

Contrôle des connaissances session 2

Exercice 1. Approximation de l'exponentielle

On interpole la fonction exponentielle sur $[-1, 1]$ en $n + 1$ points x_0, \dots, x_n de $[-1, 1]$.

1. Donner une majoration de l'erreur d'interpolation en x fonction de x, n et de x_0, \dots, x_n .
2. On choisit x_0, \dots, x_n en les racines du polynôme de Tchebyshev $T_{n+1}(\cos(x)) = \cos((n+1)x)$. Déterminer n_0 pour assurer une erreur absolue inférieure à $1\text{e-}10$ sur $[-1, 1]$. On ne demande pas la plus petite valeur de n_0 qui convient. Même question pour avoir une erreur relative inférieure à $1\text{e-}10$.
3. Déterminer n en fonction du n_0 de la question précédente pour avoir un même majorant de l'erreur avec des points x_0, \dots, x_n équidistribués sur $[-1, 1]$.
4. Quelle valeur de n faut-il choisir pour obtenir la même précision en approchant l'exponentielle par son développement de Taylor à l'ordre n en $x = 0$?

1. La dérivée $n + 1$ -ième de l'exponentielle est elle-même, elle est majorée par e sur $[-1, 1]$ donc :

$$|e^x - P_n(x)| \leq \frac{e}{(n+1)!} \prod_{j=0}^n (x - x_j)$$

2. Comme $T_{n+1}(x) = 2^n \prod_{j=0}^n (x - x_j)$, on a

$$|e^x - P_n(x)| \leq \frac{e}{(n+1)!2^n}$$

qui est plus petit que $1\text{e-}10$ à partir de $n = 10$. En erreur relative, il faut diviser par e^x , donc il suffit de majorer $e^2/(n+1)!/2^n$, $n = 11$ convient.

3. On note h le pas entre deux x_j successifs ($h = 1/5$ pour $n = 10$). On majore $\prod_{j=0}^n (x - x_j)$ par $h^{n+1} \frac{1}{2} n!$ car on est à distance au plus $h/2$ du plus proche, puis h du deuxième plus proche, puis $2h$ du suivant, etc. Avec le n de la question précédente l'erreur d'interpolation est majorée par $h^{n+1} \frac{1}{2(n+1)}$, ici

$$e^2/5^{11}/2/11 = 6.9\text{e-}9$$

majoration un peu moins bonne que la précédente mais nécessitant moins de calculs à la machine.

4. Il existe θ entre 0 et x tel que :

$$f(x) = \text{Taylor}_n(x) + e^\theta \frac{x^{n+1}}{(n+1)!}.$$

Le produit des $x - x_k$ est remplacé par x^{n+1} , il suffira donc que $e/(n+1)!$ soit plus petit que $1\text{e-}10$, ce qui se produit dès $n = 13$.

Exercice 2. Approximation au sens des moindres carrés

On cherche un polynôme du second degré $P(t) = \alpha t^2 + \beta t + \gamma$ approchant le mieux possible au sens des moindres carrés l'exponentielle aux points d'abscisses $t_1 = -1$, $t_2 = -1/2$, $t_3 = 0$, $t_4 = 1/2$ et $t_5 = 1$.

1. On pose $x = (\gamma, \beta, \alpha)$ et b le vecteur de composantes $(e^{t_1}, \dots, e^{t_5})$ le problème revient alors à minimiser $\|Ax - b\|_2$. Déterminer la matrice A en fonction des t_i .
2. Comment peut on résoudre ce problème en utilisant des outils d'algèbre linéaire?

1. `A:=tran(tran(vandermonde([-1,-1/2,0,1/2,1]))[0..2])`
2. On résoud le système normal associé $A^T A = A^T b$. On obtient un résultat plus précis avec la décomposition QR de A . On trouve `[0.994415410173,1.14859907711,0.547734459671]`.

Exercice 3. Transformée de Fourier rapide

Pour tout entier N , on appelle *transformée de Fourier discrète* l'application de \mathbb{C}^N dans \mathbb{C}^N qui envoie le N -uplet (x_0, \dots, x_{N-1}) sur le N -uplet (X_0, \dots, X_{N-1}) défini par

$$\forall j \in \{0, \dots, N-1\}, X_j = \sum_{k=0}^{N-1} x_k e^{-\frac{2i\pi}{N}jk}.$$

De façon équivalente, on a

$$\begin{pmatrix} X_0 \\ X_1 \\ X_2 \\ \vdots \\ X_{N-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \dots & \omega^{(N-1)(N-1)} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{N-1} \end{pmatrix}$$

où $\omega = \omega_N = e^{-\frac{2i\pi}{N}}$. On notera $M_N = (\omega^{(j-1)(k-1)})_{1 \leq j, k \leq N}$ la matrice ci-dessus.

1. Combien d'opérations (additions et multiplications de nombres complexes) faut-il réaliser pour calculer une transformée de Fourier discrète avec les formules ci-dessus ?
2. Montrer que $M_N^* M_N = N I_N$ où M_N^* est la transposée de la conjuguée de M_N . En déduire la norme et le conditionnement de M_N pour la norme subordonnée à la norme hermitienne de \mathbb{C}^N (on rappelle que pour toute matrice $A \in \mathcal{M}_N(\mathbb{C})$ et pour tout $v \in \mathbb{C}^N$, $\|Av\| = \sqrt{\langle v | A^* A v \rangle}$).
3. On suppose que le N -uplet $x = (x_0, \dots, x_{N-1})$ est connue avec une imprécision Δx . On note ΔX l'imprécision sur sa transformée $X = (X_0, \dots, X_{N-1})$. Montrer que (pour la norme hermitienne) $\|\Delta X\| = \sqrt{N} \|\Delta x\|$ et $\frac{\|\Delta X\|}{\|X\|} = \frac{\|\Delta x\|}{\|x\|}$.

