



PRÉPASIMPLE

prepasimple.fr

ESSEC

Mathématiques 2 — Option économique

2008

Mathématiques appliquées

Mathématiques appliquées — Prépa ECG

Annales de concours · Corrigé

Préliminaires

Dans ces préliminaires, le vecteur V satisfait l'égalité (E) . Les sommes utilisées sont finies.

P1. Si $x \geq 0$, alors $|x| - x = x - x = 0$. Si $x < 0$, alors $|x| - x = -2x > 0$. Dans tous les cas, $|x| - x \geq 0$; l'égalité a lieu exactement lorsque $x \geq 0$.

P2.a. On développe séparément les deux carrés :

$$(|v_1| + |v_2| + |v_3|)^2 = v_1^2 + v_2^2 + v_3^2 + 2|v_1||v_2| + 2|v_1||v_3| + 2|v_2||v_3|,$$

$$(v_1 + v_2 + v_3)^2 = v_1^2 + v_2^2 + v_3^2 + 2v_1v_2 + 2v_1v_3 + 2v_2v_3.$$

En soustrayant et en utilisant $|a||b| = |ab|$, on obtient bien

$$\begin{aligned} (|v_1| + |v_2| + |v_3|)^2 - (v_1 + v_2 + v_3)^2 &= 2(|v_1v_2| - v_1v_2) + 2(|v_1v_3| - v_1v_3) \\ &+ 2(|v_2v_3| - v_2v_3). \end{aligned}$$

P2.b. L'égalité (E) entraîne l'égalité des carrés de ses deux membres. Le membre de gauche de l'identité précédente vaut donc 0. Chacun des trois termes de droite est positif ou nul d'après P1. Une somme finie de nombres positifs ou nuls étant nulle uniquement si tous ses termes sont nuls, on en déduit $|v_jv_{j'}| - v_jv_{j'} = 0$ pour $j \neq j'$.

Ainsi $v_jv_{j'} \geq 0$. Si $|v_jv_{j'}| > 0$, aucun des deux facteurs n'est nul et leur produit est strictement positif; ils ont donc même signe.

P2.c. Si tous les coefficients sont nuls, les deux égalités proposées sont vraies. Sinon, choisissons un coefficient non nul. D'après P2.b, tout autre coefficient non nul a le même signe que celui-ci. Tous les coefficients sont donc positifs ou nuls, ou tous négatifs ou nuls. Dans le premier cas, $V = |V|$; dans le second, $V = -|V|$.

P3. Méthode 1 — Utiliser le signe de la somme

Posons

$$s = \sum_{j=1}^N v_j.$$

Si $s \geq 0$, alors (E) donne

$$\sum_j |v_j| = s = \sum_j v_j,$$

donc

$$\sum_{j=1}^N (|v_j| - v_j) = 0.$$

Tous les termes sont positifs ou nuls par P1 ; ils sont donc tous nuls. Ainsi $v_j \geq 0$ pour tout j , et $V = |V|$.

Si $s < 0$, alors (E) donne

$$\sum_j |v_j| = -s = -\sum_j v_j,$$

donc

$$\sum_{j=1}^N (|v_j| + v_j) = 0.$$

On applique P1 à $-v_j : |v_j| + v_j \geq 0$. Chaque terme est nul, donc $v_j \leq 0$ et $V = -|V|$.

La conclusion est valable pour tout $N \geq 2$.

Méthode 2 — Généraliser le développement par paires

En développant les carrés comme dans P2.a, on obtient

$$\left(\sum_{j=1}^N |v_j| \right)^2 - \left(\sum_{j=1}^N v_j \right)^2 = 2 \sum_{1 \leq j < k \leq N} (|v_j v_k| - v_j v_k).$$

Le membre de gauche est nul par (E). Tous les termes de droite sont positifs ou nuls par P1 ; ainsi $v_j v_k \geq 0$ pour chaque paire. Si le vecteur est non nul, fixons un coefficient non nul : tous les autres coefficients non nuls ont le même signe que lui. Cela donne encore $V = |V|$ ou $V = -|V|$; le vecteur nul vérifie aussi la conclusion.

La première méthode est plus courte et traite directement tous les N ; la seconde prolonge le calcul demandé pour trois composantes et met en évidence le rôle des produits entre paires de coefficients.

