

GRAPHES PROBABILISTES ET CHAINES DE MARKOV

1 Introduction et rappels

Ce qui suit a pour but d'expliquer le cadre général d'un certain nombre d'exercices de Terminale ES qui portent sur les graphes probabilistes ou les chaînes de Markov. On verra que le problème sous-jacent est essentiellement le calcul de la puissance n -ième d'une matrice. Comme les élèves de TES ne disposent pas de la réduction des matrices (diagonalisation, etc.), les exercices sont de deux types :

- Soit, et c'est le cas le plus souvent en dimension 3, ils se contentent d'une approche expérimentale à la calculatrice pour déterminer le n -ième itéré (avec n petit) et constater qu'il semble exister un état limite.

- Soit, et c'est parfois le cas en dimension 2, ils proposent une méthode aboutissant au calcul explicite de A^n et de sa limite, mais en utilisant une procédure *ad hoc*, dont le moins qu'on puisse dire est qu'elle n'est pas transparente.

Le but de ce papier est d'abord de présenter la théorie générale, puis son adaptation aux exemples élémentaires.

1.1 Exemple introductif

Un individu est susceptible de contracter une maladie. Il y a trois états possibles : Immunisé, Sain (mais non immunisé), Malade. On peut modéliser ce système par un graphe¹ orienté et pondéré (avec des boucles). La matrice de transition² entre ces états (de semaine en semaine) est (dans l'ordre I, S, M) :

$$A = \begin{pmatrix} 0,9 & 0,1 & 0 \\ 0 & 0,5 & 0,5 \\ 0,8 & 0 & 0,2 \end{pmatrix}.$$

Cela signifie, par exemple, que si l'on est immunisé on a 9 chances sur 10 de le rester et 1 chance sur 10 de perdre son immunisation en restant sain, etc.

1. L'apport du graphe dans ces questions n'est pas considérable et je ferai l'impasse là-dessus dans ce qui suit.

2. Je considère qu'il serait plus simple et plus conforme à l'intuition géométrique de travailler avec la matrice transposée, ce qui permettrait d'écrire l'équation d'évolution avec des matrices colonnes V_n sous la forme $V_{n+1} = AV_n$. Mais la tradition des probabilistes est d'écrire avec des matrices lignes ...

(La somme des termes des lignes doit évidemment être égale à 1.)

La matrice permet de calculer l'évolution de la population. Si on part d'une population initiale avec des proportions des trois types données par la matrice ligne $V = (x, y, z)$: $100x\%$ d'immunisés, $100y\%$ de sains, $100z\%$ de malades, on passe, en une semaine, à l'état $V' = VA$ i.e. $x' = 0,9x + 0,8z$, $y' = 0,1x + 0,5y$, $z' = 0,5y + 0,2z$. En effet, les immunisés x' proviennent soit des 90% d'immunisés qui le sont restés : $0,9x$, soit des 20% de malades qui se sont immunisés : $0,2z$ et de même pour les autres.

Notons V_n la matrice ligne des proportions des trois types à la n -ième semaine. L'évolution du système est donnée par la formule $V_{n+1} = V_n A$ et on en déduit, par une récurrence immédiate, $V_n = V_0 A^n$.

On trouve, par exemple :

$$A^2 = \begin{pmatrix} 0,81 & 0,14 & 0,05 \\ 0,4 & 0,25 & 0,35 \\ 0,88 & 0,08 & 0,04 \end{pmatrix}.$$

Cela signifie, par exemple, que si l'on est immunisé au départ, au bout de deux semaines, on le reste avec une probabilité de $0,81$, mais qu'on peut perdre son immunité avec une probabilité de $0,14$ et qu'on peut même tomber malade avec une petite probabilité ($0,05$).

