

Maths – Bio

Modèles matriciels de population

François BIENVENU

Séances des 30 octobre et 6 novembre 2013

Remarques préliminaires

Le contenu de ce document est en grande partie inspiré d'un ouvrage de référence sur le sujet : *Matrix Population Models : Construction, Analysis and Interpretation*, de Hal Caswell [[Caswell, 2001](#)]. Le lecteur est donc invité à s'y reporter pour plus d'informations sur le sujet. Vous pouvez également me contacter (pour me signaler une erreur dans ce document, pour m'emprunter mon Caswell...) à l'adresse :

francois -point- bienvenu -arobase- free -point- fr

Les séances du GT sur les modèles matriciels de population étaient organisées sous forme de questions - réponses. Pour tenter de rendre compte de cela, des questions ont été incluses dans ce document. Le lecteur est invité à tenter d'y répondre avant de poursuivre sa lecture (sans quoi c'était bien la peine que je m'embête à les mettre!).

1 Introduction

1.1 A propos des modèles matriciels de population

Les modèles matriciels de population (MMP) se rattachent au domaine de la dynamique des populations, c'est-à-dire l'étude de modèles mathématiques permettant de décrire / prédire la dynamique de l'effectif (et éventuellement d'autres caractéristiques) d'un groupe d'individus au cours du temps – pour ceux qui se demanderaient quelle est la différence avec la démographie : je ne crois pas qu'il y en ait vraiment ; mais on tend à réserver le mot 'démographie' aux études concrètes de la dynamique des populations humaines.

Du fait des questions de biologie sous-jacentes (ou parfois du fait de leur histoire, qui au passage commence en 1945 avec un article de Patrick H. Leslie [[Leslie, 1945](#)]), l'étude des MMP se fait avec certaines conventions et notations – mais nous n'allons pas trop nous en préoccuper ici car nous allons plus nous intéresser aux objets mathématiques utilisés qu'aux questions de biologie. En conséquence, ce document n'est *pas* une introduction aux MMP – pour cela, se

référer à l'ouvrage de Caswell [Caswell, 2001] ou au cours de Stéphane Legendre.

1.2 Une digression sur les modèles en général

Avant d'attaquer le vif du sujet, une petite digression sur deux différents types de modèles :

Question : Quelle grande opposition entre deux types de modèles peut-on faire ?

On peut opposer les modèles empiriques aux modèles mécanistiques : les modèles empiriques ont pour but premier de décrire aussi précisément que possible des observations, le plus souvent dans un but prédictif. Beaucoup de modèles ont d'abord été de ce type, i.e. n'étaient initialement justifiés que par leur bonne adéquation aux observations (loi de Kepler, loi des gaz parfaits...). A l'inverse, les modèles mécanistiques n'ont généralement pas pour but de donner des résultats aussi proches de la réalité que possible, mais plutôt de comprendre comment un phénomène est possible en le faisant émerger à partir d'éléments "intuitifs" – on s'attache alors plus

à l'aspect qualitatif que quantitatif des résultats obtenus.

Bien sûr, cette distinction n'est pas une réelle dichotomie ; il s'agit plutôt d'un continuum... sur lequel un même modèle peut d'ailleurs occuper différentes positions en fonction du point de vue adopté ! Ainsi, un modèle comme la loi des gaz parfaits a d'abord été introduit et utilisé comme un modèle purement empirique, avant que la thermodynamique statistique n'en fournisse une interprétation mécanistique.

La raison pour cette digression dès l'introduction est que les MMP peuvent être placés quasiment aux deux extrémités de ce continuum, et qu'il est important de l'avoir en tête car la complémentarité de ces deux aspects contribue à l'intérêt de ces modèles : on peut les utiliser pour faire de la prédiction à court terme, tout en sachant qu'il y a une interprétation derrière et que les résultats ne sortent pas de nulle part, ou bien les utiliser pour faire de la théorie tout en sachant que malgré les hypothèses abusivement simplistes qui les sous-tendent, on obtient généralement des résultats raisonnablement réalistes à court terme.

2 Construction du modèle

2.1 Un premier modèle... et ses limitations

Partons d'un des modèles de dynamique des populations les plus simples qui soient : le modèle géométrique de Malthus,

$$n(t+1) = \lambda n(t),$$

où $n(t)$ est la taille de la population au temps t , et λ est appelé taux de croissance de la population. Sous cette forme, il s'agit plutôt d'un modèle empirique ; mais on peut le rendre plus mécanistique en l'écrivant

$$n(t+1) = (f + s) n(t)$$

et en interprétant s comme la probabilité de survie d'un individu entre t et $t+1$, et f comme le nombre moyen de descendants produits (et ayant survécu) par chaque individu dans le même intervalle de temps (qu'on appellera dans la suite intervalle de projection ou, plus simplement, pas de temps).

Question : Quelles sont les hypothèses sous-tendant cette interprétation mécanistique du modèle ?

- Tous les individus sont identiques, i.e. ont la même probabilité de survie et produisent en moyenne le même nombre de descendants à chaque pas de temps.
- La taille de la population est suffisamment importante pour que, du fait de la loi des grand nombre, la dynamique soit déterministe : la valeur prise par la taille de la population à chaque pas de temps est égale à son espérance.
- Il n'y a pas d'interactions des individus entre eux ou des individus avec leur environnement.

Les deux dernières hypothèses peuvent être relâchées, mais à un certain prix : une population de petite taille sera bien mieux décrite par un modèle stochastique, et la prise en compte des interactions entre les individus (directe ou via leur environnement) conduirait à un modèle non linéaire avec de la fréquence / densité dépendance ; or les modèles non linéaire deviennent vite complexes, surtout en temps discret (cf équation logistique en temps discret). Nous allons donc nous intéresser à la première hypothèse.

2.2 La prise en compte des différences entre individus

Relâcher l'hypothèse selon laquelle les individus sont tous identiques revient à introduire une *structure* de la population. Pour ce faire, le plus naturel est de considérer que les traits démographiques (probabilité de survie, fertilité...) des individus dépendent de *variables d'état*. Comme leur nom l'indique, ces variables renseignent sur l'état (comme dans "en bon / mauvais état", "en l'état actuel", "états d'âme", etc...) dans lequel un individu se trouve, et c'est de cet état que vont dépendre les traits démographiques.

La notion d'état permet de prendre en compte l'influence de l'histoire d'un individu sur ses traits démographiques, puisque l'état dans lequel un individu se trouve dépend de son passé; néanmoins, une fois que l'état d'un individu est connu, son passé n'importe plus : on considère que l'état contient *toute* l'information sur les caractéristiques d'un individu (y compris sur les probabilités de ses états futurs).

Cette idée d'état est satisfaisante car elle résout le problème des différences de traits démographiques

entre individus, au moins de façon conceptuelle : s'il subsiste des différences (non liées à des effets aléatoires) entre les individus se trouvant dans un même état, celles-ci doivent avoir une cause. Il doit donc exister une différence mesurable dans un des traits impliqués dans cette cause. Il suffit alors de prendre en compte ce trait comme variable d'état. En itérant le processus autant que souhaité, jusqu'à ce qu'il n'y ait plus de différences entre individus au sein d'un état ou du moins jusqu'à ce que celles-ci soient négligeables, on résout notre problème.

Mais même lorsqu'on a choisi les variables d'état, on peut construire deux types de modèles, selon que l'on choisit des variables d'état continues ou discrètes. Choisir des variables d'état continues conduit à utiliser un type de modèle (à la mode actuellement) appelés *Integral Projection Models* (IPM). Choisir des variables d'état discrètes conduit aux MMP, et c'est à ces modèles qu'on va s'intéresser ici.

Notons que le choix de variables d'état discrètes pour construire le modèle n'est pas incompatible avec le fait que ces variables soient en réalité continues : on peut utiliser ces variables continues pour définir

des classes discrètes. Par exemple, on peut définir des classes de tailles (ex : “petits”, “moyens”, “grands”) au moyen de fourchettes pour une variables comme la longueur du corps. Cette situation est très fréquente, les deux variables d’état les plus utilisées étant l’âge et la taille... Mais il existe tout de même de nombreuses situations où une variable discrète est un choix évident – par exemple lors qu’on considère des stades développementaux (ex : “larves”, “nymphe”, “adultes” chez les insectes). Enfin, notons également qu’il n’y a pas de restriction sur la nature des variables d’état et des classes : dans la plupart des cas, il s’agira de variables classiques (âge, taille...), mais on peut également choisir de prendre en compte un position géographique (populations structurées en dèmes), un statu social, etc...

2.3 Forme générale des modèles matriciels de population

Une fois les individus groupés en classes, il est naturel de définir le modèle en termes de *flux* entre ces classes. On distingue alors deux types de flux : ceux qui correspondent au “déplacement” d’un individu d’une classe vers une autre (ex : croissance) –

qui sont donc des probabilités de survie – et ceux qui correspondent à la production d’individus – qui correspondent à des fertilités. En fait, ces deux types de flux seront traités de la même manière, la plus simple qui soit : on considère que le flux sortant d’une classe est proportionnel au nombre d’individus dans cette classe, avec un coefficient de proportionnalité (probabilité de survie, fertilité, ou somme des deux lorsque les deux types de flux coexistent) constant d’un pas de temps sur l’autre. Ainsi, le modèle s’écrit dans sa forme générale :

$$n_i(t + 1) = \sum_j a_{ij} n_j(t),$$

où $n_i(t)$ est le nombre d’individus dans la classe i au temps t , et $a_{ij} = f_{ij} + s_{ij}$ est la somme de la fraction s_{ij} des individus de la classe j au temps t qui se trouveront dans la classe i au temps $t + 1$ et du nombre f_{ij} d’individus de classe i que produira un individu de la classe j entre t et $t + 1$.

Ce système peut s’écrire sous forme matricielle :

$$\mathbf{n}(t + 1) = \mathbf{A}\mathbf{n}(t),$$

où \mathbf{n} est le vecteur population, et $\mathbf{A} = (a_{ij})$ la matrice de projection de la population.

Question : Et le sexe dans tout ça ?

