COURS 23

Solutions rationnelles de récurrences et sommation hypergéométrique indéfinie

Résumé

Un petit noyau d'algorithmes relativement simples permet de trouver les solutions polynomiales et rationnelles de récurrences linéaires. Nous les appliquons dans ce cours pour trouver les sommes indéfinies de suites hypergéométriques. Au cours suivant (Cours 24), ils resserviront pour la recherche de solutions hypergéométriques de récurrences linéaires et pour trouver des sommes définies de suites hypergéométriques.

Le début du cours porte sur la résolution d'une relation de récurrence linéaire en ses solutions polynomiales et rationnelles, c'est-à-dire en ses suites solutions dont le terme général est donné par l'évaluation d'un polynôme, respectivement d'une fraction rationnelle, en l'indice de la suite. La résolution dans ces classes « élémentaires » est la base de toute une algorithmique sur les suites. Dans la suite du cours, l'algorithme de Gosper pour la sommation indéfinie se ramène à des résolutions en solutions rationnelles. Il en est de même au cours 24 pour d'autres applications qui ne sont pas traitées dans le cours, telles la résolution de récurrences dans des classes de solutions plus complexes (sommes emboîtées, quotients de sommes), ou encore la désingularisation et la factorisation d'une récurrence.

Les questions de complexité ne sont pas abordées ici. Le corps $\mathbb{K}$ qui est utilisé dans certains énoncés a toujours caractéristique nulle, et on peut penser que $\mathbb{K}$ est le corps $\mathbb{Q}$ sans que cela limite la portée des idées de ce cours. Le terme « constante » est employé pour désigner un élément du corps $\mathbb{K}$: une constante est alors indépendante de l'indice de sommation (k sur les exemples donnés plus haut).

Pour fixer la notation, la récurrence dont on cherche les zéros est

$$Lu(n) = \sum_{k=0}^{m} a_k(n)u(n+k), \tag{1}$$

où les $a_k(n)$ sont des polynômes de $\mathbb{K}[n]$, et on note d le maximum des degrés des a_k, $k = 0, \ldots, m$.

1. Solutions polynomiales

Le cœur technique de ce cours est dans cette section, organisée par généralité croissante. Les sections qui suivent, bien plus courtes, montrent l'application de la recherche de solutions polynomiales à la recherche de solutions rationnelles et à la sommation.

1.1. Équation homogène. Il s'agit ici de trouver les polynômes $P(n)$ tels que $LP(n) = 0$. Une première observation simple est que L est une application linéaire de $\mathbb{K}[n]_D$, l'espace vectoriel des polynômes de degré au plus D, dans $\mathbb{K}[n]_{D+d}$, pour tout $D \in \mathbb{N}$. Les noyaux des restrictions de L à $\mathbb{K}[n]_D$ pour $D = 0, 1, \ldots$ forment une suite croissante d'espaces vectoriels stationnaire à partir d'un certain indice D_0. L'algorithme consiste donc à trouver une borne sur cet indice (c'est-à-dire sur le degré maximal des polynômes solutions) et à calculer une base de l'espace correspondant.

Exemple 1. Observons que la récurrence

$$Lu(n) = nu(n+1) - (n+100)u(n) = 0$$

a pour solution le polynôme $u(n) = n(n+1)\cdots(n+99)$ de degré 100. Notre but dans cet exemple est de montrer que l'ensemble des solutions polynomiales de cette récurrence est exactement l'ensemble des multiple de u par une constante.

L'application de l'opérateur L sur un monôme n^d donne

$$Ln^d = (d-100)n^d + \cdots,$$

où les points de suspension correspondent à des termes de degré au plus $d-1$. Par linéarité, un polynôme f de degré d autre que 100 est tel que Lf a aussi degré d. Il s'ensuit que les solutions polynomiales ne peuvent avoir que degré 100, d'où le résultat annoncé.