Ce qui est intéressant c'est de regarder l'évolution à plus long terme. On constate que la matrice A^n semble avoir une limite, limite vers laquelle elle converge vite (dès $n = 5$ on la perçoit clairement). De plus, cette matrice limite L a ses trois lignes égales, ce qui signifie que, quel que soit l'état initial $V_0 = (I, S, M)$, on a les mêmes chances d'être, au bout d'un temps assez long, dans l'un des trois états.

Dans notre cas, la matrice limite est environ :

$$A^2 = \begin{pmatrix} 0,754716 & 0,1509433 & 0,0943396 \\ 0,754716 & 0,1509433 & 0,0943396 \\ 0,754716 & 0,1509433 & 0,0943396 \end{pmatrix}.$$

L'état limite est donc $(0,754716, 0,1509433, 0,0943396)$.

Les deux questions essentielles suggérées par cet exemple sont les suivantes :

1) Pourquoi la matrice A^n admet-elle une limite et qu'est-ce qui explique la forme de cette limite ?

2) Comment trouver *a priori* ce mystérieux état limite ?

1.2 Rappels sur les suites de matrices

Une matrice $p \times p$ réelle est donnée par p^2 coefficients et elle peut donc se voir comme un élément de \mathbf{R}^{p^2} . On peut munir l'espace des matrices d'une norme qui définit la topologie usuelle de \mathbf{R}^{p^2} (par exemple le *sup* des valeurs absolues des coefficients, ou n'importe quelle autre norme). On a ainsi une notion de convergence des suites de matrices : la suite A_n , de coefficients $a_{i,j;n}$, converge vers la matrice A de coefficients $a_{i,j}$ si et seulement si, pour tous i, j , la suite $(a_{i,j;n})$ converge vers $a_{i,j}$.

On note que les opérations usuelles sur les matrices (produit, somme, inverse) sont continues là où elles sont définies.

2 La théorie

2.1 Définitions

2.1 Définition. Soit p un entier positif et soit A une matrice $p \times p$. On dit que A est une matrice **stochastique** (resp. **anti-stochastique**) si ses coefficients sont ≥ 0 et si la somme des termes de chaque ligne (resp. de chaque colonne) de A est égale à 1. Si les coefficients de A sont > 0 on dit que A est **strictement positive**. Un vecteur ligne ou colonne sera dit **stochastique** si ses coefficients sont ≥ 0 et de somme égale à 1.

2.2 Remarques.

- 1) On notera que la limite d'une suite de matrices stochastiques (resp. anti-stochastiques) est encore une matrice stochastique (resp. anti-stochastique). En effet, l'ensemble des matrices stochastiques est défini par des égalités (sommes égales à 1) et des inégalités larges (coefficients ≥ 0) et c'est donc un fermé.
- 2) Les coefficients a, b, c, d d'une matrice anti-stochastique sont compris entre 0 et 1, strictement si la matrice est strictement positive.
- 3) Deux vecteurs stochastiques proportionnels sont nécessairement égaux.

2.2 Le théorème de Perron-Frobenius

Le résultat qui explique les phénomènes entrevus ci-dessus est le suivant :

2.3 Théorème. (Perron-Frobenius) Soit A une matrice stochastique strictement positive $p \times p$. Alors, le nombre 1 est valeur propre de A , un vecteur propre associé étant $(1, 1, \dots, 1)$. C'est une valeur propre simple. Toutes les autres valeurs propres complexes de A sont de module < 1 .

2.4 Remarque. Le même résultat (à l'exception du vecteur propre) vaut évidemment pour une matrice anti-stochastique. En effet, si A est anti-stochastique, sa transposée est stochastique et admet le même polynôme caractéristique, voir 2.10 ci-dessous.

Démonstration. Que 1 soit valeur propre avec le vecteur propre annoncé est une conséquence immédiate du fait que A est stochastique.