Ce modèle, du fait de la règle simpliste utilisée pour décrire les flux, est linéaire (i.e. $\mathbf{A}(x \mathbf{n} + \mathbf{m}) = x \mathbf{A} \mathbf{n} + \mathbf{A} \mathbf{m}$). Il ne prend pas en compte les interactions entre les individus. Or, s'il y a deux classes d'individus interagissant ensemble qu'on peut avoir envie d'opposer en dynamique de population (et pas en général ☺), il s'agit bien des femelles et des mâles. Cela pose un problème, car on s'attend à ce que le nombre d'individus produits dépende du sex-ratio dans la population et ses sous-classes de manière complexe. Heureusement, le constat que, dans beaucoup de populations, toutes femelles sont fécondées (les femelles sont limitantes – cf lien avec sélection sexuelle et anisogamie) permet de faire une simplification non aberrante en écrivant

$$n_i(t+1) = a_{ij} n_j^{\text{♀}}(t).$$

Mais puisque, sous cette hypothèse, seul le nombre de femelles importe pour la dynamique de la population, on n'a pas besoin de suivre le nombre de mâles et l'on peut redéfinir $a_{ij} = s_{ij} + f_{ij}$ de manière à ce que s_{ij} soit la probabilité de survie des femelles et que f_{ij} corresponde au nombre de femelles produites –

auquel cas on a

$$n_i^{\text{♀}}(t + 1) = a_{ij} n_j^{\text{♀}}(t).$$

On dit alors que le modèle est *female-based* et l'on arrête de préciser à chaque fois qu'on ne travaille sur les femelles. A noter que toutes ces questions ne se posent pas pour des populations asexuées – le modèle est alors dit *no-sex*.

2.4 Le graphe du cycle de vie

On peut représenter les relations entre les classes du modèle et ses flux par un graphe dont les sommets correspondent aux classes et les arêtes aux flux. Les flux étant quantitatifs et asymétriques, le graphe résultant est pondéré et orienté. Bien entendu, lorsqu'un flux est nul, plutôt que d'ajouter une arête de poids nul, on ne le prend pas en compte. On obtient généralement des graphes comme ceux représentés ci-dessous (figure 1) – mais bien que ces formes soient de loin les plus courantes, tout est possible. Le graphe obtenu est appelé graphe du cycle de vie.

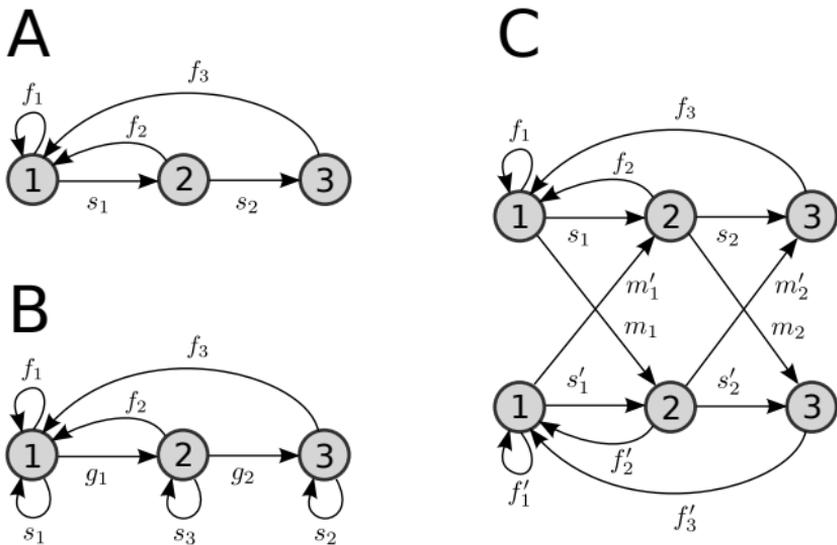


FIGURE 1 – Exemples typiques de cycles de vie. (A) Modèle en classes d'âge (aussi appelé modèle de Leslie). (B) Modèle en classes de taille. (C) Modèle de métapopulation (correspondant ici à un modèle en classes d'âge pour une population structurée en deux dèmes).

Question : Quel est le lien entre le graphe du cycle de vie et la matrice de projection de la population ?

A est la matrice d'adjacence (ou, si l'on suit la convention la plus fréquemment employée pour définir les matrices d'adjacence, la *transposée* de cette matrice), i.e. a_{ij} est le poids de l'arête allant

de j à i si celle-ci existe, et 0 sinon. Il s'agit d'une des diverses façons de représenter un graphe.

Nous avons terminé notre construction des MMP. Il s'agit de modèles simples dont on voit déjà qu'ils sont très faciles à simuler (en itérant le produit matriciel). Mais la simulation n'est pas ce qui rend ces modèles intéressants – il s'agit plutôt du fait que leur simplicité permet d'obtenir un grand nombre de résultats analytiques, et c'est ce à quoi on va s'intéresser maintenant.

3 Analyse du modèle

Question : Quelles sont les questions qu'on pourrait se poser pour ce type de modèle ?

Voici une liste de questions assez générales qu'on peut avoir envie de se poser. Toutes ne seront pas traitées en détail ici (mais toutes seront abordées).

- Comportement à long terme (régime asymptotique) ?
- Comportement à court terme (régime transitoire) ?

- Sensibilité des résultats obtenus aux détails d'un modèle particulier ?

3.1 Régime asymptotique

Question : Essayer de voir ce qui se passe en $t = +\infty$ (indice : se rappeler les résultats de réduction des endomorphismes, et ne pas hésiter à faire des disjonctions de cas / introduire des hypothèses supplémentaires).

3.1.1 Rappels sur la réduction des endomorphismes

Un endomorphisme (i.e. une application linéaire d'un espace vectoriel dans lui-même) d'un espace vectoriel de dimension finie peut-être représenté au moyen d'une matrice carrée. Cela présuppose néanmoins le choix de deux bases : en effet, la représentation se fait en prenant pour colonnes de la matrice les coordonnées de l'image des vecteurs de la première base dans la seconde, ce qui permet d'obtenir l'image d'un vecteur quelconque au moyen d'un simple produit matriciel. En résumé,

$$y = f(x) \quad \longleftrightarrow \quad \mathbf{y} = \mathbf{M}\mathbf{x},$$

où \mathbf{x} (resp. \mathbf{y}) est le vecteur colonne des coordonnées de x (resp. y) dans la base \mathcal{B} (resp. \mathcal{B}') et \mathbf{M} est la matrice représentant f (dont les colonnes sont les coordonnées des images par f des vecteurs de \mathcal{B} dans \mathcal{B}').

En général, une base – qu'on appelle alors base *canonique* – s'impose comme étant plus naturelle que les autres. Dans \mathbb{R}^n , s'il s'agit de la base $\mathcal{C} = (e_1, \dots, e_n)$, où e_k est le vecteur dont toutes les composantes sont nulles à l'exception de la k -ième, qui vaut 1. On choisit alors généralement $\mathcal{B} = \mathcal{B}' = \mathcal{C}$ pour représenter les endomorphismes.

Mais rien ne garantit que ce choix soit toujours très judicieux – il se pourrait que, pour certains endomorphismes, d'autres bases soient plus adaptées (au sens qu'elles permettront de représenter l'endomorphisme par une matrice ayant des propriétés particulières facilitant la résolution d'un problème donné). La réduction des endomorphismes est la recherche de ces bases.

Intuitivement, on aimerait avoir une base dans laquelle l'endomorphisme “en fait le moins possible” (par exemple, que sa matrice se rapproche de celle de

l'endomorphisme le plus fainéant du monde : l'identité). On va donc tenter de trouver des “invariants” de l'endomorphisme.

On appelle vecteur propre à droite (*right eigenvector*, en anglais) un vecteur \mathbf{w} non-nul – sinon c'est trop facile – tel que :

$$\exists \lambda, \quad \mathbf{M}\mathbf{w} = \lambda\mathbf{w}.$$

En gros, dans cette direction, l'endomorphisme de réalise qu'une simple homothétie d'un facteur λ appelé valeur propre (*eigenvalue*) associée à ce vecteur. De même, on introduit la notion de vecteur propre à gauche :

$$\mathbf{v} \text{ tq } \exists \lambda, \quad \mathbf{v}\mathbf{M} = \lambda\mathbf{v}.$$

Si on trouve un système de vecteurs propres formant une base de l'espace vectoriel, on est très contents car cela veut dire qu'en fin de compte notre endomorphisme n'est guère qu'une sorte d'homothétie non-isotrope (i.e. avec des facteurs différents dans différentes directions – ce qui explique la composante rotationnelle observée pour les vecteurs non-alignés dans cette direction), et que si on le représente dans la base des vecteurs propres on obtient une matrice diagonale (dont les éléments diagonaux sont bien entendu les valeurs propres), type de matrices sur lequel

on sait plein de choses (on sait qu'elles commutent, on sait les inverser, les élever à la puissance k , les exponentier, les traiter de façon efficace en informatique, etc...). On dit alors que l'endomorphisme est diagonalisable, ce qui traduit le fait qu'il peut être représenté par une matrice diagonale – ou encore que la matrice \mathbf{M} est semblable à une matrice diagonale, i.e.

$$\exists \mathbf{D} \text{ diagonale, } \mathbf{P} \text{ inversible tq } \mathbf{M} = \mathbf{PDP}^{-1}.$$

Au passage, \mathbf{P} est appelée matrice du changement de base. On peut l'obtenir en prenant pour colonnes les vecteurs propres formant la base $(\mathbf{w}_1, \dots, \mathbf{w}_n)$ trouvée (pris dans le même ordre que les valeurs propres formant la diagonale de \mathbf{D}).

Des fois, on ne peut pas trouver de base de vecteurs propres... Dans ce cas, la situation est plus compliquée, mais on peut quand même faire des choses (par exemple, on peut toujours trouver un couple de matrices \mathbf{P} inversible et \mathbf{T} triangulaire (contenant éventuellement des entrées complexes) telles que $\mathbf{M} = \mathbf{PTP}^{-1}$, ce qui permet d'exploiter les propriétés des matrices triangulaires).

