UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES
Section de mathématiques

Analyse Numérique
Corrigé Série 4

17 octobre 2019

1. (**Algorithme de Clenshaw**) On note $T_k(x)$ le k-ème polynôme de Chebyshev.
   (a) Étant donné $p \in \mathbb{P}_n$, montrer que celui-ci satisfait $p(x) = \sum_{k=0}^{N} c_k T_k(x)$ pour $N \in \mathbb{N}$. Trouver le N minimal. Est-ce que le choix des coefficients c_k est unique pour ce N minimal ?
   Sol.: *Comme $p \in \mathbb{P}_n$, on a que $p \in \mathrm{Vect}\{1, x, x^2, \ldots, x^n\}$. D'autre part, comme $\{T_k\}_{k=0}^{\infty}$ est une suite de polynômes telle que T_k est exactement de degré k, le sous-espace vectoriel engendré $\mathbb{P}_n = \mathrm{Vect}\{T_0, T_1, T_2, \ldots, T_n\}$ contient tous les polynômes de degré n. Puisque les T_k sont orthogonaux, ils forment une base orthogonale de $\mathbb{P}_n$. Donc nous avons que $N = n$. Comme $\{T_0, T_1, T_2, \ldots, T_N\}$ est une base de $\mathbb{P}_n$, les coefficients c_k sont uniques dans cette base.*
   (b) Étant donné $x \in \mathbb{R}$, montrer que $p(x) = \sum_{k=0}^{N} c_k T_k(x)$ peut être évalué par le procédé suivant. On fixe $u_{n+1} = 0$ et $u_n = c_n$ et puis on calcule
$$u_k = 2x\, u_{k+1} - u_{k+2} + c_k, \quad k = n-1, n-2, \ldots, 0.$$
   Alors
$$p(x) = \frac{1}{2}(c_0 - u_2 + u_0).$$
   Sol.: *On rappelle la formule de récurrence sur les polynômes de Chebyshev pour $N \geq 1$,*
$$T_{N+1}(x) = 2x\, T_N(x) - T_{N-1}(x).$$
   On a :
$$\begin{aligned}
p(x) &= c_0T_0 + c_1T_1 + \cdots + c_{N-2}T_{N-2} + c_{N-1}T_{N-1} + c_NT_N \\
&= c_0T_0 + c_1T_1 + \cdots + c_{N-2}T_{N-2} + c_{N-1}T_{N-1} + u_N(2x\, T_{N-1} - T_{N-2}) \\
&= c_0T_0 + c_1T_1 + \cdots + c_{N-2}\boldsymbol{T_{N-2}} + c_{N-1}\boldsymbol{T_{N-1}} + u_N 2x\, \boldsymbol{T_{N-1}} - u_N\boldsymbol{T_{N-2}} \\
&= c_0T_0 + c_1T_1 + \cdots + (c_{N-2} - u_N)\, \boldsymbol{T_{N-2}} + (c_{N-1} + 2x\, u_N)\, \boldsymbol{T_{N-1}},
\end{aligned}$$
   on utilise $c_{N-1} = u_{N-1} - 2x\, u_N + u_{N+1}$ et $u_{N+1} = 0$, et on continue
$$\begin{aligned}
p(x) &= c_0T_0 + c_1T_1 + \cdots + (c_{N-2} - u_N)\, T_{N-2} + u_{N-1}T_{N-1} \\
&= c_0T_0 + c_1T_1 + \cdots + (c_{N-2} - u_N)\, T_{N-2} + u_{N-1}(2x\, T_{N-2} - T_{N-3}) \\
&\;\vdots \\
&= (c_0 - u_2)T_0 + u_1T_1 \\
&= c_0 - u_2 + x\, u_1 \\
&= \frac{1}{2}(c_0 - u_2 + u_0),
\end{aligned}$$
   où on a utilisé $T_1 = x$ et $u_0 = 2x\, u_1 - u_2 + c_0$ dans les dernières deux étapes.

2. ($\star$, tout l'exercice) (**Précision double IEEE**) Dans cet exercice, on considère la précision double comme définie selon le standard IEEE.
   (a) Représenter en format décimal les nombres suivants

	Signe	Exposant	Mantisse
i.	0	11111111111	00000000 . . . 0
ii.	1	11111111111	00100010 . . . 0
iii.	0	00000000000	01011011 . . . 0
iv.	0	10000110010	00101011 . . . 0
v.	1	10000010110	10011101 . . . 0
vi.	0	00000000000	00000000 . . . 0

en spécifiant s'ils sont normalisés, dénormalisés, ***NaN*** (not a number) ou infinis. Justifier vos réponses.

Sol.:

i. *Le signe est positif. L'exposant est* 11111111111 *et la mantisse est nulle. La réponse est donc* ***+Inf****.*

ii. *L'exposant est* 11111111111 *et la mantisse est non nulle. Il s'agit donc de* ***NaN****.*

iii. *Le signe est positif. L'exposant est* 00000000000 *et la mantisse non nulle. On a donc un nombre dénormalisé. L'exposant est donc* $00000000000 = -1022$*. On obtient* $0.01011011\ldots 0 \times 2^{-1022} = (2^{-2} + 2^{-4} + 2^{-5} + 2^{-7} + 2^{-8}) \times 2^{-1022} = 2^{-1024} + 2^{-1026} + 2^{-1027} + 2^{-1029} + 2^{-1030} \approx 7.9094 \times 10^{-309}$ *.*

iv. *Le signe est positif.* $10000110010 = 2^{10}+2^5+2^4+2^1 = 1074$*. L'exposant est donc* $1074-1023 = 51$*. On obtient* $1.00101011\ldots 0 \times 2^{51} = (2^0 + 2^{-3} + 2^{-5} + 2^{-7} + 2^{-8}) \times 2^{51} = 2^{51} + 2^{48} + 2^{46} + 2^{44} + 2^{43} \approx 2.6300 \times 10^{15}$*.*

v. *Le signe est négatif.* $10000010110 = 2^{10}+2^4+2^2+2^1 = 1046$*. L'exposant est donc* $1046-1023 = 23$*. On obtient* $-1.10011101\ldots 0 \times 2^{23} = -(2^0 + 2^{-1} + 2^{-4} + 2^{-5} + 2^{-6} + 2^{-8}) \times 2^{23} = -(2^{23} + 2^{22} + 2^{19} + 2^{18} + 2^{17} + 2^{15}) = -13\,533\,184$*.*

vi. *Le signe est positif. L'exposant et la mantisse sont tous deux nuls. On obtient donc* $+0$*.*

(b) Quelle est la valeur de la précision de la machine $\varepsilon_{\text{mach}}$?

Sol.: *Le nombre de bits de la mantisse est* $p = 52$*, et donc la précision de la machine est donnée par* $2^{-52} \approx 2.2204 \times 10^{-16}$*.*