Soit λ une valeur propre complexe de A et soit x un vecteur propre non nul associé. On a donc $Ax = \lambda x$, c'est-à-dire $\lambda x_i = \sum_{j=1}^p a_{ij}x_j$. En prenant les

modules, on a $|\lambda||x_i| \leq \sum_{j=1}^p a_{ij}|x_j|$. Si x_i désigne la coordonnée de plus grand

module, on en déduit, en tenant compte de $\sum_{j=1}^p a_{ij} = 1$, qu'on a $|\lambda| \leq 1$. De plus l'égalité ne peut avoir lieu que si tous les x_i sont de même module, et même si tous les x_i sont égaux, en vertu du lemme suivant :

2.5 Lemme. Soient x_1, \dots, x_n des complexes de modules $\leq r$ et soient $\alpha_1, \dots, \alpha_n$ des réels > 0 de somme 1. Si on a $|\alpha_1 x_1 + \dots + \alpha_n x_n| = r$, les x_i sont tous égaux.

Démonstration. Si deux des x_i sont distincts le barycentre de ces deux là est intérieur au disque de centre 0 et de rayon r , et par associativité le barycentre total aussi.

Cela montre que le seul vecteur propre possible pour une valeur propre λ de module 1 est (x_i, \dots, x_i) , qui n'est autre, à un scalaire près, que $(1, \dots, 1)$ et cela impose $\lambda = 1$. On a donc montré que les valeurs propres différentes de 1 sont de module < 1 .

Il reste à montrer que la valeur propre 1 est simple. On note déjà que l'argument précédent montre que le seul vecteur propre pour 1 est $e = (1, 1, \dots, 1)$. Si l'ordre de la valeur propre était supérieur ou égal à 2 il y aurait un vecteur réel $f = (x_1, \dots, x_p)$ tel que $Af = f + e$ (cela résulte de la forme de Jordan, ou du lemme 2.6 ci-dessous). On aurait donc $\sum_{j=1}^p a_{ij}x_j = x_i + 1$.

Appliquant ceci avec le plus grand des x_i on a une contradiction (toujours parce que la somme des termes des lignes vaut 1).

2.6 Lemme. Si λ est valeur propre d'ordre ≥ 2 de la matrice A et si l'espace propre est de dimension 1 engendré par e , il existe x tel que $Ax = \lambda x + e$.

Démonstration. Quitte à considérer $A - \lambda I$ on peut supposer $\lambda = 0$, donc $\text{Ker } A = (e)$. Il s'agit de montrer qu'on a $\text{Ker } A \subset \text{Im } A$. Sinon, ces espaces

seraient en somme directe et si on écrit A dans une base adaptée à cette somme elle a une décomposition par blocs $A = \begin{pmatrix} 0 & 0 \\ 0 & A' \end{pmatrix}$ avec A' inversible. Mais alors, le calcul de $\det(A - \lambda I)$ montre que 0 est valeur propre simple de A .

2.7 Corollaire. *Les résultats du théorème 2.3 valent encore si l'on suppose seulement qu'une certaine puissance A^r est strictement positive.*

Démonstration. On applique Perron-Frobenius à A^r , dont les valeurs propres sont les λ_i^r . Elle admet la valeur propre 1 simple et ses autres valeurs propres sont de module < 1 . Il en résulte que 1 est valeur propre simple de A (si elle était de multiplicité ≥ 2 dans A cela serait encore vrai pour A^r). Les autres λ_i^r sont de module < 1 , donc aussi les λ_i .

2.3 Conséquences

La première conséquence de Perron-Frobenius est la convergence de la suite A^n , avec la forme repérée particulière de la limite :

2.8 Corollaire. *Soit A une matrice stochastique dont une puissance est strictement positive. Lorsque n tend vers $+\infty$ la matrice A^n tend vers une matrice*

L de rang 1 de la forme $L = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_p \\ \alpha_1 & \alpha_2 & \cdots & \alpha_p \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_1 & \alpha_2 & \cdots & \alpha_p \end{pmatrix}$ avec $\alpha_1 + \alpha_2 + \cdots + \alpha_p = 1$.

