Un corrigé de l'épreuve ENS 2015, filière BCPST

Comme pour les sessions précédentes, on peut observer quelques imprécisions dans l'énoncé proposé aux candidats.

- *L'ensemble $\mathcal{M}_n(\mathbb{R})$ des matrices carrées réelles d'ordre n est rebaptisé, de façon impropre, « l'ensemble des matrices carrées réelles de dimension n par n ». De même, « de taille n », deux lignes plus bas, est une expression qu'il est curieux de trouver à l'écrit.*
- *En (1.2), les notations z_1 et z_2 ne sont pas définies ; il faut comprendre que l'on conserve celles de la question (1.1)*
- *La définition du rayon spectral, après la question (2.2), pose problème. En effet, le théorème garantissant l'existence d'une valeur propre complexe pour n'importe quelle matrice carrée à coefficients réels (ou complexes) ne figure plus explicitement au programme...*
- *En (3.1), il est écrit « valeur absolue » au lieu de « module ».*
- *En (3.3) et (3.4), la propriété à démontrer n'est pas vraie pour tout entier naturel N, mais pour tout entier naturel N non nul.*
- *Dans les questions (4.2) à (4.5), l'énoncé a « oublié » de préciser qu'on supposait la matrice S stochastique...*
- *On ne sait pas qui est i dans les questions (5.1) et (5.2). Il n'est pas précisé que l'entier N doit ici être non nul.*
- *Dans la partie 5, l'énoncé ne précise pas que les mutations ont lieu indépendamment du numéro du parent d'où provient un allèle donné.*
- *En (5.7), « la » matrice B est impropre ; pour certains choix du vecteur π^0, plusieurs matrices B peuvent convenir.*
- *L'espérance conditionnelle, qui apparaît à la question (5.9), est une notion qui ne figure pas au programme de BCPST.*

On peut aussi relever (liste non exhaustive !) quelques points qui relèvent davantage du détail :

- *La première ligne de l'énoncé fait allusion à un « examen » au lieu d'un concours.*
- *La lettre n n'est pas définie.*
- *Il n'est pas d'usage de commencer une phrase par un symbole $|z|$ comme dans « $|z|$ est le module de z ».*
- *Il n'est pas d'usage de commencer une phrase par le symbole $\forall$ (« $\forall A \in \mathcal{M}_n(\mathbb{R})$, ${}^tA$ est... »).*
- *Une coquille qui a échappé à la relecture : « le i-ème composante de v »*
- *En (1.5), on ne comprend pas le sens de la virgule.*

1. Résultats préliminaires 1 : matrices strictement positives

(1.1) Pour tout nombre complexe z, on a par définition du module :

$$|z|^2 = \Re(z)^2 + \Im(z)^2.$$

Soient z_1 et z_2 deux nombres complexes. Écrivons cette identité pour le nombre complexe $z_1\overline{z_2}$:

$$|z_1\overline{z_2}|^2 = \Re(z_1\overline{z_2})^2 + \Im(z_1\overline{z_2})^2.$$

Or le module d'un produit est égal au produit des modules, et un nombre complexe et son conjugué ont même module, d'où l'identité cherchée :

$$\boxed{|z_1|^2|z_2|^2 = \Re(z_1\overline{z_2})^2 + \Im(z_1\overline{z_2})^2}.$$

(1.2) Le résultat de la question (1.1) implique :

$$|z_1|^2|z_2|^2 \geqslant \Re(z_1\overline{z_2})^2$$

donc en particulier

$$|z_1||z_2| \geqslant \Re(z_1\overline{z_2}).$$

La partie réelle d'une différence de deux nombres complexes est égale à la différence des parties réelles. On en déduit, pour tous nombres complexes z_1 et z_2 :

$$\Re(|z_1||z_2| - z_1\overline{z_2}) = \Re(|z_1||z_2|) - \Re(z_1\overline{z_2}) = |z_1||z_2| - \Re(z_1\overline{z_2}) \geqslant 0.$$

Examinons le cas d'égalité. On a, en utilisant l'identité de (1.1) pour écrire la deuxième équivalence :

$$\begin{aligned}
\Re(|z_1||z_2| - z_1\overline{z_2}) = 0 \quad &\Longleftrightarrow \quad |z_1||z_2| = \Re(z_1\overline{z_2}) \\
&\Longleftrightarrow \quad \left\{ \begin{array}{l} |z_1||z_2| = \Re(z_1\overline{z_2}) \\ \Im(z_1\overline{z_2}) = 0 \end{array} \right. \\
&\Longleftrightarrow \quad \left\{ \begin{array}{l} |z_1||z_2| = \Re(z_1\overline{z_2}) \\ z_1\overline{z_2} = \Re(z_1\overline{z_2}) \end{array} \right. \\
&\Longleftrightarrow \quad |z_1||z_2| = z_1\overline{z_2}.
\end{aligned}$$

(1.3) Le nombre complexe $\sum_{i=1}^n \sum_{j=1}^n B_{i,j}(|v_i||v_j| - v_i\overline{v_j})$ est nul par hypothèse, donc sa partie réelle est nulle. On utilise le fait que la partie réelle d'une somme de nombres complexes est égale à la somme des parties réelles. De plus, si z est un nombre complexe et a un nombre réel, on a $\Re(az) = a\Re(z)$. On obtient ainsi :

$$\sum_{i=1}^n \sum_{j=1}^n B_{i,j}\Re(|v_i||v_j| - v_i\overline{v_j}) = 0.$$

D'après le résultat de la question (1.2), chacun des nombres réels $\Re(|v_i||v_j| - v_i\overline{v_j})$ est positif ; par hypothèse, les nombres $B_{i,j}$ sont tous positifs. Or une somme de réels positifs est nulle si et seulement si tous ces réels positifs sont nuls. Ainsi, on a, pour tout couple (i,j) d'entiers compris entre 1 et n :

$$\Re(|v_i||v_j| - v_i\overline{v_j}) = 0.$$

D'après le cas d'égalité étudié à la question (1.2), on obtient, pour tout couple (i,j) d'entiers compris entre 1 et n :

$$\boxed{|v_i||v_j| = v_i\overline{v_j}}.$$

(1.4) Démontrer l'inégalité $Au < Av$ équivaut à démontrer l'inégalité $A(v-u) > 0$. Pour tout entier i compris entre 1 et n, notons $(A(v-u))_i$ la $i$ème composante du vecteur $A(v-u)$. La définition du produit matriciel fournit, pour tout entier i compris entre 1 et n :

$$(A(v-u))_i = \sum_{j=1}^{n} A_{i,j}(v_j - u_j).$$

Par hypothèse, chacun des réels $A_{i,j}$ est strictement positif, chacun des réels $v_j - u_j$ est positif, et l'un au moins des réels $v_j - u_j$ est strictement positif (puisque l'on a supposé $u \leqslant v$ et u et v distincts). Ainsi, la somme ci-dessus est formée de réels tous positifs dont l'un au moins est strictement positif. Par conséquent, cette somme est strictement positive. On a donc bien démontré :

$$\boxed{Au < Av}.$$

(1.5) Soit i un entier compris entre 1 et n.

- Observons que, si le réel $(Au)_i$ est négatif ou nul, tout réel strictement positif ε vérifie l'inégalité
$$((1+\varepsilon)Au)_i = (1+\varepsilon)(Au)_i \leqslant (Au)_i \leqslant (Av)_i.$$
- Si le réel $(Au)_i$ est strictement positif, on a pour tout réel strictement positif ε :
$$((1+\varepsilon)Au)_i \leqslant (Av)_i \iff \varepsilon \leqslant \frac{(Av)_i}{(Au)_i} - 1.$$

Considérons l'ensemble

$$\mathcal{E} = \left\{ \frac{(Av)_i}{(Au)_i} - 1 \,; \quad i \in \{1, \dots, n\}\,, \, (Au)_i > 0 \right\}.$$

Il suffit de poser :

$$\boxed{\varepsilon = \begin{cases} 1 & \text{si l'ensemble } \mathcal{E} \text{ est vide} \\ \min \mathcal{E} & \text{si l'ensemble } \mathcal{E} \text{ est non vide (il est alors fini).} \end{cases}}$$

