



HAL
open science

Convergence de la variabilité dans les modèles polygéniques gaussiens

S Lessard, S Mahdi

► **To cite this version:**

S Lessard, S Mahdi. Convergence de la variabilité dans les modèles polygéniques gaussiens. *Genetics Selection Evolution*, 1995, 27 (5), pp.395-421. hal-00894105

HAL Id: hal-00894105

<https://hal.science/hal-00894105>

Submitted on 11 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Convergence de la variabilité dans les modèles polygéniques gaussiens

S Lessard, S Mahdi

Université de Montréal, département de mathématiques et de statistique,
CP 6128, succursale Centre-ville, Montréal, PQ H3C 3J7, Canada

(Reçu le 3 mai 1994; accepté le 4 mai 1995)

Résumé – Dans les modèles polygéniques gaussiens développés par Lande, la contribution génétique à la variabilité d'un caractère phénotypique est celle d'une somme d'effets alléliques à plusieurs locus sujets à mutation, la contribution environnementale à cette variabilité étant le résultat d'un écart aléatoire, indépendant des effets alléliques, de moyenne nulle et de variance constante de génération en génération. La dynamique de la variabilité génétique est décrite par des équations de récurrence, d'une génération à la suivante, pour la matrice de covariance du vecteur multinormal des effets alléliques. Ces équations admettent un seul équilibre et la convergence globale vers cet équilibre est vérifiée par des itérations numériques. Dans cet article, nous présentons un modèle polygénique gaussien dans lequel les forces de sélection (viabilité, accouplement préférentiel, fertilité) agissent de façon générale sur le vecteur des effets alléliques, ceux-ci étant sujets à mutation, ségrégation et recombinaison. Pour étudier la convergence de la variabilité, nous proposons une condition pour la convergence des itérés de transformations de matrices symétriques définies semi-positives sous l'hypothèse d'existence d'un point fixe unique. Il s'agit essentiellement d'une condition de concavité qui, combinée à une condition de monotonie, a déjà été utilisée par Karlin pour démontrer la convergence de la variabilité dans des modèles phénotypiques gaussiens sans recombinaison ni ségrégation. Nous montrons que la condition de monotonie est redondante et nous donnons une preuve directe du résultat de convergence, en affaiblissant légèrement les hypothèses. Enfin, nous montrons que ce résultat s'applique aux modèles polygéniques gaussiens sans distinction entre les sexes ou sans linkage entre les locus. La robustesse des résultats en l'absence de ces hypothèses et la vitesse de la convergence sont étudiées par des itérations numériques.

modèle polygénique / modèle gaussien / variabilité génétique / matrice de covariance / concavité

Summary – **Convergence of variability in Gaussian polygenic models.** *In the Gaussian polygenic models developed by Lande, the genetic contribution to the variability of a phenotypic character is given by a sum of allelic effects at several loci subject to mutations. The environmental contribution to that variability is the result of a random deviation, independent of the allelic effects, with a zero mean and a constant variance from generation*

to generation. The dynamics of the genetic variability is described by recurrence equations, from one generation to the next, for the covariance matrix of the multinormal vector corresponding to the allelic effects. These equations admit only one equilibrium and global convergence toward that equilibrium is checked by numerical iterations. In this article we present a Gaussian polygenic model in which the selection forces (viability, preferential mating, fertility) act generally on the vector of allelic effects, these being subject to mutation, segregation and recombination. In order to study the convergence of variability, we establish a criterion for the convergence of the iterates of transformations on semi-positive definite matrices under the condition that there exists a unique fixed point. This criterion is essentially a concavity property which, combined with a monotonicity property, has previously been used by Karlin to demonstrate the convergence of variability in Gaussian phenotypic models without recombination or segregation. We show that the monotonicity property does not have to be assumed and we give a direct proof of the convergence result under slightly weakened hypotheses. Finally, we show that this result applies to Gaussian polygenic models without differences between the sexes or without linkage between the loci. The robustness of the results without these hypotheses and the rate of convergence are studied by numerical iterations.

polygenic model / Gaussian model / genetic variability / covariance matrix / concavity