Démonstration. Supposons d'abord A diagonalisable sur \mathbf{C} . On a donc $A = P^{-1}DP$ avec D diagonale de coefficients $\lambda_1 = 1, \lambda_2, \dots, \lambda_p$ et on a $A^n = P^{-1}D^nP$. Comme les λ_i pour $i > 1$ sont de modules < 1 , λ_i^n tend vers 0 quand n tend vers $+\infty$ et la limite de D^n est une matrice diagonale Δ de coefficients $(1, 0, \dots, 0)$. Il en résulte que A^n tend vers $P\Delta P^{-1} = L$ qui est de rang 1 (comme Δ) et stochastique (car A^n l'est et que c'est une condition fermée). Cette matrice a toutes ses lignes proportionnelles et stochastiques, donc égales. Elle est donc de la forme annoncée.

Si A n'est pas diagonalisable elle est semblable à $D + N$ où N est une matrice nilpotente qui commute avec D . On voit facilement que $(D + N)^n$ tend encore vers Δ (appliquer la formule du binôme en tenant compte de $N^p = 0$, et noter qu'on a $\Delta N = 0$).

2.9 Remarque. Une matrice stochastique strictement positive n'est pas nécessairement

diagonalisable. C'est le cas par exemple, de la matrice : $\begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{6} & \frac{5}{12} & \frac{5}{12} \end{pmatrix}$ qui est

stochastique et strictement positive. On vérifie qu'elle est de rang 2, avec la valeur propre 0 double, donc non diagonalisable.

Le dernier résultat précise l'état limite :

2.10 Corollaire. *Le vecteur $e = (\alpha_1, \dots, \alpha_p)$ est l'unique vecteur propre stochastique de la matrice tA relatif à la valeur propre 1.*

Démonstration. La matrice tA a le même polynôme caractéristique que A , donc les mêmes valeurs propres, avec la même multiplicité. Soit e l'unique vecteur propre stochastique de tA relatif à 1. On a donc ${}^tA(e) = e$, donc aussi ${}^tA^n(e) = e$ et, à la limite, ${}^tL(e) = e$. Le vecteur e est donc dans l'image de tL . Mais cette image est engendrée par le vecteur $(\alpha_1, \dots, \alpha_p)$ qui est son unique vecteur stochastique.

2.4 Retour à l'exemple introductif

La matrice A^2 étant strictement positive, le théorème de Perron-Frobenius s'applique. L'état limite est un vecteur ligne V qui vérifie $VA = V$, ou encore, le vecteur colonne tV est un vecteur propre stochastique pour la matrice tA et la valeur propre 1. On calcule aisément un tel vecteur : $(\frac{40}{53}, \frac{8}{53}, \frac{5}{53})$. On retrouve les valeurs approchées vues ci-dessus.

Dans le cas d'une matrice 3×3 , il est difficile de prouver les résultats ci-dessus au niveau TES. En revanche, pour les matrices 2×2 , on peut le faire et c'est l'objet des sections suivantes.

3 Le cas $p = 2$: adaptation de la théorie

Dans ce qui suit, on utilisera un langage géométrique et non probabiliste (autrement dit on agit sur les vecteurs colonnes et pas sur les vecteurs lignes). L'objectif de ce paragraphe est de montrer que le théorème de Perron-Frobenius est complètement élémentaire (niveau L1-L2) dans le cas 2×2 .

3.1 La situation

Rappelons la définition dans le cas $p = 2$:

3.1 Définition.

1) Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice 2×2 à coefficients réels. On dit que A est anti-stochastique (resp. anti-stochastique strictement positive) si ses

coefficients sont ≥ 0 (resp. > 0) et si on a $a + c = b + d = 1$.

2) Un vecteur colonne $U = \begin{pmatrix} x \\ y \end{pmatrix}$ est dit **stochastique** si on a $x, y \geq 0$ et $x + y = 1$.