Dans tous les cas, le réel ε choisi est strictement positif et vérifie l'inégalité :

$$\boxed{(1+\varepsilon)Au \leqslant Av}.$$

2. Résultats préliminaires 2 : algèbre linéaire

(2.1) Par hypothèse, les droites vectorielles engendrées par les vecteurs v et v' coïncident. En particulier, le vecteur v appartient à la droite vectorielle engendrée par le vecteur v'. Il existe donc un réel m vérifiant :

$$v = mv'.$$

Calculons la somme des composantes de chacun de ces vecteurs :

$$\sum_{i=1}^{n} v_i = m \sum_{i=1}^{n} v'_i.$$

Par hypothèse, les vecteurs v et v' sont des lois de probabilités, donc chacune des sommes $\sum_{i=1}^{n} v_i$ et $\sum_{i=1}^{n} v'_i$ est égale à 1. Il en résulte que le réel m est égal à 1. On conclut donc :

$$\boxed{v = v'}.$$

(2.2) Soient A une matrice de $\mathcal{M}_n(\mathbb{R})$, v un vecteur de $\mathbb{C}^n$, i un entier compris entre 1 et n. On a d'après l'inégalité triangulaire :

$$\left| \sum_{j=1}^{n} A_{i,j} v_j \right| \leqslant \sum_{j=1}^{n} |A_{i,j} v_j| = \sum_{j=1}^{n} |A_{i,j}||v_j|.$$

On peut majorer tous les réels $|v_j|$ par leur maximum $\|v\|_\infty$ d'où, par positivité des réels $|A_{i,j}|$:

$$\left| \sum_{j=1}^{n} A_{i,j} v_j \right| \leqslant \|v\|_\infty \sum_{j=1}^{n} |A_{i,j}|.$$

On peut majorer chacune des sommes $\sum_{j=1}^{n} |A_{i,j}|$ par la norme infinie de la matrice A, telle qu'elle a été définie dans l'énoncé ; on obtient donc :

$$\boxed{\left| \sum_{j=1}^{n} A_{i,j} v_j \right| \leqslant \|A\|_\infty \, \|v\|_\infty}.$$

(2.3) Soit $\lambda \in \mathbb{C}$ une valeur propre de la matrice A. Soit $v \in \mathbb{C}^n$ un vecteur propre associé. En particulier, le vecteur v et non nul, et il vérifie $Av = \lambda v$, c'est-à-dire, par définition du produit matriciel :

$$\forall i \in \{1, \dots, n\}\,, \quad \sum_{j=1}^{n} A_{i,j} v_j = \lambda v_i.$$

En utilisant l'inégalité de la question (2.2), on obtient :

$$\forall i \in \{1, \dots, n\}\,, \quad |\lambda| \, |v_i| \leqslant \|A\|_\infty \, \|v\|_\infty.$$

Par définition de la norme infinie $\|v\|_\infty$, il existe un entier i_0 compris entre 1 et n (pas nécessairement unique) vérifiant $|v_{i_0}| = \|v\|_\infty$. L'inégalité précédente écrite pour cet entier i_0 devient :

$$|\lambda| \, |v_{i_0}| = |\lambda| \, \|v\|_\infty \leqslant \|A\|_\infty \, \|v\|_\infty.$$

Comme v est un vecteur propre, il est en particulier non nul, donc sa norme infinie $\|v\|_\infty$ est strictement positive. L'inégalité précédente équivaut donc à :

$$|\lambda| \leqslant \|A\|_\infty.$$

On vient de démontrer que le module de toute valeur propre λ de la matrice A est inférieur ou égal à la norme infinie $\|A\|_\infty$. Or, par définition du rayon spectral $\rho(A)$, il existe une valeur propre λ de A vérifiant $|\lambda| = \rho(A)$. On a donc bien démontré l'inégalité :

$$\boxed{\rho(A) \leqslant \|A\|_\infty}.$$

(2.4) Notons $B_{i,j}$ (respectivement $C_{i,j}$, $(BC)_{i,j}$) les coefficients de la matrice B (resp. C, BC). Par définition de la norme infinie de la matrice BC, il existe un entier i_0 compris entre 1 et n vérifiant :

$$\|BC\|_\infty = \left|\sum_{j=1}^{n}(BC)_{i_0,j}\right|.$$

Par définition du produit matriciel, on a pour tout couple (i,j) d'entiers compris entre 1 et n :

$$(BC)_{i,j} = \sum_{k=1}^{n} B_{i,k}C_{k,j}.$$

Ainsi, on a :

$$\|BC\|_\infty = \left|\sum_{j=1}^{n}\sum_{k=1}^{n} B_{i_0,k}C_{k,j}\right| = \left|\sum_{k=1}^{n}\sum_{j=1}^{n} B_{i_0,k}C_{k,j}\right| = \left|\sum_{k=1}^{n} B_{i_0,k}\sum_{j=1}^{n} C_{k,j}\right|.$$

D'après l'inégalité triangulaire, on a :

$$\|BC\|_\infty \leqslant \sum_{k=1}^{n}|B_{i_0,k}|\left|\sum_{j=1}^{n} C_{k,j}\right| \leqslant \sum_{k=1}^{n}|B_{i_0,k}|\sum_{j=1}^{n}|C_{k,j}|.$$

Par définition de la norme infinie de la matrice C, on a (en utilisant aussi la positivité des réels $|B_{i_0,k}|$) :

$$\|BC\|_\infty \leqslant \|C\|_\infty \sum_{k=1}^{n}|B_{i_0,k}|.$$

Par définition de la norme infinie de la matrice B, on a (en utilisant aussi la positivité du réel $\|C\|_\infty$) :

$$\boxed{\|BC\|_\infty \leqslant \|B\|_\infty \|C\|_\infty}.$$

(2.4) En utilisant deux fois successivement l'inégalité que l'on vient de démontrer, on obtient :

$$\|S^{-1}MS\|_\infty \leqslant \|S^{-1}M\|_\infty\,\|S\|_\infty \leqslant \|S^{-1}\|_\infty\,\|M\|_\infty\,\|S\|_\infty.$$

autrement dit

$$\boxed{\|S^{-1}MS\|_\infty \leqslant \|S\|_\infty\,\|S^{-1}\|_\infty\,\|M\|_\infty}.$$

D'autre part, comme la matrice S est inversible, elle est en particulier non nulle, donc sa norme infinie est non nulle. De la même manière, la norme infinie de la matrice S^{-1} est non nulle. On peut donc légitimement écrire :

$$\|S^{-1}MS\|_\infty = \frac{1}{\|S^{-1}\|_\infty\|S\|_\infty}\|S\|_\infty\,\|S^{-1}MS\|_\infty\,\|S^{-1}\|_\infty.$$

L'inégalité de la question (2.4) utilisée deux fois successivement implique :

$$\|S\|_\infty\,\|S^{-1}MS\|_\infty\,\|S^{-1}\|_\infty \geqslant \|SS^{-1}MS\|_\infty\,\|S^{-1}\|_\infty \geqslant \|MSS^{-1}\|_\infty = \|M\|_\infty.$$

On vient de démontrer l'inégalité souhaitée :

$$\boxed{\|S^{-1}MS\|_\infty \geqslant (\|S^{-1}\|_\infty\|S\|_\infty)^{-1}\|M\|_\infty}.$$

(2.6) Pour tout entier naturel N, une récurrence immédiate démontre l'identité :

$$A^N = PD^NP^{-1}.$$

D'après les inégalités démontrées à la question (2.5) en posant $S = P^{-1}$, on a pour tout entier naturel N :

$$(\|P^{-1}\|_\infty\|P\|_\infty)^{-1}\|D^N\|_\infty \leqslant \|A^N\|_\infty \leqslant \|P\|_\infty\,\|P^{-1}\|_\infty\,\|D^N\|_\infty.$$

Par croissance de la fonction $x \mapsto x^{\frac{1}{N}}$ sur $\mathbb{R}^+$, on obtient pour tout entier naturel non nul N :

$$(\|P^{-1}\|_\infty\|P\|_\infty)^{-\frac{1}{N}}\|D^N\|_\infty^{\frac{1}{N}} \leqslant \|A^N\|_\infty^{\frac{1}{N}} \leqslant \|P\|_\infty^{\frac{1}{N}}\,\|P^{-1}\|_\infty^{\frac{1}{N}}\,\|D^N\|_\infty^{\frac{1}{N}} \quad (*)$$

Soit a un nombre réel. Rappelons le fait suivant

- si a est nul, la suite de terme général $a^{\frac{1}{N}}$ est évidemment constamment nulle, donc converge vers 0 ;
- si a est strictement positif, on a pour tout entier naturel non nul N :

  $$a^{\frac{1}{N}} = \exp\left(\frac{1}{N}\ln a\right)$$

  donc la suite de terme général $a^{\frac{1}{N}}$ tend vers 1.