Cet exemple se généralise. Pour y voir plus clair, il est plus commode de récrire les décalages $u(n+k)$ de la suite initiale en terme de différences finies. Ainsi, on note $\Delta u(n) = u(n+1) - u(n)$ et, par récurrence, $\Delta^{k+1}u(n) = (\Delta^k u)(n+1) - (\Delta^k u)(n)$. La récurrence à annuler prend la forme

$$Lu = \sum_{k=0}^{m} b_k(n)\Delta^k u = 0$$

pour de nouveaux polynômes b_k. L'opérateur Δ fait décroître de 1 exactement le degré des polynômes. Ainsi

$$\deg(\Delta^k P) \leq \max(\deg P - k, 0),$$

avec égalité tant que $\Delta^k P$ n'est pas nul, et donc $\deg(LP) \leq \deg P + \max_k\{\deg(b_k) - k\}$. Soient alors

$$b := \max_k\{\deg(b_k) - k\}, \qquad E := \{\, k \mid \deg(b_k) - k = b \,\}.$$

Ces quantités vont servir à fournir une borne sur le degré des solutions.

Exemple 2. La récurrence de l'exemple précédent se récrit

$$(n\Delta - 100)(u_n) = 0;$$

l'entier b vaut 0 et l'ensemble E est $\{0, 1\}$.

Soit D le degré d'une solution. La discussion distingue deux cas :
— soit $D + b < 0$, et alors $-(b+1)$ est une borne sur D ;
— sinon le coefficient de degré $D + b$ dans $L(n^D + \cdots)$ vaut

$$\sum_{k\in E} \mathrm{lc}(b_k)D(D-1)\cdots(D-k+1),$$

où lc désigne le coefficient de tête (*leading coefficient*). Cette expression, vue comme un polynôme en D, s'appelle *le polynôme indiciel* de la récurrence, lequel est non nul.

Cette discussion mène au résultat suivant.

PROPOSITION 1. *Une borne sur le degré des solutions polynomiales de l'opérateur* $L = \sum b_k(n)\Delta^k$ *est donnée par le maximum de* $-(b+1)$ *et de la plus grande racine entière positive du polynôme indiciel de* L.

Un algorithme simple consiste alors à calculer cette borne et à rechercher ensuite les solutions par un calcul d'algèbre linéaire.

EXERCICE 1. Trouver les solutions polynomiales de la récurrence

$$3u(n+2) - nu(n+1) + (n-1)u(n) = 0.$$

1.2. Équation inhomogène. L'opérateur L étant un endomorphisme linéaire de $\mathbb{K}[n]$, l'équation inhomogène $Lu(n) = Q(n)$ ne peut avoir de solutions polynomiales que si Q est un polynôme. La discussion qui précède montre qu'une borne sur le degré de ces solutions est donnée par le maximum de $\deg(Q) - b$ et de la borne obtenue pour la partie homogène. Le reste du calcul est à nouveau réduit à de l'algèbre linéaire en dimension finie. Une autre manière d'aboutir à cette borne consiste à appliquer $\Delta^{\deg Q+1}$ aux deux membres de l'équation inhomogène pour la rendre homogène et appliquer le calcul précédent.

EXERCICE 2. Montrer que pour $\alpha = 0$ la récurrence

$$3u(n+2) - nu(n+1) + (n-1)u(n) = -2(n-\alpha)^3$$

n'a pas de solution, alors que pour $\alpha = 5$, elle en a.

1.3. Équation inhomogène paramétrée. Le problème est ici de trouver s'il existe un polynôme $u(n)$ et des constantes $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ tels que

$$Lu(n) = \lambda_1 Q_1(n) + \cdots + \lambda_k Q_k(n),$$

les polynômes Q_i étant donnés. La borne sur le degré de $u(n)$ vient d'être donnée. Il ne reste qu'à observer que les équations qu'il faut ensuite résoudre sont linéaires non seulement en les coefficients du polynôme, mais aussi en les λ_i. Une fois encore, le problème est ainsi réduit à un calcul d'algèbre linéaire.

EXERCICE 3. Résoudre en (u, λ, μ) la récurrence

$$3u(n+2) - nu(n+1) + (n-1)u(n) = \lambda n^3 + \mu n^2.$$

2. Solutions rationnelles : Algorithme d'Abramov

2.1. Équation homogène. Le problème est maintenant de trouver les solutions rationnelles de l'équation $Lu(n) = 0$. L'algorithme procède en deux temps : d'abord le calcul d'un multiple des dénominateurs des solutions, ensuite un changement de fonction inconnue pour ramener la recherche du numérateur à celle de solutions polynomiales d'une nouvelle équation linéaire, problème qui vient d'être traité.