INTRODUCTION

L'une des questions majeures de la génétique quantitative est le maintien de la variabilité de caractères héréditaires. Cette question remonte aux travaux de Galton (1887) et de Pearson (1903) pour qui le caractère espéré d'un individu était donné par une moyenne (mélange) des caractères des parents, la variabilité étant maintenue par des écarts dus aux effets aléatoires de l'environnement. Ceux-ci devaient cependant être très importants pour contrecarrer les effets réducteurs du mélange et expliquer la variabilité observée. Les mécanismes génétiques (ségrégation et recombinaison), même en présence de sélection, devaient par la suite permettre d'expliquer le maintien de polymorphismes sans même l'intervention de mutations (voir, par exemple, Fisher, 1918). Mais cela n'est pas le cas pour des modèles génétiques de sélection additifs, modèles approximatifs où chaque gène apporte une contribution additive à la valeur sélective, qui conduisent, sans mutation, à une diminution continue de variabilité. On propose dans ce cas que sélection et mutation doivent s'opposer et s'équilibrer pour expliquer le maintien de la variabilité. Les questions qui retiennent alors l'attention sont celles de l'existence et de la caractérisation (moyennes, variances, covariances) des états d'équilibres, de l'unicité de ces états (un ou plusieurs états d'équilibre possibles), de leur stabilité locale (états d'équilibre observables ou pas) et finalement de leur stabilité globale (états d'équilibre atteignables ou pas), qui exclut la possibilité de cycles limites ou même de phénomènes de chaos. C'est à cette dernière question que nous nous intéressons dans cet article.

Encore récemment, des modèles gaussiens ont été proposés pour des caractères quantitatifs : le modèle polygénique de Lande (1976) et le modèle phénotypique

(multifactoriel) de Karlin (1979). Dans ces modèles, le caractère se décompose en 2 parties additives représentées par 2 variables indépendantes de loi normale : une partie héréditaire (transmise par les parents) et une partie environnementale (associée aux fluctuations du milieu et à la mutation). Dans le modèle de Karlin (1979), la partie héréditaire est une combinaison linéaire de type galtonien des caractères des parents, ce qui a l'avantage de préserver la normalité d'une génération à la suivante si les forces d'évolution (sélection, mutation) s'expriment par des fonctions gaussiennes. Dans le modèle de Lande (1976), la partie héréditaire est une somme d'effets alléliques chez les parents dont la transmission obéit à des mécanismes génétiques (ségrégation mendélienne, recombinaison à plusieurs locus), ce qui résulte en un mélange de distributions statistiques (voir, par exemple, Felsenstein, 1977) qui ne peut généralement qu'être approché par une loi normale. La validité d'une telle approximation a été établie pour la première fois par Kimura (1965) pour les modèles à plusieurs locus libres (indépendants) dans le cas de sélection faible et de mutation faible (voir aussi Bulmer, 1980, pour le cas d'un grand nombre de locus indépendants), et elle a été étendue par la suite aux modèles à plusieurs locus liés dans le cas où les différences des valeurs sélectives et les taux de mutation sont suffisamment faibles par rapport aux taux de recombinaison (Fleming, 1979). Si cela est le cas, il suffit de suivre les moyennes, les variances et les covariances de génération en génération, comme dans le modèle de Karlin, pour décrire la dynamique de la distribution du caractère. Mais alors que la convergence de la variabilité (matrice de covariance) a été démontrée rigoureusement dans le modèle de Karlin, elle n'a été vérifiée jusqu'à maintenant que numériquement dans le modèle de Lande.

Nous proposons dans cet article un modèle polygénique général dans lequel les forces de sélection (viabilité, accouplement préférentiel, fertilité) agissent de façon quelconque sur des effets alléliques à plusieurs locus sujets à la mutation, à la recombinaison et à la ségrégation mendélienne dans une population d'effectif infini. Nous faisons cependant l'hypothèse très restrictive (voir la discussion plus loin) que les facteurs d'évolution sont tels que les effets alléliques suivent une loi multinormale en tout temps. Nous établissons les équations de récurrence pour la moyenne et la matrice de covariance d'une génération à la suivante et nous analysons la convergence des itérés de la transformation ainsi obtenue.