Enfin, il est important de voir ce que la diagonal-

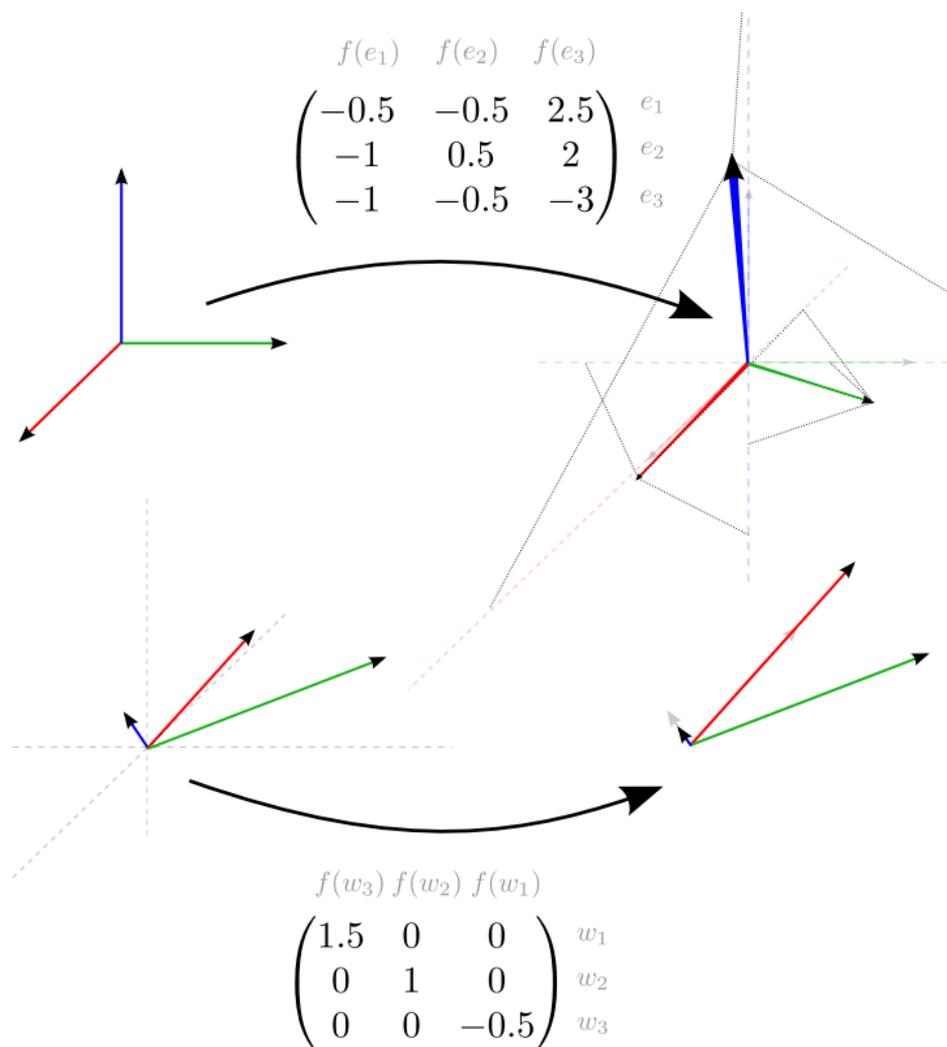


FIGURE 2 – Interprétation géométrique de la diagonalisation : en haut, l'action d'un endomorphisme sur la base canonique est représentée. La correspondance entre les vecteurs de la base et leur image n'est pas évidente, rendant peu intuitive l'action de l'endomorphisme ; en bas, l'action du même endomorphisme, mais cette fois sur une base de vecteurs propres. Cette fois, on voit bien que cet endomorphisme est en fait une homothétie non-isotrope, et l'on voit même les directions dans lesquelles il y a dilatation ou contraction.

isation représente géométriquement si l'on veut pouvoir développer une intuition sur ce que représentent les valeurs propres et les vecteurs propres. La figure 2 est une tentative de représenter cela (bonne chance pour le voir en 3D – même moi qui l'ai dessinée je n'y arrive pas trop ; c'est d'ailleurs dommage car je pense que c'est un bon exemple. Les matrices sont là pour en témoigner).

3.1.2 Application à l'étude du régime asymptotique

Avant de traiter le problème de façon formelle, essayons de nous former une intuition géométrique. Commençons par supposer que la matrice de projection de la population, \mathbf{A} , est diagonalisable. Comme on vient de le rappeler, cela signifie qu'à chaque pas de temps, le vecteur population \mathbf{n} est dilaté / contracté dans différentes directions (avec un changement de sens si la valeur propre associée est négative). Comme l'illustre la figure 3, on peut donc s'attendre à ce que le vecteur population tende à s'aligner dans la direction du vecteur propre associé à la valeur propre la plus grande – du moins si celle-ci (1) existe et (2) est simple, c'est-à-dire a un sous-espace propre associé de dimension 1.



FIGURE 3 — Illustration en dimension 2 du fait qu’au fil des projections de la population (i.e. des applications de l’endomorphisme représenté par \mathbf{A}), le vecteur population tend à s’aligner dans la direction du vecteur propre associé à la valeur propre la plus grande. Les deux vecteurs propres sont représentés en couleur sur la droite, avec leur valeur propre associée. Les vecteurs noirs correspondent à la séquence des vecteurs population.

Afin de tenter de confirmer notre intuition, supposons donc non seulement que \mathbf{A} est diagonalisable, mais aussi qu’elle possède une valeur propre réelle strictement supérieure en module à toutes les autres et de multiplicité 1. On peut tout à fait utiliser le fait que $\mathbf{A}^t = \mathbf{P}\mathbf{D}^t\mathbf{P}^{-1}$ pour voir ce qui se passe, mais il est équivalent et légèrement plus élégant d’écrire que, puisque la famille $(\mathbf{w}^{(k)})$ des vecteurs propres à droite forme une base de l’espace vectoriel, il existe une unique combinaison de réels c_1, \dots, c_n tels que

$$\mathbf{n}(0) = c_1\mathbf{w}^{(1)} + \dots + c_n\mathbf{w}^{(n)}.$$

Ainsi, par linéarité du produit matriciel et en utilisant le fait que $(\mathbf{w}^{(k)})$ est un vecteur propre de \mathbf{A} associé à la valeur propre λ_k , on obtient :

$$\mathbf{n}(1) = \mathbf{A}\mathbf{n}(0) = c_1\mathbf{A}\mathbf{w}^{(1)} + \dots + c_n\mathbf{A}\mathbf{w}^{(n)} = c_1\lambda_1\mathbf{w}^{(1)} + \dots + c_n\lambda_n\mathbf{w}^{(n)}.$$

Par récurrence,

$$\mathbf{n}(t) = c_1 \lambda_1^t \mathbf{w}^{(1)} + \dots + c_n \lambda_n^t \mathbf{w}^{(n)} .$$

Intéressons-nous à la i -ième composante de ce vecteur,

$$n_i(t) = c_1 \lambda_1^t w_i^{(1)} + \dots + c_n \lambda_n^t w_i^{(n)} .$$

On peut supposer sans perte de généralité que les vecteurs $(\mathbf{w}^{(k)})$ ont été ordonnés par ordre décroissant du module de la valeur propre associée, si bien que λ_1 est la valeur propre de plus grand module, qu'on a supposée réelle. On voit que si $w_i^{(1)} \neq 0$ et $c_1 \neq 0$, en $t \rightarrow +\infty$:

$$n_i(t) \sim c_1 \lambda_1^t w_i^{(1)} .$$

Si toutes les composantes de $\mathbf{w}^{(1)}$ sont non-nulles, et si la composante du vecteur population initial selon $\mathbf{w}^{(1)}$ dans la base des vecteurs propres n'est pas nulle, on a donc

$$\mathbf{n}(t) \sim c_1 \lambda_1^t \mathbf{w}^{(1)} .$$

Ceci confirme notre intuition initiale sur le comportement asymptotique : au fil des itérations, le vecteur population vient s'aligner dans la direction de $\mathbf{w}^{(1)}$ puis se dilate / contracte d'un facteur λ_1 à chaque pas

de temps. On peut donc interpréter λ_1 comme le taux de croissance asymptotique de la population (attention néanmoins au fait qu'il s'agit d'un taux de croissance *géométrique* – le taux de croissance *exponentiel* correspondant étant $r = \log \lambda_1$). Le vecteur $\mathbf{w}^{(1)}$ est quant à lui interprété comme la structure stable de la population : on peut le rescaler de manière à ce que ses entrées correspondent aux fractions d'individus dans les classes du modèle à temps long ($\frac{n_i(t)}{n(t)} \rightarrow w_i^{(1)}$ quand $t \rightarrow \infty$, où $n(t) = \sum_i n_i(t)$ est la taille totale de la population).

Néanmoins, à ce stade, on a dû faire un certain nombre d'hypothèses pour obtenir ces résultats :

1. Il existe une valeur propre réelle λ_1 , supérieure à toutes les autres en module et de multiplicité 1.
2. Les entrées du vecteur propre $\mathbf{w}^{(1)}$ associé à λ_1 sont toutes non-nulles.
3. \mathbf{A} est diagonalisable.
4. La composante du vecteur population initial selon $\mathbf{w}^{(1)}$ dans la base des vecteurs propres n'est pas nulle.

En fait, on va voir qu'il existe un théorème justifiant les deux premières hypothèses. Mais avant cela, intéressons-nous à l'hypothèse selon laquelle \mathbf{A} est diagonalisable. Nous reviendrons également sur la dernière hypothèse, qui dans sa formulation actuelle présuppose que \mathbf{A} est diagonalisable.

Remarque : le paragraphe qui suit utilise des résultats qui ne sont pas au programme de prépa BCPST. Néanmoins, ceux-ci seront rappelés. De plus, il s'agit de résultats classiques sur lesquels le lecteur trouvera facilement plus d'information (des liens seront donnés). La lecture de ce paragraphe ne devrait donc pas poser de problème, mais il est tout à fait possible de passer directement à la [section suivante](#).

Pour simplifier les notations, notons simplement $\lambda = \lambda_1$, $\mathbf{w} = \mathbf{w}^{(1)}$ et $\mathbf{v} = \mathbf{v}^{(1)}$. Considérons maintenant l'application définie par la matrice :

$$\mathbf{P} = \frac{\mathbf{w}\mathbf{v}}{\mathbf{v}\mathbf{w}} = \frac{1}{\sum_k v_k w_k} \begin{pmatrix} w_1 v_1 & \cdots & w_1 v_n \\ \vdots & & \vdots \\ w_n v_1 & \cdots & w_n v_n \end{pmatrix}.$$

Puisque toutes ses colonnes sont proportionnelles à \mathbf{w} , il est clair que cette matrice est de rang 1 et que

$\text{Im}(\mathbf{P}) = \text{Vect}(\mathbf{w})$. De plus,

$$\mathbf{P}^2 = (c_{ij}) \quad \text{où} \quad c_{ij} = \sum_k \frac{w_i v_k}{\mathbf{v}\mathbf{w}} \frac{w_k v_j}{\mathbf{v}\mathbf{w}} = \frac{w_i v_j}{\mathbf{v}\mathbf{w}},$$

c'est-à-dire $\mathbf{P}^2 = \mathbf{P}$.

Théorème : il est équivalent de dire d'un endomorphisme f de E :

1. f est idempotent : $f \circ f = f$
2. Le noyau et l'image de f sont supplémentaires :
 $E = \text{Ker}(f) \oplus \text{Im}(f)$

Un tel endomorphisme est appelé un projecteur, ou encore la projection sur $\text{Im}(f)$ parallèlement à $\text{Ker}(f)$.