Pour trouver un multiple du dénominateur, on peut observer que les pôles de $u(n), u(n+1), \dots, u(n+m)$ sont décalés les uns des autres. Si u n'a pas deux pôles différant d'un entier, il ne peut donc y avoir de solution rationnelle que si le dénominateur Q de u vérifie

$$Q(n) \mid a_0(n), \quad Q(n+1) \mid a_1(n), \dots, \quad Q(n+m) \mid a_m(n),$$

et donc dans ce cas un multiple du dénominateur est donné par

$$\mathrm{pgcd}\big(a_0(n), a_1(n-1), \dots, a_m(n-m)\big).$$

En général cependant, il peut y avoir des racines de Q qui diffèrent d'un entier et cet argument n'est plus valable. Cependant, si α et β sont deux racines de Q telles que $\alpha - \beta = k \in \mathbb{N}$ est maximal, alors nécessairement $a_0(\alpha) = 0$ et $a_m(\beta - m) = 0$. L'algorithme suivant utilise cette idée pour fournir un multiple de Q.

Multiple du dénominateur

Entrée : la récurrence $Lu(n) = 0$, avec L donné par l'équation (1) ;

Sortie : un multiple du dénominateur des solutions rationnelles.

1. Calculer le polynôme
$$R(h) = \mathrm{Res}_n\big(a_0(n+h), a_m(n-m)\big).$$
2. Si R n'a pas de racines dans $\mathbb{N}$, alors renvoyer le polynôme 1 ; sinon, soit $h_1 > h_2 > \dots > h_m \geq 0$ ses racines entières positives. Initialiser Q à 1, A à $a_0(n)$, B à $a_m(n-m)$;
3. Pour $i = 1, \dots, m$ faire
   $g(n) := \mathrm{pgcd}\big(A(n+h_i), B(n)\big)$;
   $Q(n) := g(n)g(n-1)\cdots g(n-h_i)Q(n)$;
   $A(n) := A(n)/g(n-h_i)$;
   $B(n) := B(n)/g(n)$.
4. Renvoyer Q.

Le polynôme dans l'étape (1) est un résultant particulier qui peut se calculer efficacement comme une somme composée. Ce calcul a été présenté dans le cours 3 comme application du calcul rapide de l'exponentielle des séries.

Il est possible de raffiner un peu cet algorithme pour obtenir des multiples de degré moindre du dénominateur, c'est ce que fait Abramov dans [2] et aussi dans certains cas en tenant compte de tous les a_i dans [1]. Une manière plus directe d'aboutir à ces raffinements à été donnée par van Hoeij [4].

Exercice 4. Trouver les solutions rationnelles des récurrences suivantes :

$$(n+1)u_{n+1} - nu_n = 0,$$
$$(n+1)(n+3)u_{n+2} - 2(n+2)nu_{n+1} + (n-1)(n+1) = 0,$$
$$(n+3)(n+2)(n^2+6n+4)u(n+2) - (n+1)(3n^3+27n^2+64n+48)u(n+1) \\ + 2n^2(n^2+8n+11)u(n) = 0.$$

2.2. Équation inhomogène. Comme pour les solutions polynomiales, l'image d'une fraction rationnelle par L étant une fraction rationnelle, il ne peut y avoir solution rationnelle de l'équation inhomogène que si le membre droit est rationnel. Dans ce cas, réduire l'équation au même dénominateur mène à une équation pour laquelle un multiple du dénominateur des solutions rationnelles est obtenu en considérant la partie homogène. Après changement de fonction inconnue, le calcul se ramène à la recherche de solutions polynomiales d'une équation inhomogène.

2.3. L'ordre 1. Lorsque la récurrence est d'ordre 1, il est en outre possible de prédire des facteurs du numérateur. En effet, dans la récurrence

$$a(n)u(n+1)+b(n)u(n)=c(n),$$

si b et c ont une racine commune, qui n'est pas un pôle de u et qui n'est pas une racine de a, alors elle est nécessairement racine de $u(n+1)$. De même, si $a(n)$ et $c(n)$ ont une racine commune qui n'est pas un pôle de $u(n+1)$ et qui n'est pas racine de b, alors elle est racine de $u(n)$.

Ces calculs préalables permettent de réduire le degré des coefficients de l'équation dont on recherche ensuite les solutions polynomiales.

2.4. Équation inhomogène paramétrée. Le même argument que ci-dessus mène à une conclusion similaire : un multiple du dénominateur s'obtient par les méthodes du cas homogène, et le changement de fonction inconnue ramène à la résolution d'une équation inhomogène paramétrée.

