

ESSEC II 2023

On s'intéresse dans ce problème aux processus de Markov finis homogènes à temps continu et on étudie deux exemples de modélisation en lien avec les crédits bancaires.

Le problème comporte quatre parties. Les parties 2 et 3 sont indépendantes de la partie 4.

Soit $n \in \mathbb{N}^*$, $n \geq 2$. On considère, dans la suite du problème, une famille de variables aléatoires X_t , pour $t \in \mathbb{R}^+$, sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, vérifiant les propriétés suivantes :

(H₁) Pour tout $t \geq 0$, $X_t(\Omega) = \{1, \dots, n\}$.

(H₂) Pour tout $r \in \mathbb{N}^*$ et $t_1 < t_2 < \dots < t_r$ des réels positifs, i_1, \dots, i_{r+1} des éléments de $\{1, \dots, n\}$ et s un réel positif, si $\mathbb{P}([X_{t_1} = i_1] \cap \dots \cap [X_{t_r} = i_r]) \neq 0$,

$$\mathbb{P}_{[X_{t_1}=i_1] \cap \dots \cap [X_{t_r}=i_r]}([X_{t_r+s} = i_{r+1}]) = \mathbb{P}_{[X_{t_r}=i_r]}([X_{t_r+s} = i_{r+1}])$$

(H₃) Pour tout $i \in \{1, \dots, n\}$, la fonction $f_i : t \mapsto \mathbb{P}([X_t = i])$ est définie, dérivable sur \mathbb{R}^+ et n'est pas la fonction nulle. On note S_i l'ensemble des réels positifs t tels que $f_i(t) \neq 0$.

(H₄) Pour tout $(i, j) \in \{1, \dots, n\}^2$, $i \neq j$ et $h \geq 0$, la fonction $t \mapsto \mathbb{P}_{[X_t=i]}([X_{t+h} = j])$ est constante sur son ensemble de définition S_i et il existe un réel positif que l'on note $\alpha_{i,j}$, tel que, si $t \in S_i$ et $h \in \mathbb{R}^+$,

$$\mathbb{P}_{[X_t=i]}([X_{t+h} = j]) = \alpha_{i,j}h + o_{h \rightarrow 0}(h)$$

(H₅) Pour tout $i \in \{1, \dots, n\}$ et $h \geq 0$, la fonction $t \mapsto \mathbb{P}_{[X_t=i]}([X_{t+h} = i])$ est constante sur son ensemble de définition S_i et il existe un réel négatif que l'on note $\alpha_{i,i}$, tel que, si $t \in S_i$ et $h \in \mathbb{R}^+$,

$$\mathbb{P}_{[X_t=i]}([X_{t+h} = i]) = 1 + \alpha_{i,i}h + o_{h \rightarrow 0}(h)$$

Partie 1 - Matrice génératrice et système différentiel associés

On note L_t la matrice ligne d'ordre n , $(\mathbb{P}([X_t = 1]) \dots \mathbb{P}([X_t = n])) = (f_1(t) \dots f_n(t))$ et on note G la matrice carrée d'ordre n dont les coefficients sont les $\alpha_{i,j}$, appelée **matrice génératrice du processus**.

On note aussi pour tout $t \in \mathbb{R}^+$, $L'_t = (f'_1(t) \dots f'_n(t))$.

L'objectif des trois premières questions est d'établir que pour tout $t \geq 0$, $L'_t = L_t G$.

1. Montrer que, pour tous $j \in \{1, \dots, n\}$ et $(t, h) \in (\mathbb{R}^+)^2$,

$$\mathbb{P}([X_{t+h} = j]) = \sum_{i=1}^n \mathbb{P}([X_{t+h} = j] \cap [X_t = i])$$

2. Soit $i \in \{1, \dots, n\}$, $t \in S_i$ et $h \in \mathbb{R}^+$, justifier que $\sum_{j=1}^n \mathbb{P}_{[X_t=i]}([X_{t+h} = j]) = 1$. En déduire que, pour tout $i \in \{1, \dots, n\}$ et $h \in \mathbb{R}^+$, on a l'égalité :

$$1 = 1 + \left(\sum_{j=1}^n \alpha_{i,j} \right) h + o_{h \rightarrow 0}(h)$$

En conclure que $\sum_{j=1}^n \alpha_{i,j} = 0$.