3.2 Valeurs et vecteurs propres

3.2 Proposition. Soit A une matrice anti-stochastique. Elle admet la valeur propre 1 et une autre valeur propre λ avec $-1 \leq \lambda \leq 1$. Si la matrice est strictement positive, on a $-1 < \lambda < 1$. Dans tous les cas, la matrice A est diagonalisable.

Démonstration. On considère tA qui a le même polynôme caractéristique, donc les mêmes valeurs propres que A . Elle admet la valeur propre 1, associée au vecteur propre $(1, 1)$. Il en résulte que A admet la valeur propre 1. Elle admet donc une deuxième valeur propre réelle $\lambda = \text{Tr } A - 1 = a + d - 1$. Comme on a $a + b + c + d = 2$ et que les coefficients sont ≥ 0 , on obtient $0 \leq a + d \leq 2$, d'où $-1 \leq \lambda \leq 1$. Si la matrice est strictement positive, les inégalités sont strictes.

Montrons que A est diagonalisable. C'est vrai si ses valeurs propres sont distinctes et c'est le cas, en particulier, si elle est strictement positive. Si les valeurs propres sont toutes deux égales à 1 on a $\text{Tr } A = a + d = 2$ et comme les coefficients sont ≤ 1 , cela impose $a = d = 1$, donc $b = c = 0$, de sorte que A est la matrice identité, qui est diagonale.

3.3 Itération

Le théorème suivant résume l'ensemble des propriétés des itérées de A :

3.3 Théorème. Soit A une matrice anti-stochastique strictement positive, et soit λ sa valeur propre distincte de 1. On pose $D = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}$ et on note P une matrice de passage qui diagonalise A .

1) On a $A = P^{-1}DP$, $D^n = \begin{pmatrix} 1 & 0 \\ 0 & \lambda^n \end{pmatrix}$ pour $n \in \mathbf{N}$ et cette matrice tend, quand n tend vers l'infini, vers $D_\infty = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

2) On a $A^n = P^{-1}D^nP$ et cette matrice tend, quand n tend vers l'infini vers la matrice $B = P^{-1}D_\infty P$. La matrice B est anti-stochastique et de rang 1. Elle est de la forme $B = \begin{pmatrix} e & e \\ f & f \end{pmatrix}$ avec $e, f \geq 0$ et $e + f = 1$.

3) Le vecteur stochastique $L = \begin{pmatrix} e \\ f \end{pmatrix}$ est l'unique vecteur propre stochastique de A relatif à la valeur propre 1. Les nombres e et f sont > 0 , précisément on a $e = \frac{b}{b+c}$ et $f = \frac{c}{b+c}$.

4) Si U est un vecteur colonne stochastique : $U = \begin{pmatrix} x \\ y \end{pmatrix}$, $A^n U$ tend vers L quand n tend vers l'infini.

Démonstration. 1) L'expression de D^n est évidente et sa limite aussi.

2) La convergence de A^n vient de la continuité des opérations. Comme l'ensemble des matrices anti-stochastiques est un fermé, la matrice A est anti-stochastique. Comme D_∞ est de rang 1 il en est de même de B . On a donc $B = \begin{pmatrix} e & e' \\ \alpha e & \alpha e' \end{pmatrix}$, avec $e(1 + \alpha) = e'(1 + \alpha) = 1$, donc $e = e'$ et la forme annoncée. Les coefficients e et f sont ≥ 0 comme limites de suites de nombres positifs.

3) Soit V un vecteur colonne propre de A relatif à la valeur propre 1. On a donc $AV = V$, donc aussi $A^n V = V$ et, par passage à la limite, $BV = V$. Le vecteur V est donc proportionnel à L et, s'il est stochastique, il lui est égal. Cela montre que e et f sont > 0 . Sinon, on aurait par exemple $e = 1$ et $f = 0$. Comme V est vecteur propre de A relatif à 1, cela impose $a = 1$ et $c = 0$, de sorte que A n'est pas strictement positive et c'est absurde. Un calcul immédiat donne les valeurs de e et f .