Pour accélérer les calculs, particulièrement dans le cas $N = 2^n$, l'algorithme de Cooley-Tukey utilise une stratégie de type « diviser pour régner ». On commence par calculer (récursivement) deux

transformées de Fourier discrètes de taille moitié : le $N/2$ -uplet $(A_0, \dots, A_{N/2-1})$, transformée des termes d'indice pair $(x_0, x_2, \dots, x_{N-2})$, et le $N/2$ -uplet $(B_0, \dots, B_{N/2-1})$, transformée des termes d'indice impair $(x_1, x_3, \dots, x_{N-1})$.

4. Montrer que pour tout $j \in \{0, \dots, N-1\}$,

$$X_j = \sum_{k=0}^{N/2-1} x_{2k} e^{-\frac{4i\pi}{N}jk} + e^{-\frac{2i\pi}{N}j} \sum_{k=0}^{N/2-1} x_{2k+1} e^{-\frac{4i\pi}{N}jk}.$$

5. En déduire que pour tout $j \in \{0, \dots, N/2-1\}$, $X_j = A_j + e^{-\frac{2i\pi}{N}j} B_j$ et $X_{j+N/2} = A_j - e^{-\frac{2i\pi}{N}j} B_j$ (on fera attention que A_j et B_j ne sont a priori définis que pour $0 \leq j \leq N/2-1$).

6. En déduire une méthode de calcul de la transformée de Fourier discrète utilisant seulement $O(n2^n) = O(N \log(N))$ additions et multiplications de nombres complexes.

Question bonus : Une multiplication de deux nombres complexes stockés sous forme algébrique revient à faire 4 multiplications réelles, à moins que l'un des facteurs ne soit réel ou imaginaire pur. Pour simplifier les calculs, Rader et Brenner ont proposé de remplacer $(B_0, \dots, B_{N/2-1})$ par $(C_0, \dots, C_{N/2-1})$, transformée de Fourier discrète de $(x_1 - x_{N-1} - Q, x_3 - x_1 - Q, \dots, x_{N-1} - x_{N-3} - Q)$ avec $Q = \frac{2}{N} \sum_{k=0}^{N/2-1} x_{2k+1}$. On obtient alors les formules suivantes, que l'on ne demande pas de justifier :

$$X_0 = A_0 + C_0, \quad X_{N/2} = A_0 - C_0, \quad \text{et pour tout } j \neq 0, N/2,$$

$$X_j = A_j + \frac{C_j}{2i \sin(\frac{2\pi}{N}j)}$$

7. Expliquer pourquoi cette méthode n'est pas satisfaisante en terme de stabilité numérique.

1. Si on suppose que M_N a été calculée auparavant, il faut N multiplications et $N-1$ additions par composante de X , donc $O(N^2)$ opérations.

2.

$$(M_N)_{j,k} = \omega^{(j-1)(k-1)}, \quad (M_N^*)_{j,k} = \overline{\omega}^{(j-1)(k-1)} = \omega^{-(j-1)(k-1)}$$

donc

$$\begin{aligned} (M_N^* M_N)_{j,l} &= \sum_{k=0}^{N-1} (M_N^*)_{j,k} (M_N)_{k,l} \\ &= \sum_{k=0}^{N-1} \omega^{-(j-1)(k-1)} \omega^{(k-1)(l-1)} \\ &= \sum_{k=0}^{N-1} \omega^{(l-j)(k-1)} \\ &= \sum_{k=0}^{N-1} \left(\omega^{(l-j)} \right)^{k-1} \\ &= \frac{1 - \omega^{(l-j)N}}{1 - \omega^{(l-j)}} \quad (\text{pour } j \neq l) \\ &= 0 \text{ si } j \neq l \end{aligned}$$

Si $j = l$, on obtient N Donc $\|M_N v\|^2 = N\|v\|^2$, la norme de M_N est \sqrt{N} et le conditionnement est 1.

3. C'est une conséquence immédiate de $\|M_N v\|^2 = N\|v\|^2$ et de $\Delta X = M_N \Delta x$.
4. On remplace ω par sa valeur, on décompose la somme en indices pairs et impairs et on factorise un ω dans la somme d'indices impairs.
5. Pour faire apparaitre les A_j et B_j , lorsque $j \in [0, N/2 - 1]$, on utilise que ω^2 est une racine primitive 2^{n-1} -ième de 1. Pour les indices j supérieurs, remarquer que $e^{-\frac{2i\pi}{N}(j+N/2)} = -e^{-\frac{2i\pi}{N}j}$ et $e^{-\frac{4i\pi}{N}(j+N/2)k} = e^{-\frac{4i\pi}{N}jk}$.
6. Si on utilise le procédé, le temps de calcul $T(n)$ pour $N = 2^n$ est égal à celui des A_N et B_N et de leur combinaison soit 2 fois

le temps de calcul pour $N = 2^{n-1}$ plus un $O(N)$

$$\begin{aligned} T(n) &= 2T(n-1) + C2^n = 2(2T(n-2) + C2^{n-1}) + C2^n = 2^2T(n-2) + 2C2^n = \\ &= 2^2(2T(n-3) + C2^{n-2}) + 2C2^n = 2^3T(n-3) + 3C2^n = \dots = 2^nT(0) + nC2^n \\ &\text{ce qui est bien un } O(n2^n) \end{aligned}$$