



PRÉPASIMPLE

prepasimple.fr

ESSEC

Mathématiques 2 — Option économique

2008

Mathématiques appliquées

Mathématiques appliquées — Prépa ECG

Annales de concours · Sujet

Tout au long du sujet, N désigne un entier naturel supérieur ou égal à 2.

Notations

Pour $x \in \mathbb{R}$, on note $|x|$ la valeur absolue de x .

Pour tout vecteur

$$V = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} \in \mathcal{M}_{N,1}(\mathbb{R}),$$

on note $|V| = \begin{pmatrix} |v_1| \\ \vdots \\ |v_N| \end{pmatrix}$ et

$$\|V\|_1 = \sum_{j=1}^N |v_j|.$$

On notera I_N la matrice identité de $\mathcal{M}_N(\mathbb{R})$. Pour $A \in \mathcal{M}_N(\mathbb{R})$, on note ${}^tA$ sa matrice transposée.

Résultat admis : pour $(A, B) \in \mathcal{M}_N(\mathbb{R})^2$, on a ${}^t(AB) = {}^tB {}^tA$.

Définitions

Une matrice est dite positive si tous ses coefficients sont des nombres réels positifs ou nuls ; elle est strictement positive si tous ses coefficients sont des nombres réels strictement positifs.

Un vecteur colonne $V \in \mathcal{M}_{N,1}(\mathbb{R})$ est dit de probabilité si V est positif et si $\|V\|_1 = 1$.

Une matrice $Q = (Q(i, j))_{1 \leq i, j \leq N} \in \mathcal{M}_N(\mathbb{R})$ est dite stochastique si elle est positive et si

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

pour tout $1 \leq j \leq N$.

Un vecteur $V \in \mathcal{M}_{N,1}(\mathbb{R})$ est dit invariant par $Q \in \mathcal{M}_N(\mathbb{R})$ si $QV = V$.

Soit $Q \in \mathcal{M}_N(\mathbb{R})$. La suite de matrices $(Q^n)_{n \geq 0}$ est convergente vers la matrice Q_∞ si, pour tout $(i, j) \in \{1, \dots, N\}^2$, la suite $((Q^n)(i, j))_{n \geq 0}$ converge vers $Q_\infty(i, j)$.

On utilise ici la convention des vecteurs colonnes et des matrices dont les colonnes sont de somme 1. Dans la convention des vecteurs lignes, la matrice de transition correspondante est $M = {}^tQ$ et la distribution ligne associée à V est ${}^tV$; la relation $V' = QV$ s'écrit alors ${}^tV' = {}^tV M$.

Préliminaires

Soit

$$V = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} \in \mathcal{M}_{N,1}(\mathbb{R})$$

tel que

$$\left| \sum_{j=1}^N v_j \right| = \sum_{j=1}^N |v_j|. \quad (E)$$

P1. Montrer que, pour tout x réel, $|x| - x \geq 0$.

P2. Étude du cas $N = 3$. On supposera donc, dans cette question P2 uniquement, $N = 3$.

P2.a. Montrer que

$$(|v_1| + |v_2| + |v_3|)^2 - (v_1 + v_2 + v_3)^2 = 2(|v_1 v_2| - v_1 v_2) + 2(|v_1 v_3| - v_1 v_3) + 2(|v_2 v_3| - v_2 v_3).$$

P2.b. Montrer à l'aide de (E) que, si $|v_j v_{j'}| > 0$ pour $(j, j') \in \{1, 2, 3\}^2$ tel que $j \neq j'$, alors v_j et $v_{j'}$ ont même signe.

P2.c. Conclure que $V = |V|$ ou $V = -|V|$.

P3. Montrer que $V = |V|$ ou $V = -|V|$ dans le cas général où N est un entier quelconque vérifiant $N \geq 2$.

Partie I — Google et PageRank

En 1998, Sergey Brin et Larry Page, cofondateurs de Google, ont introduit la notion de PageRank, un indice mesurant la notoriété de chacune des pages Web référencées dans Google. Le principe mathématique étudié ici repose sur le modèle suivant.

On numérote de 1 à N les pages Web référencées dans Google ; on pense que $N = 10^9$ est un bon ordre de grandeur dans le contexte du modèle. On dira qu'une page $j \in \{1, \dots, N\}$ pointe vers une autre page $i \in \{1, \dots, N\}$ s'il existe un lien dans la page j permettant de rejoindre la page i en cliquant dessus.