[http://fr.wikipedia.org/wiki/Projecteur_\(mathématiques\)](http://fr.wikipedia.org/wiki/Projecteur_(mathématiques))

Ainsi, de $\mathbf{P}^2 = \mathbf{P}$ on déduit que $E = \text{Ker}(\mathbf{P}) \oplus \text{Im}(\mathbf{P})$. Or, on remarque que \mathbf{P} et \mathbf{A} commutent. En effet,

$$\mathbf{A}\mathbf{P} = \frac{1}{\mathbf{v}\mathbf{w}} \mathbf{A}\mathbf{w}\mathbf{v} = \frac{1}{\mathbf{v}\mathbf{w}} (\lambda\mathbf{w})\mathbf{v} = \frac{1}{\mathbf{v}\mathbf{w}} \mathbf{w}(\lambda\mathbf{v}) = \frac{1}{\mathbf{v}\mathbf{w}} \mathbf{w}\mathbf{v}\mathbf{A} = \mathbf{P}\mathbf{A}$$

Théorème : Si \mathbf{A} commute avec \mathbf{B} , alors le noyau et l'image de tout polynôme en \mathbf{B} sont stables par \mathbf{A} .

https://fr.wikipedia.org/wiki/Paire_de_matrices_commutantes

On en déduit donc que $\text{Ker}(\mathbf{P})$ et $\text{Im}(\mathbf{P}) = \text{Vect}(\mathbf{w})$ sont stables par \mathbf{A} . Pour faire le parallèle avec la démonstration dans le cas diagonalisable : en choisissant $\mathbf{w}^{(1)}, \dots, \mathbf{w}^{(n)}$ comme base, nous avons en fait décomposé E en une somme de sous-espaces $\text{Vect}(\mathbf{w}^{(1)}), \dots, \text{Vect}(\mathbf{w}^{(n)})$ tous stables par \mathbf{A} . Ceci nous a permis de montrer que la composante selon $\mathbf{w}^{(1)}$ grandissait plus vite que les autres, qui devenaient négligeables. On voit donc qu'en réalité, seule la décomposition de E en $\text{Vect}(\mathbf{w}^{(1)})$ et un autre sous-espace stable par \mathbf{A} importait vraiment – or c'est précisément cette décomposition que nous venons d'obtenir grâce à \mathbf{P} .

Considérons maintenant l'image de $\mathbf{n}(0)$ par $(\frac{1}{\lambda}\mathbf{A})^t$. On commence par décomposer $\mathbf{n}(0)$ selon \mathbf{P} :

$$\begin{aligned} \mathbf{n}(0) &= \underbrace{\mathbf{P}\mathbf{n}(0)}_{\in \text{Im}(\mathbf{P})} + \underbrace{\mathbf{n}(0) - \mathbf{P}\mathbf{n}(0)}_{\in \text{Ker}(\mathbf{P})} \\ &= c\mathbf{w} + c'\mathbf{x} \end{aligned}$$

Appliquons maintenant $\frac{1}{\lambda}\mathbf{A}$ une première fois :

$$\begin{aligned} \frac{1}{\lambda}\mathbf{A}\mathbf{n}(0) &= c\frac{1}{\lambda}\mathbf{A}\mathbf{w} + c'\frac{1}{\lambda}\mathbf{A}\mathbf{x} \\ &= c\mathbf{w} + c'\mathbf{B}\mathbf{x}, \end{aligned}$$

où \mathbf{B} est la restriction de $\frac{1}{\lambda}\mathbf{A}$ à $\text{Ker}(\mathbf{P})$. Par récurrence,

$$\left(\frac{1}{\lambda}\mathbf{A}\right)^t \mathbf{n}(0) = c\mathbf{w} + c'\mathbf{B}^t\mathbf{x}.$$

Or les valeurs propres de \mathbf{B} sont toutes de module strictement inférieur à 1, puisque par hypothèse λ est strictement supérieure en module à toutes les autres valeurs propres de \mathbf{A} . Par conséquent, $\mathbf{B}^t \rightarrow (0)$ et donc

$$\left(\frac{1}{\lambda}\mathbf{A}\right)^t \mathbf{n}(0) \rightarrow c\mathbf{w},$$

c'est-à-dire, si $c \neq 0$,

$$\mathbf{n}(t) \rightarrow c\lambda^t\mathbf{w}.$$

On retrouve bien le résultat que nous avons montré dans le cas diagonalisable, et on peut donc faire la même interprétation de λ (taux de croissance asymptotique) et de \mathbf{w} (structure stable) que précédemment.

Question : Que peut-on dire de la constante c ? Qu'en déduit-on ?

$$c\mathbf{w} = \mathbf{P}\mathbf{n}(0) = \begin{pmatrix} \mathbf{wv} \\ \mathbf{vw} \end{pmatrix} \mathbf{n}(0) = \begin{pmatrix} \mathbf{vn}(0) \\ \mathbf{vw} \end{pmatrix} \mathbf{w},$$

d'où

$$c = \frac{\mathbf{v}\mathbf{n}(0)}{\mathbf{v}\mathbf{w}}.$$

Cela signifie que, lorsque celui-ci est choisi de manière à ce que $\mathbf{v}\mathbf{w} = 1$, les entrées du vecteur \mathbf{v} sont des poids par lesquels les nombres initiaux d'individus dans chaque classe sont pondérés pour obtenir la taille de la population lorsque $t \rightarrow \infty$. On peut donc interpréter v_i comme la contribution relative d'un individu de la classe i à la taille de totale de la population à temps longs. C'est pourquoi on appelle \mathbf{v} le vecteur des *valeurs reproductives* (il existe d'autres arguments pour cette interprétation). A noter qu'on appelle généralement v_i la "valeur reproductrice d'un individu de classe i ", tandis que l'expression "valeur reproductrice de la classe i " fait référence au produit $v_i w_i$.

Dans ce qui précède, on a supposé $c = \frac{\mathbf{v}\mathbf{n}(0)}{\mathbf{v}\mathbf{w}} \neq 0$. En réalité, cette hypothèse est très naturelle car il est extrêmement peu probable que le vecteur population initial ait une composante selon \mathbf{w} qui soit exactement nulle.

Enfin, on a également fait l'hypothèse que \mathbf{A} possédait une valeur propre réelle positive et stricte-

ment supérieure en module à toutes les autres, de multiplicité 1 et telle que le vecteur propre à droite associé ne contienne pas d'entrées nulles. On va maintenant justifier cette hypothèse.

3.1.3 Le théorème de Perron-Frobenius

Derrière l'expression "théorème de Perron-Frobenius" sont regroupés un ensemble de résultats dus à Oskar Perron et Georg Frobenius concernant les éléments propres de certaines matrices. Nous n'énonçons ici que les résultats dont nous avons besoin.

On dit qu'une matrice \mathbf{M} est positive au sens de Perron-Frobenius, ce qu'on note $\mathbf{M} \geq 0$, si toutes ses entrées sont positives. De même, une matrice est dite positive stricte lorsque toutes ses entrées sont strictement supérieures à 0, et l'on note $\mathbf{M} > 0$. Enfin, une matrice $\mathbf{M} \geq 0$ est dite primitive lorsqu'il existe un entier k pour lequel $\mathbf{M}^k > 0$. (*Attention : (1) il existe une autre notion de positivité d'une matrice, liée à la forme bilinéaire associée (2) la convention anglo-saxonne est de réserver l'usage du mot "positive" à la notion de positivité stricte et d'employer le terme "non-negative" pour désigner la positivité large.*)

Théorème de Perron-Frobenius : soit M une matrice primitive.

- M possède une valeur propre réelle strictement positive, supérieure en module à toutes les autres et de multiplicité 1, appelée *valeur propre dominante*.
- Les vecteurs propres à gauche et à droite associée à la valeur propre dominante, appelés vecteurs propres dominants, sont uniques à un facteur multiplicatif près et peuvent être choisis de manière à avoir toutes leurs entrées strictement positives.

Les conclusions du théorème de Perron-Frobenius correspondent donc exactement aux hypothèses que nous avons faites pour étudier le comportement asymptotique de $\mathbf{n}(t)$. Néanmoins, ce théorème reste un peu mystérieux, notamment du fait de la notion de primitivité. Heureusement, il existe une autre caractérisation – équivalente mais beaucoup plus intuitive – de cette propriété.

Caractérisation de la primitivité :

- **Irréductibilité** : Une matrice est dite irréductible lorsque le graphe correspondant est fortement

connexe, c'est-à-dire que pour tout couple (i, j) de sommets, il existe un chemin allant de i à j (i.e. il est possible d'aller de i à j en suivant les arêtes du graphe).

- **Périodicité** : La période d'une matrice est le PGCD (plus grand diviseur commun) de la longueur des cycles du graphe correspondant. Lorsque la période d'une matrice est égale à 1, celle-ci est dite apériodique.
- **Primitivité** : Une matrice positive est primitive si et seulement si elle est irréductible et apériodique.

A première vue, cette caractérisation ne paraît pas vraiment plus simple. Mais on peut déjà remarquer qu'elle rend plus apparent le fait que la primitivité est une propriété de la *structure* de la matrice, au sens qu'elle ne dépend pas des valeurs numériques de ses entrées mais simplement de la position des entrées non-nulles (ceci peut aussi se voir avec la caractérisation $\exists k$ tq $\mathbf{M}^k > 0$, mais de façon légèrement moins évidente). Mais surtout, cette caractérisation est beaucoup plus simple à utiliser en pratique : l'irréductibilité d'un graphe de taille raisonnable se voit de façon quasi-instantanée ;

quant à l'apériodicité, elle peut sembler plus difficile à vérifier, mais en pratique il existe des “astuces” la rendant simple à repérer.

Question : Quelles pourraient être deux de ces astuces ?

Une première situation dans laquelle l'apériodicité est évidente est lorsqu'une boucle (arête pointant sur le sommet dont elle part) est présente dans le graphe du cycle de vie. En effet, cette boucle forme un cycle de longueur 1 ; or le PGCD de tout ensemble d'entiers contenant 1 est égal à 1. De même, une autre situation courante correspond à la présence d'un cycle de longueur p et d'un cycle de longueur $p + 1$: si un nombre divise p et $p + 1$, alors il divise toute combinaison linéaire (à coefficients entiers) de ces deux nombres, dont $p + 1 - p = 1$. Le PGCD d'un ensemble d'entiers contenant p et $p + 1$ est donc égal à 1. Cette situation est très fréquente et se repère facilement dans les cycles de vies utilisés en pratique.