Comme la matrice P n'est pas nulle, sa norme infinie n'est pas nulle. Ainsi, la suite de terme général $\|P\|_\infty^{\frac{1}{N}}$ tend vers 1, de même que la suite de terme général $\|P^{-1}\|_\infty^{\frac{1}{N}}$.
Notons $\lambda_1, \dots, \lambda_n$ les coefficients diagonaux de la matrice D, qui sont aussi les valeurs propres de la matrice A (puisque les matrices A et D sont semblables), non nécessairement distinctes. Par définition de la norme infinie d'une matrice donnée par l'énoncé (la maximum de la somme des modules des coefficients de chaque ligne), on a, pour tout entier naturel non nul N,

$$\|D^N\|_\infty = \max(|\lambda_1|^N, \dots, |\lambda_n|^N)$$

d'où, en utilisant la croissance sur $\mathbb{R}^+$ de la fonction $x \mapsto x^{\frac{1}{N}}$:

$$\|D^N\|_\infty^{\frac{1}{N}} = \max(|\lambda_1|, \dots, |\lambda_n|) = \rho(A).$$

Ainsi, chacun des membres extrêmes dans les inégalités $(*)$ admet une limite quand N tend vers l'infini, et ces limites sont identiques, égales à $\rho(A)$.

D'après le théorème d'encadrement, on en déduit que la suite de terme général $\|A^N\|_\infty^{\frac{1}{N}}$ converge, et sa limite est égale à $\rho(A)$.
On a bien démontré le théorème de Gelfand dans le cas particulier d'une matrice A diagonalisable dans $\mathcal{M}_n(\mathbb{R})$:

$$\boxed{\lim_{N\to+\infty}\|A^N\|_\infty^{\frac{1}{N}} = \rho(A)}.$$

Remarquons que la démonstration ci-dessus permet, sans même avoir à l'adapter, de démontrer le théorème de Gelfand dans le cas d'une matrice A de $\mathcal{M}_n(\mathbb{R})$ non diagonalisable dans $\mathcal{M}_n(\mathbb{R})$ mais diagonalisable dans $\mathcal{M}_n(\mathbb{C})$. Resterait alors à traiter le cas d'une matrice A de $\mathcal{M}_n(\mathbb{R})$ non diagonalisable dans $\mathcal{M}_n(\mathbb{C})$...

Remarquons à ce stade que le parti pris de l'énoncé est un peu curieux : il nous propose de démontrer l'identité $\rho(A) = \rho({}^tA)$ seulement dans le cas où la matrice A est diagonalisable, alors qu'une démonstration dans le cas général est immédiate. En effet, il suffit pour cela de voir que les matrices A et ${}^tA$ ont mêmes valeurs propres, ce qui découle par exemple de l'équivalence suivante, valable pour tout nombre complexe λ :

$$\lambda \in Sp(A) \iff rg\,(A - \lambda I_n) < n \iff rg\,({}^t(A - \lambda I_n)) = rg\,({}^tA - \lambda I_n) < n \iff \lambda \in Sp({}^tA).$$

On a simplement utilisé le fait qu'une matrice et sa transposée ont même rang, ce qui est un résultat figurant au programme.
Le fait que les matrices A et ${}^tA$ aient mêmes valeurs propres implique évidemment l'identité $\rho(A) = \rho({}^tA)$.

(2.7) On a :

$$I_n = {}^t(PP^{-1}) = {}^t(P^{-1})\,{}^tP.$$

Ceci prouve que chacune des matrices ${}^t(P^{-1})$ et ${}^tP$ est inversible et qu'elles sont inverses l'une de l'autre. Autrement dit, on a :

$$\boxed{{}^t(P^{-1}) = ({}^tP)^{-1}}.$$

(2.8) La relation $A = PDP^{-1}$ équivaut à :

$${}^tA = {}^t(P^{-1})\,{}^tD\,{}^tP.$$

Une matrice diagonale et sa transposée sont évidemment égales. En utilisant l'identité vue à la question (2.7), on obtient :

$$\boxed{{}^tA = ({}^tP)^{-1}D\,{}^tP.}$$

Les matrices A et ${}^tA$ sont toutes deux semblables à une même matrice D, donc sont semblables. Par conséquent, elles ont les mêmes valeurs propres, donc le même rayon spectral. On en déduit :

$$\boxed{\rho(A) = \rho({}^tA)}.$$

3. Le théorème de Perron-Froebenius

(3.1) *L'énoncé parle de « valeurs absolues », mais il s'agit bien de modules ici : les vecteurs considérés sont à coefficients complexes.*
Soit i un entier compris entre 1 et n. On a, par définition du produit matriciel :

$$(A|\phi|)_i = \sum_{j=1}^{n} A_{i,j}|\phi|_j = \sum_{j=1}^{n} A_{i,j}|\phi_j|.$$

D'après l'inégalité triangulaire, on a :

$$(A|\phi|)_i = \sum_{j=1}^{n} |A_{i,j}\phi_j| \geqslant \left|\sum_{j=1}^{n} A_{i,j}\phi_j\right| = |(A\phi)_i|\,.$$

Comme ϕ est un vecteur propre de la matrice A associé à la valeur propre λ, on a $A\phi = \lambda\phi$, donc $(A\phi)_i = \lambda\phi_i$ pour tout entier i compris entre 1 et n. On a donc, pour tout entier i compris entre 1 et n :

$$(A|\phi|)_i \geqslant |(\lambda\phi)_i| = |\lambda|\,|\phi_i| = \rho(A)\,|\phi_i|.$$

On vient de démontrer l'inégalité entre vecteurs suivante :

$$\boxed{A|\phi| \geqslant \rho(A)|\phi|}.$$

(3.2) L'énoncé a supposé, par l'absurde, que l'on a $v = A|\phi| > u = \rho(A)|\phi|$. La matrice A est strictement positive. Les hypothèses des questions (1.4) et (1.5) sont donc vérifiées. On en déduit qu'il existe un réel strictement positif ε vérifiant :

$$(1+\varepsilon)Au \leqslant Av$$

c'est-à-dire

$$\boxed{(1+\varepsilon)\rho(A)A|\phi| \leqslant A^2|\phi|}.$$

(3.3) *Remarquons une erreur d'énoncé : si l'entier N est nul, l'inégalité souhaitée n'est pas nécessairement vraie. En effet, on a $A\,|\phi| \geqslant \rho(A)|\phi|$ et on a seulement supposé ici $A|\phi|$ différent de $\rho(A)|\phi|$, ce qui n'implique pas l'inégalité stricte $|\phi| > \rho(A)|\phi|$.*
On ne peut donner ici de contre-exemple puisqu'on raisonne par l'absurde : sous les hypothèses de l'énoncé, on aura donc toujours $A|\phi| = \rho(A)|\phi|$.

Raisonnons par récurrence sur l'entier naturel non nul N. Pour tout entier naturel non nul N, posons

$$\mathcal{P}(N) : \quad A^{N+1}|\phi| \geqslant (1+\varepsilon)\rho(A)A^N|\phi|.$$

- Pour $N = 1$, il s'agit de l'inégalité (3.2). La propriété $\mathcal{P}(1)$ est donc vraie.
- Soit N un entier naturel non nul. On suppose que la propriété $\mathcal{P}(N)$ est vraie. On a donc
$$A^{N+1}|\phi| \geqslant (1+\varepsilon)\rho(A)A^N|\phi|.$$
On utilise le fait que, lorsque A est une matrice strictement positive, si u et v dont deux vecteurs de $\mathbb{R}^n$ vérifiant $u \leqslant v$, on a $Au \leqslant Av$ (ce résultat est vrai si u et v sont distincts d'après le résultat de la question (1.4), et il est évident si u et v sont égaux). On applique ce résultat aux vecteurs $u = (1+\varepsilon)\rho(A)A^N|\phi|$ et $v = A^{N+1}|\phi|$. On obtient :
$$A^{N+2}|\phi| \geqslant (1+\varepsilon)\rho(A)A^{N+1}|\phi|.$$
La propriété $\mathcal{P}(N+1)$ est donc vérifiée.