Pour tout $j \in \{1, \dots, N\}$, on note d_j le nombre de pages vers lesquelles la j -ième page pointe.

Lorsque $d_j = 0$, on pose $A(i, j) = 0$ si $i \neq j$ et $A(j, j) = 1$. Lorsque $d_j > 0$, on pose $A(i, j) = 1/d_j$ si j pointe vers i et $A(i, j) = 0$ sinon.

Si $\rho \in [0, 1[$, on définit la matrice de Google $G = (G(i, j))_{1 \leq i, j \leq N}$ par

$$G(i, j) = \rho A(i, j) + \frac{1 - \rho}{N}, \quad 1 \leq i, j \leq N. \quad (D)$$

Pour tout $j \in \{1, \dots, N\}$, le PageRank de la page j est un nombre réel positif ou nul noté $p(j)$. Les $p(j)$ sont définis par le système d'équations

$$\sum_{j=1}^N p(j) = 1 \quad \text{et} \quad \sum_{j=1}^N G(i, j)p(j) = p(i) \quad \text{pour tout } 1 \leq i \leq N. \quad (S)$$

Le but du sujet est de justifier l'existence et l'unicité de la solution positive ou nulle de (S) et de fournir une interprétation probabiliste du PageRank : la « chance » qu'un surfeur se retrouve sur une page donnée.

A. Étude de la matrice G de Google

On démontre dans cette section quelques propriétés simples de la matrice G .

I.A.1. Montrer que $G \in \mathcal{M}_N(\mathbb{R})$ est une matrice strictement positive.

I.A.2. Soit $j \in \{1, \dots, N\}$ tel que $d_j = 0$. Montrer que

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

I.A.3. Soit $j \in \{1, \dots, N\}$ tel que $d_j > 0$. En écrivant que

$$\sum_{i=1}^N G(i, j) = \sum_{\substack{i \in \{1, \dots, N\} \\ j \text{ pointe vers } i}} G(i, j) + \sum_{\substack{i \in \{1, \dots, N\} \\ j \text{ ne pointe pas vers } i}} G(i, j),$$

montrer que

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

I.A.4. Que peut-on en déduire pour G ?

I.A.5. Montrer que le vecteur $P_* = \begin{pmatrix} p(1) \\ \vdots \\ p(N) \end{pmatrix}$ défini en (S), en admettant

qu'il existe, est invariant par G .

B. Modèle du surfeur sur le Web

Dans toute cette partie, $(\Omega, \mathcal{F}, P)$ désigne un espace probabilisé et toutes les variables aléatoires utilisées sont définies sur cet espace. On rappelle que l'on numérote de 1 à N les pages référencées dans Google.

On considère un internaute surfant sur Google. On note X_0 la première page visitée et X_n la page sur laquelle il se retrouve au bout de n opérations, soit de clic sur un lien dans une page, soit d'abandon au profit d'une autre adresse. Pour tout $n \in \mathbb{N}$, X_n est à valeurs dans $\{1, \dots, N\}$.

On admet que le mécanisme de transition depuis la page j choisit la page i avec la probabilité $G(i, j)$. Lorsque $P(X_{n-1} = j) > 0$, cela s'écrit

$$P(X_n = i \mid X_{n-1} = j) = G(i, j).$$

Les probabilités du mécanisme de transition sont données par le modèle, même si l'événement $[X_{n-1} = j]$ a une probabilité nulle; aucune probabilité conditionnelle par un tel événement n'est définie. On considère un entier naturel n strictement positif.

I.B.1. On note V_n le vecteur de $\mathcal{M}_{N,1}(\mathbb{R})$ dont, pour tout $i \in \{1, \dots, N\}$, la i -ième composante est $P(X_n = i)$. Vérifier que V_n est bien un vecteur de probabilité.

I.B.2. Exprimer, pour tout $(i, j) \in \{1, \dots, N\}^2$, $P([X_n = i] \cap [X_{n-1} = j])$ à l'aide de G et de V_{n-1} .

I.B.3. En déduire que $V_n = GV_{n-1}$.

I.B.4. Montrer que, pour tout k entier naturel, $V_k = G^k V_0$.