Notons que des généralisations du modèle de Lande ont aussi été proposées par Chevalet (1988) et par Perrin (1993) pour le cas où le caractère est une combinaison linéaire d'effets génétiques additifs sans mutation, mais dans une population d'effectif fini. Cependant, ces auteurs portent principalement leur attention sur la dynamique de la variance totale et sur la structure de la covariance des effets alléliques.

LE MODÈLE

On considère une population sexuée, d'effectif infini et à générations séparées, dans laquelle les individus sont diploïdes, c'est-à-dire qu'ils sont le résultat de l'union de 2 gamètes haploïdes dont l'un est d'origine maternelle (ovule) et l'autre d'origine paternelle (spermatozoïde). Le *génotype* d'un individu est déterminé par des effets alléliques à n locus. Au début de la génération t , ces effets sont représentés par

le couple de vecteurs $\mathbf{X}_f(t) = (\mathbf{X}_1(t), \mathbf{X}_2(t))$ chez les femelles et par le couple de vecteurs $\mathbf{X}_m(t) = (\mathbf{X}_3(t), \mathbf{X}_4(t))$ chez les mâles. Les vecteurs $\mathbf{X}_1(t), \mathbf{X}_3(t)$ désignent les effets alléliques des n locus sur les gamètes d'origine maternelle et les vecteurs $\mathbf{X}_2(t), \mathbf{X}_4(t)$, les effets alléliques sur les gamètes d'origine paternelle. Les couples de vecteurs $\mathbf{X}_f(t)$ et $\mathbf{X}_m(t)$ sont supposés de loi multinormale de moyenne $\bar{\mathbf{x}}_f(t)$ et $\bar{\mathbf{x}}_m(t)$, respectivement, et de matrice de covariance $\Sigma_f(t)$ et $\Sigma_m(t)$, respectivement. Les paramètres de ces lois se modifient sous l'action des facteurs d'évolution que sont la sélection (viabilité, accouplement, fertilité), et la transmission (recombinaison, ségrégation, mutation).

Viabilité

Les jeunes de la génération t peuvent avoir des viabilités différentes selon leur génotype avant de devenir adultes. On caractérise la viabilité chez les femelles par une fonction des effets alléliques, notée λ_f , que l'on suppose gaussienne et qui est entièrement déterminée par un point \mathbf{c}_f , qui correspond à un *maximum*, et une matrice symétrique définie semi-positive Λ_f , qui décrit les *interactions* mutuelles des effets alléliques en ce qui a trait au facteur considéré. (Dans le cas où Λ_f est définie positive, cette matrice correspond à l'inverse d'une matrice de covariance d'une loi de probabilité multinormale et \mathbf{c}_f à la moyenne de cette loi. Dans le cas où Λ_f n'est pas inversible, le point de maximum \mathbf{c}_f n'est pas unique.) De même chez les mâles, la viabilité est décrite par une fonction gaussienne des effets alléliques, notée λ_m , de point de maximum \mathbf{c}_m et de matrice d'interaction Λ_m . Les fonctions de viabilité sont alors de la forme

$$\lambda_f(\mathbf{x}_f) \propto \exp \left[-\frac{1}{2}(\mathbf{x}_f - \mathbf{c}_f)^T \Lambda_f (\mathbf{x}_f - \mathbf{c}_f) \right] \quad [1]$$

et

$$\lambda_m(\mathbf{x}_m) \propto \exp \left[-\frac{1}{2}(\mathbf{x}_m - \mathbf{c}_m)^T \Lambda_m (\mathbf{x}_m - \mathbf{c}_m) \right] \quad [2]$$

(La notation T est utilisée pour désigner la transposée.) Les densités de probabilité des vecteurs $\mathbf{X}_f^*(t)$ et $\mathbf{X}_m^*(t)$, qui représentent les effets alléliques chez les adultes femelles et les adultes mâles de la génération t respectivement, sont données par

$$p_t^*(\mathbf{x}_f) = \frac{\lambda_f(\mathbf{x}_f)p_t(\mathbf{x}_f)}{\int \int \lambda_f(\zeta)p_t(\zeta)d\zeta} \quad [3]$$

et

$$q_t^*(\mathbf{x}_m) = \frac{\lambda_m(\mathbf{x}_m)q_t(\mathbf{x}_m)}{\int \int \lambda_m(\zeta)q_t(\zeta)d\zeta} \quad [4]$$

