REVUE FRANÇAISE D'AUTOMATIQUE INFORMATIQUE
RECHERCHE OPÉRATIONNELLE. MATHÉMATIQUE

FRANÇOISE VEILLON

Brève communication. Une nouvelle méthode de calcul de la transformée inverse d'une fonction au sens de Laplace et de la déconvolution de deux fonctions

Revue française d'automatique informatique recherche opérationnelle. Mathématique, tome 6, n° R2 (1972), p. 91-98

<http://www.numdam.org/item?id=M2AN_1972__6_2_91_0>

R.A.I.R.O.
(6e année, juillet 1972, R-2, p. 91-98)

UNE NOUVELLE METHODE DE CALCUL DE LA TRANSFORMEE INVERSE D'UNE FONCTION AU SENS DE LAPLACE ET DE LA DECONVOLUTION DE DEUX FONCTIONS.

par Françoise VEILLON (1)

Résumé. — *Présentation d'une amélioration de la méthode de Dubner et Abate publiée dans le journal de l'ACM (janvier 1968, vol. 15, n° 1) et d'une application au problème de déconvolution numérique de deux fonctions.*

Méthode [DA]

Soit une fonction à valeurs réelles $f(t)$, définie sur $]-\infty, +\infty[$, nulle sur $]-\infty, 0[$, que l'on veut calculer connaissant sa transformée de Laplace $F(p)$. On suppose que $F(p)$ n'a pas de pôles à droite de l'origine (un changement de variable sur p peut toujours nous ramener à ce cas).

On définit la suite de fonctions $g_n(t)$ telle que $g_n(t)$ soit définie sur $]-\infty, +\infty[$, périodique, de période $2T$, coïncide avec $e^{-at} f(t)$ sur le segment $[nT,(n+1)T]$ et soit symétrique par rapport à la droite des ordonnées passant par le point d'abscisse nT.

La définition de la suite $g_n(t)$ permet d'écrire :

$$\sum_{n=0}^{\infty} e^{at} g_n(t) = \sum_{n=0}^{\infty} e^{at}\, e^{-a(2nT+t)} f(2nT+t) + \sum_{n=1}^{\infty} e^{at}\, e^{-a(2nT-t)} f(2nT-t)$$

En isolant le terme correspondant à $n = 0$, il vient :

$$\sum_{n=0}^{\infty} e^{at} g_n(t) = f(t) + \sum_{n=1}^{\infty} e^{-2aTn}[f(2nT+t) + e^{2at} f(2nT-t)]$$

(1) Mathématiques Appliquées, Informatique, Université de Grenoble.

En identifiant cette expression avec son développement en série de Fourier :

$$\sum_{n=0}^{\infty} e^{at} g_n(t) = \frac{2\,e^{at}}{T}\left[\frac{1}{2}\,ReF(a) + \sum_{k=1}^{\infty} ReF\left(a + \frac{ik\pi}{T}\right)\cos\frac{k\pi t}{T}\right]$$

il vient :

$$f(t) = \frac{2\,e^{at}}{T}\left[\frac{1}{2}\,ReF(a) + \sum_{k=1}^{\infty} ReF\left(a + \frac{ik\pi}{T}\right)\cos\frac{k\pi t}{T}\right] + Err$$

$$Err = \sum_{n=1}^{\infty} e^{-2aTn}\left[f(2nT + t) + e^{2at} f(2nT - t)\right]$$

Si t tend vers T, Err est de l'ordre de $f(T)$, ce qui n'est pas acceptable. Dubner et Abate ont montré que si t_{max} est la valeur maximum pour laquelle on veut calculer $f(t)$, et si $t_{max} = \frac{T}{2}$, Err peut être rendu aussi petit que possible en jouant sur a. En effet, puisque $F(p)$ n'a pas de singularités à droite de l'origine alors $f(t)$ est bornée à l'infini par une fonction de la forme Ct^m (C constante et m entier). Dans ce cas Err est de l'ordre de :

$$C(1{,}5T)^m\,e^{-aT} \text{ ou } C\,e^{-at} \text{ si } m = 0$$

expressions qui sont décroissantes en fonction de a.

