kömig Hyghens :
- $E((X-a)^2) = V(X) + (E(X) - a)^2, \forall a \in \mathbb{R}$
- $V(X) = E(X^2) - E^2(X)$

Preuve :
- Supposons X une variable aléatoire continue ($f(x)$ densité de X). On a

$$\begin{aligned}
E((X-a)^2) &= E((X - E(X) + E(X) - a)^2) \\
&= E((X - E(X)^2) + (E(X) - a)^2 + 2(E(X) - a)(X - E(X))) \\
&= \int_{\mathbb{R}} (x - E(X))^2 f(x)dx + (E(X) - a)^2 \int_{\mathbb{R}} f(x)dx + 2(E(X) - a) \int_{\mathbb{R}} (x - E(x))f(x)dx \\
&= V(X) + (E(X) - a)^2 + 2(E(X) - a) \left(\underbrace{\int_{\mathbb{R}} x f(x)dx}_{E(X)} - E(X) \underbrace{\int_{\mathbb{R}} f(x)dx}_{1} \right) \\
&= V(X) + (E(X) - a)^2
\end{aligned}$$

- Idem que précédemment avec $a = 0$.

►*Remarque :*

$$\begin{aligned}
V(X+Y) &= E((X+Y)^2) - E^2(X+Y) \\
&= E(X^2) + E(Y^2) + 2E(XY) - E^2(X) - E^2(Y) - 2E(X)E(Y) \\
&= V(X) + V(Y) + 2(E(XY) - E(X)E(Y)) \\
&= V(X) + V(Y) + 2cov(X,Y)
\end{aligned}$$

Lorsque X et Y sont indépendantes, la covariance est nulle.

1.5 Lois discrètes

1.5.1 Loi discrète uniforme

Soit $X = \{1, 2, 3, \ldots, n\}$, $P[X = k] = \frac{1}{n}, \forall k \in [\![1..n]\!]$. On a

$$E(X) = \sum_k x_k P[X = x_k] = \sum_{k=1}^{n} kP[X = k] = \frac{1}{n}\sum_{k=1}^{n} k = \frac{n(n+1)}{2n} = \frac{n+1}{2}$$

On a $V(X) = E(X^2) - E^2(X)$. Or

$$E(X^2) = \sum_k x_k P[X = k] = \frac{1}{n}\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6n} = \frac{(n+1)(2n+1)}{6}$$

Donc

$$V(X) = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}$$

1.5.2 Loi de Bernouilli de paramètre p

C'est la loi d'une variable aléatoire X ne pouvant prendre que les valeurs 1 ou 0 avec les probabilités p et $1-p$. X représente la fonction indicatrice de A : $X = \begin{cases} 1 & \text{si } A \text{ se réalise avec la probabilité } p \\ 0 & \text{sinon} \end{cases}$.
On a

$$E(X) = \sum_{k=0}^{1} kP[X = k] = 1P[X = 1] = p \quad \text{et} \quad E(X^2) = \sum_{k=0}^{1} k^2 P[X = k] = 1P[X = 1] = p$$

Donc

$$V(X) = E(X^2) - E^2(X) = p - p^2 = p(1-p)$$

1.5.3 Loi binômiale $B(n, p)$

Supposons que l'on répète n fois dans des conditions identiques une expérience aléatoire dont l'issue se traduit par la réalisation (ou la non-réalisation) de l'événement A avec la probabilité p, le résultat d'une expérience étant indépendant des résultats précédents.

Soit X une variable aléatoire "nombre de réalisation de A sur n expériences". Soit X_i variable aléatoire de Bernouilli associée à l'expérience numéro i $(1 \leq i \leq n)$.

Donc $X = \sum_{i=1}^{n} X_i$ (somme de variable aléatoire indépendantes de Bernouilli). D'où

$$E(X) = \sum_{i=1}^{n} E(X_i) = np$$

et

$$V(X) = V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) = np(1-p)$$

1.5.4 Loi de Poisson de paramètre λ : $\mathscr{P}(\lambda)$

C'est la loi d'une variable aléatoire entière positive qui satisfait à

$$P[X = x] = e^{-\lambda}\frac{\lambda^x}{x!}, \forall x \in \mathbb{N}$$

On a

$$\sum_{x=0}^{+\infty} P[X = x] = e^{-\lambda}\sum_{x=0}^{+\infty}\frac{\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1$$

On obtient la loi de Poisson comme approximation de la loi binômiale dans le schéma suivant :