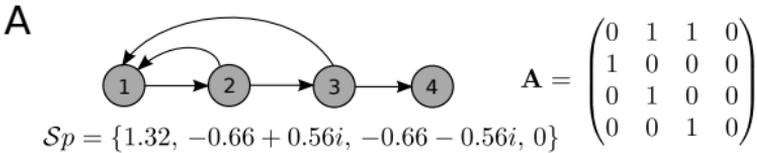
Le lecteur peut se convaincre de la facilité de la vérification de la primitivité au moyen de la caractérisation “positive, irréductible, apériodique” en l'appliquant aux graphes présentés en figure 1 pour

montrer qu'ils correspondent bien à des matrices de projection primitives.

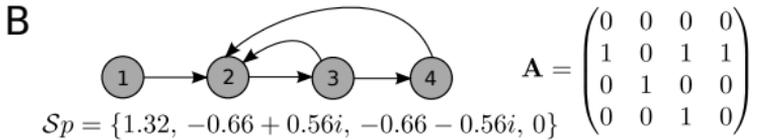
Nous n'allons pas démontrer le théorème de Perron-Frobenius. En revanche, afin de le rendre un peu plus intuitif, nous allons tenter de voir – de façon informelle – ce qu'apportent les hypothèses d'irréductibilité et d'apériodicité. Pour cela, nous allons explorer ce qui peut se produire lorsque ces hypothèses ne sont pas vérifiées. Commençons par l'irréductibilité.

Question : Donner différents exemples de matrices réductibles et donner la conséquence sur les conclusions du théorème de Perron-Frobenius.

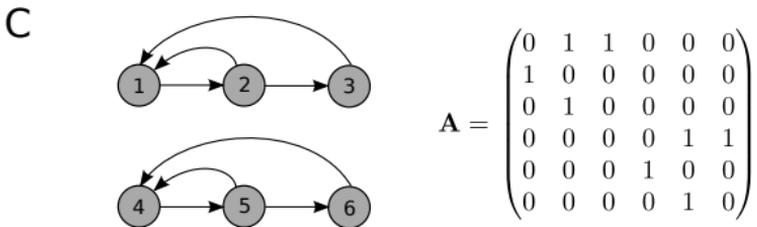
La figure 4 donne des exemples de modèles réductibles et apériodiques. Comme on peut le voir, divers comportements sont possibles. Par exemple, dans la figure 4.A, l'irréductibilité n'est pas vérifiée à cause de la présence d'une classe "puits", à laquelle on peut arriver mais dont on ne peut pas repartir (classe n° 4. Il s'agit sans doute de la cause la plus fréquente de réductibilité, car elle peut correspondre à présence d'une classe post-reproductive, c'est-à-dire d'individus qui ne se reproduiront plus –



$$w = \begin{pmatrix} 0.36 \\ 0.27 \\ 0.21 \\ 0.16 \end{pmatrix} \quad v = (1.13 \quad 1.50 \quad 0.86 \quad 0)$$



$$w = \begin{pmatrix} 0 \\ 0.43 \\ 0.32 \\ 0.25 \end{pmatrix} \quad v = (0.72 \quad 0.96 \quad 1.27 \quad 0.72)$$



$$Sp = \{1.32, -0.66 + 0.56i, -0.66 - 0.56i, \}$$

$$w = \begin{pmatrix} 0.43 \\ 0.32 \\ 0.25 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0.43 \\ 0.32 \\ 0.25 \end{pmatrix} \quad v = (0.96 \quad 1.27 \quad 0.72 \quad 0 \quad 0 \quad 0)$$

$$\text{et} \quad (0 \quad 0 \quad 0 \quad 0.96 \quad 1.27 \quad 0.72)$$

FIGURE 4 – Exemples de modèles réductibles (mais apériodiques). L'ensemble des valeurs propres (spectre), ainsi que les vecteurs propres associés à la valeur propre de plus grand module sont représentés. Cf texte principal pour plus de détails.

et n'ont donc pas d'impact sur la dynamique de la population – mais qu'on choisit tout de même d'inclure dans le modèle car on veut suivre leur effectif. L'absence d'impact sur la dynamique de la population des classes post-reproductives se traduit par une valeur reproductive nulle, comme en témoigne $v_4 = 0$ dans l'exemple présenté. Néanmoins, en pratique, les classes post-reproductives sont peu courantes dans la nature, et il y a une bonne raison à cela : puisque ces individus n'ont plus la possibilité de se reproduire, ceux qui survivraient plus longtemps ne seraient pas pour autant avantagés du point de vue de la sélection naturelle. Pire, tout individu qui allouerait ses ressources de manière à atteindre une classe post-reproductive serait désavantagé par rapport à un individu qui investirait les mêmes ressources dans la production d'individus et mourrait juste avant d'atteindre une classe post-reproductive. Ainsi, la présence d'une classe post-reproductive est soit une contingence, soit le résultat d'un effet des individus de la classe post-reproductive sur la dynamique de la population (cf ménopauses des femmes et hypothèses des grand-mères).

La figure 4.B présente la situation opposée, à savoir la présence d'une classe “source”, dont on peut

partir mais à laquelle on ne peut pas revenir (classe n° 1). Cette fois, la valeur reproductive de cette classe n'est pas nulle ($v_1 \neq 0$), puisque les individus initialement présent dans cette classes pourront contribuer à la dynamique de la population. En revanche, il est clair que cette classe va se “vider”, puisqu'aucun flux entrant ne permet de la repeupler. C'est ce que traduit $w_1 = 0$. Cette situation est très rare. On pourrait éventuellement l'envisager, par exemple, dans un modèle de conservation où les animaux réintroduits par l'homme auraient (soit dans un premier temps, soit tout au long de leur vie) des traits démographiques différents des animaux ayant grandi dans la nature, et où le modèle serait utilisé pour suivre la dynamique d'une population suite à une réintroduction.

La figure 4.C représente la situation où le graphe du cycle de vie contient deux composantes fortement connexes complètement isolées. Comme on aurait pu s'y attendre, on voit que le sous-espace associé à la valeur propre dominante n'est pas de dimension 1. En effet, selon le sous-modèle dans lequel la population est initialement placée (classes n° 1, 2 et 3 ou n° 4, 5 et 6), l'un des sous-modèle restera vide et la structure de la population convergera vers l'un des

vecteurs propres à droites associés à la valeur propre dominante – tout se passe comme si les classes de la composante fortement connexe dans laquelle aucun individu n'a été introduit n'existaient pas. A noter que même si l'on introduit des individus dans les deux composantes fortement connexes, à moins que les sous-modèles correspondant n'aient exactement la même valeur propre dominante (comme c'est le cas ici) – auquel cas la structure de la population tendra vers une combinaison linéaire des deux vecteurs propres associés à la valeur propre dominante – la structure de la population convergera vers le vecteur propre obtenu en complétant celui du sous-modèle de plus grande valeur propre dominante par des zéros pour les classes de l'autre sous-modèle : en effet, la sous-population ayant le taux de croissance le plus élevé grandira plus vite que l'autre sous-population. Les croissances étant géométriques, l'effectif de la population grandissant le moins vite deviendra négligeable devant celui de l'autre sous-population. Ce type de modèle, à composantes fortement connexes isolées, ne se rencontre jamais, même dans les modèles à méta-population, car il est alors équivalent et bien plus simple d'étudier les deux modèles séparément.

Enfin, des situations plus complexes que celles présentées dans la figure 4 peuvent parfois survenir. Par exemple, dans une dynamique “source-puit”, on peut avoir un modèle de méta-population où l’une des sous-populations a un taux de croissance supérieur à 1 et envoie des individus à une autre sous-population qui, livrée à elle-même, aurait un taux de croissance inférieur à 1 et s’éteindrait. Il est alors difficile de formuler des résultats généraux. Néanmoins, ces modèles sont très rares, et même lorsqu’ils se présentent, on peut toujours se ramener à un modèle habituel en rajoutant un taux de migration extrêmement faible depuis le “puits” vers la “source (ce qui n’est pas forcément moins réaliste, au contraire).

En conclusion, lorsque l’irréductibilité n’est pas vérifiée, des situations très diverses peuvent se produire. Nous avons vu des situations où les vecteurs propres dominants existaient mais n’étaient pas strictement positifs (figures 4.A et 4.B) mais aussi des situations où les sous-espaces propres associés à la valeur propre dominante n’étaient pas de dimension 1 (figure 4). On en déduit que l’irréductibilité est vraiment une hypothèse importante du théorème de Perron-Frobenius. Intéressons-nous maintenant

à l'apériodicité et, paradoxalement, commençons par nous intéresser de plus près au lien entre irréductibilité et primitivité, de manière à voir ce qu'apporte l'apériodicité.

Question : Donner une caractérisation matricielle de l'irréductibilité.

Il faut bien voir que l'entrée (i, j) de \mathbf{A}^k quantifie le flux de j vers i en k intervalles de temps exactement (i.e. la quantité de matière qui arrive en i à $t = k$ et dont l'origine peut-être attribuée à la matière se trouvant en j à $t = 0$ est égale à $(\mathbf{A}^k)_{ij}n_j(0)$). A partir de là, il devient évident que, puisqu'on ne travaille qu'avec des flux positifs et que l'existence d'un chemin est équivalente à l'existence d'un flux, une caractérisation matricielle de l'irréductibilité est :

$$\forall(i, j), \quad \exists k \text{ tq } (\mathbf{A}^k)_{ij} > 0.$$

A noter que la propriété ci-dessus est différente de celle garantissant la primitivité :

$$\exists k \text{ tq } \forall(i, j), \quad (\mathbf{A}^k)_{ij} > 0.$$

Dans le cas de l'irréductibilité, on ne demande que l'existence d'un chemin (de longueur quelconque) de

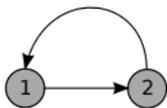
j à i , tandis que la primitivité requiert l'existence d'une longueur de chemin telle que pour tous couple (i, j) de sommets il existe un chemin (repassant éventuellement plusieurs fois par une même arête) de cette longueur exactement allant de i à j . Cela permet de se former une première intuition sur ce qu'apporte l'apériodicité à l'irréductibilité.

Question : Donner un exemple de matrice irréductible périodique et donner la conséquence sur les conclusions du théorème de Perron-Frobenius.

Les figures 5.A et 5.B sont des exemples de modèles irréductibles périodiques (de période égale à 2). La figure 5.C n'est donnée qu'à titre de comparaison avec la figure 5.B.