4) Enfin, comme A^n tend vers B , la continuité des opérations montre que $A^n U$ tend vers BU et cette matrice a pour coefficients $(x + y)e = e$ et $(x + y)f = f$.

4 Le cas $p = 2$: version élémentaire

Il s'agit maintenant de montrer comment on peut adapter ce qui précède pour bâtir un exercice au niveau terminale ES, sans trop parachuter les indications et de comprendre pourquoi l'exercice fonctionne.

On part de la matrice A anti-stochastique strictement positive. On calcule ses puissances A^2, A^3 , etc. à la main, puis A^{10}, A^{20} , etc. à l'aide de la calculatrice. On voit apparaître la matrice limite B . On constate qu'elle a deux colonnes égales et stochastiques et cela montre qu'on a $B^2 = B$ (soit par un calcul direct, soit en regardant les images des vecteurs de base). Bien entendu, cela n'étonne personne puisque B est semblable à D_∞ qui est évidemment idempotente (c'est la projection sur le premier vecteur de base).

On considère alors la matrice différence $C = A - B$. On a $A = B + C$ et, comme on veut calculer les puissances de A , on calcule d'abord C^2 . On trouve $C^2 = \lambda C$ où λ est un nombre strictement compris entre -1 et 1 . Là encore, c'est évident si l'on a vu la théorie, puisque C est semblable à $D - D_\infty = \begin{pmatrix} 0 & 0 \\ 0 & \lambda \end{pmatrix}$ et c'est donc λ fois un projecteur.

On commence le calcul de $A^2 = (B + C)^2 = B^2 + BC + CB + C^2$ en faisant évidemment attention à l'ordre des facteurs. On doit donc calculer BC et CB et on constate avec ravissement qu'ils sont tous les deux nuls. Bien entendu, c'est normal, puisque les matrices D_∞ et $D - D_\infty$ vérifient aussi cette propriété.

On a alors, par une récurrence facile, $A^n = B + \lambda^n C$. Évidemment, quand n tend vers l'infini, A^n tend vers B puisque λ^n tend vers 0 (au niveau TES, on le constate sur chacun des termes de A^n).

4.1 Remarque. Dans certains exercices (voir l'exemple ci-dessous) on introduit $C' = \lambda^{-1}C$ qui est idempotente. On constate alors que la matrice C' a presque les mêmes coefficients que B , précisément, si on a $B = \begin{pmatrix} e & e \\ f & f \end{pmatrix}$, on a $C' = \begin{pmatrix} f & -e \\ -f & e \end{pmatrix}$.

L'explication de ce phénomène est triple. D'abord, il est clair que la matrice C' est un projecteur. En effet, elle est de rang 1 et sa trace est la même que celle de B , donc égale à 1 .

Ensuite, on a $BC' = CB' = 0$ car la matrice C' est la transposée de la matrice des cofacteurs de B et, comme on a $\det B = 0$, le produit est bien la matrice nulle.

Enfin, on montre aisément (en se plaçant dans une base où B est diagonale) que si B est un projecteur, il y a un unique projecteur C' qui vérifie $BC' = CB' = 0$ et c'est donc celui rencontré ci-dessus.

4.1 Un exemple

Voici un exercice de niveau TES, utilisant les techniques précédentes :

Le glucose admet deux isomères, l'un dextrogyre (c'est-à-dire qui dévie la lumière vers la droite), le dextrose, l'autre lévogyre, le lévulose. Dans une certaine réaction chimique, dite de Pierre Landin, une partie du dextrose se transforme en lévulose et vice-versa. Précisément, au cours d'une unité de temps, 60% des molécules de dextrose se transforment en lévulose et les 40% restant demeurent sous forme dextrose, tandis que 35% des molécules de lévulose restent sous forme lévulose et que les 65% autres deviennent dextrose.