Partie II — Matrices stochastiques

Le but de cette deuxième partie est de prouver l'existence et l'unicité d'un vecteur de probabilité invariant pour une matrice stochastique strictement positive.

A. Étude d'un exemple

On considère une matrice $Q \in \mathcal{M}_2(\mathbb{R})$ stochastique et strictement positive. Elle peut se mettre sous la forme

$$Q = \begin{pmatrix} 1-q & q' \\ q & 1-q' \end{pmatrix}, \quad q \in]0, 1[, \quad q' \in]0, 1[.$$

II.A.1. Déterminer l'ensemble des vecteurs $V \in \mathcal{M}_{2,1}(\mathbb{R})$ vérifiant $QV = V$.

II.A.2. Montrer qu'il existe un unique vecteur de probabilité

$V_\infty \in \mathcal{M}_{2,1}(\mathbb{R})$ invariant par Q et le calculer.

II.A.3. Prouver que, pour tout entier $n \geq 1$,

$$Q^n = \frac{1}{q+q'} \begin{pmatrix} q' & q' \\ q & q \end{pmatrix} + \frac{(1-q-q')^n}{q+q'} \begin{pmatrix} q & -q' \\ -q & q' \end{pmatrix}.$$

II.A.4. En déduire que $(Q^n)_{n \geq 1}$ converge, lorsque n tend vers l'infini, vers une matrice dont les deux vecteurs colonnes sont égaux à V_∞ . On cherchera à généraliser ce résultat dans la partie III.

B. Existence d'un vecteur de probabilité invariant

Dans cette section II.B, $Q \in \mathcal{M}_N(\mathbb{R})$ est une matrice stochastique.

II.B.1. On note U le vecteur de $\mathcal{M}_{N,1}(\mathbb{R})$ dont toutes les composantes valent 1. Calculer ${}^tQU$.

II.B.2. Montrer que, si $Q - I_N$ est inversible, alors ${}^tQ - I_N$ est aussi inversible.

II.B.3. Dédurre des deux questions précédentes que 1 est valeur propre de Q .

Soit λ une valeur propre réelle de Q telle que $|\lambda| = 1$ et $V \in \mathcal{M}_{N,1}(\mathbb{R})$ un vecteur propre de Q associé à λ .

II.B.4. Prouver que le vecteur $Q|V| - |V|$ est positif.

II.B.5. Montrer que le vecteur $|V|$ est invariant par Q .

Indication : on pourra sommer les composantes du vecteur $Q|V| - |V|$.

II.B.6. Prouver l'existence d'au moins un vecteur de probabilité invariant par Q .

C. Unicité d'un vecteur de probabilité invariant

Dans cette section II.C, $Q \in \mathcal{M}_N(\mathbb{R})$ est une matrice stochastique strictement positive. On sait d'après II.B.6 qu'il existe au moins un vecteur de probabilité invariant par Q , noté

$$V_\infty = \begin{pmatrix} (v_\infty)_1 \\ \vdots \\ (v_\infty)_N \end{pmatrix}.$$

Cette section permettra de démontrer l'unicité d'un tel vecteur.

II.C.1. Prouver que, si V est un vecteur positif invariant par Q , alors soit $V = 0_{\mathcal{M}_{N,1}(\mathbb{R})}$, soit V est strictement positif.

II.C.2. Justifier que V_∞ est strictement positif.

On considère à présent un autre vecteur de probabilité

$$W_\infty = \begin{pmatrix} (w_\infty)_1 \\ \vdots \\ (w_\infty)_N \end{pmatrix}$$

invariant par Q . Puis on définit

$$\alpha = \min \left\{ \frac{(w_\infty)_i}{(v_\infty)_i} : 1 \leq i \leq N \right\} = \frac{(w_\infty)_{i_0}}{(v_\infty)_{i_0}} \quad \text{et} \quad V = W_\infty - \alpha V_\infty.$$

II.C.3. Montrer que V est invariant par Q .

II.C.4. Montrer que V est positif mais pas strictement positif.

II.C.5. En déduire que $W_\infty = \alpha V_\infty$.

II.C.6. En conclure que $W_\infty = V_\infty$.

On reprend jusqu'à la fin de cette partie les notations de la partie I sur Google et la notion de PageRank.