Partie I — Google et PageRank

A. Étude de la matrice G de Google

I.A.1. Tous les coefficients de A sont positifs ou nuls. Comme $0 \leq \rho < 1$ et $N \geq 2$, on a, pour tous i, j ,

$$G(i, j) = \rho A(i, j) + \frac{1 - \rho}{N} \geq \frac{1 - \rho}{N} > 0.$$

La matrice réelle G est donc strictement positive.

I.A.2. Lorsque $d_j = 0$, la colonne j de A est la colonne j de I_N . Ainsi

$$\sum_i A(i, j) = 1,$$

et

$$\sum_{i=1}^N G(i, j) = \rho \sum_{i=1}^N A(i, j) + \sum_{i=1}^N \frac{1-\rho}{N} = \rho + 1 - \rho = 1.$$

I.A.3. Parmi les N pages, exactement d_j sont visées par un lien de la page j . Pour chacune d'elles, $G(i, j) = \rho/d_j + (1-\rho)/N$. Pour chacune des $N - d_j$ autres, $G(i, j) = (1-\rho)/N$. La partition indiquée donne donc

$$\sum_{i=1}^N G(i, j) = d_j \left(\frac{\rho}{d_j} + \frac{1-\rho}{N} \right) + (N - d_j) \frac{1-\rho}{N} = \rho + 1 - \rho = 1.$$

La division par d_j est légitime puisque cette question suppose $d_j > 0$.

I.A.4. Les questions I.A.2 et I.A.3 couvrent toutes les colonnes. Celles-ci sont de somme 1, et les coefficients sont positifs d'après I.A.1. Par définition, G est une matrice stochastique strictement positive, dans la convention par colonnes du sujet.

I.A.5. Par définition du produit matriciel, la composante i de GP_* vaut

$$\sum_{j=1}^N G(i, j)p(j).$$

Le système (S) affirme que cette somme est $p(i)$ pour tout i . Donc $GP_* = P_*$: le vecteur P_* est invariant par G . Il est aussi un vecteur de probabilité, puisque ses coefficients sont positifs ou nuls et leur somme vaut 1.

B. Modèle du surfeur sur le Web

I.B.1. Chaque $P(X_n = i)$ est positif ou nul. Comme X_n prend ses valeurs dans $\{1, \dots, N\}$, les événements $[X_n = i]$ forment un système complet, et

$$\|V_n\|_1 = \sum_{i=1}^N P(X_n = i) = 1.$$

Ainsi V_n est un vecteur de probabilité. Le même raisonnement s'applique à V_0 .

La convention colonne préservée ici correspond à la convention ligne avec $M = {}^tG$ et $W_n = {}^tV_n$. Les coefficients $M(j, i) = G(i, j)$ décrivent la transition de j vers i .

I.B.2. Question adaptée au programme actuel. Le mécanisme de transition est précisé dans le sujet pour éviter tout conditionnement par un événement de probabilité nulle.

Si $P(X_{n-1} = j) > 0$, la formule des probabilités composées donne

$$P([X_n = i] \cap [X_{n-1} = j]) = P(X_n = i \mid X_{n-1} = j)P(X_{n-1} = j) = G(i, j)(V_{n-1})_j.$$

Si $P(X_{n-1} = j) = 0$, l'intersection est incluse dans un événement de probabilité nulle ; sa probabilité vaut donc 0. Le produit $G(i, j)(V_{n-1})_j$ vaut lui aussi 0, sans utiliser de probabilité conditionnelle.

Dans tous les cas,

$$P([X_n = i] \cap [X_{n-1} = j]) = G(i, j)(V_{n-1})_j.$$

I.B.3. Pour chaque i , on décompose $[X_n = i]$ selon le système complet $([X_{n-1} = j])_{1 \leq j \leq N}$. Les intersections sont disjointes; d'après I.B.2,

$$P(X_n = i) = \sum_{j=1}^N P([X_n = i] \cap [X_{n-1} = j]) = \sum_{j=1}^N G(i, j)(V_{n-1})_j.$$

Cette dernière expression est la composante i de GV_{n-1} . Donc $V_n = GV_{n-1}$. En convention ligne, cette relation s'écrit $W_n = W_{n-1}^t G$.

