Agrégation marocaine de Mathématiques, session 2025

Épreuve de modélisation et Calcul Scientifique

Résolution de numérique de systèmes linéaires par chaînes de M

- *Il est rappelé que le jury n'exige pas une compréhension exhaustive du texte. Vous êtes laissé(e) libre d'organiser votre discussion comme vous l'entendez. Des suggestions de développement, largement indépendantes les unes des autres, vous sont proposées en fin de texte. Vous n'êtes pas tenu(e) de les suivre. Il vous est conseillé de mettre en lumière vos connaissances à partir du fil conducteur constitué par le texte. Le jury appréciera que la discussion soit accompagnée d'exemples traités sur ordinateur.*

I. Résolution d'un système d'équations linéaires par chaînes de Markov

Considérons un système d'équations linéaires $Ax = b$ pour lequel on souhaite calculer une estimation de la solution par une méthode aléatoire utilisant les chaînes de Markov. On suppose que le système est complexe au point que les méthodes déterministes sont très couteuses en espace et en temps. Ici on parle de la complexité dont les sources peuvent êtres la grande dimension de l'espace de travail et/ou la raideur de la matrice (i.e. des coefficients trop faibles par rapport à d'autres).

Dans le système $Ax = b$, $A = (A_{ij})_{1\leq i,j\leq r}$ est une matrice réelle et carrée d'ordre r, $x = (x_1, \dots, x_r)$ et $b = (b_1, \dots, b_r)$ sont des vecteurs de $I\!R^r$. Pour simplifier on adopte la notation horizontale (ligne) aussi bien pour les vecteurs que pour leurs tranposées (le contexte fera la différence entre ligne et colonne).

En réecrit le système sous la forme $x = Bx + b$ et on suppose que

$$\max_{1\leq i\leq r} \sum_{j=1}^{r} |B_{ij}| < 1 \qquad avec \qquad B = (B_{ij})_{1\leq i,j\leq r}.$$

Maintenant à partir de cette dernière écriture, on propose l'équation récurssive suivante

$$x^{(k+1)} = Bx^{(k)} + b, \quad k = 0, 1, \dots,$$

avec $x^{(0)} = 0$ et $B^0 = I$ (identité de $I\!R^r$).

Nous avons

$$\lim_{k\longrightarrow+\infty} x^{(k)} = x \quad (x \text{ solution du système } Ax = b)$$

La $j^{ième}$ composante de $x^{(k+1)}$ est donnée par

$$x_j^{(k+1)} = b_j + \sum_{1\leq i_1\leq r} B_{ji_1}b_{i_1} + \sum_{1\leq i_1,i_2\leq r} B_{ji_1}B_{i_1i_2}b_{i_2} + \cdots + \sum_{1\leq i_1,i_2,\dots,i_k\leq r} B_{ji_1}B_{i_1i_2}\cdots B_{i_{k-1}i_i}b_{i_k} \tag{1}$$

Considérons maintenant une chaîne de Markov arbitraire $X = \{X_n; n \in I\!N\}$, d'espace d'état $E = \{1, \dots, r\}$ et de matrice de transition $K = (K_{ij})_{i,j\in E}$ telle que

$$(H1) \qquad K_{ij} > 0 \quad \text{si} \quad B_{ij} > 0,$$

et de loi initiale $\mu = (\mu_i)_{i\in E}$.
Considérons le problème d'estimation du produit scalaire $\langle h, x^{(k+1)}\rangle$ avec $h \in I\!R^r$ un vecteur de paramètres. Suppons maintenant que l'on dispose d'une trajectoire $X_0, X_1, \ldots, X_k$ de la chaîne X et définissons la variable aléatoire W_m par

$$W_m = \frac{B_{X_0X_1}B_{X_1X_2}\cdots B_{X_{m-1}X_m}}{K_{X_0X_1}K_{X_1X_2}\cdots K_{X_{m-1}X_m}},$$

qui peut s'écrire récurssivement

$$W_m = W_{m-1}\frac{B_{X_{m-1}X_m}}{K_{X_{m-1}X_m}}, \qquad W_0 = 1.$$

On définit aussi la variable aléatoire

$$Y_k(h) = \frac{h_{X_0}}{\mu_{X_0}}\sum_{m=0}^{k} W_m b_{i_m}$$

Le résultat principal de cette modélisation est donné par la proposition suivante :