Soit A un événement de probabilité p très faible (en pratique, $p < 0,1$) que l'on essaie d'obtenir quelques fois en répétant l'expérience un grand nombre de fois (en pratique $n > 50$).
Soit X une variable aléatoire représentant le nombre de réalisations de A.

$$B(n,p) \approx e^{-\lambda}\frac{\lambda^k}{k!} \quad \text{où} \quad \lambda = np$$

THEOREME : Soit (X_n) une suite de variables binôminales $B(n,p)$ telle que $n \to +\infty$ et $p \to 0$ de manière à ce que np tende vers une limite λ (finie), alors X_n converge en loi vers $\mathscr{P}(\lambda)$.
Preuve :

$$\begin{aligned}
P[X_n = k] &= C_n^k p^k (1-p)^{n-p} \\
&= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= \frac{n(n-1)\dots(n-k+1)}{k!} p^k (1-p)^{n-k} \\
&= \frac{(np)^k}{k!} \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right)(1-p)^{n-k}
\end{aligned}$$

On a

$$\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right) \xrightarrow[n\to+\infty]{} 1$$

et

$$(1-p)^{-k} \xrightarrow[p\to 0]{} 1 \quad \text{et} \quad (1-p)^n \underset{np\to\lambda}{\sim} \left(1-\frac{\lambda}{n}\right)^n \xrightarrow[n\to+\infty]{} e^{-\lambda}$$

Donc

$$C_n^k p^k (1-p)^{n-p} \xrightarrow[n\to+\infty, p\to 0]{} \frac{\lambda^k}{k!} e^{-\lambda}$$

►*Exemple :*
Calculer $E(X)$ et $V(X)$, X suivant la loi de Poisson. On a

$$\begin{aligned}
E(X) &= \sum_{k=0}^{+\infty} k P[X = k] \\
&= \sum_{k=0}^{+\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\
&= e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!} \\
&= e^{-\lambda} \lambda \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= e^{-\lambda} \lambda e^{\lambda} = \lambda
\end{aligned}$$

et

$$
\begin{aligned}
E(X^2) &= \sum_{k=0}^{+\infty} k^2 P[X=k] \\
&= \sum_{k=0}^{+\infty} k^2 e^{-\lambda}\frac{\lambda^k}{k!} \\
&= e^{-\lambda}\sum_{k=1}^{+\infty} k\frac{\lambda^k}{(k-1)!} \\
&= e^{-\lambda}\sum_{k=1}^{+\infty}(k-1+1)\frac{\lambda^k}{(k-1)!} \\
&= e^{-\lambda}\left(\sum_{k=1}^{+\infty}\frac{\lambda^k}{(k-2)!}+\sum_{k=1}^{+\infty}\frac{\lambda^k}{(k-1)!}\right) \\
&= e^{-\lambda}\left(\lambda^2\sum_{k=1}^{+\infty}\frac{\lambda^{k-2}}{(k-2)!}+\lambda\sum_{k=1}^{+\infty}\frac{\lambda^{k-1}}{(k-1)!}\right) \\
&= e^{-\lambda}(\lambda^2 e^{\lambda}+\lambda e^{\lambda}) \\
&= \lambda^2+\lambda
\end{aligned}
$$

Donc

$$V(X) = E(X^2) - E^2(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

1.5.5 Loi hypergéométrique $H(N, n, p)$

Soit une population de N individus parmi lesquels une proportion p possède un certain caractère. On prélève un échantillon de n individus parmi cette population, le tirage pouvant s'effectuer d'un seul coup, ou au fur et à mesure, mais sans remise.
Soit X le nombre d'individus de l'échantillon possédant le caractère. On a

$$P[X=x] = \frac{C_{Np}^{x} C_{N-Np}^{n-x}}{C_N^n}$$

2 Approximations

Le but de ce chapitre est de donner les premières notions de la théorie de l'approximation permettant d'aborder la résolution de problèmes tels que :

1. Etant donnée une fonction continue sur $[a,b]$, déterminer dans l'espace des polynômes de degré n celui qui rend la quantité $\max_{a\leq X\leq b}|f(x)-P(x)|$ la plus petite possible.