I.B.4. On raisonne par récurrence sur $k \in \mathbb{N}$. Au rang 0, $G^0 V_0 = I_N V_0 = V_0$. Si $V_k = G^k V_0$, la relation de I.B.3 au rang $k + 1$ donne

$$V_{k+1} = GV_k = GG^k V_0 = G^{k+1} V_0.$$

Ainsi $V_k = G^k V_0$ pour tout entier $k \geq 0$.

Partie II — Matrices stochastiques

A. Étude d'un exemple

II.A.1. Écrivons $V = \begin{pmatrix} x \\ y \end{pmatrix}$. L'égalité $QV = V$ équivaut au système

$$-qx + q'y = 0, \quad qx - q'y = 0.$$

Les deux équations sont opposées. Comme $q' > 0$, la première donne $y = (q/q')x$. En posant $t = x/q'$, on obtient exactement les vecteurs

$$V = t \begin{pmatrix} q' \\ q \end{pmatrix}, \quad t \in \mathbb{R}.$$

II.A.2. Le vecteur précédent est positif ou nul exactement lorsque $t \geq 0$, car $q, q' > 0$. Il est alors de probabilité si et seulement si $t(q + q') = 1$. Comme $q + q' > 0$, il existe un unique choix, $t = 1/(q + q')$. Par conséquent,

$$V_{\infty} = \frac{1}{q + q'} \begin{pmatrix} q' \\ q \end{pmatrix}.$$

Ses deux composantes sont strictement positives et leur somme vaut 1.

II.A.3. Posons $s = q + q' > 0$ et $r = 1 - q - q'$. Notons

$$L = \frac{1}{s} \begin{pmatrix} q' & q' \\ q & q \end{pmatrix}, \quad K = \frac{1}{s} \begin{pmatrix} q & -q' \\ -q & q' \end{pmatrix}.$$

Méthode 1 — Récurrence à partir de deux matrices simples

On vérifie d'abord $L + K = I_2$. En multipliant les matrices, on obtient

$$L^2 = \frac{1}{s^2} \begin{pmatrix} q's & q's \\ qs & qs \end{pmatrix} = L.$$

Comme $K = I_2 - L$, on en déduit $LK = KL = 0$ et $K^2 = K$. Par ailleurs,

$$L + rK = rI_2 + (1 - r)L = rI_2 + sL = \begin{pmatrix} r + q' & q' \\ q & r + q \end{pmatrix} = Q.$$

Pour $n = 1$, on a bien $Q = L + rK$. Si $Q^n = L + r^n K$, alors

$$Q^{n+1} = (L + r^n K)(L + rK) = L^2 + rLK + r^n KL + r^{n+1} K^2 = L + r^{n+1} K.$$

La récurrence établit $Q^n = L + r^n K$ pour tout $n \geq 1$, ce qui est la formule demandée. Le cas $r = 0$ est inclus, sans division par r .

Méthode 2 — Utiliser une base de vecteurs propres

Les deux vecteurs $u = \begin{pmatrix} q' \\ q \end{pmatrix}$ et $w = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ vérifient

$$Qu = u, \quad Qw = rw.$$

Ils sont indépendants : si $au + bw = 0$, l'addition des deux composantes donne $as = 0$,

donc $a = 0$, puis $b = 0$. La matrice $S = \begin{pmatrix} q' & 1 \\ q & -1 \end{pmatrix}$ est donc inversible. Un produit direct

vérifie

$$S^{-1} = \frac{1}{s} \begin{pmatrix} 1 & 1 \\ q & -q' \end{pmatrix}.$$

On a $Q = S \begin{pmatrix} 1 & 0 \\ 0 & r \end{pmatrix} S^{-1}$, d'où, pour $n \geq 1$,

$$Q^n = S \begin{pmatrix} 1 & 0 \\ 0 & r^n \end{pmatrix} S^{-1} = \frac{1}{s} \begin{pmatrix} q' + qr^n & q' - q'r^n \\ q - qr^n & q + q'r^n \end{pmatrix} = L + r^n K.$$

La première méthode exploite directement la forme de la réponse et une récurrence ; la seconde explique les deux termes de cette forme par les deux directions propres. Aucun calcul de déterminant pour chercher les valeurs propres n'est nécessaire.

II.A.4. Comme $0 < q, q' < 1$, on a $0 < q + q' < 2$, donc $-1 < r < 1$. La suite géométrique (r^n) tend vers 0. La formule de II.A.3 donne ainsi, coefficient par coefficient,

$$Q^n \rightarrow L = \frac{1}{q + q'} \begin{pmatrix} q' & q' \\ q & q \end{pmatrix}.$$

Les deux colonnes de cette limite sont exactement V_∞ .