D'après le principe de récurrence, on a donc démontré :

$$\boxed{\forall N \in \mathbb{N}^*, \quad A^{N+1}|\phi| \geqslant (1+\varepsilon)\rho(A)A^N|\phi|}.$$

(3.4) *De même qu'en (3.3), si l'entier N est nul, l'inégalité souhaitée n'est pas nécessairement vraie.*
Montrons ce résultat par récurrence sur l'entier naturel non nul N. Pour tout entier naturel non nul N, posons

$$\mathcal{Q}(N) : \quad A^{N+1}|\phi| \geqslant ((1+\varepsilon)\rho(A))^N|\phi|.$$

- Pour $N = 1$, il s'agit de l'inégalité (3.2). La propriété $\mathcal{Q}(1)$ est donc vraie.
- Soit N un entier naturel non nul. On suppose que la propriété $\mathcal{Q}(N)$ est vraie.
Le résultat de la question (3.3) fournit l'inégalité :
$$A^{N+2}|\phi| \geqslant (1+\varepsilon)\rho(A)A^{N+1}|\phi|.$$

Cette dernière inégalité et l'hypothèse de récurrence $\mathcal{Q}(N)$ impliquent l'inégalité (on pourrait en effet facilement vérifier que la transitivité des inégalités est encore vraie pour ce nouveau type d'inégalité entre vecteurs) :

$$A^{N+2}|\phi| \geqslant (1+\varepsilon)\rho(A)((1+\varepsilon)\rho(A))^N|\phi| = ((1+\varepsilon)\rho(A))^{N+1}|\phi|.$$

L'assertion $\mathcal{Q}(N+1)$ est donc vraie.
D'après le principe de récurrence, on a donc démontré :

$$\boxed{\forall N \in \mathbb{N}^*, \quad A^{N+1}|\phi| \geqslant ((1+\varepsilon)\rho(A))^N|\phi|}.$$

(3.5) Soit N un entier naturel non nul. L'inégalité démontrée en (3.4) implique clairement l'inégalité suivante sur les normes infinies des vecteurs :

$$\|A^{N+1}|\phi|\|_\infty \geqslant \|((1+\varepsilon)\rho(A))^N|\phi|\|_\infty = ((1+\varepsilon)\rho(A))^N\||\phi|\|_\infty = ((1+\varepsilon)\rho(A))^N\|\phi\|_\infty \quad (1)$$

L'inégalité obtenue à la question (2.2) prouve que l'on a, pour toute matrice $M \in \mathcal{M}_n(\mathbb{R})$ et tout vecteur $v \in \mathbb{C}^n$:

$$\|Mv\|_\infty \leqslant \|M\|_\infty \, \|v\|_\infty.$$

On a donc ici en particulier :

$$\|A^{N+1}|\phi|\|_\infty \leqslant \|A^{N+1}\|_\infty \, \||\phi|\|_\infty = \|A^{N+1}\|_\infty \, \|\phi\|_\infty \quad (2)$$

Les inégalités (1) et (2) impliquent :

$$\|A^{N+1}\|_\infty \, \|\phi\|_\infty \geqslant ((1+\varepsilon)\rho(A))^N\|\phi\|_\infty.$$

La fonction $x \mapsto x^{\frac{1}{N+1}}$ est croissante sur $\mathbb{R}^+$. On en déduit par composition des inégalités :

$$\forall N \in \mathbb{N}^*, \quad \|A^{N+1}\|_\infty^{\frac{1}{N+1}} \, \|\phi\|_\infty^{\frac{1}{N+1}} \geqslant ((1+\varepsilon)\rho(A))^{\frac{N}{N+1}}\|\phi\|_\infty^{\frac{1}{N+1}} \quad (3)$$

Le vecteur ϕ étant non nul, sa norme est non nulle. Ainsi, la suite de terme général $\|\phi\|_\infty^{\frac{1}{N+1}}$ tend vers 1.
La suite de terme général $((1+\varepsilon)\rho(A))^{\frac{N}{N+1}}$ tend vers $(1+\varepsilon)\rho(A)$.
D'après le théorème de Gelfand, la suite de terme général $\|A^{N+1}\|_\infty^{\frac{1}{N+1}}$ tend vers $\rho(A)$.
Chacun des membres de l'inégalité (3) admet donc une limite finie. Le passage à la limite dans ces inégalités fournit :

$$\rho(A) \geqslant (1+\varepsilon)\rho(A)$$

c'est-à-dire

$$0 \geqslant \varepsilon\rho(A).$$

Le réel ε étant strictement positif, on constate que le réel positif $\rho(A)$ est nécessairement nul :

$$\rho(A) = 0 \quad (4)$$

D'autre part, comme la matrice A est par hypothèse strictement positive, tous ses coefficients sont strictement positifs. Notons $a_{1,1}^{(1)} > 0$ le coefficient situé à la première ligne et première colonne de la matrice A. Le fait que tous les coefficients de A soient positifs ainsi que la formule du produit matriciel montrent que le coefficient $a_{1,1}^{(2)}$ situé à la première ligne et à la première colonne de la matrice A^2 vérifie :

$$a_{1,1}^{(2)} \geqslant (a_{1,1}^{(1)})^2.$$

Par une récurrence immédiate, on démontre que, pour tout entier naturel non nul N, le coefficient $a_{1,1}^{(N)}$ situé à la première ligne et à la première colonne de la matrice A^N vérifie :

$$a_{1,1}^{(N)} \geqslant (a_{1,1}^{(1)})^N.$$

D'après la définition donnée dans l'énoncé de la norme infinie d'une matrice, on a, pour tout entier naturel non nul N :

$$\|A^N\|_\infty \geqslant a_{1,1}^{(N)} \geqslant (a_{1,1}^{(1)})^N$$

d'où, par croissance sur $\mathbb{R}^+$ de la fonction $x \mapsto x^{\frac{1}{N}}$:

$$\|A^N\|_\infty^{\frac{1}{N}} \geqslant a_{1,1}^{(1)}.$$

D'après le théorème de Gelfand, le membre de gauche admet une limite lorsque N tend vers l'infini, égale au rayon spectral $\rho(A)$. Le passage à la limite dans l'inégalité fournit :

$$\rho(A) \geqslant a_{1,1}^{(1)} > 0 \quad (5)$$

Les inégalités (4) et (5) se contredisent. Par conséquent, l'hypothèse « $A|\phi|$ n'est pas égal à $\rho(A)|\phi|$ » est absurde. On a donc démontré :

$$\boxed{A|\phi| = \rho(A)|\phi|}.$$

Comme le vecteur ϕ a été supposé non nul, il en est de même du vecteur $|\phi|$.
On vient donc de démontrer que
$\boxed{\text{le vecteur } |\phi| \text{ est un vecteur propre de la matrice } A\text{, associé à la valeur propre } \rho(A)}$.

(3.6) Appliquons le résultat de la question (1.4) aux vecteurs $u = 0_{\mathbb{R}^n}$ et $v = |\phi|$. On a bien $u \leqslant v$, u et v distincts (ϕ est un vecteur propre, donc est non nul) et A strictement positive, d'où $Au < Av$, c'est-à-dire $0 < A|\phi| = \rho(A)|\phi|$.
Cette dernière inégalité indique en particulier que le réel positif $\rho(A)$ n'est pas nul. On en déduit : $|\phi| > 0$.
On vient de démontrer que $\boxed{\text{le vecteur } |\phi| \text{ est strictement positif}}$.

(3.7) Soit i un entier compris entre 1 et n D'après les règles du produit matriciel, on obtient, comme à la question précédente :

$$\sum_{j=1}^{n} A_{i,j}\phi_j = (A\phi)_i = (\lambda\phi)_i = \lambda\phi_i.$$

d'où

$$\left|\sum_{j=1}^{n} A_{i,j}\phi_j\right| = |\lambda| \, |\phi_i| = \rho(A) \, |\phi_i|.$$