Proposition

$$I\!E(Y_k(h)) = \langle h, \sum_{m=1}^{k} B^m b\rangle = \langle h, x^{(k+1)}\rangle \tag{2}$$

Cela signigie que $Y_k(h)$ est un estimateur (appelé estimateur sans biais) de $\langle h, x^{(k+1)}\rangle$. Une réalisation de la variable aléatoire $Y_k(h)$ constitue donc une estimation de notre produit scalaire $\langle h, x^{(k+1)}\rangle$ et en chosissant $h = (h_1, \ldots, h_r)$ tel que $h_i = \delta_j(i)$ (symbole de kroneker), on déduit une estimation de $x_j^{(k+1)}$.
Remarque : Nous aurons besoin d'ajouter une hypothèse sur la loi initiale μ qui est :

$$(H2) \qquad\qquad \mu_i > 0 \quad \text{si} \quad h_i \neq 0.$$

Algorithme d'estimation de $x_j^{(k+1)}$

Etape 0. Choix de la composante j et d'un entier k reltivement grand
Choix de $\mu_1, \ldots, \mu_r$ et de $K = (K_{ij})_{1\leq i,j\leq r}$ satisfaisant $(H1)$ et $(H2)$.

Etape 1. Génération de N trajectoires, de la chaîne X, indépendantes et de même longueur $k+1$: $i_0^{(s)}, i_1^{(s)}, \ldots, i_k^{(s)},\ s = 1, 2, \ldots, N$

Etape 2. Calcul des N réalisations de la variable $Y_k(h)$:

$$y_k^{(s)}(h) = \frac{h_{i_0^{(s)}}}{\mu_{i_0}^{(s)}}\sum_{m=0}^{k} W_m^{(s)} b_{i_m^{(s)}},\ s = 1, 2, \ldots, N$$

où

$$W_m^{(s)} = \frac{B_{i_0^{(s)}i_1^{(s)}}B_{i_1^{(s)}i_2^{(s)}}\cdots B_{i_{m-1}^{(s)}i_m^{(s)}}}{K_{i_0^{(s)}i_1^{(s)}}K_{i_1^{(s)}i_2^{(s)}}\cdots K_{i_{m-1}^{(s)}i_m^{(s)}}}$$

Etape 3. Calcul d'une estimation de $I\!E(Y_k(h)) = \langle h, x^{(k+1)}\rangle$:

$$\frac{1}{N}\sum_{s=1}^{N} y_k^{(s)}(h)$$

N.B : La section suivante sur les chaînes de Markov est présentée uniquement à titre de rappel et ne fera pas l'objet d'un développement mathématique par le candidat.

II- Un rappel sur les chaînes de Markov à temps discret
Soit $(\Omega, \mathcal{F}, I\!\!P)$ un espace probabilisé dans lequel sont définies toutes les variables aléatoires considérées dans la modélisation proposée.
Soit le processus $X = (X_n)_{n\geq 0}$ défini par une suite de variables aléatoires à valeurs dans un ensemble fini E appelé espace d'état, et soit $\mu = (\mu(i))_{i\in E}$ une loi de probabilités sur E. Pour $i \in E$, la notation $X_n = i$ signifie que le processus X est dans l'état i à l'instant n.

II.1 Définition 1. On dit que le processus $X = (X_n)_{n\geq 0}$ est une chaîne de Markov à temps discret ($n \in I\!\!N$), à espace d'état E et de loi initiale μ si et seulement si

$$\forall\, n \geq 0,\ \forall\, i, i_0, \ldots, i_{n-1}, j \in E,$$

$$I\!\!P(X_{n+1} = j / X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) \quad = \quad I\!\!P(X_{n+1} = j / X_n = i)$$

$$I\!\!P(X_0 = i) = \mu(i)$$

La probabilité $I\!\!P(X_{n+1} = j / X_n = i)$ est appelée probabilité de transition, de la chaîne X, de l'état i à l'état j entre les instants n et $n+1$.