B. Existence d'un vecteur de probabilité invariant

Dans cette section, seule la stochasticité de Q est supposée ; ses coefficients peuvent être nuls.

II.B.1. Pour chaque j , la composante j de ${}^tQU$ est

$$({}^tQU)_j = \sum_{i=1}^N Q(i, j) \times 1 = 1,$$

puisque la colonne j de Q a pour somme 1. Par conséquent, ${}^tQU = U$.

II.B.2. Supposons $Q - I_N$ inversible et notons $B = (Q - I_N)^{-1}$. Les deux identités $B(Q - I_N) = I_N$ et $(Q - I_N)B = I_N$, après transposition, donnent respectivement

$$({}^tQ - I_N){}^tB = I_N \quad \text{et} \quad {}^tB({}^tQ - I_N) = I_N.$$

Ainsi ${}^tQ - I_N$ est inversible, d'inverse ${}^tB$. La preuve utilise uniquement la transposition d'un produit.

II.B.3. Le vecteur U est non nul et $({}^tQ - I_N)U = 0$ d'après II.B.1. Donc ${}^tQ - I_N$ n'est pas inversible. Par la contraposée de II.B.2, $Q - I_N$ n'est pas inversible non plus.

Pour une matrice carrée, la non-inversibilité équivaut à l'existence d'un vecteur non nul dans son noyau. Il existe donc $V \neq 0$ tel que $(Q - I_N)V = 0$, soit $QV = V$. Cela prouve que 1 est valeur propre de Q .

II.B.4. On utilise maintenant le vecteur propre et la valeur propre λ donnés dans le sujet. Pour chaque ligne i , l'inégalité triangulaire et la positivité des coefficients de Q donnent

$$|(QV)_i| = \left| \sum_{j=1}^N Q(i, j)v_j \right| \leq \sum_{j=1}^N Q(i, j)|v_j| = (Q|V|)_i.$$

Or $QV = \lambda V$ et $|\lambda| = 1$, donc $|(QV)_i| = |v_i|$. Ainsi $(Q|V| - |V|)_i \geq 0$ pour chaque i . Le vecteur demandé est positif.

II.B.5. Posons $W = Q|V| - |V|$. Il est positif d'après II.B.4. On somme ses composantes, en intervertissant seulement deux sommes finies :

$$\sum_{i=1}^N w_i = \sum_{i=1}^N \sum_{j=1}^N Q(i, j)|v_j| - \sum_{i=1}^N |v_i| = \sum_{j=1}^N |v_j| \left(\sum_{i=1}^N Q(i, j) \right) - \sum_{i=1}^N |v_i| = 0.$$

Une famille finie de réels positifs ou nuls ayant une somme nulle est composée de zéros. Donc $W = 0$, ce qui signifie $Q|V| = |V|$.

II.B.6. D'après II.B.3, choisissons un vecteur propre $E \neq 0$ associé à la valeur propre 1. Les questions II.B.4 et II.B.5 s'appliquent avec $\lambda = 1$ et donnent $Q|E| = |E|$.

Comme $E \neq 0$, on a $\|E\|_1 > 0$. Le vecteur

$$V_\infty = \frac{|E|}{\|E\|_1}$$

est positif, sa somme de composantes vaut 1 et $QV_\infty = V_\infty$. C'est donc un vecteur de probabilité invariant. L'existence est établie pour toute matrice stochastique du sujet, même lorsqu'elle n'est pas strictement positive.

C. Unicité d'un vecteur de probabilité invariant

Dans cette section, tous les coefficients de Q sont strictement positifs.

II.C.1. Soit V positif tel que $QV = V$. Si $V = 0$, la première possibilité de l'énoncé est vérifiée. Sinon, une composante, notée v_{j_0} , est strictement positive. Pour tout i ,

$$v_i = (QV)_i = \sum_{j=1}^N Q(i, j)v_j \geq Q(i, j_0)v_{j_0} > 0.$$

Le vecteur V est donc strictement positif.

II.C.2. Le vecteur V_∞ est positif, invariant et non nul, car ses composantes sont de somme 1. La question II.C.1 entraîne que toutes ses composantes sont strictement positives.