Accouplement

On suppose que les adultes femelles et les adultes mâles de la génération t se rencontrent au hasard et que la probabilité d'accouplement d'une femelle et d'un mâle est proportionnelle à une fonction gaussienne des effets alléliques chez la femelle et chez le mâle, notée ψ , qui atteint un maximum à un point noté $\mathbf{c}^{(1)} = (\mathbf{c}_f^{(1)}, \mathbf{c}_m^{(1)})$ et dont la matrice d'interaction symétrique définie semi-positive est notée Ψ . Dans le cas d'homogamie, cette fonction peut être, par exemple, de la forme $\psi(\mathbf{x}_f, \mathbf{x}_m) = \exp \left[-\frac{1}{2} (\|\mathbf{x}_f - \mathbf{x}_m\|)^2 \right]$, en prenant

$$\mathbf{c}_f^{(1)} = \mathbf{c}_m^{(1)} = \mathbf{0} \quad \text{et} \quad \Psi = \begin{pmatrix} \mathbf{I} & -\mathbf{I} \\ -\mathbf{I} & \mathbf{I} \end{pmatrix} \tag{5}$$

où $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}}$ et \mathbf{I} est la matrice identité de dimension $2n \times 2n$. La panmixie correspond au cas $\Psi = \mathbf{0}$, c'est-à-dire $\psi(\mathbf{x}_f, \mathbf{x}_m) = 1$ pour tout $(\mathbf{x}_f, \mathbf{x}_m)$. Le cas général décrit un accouplement préférentiel.

La densité de probabilité du couple de vecteurs $(\mathbf{X}_f^{**}(t), \mathbf{X}_m^{**}(t))$, qui représente les effets alléliques chez un couple formé à la génération t , est alors donnée par

$$p_t^{**}(\mathbf{x}_f, \mathbf{x}_m) = \frac{\psi(\mathbf{x}_f, \mathbf{x}_m) p_t^*(\mathbf{x}_f) q_t^*(\mathbf{x}_m)}{\int \int \psi(\zeta_f, \zeta_m) p_t^*(\zeta_f) q_t^*(\zeta_m) d\zeta_f d\zeta_m} \tag{6}$$

Fertilité

Les couples peuvent aussi différer par leur fertilité. On considère donc une fonction ϕ qui attribue aux couples, selon les effets alléliques qu'ils portent, des poids qui sont proportionnels au nombre moyen de leurs descendants. On suppose encore une fois que cette fonction est gaussienne, de point de maximum $\mathbf{c}^{(2)} = (\mathbf{c}_f^{(2)}, \mathbf{c}_m^{(2)})$ et de matrice d'interaction Φ , symétrique définie semi-positive.

La densité de probabilité des effets alléliques chez les couples de parents de la génération t représentés autant de fois qu'ils ont de descendants devient alors

$$p_t^\omega(\mathbf{x}_f, \mathbf{x}_m) = \frac{\phi(\mathbf{x}_f, \mathbf{x}_m) p_t^{**}(\mathbf{x}_f, \mathbf{x}_m)}{\int \int \phi(\zeta_f, \zeta_m) p_t^{**}(\zeta_f, \zeta_m) d\zeta_f d\zeta_m} \tag{7}$$

Transformation par la sélection

Comme les différents facteurs ci-dessus sont tous représentés par des fonctions gaussiennes qui pondèrent des fonctions de densité d'effets alléliques, on peut les regrouper en un seul que l'on appellera *sélection*. Cette sélection sera décrite par

une fonction gaussienne, notée ω . L'équation [7] devient alors

$$p_t^w(\mathbf{x}_f, \mathbf{x}_m) = \frac{\omega(\mathbf{x}_f, \mathbf{x}_m)p_t(\mathbf{x}_f)q_t(\mathbf{x}_m)}{\int \int \omega(\zeta_f, \zeta_m)p_t(\zeta_f)q_t(\zeta_m)d\zeta_f d\zeta_m} \quad [8]$$

On note $\boldsymbol{\theta} = (\boldsymbol{\theta}_f, \boldsymbol{\theta}_m)$ le point de maximum de la fonction de sélection et $\boldsymbol{\Omega}$ la matrice symétrique définie semi-positive décrivant les interactions totales des effets alléliques en ce qui a trait à la sélection.