2. Construire un polynôme d'interpolation P tel que $P(x_i)=f(x_i), \forall i\in [\![0..n]\!]$.

2.1 Approximation uniforme

Soit $E=\mathscr{C}([a,b])=\{f:[a,b]\to\mathbb{R} \text{ continue }\}$. E est un espace vectoriel normé avec $\|f\|=\max_{a\leq x\leq b}|f(x)|$. Si ε_n est un sous-espace vectoriel de dimension n de E, la meilleure approximation uniforme $\Phi^*\in\varepsilon_n$ de $f\in E$ est donc la fonction définie par :

$$\|f-\Phi^*\|=\min_{\Phi\in\varepsilon_n}\left(\max_{a\leq x\leq b}|f(x)-\Phi(x)|\right)=\min_{\Phi\in\varepsilon_n}\|f-\Phi\|$$

Soit $\{\varphi_1,\ldots,\varphi_n\}$ une base de ε_n, $\Phi^*=\sum_{i=1}^{n}a_i^*\varphi_i$. Déterminer Φ^* revient a calculer les cœfficients a_i^*.

2.1.1 Polynômes de Chebyshev

Ces polynômes sont définies par

$$\begin{cases} T_{n+1}(x) & = & 2xT_n(x)-T_{n-1}(x) \\ T_0(x)=1 & , & T_1(x)=x \end{cases}$$

Démontrer les propriétés suivantes :

1. $T_n(x)=\cos(n\arccos(x))=\cos(n\Theta)$ où $\Theta=\arccos(x),\ -1\leq x\leq 1$.

   Par récurrence, montrons que $T_n(x)=\cos(n\arccos(x))=\cos(n\Theta)$ où $\Theta=\arccos(x), -1\leq x\leq 1$:
   – $n=0\Rightarrow T_n(x)=1=\cos(0)$ et $n=1\Rightarrow T_n(x)=x=\cos(\Theta)$.

   – Supposons que $T_k(x)=\cos(k\Theta), \forall 0\leq k\leq n$.

$$\begin{aligned} T_{n+1}(x) &= 2xT_n(x)-T_{n-1}(x) \\ &= 2x\cos(n\Theta)-\cos((n-1)\Theta) \\ &= 2\cos(\Theta)\cos(n\Theta)-(\cos(n\Theta)\cos(\Theta)+\sin(n\Theta)\sin(\Theta)) \\ &= \cos(\Theta)\cos(n\Theta)-\sin(n\Theta)\sin(\Theta) \\ &= \cos((n+1)\Theta) \end{aligned}$$

2. Le cœfficient dominant est $a_n=2^{n-1}$.

   Montrons que $a_n=2^{n-1}, \forall n\geq 1$
   – $T_1=x\Rightarrow a_1=2^{1-1}=1$
   – Supposons $a_k=2^{k-1}$

$$\begin{aligned} T_{n+1}(x) &= 2xT_n(x)-\underbrace{T_{n-1}(x)}_{d^\circ=n-1} \text{ avec } T_n(x)=2^{n-1}x^n+\underbrace{Q_{n-1}(x)}_{d^\circ\leq n-1} \\ &= 2x(2^{n-1}x^n+Q_{n-1}(x))-T_{n-1}(x) \\ &= 2^nx^{n+1}+\underbrace{2xQ_{n-1}(x)-T_{n-1}(x)}_{d^\circ\leq n} \end{aligned}$$

   Donc $a_{n+1}=2^n=2^{n+1-1}$.

3. La famille des polynômes de Chebyshev $\{T_0, T_1, \dots, T_n, \dots\}$ est une famille de polynômes orthogonaux 2 à 2 sur $[-1,1]$ relativement à la fonction poids de Chebyshev : $\omega(x) = \dfrac{1}{\sqrt{1-x^2}}$, i.e.

$$< T_n, T_m >= \int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}dx = 0,\ \forall n \neq m$$

Montrons que $\displaystyle\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}dx = 0, \forall n \neq m$ (orthogonalisation des polynômes de Chebyshev relativement à la fonction poids $\omega(x) = \dfrac{1}{\sqrt{1-x^2}}$).