II.C.3. Les vecteurs W_∞ et V_∞ étant invariants, la linéarité du produit matriciel donne

$$QV = Q(W_\infty - \alpha V_\infty) = QW_\infty - \alpha QV_\infty = W_\infty - \alpha V_\infty = V.$$

Le vecteur V est invariant par Q .

II.C.4. Les dénominateurs $(v_\infty)_i$ sont strictement positifs d'après II.C.2.

La définition du minimum donne

$$\alpha \leq \frac{(w_\infty)_i}{(v_\infty)_i},$$

donc $v_i = (w_\infty)_i - \alpha(v_\infty)_i \geq 0$ pour chaque i . Le minimum d'une famille finie non vide est atteint : pour l'indice i_0 du sujet, on a $v_{i_0} = 0$. Le vecteur V est donc positif, mais il n'est pas strictement positif.

II.C.5. Le vecteur V est positif et invariant, mais pas strictement positif. La dichotomie démontrée en II.C.1 impose alors $V = 0$. Par définition de V , cela donne $W_\infty = \alpha V_\infty$.

II.C.6. Les deux vecteurs sont de probabilité. En sommant les composantes de $W_\infty = \alpha V_\infty$, on obtient $1 = \alpha \times 1$, donc $\alpha = 1$. Par conséquent, $W_\infty = V_\infty$: le vecteur de probabilité invariant est unique.

II.C.7. D'après I.A.4, la matrice G est stochastique strictement positive. La question II.B.6 donne l'existence d'un vecteur de probabilité invariant par G ; II.C.6 en assure l'unicité. Or le système (S) , avec $p(j) \geq 0$, exprime exactement qu'un vecteur de composantes $p(j)$ est de probabilité et invariant par G . Il admet donc une et une seule solution positive ou nulle.

II.C.8. On utilise l'équation de la composante i dans (S) et la définition (D) :

$$p(i) = \sum_{j=1}^N \left(\rho A(i, j) + \frac{1-\rho}{N} \right) p(j) = \rho \sum_{j=1}^N A(i, j) p(j) + \frac{1-\rho}{N} \sum_{j=1}^N p(j).$$

Le premier terme est positif ou nul et la dernière somme vaut 1. Il en résulte

$$p(i) \geq \frac{1-\rho}{N} > 0.$$

Chaque page reçoit ainsi une probabilité strictement positive, même si aucun lien provenant d'une autre page ne pointe vers elle.

II.C.9. La valeur $\rho = 1$ est étudiée ici comme un cas limite distinct du domaine $[0, 1[$

utilisé précédemment. Elle supprime le terme positif $(1 - \rho)/N$: on a alors $G = A$, qui n'est plus nécessairement strictement positive.

Prenons $N = 3$ et trois pages ne pointant vers aucune page. Pour chaque j , $d_j = 0$; par la convention du sujet, $A(j, j) = 1$ et $A(i, j) = 0$ si $i \neq j$. Donc $G = A = I_3$.

Le système (S) se réduit à $p(1) + p(2) + p(3) = 1$ avec $p(1), p(2), p(3) \geq 0$, puisque

$$I_3 P_* = P_* \text{ pour tout vecteur. Par exemple, } \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ et } \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ sont deux solutions distinctes ; il}$$

y en a même une infinité.

L'unicité n'est donc plus garantie lorsque $\rho = 1$. Ce contre-exemple n'affirme pas qu'elle échoue pour tous les réseaux à cette valeur ; il montre pourquoi l'hypothèse $\rho < 1$ assure une garantie valable pour tout réseau du modèle.

Partie III — Validation du PageRank

A. Valeurs propres de Q

III.A.1. Pour tout V , on applique l'inégalité triangulaire sur chaque ligne puis la stochasticté des colonnes :

$$\|QV\|_1 = \sum_{i=1}^N \left| \sum_{j=1}^N Q(i, j)v_j \right| \leq \sum_{i=1}^N \sum_{j=1}^N Q(i, j)|v_j| = \sum_{j=1}^N |v_j| \sum_{i=1}^N Q(i, j) = \|V\|_1.$$

Si V est positif, chaque $Q(i, j)v_j$ est positif ou nul. L'inégalité triangulaire ci-dessus est alors une égalité sur chaque ligne ; donc $\|QV\|_1 = \|V\|_1$.

