Approximation des fonctions

Préparation agrégation, option C

05/06

Ce texte présente quelques méthodes d'approximation de fonctions qui servent en particulier à calculer les fonctions classiques en utilisant des fonctions plus simples (addition, multiplication, division, évaluation de polynômes).

1 La méthode de Newton

Pour calculer $F(a)$, on résoud une équation du type $f(x) = 0$ dont $x = F(a)$ est solution, et telle que le calcul de g soit possible (ou plus efficace). Par exemple pour calculer la racine n-ième d'un réel positif a, on prend $f(x) = x^n - a$.

On utilise souvent des fonctions dont la convexité est connue, pour pouvoir appliquer le

Theorem 1 *Soit f de classe C^2 telle que $f(r) = 0, f'(r) > 0$ et $f'' \geq 0$ sur $[r, b]$. Alors pour tout $u_0 \in [r, b]$ la suite de la méthode de Newton*

$$u_{n+1} = u_n - \frac{f(u_n)}{f'(u_n)},$$

est définie, décroissante, minorée par r et converge vers r.

Le théorème des accroissements finis permet alors de calculer la précision de la racine approchée trouvée en fonction de l'image de cette racine :

$$0 \leq u_n - r \leq \frac{f(u_n)}{f'(r)}$$

Typiquement le nombre de chiffres significatifs corrects est multiplié par 2 à chaque itération. C'est par exemple cette méthode qui permet de calculer la racine carrée de $a > 1$:

$$u_{n+1} = \frac{1}{2}(u_n + \frac{a}{u_n})$$

Le calcul de la valeur initiale u_0 de la suite a une grande importance pour le nombre d'itérations à effectuer, on utilise diverses méthodes, par exemple une des méthodes présentées ci-dessous, la méthode de Newton pouvant servir à améliorer la précision (en particulier dans le cadre des flottants multiprécision).

2 Les polynomes

2.1 Les séries entières.

A l'intérieur du rayon de convergence d'une fonction analytique, on peut utiliser le développement en séries entières de la fonction. On calcule une majoration à priori du reste sur la zone considérée, puis on cherche le rang à partir duquel ce reste est plus petit que la précision souhaitée. Il y a trois grandes méthodes de majoration : le reste de Taylor (lorsque la dérivée n-ième de la fonction est facile à calculer et majorer), les séries alternées ou l'utilisation de la majoration suivante : s'il existe un point x_0 tel que $|a_n x_0^n|$ est borné par M, alors pour $|x| < |x_0|$ on peut majorer le reste de la série au rang n par

$$|R_n| \leq M \frac{|\frac{x}{x_0}|^n}{1 - |\frac{x}{x_0}|}$$

Le calcul de f à la précision souhaitée se ramène ainsi à une évaluation de polynome.

Certaines séries se pretent bien à cette méthode, par exemples les fonctions transcendantes directes (exp, sin, cos), d'autres moins bien (atan, ln) car la convergence est un peu plus lente (pour calculer n bits de mantisse, il faut $O(n)$ termes du développement au lieu de $O(n/\ln(n))$ termes). Les systèmes de calcul en flottant multiprécision utilisent les séries entières (aprés réduction d'argument) pour l'exponentielle ou le calcul de $\ln(2)$. Richard Brent a montré en 1976 qu'on pouvait évaluer les fonctions usuelles en $O(M(n)\ln(n))$ opérations, où $M(n)$ désigne la complexité de la multiplication des entiers de taille n (par exemple $M(n) = O(n \ln(n) \ln(\ln(n)))$ pour la méthode de Schönhage-Strassen). Cela fait intervenir

2.2 Interpolation polynomiale (Lagrange, ...)

Cette fois, on approche f par un polynôme qui est égal à f en certains points (on peut aussi imposer que les dérivées de f et P soient égales en certains points). C'est la méthode d'interpolation de Lagrange, pour laquelle on a une majoration de la différence entre f et P : Soit f une fonction $n+1$ fois dérivable sur un intervalle $I = [a, b]$ de $\mathbb{R}$, $x_0, ..., x_n$ des réels distincts de I. Soit P le polynome de Lagrange donné par les x_j et $y_j = f(x_j)$. Pour tout réel $x \in I$, il existe un réel $\xi_x \in [a, b]$ (qui dépend de x) tel que :

$$f(x) - P(x) = \frac{f^{[n+1]}(\xi_x)}{(n+1)!} \prod_{j=0}^{n} (x - x_j)$$

Le choix des points x_i (à n fixé) rendant la majoration ci-dessus optimale dépend de l'utilisation que l'on veut faire de P. On peut prendre des points équirépartis sur l'intervalle, puis s'en servir pour calculer une valeur approchée d'intégrale, ce sont les méthodes de Newton-Cotes. Ce choix n'est pas optimal pour l'ordre de la méthode de quadrature, si on laisse le choix des x_i libres, on montre qu'il existe un choix rendant la méthode d'intégration exacte pour tous les polynômes de degré plus petit que $2n+1$ (méthodes de Gauss).

On peut aussi souhaiter qu'une norme issue d'un produit scalaire (par exemple la norme L^2 avec un poids) de $f-P$ soit minimale, ce qui revient à projeter f sur un sous-espace engendré par des polynômes de degré plus petit qu'un degré fixé, il s'agit alors d'en calculer une base (de polynomes orthogonaux). Le même principe permet de construire les séries de Fourier en prenant des polynômes trigonométriques.

3 Les approximants de Padé

Pour approcher f on utilise une fraction rationnelle (dont le degré du numérateur et du dénominateur sont bornés) à la place de polynomes. Le calcul du numérateur et du dénominateur fait intervenir l'algorithme de Bézout.