3. a) Montrer que, pour tous $j \in \{1, \dots, n\}$ et $(t, h) \in (\mathbb{R}^+)^2$, on a alors :

$$\mathbb{P}([X_{t+h} = j]) = \mathbb{P}([X_t = j]) + \sum_{i=1}^n \left(\alpha_{i,j} h + o_{h \rightarrow 0}(h) \right) \mathbb{P}([X_t = i])$$

b) En déduire que pour tous $j \in \{1, \dots, n\}$, $t \geq 0$ et $h > 0$:

$$\frac{f_j(t+h) - f_j(t)}{h} = \sum_{i=1}^n f_i(t) \alpha_{i,j} + o_{h \rightarrow 0}(1)$$

En conclure que $f'_j(t) = \sum_{i=1}^n f_i(t) \alpha_{i,j}$.

c) Vérifier $L'_t = L_t G$.

4. Probabilité moyenne d'être dans un état

Soit $T > 0$ et U_T une variable aléatoire à valeurs dans $[0, T]$ qui suit la loi uniforme sur cet intervalle. On pose $Z_{i,T} = f_i(U_T)$.

Montrer que $\mathbb{E}(Z_{i,T})$ existe et vaut $\frac{1}{T} \int_0^T f_i(t) dt$. On note $e_i(T)$ cette espérance.

5. On suppose dans cette question que $n = 2$ et que $G = \begin{pmatrix} -a & a \\ b & -b \end{pmatrix}$ où a et b sont deux réels strictement positifs. On pose $p = \frac{b}{a+b}$, $q = 1 - p$, $\alpha = f_1(0)$.

a) Montrer que f_1 vérifie l'équation différentielle d'ordre 1 sur \mathbb{R}^+ , $y' + (a+b)y = b$.

b) En conclure que pour tout $t \geq 0$,

$$f_1(t) = p + (\alpha - p) \exp(-(a+b)t) \quad \text{et} \quad f_2(t) = q - (\alpha - p) \exp(-(a+b)t)$$

c) Montrer que pour tout $t \in \mathbb{R}^+$, $f_1(t) \in [\min(p, \alpha), \max(p, \alpha)]$ et que $\lim_{t \rightarrow +\infty} f_1(t) = p$.

d) Déterminer $\lim_{T \rightarrow +\infty} e_1(T)$.

6. On suppose dans cette question que $n = 3$ et $G = \frac{1}{30} \begin{pmatrix} -3 & 1 & 2 \\ 1 & -2 & 1 \\ 2 & 1 & -3 \end{pmatrix}$.

Pour tout $t \geq 0$, on note C_t (respectivement C'_t) la transposée de la matrice ligne L_t (respectivement L'_t).

a) Montrer que $-\frac{1}{6}$, $-\frac{1}{10}$ et 0 sont des valeurs propres de G .

b) On pose $P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -2 & 0 \\ 1 & 1 & -1 \end{pmatrix}$. Justifier que $G = \frac{1}{30} P \begin{pmatrix} 0 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -5 \end{pmatrix} P^{-1}$.

c) Calculer ${}^t P P$. En déduire que $P^{-1} = \frac{1}{6} \begin{pmatrix} 2 & 2 & 2 \\ 1 & -2 & 1 \\ 3 & 0 & -3 \end{pmatrix}$.

d) On pose, pour tout $t \in \mathbb{R}^+$, $P^{-1} C_t = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix}$.

Montrer que pour tout $t \in \mathbb{R}^+$, $y'_1(t) = 0$, $y'_2(t) = -\frac{1}{10} y_2(t)$, $y'_3(t) = -\frac{1}{6} y_3(t)$.

e) En conclure que, pour tout $t \geq 0$, $C_t = P \begin{pmatrix} \alpha \\ \beta e^{-\frac{1}{10}t} \\ \gamma e^{-\frac{1}{6}t} \end{pmatrix}$ où $\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = P^{-1} C_0$, puis que pour $i \in \{1, 2, 3\}$,

$$\lim_{t \rightarrow +\infty} \mathbb{P}([X_t = i]) = \frac{1}{3}.$$

7. *Temps initial passé dans un état* - On pose pour tout $i \in \{1, \dots, n\}$, $\beta_i = -\alpha_{i,i}$ et on suppose dans cette question que, si $\mathbb{P}([X_0 = i]) \neq 0$, alors $\beta_i \neq 0$.