On voit qu'il se produit un phénomène nouveau : il existe des valeurs propres de même module que la valeur propre "dominante" : -1 en (A) et -1.27 en (B). En fait, il s'agit d'un phénomène général et non d'un hasard dû au choix des modèles : on peut montrer que lorsqu'une matrice \mathbf{M} est positive et irréductible, le réel égal au maximum du module de ses valeurs propres (appelé rayon spectral et souvent noté $\rho(\mathbf{M})$) est une valeur propre de \mathbf{M} de multiplicité 1 et qu'il

A



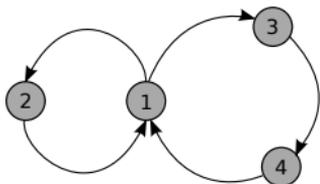
$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\mathcal{S}p = \{1, -1\}$$

$$\mathbf{v} = (1 \quad 1)$$

$$\mathbf{w} = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$

C

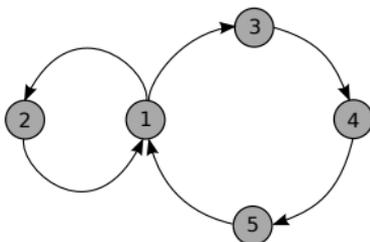


$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\mathcal{S}p = \{1.32, -0.66 + 0.56i, -0.66 - 0.56i, 0\}$$

$$\mathbf{v} = (1.27 \quad 0.96 \quad 0.72 \quad 0.96) \quad \mathbf{w} = \begin{pmatrix} 0.32 \\ 0.25 \\ 0.25 \\ 0.19 \end{pmatrix}$$

B



$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\mathcal{S}p = \{1.27, -1.27, 0.79i, -0.79i, 0\}$$

$$\mathbf{v}_1 = (1.33 \quad 1.05 \quad 0.65 \quad 0.82 \quad 1.05) \quad \mathbf{w}_1 = \begin{pmatrix} 0.27 \\ 0.21 \\ 0.21 \\ 0.17 \\ 0.13 \end{pmatrix} \quad \mathbf{w}_2 = \begin{pmatrix} 0.27 \\ -0.21 \\ -0.21 \\ 0.17 \\ -0.13 \end{pmatrix}$$

$$\mathbf{v}_2 = (1.33 \quad -1.05 \quad -0.65 \quad 0.82 \quad -1.05)$$

existe $h - 1$ autres valeurs propres de \mathbf{M} de même module et également de multiplicité 1, où h est la période de la matrice (PGCD de la longueur des cycles du graphe associé). \mathbf{M} admet donc h valeurs propres de module $\rho(\mathbf{M})$, qui sont données par les racines h -ièmes de l'unité multipliées par $\rho(\mathbf{M})$, soit les complexes :

$$\rho(\mathbf{M}) e^{\frac{2ik\pi}{h}}, \quad k \in \{0, \dots, h - 1\}$$

Nous ne justifierons pas ce résultat, mais on peut s'en convaincre en le vérifiant sur des exemples simples, comme les matrices de la forme :

$$\begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

Enfin, nous ne le détaillerons pas plus mais ce genre de modèle conduit à des dynamiques exhibant une composante cyclique de période h .

3.1.4 Bilan sur le régime asymptotique

Cette section étant un peu longue, nous la terminons par un résumé des divers résultats que nous avons démontrés ou admis concernant le régime asymptotique :

La grande majorité des matrices de projection utilisés dans les MMP sont primitives, c'est-à-dire positives, irréductibles et aperiodiques. Dans ce cas, Le théorème de Perron-Frobenius garantit l'existence d'une valeur propre réelle et positive, plus grande en module que toutes les autres, appelée valeur propre dominante. De plus, les sous-espaces propres (à droite et à gauche) associés à cette valeur propre sont de dimension 1 et les vecteurs propres engendrant ces sous-espaces propres peuvent être choisis de manière à avoir toutes leurs composantes strictement positives. Dans ce cas,

- La valeur propre dominante, λ , est interprétée comme le taux de croissance de la population. Si $\lambda > 1$, la population croît géométriquement. Si $\lambda < 1$, elle décroît jusqu'au moment où la taille de la population devient trop faible pour que la population puisse être décrite par un modèle matriciel de population. Enfin, si

$\lambda = 1$, la population finit par se stabiliser. Dans la pratique, on s'attend souvent à trouver λ proche de 1, ce qui traduit le fait qu'on s'attend à ce que les population naturelles ne soient généralement pas trop loin de l'équilibre démographique (sauf perturbation récente).

- Le vecteur propre à droite associé à λ , \mathbf{w} , est interprété comme la distribution stable de la population : ses entrées sont proportionnelles à la fraction d'individus dans chaque classe lorsque la population a atteint son régime asymptotique.
- Le vecteur propre à gauche associé à λ , \mathbf{v} , peut-être interprété comme représentant les valeurs reproductives des individus de chaque classe, i.e. leur contribution relative à la dynamique de la population.
- L'équation à retenir est :

$$\mathbf{n}(t) \rightarrow \mathbf{v}\mathbf{n}(0) \lambda^t \mathbf{w},$$

où \mathbf{v} a été choisi de manière à ce que $\mathbf{v}\mathbf{w} = 1$.

Certains modèles n'ont pas des matrices de projection primitives. Lorsque la matrice est réductible,

cela peut conduire à différents comportements. Le cas le plus fréquent correspond à la présence d'une classe post-reproductive, qui ne pose pas de problème (les interprétations ci-dessus restent valables – la seule différence avec un modèle habituel sera que la composante de \mathbf{v} correspondant à la classe post-reproductive sera nulle).

Lorsque la matrice de projection est irréductible mais n'est pas aperiodique, elle admet h valeurs propres de module égal à celui de la valeur propre dominante, où h est le PGCD des longueurs des cycles du graphe du cycle de vie. La dynamique du modèle exhibe alors une composante "cyclique" de période h .

Une dernière remarque avant de passer à l'étude du régime transitoire : j'attire l'attention du lecteur sur le fait que lorsque la matrice de projection de la population est primitive, le comportement asymptotique de la population ne dépend pas des conditions initiales. Il s'agit d'un résultat très puissant. Au passage, Caswell fait usage du terme "ergodicité" pour faire référence à cette propriété. Bien entendu, cet usage du mot n'est pas injustifié. Mais en général, le terme "ergodique" est employé pour indiquer qu'on obtient la même moyenne en suivant

un petit échantillon pendant une longue période de temps qu'en suivant un grand échantillon pendant une courte période de temps (par exemple, on pourrait dire d'une population qu'elle est ergodique pour dire qu'on obtiendrait la même moyenne en suivant un unique individu pendant plusieurs années qu'en suivant plusieurs d'individus pendant une année).

3.2 Régime transitoire

Question : Quel paramètre simple pourrait quantifier la vitesse de convergence du modèle vers le régime asymptotique ?

Comme l'illustre la figure 3, la vitesse de convergence vers le régime asymptotique va essentiellement dépendre du rapport entre la valeur propre dominante et le module des autres valeurs propres. C'est pourquoi on définit le coefficient d'amortissement :

$$\alpha = \frac{\lambda_1}{|\lambda_2|},$$

où λ_1 est la valeur propre dominante et λ_2 une des valeurs propres dont le module est le plus grand après

celui de λ_1 . Plus α est grand, plus l'effet des conditions initiales sera gommé rapidement. On peut bien sûr définir d'autres coefficients d'amortissement en considérant les rapports $\lambda_1/|\lambda_k|$. Enfin, nous ne les présenterons pas ici, mais il existe de nombreux autres indicateurs de la vitesse de convergence vers le régime asymptotique.

Question : Quels sont les effets des différentes valeurs propres sur le régime transitoire ?

La présentation qui a été faite de cette question lors du GT manquait de rigueur. Il me semble que c'est également le cas de l'ouvrage de Caswell, et même que son raisonnement est inexact. Je vais donc tenter d'être plus rigoureux, mais cette partie reste à lire avec un oeil critique.

Jusqu'à maintenant, nous avons raisonné sur les éléments propres dominants de \mathbf{A} , qui sont réels. Mais il est probable que \mathbf{A} admette également des valeurs propres complexes, auxquelles seront associés des vecteurs propres complexes. Si l'on veut interpréter les effets des différentes valeurs propres de \mathbf{A} en étant aussi général que possible, il nous faut donc travailler dans \mathbb{C}^n .

Commençons par choisir une base adaptée. Soit $(\mathbf{w}^{(1)}, \dots, \mathbf{w}^{(m)})$ une famille libre de vecteurs propres de \mathbf{A} (au passage, on sait qu'une telle famille existe puisque \mathbf{A} admet un vecteur propre dominant). On peut la compléter en une base $(\mathbf{w}^{(1)}, \dots, \mathbf{w}^{(m)}, \mathbf{x}^{(m+1)}, \dots, \mathbf{x}^{(n)})$ de \mathbb{C}^n . Dans cette base, le vecteur population peut être écrit :

$$\mathbf{n}(t) = c_1 \lambda_1^t \mathbf{w}^{(1)} + \dots + c_m \lambda_m^t \mathbf{w}^{(m)} + c_{m+1} \mathbf{A}^t \mathbf{x}^{(m+1)} + \dots + c_n \mathbf{A}^t \mathbf{x}^{(n)}$$

Cette écriture permet d'isoler les contributions des différentes valeurs propres de \mathbf{A} à la composition de la population. Mais avant de discuter cela, faisons quelques remarques quant au choix des vecteurs $\mathbf{w}^{(k)}$:

Tout d'abord, les vecteurs $\mathbf{w}^{(k)}$ associés à des valeurs propres réelles peuvent être choisis dans \mathbb{R}^n . En effet, soit λ_k une valeur propre réelle et $\mathbf{w}^{(k)}$ un vecteur propre associé. On peut écrire $\mathbf{w}^{(k)} = \Re(\mathbf{w}^{(k)}) + i \Im(\mathbf{w}^{(k)})$. Si $\Re(\mathbf{w}^{(k)})$ est le vecteur nul, il suffit de multiplier $\mathbf{w}^{(k)}$ par le scalaire i (on travaille avec un \mathbb{C} -espace vectoriel) pour obtenir un vecteur propre réel. Sinon, on a

$$\mathbf{A} \mathbf{w}^{(k)} = \lambda_k \mathbf{w}^{(k)} = \lambda_k \Re(\mathbf{w}^{(k)}) + i \lambda_k \Im(\mathbf{w}^{(k)})$$

et

$$\mathbf{A}\mathbf{w}^{(k)} = \mathbf{A}(\Re(\mathbf{w}^{(k)}) + i\Im(\mathbf{w}^{(k)})) = \mathbf{A}\Re(\mathbf{w}^{(k)}) + i\mathbf{A}\Im(\mathbf{w}^{(k)}).$$

Puisque λ_k et \mathbf{A} sont réels, par identification : $\mathbf{A}\Re(\mathbf{w}^{(k)}) = \lambda_k\Re(\mathbf{w}^{(k)})$. On a donc bien trouvé un vecteur propre réel de \mathbf{A} .