En utilisant cette notation, on trouve

$$p_t^w(\mathbf{x}_f, \mathbf{x}_m) \propto \exp\left\{-\frac{1}{2}\left[(\mathbf{x}_f - \bar{\mathbf{x}}_f(t), \mathbf{x}_m - \bar{\mathbf{x}}_m(t))^T \boldsymbol{\Sigma}^{-1}(t)(\mathbf{x}_f - \bar{\mathbf{x}}_f(t), \mathbf{x}_m - \bar{\mathbf{x}}_m(t)) + (\mathbf{x}_f - \boldsymbol{\theta}_f, \mathbf{x}_m - \boldsymbol{\theta}_m)^T \boldsymbol{\Omega}(\mathbf{x}_f - \boldsymbol{\theta}_f, \mathbf{x}_m - \boldsymbol{\theta}_m)\right]\right\} \quad [9]$$

où

$$\boldsymbol{\Omega} = \boldsymbol{\Lambda} + \boldsymbol{\Psi} + \boldsymbol{\Phi}, \quad [10]$$

$$\boldsymbol{\Omega}\boldsymbol{\theta} = \boldsymbol{\Lambda}\mathbf{c} + \boldsymbol{\Psi}\mathbf{c}^{(1)} + \boldsymbol{\Phi}\mathbf{c}^{(2)} \quad [11]$$

avec

$$\mathbf{c} = (\mathbf{c}_f, \mathbf{c}_m), \boldsymbol{\Sigma}(t) = \begin{pmatrix} \boldsymbol{\Sigma}_f(t) & 0 \\ 0 & \boldsymbol{\Sigma}_m(t) \end{pmatrix}, \boldsymbol{\Lambda} = \begin{pmatrix} \boldsymbol{\Lambda}_f & 0 \\ 0 & \boldsymbol{\Lambda}_m \end{pmatrix}$$

Le couple de vecteurs $(\mathbf{X}_f^w(t), \mathbf{X}_m^w(t))$ qui a la densité ci-dessus est de loi multinormale de moyenne $(\bar{\mathbf{x}}_f^w(t), \bar{\mathbf{x}}_m^w(t))$ et de matrice de covariance $\boldsymbol{\Sigma}^w(t)$ que l'on détermine à partir de la relation :

$$\begin{aligned} & (\mathbf{x}_f, \mathbf{x}_m)^T \boldsymbol{\Sigma}^{w-1}(t)(\mathbf{x}_f, \mathbf{x}_m) - 2(\mathbf{x}_f, \mathbf{x}_m)^T \boldsymbol{\Sigma}^{w-1}(t)(\bar{\mathbf{x}}_f^w(t), \bar{\mathbf{x}}_m^w(t)) \\ &= (\mathbf{x}_f, \mathbf{x}_m)^T [\boldsymbol{\Sigma}^{-1}(t) + \boldsymbol{\Omega}] (\mathbf{x}_f, \mathbf{x}_m) - 2(\mathbf{x}_f, \mathbf{x}_m)^T [\boldsymbol{\Sigma}^{-1}(t)(\bar{\mathbf{x}}_f(t), \bar{\mathbf{x}}_m(t)) + \boldsymbol{\Omega}(\boldsymbol{\theta}_f, \boldsymbol{\theta}_m)] \end{aligned}$$

Si on procède à l'identification des matrices de covariance et des vecteurs moyennes, on obtient :

$$\boldsymbol{\Sigma}^w(t) = (\boldsymbol{\Sigma}^{-1}(t) + \boldsymbol{\Omega})^{-1} \quad [12]$$

et

$$\begin{pmatrix} \bar{\mathbf{x}}_f^w(t) \\ \bar{\mathbf{x}}_m^w(t) \end{pmatrix} = (\boldsymbol{\Sigma}^{-1}(t) + \boldsymbol{\Omega})^{-1} \left[\boldsymbol{\Sigma}^{-1}(t) \begin{pmatrix} \bar{\mathbf{x}}_f(t) \\ \bar{\mathbf{x}}_m(t) \end{pmatrix} + \boldsymbol{\Omega} \begin{pmatrix} \boldsymbol{\theta}_f \\ \boldsymbol{\theta}_m \end{pmatrix} \right] \quad [13]$$