Si $m = 0$ et $aT = 10$, Err $\simeq C \times 10^{-5}$.

Dubner et Abate préconisent d'utiliser leur méthode avec aT constant valant de 8 à 15, et de prendre 500 à 1 000 termes pour calculer la somme selon les cas.

Raffinements de calcul

1° Calcul de la somme infinie

Pour mieux calculer la somme infinie, nous avons utilisé l'ε-algorithme qui est une méthode d'accélération de la convergence de la série associée à la somme [*BR*].

2° Rapport entre T et t

Il n'est pas justifié de lier T à t_{max}. Cela signifie que pour calculer $f(t)$ pour $t = 1$ par exemple, T sera différent selon que l'on tabule $f(t)$ jusqu'à $t = 10$ ou $t = 20$. Il nous a paru préférable de lier T à t de la façon suivante :

$$\frac{T}{t} = \lambda_1 \qquad \text{et} \qquad \lambda_1 \geqslant 2$$

Cette convention a été confirmée par l'étude de la fonction $f(t) = 1$. En effet il vient en isolant les termes négligés dans la sommation

$$1 = f(t) \sim \frac{2\,e^{at}}{T}\left[\frac{1}{2a} + \sum_{k=1}^{N} \frac{1}{1 + \dfrac{\pi^2 k^2}{T^2 a^2}} \cos\left(\frac{k\pi}{\lambda_1}\right)\right] + \frac{2\,e^{at}}{aT} \sum_{k=N+1}^{\infty} \frac{1}{1 + \dfrac{k^2\pi^2}{T^2 a^2}} \cos\left(\frac{k\pi}{\lambda_1}\right)$$

En prenant N assez grand, la deuxième somme peut être rendue petite et le terme négligé dans son ensemble peut rester petit en jouant sur a et λ_1. De plus, *le terme négligé est constant.*

3° Optimisation de λ_1

Avec une même valeur de aT et un même N, des essais sur diverses valeurs de λ_1 et diverses fonctions ont montré que $\lambda_1 = 8$ était la meilleure solution.

4° Influence du nombre N auquel on arrête la sommation

Posons $U_k = ReF\left(a + \dfrac{ik\pi}{T}\right) \cos\left(\dfrac{k\pi}{8}\right)$.

Si U_k et U_l sont égaux modulo 16, alors ils font intervenir le même cosinus.

Des essais où seul N variait ont montré que $N = 36$ modulo (16) (qui correspond à un cosinus nul) donnait de meilleurs résultats que $N = 30$ modulo (16) $\left(\cos\dfrac{k\pi}{8} = \dfrac{\sqrt{2}}{2}\right)$.

$N = 32$ modulo (16) est encore plus défavorable (cosinus égal à 1).

5° Blocage des termes de la somme

Nous avons préféré bloquer 8 par 8 ($m \times \lambda_1$ avec $m \geqslant 1$ en général) les termes de la somme.

Soit $V_k = \sum_{l=1}^{k} ReF\left(a + \dfrac{il\pi}{T}\right)$.

Au lieu d'appliquer l'ε-algorithme sur N termes $V_1 \ldots V_N$, nous l'avons appliqué sur M termes :

$$W_1 = \sum_{m=1}^{8} V_m \ldots W_M = \sum_{m=8(M-1)+1}^{8M} V_m$$

Si $M = N$, nous avons une plus grande précision pour un même volume de calculs car nous avons utilisé 8 fois plus de termes de la suite d'origine.

Si $M = N/8$, nous avons un même résultat pour environ 32 fois moins de calculs (il y a $\frac{N(N+1)}{2}$ termes à calculer dans l'ε-algorithme activé sur N termes). Le fait que W_k fasse intervenir chacun des 8 cosinus différents concernés a pour effet de faire de la suite W_k une suite plus « stable » que la suite V_k.