On appelle d_n et l_n les proportions de dextrose et de lévulose au temps n dans le mélange (l'unité de temps est l'heure). On suppose qu'on a, au départ, $d_0 = 1$ et $l_0 = 0$. On note U_n le vecteur colonne $\begin{pmatrix} d_n \\ l_n \end{pmatrix}$.

1) Montrer que le phénomène se traduit par une égalité matricielle de la forme $U_{n+1} = AU_n$ où A est une matrice 2×2 que l'on précisera. En déduire qu'on a $U_n = A^n U_0$.

2) Calculer U_1, U_2, U_3 . Interpréter les résultats en termes de proportions de chaque composant.

3) À l'aide de la calculatrice, calculer A^n pour $n = 10, 20, 30$. Que constate-t-on ? Qu'en déduit-on pour les proportions de dextrose et de lévulose ?

4) On pose $B = \begin{pmatrix} 0,52 & 0,52 \\ 0,48 & 0,48 \end{pmatrix}$. Calculer $B - A$ et $C = 4(B - A)$.

a) Calculer les puissances B^2, B^3, \dots, B^n .

b) Calculer C^2, C^3, \dots, C^n .

c) Calculer BC et CB .

d) Déduire de ce qui précède la formule :

$$A^n = \begin{pmatrix} 0,52 + (-0,25)^n \cdot 0,48 & 0,52(1 - (-0,25)^n) \\ 0,48(1 - (-0,25)^n) & 0,48 + (-0,25)^n \cdot 0,52 \end{pmatrix}.$$

e) Calculer d_n et l_n et déterminer les limites de d_n et l_n quand n tend vers l'infini.

5 Une variante utilisant les suites arithmético-géométriques pour le cas $n = 2$

Reprenons le cas d'une matrice anti-stochastique $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ à coefficients réels strictement positifs. On a donc $a + c = b + d = 1$.

Si l'on part d'un vecteur (colonne) stochastique $U_0 = (x_0, y_0)$, on cherche à calculer $U_n = A^n U_0 = (x_n, y_n)$ et à déterminer sa limite. Une méthode pour cela consiste à se ramener à l'étude d'une suite arithmético-géométrique.

On montre d'abord par récurrence sur n que U_n est un vecteur stochastique (i.e. qu'on a $x_n + y_n = 1$). On note ensuite qu'on a les relations de récurrence $x_{n+1} = ax_n + by_n$ et $y_{n+1} = cx_n + dy_n$. Remplaçant y_n par $1 - x_n$ dans la première relation, on obtient la relation (*) : $x_{n+1} = (a - b)x_n + b$. On est en présence d'une suite arithmético-géométrique. Parmi les solutions de cette équation il y a la solution constante $x_n = x_{n+1} = e$ (le point fixe,

qui sera la limite éventuelle de la suite). Ici, on a $e = \frac{b}{1-a+b} = \frac{b}{b+c}$ (ce nombre est bien défini car b et c sont > 0). La solution générale de l'équation (*) est alors de la forme $x_n = e + \lambda(a-b)^n$ et on calcule la valeur de λ avec le cas $n = 0$. On en déduit $y_n = f - \lambda(a-b)^n$ avec $f = \frac{c}{b+c}$.

On a l'encadrement $-1 \leq -b \leq a-b \leq a \leq 1$. Si l'on a $a-b = 1 = a+c$, on en déduit $b+c = 0$ donc $b = c = 0$ et c'est absurde et de même si $a-b = -1$. Le nombre $a-b$ est donc strictement compris entre -1 et 1 et la suite géométrique associée tend vers 0 lorsque n tend vers $+\infty$. Il en résulte que x_n tend vers e et y_n vers f et on retrouve les résultats attendus.