On définit les variables aléatoires, Y_1, \dots, Y_n et Y égales, au premier instant t où $X_t \neq i$ pour Y_i et au premier instant t où $X_t \neq X_0$ pour Y . On admet que ces instants existent. Ainsi Y est à valeurs dans $]0, +\infty[$ et si $X_0 \neq i$, $Y_i = 0$.

Soit i tel que $\mathbb{P}([X_0 = i]) \neq 0$. On admet que pour tout $x > 0$, lorsque k est un entier naturel assez grand $\mathbb{P}\left(\bigcap_{j=0}^k [X_{\frac{j}{k}x} = i]\right) \neq 0$ et que l'on a :

$$\mathbb{P}([Y_i > x]) = \lim_{k \rightarrow +\infty} \mathbb{P}\left(\bigcap_{j=0}^k [X_{\frac{j}{k}x} = i]\right)$$

a) Montrer que pour tout $x > 0$ et $k \in \mathbb{N}^*$, k assez grand :

$$\mathbb{P}\left(\bigcap_{j=0}^k [X_{\frac{j}{k}x} = i]\right) = \mathbb{P}([X_0 = i]) \prod_{j=0}^{k-1} \mathbb{P}_{[X_{\frac{j}{k}x} = i]}([X_{\frac{j+1}{k}x} = i]) = \mathbb{P}([X_0 = i]) \left(1 - \frac{\beta_i}{k}x + o_{k \rightarrow +\infty}\left(\frac{1}{k}\right)\right)^k$$

b) En déduire que pour tout $x \geq 0$, $\mathbb{P}_{[X_0=i]}([Y_i > x]) = e^{-\beta_i x}$. Quelle est la loi de Y_i pour la probabilité conditionnelle $\mathbb{P}_{[X_0=i]}$?

c) Montrer que pour tout $x \geq 0$, $\mathbb{P}([Y > x]) = \sum_{k=1}^n \mathbb{P}([X_0 = k])e^{-\beta_k x}$.

d) En conclure que Y est une variable à densité et déterminer une densité de Y .

e) On note $I = \{k \in \{1, \dots, n\} \mid \mathbb{P}([X_0 = k]) \neq 0\}$. Établir que Y admet une espérance égale à $\sum_{k \in I} \frac{\mathbb{P}([X_0 = k])}{\beta_k}$.

Partie 2 - Matrice de transition, lien avec la matrice génératrice

On utilise les notations de la partie 1.

8. *Définition de la matrice de transition* - Pour tous $(i, j) \in \{1, \dots, n\}^2$ et $s \geq 0$, si $t \in S_i$, on pose

$$m_{i,j}(s) = \mathbb{P}_{[X_t=i]}([X_{t+s} = j])$$

qui ne dépend pas de t d'après les hypothèses (H_4) et (H_5) .

On note $M(s)$ la matrice d'éléments génériques $m_{i,j}(s)$.

a) Établir que pour tout $s \geq 0$, $L_s = L_0 M(s)$.

b) Soit $i \in \{1, \dots, n\}$, $r \in S_i$. En utilisant la propriété (H_2) et en distinguant les cas où $\mathbb{P}([X_r = i] \cap [X_{r+s} = k])$ est nulle ou non, montrer que pour tous $(j, k) \in \{1, \dots, n\}^2$, s et t des réels positifs :

$$\mathbb{P}([X_r = i] \cap [X_{r+s} = k] \cap [X_{r+s+t} = j]) = \mathbb{P}([X_r = i])m_{i,k}(s)m_{k,j}(t)$$

En déduire que, pour tous $j \in \{1, \dots, n\}$ et s, t des réels positifs,

$$\mathbb{P}([X_r = i] \cap [X_{r+s+t} = j]) = \mathbb{P}([X_r = i]) \sum_{k=1}^n m_{i,k}(s)m_{k,j}(t)$$

c) En conclure que pour tous s et t , des réels positifs, $M(s+t) = M(s)M(t)$.

d) Montrer que pour tout $k \in \mathbb{N}^*$ et t réel positif, $M(kt) = (M(t))^k$.