Cette preuve n'utilise que la positivité au sens large de Q et les sommes de ses colonnes, et s'applique donc aussi à une matrice stochastique qui possède des coefficients nuls.

III.A.2. Soit $\lambda \in \mathbb{R}$ une valeur propre de Q , et soit $V \neq 0$ un vecteur propre associé. Alors

$$\|QV\|_1 = \|\lambda V\|_1 = |\lambda| \|V\|_1.$$

D'après III.A.1, $|\lambda| \|V\|_1 \leq \|V\|_1$. Comme $V \neq 0$, la quantité $\|V\|_1$ est strictement positive et l'on peut diviser : $|\lambda| \leq 1$.

III.A.3. On se place sous les hypothèses précisées après III.A.2 : $QV = \lambda V$, $|\lambda| = 1$ et $\|V\|_1 = 1$. D'après II.B.5, $Q|V| = |V|$; de plus $|V|$ est de probabilité. L'unicité démontrée en II.C donne $|V| = V_\infty$.

En prenant la première composante dans les deux égalités $QV = \lambda V$ et $Q|V| = |V|$, on obtient

$$\left| \sum_{j=1}^N Q(1, j) v_j \right| = |\lambda v_1| = |v_1| = \sum_{j=1}^N Q(1, j) |v_j|.$$

C'est l'identité demandée.

III.A.4. Appliquons P3 aux N réels $a_j = Q(1, j) v_j$. Comme $Q(1, j) > 0$, on a $|a_j| = Q(1, j) |v_j|$. L'identité de III.A.3 est donc exactement l'hypothèse (E) pour la famille (a_j) .

Par P3, tous les a_j sont positifs ou nuls, ou tous négatifs ou nuls. Les facteurs $Q(1, j)$ étant strictement positifs, tous les v_j ont le même signe au sens large. Par conséquent, $V = |V|$ ou $V = -|V|$; dans les deux cas, V est colinéaire à $|V|$.

Puisque $Q|V| = |V|$, ces deux possibilités donnent $QV = V$. Or $QV = \lambda V$ et $V \neq 0$: ainsi $\lambda = 1$.

La preuve donne en outre, après normalisation, $V = V_\infty$ ou $V = -V_\infty$. Tout vecteur propre non nul associé à 1 est donc un multiple de V_∞ . L'espace propre associé à 1 est de dimension 1.

B. Convergence

Dans cette section seulement, Q est supposée diagonalisable sur $\mathbb{R}$.

III.B.1. Par définition de la diagonalisation, il existe une matrice inversible $S \in \mathcal{M}_N(\mathbb{R})$ et une matrice diagonale réelle D telles que $Q = SDS^{-1}$. En multipliant deux fois, les facteurs $S^{-1}S$ se simplifient : $Q^2 = SD^2S^{-1}$.

Plus généralement, une récurrence donne $Q^n = SD^nS^{-1}$ pour tout $n \geq 1$. En effet, si la relation est vraie au rang n , alors $Q^{n+1} = SD^nS^{-1}SDS^{-1} = SD^{n+1}S^{-1}$.

III.B.2. Les coefficients diagonaux de D sont les valeurs propres réelles correspondant aux colonnes de S , qui forment une base de vecteurs propres. La valeur propre 1 existe d'après II.B.3 : au moins un coefficient diagonal vaut donc 1.

Il ne peut pas y en avoir deux. Sinon, les deux colonnes correspondantes de S seraient deux vecteurs propres indépendants associés à 1, alors que III.A.4 montre que tous ces vecteurs sont colinéaires à V_∞ .

Il y a ainsi exactement un coefficient diagonal égal à 1. Tout autre coefficient λ vérifie $|\lambda| \leq 1$ d'après III.A.2 et ne peut vérifier $|\lambda| = 1$ sans valoir 1, d'après III.A.4. Pour tous les autres coefficients, on a donc $|\lambda| < 1$.

III.B.3. Quitte à réordonner les colonnes de S et les coefficients correspondants de D , écrivons $D = \text{diag}(1, \lambda_2, \dots, \lambda_N)$ avec $|\lambda_j| < 1$ pour $j \geq 2$. Alors

$$D^n = \text{diag}(1, \lambda_2^n, \dots, \lambda_N^n) \longrightarrow D_\infty = \text{diag}(1, 0, \dots, 0).$$

Chaque coefficient de SD^nS^{-1} est une somme finie de termes $S(i, k)\lambda_k^n S^{-1}(k, j)$. On peut donc passer à la limite dans cette somme finie et obtenir

$$Q^n \longrightarrow Q_\infty = SD_\infty S^{-1}.$$

La convergence est bien celle, coefficient par coefficient, définie dans le sujet.