Recombinaison et ségrégation

Les individus sélectionnés peuvent transmettre à leurs descendants ou bien des copies des gamètes qu'ils ont hérités ou bien des recombinaisons de ces gamètes. On désigne le taux de recombinaison entre les locus i et j par r_{ij} , lequel correspond à la proportion de gamètes recombinés aux locus i et j pour tout $i, j = 1, \dots, n$ et

$i \neq j$. On définit aussi $r_{ii} = 0$ pour tout $i = 1, \dots, n$. À remarquer que $r_{ij} = r_{ji}$ pour tout $i, j = 1, \dots, n$. De plus, on suppose que la ségrégation est mendélienne chez les 2 sexes.

Si on désigne par $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)$ un gamète à n locus transmis par un individu de génotype aléatoire $(\boldsymbol{x}_1, \boldsymbol{x}_2)$, où $\boldsymbol{x}_1 = (\chi_{11}, \dots, \chi_{1n})$ représente le gamète d'origine maternelle de cet individu et $\boldsymbol{x}_2 = (\chi_{21}, \dots, \chi_{2n})$ celui d'origine paternelle, on a :

$$(\eta_i, \eta_j) = \begin{cases} (\chi_{1i}, \chi_{1j}) & \text{avec probabilité } (1 - r_{ij})/2 \\ (\chi_{2i}, \chi_{2j}) & \text{avec probabilité } (1 - r_{ij})/2 \\ (\chi_{1i}, \chi_{2j}) & \text{avec probabilité } r_{ij}/2 \\ (\chi_{2i}, \chi_{1j}) & \text{avec probabilité } r_{ij}/2 \end{cases}$$

On remarque que le vecteur $\boldsymbol{\eta}$ admet pour moyenne la demi-somme des moyennes des vecteurs \boldsymbol{x}_1 et \boldsymbol{x}_2 . En effet, pour tout $i = 1, \dots, n$, on a :

$$\eta_i = \begin{cases} \chi_{1i} & \text{avec probabilité } \frac{1}{2} \\ \chi_{2i} & \text{avec probabilité } \frac{1}{2} \end{cases}$$

d'où

$$\bar{\eta}_i = \frac{\bar{\chi}_{1i} + \bar{\chi}_{2i}}{2}$$

et ce quel que soit le taux de recombinaison. La recombinaison n'affecte donc pas les moyennes des effets alléliques.

Par ailleurs, le calcul des covariances entre les composantes du vecteur $\boldsymbol{\eta}$ donne, pour tout $i, j = 1, \dots, n$,

$$\begin{aligned} \text{Cov}(\eta_i, \eta_j) &= (1 - r_{ij}) \left(\frac{\text{Cov}(\chi_{1i}, \chi_{1j}) + \text{Cov}(\chi_{2i}, \chi_{2j})}{2} \right) \\ &+ r_{ij} \left(\frac{\text{Cov}(\chi_{1i}, \chi_{2j}) + \text{Cov}(\chi_{2i}, \chi_{1j})}{2} \right) - \frac{1}{2}(r_{ij} - \frac{1}{2}) \times (\bar{\chi}_{1i} - \bar{\chi}_{2i})(\bar{\chi}_{1j} - \bar{\chi}_{2j}) \end{aligned}$$

À remarquer que, dans les expressions des covariances des effets alléliques sur les gamètes transmis, interviennent les moyennes des effets alléliques, comme cela a déjà été relevé par Chevalet (communication personnelle), sauf dans le cas de la non distinction des sexes, qui garantit $(\bar{\chi}_{1i} - \bar{\chi}_{2i}) = 0$, pour tout i) ou dans celui de l'absence de linkage, c'est-à-dire celui de la recombinaison libre ($r_{ij} = \frac{1}{2}$, pour tout $i \neq j$.) *Nous analyserons dans ce qui suit la dynamique de la variabilité dans le cas de la non distinction de sexes ou de la recombinaison libre.*

Les taux de recombinaison permettent de définir une matrice, dite matrice de recombinaison : $\mathbf{R} = ||r_{ij}||_{i,j=1}^n$. On a alors :

Lemme 1 : *Les matrices $\mathbf{U} - \mathbf{R}$ et $\mathbf{U} - 2\mathbf{R}$, où \mathbf{U} est une matrice dont tous les éléments sont égaux à 1 et \mathbf{R} est une matrice de recombinaison, sont définies semi-positives.*

Le lemme 1, démontré à l'annexe A.1, fait ressortir une propriété mathématique fondamentale des matrices de recombinaison qui ne semble pas avoir été remarquée

jusqu'à maintenant. Elle intervient dans la définition de la transformation de la matrice de covariance des effets alléliques suivant les mécanismes génétiques de transmission.