Plus précisément, on cherche une fraction ayant même développement de Taylor que f en un point donné, 0 pour fixer les idées, à un ordre fixé. On doit donc résoudre

$$P = \frac{N}{D} + O(x^{p+1}) \tag{1}$$

où P est donné de degré inférieur ou égal à p, N et D sont les inconnues de degré inférieur ou égal à n et d, avec $n+d=p$ et D premier avec x. D admet donc un inverse modulo x^p, si on multiplie par D, on obtient :

$$DP = N \pmod{x^{p+1}}$$

On effectue l'algorithme de Bézout pour $A = x^{p+1}$ et $B = P$:

$$AU_k + BV_k = R_k, \quad U_0 = 1 = V_1, U_1 = 0 = V_0, R_0 = x^{p+1}, R_1 = P$$

et on s'arrête au premier rang k tel que le degré de R_k soit inférieur ou égal à n. On montre ensuite que :

$$\deg(V_k) + \deg(R_{k-1}) = ... = \deg(V_1) + \deg(R_0) = \deg(A) = p+1$$

ce qui entraine que le degré de V_k est inférieur ou égal à $p+1-(n+1)=d$. On obtient ainsi une solution de $DP = N \pmod{x^{p+1}}$ ($D = V_k$ et $N = R_k$). Si V_k est premier avec x, on conclut. On peut montrer que si V_k et R_k ne sont pas premiers entre eux, alors (1) n'admet pas de solutions avec ces conditions de degré.

Par exemple pour e^x en prenant $p=6$ et $n=3$, on trouve $(x^3+12x^2+60\times x+120)/(-x^3+12x^2-60x+120)$.

On peut montrer qu'il existe un $\varepsilon > 0$ tel que l'erreur absolue entre f et son approximant de Padé sur $[-\varepsilon, \varepsilon]$ est atteint aux bornes de l'intervalle, et on peut calculer explicitement une valeur de ε.

Cette méthode est très utile en pratique, par exemple pour le calcul de la fonction exponentielle en précision machine, car elle nécessite d'une part moins d'opérations (à cause de la symétrie numérateur/dénominateur) et fournit de plus une approximation de meilleure qualité que le développement de Taylor de e^x (à l'ordre 6 ici).

On peut aussi approcher une fonction par une fraction rationnelle interpolant la fonction en certains points (comme pour les polynômes de Lagrange mais avec une fraction au lieu d'un polynôme). Le problème à résoudre est analogue mais modulo un produit de $x-x_i$ au lieu de x^{t+1}. On parle d'approximant de Cauchy.

4 Méthodes mixtes

Pour le calcul des fonctions usuelles, on utilise la plupart du temps une combinaison des méthodes ci-dessus et des propriétés de ces fonctions.

Par exemple, pour calculer le logarithme d'un flottant en base 2, on calcule une fois pour toutes $ln(2)$ à la précision souhaitée, puis on utilise la représentation $a = 2^e(1+m)$, donc $\ln(a) = e\ln(2) + \ln(1+m)$ avec $m \in [0, 1[$. On peut alors utiliser le développement en séries de $1+m$ mais si m est proche de 1, la convergence sera très lente (controlée par la nature de série alternée). On peut alors décider de calculer une approximation en prenant le développement à l'ordre disons 9, puis on effectue quelques itérations de Newton de l'équation $e^x - a = 0$ en prenant comme valeur initiale cette valeur (par excès) de $\ln(a)$. Cette méthode pourrait en particulier être utilisée pour calculer le logarithme multi-précision à partir d'une valeur approchée par excès en simple précision (en réalité, on utilise la moyenne arithmético-géométrique qui est utilisée, elle fait intervenir un encadrement d'une intégrale elliptique par $ln(x)$ en utilisant les valeurs de $\ln(2)$ et π).

On peut aussi utiliser des méthodes ad hoc pour calculer la valeur initiale, par exemple pour la fonction arctangeante sur un microprocesseur calculant en base 10, on stocke les valeurs de $\arctan(10^{-k})$ pour les premières valeurs de k. Soit a l'angle dont on veut calculer l'arctangente, on se ramène d'abord à a dans l'intervalle $[0, 1]$. Soit $\alpha = \arctan(a)$. On pose alors $x = a$ et $y = 1$. Soit $\beta = \arctan(10^{-1})$. On calcule

$$\tan(\alpha - \beta) = \frac{x/y - 1/10}{1 + 1/10x/y} = \frac{x - y/10}{y + x/10}$$

Si ce nombre est positif, on pose $x' = x - y/10$ et $y' = y + x/10$, et on se ramène au calcul de $\arctan(x'/y')$. Si ce nombre est négatif, cela signifie que le développement décimal de $\arctan(x/y)$ a comme première décimale 0. On calcule ainsi la première décimale du développement de $\arctan(a)$ en effectuant uniquement des additions, des soustractions et des décalages (une division par 10 étant un décalage). On continue alors de la même manière pour les décimales suivantes qui sont dans la table des arctangentes. On termine en effectuant la méthode de Newton.

5 Suggestions de développement

- Complexité des opérations de base sur les flottants multiprécision, cas de la racine carrée, de l'exponentielle en utilisant le développement en séries.
- Approximation polynomiale par intervalles avec recollement C^1, C^2, etc.
- Calcul des fonctions usuelles (discussion sur les méthodes, leur efficacité...)
- Interpolation de Lagrange, calcul effectif (différences divisées), applications (par exemple pour des algorithmes efficaces : polynome caractéristique, PGCD à plusieurs variables,...)
- Calcul d'approximants de Padé ou de Cauchy.
- Comparaison entre différentes méthodes d'approximation pour une même fonction (par exemple le logarithme nepérien, la fonction sinus), dans l'optique du calcul en simple et multi-précision.