III.B.4. Montrons d'abord que le produit de deux matrices stochastiques B et C est

stochastique. Ses coefficients sont positifs ou nuls. Pour chaque colonne j ,

$$\sum_{i=1}^N (BC)(i, j) = \sum_{i=1}^N \sum_{k=1}^N B(i, k)C(k, j) = \sum_{k=1}^N C(k, j) \sum_{i=1}^N B(i, k) = \sum_{k=1}^N C(k, j) = 1.$$

Puisque Q est stochastique, une récurrence sur n montre que toutes ses puissances Q^n , $n \geq 1$, le sont.

Chaque coefficient de Q_∞ est la limite de coefficients positifs ou nuls, donc est positif ou nul. Enfin, pour chaque j , la limite d'une somme finie donne

$$\sum_{i=1}^N Q_\infty(i, j) = \lim_{n \rightarrow +\infty} \sum_{i=1}^N Q^n(i, j) = 1.$$

La matrice Q_∞ est donc stochastique.

III.B.5. Pour chaque (i, j) , le coefficient

$$(QQ^n)(i, j) = \sum_{k=1}^N Q(i, k)Q^n(k, j)$$

tend vers

$$\sum_{k=1}^N Q(i, k)Q_\infty(k, j) = (QQ_\infty)(i, j),$$

par passage à la limite dans une somme finie.

Or $QQ^n = Q^{n+1}$, et la suite décalée (Q^{n+1}) a la même limite Q_∞ que (Q^n) . Par unicité des limites coefficient par coefficient, $QQ_\infty = Q_\infty$.

III.B.6. Notons C_j la colonne j de Q_∞ . La colonne j du produit QQ_∞ est QC_j . L'égalité de III.B.5 donne donc $QC_j = C_j$ pour chaque j : chaque colonne de Q_∞ est invariante par Q .

III.B.7. D'après III.B.4, chaque colonne C_j de Q_∞ est un vecteur de probabilité. D'après III.B.6, elle est invariante par Q . L'unicité de II.C.6 impose donc $C_j = V_\infty$ pour tout j .

Sous l'hypothèse de diagonalisation de cette section, on a ainsi établi

$$Q^n \rightarrow Q_\infty = \begin{pmatrix} (v_\infty)_1 & \cdots & (v_\infty)_1 \\ \vdots & & \vdots \\ (v_\infty)_N & \cdots & (v_\infty)_N \end{pmatrix}.$$

Résultat admis avant la partie III.C

Résultat admis dans cette version adaptée. Pour la matrice G du modèle, sans supposer sa diagonalisation, la suite $(G^n)_{n \geq 1}$ converge vers la matrice G_∞ dont chaque colonne

est $P_* = \begin{pmatrix} p(1) \\ \vdots \\ p(N) \end{pmatrix}$, l'unique solution de (S) sous forme de vecteur de probabilité.

Question adaptée au programme actuel. L'admission générale du texte original est restreinte à cette seule conséquence pour G , utilisée en III.C.1. Le renvoi à la partie « II.C » est rectifié en « III.C ». La seconde méthode de III.C.1 établit aussi directement la conséquence nécessaire pour le modèle, avec une majoration élémentaire.

C. Application au modèle du surfeur

III.C.1. Méthode 1 — Utiliser la limite des puissances admise pour G

Par I.B.4, $V_n = G^n V_0$. Notons $v_{0,j} = P(X_0 = j)$; ces nombres sont positifs ou nuls et de somme 1. Pour chaque i , l'admission précédente et la finitude de la somme donnent

$$P(X_n = i) = \sum_{j=1}^N G^n(i, j) v_{0,j} \longrightarrow \sum_{j=1}^N p(i) v_{0,j} = p(i) = P(X_\infty = i).$$

Les variables X_n et X_∞ sont à valeurs entières dans le même ensemble fini. La caractérisation de la convergence en loi par les probabilités ponctuelles pour des variables à valeurs dans $\mathbb{Z}$ s'applique : X_n converge en loi vers X_∞ .