- Si $(A_k)_{k \geq 1}$ est une suite de matrices appartenant à $\mathcal{M}_n(\mathbb{R})$, A une matrice appartenant à $\mathcal{M}_n(\mathbb{R})$, si l'on note $a_{i,j}(k)$ le coefficient d'indice (i, j) de la matrice A_k , $a_{i,j}$ le coefficient d'indice (i, j) de A , alors on écrira $A = \lim_{k \rightarrow +\infty} A_k$ si pour $(i, j) \in \{1, \dots, n\}^2$, $\lim_{k \rightarrow +\infty} a_{i,j}(k) = a_{i,j}$.

On dit alors que la suite $(A_k)_{k \geq 1}$ converge vers A .

- On admet, dans la suite de cette partie et dans la partie 3, que pour tout $t \geq 0$,

$$M(t) = \lim_{k \rightarrow +\infty} \left(I_n + \frac{t}{k} G \right)^k \quad (**)$$

9. On veut simuler le processus à partir de la donnée de la matrice G et de L_0 . On admet que pour $t \in [0, 100]$, on peut considérer que $M(t) = \left(I_n + \frac{t}{1000} G \right)^{1000}$.

- On importe des bibliothèques :

`numpy as np, numpy.random as rd, matplotlib.pyplot as plt, numpy.linalg as al`

- On rappelle que si M est une matrice, représentée par un tableau `numpy`, `M[:, j]` désigne le vecteur des coefficients de la j -ème colonne de M , de même pour `M[i, :]` et la i -ème ligne de M .

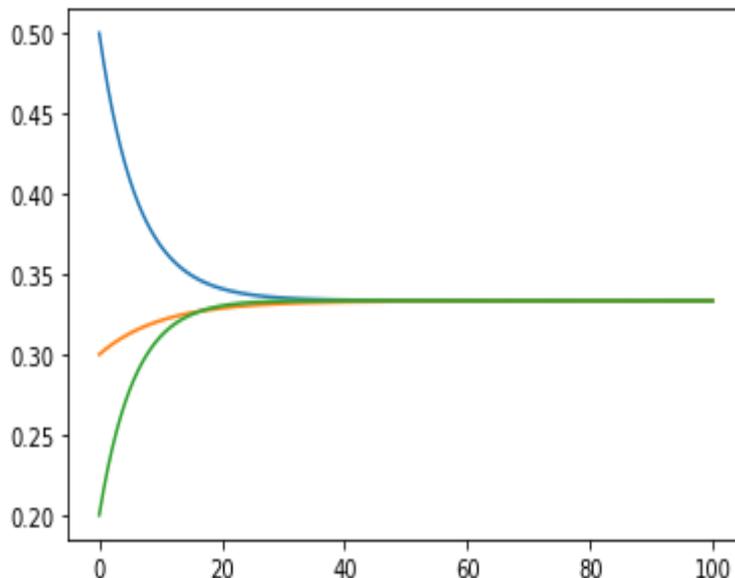
- a) Écrire une fonction **Python** `transition(t, G)` de paramètres G représentant la matrice génératrice carrée d'ordre n et t , qui renvoie la matrice $\left(I_n + \frac{t}{1000} G \right)^{1000}$.

- b) Utiliser la fonction précédente pour écrire une fonction `traceLoi2Xt(G, L0, tmax)` qui trace, sur un même graphique, les graphes des fonctions $t \mapsto \mathbb{P}([X_t = i])$ sur le segment $[0, t_{\max}]$ pour i variant de 1 à n , G et L_0 représentant, respectivement, la matrice génératrice du processus et la ligne L_0 .

On utilisera 1000 points pour les graphes.

- c) Si G est la matrice de la partie 1, question 6., l'instruction,

`traceLoi2Xt(1/30*np.array([[-3, 1, 2], [1, -2, 1], [2, 1, -3]]), 1/10*np.array([5, 3, 2]), 100)`
affiche l'image suivante :



Expliquer en quoi ce graphique est cohérent avec un résultat obtenu précédemment.