Mutation

À chaque locus i , on peut avoir mutation, c'est-à-dire modification de l'effet allélique, et ce avec une faible probabilité $\mu_i > 0$. Les mutations qui surviennent aux différents locus sont toutefois supposées indépendantes dans leur ensemble et indépendantes des effets alléliques qu'elles modifient. De plus, leurs espérances sont toutes supposées nulles. Les mutations sont donc *centrées*. Ainsi, on caractérise la mutation au locus i par une variable aléatoire ξ_i de moyenne 0 et de variance $v_i > 0$. Si on désigne par $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)$ un gamète avant la mutation et par $\boldsymbol{\eta}' = (\eta'_1, \dots, \eta'_n)$, ce gamète après la mutation, on a pour tout $i = 1, \dots, n$:

$$\eta'_i = \begin{cases} \eta_i & \text{avec probabilité } 1 - \mu_i \\ \eta_i + \xi_i & \text{avec probabilité } \mu_i \end{cases}$$

Maintenant, si on calcule les moyennes, les variances et les covariances des variables η'_1, \dots, η'_n , on obtient :

$$\begin{aligned} E(\eta'_i) &= E(\eta_i) \\ \text{Var}(\eta'_i) &= \text{Var}(\eta_i) + \mu_i v_i \\ \text{Cov}(\eta'_i, \eta'_j) &= \text{Cov}(\eta_i, \eta_j) \end{aligned}$$

pour tout $i, j = 1, \dots, n$ et $i \neq j$. À remarquer que les mutations ne changent pas les moyennes ni les covariances des effets alléliques à des locus différents sous les hypothèses ci-dessus, alors qu'elles accroissent les variances à tous les locus.

Les taux de mutation et les variances de mutation des effets alléliques permettent de définir une matrice diagonale, dite *matrice de mutation* :

$$\mathbf{D} = \begin{pmatrix} \mu_1 v_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n v_n \end{pmatrix} \quad [14]$$

Enfin, si on considère le génotype $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)$, où $\boldsymbol{\eta}_1 = (\eta_{11}, \dots, \eta_{1n})$ représente le gamète issu d'une femelle de génotype $(\boldsymbol{\chi}_1, \boldsymbol{\chi}_2) = (\chi_{11}, \dots, \chi_{1n}, \chi_{21}, \dots, \chi_{2n})$ et $\boldsymbol{\eta}_2 = (\eta_{21}, \dots, \eta_{2n})$ le gamète issu d'un mâle de génotype $(\boldsymbol{\chi}_3, \boldsymbol{\chi}_4) = (\chi_{31}, \dots, \chi_{3n}, \chi_{41}, \dots, \chi_{4n})$, alors la ségrégation mendélienne fait en sorte que

$$\begin{aligned} \text{Cov}(\eta_{1i}, \eta_{2j}) &= \frac{1}{4} \text{Cov}(\chi_{1i}, \chi_{3j}) + \frac{1}{4} \text{Cov}(\chi_{1i}, \chi_{4j}) \\ &\quad + \frac{1}{4} \text{Cov}(\chi_{2i}, \chi_{3j}) + \frac{1}{4} \text{Cov}(\chi_{2i}, \chi_{4j}) \end{aligned}$$

pour tout $i, j = 1, \dots, n$.