De même, on a :
$$\sum_{j=1}^{n} A_{i,j}|\phi_j| = (A|\phi|)_i = (\rho(A)|\phi|)_i = \rho(A)\,|\phi_i|.$$
On vient donc d'établir l'identité cherchée, pour tout entier i compris entre 1 et n :
$$\boxed{\left|\sum_{j=1}^{n} A_{i,j}\phi_j\right| = \sum_{j=1}^{n} A_{i,j}|\phi_j|}.$$
(3.8) D'après le résultat de la question (3.7) écrite pour $i=1$, on a :
$$\left|\sum_{j=1}^{n} A_{1,j}|\phi_j|\right|^2 - \left|\sum_{j=1}^{n} A_{1,j}\phi_j\right|^2 = 0 \quad (6)$$
Le carré du module d'un nombre complexe z est, par définition, égal au produit $z\overline{z}$. On a donc d'autre part, en utilisant le fait que le conjugué d'une somme (respectivement d'un produit) est la somme (resp. le produit) des conjugués :
$$\begin{aligned}\left|\sum_{j=1}^{n} A_{1,j}|\phi_j|\right|^2 - \left|\sum_{j=1}^{n} A_{1,j}\phi_j\right|^2 &= \sum_{i=1}^{n} A_{1,i}|\phi_i|\sum_{j=1}^{n} A_{1,j}|\phi_j| - \sum_{i=1}^{n} A_{1,i}\phi_i\sum_{j=1}^{n} A_{1,j}\overline{\phi_j} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} A_{1,i}A_{1,j}(|\phi_j||\phi_i| - \phi_i\overline{\phi_j}) \quad (7)\end{aligned}$$
Les identités (6) et (7) impliquent :
$$\sum_{i=1}^{n}\sum_{j=1}^{n} A_{1,i}A_{1,j}(|\phi_j||\phi_i| - \phi_i\overline{\phi_j}) = 0.$$
Considérons la matrice B de $\mathcal{M}_n(\mathbb{R})$ dont les coefficients $B_{i,j}$ sont définis pour tout couple (i,j) d'entiers compris entre 1 et n par :
$$B_{i,j} = A_{1,i}A_{1,j}.$$
Comme la matrice A est strictement positive, il en est de même de la matrice B. On peut donc appliquer le résultat de la question (1.3) :
$$\forall (i,j) \in \{1,\ldots,n\}^2, \quad |\phi_j||\phi_i| = \phi_i\overline{\phi_j}.$$
Le résultat demandé par l'énoncé en découle :
$$\boxed{\forall i \in \{1,\ldots,n\}, \quad |\phi_i||\phi_1| = \phi_i\overline{\phi_1}}.$$
(3.9) D'après le résultat de la question (3.6), chacune des composantes du vecteur $|\phi|$ est strictement positive. En particulier, le réel $|\phi_1| = |\phi|_1$ est strictement positif, donc le nombre complexe $\overline{\phi_1}$ est non nul. On peut donc écrire le résultat de la question précédente sous la forme :
$$\forall i \in \{1,\ldots,n\}, \quad \phi_i = |\phi_i|\frac{|\phi_1|}{\overline{\phi_1}}.$$
Posons
$$c = \frac{|\phi_1|}{\overline{\phi_1}}.$$
Le nombre complexe c est non nul (plus précisément, il a pour module 1) et vérifie :
$$\forall i \in \{1,\ldots,n\}, \quad \phi_i = c|\phi_i|.$$
On vient de montrer l'$\boxed{\text{existence d'un nombre complexe non nul } c \text{ vérifiant l'identité } \phi = c|\phi|}$.

(3.10) Multiplions par c membre à membre la relation $A|\phi| = \rho(A)|\phi|$, établie à la question (3.5) :
$$Ac|\phi| = \rho(A)c|\phi|.$$
À l'aide du résultat de la question (3.9), cette identité s'écrit aussi :
$$\boxed{A\phi = \rho(A)\phi}.$$
Comme le vecteur ϕ est non nul, ceci signifie que ϕ est un vecteur propre de A, associé à la valeur propre $\rho(A)$. D'autre part, par définition, le vecteur ϕ est vecteur propre de A associé à la valeur propre λ. Or un vecteur propre ne peut être associé qu'à une seule valeur propre. On en déduit :
$$\boxed{\lambda = \rho(A)}.$$
Observons que le nombre complexe λ a été défini comme une valeur propre de A quelconque de module égal à $\rho(A)$. On vient donc bien de démontrer que toute valeur propre λ de A de module égal à $\rho(A)$ est égale à $\rho(A)$.
En d'autres termes, $\boxed{\rho(A) \text{ est l'unique élément de l'ensemble } \{\lambda \in \mathrm{Sp}(A)\,;\, |\lambda| = \rho(A)\}}$.

(3.11) D'après le résultat de la question (3.10), le vecteur $|\phi|$, dont toutes les composantes sont strictement positives, est vecteur propre de A associé à la valeur propre $\rho(A)$:
$$A|\phi| = \rho(A)|\phi|.$$
Posons
$$v = \frac{1}{\sum_{i=1}^{n}|\phi|_i}|\phi|.$$
Le vecteur $\boxed{v \text{ bien strictement positif}}$, la somme de ses composantes est égale à 1, donc $\boxed{v \text{ est une loi de probabilité}}$. De plus, le vecteur v est colinéaire au vecteur $|\phi|$, qui est un vecteur propre de la matrice A associé à la valeur propre $\rho(A)$. Par conséquent, le vecteur v est aussi vecteur propre de la matrice A associé à la valeur propre $\rho(A)$:
$$\boxed{Av = \rho(A)v}.$$
(3.12) Soient v et v' deux vecteurs strictement positifs, v étant de plus une loi de probabilité, tous deux vecteurs propres de la matrice A associés à la valeur propre $\rho(A)$. Soit $\overline{c}$ un réel positif. On a :
$$v - \overline{c}v' \geqslant 0 \iff (\forall i \in \{1,\ldots,n\}, \quad v_i - \overline{c}v'_i \geqslant 0) \iff \left(\forall i \in \{1,\ldots,n\}, \quad \overline{c} \leqslant \frac{v_i}{v'_i}\right).$$

Posons

$$\boxed{\overline{c} = \min\left\{\frac{v_i}{v'_i}\,;\quad 1 \leqslant i \leqslant n\right\}}.$$

On définit ainsi un réel $\overline{c}$ positif vérifiant

$$v - \overline{c}v' \geqslant 0.$$

De plus, par définition d'un minimum, il existe (au moins) un entier i_0 compris entre 1 et n vérifiant

$$\overline{c} = \frac{v_{i_0}}{v'_{i_0}}$$

c'est-à-dire

$$v_{i_0} - \overline{c}v'_{i_0} = 0.$$

Le réel $\overline{c}$ ainsi défini satisfait donc les conditions de l'énoncé (et c'est le seul).

(3.13) Le choix du réel $\overline{c}$ effectué à la question précédente implique l'inégalité :

$$\overline{c}v' \leqslant v.$$

Supposons v différent de $\overline{c}v'$. Comme la matrice A est strictement positive, on peut appliquer le résultat de la question (1.4) :

$$A\overline{c}v' < Av$$

c'est-à-dire $\boxed{A(v - \overline{c}v') > 0 \text{ si } v \neq \overline{c}v'}$.

(3.14) Supposons, par l'absurde, que l'espace propre de la matrice A associé à la valeur propre $\rho(A)$ ne soit pas une droite vectorielle.
Il existe alors un vecteur propre $w' \in \mathbb{C}^n$ de A associé à la valeur propre $\rho(A)$ tel que w' ne soit pas colinéaire à v.
D'après le résultat de la question (3.6), le vecteur $|w'|$ est strictement positif. Posons

$$v' = |w'|.$$

On vient de construire un vecteur v' vérifiant les propriétés suivantes :

- v' est un vecteur propre de A (d'après le résultat de la question (3.5))
- v' est strictement positif (d'après le résultat de la question (3.6))
- v' n'est pas colinéaire à v (en effet, d'après le résultat de la question (3.9), il existe un nombre complexe c non nul vérifiant $w' = cv'$; si v' était colinéaire à v, alors w' serait lui aussi colinéaire à v, ce qui contredirait la définition de w').

D'après le résultat de la question (3.13), il existe un réel positif $\overline{c}$ vérifiant :

$$v - \overline{c}v' \geqslant 0 \quad \text{et} \quad \big(v \neq \overline{c}v' \Longrightarrow A(v - \overline{c}v') > 0\big).$$

La condition $v \neq \overline{c}v'$ est ici assurée par le fait que les vecteurs v et v' ne sont pas colinéaires.
Par conséquent, on a

$$A(v - \overline{c}v') > 0.$$

D'autre part, le vecteur $v - \overline{c}v'$ est un vecteur propre de la matrice A associé à la valeur propre $\rho(A)$ (puisqu'il est non nul et combinaison linéaire de deux vecteurs propres de A associés à la valeur propre $\rho(A)$). On a donc :

$$A(v - \overline{c}v') = \rho(A)(v - \overline{c}v').$$

Ces deux dernières identités impliquent :

$$\rho(A)(v - \overline{c}v') > 0.$$

Ainsi, chacune des composantes du vecteur $v - \overline{c}v'$ est strictement positive. Or on a choisi précédemment le réel $\overline{c}$ pour que l'une au moins de ces composantes soit nulle. On vient de mettre en évidence une contradiction. Il n'existe donc pas de vecteur propre de la matrice A associé à la valeur propre $\rho(A)$ qui ne soit pas colinéaire au vecteur v. L'espace propre associé à la valeur propre $\rho(A)$ est donc la droite vectorielle engendrée par le vecteur v. En d'autres termes, on vient de démontrer :

$$\boxed{\{w \in \mathbb{C}^n\,;\, Aw - \rho(A)w = 0\} = \operatorname{Ker}(A - \rho(A)I_n) = \operatorname{Vect}(v) = \{mv\,;\, m \in \mathbb{C}\}}.$$

(3.15) On sait déjà que $\boxed{\text{le vecteur } v \text{ est une loi de probabilité vérifiant } Av = \rho(A)v}$.
Un raisonnement du type de celui de la question (2.1) montre
l'$\boxed{\text{unicité d'une telle loi de probabilité}}$.