II.2 Définition 2. La chaîne de Markov X est dite homogène dans le temps si et seulement si la probabilité $I\!\!P(X_{n+1} = j / X_n = j)$ est indépendante du temps n. Dans ce cas on la note $K(i,j) := I\!\!P(X_{n+1} = j / X_n = j)$, et la matrice ainsi obtenue $K = (K(i,j))_{i,j\in E}$ est appelée matrice de transition de la chaîne de Markov homogène X.

II.3 Remarque 1. Une chaîne de Markov à temps discret et homogène X est caractérisée par le triplet (E, μ, K), où E est son espace d'état, $\mu = (\mu(i))_{i\in E}$ est sa loi initiale, $\mu(i) = I\!\!P(X_0 = i)$, et $K = (K(i,j))_{i,j\in E}$ est sa matrice de transition.

II.4 Remarque 2. Pour tout état i fixé dans E, $\{K(i,j),\ \ j \in E\}$ est une loi de probabilité sur E. Les lignes de la matrice de transitions sont donc des lois de probabilité sur E.

II.5 Notations. On note $X \sim (E, \mu, K)$ pour désigner une chaîne de Markov $X = (X_n)_{n\geq 0}$ à temps discret, à espace d'état E, de loi initiale $\mu = (\mu(i))_{i\in E}$, et de matrice de transition $K = (K(i,j))_{i,j\in E}$.

II.6 Graphe de transition. Le graphe de transition d'une chaîne de Markov $X \sim (E, \mu, K)$ est un graphe orienté et évalué, dont les sommets sont les états (éléments de E) et les arcs sont des flèches dont les valeurs sont les coefficients de la matrice K (la valeur de l'arc qui se dirige de l'état i vers l'état j est $K(i,j)$). Les arcs de valeurs nulles ne sont pas dessinés.

II.7 Définition 3. On dit qu'une chaîne de Markov $X \sim (E,\mu,K)$ est irréductible si et seulement si

$$\forall\, i,j \in E, \;\; \exists m \in I\!N^* \quad K^{(m)}(i,j) > 0$$

où $K^{(m)}(i,j)$ est l'entrée (i,j) de la matrice K^m. Cela représente la probabilité de passer de l'état i à l'état j en m transitions.
Nous avons donc pour tout $i,j \in E$,

$$K^{(m)}(i,j) = \sum_{i_1,\ldots,i_{m-1}\in E} K(i,i_1)K(i_1,i_2)\cdots K(i_{m-1},j) \tag{3}$$

II.8 Proposition 1. Soit une chaîne de Markov homogène $X \sim (E,\mu,K)$.
Alors $\forall\, n \geq 0$, $\forall\, i_0,i_1,\ldots,i_n \in E$

$$I\!P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \mu(i_0)K(i_0,i_1)K(i_1,i_2)\cdots K(i_{n-1},i_n)$$

II.9 Proposition 2. Soit une chaîne de Markov homogène $X \sim (E,\mu,K)$. Alors

$$\forall\, m > 0, \quad I\!P(X_m = j) = \sum_{i\in E} \mu(i)K^{(m)}(i,j), \quad \forall j \in E$$

II.10 Proposition 3. Soit une chaîne de Markov homogène $X \sim (E,\mu,K)$. Alors

$$\forall\, n \geq 0,\; \forall\, m > 0,\; \forall\, i,j \in E, \quad I\!P(X_{n+m} = j / X_n = i) = I\!P(X_m = j/X_0 = i) = K^{(m)}(i,j)$$

II.11 Proposition 4. Soit une chaîne de Markov homogène $X \sim (E,\mu,K)$. Nous avons alors les équations de Chapman-Kolmogorov suivantes :

$$\forall m \geq 2 \;,\; \forall\, 0 < r < m \;,\; \forall\, i,j \in E, \quad K^{(m)}(i,j) = \sum_{k\in E} K^{(m-r)}(i,k)K^{(r)}(k,j)$$