Ensuite, rappelons-nous que les valeurs propres complexes viennent par paires de valeurs propres complexes conjuguées. Or, si $\mathbf{w}^{(k)}$ est un vecteur propre associé à la valeur propre λ_k , alors $\overline{\mathbf{w}^{(k)}}$ est un vecteur propre associé à $\overline{\lambda_k}$. On peut donc choisir les vecteurs propres de la base de \mathbb{C}^n de manière à ce que les vecteurs propres non-réels y soient présents par paires de vecteurs propres conjugués. On admet que les vecteurs $\mathbf{x}^{(m)}, \dots, \mathbf{x}^{(n)}$ qui servent à compléter la base de \mathbb{C}^n lorsque \mathbf{A} n'est pas diagonalisable peuvent également être choisis soit réels, soit de manière à être présents par paires de vecteurs conjugués.

Revenons alors sur l'expression du vecteur population :

$$\mathbf{n}(t) = c_1 \lambda_1^t \mathbf{w}^{(1)} + \dots + c_m \lambda_m^t \mathbf{w}^{(m)} + c_{m+1} \mathbf{A}^t \mathbf{x}^{(m+1)} + \dots + c_n \mathbf{A}^t \mathbf{x}^{(n)}$$

On admet que si $\lambda_k \in \mathbb{R}$, $c_k \in \mathbb{R}$ et que pour les valeurs propres non-réelles, si $\lambda_{k+1} = \overline{\lambda_k}$, alors $c_{k+1} =$

$\overline{c_k}$. Ceci permet de de s'assurer que tous les termes imaginaires s'annulent (ce qui est nécessaire puisque $\mathbf{n}(t)$ est réel). En effet, on a alors :

$$\begin{aligned} c_k \lambda_k^t \mathbf{w}^{(k)} + c_{k+1} \lambda_{k+1}^t \mathbf{w}^{(k+1)} &= c_k \lambda_k^t \mathbf{w}^{(k)} + \overline{c_k} \overline{\lambda_k}^t \overline{\mathbf{w}^{(k)}} \\ &= c_k \lambda_k^t \mathbf{w}^{(k)} + \overline{c_k \lambda_k^t \mathbf{w}^{(k)}} \\ &= 2 \Re(c_k \lambda_k^t \mathbf{w}^{(k)}) \end{aligned}$$

On peut isoler la contribution de la valeur propre λ_k :

$$c_k \lambda_k^t \mathbf{w}^{(k)} .$$

Cette écriture n'induit pas de perte de généralité dans le cas où le sous-espace propre associé à λ_k est de dimension d : dans ce cas, λ_k^t apparaît d fois dans l'écriture de $\mathbf{n}(t)$ mais on peut la factoriser pour obtenir un terme de la forme $c_k \lambda_k^t \mathbf{w}^{(k)}$, où $c_k \mathbf{w}^{(k)}$ est la somme des $c_{k_i} \mathbf{w}^{(k_i)}$ associés à la valeur propre commune λ_k .

Clairement, il a une première distinction à faire selon que $|\lambda_k|$ est plus grand ou pas que 1. Si $|\lambda_k| < 1$, cette contribution va décroître rapidement. Ce n'est pas le cas si $|\lambda_k| \geq 1$; mais comme la contribution de la valeur propre λ_1 s'accroîtra encore plus rapidement, cette contribution deviendra tout de même

négligeable.

Ensuite, une seconde distinction existe selon que λ_k est réelle ou non. Si λ_k est réelle, on a deux possibilités : soit $\lambda_k > 0$, auquel cas son effet sera monotone, c'est-à-dire qu'il ne fera qu'augmenter ou diminuer. En revanche, si $\lambda_k < 0$, le vecteur $c_k \lambda_k^t \mathbf{w}^{(k)}$ changera de sens à chaque pas de temps. On aura donc un effet 2-périodique.

Enfin, si λ_k est complexe, la situation l'est également. On a vu que λ_k serait couplée à une autre valeur propre $\lambda_{k+1} = \overline{\lambda_k}$, et que leur contribution serait

$$2 \Re(c_k \lambda_k^t \mathbf{w}^{(k)}) .$$

En écrivant λ_k sous la forme $|\lambda_k|(\cos(\theta) + i \sin(\theta))$, cela devient

$$2 |\lambda_k|^t (\cos(\theta t) \Re(c_k \mathbf{w}^{(k)}) - \sin(\theta t) \Im(c_k \mathbf{w}^{(k)})) .$$

On voit qu'en plus du côté croissant ou décroissant de cette contribution lié à $|\lambda_k|$, s'ajoute une composante oscillatoire de période $\frac{2\pi}{\theta}$ liée à l'argument de λ_k . En remarquant que $\Re(c_k w_i^{(k)}) = |c_k w_i^{(k)}| \cos(\psi)$ et $\Im(c_k w_i^{(k)}) = |c_k w_i^{(k)}| \sin(\psi)$, où $\psi = \text{Arg}(c_k w_i^{(k)})$, et en utilisant l'identité $\cos(a) \cos(b) - \sin(a) \sin(b) =$

$\cos(a + b)$, on peut également écrire que la contribution de λ_k à $n_i(t)$ est

$$2|c_k w_i^{(k)}| \times |\lambda_k|^t \cos(\theta t + \psi).$$

Mais cela ne sert à rien sinon à nous rendre contents d'avoir utilisé une formule trigo.

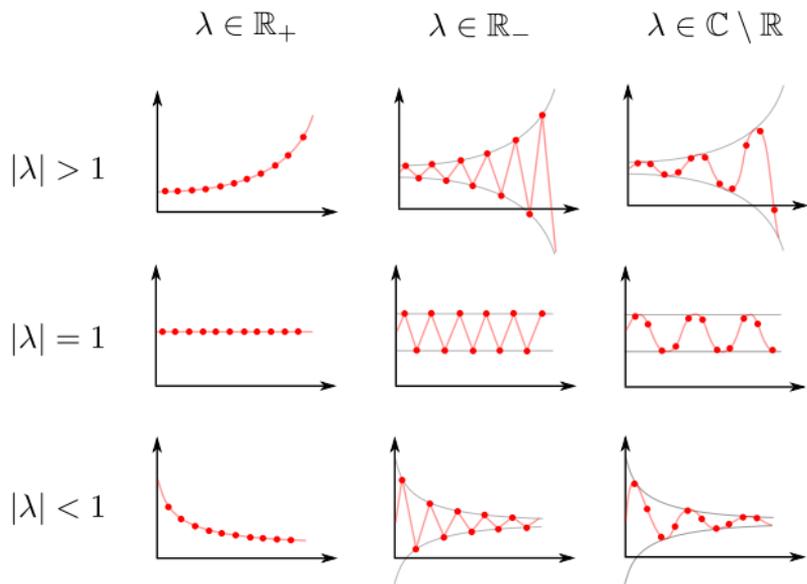


FIGURE 6 — Récapitulatif des effets des différentes valeurs propres en fonction de leur nature. Le temps est représenté en abscisse et la contribution en ordonnée (attention, le zéro ne se trouve pas au niveau de l'axe des abscisses. Les contributions ne sont pas positives et changent de signe lorsque $\lambda \notin \mathbb{R}_+$).

La figure 6 résume les différents types de contributions des valeurs propres. Cela permet de se for-

mer une idée de la dynamique du système à partir du spectre de \mathbf{A} . Mais attention, il ne faut pas perdre de vue que bien que les contributions des différentes valeurs propres s'ajoutent, elles deviennent négligeables devant celle de la valeur propre de plus grand module (et ce d'autant plus vite que le ratio $\frac{\lambda_1}{|\lambda_k|}$ est élevé).

3.3 Sensibilité aux paramètres

Nous avons obtenu de nombreux résultats, dont la plupart se prêtent bien à des calculs numériques. Néanmoins, les matrices utilisées dans la pratique ne correspondent jamais à la “vraie” (en supposant qu'elle existe...) matrice décrivant la dynamique de la population, car les erreurs de mesure sont inévitables. On peut donc se demander à quel point les différents résultats que nous avons présentés seraient modifiés si les éléments servant à la construction du modèle étaient légèrement différents. Ces questions sont souvent regroupées sous le terme “analyse de sensibilité”, même si le terme “sensibilité” désigne un objet mathématique bien précis.

Question : Quel est l'indicateur le plus naturel pour exprimer la sensibilité des résultats aux paramètres ?

Tenter d'en trouver une expression au moyen d'un calcul "à la physicienne".

L'idée la plus naturelle est d'utiliser une dérivée. Puisqu'on a affaire à des fonctions de plusieurs variables et qu'il est souvent souhaitable d'isoler l'effet du changement d'un paramètre, on utilise des dérivées partielles. On définit alors la *sensibilité* de f au paramètre α :

$$s_f(\alpha) = \frac{\partial f}{\partial \alpha}$$

On peut bien sûr appliquer cette définition à n'importe quel descripteur f et à n'importe quel paramètre α , mais certains choix semblent plus pertinents. Par exemple, puisque le descripteur le plus important d'une population est probablement son taux de croissance λ , et puisque la construction de tout modèle matriciel conduit à l'obtention d'une matrice de projection de la population \mathbf{A} , il est naturel de s'intéresser à des indicateurs comme

$$s_\lambda(a_{ij}) = \frac{\partial \lambda}{\partial a_{ij}},$$

qu'on appelle souvent "les sensibilités du modèle" sans précision supplémentaire.