Posons $x = \cos(\Theta)$, $dx = d\cos(\Theta) = -\sin(\Theta)d\Theta$. On a

$$\begin{aligned}
\int_{\pi}^{0} \frac{T_n(\cos(\Theta))T_m(\cos(\Theta))}{\sqrt{1-\cos^2(\Theta)}} - \sin(\Theta)d\Theta &= \int_{0}^{\pi} T_n(\cos(\Theta))T_m(\cos(\Theta)\frac{\sin(\Theta)}{|\sin(\Theta)|}d\Theta \\
&= \int_{0}^{\pi} \cos(n\Theta)\cos(m\Theta)d\Theta \\
&= \int_{0}^{\pi} \frac{\cos((n+m)\Theta) + \cos((n-m)\Theta)}{2}d\Theta \\
&= \frac{1}{2}\left[\frac{\sin((n+m)\Theta)}{n+m} + \frac{\sin((n-m)\Theta)}{n-m}\right]_0^{\pi} \\
&= 0
\end{aligned}$$

Calcul des normes :

$$\|T_n\| = \left(\int_{-1}^{1} T_n^2(x)\frac{1}{\sqrt{1-x^2}}dx\right)^{\frac{1}{2}}$$

– 1^{er} cas : $n \neq 0$

$$\begin{aligned}
\|T_n\|^2 &= \frac{1}{2}\int_0^{\pi}(\cos(2n\Theta)+1)d\Theta \\
&= \frac{1}{2}\int_0^{\pi}\cos(2n\Theta)d\Theta + \frac{1}{2}\int_0^{\pi}d\Theta \\
&= \frac{\pi}{2} \Rightarrow \|T_n\| = \sqrt{\frac{\pi}{2}}
\end{aligned}$$

– $2^{\text{ème}}$ cas : $n = 0$

$$I = \frac{1}{2}\int_0^{\pi}(1+1)d\Theta = \pi \Rightarrow \|T_0\| = \sqrt{\pi}$$

4. $T_n(x) = +1, -1, +1, -1, \dots$ pour $x_0 = 1, x_1 = \cos\left(\dfrac{\pi}{n}\right), x_2 = \cos\left(\dfrac{2\pi}{n}\right), \dots, x_k = \cos\left(\dfrac{k\pi}{n}\right)$ (propriété d'alternance).

   $T_n(x) = 1; -1; 1; -1; 1; -1 \dots$ pour $x_k = \frac{k\pi}{n}$

$$T_n(x_0) = \cos(0) = 1, \quad T_n(x_1) = \cos\left(\frac{n\pi}{n}\right) = \cos(\pi) = -1, \quad T_n(x_2) = \cos\left(\frac{2\pi}{n}\right) = 1 \dots$$

2.1.2 Meilleure approximation uniforme de la fonction nulle

THEOREME Chebyshev: Dans l'ensemble des polynômes de degré n ayant le cœfficient de tête égal à 1, c'est $T_n^* = \dfrac{T_n}{2^{n-1}}$ qui réalise la meilleure approximation uniforme de la fonction nulle sur $[-1,1]$. Ici,

$$\|T_n^*\| = \max_{-1 \leq x \leq 1} |T_n^*(x)|$$

Le théorème traduit que

$$\|T_n^*\| = \min_{R_n \in \mathscr{P}_n} \|R_n\| = \frac{1}{2^{n-1}} \quad \text{où} \quad \mathscr{P}_n = \{X^n + a_{n-1}X^{n-1} + \dots + a_0\}$$

Preuve Chebyshev: Supposons le contraire, i.e. $\exists R_n \in \mathscr{P}_n$ tel que $\|R_n\| < \|T_n^*\| \Rightarrow \|R_n\| < \frac{1}{2^{n-1}}$.

Soit $T_n^* - R_n = P_{n-1}, d°(P_{n-1}) \leq n-1$.

$$\begin{array}{rcl} P_{n-1}(x_0 = 1) = T_n^*(1) - R_n(1) = \frac{1}{2^{n-1}} - R_n(1) & > & 0 \\ P_{n-1}\left(x_1 = \cos\left(\frac{\pi}{n}\right)\right) = T_n^*(x_1) - R_n(x_1) = -\frac{1}{2^{n-1}} - R_n(x_1) & < & 0 \\ \vdots & > & 0 \\ \vdots & < & 0 \\ P_{n-1}(x_n) & \dots & 0 \end{array}$$

Donc P_{n-1} possède au moins n racines dans l'intervalle $[-1, 1]$ en utilisant la propriété d'alternance. Ce qui est absurde car son degré ne dépasse pas $n-1$. Donc c'est Chebyshev normalisé qui a la plus petite norme.