Les matrices stochastiques

(4.1) Démontrons que
$\boxed{\text{le produit de deux matrices de } \mathcal{M}_n(\mathbb{R}) \text{ stochastiques est une matrice stochastique}}$.
Soient A et B deux telles matrices. Avec les mêmes notations qu'au début du problème, on a, pour tout couple (i,j) d'entiers compris entre 1 et n, par définition du produit matriciel :

$$(AB)_{i,j} = \sum_{k=1}^{n} A_{i,k}B_{k,j}.$$

Chacun des réels $A_{i,k}$ et $B_{k,j}$ est positif, donc chacun des réels $(AB)_{i,j}$ est positif.
De plus, on a pour tout entier i compris entre 1 et n :

$$\sum_{j=1}^{n}(AB)_{i,j} = \sum_{j=1}^{n}\sum_{k=1}^{n} A_{i,k}B_{k,j} = \sum_{k=1}^{n} A_{i,k}\sum_{j=1}^{n} B_{k,j}.$$

Comme la matrice B est stochastique, chacune des sommes $\sum_{j=1}^{n} B_{k,j}$ est égale à 1. On a donc, pour tout entier i compris entre 1 et n :

$$\sum_{j=1}^{n}(AB)_{i,j} = \sum_{k=1}^{n} A_{i,k}.$$

Comme la matrice A est stochastique, chacune des sommes $\sum_{k=1}^{n} A_{i,k}$ est égale à 1. On a donc pour tout entier i compris entre 1 et n :

$$\sum_{j=1}^{n}(AB)_{i,j} = 1.$$

On vient de démontrer que le produit AB est stochastique.

Soit $S \in \mathcal{M}_n(\mathbb{R})$ une matrice stochastique. Pour tout entier naturel k, considérons l'assertion :

$$\mathcal{P}(k) : \text{la matrice } S^k \text{ est stochastique.}$$

- La matrice S^0 est égale à l'identité I_n qui est évidemment stochastique. La propriété $\mathcal{P}(0)$ est donc vraie.
- La matrice S est stochastique par hypothèse. La propriété $\mathcal{P}(1)$ est donc vraie.
- Soit k un entier naturel. On suppose la propriété $\mathcal{P}(k)$ vérifiée, *i.e.* la matrice S^k stochastique. La matrice S^{k+1} est produit des matrices S^k et S qui sont toutes deux stochastiques. D'après la propriété que l'on vient de démontrer, la matrice S^{k+1} est elle aussi stochastique. La propriété $\mathcal{P}(k+1)$ est vérifiée.

On vient de démontrer par récurrence sur l'entier k que,
$\boxed{\text{pour tout entier naturel } k\text{, la matrice } S^k \text{ est stochastique}}$.

(4.2) *Dans cette question, l'énoncé a « oublié » de préciser qu'on supposait la matrice S stochastique...*
Soit $v \in \mathbb{R}^n$ une loi de probabilité. Ainsi, toutes les composantes v_i du vecteur v sont positives, et leur somme $\sum_{i=1}^n v_i$ vaut 1.
Soit k un entier naturel. La matrice $({}^tS)^k$ est évidemment égale à la matrice ${}^t(S^k)$. Or on a démontré en (4.1) que la matrice S^k est stochastique. Posons $w = ({}^tS)^k v$. Les composantes w_i du vecteur w sont toutes clairement positives :

$$\forall i \in \{1, \dots, n\}, \quad w_i = \sum_{j=1}^n (({}^tS)^k)_{i,j} v_j = \sum_{j=1}^n ({}^t(S^k))_{i,j} v_j = \sum_{j=1}^n (S^k)_{j,i} v_j \geqslant 0.$$

Vérifions que la somme des composantes du vecteur w est égale à 1. On a

$$\sum_{i=1}^n w_i = \sum_{i=1}^n \sum_{j=1}^n (S^k)_{j,i} v_j = \sum_{j=1}^n \sum_{i=1}^n (S^k)_{j,i} v_j = \sum_{j=1}^n v_j \sum_{i=1}^n (S^k)_{j,i}.$$

Comme la matrice S^k est stochastique, chacune des sommes $\sum_{i=1}^n (S^k)_{j,i}$ est égale à 1. On obtient donc :

$$\sum_{i=1}^n w_i = \sum_{j=1}^n v_j.$$

Comme le vecteur v est une loi de probabilité, la somme $\sum_{j=1}^n v_j$ est égale à 1. On conclut donc :

$$\sum_{i=1}^n w_i = 1.$$

Finalement, $\boxed{\text{le vecteur } ({}^tS)^k v \text{ est une loi de probabilité pour tout entier naturel } k}$.

(4.3) *Dans cette question, l'énoncé a « oublié » de préciser qu'on supposait la matrice S stochastique...*
Notons

$$U = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$$

la matrice colonne à n lignes dont tous les coefficients sont égaux à 1. Pour tout entier i compris entre 1 et n, on a :

$$(SU)_i = \sum_{j=1}^n S_{i,j} U_j = \sum_{j=1}^n S_{i,j}.$$

Comme la matrice S est stochastique, les sommes $\sum_{j=1}^n S_{i,j}$ sont toutes égales à 1. Il en résulte, pour tout entier i compris entre 1 et n :

$$(SU)_i = 1 = U_i.$$

Ceci équivaut à :

$$SU = U.$$

Ainsi, $\boxed{\text{le vecteur } U \text{ est un vecteur propre de la matrice } S \text{ associé à la valeur propre } 1}$.

(4.4) *Dans cette question, l'énoncé a « oublié » de préciser qu'on supposait la matrice S stochastique...*
Comme la matrice S est stochastique, tous ses coefficients sont positifs et la somme des coefficients de chacune des lignes est égale à 1. Par conséquent, on a, pour tout entier i compris entre 1 et n :

$$\sum_{j=1}^n |S_{i,j}| = \sum_{j=1}^n S_{i,j} = 1.$$

Par définition de la norme infinie de la matrice S, on a donc :

$$\|S\|_\infty = 1.$$

D'après le résultat de la question (2.3), on a :

$$\|S\|_\infty \geqslant \rho(S)$$

c'est-à-dire ici :

$$1 \geqslant \rho(S).$$

D'autre part, on a mis en évidence à la question précédente le fait que 1 est valeur propre de la matrice S. Par conséquent, on a :

$$1 \leqslant \rho(S).$$

On conclut donc :

$$\boxed{\rho(S) = 1}.$$

(4.5) Si la matrice S est strictement positive et stochastique (*ce que l'énoncé oublie encore de préciser*), alors la matrice ${}^tS$ est elle aussi strictement positive, donc on peut lui appliquer le théorème de Perron-Froebenius démontré à la partie 3 :
$\boxed{\text{il existe une unique loi de probabilité } \pi \text{ vérifiant } ({}^tS)\pi = \pi\text{, et de plus on a } \pi > 0}$.

5. Le modèle de Wright-Fisher avec mutation

(5.1) On suppose dans cette question que l'on a $X_0 = i$, où i est un entier compris entre 0 et N. La variable X_1 peut alors être vue comme

$\boxed{\text{la somme de } N \text{ variables aléatoires indépendantes de loi de Bernoulli de même paramètre } \frac{i}{N}}$.

(5.2) D'après le résultat de la question (5.1), la loi conditionnelle de X_1 sachant $X_0 = i$ est $\boxed{\text{la loi binomiale de paramètres } N \text{ et } \frac{i}{N}}$.

(5.3) Rappelons l'énoncé du théorème central limite. Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes de même loi admettant une espérance m et une variance σ^2. Notons

$$\forall n \in \mathbb{N}^*, \quad M_n = \frac{Y_1 + \cdots + Y_n}{n}$$

leur moyenne empirique et

$$\forall n \in \mathbb{N}^*, \quad M_n^* = \frac{M_n - E(M_n)}{\sigma(M_n)} = \sqrt{n}\frac{M_n - m}{\sigma} = \frac{Y_1 + \cdots + Y_n - nm}{\sqrt{n}\sigma}$$

la variable centrée réduite obtenue à partir de la variable M_n.
Alors, pour tout couple $(\alpha, \beta) \in \mathbb{R}^2 \cup \{-\infty, +\infty\}$ vérifiant $\alpha < \beta$, on a :

$$\lim_{n\to+\infty} P(M_n^* \in [\alpha, \beta]) = \int_\alpha^\beta \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t.$$

Ainsi, lorsque l'entier n est « grand », on peut approcher la probabilité $P(M_n^* \in [\alpha, \beta]) = P\left(\frac{Y_1 + \cdots + Y_n - nm}{\sqrt{n}\sigma} \in [\alpha, \beta]\right)$ par le réel $\int_\alpha^\beta \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t$.