- d) On veut simuler et représenter, sur un même graphique, les valeurs de X_0, X_t, \dots, X_{kt} , pour $t > 0$ et $k \in \mathbb{N}^*$, à partir de la loi de X_0 donnée dans une ligne L_0 . Compléter la fonction suivante pour qu'elle réalise cette tâche :

```

1  def simulX(t,k,L0,G):
2      listeDesT=[] ; listeDesX=[]
3      Mt=transition(t,G) ; Lt = L0
4      for i in range(k+1):
5          listeDesT.append(i*t)
6          p=rd.random()
7          s=...
8          j=0
9          while p>...:
10             j+=1
11             s+=Lt[j]
12             Lt=...
13             listeDesX.append(j+1)
14             plt.plot(listeDesT,listeDesX) ; plt.show()

```

Partie 3 - Deux exemples de modélisations

On conserve les notations des deux premières parties.

10. On considère trois états pour le recouvrement d'un crédit bancaire après un défaut de paiement et un accord entre le débiteur et l'organisme de crédit sur la somme à recouvrer :

- 1 - en cours de recouvrement, lorsque le débiteur est en train de régulariser sa créance ;
- 2 - recouvré, lorsque le débiteur a honoré la totalité du montant dû ;
- 3 - non recouvré, lorsque l'organisme de crédit considère que l'argent est définitivement perdu.

La matrice génératrice G du processus de Markov modélisant ce phénomène est $\begin{pmatrix} -\alpha - \beta & \alpha & \beta \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

et $L_0 = (1 \ 0 \ 0)$, α et β étant des réels strictement positifs.

a) Montrer que pour tout $i \in \mathbb{N}^*$, $G^i = (-\alpha - \beta)^{i-1}G$.

b) En déduire que pour tous $k \in \mathbb{N}^*$ et t réel : $(I_3 + \frac{t}{k}G)^k = I_3 + \left(\sum_{i=1}^k \binom{k}{i} \left(\frac{t}{k}\right)^i (-\alpha - \beta)^{i-1} \right) G$.

c) Montrer que pour tous $k \in \mathbb{N}^*$ et t réel, $\sum_{i=1}^k \binom{k}{i} \left(\frac{t}{k}\right)^i (-\alpha - \beta)^{i-1} = \frac{1 - (1 - (\alpha + \beta)\frac{t}{k})^k}{\alpha + \beta}$ et en déduire que pour tout $t \geq 0$,

$$M(t) = I_3 + \frac{1 - \exp(-(\alpha + \beta)t)}{\alpha + \beta} G$$

d) En conclure que pour tout $t \geq 0$, $\mathbb{P}([X_t = 1]) = \exp(-(\alpha + \beta)t)$,

$$\mathbb{P}([X_t = 2]) = \frac{\alpha}{\alpha + \beta}(1 - \exp(-(\alpha + \beta)t)) \text{ et } \mathbb{P}([X_t = 3]) = \frac{\beta}{\alpha + \beta}(1 - \exp(-(\alpha + \beta)t))$$

e) En utilisant les résultats de la question 7. de la partie 1, montrer que le temps aléatoire passé en recouvrement suit la loi exponentielle de paramètre $\alpha + \beta$.

11. On distingue, pour l'accès au crédit d'une organisation, trois niveaux de solvabilité :

- 1 - niveau C ;
- 2 - niveau B ;
- 3 - niveau A.

On suppose que ce niveau évolue dans le temps suivant un processus de Markov avec

$$G = \frac{1}{3} \begin{pmatrix} -\alpha & \alpha & 0 \\ 0 & -\alpha & \alpha \\ 4\alpha & 0 & -4\alpha \end{pmatrix} \text{ et } L_0 = (1 \ 0 \ 0), \alpha > 0. \text{ On note aussi } A = \frac{1}{3} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -4 & 0 & 4 \end{pmatrix}$$

a) On admet que $A^3 = \frac{1}{27} \begin{pmatrix} -3 & -3 & 6 \\ 24 & -3 & -21 \\ -84 & 24 & 60 \end{pmatrix}$. Calculer $A^3 - 2A^2 + A$ (on explicitera A^2). Que peut-on dire du polynôme $U(x) = x^3 - 2x^2 + x$?