2.1.3 Meilleure approximation uniforme d'une fonction continue avec $f \neq 0$

$P_n^* \in \mathbb{P}_n = \{$ polynômes de $d° \leq n\}$. P_n^* est la meilleure approximation uniforme de $f \in E$ ssi

$$\|f - P_n^*\| = \min_{P_n \in \mathbb{P}} \|f - P_n\|$$

THEOREME : Si $P_n \in \mathbb{P}$ est tel que la fonction erreur $\varepsilon_n = f - P_n$ atteint les valeurs extrêmes alternées $M, -M, M, -M, M, \dots,$ (où $M = \|\varepsilon_n\|$) en au moins $n+2$ points $x_1, x_2, \dots, x_n+2$ de $[a, b]$ alors $P_n = P_n^*$.

Preuve : Supposons $\exists Q_n \in \mathbb{P}_n$ tel que $\|f - Q_n\| < \|f - P_n\| = \|\varepsilon_n\| = M$. Soit $R_n = Q_n - P_n$, $d°R_n \leq n$. On a $R_n = f - P_n + Q_n - f = \varepsilon_n + Q_n - f$.

$$\begin{array}{rcl} R_n(x_1) = \varepsilon_n(x_1) + Q_n(x_1) - f(x_1) = M + Q_n(x_1) - f(x_1) & > & 0 \\ R_n(x_2) = \varepsilon_n(x_2) + Q_n(x_2) - f(x_2) = -M + Q_n(x_2) - f(x_2) & < & 0 \\ \dots & \dots & \dots \\ R_n(x_{n+2}) = \varepsilon_n(x_{n+2}) + Q_n(x_{n+2}) - f(x_{n+2}) = M + Q_n(x_{n+2}) - f(x_{n+2}) & > & 0 \end{array}$$

R_n change au moins $n+1$ fois de signe sur $[a, b]$ en raison de l'alternance de ε_n, donc $R_n(x)$ possède au moins $n+1$ racines, ce qui est ABSURDE car $d°(R_n) \leq n$.

2.2 Approximation des moindres carrés

Soit E espace vectoriel muni d'un produit scalaire $< f, g >$. E est normé $\|f\| = \sqrt{< f, f >}$.
THEOREME : Une condition nécessaire et suffisante pour que $\Phi^* \in F$ (où $F \subset E$ dimension finie) soit une meilleure approximation de $f \in E$ est que $< f - \Phi^*, \Phi >= 0,\ \forall \Phi \in F$.

2.3 Interpolation

Soit $f(x)$ une fonction continue sur $[a, b]$, on suppose que l'on connaît seulement sa valeur en $n+1$ points x_i de $[a, b]$, $i \in [\![1..n]\!]$. Chercher une fonction φ d'un type choisi à l'avance qui interpole $f(x)$ sur $[a, b]$, c'est déterminer une fonction φ telle que $\varphi(x_i) = f(x_i), \forall i \in [\![1..n]\!]$. En général, on cherche φ dans l'espace des polynômes.

2.3.1 Polynôme d'interpolation de Lagrange

Soit $P_n(x) = \sum_{i=1}^{n+1} L_i(x) f(x_i)$ où les fonctions $L_i(x)$ sont des polynômes de degré n telles que

- $L_i(x_1) = 1$.
- $L_i(x_j) = 0, \forall j \neq i$.

c'est-à-dire $L_i(x_j) = \delta_{i,j}$.

Puisque $L_i(x_j) = 0, \forall j \neq i$, alors $L_i(x) = C(x - x_1)(x - x_2) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_{n+1})$. Or

$$L_i(x_i) = 1 \Rightarrow C = \frac{1}{(x_i - x_1)(x_i - x_2) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_{n+1})}$$

Donc

$$L_i(x) = \prod_{j=1, j\neq i}^{n+1} \frac{x - x_j}{x_i - x_j} \Rightarrow P_n(x) = \sum_{i=1}^{n+1} \prod_{j=1, j\neq i}^{n+1} \frac{x - x_j}{x_i - x_j} f(x_i)$$

On a bien

$$P_n(x_j) = \sum_{i=1}^{n+1} L_i(x_j) f(x_i) = f(x_i)$$

Donc $P_n(x)$ interpole $f(x)$.