Soient a et b deux réels vérifiant $a < b$. Soient N un entier naturel non nul et $x \in]0, 1[$ tel que le produit Nx soit entier. D'après le résultat de la question (5.1), sachant $X_0 = Nx$, on peut voir la loi de la variable X_1 comme la loi de la somme de N variables indépendantes de loi de Bernoulli de même paramètre $\frac{Nx}{N} = x$, notées $B_1, \dots, B_N$.
Ainsi, on a :

$$\begin{aligned}
P\left(\frac{X_1 - Nx}{\sqrt{N}} \in [a, b] \,\Big|\, X_0 = Nx\right) &= P_{[X_0=Nx]}\left(\frac{X_1 - Nx}{\sqrt{N}} \in [a, b]\right) \\
&= P\left(\frac{B_1 + \cdots + B_N - NE(B_1)}{\sqrt{N}} \in [a, b]\right) \\
&= P\left(\frac{\sum_{i=1}^N B_i - NE(B_1)}{\sqrt{N}\sigma(B_1)} \in \left[\frac{a}{\sigma(B_1)}, \frac{b}{\sigma(B_1)}\right]\right)
\end{aligned}$$

La variance d'une variable de loi de Bernoulli de paramètre x vaut $x(1-x)$. Par conséquent, on a :

$$P\left(\frac{X_1 - Nx}{\sqrt{N}} \in [a, b] \,\Big|\, X_0 = Nx\right) = P\left(\frac{\displaystyle\sum_{i=1}^N B_i - NE(B_1)}{\sqrt{N}\sigma(B_1)} \in \left[\frac{a}{\sqrt{x(1-x)}}, \frac{b}{\sqrt{x(1-x)}}\right]\right).$$

D'après l'approximation (qui découle du théorème central limite) rappelée ci-dessus, $\boxed{\text{on peut approcher la probabilité } P\left(\frac{X_1 - Nx}{\sqrt{N}} \in [a, b] \,\Big|\, X_0 = Nx\right) \text{ par le réel } \int_{a/\sqrt{x(1-x)}}^{b/\sqrt{x(1-x)}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t \text{ lorsque } N \text{ est « grand ».}}$

Dans le cas particulier où les variables Y_i sont indépendantes de même loi de Bernoulli, le théorème central limite s'appelle théorème de Moivre-Laplace, que l'on pouvait aussi invoquer ici directement.

(5.4) D'après la formule des probabilités totales appliquée au système complet d'événements $([X_0 = i])_{0\leqslant i\leqslant N}$, on a, pour tout entier j compris entre 0 et N :

$$P(X_1 = j) = \sum_{i=0}^N P(X_1 = j \,|\, X_0 = i) P(X_0 = i).$$

Remarquons que l'on peut, comme d'habitude, donner un sens à cette formule même dans le cas où l'un au moins des événements $[X_0 = i]$ est de probabilité nulle ; dans ce cas, la probabilité conditionnelle $P(X_1 = j \,|\, X_0 = i)$ n'est pas définie, mais on peut poser par convention $P(X_1 = j \,|\, X_0 = i) P(X_0 = i) = 0$.
Avec les notations de l'énoncé, on a donc, pour tout entier j compris entre 0 et N :

$$(\pi^1)_j = \sum_{i=0}^N P_{i,j}(\pi^0)_i = \sum_{i=0}^N ({}^tP)_{j,i}(\pi^0)_i$$

D'après la définition du produit matriciel, ceci signifie :

$$\boxed{\pi^1 = ({}^tP)\pi^0}$$

(on sous-entend ici que les vecteurs de $\mathbb{R}^n$ qui interviennent, π^0 et π^1, sont identifiés à des matrices colonnes de $\mathcal{M}_{n,1}(\mathbb{R})$).

(5.5) Les coefficients de la matrice P sont des probabilités, donc ce sont en particulier des réels positifs.
Pour tout entier i compris entre 0 et N, on a :

$$\sum_{j=0}^N P_{i,j} = \sum_{j=0}^N P(X_1 = j \,|\, X_0 = i) = \sum_{j=0}^N P_{[X_0=i]}(X_1 = j) = P_{[X_0=i]}(\Omega) = 1.$$

On vient de démontrer que $\boxed{\text{la matrice } P \text{ est stochastique}}$.

Soit i un entier compris entre 0 et N. Sachant $X_0 = i$, la variable X_1 suit la loi binomiale de paramètres N et $\frac{i}{N}$. Ainsi, on a :

$$\boxed{\forall (i, j) \in \{0, \dots, N\}^2, \quad P_{i,j} = P_{[X_0=i]}(X_1 = j) = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}}.$$

(5.6) Dans ce nouveau cadre, pour tout entier i compris entre 0 et N, la loi de X_1 sachant $X_0 = i$ est la loi de la somme

$$\sum_{k=1}^{N} (B_{i,k}(1 - M_k) + (1 - B_{i,k})M_k)$$

où

- $B_{i,1}, \ldots, B_{i,N}, M_1, \ldots, M_N$ sont des variables aléatoires mutuellement indépendantes.
- Les variables $B_{i,1}, \ldots, B_{i,N}$ suivent toutes la même loi de Bernoulli de paramètre $\frac{i}{N}$; la variable $B_{i,k}$ prend la valeur 1 si l'allèle du $k^{\text{ème}}$ individu de la génération 1 (c'est-à-dire la deuxième génération) est l'allèle n ; la variable $B_{i,k}$ prend la valeur 0 si l'allèle du $k^{\text{ème}}$ individu de la génération 1 est l'allèle b.
- Pour tout entier k compris entre 1 et N, la variable M_k prend la valeur 1 si l'allèle dont hérite le $k^{\text{ème}}$ individu de la génération 1 a subi une mutation, et elle prend la valeur 0 si cet allèle n'a pas subi de mutation. Ainsi, les variables $M_1, \ldots, M_N$ suivent toutes la même loi de Bernoulli de paramètre p.

Pour tout entier k compris entre 1 et N, notons

$$Z_{i,k} = B_{i,k}(1 - M_k) + (1 - B_{i,k})M_k.$$

Ainsi, la variable aléatoire $Z_{i,k}$ est

- égale à 1 si
  - ou bien l'allèle du $k^{\text{ème}}$ individu de la génération 1 est l'allèle n du parent et cet allèle n'a pas muté (dans ce cas, on a $B_{i,k}(1-M_k)+(1-B_{i,k})M_k = 1(1-0)+(1-1)0 = 1$)
  - ou bien l'allèle du $k^{\text{ème}}$ individu de la génération 1 est l'allèle b du parent et cet allèle a muté en allèle n (dans ce cas, on a $B_{i,k}(1-M_k)+(1-B_{i,k})M_k = 0(1-1)+(1-0)1 = 1$).
- égale à 0 sinon.

Autrement dit, pour tout entier k compris entre 1 et N, en utilisant l'incompatibilité des événements $[B_{i,k} = 1]$ et $[B_{i,k} = 0]$ puis l'indépendance des variables $B_{i,k}$ et M_k, la variable aléatoire $Z_{i,k}$ suit la loi de Bernoulli de paramètre

$$\begin{aligned}
P(Z_{i,k} = 1) &= P\left(([B_{i,k} = 1] \cap [M_k = 0]) \cup ([B_{i,k} = 0] \cap [M_k = 1])\right) \\
&= P([B_{i,k} = 1] \cap [M_k = 0]) + P([B_{i,k} = 0] \cap [M_k = 1]) \\
&= P(B_{i,k} = 1)P(M_k = 0) + P(B_{i,k} = 0)P(M_k = 1) \\
&= \frac{i}{N}(1 - p) + (1 - \frac{i}{N})p.
\end{aligned}$$

Ainsi, pour tout couple (i, j) d'entiers compris entre 0 et N, on a :

$$P(X_1 = j \mid X_0 = i) = P\left(\sum_{k=1}^{N} Z_{i,k} = j\right)$$

où $Z_{i,1}, \ldots, Z_{i,N}$ sont des variables aléatoires mutuellement indépendantes, chacune des variables $Z_{i,k}$, k compris entre 1 et N, suivant la loi de Bernoulli de paramètre $\frac{i}{N}(1-p)+(1-\frac{i}{N})p$.