II.C.7. Montrer que le système (S) définissant le PageRank admet bien une et une seule solution positive ou nulle.

II.C.8. Démontrer que $p(i) \geq (1 - \rho)/N$ pour tout $i \in \{1, \dots, N\}$.

II.C.9. Le rôle du paramètre ρ est essentiel pour assurer l'unicité de la solution du système (S) . Que se passerait-il pour $\rho = 1$ et, disons, $N = 3$ pour simplifier à l'extrême ? Songer qu'il peut exister des pages Web qui ne pointent vers aucune autre.

Partie III — Validation du PageRank

Dans toute cette partie III, on considère $Q \in \mathcal{M}_N(\mathbb{R})$ stochastique et strictement positive. On notera V_∞ l'unique vecteur de probabilité invariant par Q .

A. Valeurs propres de Q

Il s'agit dans cette section de localiser les valeurs propres réelles de Q .

III.A.1. Vérifier que, pour tout vecteur $V \in \mathcal{M}_{N,1}(\mathbb{R})$, on a $\|QV\|_1 \leq \|V\|_1$, avec de plus $\|QV\|_1 = \|V\|_1$ lorsque V est positif.

III.A.2. En déduire que, pour toute valeur propre réelle λ de Q , on a $|\lambda| \leq 1$.

Soient λ une valeur propre réelle de Q telle que $|\lambda| = 1$ et $V \in \mathcal{M}_{N,1}(\mathbb{R})$ un vecteur propre de Q associé à λ tel que $\|V\|_1 = 1$. On sait grâce à II.B.5 et à l'unicité que $|V| = V_\infty$.

III.A.3. Établir l'identité

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

où v_i est la i -ième composante de V pour $i \in \{1, \dots, N\}$.

III.A.4. En déduire en utilisant P3 que V est colinéaire à $|V|$, puis que $\lambda = 1$.

B. Convergence

On fera dans cette section III.B l'hypothèse supplémentaire que Q est diagonalisable sur $\mathbb{R}$.

Le but de ce qui suit est d'établir que la suite $(Q^n)_{n \geq 1}$ converge, lorsque n tend vers l'infini, vers la matrice Q_∞ dont les vecteurs colonnes sont tous égaux à V_∞ .

III.B.1. Montrer qu'il existe une matrice diagonale $D \in \mathcal{M}_N(\mathbb{R})$ et une matrice inversible $S \in \mathcal{M}_N(\mathbb{R})$ telles que, pour tout $n \geq 1$, $Q^n = SD^nS^{-1}$.

III.B.2. Prouver que, parmi les N coefficients diagonaux de D , un et un seul est égal à 1, alors que les autres sont de valeur absolue strictement inférieure à 1.

III.B.3. En déduire que la suite $(Q^n)_{n \geq 1}$ converge vers une matrice Q_∞ .

III.B.4. Prouver que, pour tout $n \geq 1$, Q^n est stochastique, puis que Q_∞ est stochastique.

III.B.5. Démontrer que $QQ_\infty = Q_\infty$.

III.B.6. En déduire que chacun des vecteurs colonnes de Q_∞ est invariant par Q .

III.B.7. Conclure.

Pour la partie III.C, on admet le résultat particulier suivant, sans hypothèse de diagonalisation de G : la suite $(G^n)_{n \geq 1}$ converge vers la matrice G_∞ dont toutes les

colonnes sont égales au vecteur $\begin{pmatrix} p(1) \\ \vdots \\ p(N) \end{pmatrix}$ défini par (S) .

C. Application au modèle du surfeur

On reprend pour la fin du sujet les notations du PageRank de Google décrit dans l'introduction de la partie I et du modèle du surfeur décrit dans la partie I.B. On considère X_∞ , variable aléatoire à valeurs dans $\{1, \dots, N\}$ dont la loi est définie par

$$P(X_\infty = i) = p(i), \quad i \in \{1, \dots, N\}.$$

III.C.1. Montrer que $(X_n)_{n \geq 0}$ converge en loi vers X_∞ lorsque n tend vers l'infini.

III.C.2. En quoi ce résultat donne-t-il du sens à l'assertion un peu vague : « Le PageRank d'une page donnée représente la chance qu'un internaute se retrouve sur la page en question lorsqu'il surfe » ?