Soit $\theta \in \mathbb{R}$ et $k \in \mathbb{N}^*$, on admet qu'il existe un polynôme Q et des réels a, b, c tels que, pour tout x réel : $(1 + \frac{\theta}{k}x)^k = Q(x)U(x) + ax^2 + bx + c$ (*).

- b) Déterminer une factorisation de $U(x)$ et en déduire que $c = 1$ et $(1 + \frac{\theta}{k})^k = a + b + c$.
- c) En dérivant la relation (*), montrer que, $\theta (1 + \frac{\theta}{k})^{k-1} = 2a + b$.
En déduire que $a = \theta (1 + \frac{\theta}{k})^{k-1} - (1 + \frac{\theta}{k})^k + 1$, $b = 2 (1 + \frac{\theta}{k})^k - \theta (1 + \frac{\theta}{k})^{k-1} - 2$.
- d) En conclure que pour tout $t \geq 0$,

$$M(t) = (1 - (1 + \alpha t)e^{-\alpha t}) A^2 + ((2 + \alpha t)e^{-\alpha t} - 2)A + I_3$$

puis préciser la loi de X_t .

Partie 4 - Démonstration de l'égalité (***) admise dans la partie 2

On utilise les notations et définitions des deux premières parties.

- On définit pour $A = (a_{i,j})_{1 \leq i,j \leq n}$ appartenant à $\mathcal{M}_n(\mathbb{R})$, $\|A\| = \max_{1 \leq i \leq n} \left(\sum_{j=1}^n |a_{i,j}| \right)$ c'est-à-dire la plus grande valeur que prend $\sum_{j=1}^n |a_{i,j}|$ lorsque i décrit $\{1, \dots, n\}$.
- On admet que si $(A_k)_{k \geq 1}$ est une suite de matrices appartenant à $\mathcal{M}_n(\mathbb{R})$ et A appartenant aussi à $\mathcal{M}_n(\mathbb{R})$, $A = \lim_{k \rightarrow +\infty} A_k$ si et seulement si $\lim_{k \rightarrow +\infty} \|A_k - A\| = 0$.

12. Un exemple - Si $A = \frac{1}{3} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ -4 & 0 & 2 \end{pmatrix}$, montrer que $\|A\| = 2$.

13. Soit $t \geq 0$.

- a) Établir $\|M(t)\| = 1$.
- b) En utilisant la question 2. de la partie 1, montrer que pour $k \in \mathbb{N}^*$ assez grand, $\|I_n + \frac{t}{k}G\| = 1$.

14. Soient $A = (a_{i,j})_{1 \leq i,j \leq n}$ et $B = (b_{i,j})_{1 \leq i,j \leq n}$ deux matrices appartenant à $\mathcal{M}_n(\mathbb{R})$.

- a) Établir que $\|A + B\| \leq \|A\| + \|B\|$.
- b) Montrer que $\|A\| \leq \sum_{i=1}^n \sum_{j=1}^n |a_{i,j}|$.
- c) Démontrer que, $\|AB\| \leq \|A\|\|B\|$ puis que pour tout entier naturel n , $\|A^n\| \leq \|A\|^n$.
- d) Vérifier que pour tout $k \in \mathbb{N}^*$, $A^{k+1} - B^{k+1} = A(A^k - B^k) + (A - B)B^k$.

e) On pose $c = \max(\|A\|, \|B\|)$. Montrer, par récurrence sur k , que pour tout $k \in \mathbb{N}^*$,

$$\|A^k - B^k\| \leq kc^{k-1}\|A - B\|$$

15. Soit t un réel positif et $k \in \mathbb{N}^*$.

a) Justifier que $\left\| M\left(\frac{t}{k}\right) - \left(I_n + \frac{t}{k}G\right) \right\| = o_{k \rightarrow +\infty}\left(\frac{1}{k}\right)$.

b) Montrer que pour tout k assez grand,

$$\left\| M\left(\frac{t}{k}\right)^k - \left(I_n + \frac{t}{k}G\right)^k \right\| \leq k \left\| M\left(\frac{t}{k}\right) - \left(I_n + \frac{t}{k}G\right) \right\|$$

c) En conclure que $M(t) = \lim_{k \rightarrow +\infty} \left(I_n + \frac{t}{k}G\right)^k$.