(5.7) *Remarquons que l'article défini « la » utilisé par l'énoncé dans « la » matrice B est impropre ; pour certains choix du vecteur π^0, plusieurs matrices B peuvent convenir. On va se contenter de donner une matrice B convenable.*
Comme à la question (5.7), on constate qu'il suffit de poser $B = (b_{i,j})_{0 \leqslant i,j \leqslant N}$ avec

$$\forall (i, j) \in \{0, \ldots, N\}^2, \quad b_{i,j} = P(X_1 = j \mid X_0 = i) = P\left(\sum_{k=1}^{N} Z_{i,k} = j\right)$$

pour obtenir la relation :

$$\pi^1 = ({}^tB)\pi^0.$$

(5.8) On suppose le réel p strictement compris entre 0 et 1. Chacune des probabilités $P(X_1 = j \mid X_0 = i)$, $(i, j) \in \{0, \ldots, N\}^2$, est alors non nulle. En effet, même dans le cas où i est nul, par exemple, c'est-à-dire dans le cas où aucun des parents n'est porteur de l'allèle n, le fait que la probabilité p de mutation soit non nulle montre que l'allèle n est présent à la génération 1 avec une probabilité non nulle.
Par conséquent, la matrice B est strictement positive.
Comme en (5.5), on constate que la matrice B est stochastique.
Les hypothèses de la question (4.5) sont donc vérifiées ; on en déduit
l'$\boxed{\text{existence d'une unique loi de probabilité } \pi \text{ vérifiant } \pi = ({}^tB)\pi}$.

Remarquons que si le vecteur π^0 est égal à π, alors on a

$$\pi^1 = ({}^tB)\pi^0 = ({}^tB)\pi = \pi$$

autrement dit le vecteur π^1 est aussi égal à π. On constate facilement que, pour tout entier naturel k, le vecteur π^k est égal à π. Ceci justifie la terminologie « loi invariante du modèle » pour qualifier la loi de probabilité π.

(5.9) *La notion d'espérance conditionnelle ne figure pas au programme de BCPST ! On peut malgré tout essayer d'en donner une définition qui semble intuitive ici...*
Soit i un entier compris entre 0 et N. On a par définition de l'espérance conditionnelle de X_1 sachant $X_0 = i$:

$$\begin{aligned}
E(X_1 \mid X_0 = i) &= \sum_{j=0}^{N} jP(X_1 = j \mid X_0 = i) \\
&= \sum_{j=0}^{N} jP\left(\sum_{k=1}^{N} Z_{i,k} = j\right) \\
&= E\left(\sum_{k=1}^{N} Z_{i,k}\right).
\end{aligned}$$

On utilise la linéarité de l'espérance :

$$E(X_1 \mid X_0 = i) = \sum_{k=1}^{N} E(Z_{i,k}).$$

Or chacune des variables $Z_{i,k}$ suit la loi de Bernoulli de paramètre $\frac{i}{N}(1-p)+(1-\frac{i}{N})p$. L'espérance conditionnelle de X_1 sachant $X_0 = i$ est donc égale à :

$$E(X_1 \,|\, X_0 = i) = \sum_{k=1}^{N} \left(\frac{i}{N}(1-p) + (1-\frac{i}{N})p \right)$$

c'est-à-dire

$$\boxed{E(X_1 \,|\, X_0 = i) = N \left(\frac{i}{N}(1-p) + (1-\frac{i}{N})p \right)}.$$

(5.10) On a, en utilisant de nouveau la formule des probabilités totales :

$$\begin{aligned}
E(X_1) &= \sum_{j=0}^{N} jP(X_1 = j) \\
&= \sum_{j=0}^{N} j \sum_{i=0}^{N} P(X_0 = i)P(X_1 = j \,|\, X_0 = i) \\
&= \sum_{i=0}^{N} P(X_0 = i) \sum_{j=0}^{N} jP(X_1 = j \,|\, X_0 = i) \\
&= \sum_{i=0}^{N} P(X_0 = i)E(X_1 \,|\, X_0 = i) \\
&= \sum_{i=0}^{N} P(X_0 = i)N \left(\frac{i}{N}(1-p) + (1-\frac{i}{N})p \right)
\end{aligned}$$

À l'aide du théorème de transfert, on reconnaît dans le membre de droite une espérance :

$$E(X_1) = E\left(N \left(\frac{X_0}{N}(1-p) + \left(1 - \frac{X_0}{N}\right) p \right) \right)$$

c'est-à-dire, par linéarité de l'espérance,

$$\boxed{E(X_1) = Np + (1-2p)E(X_0)}.$$

(5.11) Supposons que le variable X_0 suive la loi π (l'unique loi de probabilité vérifiant $\pi = ({}^tB)\pi$). La variable X_1 suit alors la loi $\pi^1 = \pi$. En particulier, les variables X_0 et X_1 ont même espérance. On a donc :

$$E(X_0) = Np + (1-2p)E(X_0)$$

c'est-à-dire

$$E(X_0) = \frac{N}{2}.$$

L'espérance de la variable X_0 étant égale à l'espérance $\sum_{i=0}^{N} i\pi_i$ de la loi π, on a démontré :

$$\boxed{\sum_{i=0}^{N} i\pi_i = \frac{N}{2}}.$$

(5.12) Pour tout entier naturel k, on a :

$$\pi^{k+1} = ({}^tB)\pi^k.$$

Les calculs effectués en (5.9) et (5.10) sont encore valables ici. On a donc :

$$\boxed{\forall k \in \mathbb{N}, \quad E(\pi^{k+1}) = Np + (1-2p)E(\pi^k)}.$$

Ainsi, la suite de terme général $E(\pi^k)$ est arithmético-géométrique. Cherchons un point fixe de cette relation de récurrence, c'est-à-dire un réel α vérifiant

$$\alpha = Np + (1-2p)\alpha.$$

Cette dernière relation équivaut à :

$$\alpha = \frac{N}{2}.$$

On a pour tout entier naturel k :

$$\begin{aligned}
E(\pi^{k+1}) &= Np + (1-2p)E(\pi^k) \\
\tfrac{N}{2} &= Np + (1-2p)\tfrac{N}{2}
\end{aligned}$$

c'est-à-dire, en retranchant membre à membre ces deux lignes :

$$\forall k \in \mathbb{N}, \quad E(\pi^{k+1}) - \frac{N}{2} = (1-2p)\left(E(\pi^k) - \frac{N}{2} \right).$$

Ceci signifie que la suite de terme général $E(\pi^k) - \frac{N}{2}$ est géométrique de raison $1-2p$. On a donc :

$$\forall k \in \mathbb{N}, \quad E(\pi^k) - \frac{N}{2} = (1-2p)^k \left(E(\pi^0) - \frac{N}{2} \right).$$

On a choisi p strictement compris entre 0 et 1. On a donc :

$$-1 < 1-2p < 1$$

donc la suite de terme général $(1-2p)^k$ tend vers 0.

On conclut donc que $\boxed{\text{la suite de terme général } E(\pi^k) \text{ converge vers } \frac{N}{2} = \sum_{i=0}^{N} i\pi_i = E(\pi)}$.

(5.13) Dans le cas où p vaut 0 ou 1, la relation

$$\forall k \in \mathbb{N}, \quad \pi^{k+1} = ({}^tB)\pi^k$$

est encore valable, de même que la relation

$$\forall k \in \mathbb{N}, \quad E(\pi^{k+1}) = Np + (1-2p)E(\pi^k)$$

et donc la relation

$$\forall k \in \mathbb{N}, \quad E(\pi^k) - \frac{N}{2} = (1-2p)^k \left(E(\pi^0) - \frac{N}{2} \right).$$

- Dans le cas $\boxed{p=0}$, on a
$$\forall k \in \mathbb{N}, \quad E(\pi^k) = E(\pi^0).$$
$\boxed{\text{La suite de terme général } E(\pi^k) \text{ est constante, donc elle converge vers } E(\pi^0)}$.

- Dans le cas $\boxed{p=1}$, on a :
$$\forall k \in \mathbb{N}, \quad E(\pi^k) = \frac{N}{2} + (-1)^k \left(E(\pi^0) - \frac{N}{2} \right).$$
On est amené à considérer deux cas :

  - $\boxed{\text{Si } E(\pi^0) \text{ est égal à } \frac{N}{2}}$, on a
  $$\forall k \in \mathbb{N}, \quad E(\pi^k) = \frac{N}{2}\cdot$$
  $\boxed{\text{La suite de terme général } E(\pi^k) \text{ est constante, donc elle converge vers } \frac{N}{2}}$.

  - $\boxed{\text{Si } E(\pi^0) \text{ est différent de } \frac{N}{2}}$, la suite extraite $(E(\pi^{2k}))_{k\in\mathbb{N}}$ converge vers $E(\pi^0)$, et la suite extraite $(E(\pi^{2k+1}))_{k\in\mathbb{N}}$ converge vers $N - E(\pi^0)$. Comme les deux réels $E(\pi^0)$ et $N - E(\pi^0)$ sont distincts ici, $\boxed{\text{la suite de terme général } E(\pi^k) \text{ diverge}}$.