On peut également lire cette conclusion sur les fonctions de répartition : pour tout réel x ,

$$P(X_n \leq x) = \sum_{1 \leq i \leq N, i \leq x} P(X_n = i)$$

tend vers la somme analogue pour X_∞ . En particulier, la convergence est vraie aux points de continuité de sa fonction de répartition.

Méthode 2 — Majorer directement l'écart à la distribution invariante

La matrice A est stochastique : si $d_j = 0$, sa colonne j a un seul coefficient égal à 1 ; si $d_j > 0$, elle a d_j coefficients égaux à $1/d_j$. Ses coefficients sont positifs ou nuls. La preuve de III.A.1, qui ne demande pas la stricte positivité, donne donc $\|AZ\|_1 \leq \|Z\|_1$ pour tout vecteur colonne réel Z .

Notons J la matrice dont tous les coefficients valent 1. La définition de G s'écrit $G = \rho A + (1 - \rho)J/N$. Si

$$\sum_j z_j = 0,$$

chaque composante de JZ est nulle. On a alors

$$GZ = \rho AZ \quad \text{et} \quad \|GZ\|_1 \leq \rho \|Z\|_1.$$

Le vecteur $Z_n = V_n - P_*$ a une somme de composantes égale à $1 - 1 = 0$. De plus, $GP_* = P_*$ et $V_{n+1} = GV_n$, donc $Z_{n+1} = GZ_n$. Si $0 < \rho < 1$, une récurrence donne

$$\|V_n - P_*\|_1 \leq \rho^n \|V_0 - P_*\|_1 \leq 2\rho^n,$$

la dernière inégalité venant de

$$\sum_j |v_{0,j} - p(j)| \leq \sum_j v_{0,j} + \sum_j p(j) = 2.$$

Par conséquent, pour chaque i ,

$$|P(X_n = i) - p(i)| \leq \|V_n - P_*\|_1 \leq 2\rho^n \rightarrow 0.$$

Si $\rho = 0$, la matrice G a tous ses coefficients égaux à $1/N$. Elle envoie tout vecteur de probabilité sur le vecteur uniforme. Ainsi P_* est ce vecteur uniforme et $V_n = P_*$ pour tout $n \geq 1$; la conclusion est encore vraie.

Dans tous les cas $0 \leq \rho < 1$, la convergence des probabilités ponctuelles entraîne la convergence en loi vers X_∞ .

La première méthode utilise directement le résultat matriciel fourni par le sujet. La seconde exploite la forme particulière de la matrice de Google : elle dispense de toute hypothèse de diagonalisation et fournit une borne explicite sur l'erreur à un rang donné.

III.C.2. Pour chaque page i et toute distribution initiale V_0 , la question précédente donne

$$\lim_{n \rightarrow +\infty} P(X_n = i) = p(i).$$

Le PageRank de la page i est donc la limite de la probabilité qu'un internaute se trouve sur cette page après un grand nombre d'opérations, dans le modèle de navigation défini par G . La limite ne dépend pas de la première page visitée ni, plus généralement, de sa distribution initiale.

Il s'agit ici d'une probabilité de présence à un instant éloigné. Aucune convergence de la fréquence des visites le long d'une trajectoire individuelle n'a été demandée ni démontrée dans le sujet ; cette interprétation plus forte ne se déduit pas de la seule convergence en loi obtenue.

Relevé des adaptations

I.B, préambule et I.B.1–I.B.4 : la convention colonne de l'annale est conservée et son lien avec la convention ligne du programme est explicité par transposition. Les probabilités de transition sont décrites comme des données du mécanisme ; l'écriture conditionnelle est limitée aux événements de probabilité strictement positive. En I.B.2, les cas de probabilité nulle sont traités directement, sans conditionnement indéfini.

III.B, résultat admis entre III.B.7 et III.C.1 : l'admission concernant toutes les matrices stochastiques strictement positives est remplacée par la seule limite des puissances de la matrice G du modèle, nécessaire à III.C.1. La preuve générale sous hypothèse de diagonalisation, questions III.B.1–III.B.7, est conservée intégralement. Une preuve élémentaire supplémentaire pour G est donnée en III.C.1.

III.B, même résultat admis : le renvoi erroné à la partie « II.C » du document original est rectifié en « III.C », partie qui utilise la convergence pour interpréter le PageRank.