►*Exemple :*

Construire le polynôme d'interpolation $P_3(x)$ de f

x_i	0	1	3	4
$f(x_i)$	1	3	2	5

$$P_3(x) = \frac{(x-1)(x-3)(x-4)}{(0-1)(0-3)(0-4)} \times 1 + \frac{(x-0)(x-3)(x-4)}{1(1-3)(1-4)} \times 3 + \frac{x(x-1)(x-4)}{3(3-1)(3-4)} \times 2 + \frac{x(x-1)(x-3)}{4(4-1)(4-3)} \times 5 = \frac{x^3}{2} - \frac{17}{6}x^2 + \frac{1}{3}$$

2.3.2 Erreur d'interpolation

$\varepsilon(x) = f(x) - P_n(x)$. Si la fonction f est $n+1$ fois continûment différentiable sur $[a, b]$ ($f \in \mathscr{C}^{n+1}[a, b]$),

$$\forall x \in [a, b],\ \exists \eta_x \in]a, b[, \varepsilon = \prod_{i=1}^{n+1} (x - x_i) \frac{f^{(n+1)}(\eta_x)}{(n+1)!}$$

Si on pose $M = \max_{a \leq x \leq b} |f^{(n+1)}(x)|$, alors $|\varepsilon(x)| \leq \left| \prod_{i=1}^{n+1} (x - x_i) \right| \frac{M}{(n+1)!}$

2.3.3 Interpolation de Newton

Différences divisées **Définition :** Soit $f(x)$ une fonction dont on connaît les valeurs en $n+1$ points $x_1, x_2, \dots, x_{n+1}$ distincts. On appelle différences divisées d'ordre $0, 1, \dots, n$ de f les expressions suivantes :

$Ordre$	$Notation$	$Dfinition$
0	$f[x_i]$	$f(x_i)$
1	$f[x_i, x_j]$	$\frac{f[x_j] - f[x_i]}{x_j - x_i}$
2	$f[x_i, x_j, x_k]$	$\frac{f[x_j, x_k] - f[x_i, x_j]}{x_k - x_i}$
$\vdots$	$\vdots$	$\vdots$
n	$f[x_1, \dots, x_{n+1}]$	$\frac{f[x_2, \dots, x_{n+1}] - f[x_1, \dots, x_n]}{x_{n+1} - x_1}$

Polynôme de Newton On a

$$f[x, x_1] = \frac{f(x) - f(x_i)}{x - x_1} \Rightarrow f(x) = f(x_1) + (x - x_1) f[x, x_1]$$

puis

$$f[x, x_1, x_2] = \frac{f[x, x_1] - f[x_1, x_2]}{x - x_2} \Rightarrow f[x_1, x_2] = f[x_1, x_2] + (x - x_2) f[x, x_1, x_2]$$

Donc $f(x) = \underbrace{f(x_1) + (x - x_1)f[x_1, x_2]}_{d^\circ = 1} + \underbrace{(x - x_1)(x - x_2)f[x, x_1, x_2]}_{erreur}$. Soit $P_1(x) = f(x_1) + (x - x_1)f[x_1, x_2]$, on a $P_1(x_1) = f(x_1)$, $P_1(x_2) = f(x_1) + (x_2 - x_1)\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f(x_2)$. Donc $P_1(x)$ interpole $f(x)$ aux points x_1 et x_2.

En réitérant le procédé, on obtient enfin

$$f(x) = f(x_1) + (x - x_1)f[x_1, x_2] + (x - x_1)(x - x_2)f[x_1, x_2, x_3] + \cdots$$

$$+(x - x_1)(x - x_2)\dots(x - x_n)f[x_1, \dots, x_{n+1}] + \underbrace{\prod_{i=1}^{n+1}(x - x_i)f[x_1, \dots, x_{n+1}]}_{erreur}$$

Donc

$$f(x) = P_n(x) + \prod_{i=1}^{n+1}(x - x_i)f[x_1, \dots, x_{n+1}]$$

avec

$$P_n(x) = f(x_1) + (x - x_1)f[x_1, x_2] + (x - x_1)(x - x_2)f[x_1, x_2, x_3] + \dots + (x - x_1)(x - x_2)\dots(x - x_n)f[x_1, \dots, x_{n+1}]$$

►*Remarque :*
Si f est un polynôme de degré n, l'erreur sera nulle.

Algorithme de Newton

$$P_n(x) - f(x_1) + \sum_{k=1}^{n}(x - x_1)\dots(x - x_k)f[x_1, \dots, x_{k+1}]$$

On pose $P_0(x) = f(x_1)$ et pour $m \in [\![0..n-1]\!]$,

$$P_{m+1}(x) = P_m(x) + (x - x_1)\dots(x - x_{m+1})f[x_1, \dots, x_{m+2}] \quad \text{avec} \quad f[x_1, \dots, x_{m+1}] = \frac{f[x_1, \dots, x_m] - f[x_2, \dots, x_{m+1}]}{x_1 + x_{m+1}}$$