6° Optimisation du paramètre *a*

La valeur approchée de $f(t)$ est théoriquement indépendante de a. Elle en est en fait dépendante à cause de :

a) l'erreur de principe que l'on néglige,

b) le calcul numérique, donc approché, de $f(t)$.

Nous pouvons donc l'écrire $f(t, a)$ et le problème est de chercher a tel que $f(t, a)$ dépende le moins de a.

Pour t fixé, nous calculons $f(t, a_i)$ pour un ensemble de valeurs a_i de a. La valeur a_{opt} retenue est soit la plus petite (car elle minimise Err) des valeurs de a pour laquelle la dérivée de la spline fonction d'ordre deux passant par les points $(f(t, a_i), a_i)$ s'annule ou à défaut la valeur de a où cette dérivée est la plus petite en valeur absolue.

Résultats

1° Tous les exemples ont été calculés avec $M = 8$.

2° De nombreux essais ont permis de mettre au point l'ensemble de a suivant :

$$a_1 = 1{,}15 \quad a_2 = 1{,}20 \quad a_3 = 1{,}25 \quad a_4 = 1{,}30 \quad a_5 = 1{,}35$$

et tous les résultats sont calculés avec cet ensemble mais le nombre et la valeur des a_i sont des paramètres de la procédure de calcul.

3° Le temps de calcul est de l'ordre de 1 seconde par valeur de t.

Application à la déconvolution

Étant donné l'équation de convolution :

$$s(t) = \int_0^t \mathcal{C}(t - \tau)\, e(\tau)\, d\tau$$

qui s'écrit après transformation de Laplace

$$S(p) = T(p) \times E(p).$$

où S, T, E sont les transformées respectives de s, $\mathcal{C}$ et e, on veut calculer $\mathcal{C}$ connaissant s et e.

Fonction	Intervalle	Erreur relative méthode de Papoulis [PA]		Erreur méthc Gaver Ste
		min	max	min
$\frac{\pi}{4} e^{-0,2t} \sin t$	$1.906\ 10^{-2} \leqslant t \leqslant 1.220\ 10^{-1}$	$6\ 10^{-5}$	1,7	$0,25\ 10^{-4}$

Fonction	Intervalle	Erreur relative méthode de régularisation [RI]		Erreur méthc Gaver S
		min	max	min
$f(t) = \begin{cases} 1 + \sin(\pi t) \\ \text{si } 0 \leqslant t \leqslant 2 \\ 0 \text{ si } t > 2 \end{cases}$	$0 \leqslant t \leqslant 2$	$7,5\ 10^{-3}$	1,8	$1,2\ 10^{-3}$

Fonction	Intervalle	Erreur relative méthode de Bellmann [BKL]		Erreur méthc Gaver S
		min	max	min
$f(t) = e^{-t/2}$	$4,14 \leqslant t \leqslant 0,016$	$0,7\ 10^{-4}$	$4,5\ 10^{-2}$	$< 3\ 10^{-6}$