II.12 Définition 4 (loi stationnaire ou invariante). Soit une chaîne de Markov homogène $X \sim (E,\mu,K)$. On dit qu'une loi de probabilité, $\pi = (\pi(i))_{i\in E}$ sur E, est une loi stationnaire de X (ou loi invariante par K) si et seulement si

$$\sum_{i\in E} \pi(i)K(i,j) = \pi(j), \quad \forall\, j \in E \tag{4}$$

Les équations (4) sont appelées **équations de balance ou d'équilibre**. Elles peuvent être réécrites matriciellement sous la forme

$$\pi^t K = \pi^t \qquad (\pi^t \text{ désigne la transposée du vecteur colonne } \pi)$$

II.13 Théorème 1. Si $X \sim (E, \mu, K)$ est une chaîne de Markov homogène et irréductible avec E fini, alors il existe une loi stationnaire $\pi = (\pi(i))_{i \in E}$ unique de X.

II.14 Génération d'une trajectoire d'une chaîne de Markov
Soit $X \sim (E, \mu, K)$ une chaîne de Markov homogène. Une procédure de génération d'une trajectoire $x = (x_0, x_1, \ldots, x_p)$, de longueur $p+1$, de la chaîne X est la suivante :

on commence par générer la réalisation x_0 selon la loi de probabilité initiale μ définie sur E, puis à partir de x_0, on génère de manière récurssive pour $k = 1, \ldots, p$, chaque réalisation x_k par la loi de probabilité $K(x_{k-1}, .)$ définie sur E.

II.14.1 Génération selon une loi de probabilité discrète.
Pour une variable aléatoire réelle discrète Y à valeurs dans $E = \{y_k; k = 1, 2, \ldots\}$ de loi de probabilités discrète $p_k = \mathbb{P}(Y = y_k),\ k = 1, 2, \ldots$, on génère une réalisation y de Y, en utilisant un nombre uniforme $u \in]0, 1[$, donnée par un générateur de nombres aléatoires (random), de la façon suivante :

$$\begin{cases} \text{Si } u \leq p_1 & \text{on décide} \quad y = y_1 \\ \text{Sinon on décide } y = y_k & \text{telle que} \quad \left(\displaystyle\sum_{i=1}^{k-1} p_i < u \leq \sum_{i=1}^{k} p_i \right) \end{cases}$$

III- Suggestions de développement

Ce paragraphe ne contient qu'un petit nombre de suggestions. Vous pouvez choisir d'étudier certains points seulement, de façon plus ou moins approfondie, et pas nécessairement dans l'ordre. Vous pouvez aussi vous poser d'autres questions que celles indiquées ci dessous. Il est vivement souhaité que vos investigations comportent une partie traitée sur ordinateur et, si possible, des représentations graphiques.

III.1 Aspect mathématique

1.) Retrouvez l'expression (1)
2.) Justifiez l'importance de chacune des hypothèses $(H1)$ et $(H2)$.
3.) Proposez une preuve pour la proposition donnant le résultat principal (2).

III.2 Aspect enseignement

1.) Expliquer des difficultés éventuelles que peut rencontrer cette modélisation.
2.) Expliquer les problèmes qui peuvent provenir de la raideur de la matrice (présence de coefficients très petits par rapport à d'autres qui sont très grands).
3.) Dans le cadre d'approximation de solutions de systèmes linéaires, expliquer quand est ce qu'une méthode aléatoire peut devenir compétitive par rapport à une méthode déterministe.

III.3 Aspect modélisation, calcul numérique et algorithmique

1.) Implémenter l'algorithme d'estimation de $x_j^{(k+1)}$
2.) Faites tourner l'algorithme sur un ou deux exemples et afficher les graphes d'évolution de l'estimation de $x_j^{(k+1)}$ en fonction des itérations k (on peut prendre par exemple $k = 5, 10, 15, 20, 25, 30$) pour différentes composantes j (examiner par exemple les trois premières composantes x_1, x_2, x_3) .
3.) Reprendre cette étude de simulation pour d'autres exemples en faisant varier r.