3. Sommation hypergéométrique indéfinie. Algorithme de Gosper

La recherche de formes closes pour la sommation est souvent naturellement posée en terme de suites hypergéométriques. Après quelques définitions et propriétés de ces suites, nous abordons leur sommation.

3.1. Suites hypergéométriques.

Definition 1. On appelle *suite hypergéométrique* une suite P-récursive vérifiant une récurrence linéaire *d'ordre 1*.

Une telle récurrence, de la forme

$$u_{n+1}=\frac{P(n)}{Q(n)}u_n, \tag{2}$$

où P et Q sont des polynômes admet une solution explicite à l'aide de la fonction Γ. Cette fonction est classiquement définie pour $\Re(z)>0$ par l'intégrale d'Euler

$$\Gamma(z)=\int_0^{+\infty}t^{z-1}e^{-t}\,dt.$$

Une intégration par parties montre l'équation fonctionnelle

$$\Gamma(z+1)=z\Gamma(z).$$

Une première conséquence de cette équation est de fournir un prolongement méromorphe de Γ à $\mathbb{C}\setminus\mathbb{Z}^-$. Grâce à la valeur facile à calculer $\Gamma(1)=1$, cette équation montre aussi que pour tout entier positif n, $\Gamma(n+1)=n!$. Enfin, l'équation fonctionnelle permet de résoudre la récurrence (2) sous la forme

$$u_n=u_0\left(\frac{\mathrm{lc}(P)}{\mathrm{lc}(Q)}\right)^n\prod_{P(\alpha)=0}\frac{\Gamma(n-\alpha)}{\Gamma(-\alpha)}\prod_{Q(\beta)=0}\frac{\Gamma(-\beta)}{\Gamma(n-\beta)},$$

où $\mathrm{lc}(p)$ désigne le coefficient de tête du polynôme p.

Cette formule est valable tant que les valeurs des points où est évaluée Γ ne sont pas des entiers négatifs ou nuls, c'est-à-dire pour $\alpha\notin\mathbb{N}$ et $\beta\notin\mathbb{N}$. Elle continue d'être valable lorsque $\alpha\in\mathbb{N}$ à condition d'interpréter le premier quotient comme une limite : d'après l'équation fonctionnelle, pour k et n deux entiers positifs ou nuls,

$$\lim_{s\to k} \frac{\Gamma(n-s)}{\Gamma(-s)} = \begin{cases} 0, & \text{si } n > k\,; \\ (-1)^n \frac{k!}{(k-n)!}, & \text{si } n \leq k. \end{cases}$$

Cette limite correspond bien à ce qui est attendu : si $k \in \mathbb{N}$ est une racine de P, alors $u_{k+1} = 0$ et par suite $u_n = 0$ pour $n > k$.

Lorsque $\beta \in \mathbb{N}$, la même limite peut être utilisée tant que $n \leq \beta$, et la suite cesse d'être définie pour $n > \beta$.

Les suites hypergéométriques jouent un rôle important en analyse classique, où elles apparaissent comme coefficients de Taylor des séries introduites par la définition suivante.

Definition 2. On appelle *série hypergéométrique généralisée* et on note

$${}_pF_q\left(\begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \middle| \, x\right)$$

la série

$$\sum_{n\geq 0} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \frac{x^n}{n!},$$

où la notation $(a)_n$ représente le produit $a(a-1)\cdots(a-n+1)$.

Des cas particuliers de séries qui s'expriment à l'aide de séries hypergéométriques généralisées sont les séries $\exp(x)$, $\log(1+x)$, les dilogarithmes et polylogarithmes, les fonctions J_ν de Bessel, les fonctions d'Airy, etc.

Exercice 5. Récrire l'identité de Dixon en terme de valeur d'une ${}_3F_2$ en 1.

3.2. Algorithme de Gosper. Étant donnée une suite hypergéométrique $u(n)$, le problème de sommation indéfinie hypergéométrique de $u(n)$ consiste à déterminer s'il existe une autre suite hypergéométrique $U(n)$ telle que $U(n+1) - U(n) = u(n)$ et si oui, la calculer.

Une observation simple est formulée dans le lemme suivant.