Fonction	Intervalle	Erreur relative pour la méthode de Gaver Stehfest		Erreu pour la Dubner Al
		min	max	min
$\frac{1}{\sqrt{\pi t}}$	$0 \leqslant t \leqslant 100$	$1{,}24\ 10^{-6}$	$1{,}24\ 10^{-6}$	$1{,}12\ 10^{-6}$
$-C - \mathrm{Ln}(t)$ $C = 0{,}5772157$	$0 \leqslant t \leqslant 100$	$4\ 10^{-6}$	$4\ 10^{-5}$	$1{,}5\ 10^{-8}$
$\frac{t^3}{6}$	$0 \leqslant t \leqslant 100$	$6\ 10^{-3}$	$6\ 10^{-3}$	$2{,}3\ 10^{-5}$
e^{-t}	$0 \leqslant t \leqslant 10$	$2{,}5\ 10^{-4}$	> 1	$5\ 10^{-6}$
$\frac{L_3(t)}{6}$ $L_3 = -t^3 + 9t^2 - 18t + 6$	$0 \leqslant t \leqslant 100$	$1{,}2\ 10^{-3}$	$2{,}1\ 10^{-1}$	$1{,}1\ 10^{-6}$
$e^{-t}\sqrt{t} + t(1 - e^{-t})$	$0 \leqslant t \leqslant 100$	$1{,}2\ 10^{-5}$	$5\ 10^{-4}$	$7\ 10^{-9}$
1	$0 \leqslant t \leqslant 100$	$5\ 10^{-7}$	$5\ 10^{-7}$	$1{,}5\ 10^{-11}$

Sachant calculer $\mathcal{G}(t)$ à partir de $Re[T(p)]$, il suffit donc de calculer $Re[T(p)]$.

Or, cette fonction est égale à

$$Re\left[\frac{S_r(p) + iS_i(p)}{E_r(p) + iE_i(p)}\right]$$

avec $S(p) = S_r(p) + iS_i(p)$

$E(p) = E_r(p) + iE_i(p)$

Or S_r, S_i, E_r, E_i sont des transformées de Fourier de s et e que l'on calcule par le procédé de Cooley et Tukey [*CT*].

Résultats

Fonction	Intervalle	Erreur relative	
		Min	Max
$t = \int_0^t dt$	$0{,}25 < t < 3$	10^{-8}	10^{-8}
$\begin{cases} t\,e^{-\alpha t} = \int_0^t e^{\alpha(t-\tau)}\,e^{\tau\alpha}\,d\tau \\ \alpha = 0{,}5 \end{cases}$	«	$1.4\ 10^{-8}$	$1.3\ 10^{-7}$
$\frac{t^3}{6} + \frac{t^2}{2} = \int_0^t \tau(1 + t - \tau)\,d\tau$	«	$1.07\ 10^{-5}$	$1.7\ 10^{-4}$
$\frac{\cos t - \cos 3t}{8} = \int_0^t \sin(\tau) \cos 3(t - \tau)\,d\tau$	«	$1.96\ 10^{-4}$	$7.35\ 10^{-2}$

N.B. Une procédure d'inversion et une procédure de déconvolution en algol se trouvent à la référence [*VE*].

[BKL] R. Bellmann, R. Kalaba et J. Lockett, *Numerical inversion of the Laplace Transform*, Elsevier, New York-Londres, 1966.

[BR] C. Brezinski, *Méthodes d'accélération de la convergence en analyse numérique.* Thèse présentée à la Faculté des Sciences de Grenoble le 26 avril 1971.

[CT] Cooley et Tukey, « An algorithm for machine calculation of complex Fourier series », *Math. Comput.* 19 (1965), pages 297-301.

[DA] H. DUBNER et J. ABATE, « Numerical inversion of Laplace transform and the finite Fourier cosine transform », *Journal de l'ACM*, 1968, vol. 15, n° 1, pages 115-123.

[PA] A. PAPOULIS, « A new method of inversion of the Laplace transform », *Quaterly of applied mathematics*, vol. 14 (1956), pages 405-414.

[RI] G. RIBIÈRE, *Amélioration du résidu dans la résolution des systèmes linéaires au sens des moindres carrés*. Thèse de 3e cycle, Paris, 1966.

[ST] H. STEHFEST, Numerical inversion of Laplace transforms. *Communications de l'ACM*. Vol. 13, n° 1 (1970), pages 47-49.

[VE] F. VEILLON, *Quelques nouvelles méthodes pour le calcul numérique de la transformée inverse de Laplace*. Thèse présentée le 11 mars 1972 à l'Université Scientifique et Médicale de Grenoble.