Il s'avère qu'il existe une formule simple permettant de calculer les sensibilités. Il s'agit d'un résultat classique d'analyse de perturbation de matrices. Nous n'allons néanmoins pas le démontrer ici rigoureusement mais nous contenter d'un calcul "intuitif" n'ayant pour but que de montrer que ce résultat ne sort pas de nulle part. Supposons que la matrice \mathbf{A} subisse une perturbation $d\mathbf{A}$ ($d\mathbf{A} = (da_{ij})$). En conséquence de cela, la valeur propre dominante et le vecteur propre dominant associé de \mathbf{A} sont également perturbés – de perturbations respectives $d\lambda$ et $d\mathbf{w}$. On obtient donc une nouvelle matrice $\mathbf{A} + d\mathbf{A}$ de valeur propre dominante $\lambda + d\lambda$ et de vecteur propre dominant $\mathbf{w} + d\mathbf{w}$. On a donc :

$$(\mathbf{A} + d\mathbf{A})(\mathbf{w} + d\mathbf{w}) = (\lambda + d\lambda)(\mathbf{w} + d\mathbf{w})$$

$$\mathbf{A}\mathbf{w} + \mathbf{A}d\mathbf{w} + d\mathbf{A}\mathbf{w} + d\mathbf{A}d\mathbf{w} = \lambda\mathbf{w} + \lambda d\mathbf{w} + d\lambda\mathbf{w} + d\lambda d\mathbf{w}$$

soit, en négligeant les termes de second ordre et en utilisant le fait que $\mathbf{A}\mathbf{w} = \lambda\mathbf{w}$,

$$\mathbf{A}d\mathbf{w} + d\mathbf{A}\mathbf{w} = \lambda d\mathbf{w} + d\lambda\mathbf{w}$$

On multiplie maintenant les deux membres de l'égalité par \mathbf{v} :

$$\mathbf{v}\mathbf{A}d\mathbf{w} + \mathbf{v}d\mathbf{A}\mathbf{w} = \lambda\mathbf{v}d\mathbf{w} + d\lambda\mathbf{v}\mathbf{w}$$

Enfin, on utilise le fait que $\mathbf{vA} = \lambda\mathbf{A}$ et on réarrange pour obtenir :

$$d\lambda = \frac{\mathbf{vdAw}}{\mathbf{vw}}.$$

Cette formule permet d'obtenir simplement le changement de λ en réponse à un changement dans les entrées de \mathbf{A} . Si l'on veut quantifier la réponse à un changement dans une unique entrée a_{ij} de \mathbf{A} , on peut utiliser le fait que dans ce cas toutes les entrées de $d\mathbf{A}$ sont nulles à l'exception de l'entrée (i, j) , qui vaut da_{ij} . Ceci permet de simplifier \mathbf{vdAw} en $v_i(da_{ij})w_j$. On obtient alors

$$\frac{d\lambda}{da_{ij}} = \frac{v_i w_j}{\mathbf{vw}}.$$

Puisque cette expression a été obtenue en ne faisant varier qu'un seul a_{ij} à la fois (c'est-à-dire en supposant $\forall (k, l) \neq (i, j), \frac{\partial a_{ij}}{\partial a_{kl}} = 0$), on a $\frac{d\lambda}{da_{ij}} = \frac{\partial \lambda}{\partial a_{ij}}$ et on peut donc écrire

$$\frac{\partial \lambda}{\partial a_{ij}} = \frac{v_i w_j}{\mathbf{vw}}.$$

Question : Quels sont les problèmes des sensibilités que nous venons de définir en tant qu'indicateurs ? Comment pourrait-on les résoudre simplement ?

Les sensibilités sont de précieux indicateurs, mais elles ont une limitation majeure : elles ne rendent pas compte de la propension des entrées de la matrice à changer. Par exemple, l'amplitude (absolue) des variations dans une fertilité élevée sera sans doute plus grande que celle dans une probabilité de survie faible. Ainsi, un même da_{ij} n'aura pas du tout la même signification biologique selon l'entrée concernée. Il en résulte qu'on ne peut pas pas comparer les sensibilités. En effet, une sensibilité très élevée peut porter sur un paramètre qui, concrètement, ne changera que très peu (en conséquence de quoi ce paramètre ne jouera pas un rôle important dans les perturbation du descripteur étudié). Et inversement, une sensibilité très faible peut porter sur un paramètre susceptible de varier énormément (et donc jouant un rôle important dans les perturbation du descripteur).

Un bon exemple de cela est donné par les entrées nulles de la matrice de projection : certaines de ces entrées peuvent être qualifiées de “zéros structurels”, au sens qu'on sait qu'elles sont exactement égales à zéro et qu'il ne peut pas en être autrement. C'est le cas de transitions qui permettrait à des individus de “rajeunir” dans des modèles en classe d'âge. Bien que les entrées correspondantes ne puissent pas varier

et ne puissent donc pas perturber les indicateurs du modèle, elles peuvent avoir des sensibilités très élevées.

Un moyen simple de contourner ce problème est de quantifier l'impact d'un changement *relatif* dans le paramètre. Autrement dit, d'utiliser

$$\frac{\Delta\%f}{\Delta\%\alpha} \quad \text{plutôt que} \quad \frac{\Delta f}{\Delta\alpha},$$

l'idée derrière cela étant qu'il est plus pertinent de comparer les conséquence d'un changement de $x\%$ dans deux entrées que celles d'un changement de x . En effet, si une fertilité est égale à 200 (cas plausibles pour les espèces comme l'huître ou certaines plantes, qui produisent un nombre énorme de descendants dont une fraction infime atteindra l'âge adulte), un changement de $\pm 10 = \pm 0.05\%$ est tout à fait possible. Alors que seul un changement de $\pm 0.05\%$ (mais certainement pas de ± 10) est plausible pour une probabilité de survie. A noter qu'on aurait très bien pu se contenter d'utiliser $\frac{\Delta f}{\Delta\%\alpha}$ plutôt que $\frac{\Delta\%f}{\Delta\%\alpha}$ – cela suffit à pouvoir comparer l'importance des perturbations de deux paramètres α et β sur un même descripteur f . Mais comme nous le verrons cette normalisation présente d'autres avantages.

Cela nous conduit à introduire la quantité

$$e_{\lambda}(a_{ij}) = \frac{a_{ij}}{\lambda} \frac{\lambda}{a_{ij}},$$

appelée élasticité de λ à a_{ij} (au passage, cette définition est valide puisqu'on a bien $\lambda > 0$ – à noter qu'on a également $e_{\lambda}(a_{ij}) = \frac{\partial \log \lambda}{\partial \log a_{ij}}$). On peut alors remarquer que puisque

$$e_{\lambda}(a_{ij}) = \frac{v_i a_{ij} w_j}{\lambda \mathbf{v} \mathbf{w}},$$

on en déduit

$$\sum_{i,j} e_{\lambda}(a_{ij}) = \sum_{i,j} \frac{v_i a_{ij} w_j}{\lambda \mathbf{v} \mathbf{w}} = \frac{\mathbf{v} \mathbf{A} \mathbf{w}}{\lambda \mathbf{v} \mathbf{w}} = 1.$$

En réalité, il n'est pas nécessaire de disposer d'une expression de $e_{\lambda}(a_{ij})$ pour établir ce résultat, qui est une conséquence directe du théorème d'Euler pour les fonction homogènes de degré 1.

[http://fr.wikipedia.org/wiki/Théorème_d'Euler_\(fonctions_de_plusieurs_variables\)](http://fr.wikipedia.org/wiki/Théorème_d'Euler_(fonctions_de_plusieurs_variables))

On peut se servir du fait que les élasticités somment à 1 pour les interpréter comme la part de contribution d'une transition à la dynamique de la population. Une autre interprétation, plus rigoureuse mais

que nous ne détaillons pas, est de remarquer qu'elles correspondent à une distribution de probabilité.

Il existe d'autres indicateurs que les sensibilités et les élasticités pour quantifier l'impact de perturbations ou d'incertitudes dans les paramètres du modèle sur des descripteurs, mais les sensibilités et les élasticités sont de loin les plus simples et les plus utilisés.

4 Conclusion

Nous avons vu les principaux résultats concernant les modèles matriciels de population. Néanmoins, certains résultats importants n'ont pas été mentionnés (équation caractéristique, résultats propres aux modèles les plus courants, etc) et nous n'avons pas toujours détaillé l'aspect biologique autant qu'il aurait pu l'être. Tout ceci justifie le mot d'avertissement dans l'introduction disant que ce document ne prétend pas être une introduction aux modèles matriciels de population. De même, le niveau de rigueur et de détail varie d'une section à l'autre – par exemple, la section sur le théorème de Perron-Frobenius ou le régime transitoire sont extrêmement incomplètes,

et le lecteur doit bien l'avoir à l'esprit.

Enfin, pour conclure ce document, on peut s'interroger sur l'*utilité* des modèles matriciels de population. En effet, les hypothèses simplistes de ces modèles conduisent à certains résultats aberrants. Par exemple, l'absence de densité-dépendance conduit à une croissance (ou décroissance) exponentielle de la population, dont on sait qu'elle ne peut pas être maintenue même à moyen terme. Finalement, les modèles matriciels que nous avons construit ne sont donc pas très différents du modèle de Malthus dont nous étions partis... Des modèles aussi simplistes ont-ils un quelconque intérêt ?

Outre la prédiction à court terme, les modèles matriciels de population présentent bien un intérêt. Je vais l'illustrer en reprenant (sans vergogne) l'analogie de Caswell avec le compteur de vitesse : un compteur de vitesse ne sert pas à prédire à quelle heure on arrivera. Il n'est pas adapté pour cela car on *sait* qu'on ne pourra pas rester à 110 jusqu'à destination, car il faudra forcément sortir de l'autoroute à un moment... Sans compter tous les imprévus comme les pannes, dont on ne peut pas prédire le moment exact mais dont on sait qu'ils se produiront sur une

échelle de temps suffisamment longue. Ce que fait un compteur de vitesse, c'est qu'il nous fournit une information *sur le présent*. Cette information est utile, mais pour des questions différentes : par exemple, savoir qu'étant donné la vitesse à laquelle on va, il nous faudra telle distance pour nous arrêter, et qu'on aura telle probabilité de bousiller la voiture si on se prend un sanglier, ou de tuer le gars dans lequel on rentrera.

Un modèle matriciel de population est exactement pareil. On sait que les probabilités de survies et les fertilités finiront par changer, et donc qu'on ne peut pas tenter de prédire ce que sera la taille de la population dans 10 ans (d'autant que même si on ne peut pas forcément les prédire, on sait bien que des aléas environnementaux finiront par se produire). En revanche, on peut utiliser les indicateurs fournis par les modèles matriciels pour estimer l'état actuel de la population : savoir si elle est en train d'exploser ou de disparaître, repérer les individus les plus importants pour la démographie (et donc qu'il faudra cibler en priorité si l'on veut manipuler la dynamique de la population), etc...

Références

- H. Caswell. *Matrix Population Models : Construction, Analysis, and Interpretation*. Sinauer Associates, 2nd edition, 2001.
- P. H. Leslie. On the use of matrices in certain population mathematics. *Biometrika*, 33, 1945.