Lemme 1. *Si $U(n)$ est une suite hypergéométrique et L un opérateur de récurrence linéaire à coefficients polynomiaux, alors il existe une fraction rationnelle $r(n)$ telle que $LU(n) = r(n)U(n)$.*

Démonstration. $r(n)$ n'est autre que le reste de la division euclidienne de L par le polynôme unitaire de degré 1 donnant la récurrence qui annule $U(n)$. Plus explicitement, si $U(n)$ est hypergéométrique, par définition, il existe une fraction rationnelle $R(n)$ telle que $U(n+1) = R(n)U(n)$ et donc par récurrence $U(n+k) = R(n+k-1)\cdots R(n)U(n)$ pour tout k. Le résultat s'en déduit additionnant les contributions. □

Ce lemme entraîne qu'une somme hypergéométrique $U(n)$ de $u(n)$ doit être le produit de $u(n)$ par une fraction rationnelle $R(n)$. Diviser la récurrence $U(n+1) - U(n) = u(n)$ par $u(n)$ réduit alors le calcul à celui de la recherche de solutions rationnelles de l'équation inhomogène

$$R(n+1)\frac{u(n+1)}{u(n)} - R(n) = 1, \tag{3}$$

d'ordre 1, problème traité dans la section précédente, particulièrement en 2.3. L'algorithme ainsi obtenu est connu sous le nom d'algorithme de Gosper [3].

Exercice 6. Calculer une somme indéfinie de $2^k(k-1)/k/(k+1)$.

Exemple 3. L'algorithme de Gosper permet également de donner des réponses négatives. Voici en détail comment il permet de prouver par l'absurde que $\sum_{k=1}^n 1/k!$ n'est pas hypergéométrique.

Si $S(n)$ est une telle somme, elle doit être le produit de $1/n!$ par une fraction rationnelle $r(n)$. Cette fraction satisfait donc la récurrence

$$S(n+1) - S(n) = \frac{1}{(n+1)!} = \frac{r(n+1)}{(n+1)!} - \frac{r(n)}{n!}.$$

En chassant les dénominateurs, il reste

$$r(n+1) - (n+1)r(n) = 1.$$

Le résultat de l'algorithme « Multiple du dénominateur » d'Abramov montre que r doit être un polynôme : les pôles de plus petite partie réelle de r doivent être annulés par 1. Enfin, r ne peut pas non plus être un polynôme : son terme de plus haut degré ne disparaît pas par évaluation de la récurrence.

3.3. Sommation paramétrée. Étant données des suites $u_1(n), \dots, u_k(n)$ hypergéométriques, il s'agit de déterminer, si elles existent, une suite hypergéométrique $S(n)$ et des constantes $\lambda_1, \dots, \lambda_k$ tels que

$$S(n+1) - S(n) = \lambda_1 u_1(n) + \cdots + \lambda_k u_k(n).$$

D'après la discussion précédente, si $\lambda_i \lambda_j \neq 0$ alors u_i et u_j sont similaires au sens où $u_i(n)/u_j(n)$ est rationnel. Il suffit donc de considérer le cas où tous les u_i sont similaires à une même suite $u(n)$. Dans ce cas, le membre droit de (3) est remplacé par une combinaison linéaire des λ_i à coefficients des fractions rationnelles et la méthode de la section précédente s'applique.

Bibliographie

[1] Abramov (S. A.). – Rational solutions of linear differential and difference equations with polynomial coefficients. *USSR Computational Mathematics and Mathematical Physics*, vol. 29, n° 11, 1989, pp. 1611–1620. – Translation of the Zhurnal vychislitel'noi matematiki i matematichesckoi fiziki.

[2] Abramov (S. A.). – Rational solutions of linear difference and q-difference equations with polynomial coefficients. In Levelt (A. H. M.) (editor), *Symbolic and Algebraic Computation.* pp. 285–289. – ACM Press, New York, 1995. Proceedings of ISSAC'95, July 1995, Montreal, Canada.

[3] Gosper (R. William). – Decision procedure for indefinite hypergeometric summation. *Proceedings of the National Academy of Sciences USA*, vol. 75, n° 1, January 1978, pp. 40–42.

[4] van Hoeij (Mark). – Rational solutions of linear difference equations. In *Proceedings of the 1998 International Symposium on Symbolic and Algebraic Computation (Rostock).* pp. 120–123. – ACM, New York, 1998.