Transformation par la transmission

Suite aux facteurs de recombinaison, de ségrégation et de mutation, on aboutit aux jeunes individus de la génération $t + 1$, décrits par les vecteurs d'effets alléliques $\mathbf{X}_f(t + 1) = (\mathbf{X}_1(t + 1), \mathbf{X}_2(t + 1))$ et $\mathbf{X}_m(t + 1) = (\mathbf{X}_3(t + 1), \mathbf{X}_4(t + 1))$ chez les femelles et chez les mâles, respectivement. Puisque la ségrégation est mendélienne, que les mutations sont centrées et que la recombinaison ne modifie pas les moyennes des effets alléliques, le couple de vecteurs $(\mathbf{X}_f(t + 1), \mathbf{X}_m(t + 1))$ admet pour moyenne

$$\begin{pmatrix} \bar{\mathbf{x}}_f(t + 1) \\ \bar{\mathbf{x}}_m(t + 1) \end{pmatrix} = \mathbf{J} \begin{pmatrix} \bar{\mathbf{x}}_f^w(t) \\ \bar{\mathbf{x}}_m^w(t) \end{pmatrix} \tag{15}$$

où $(\bar{\mathbf{x}}_f^w(t), \bar{\mathbf{x}}_m^w(t))$ est la moyenne du couple de vecteurs $(\mathbf{X}_f^w(t), \mathbf{X}_m^w(t))$ donné en [13] et

$$\mathbf{J} = \begin{pmatrix} \mathbf{I}/2 & \mathbf{I}/2 & 0 & 0 \\ 0 & 0 & \mathbf{I}/2 & \mathbf{I}/2 \\ \mathbf{I}/2 & \mathbf{I}/2 & 0 & 0 \\ 0 & 0 & \mathbf{I}/2 & \mathbf{I}/2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \otimes \mathbf{I} \tag{16}$$

dans laquelle \mathbf{I} représente la matrice identité de dimension $n \times n$ et \otimes désigne le produit tensoriel. La matrice de covariance $\Sigma(t + 1) = \|\Sigma_{ij}(t + 1)\|_{i,j=1}^4$, où $\Sigma_{ij}(t + 1)$ représente la matrice de covariance de $\mathbf{X}_i(t + 1)$ et $\mathbf{X}_j(t + 1)$ pour $i, j = 1, \dots, 4$, s'exprime à partir de la matrice de covariance du vecteur $(\mathbf{X}_f^w(t), \mathbf{X}_m^w(t)) = (\mathbf{X}_1^w(t), \dots, \mathbf{X}_4^w(t))$ désignée $\Sigma^w(t)$ et donnée en [12]. En écrivant $\Sigma^w(t) = \|\Sigma_{ij}^w(t)\|_{i,j=1}^4$ où $\Sigma_{ij}^w(t)$ représente la matrice de covariance de $\mathbf{X}_i^w(t)$ et $\mathbf{X}_j^w(t)$ pour $i, j = 1, \dots, 4$, on obtient, après la recombinaison, la ségrégation mendélienne et la mutation, les équations de récurrence suivantes :

$$\Sigma_{11}(t + 1) = (\mathbf{U} - \mathbf{R}_f) \odot \left(\frac{\Sigma_{11}^w(t) + \Sigma_{22}^w(t)}{2} \right) + \mathbf{R}_f \odot \left(\frac{\Sigma_{12}^w(t) + \Sigma_{21}^w(t)}{2} \right) + \mathbf{D}_f \tag{17}$$

$$\Sigma_{12}(t + 1) = \frac{1}{4} (\Sigma_{13}^w(t) + \Sigma_{14}^w(t) + \Sigma_{23}^w(t) + \Sigma_{24}^w(t)) \tag{18}$$

$$\Sigma_{22}(t + 1) = (\mathbf{U} - \mathbf{R}_m) \odot \left(\frac{\Sigma_{33}^w(t) + \Sigma_{44}^w(t)}{2} \right) + \mathbf{R}_m \odot \left(\frac{\Sigma_{34}^w(t) + \Sigma_{43}^w(t)}{2} \right) + \mathbf{D}_m \tag{19}$$

$$\Sigma_{33}(t + 1) = \Sigma_{11}(t + 1) \tag{20}$$

$$\Sigma_{34}(t + 1) = \Sigma_{12}(t + 1) \tag{21}$$

$$\Sigma_{44}(t + 1) = \Sigma_{22}(t + 1) \tag{22}$$

$$\Sigma_{21}(t + 1) = \Sigma_{12}^T(t + 1) \tag{23}$$

$$\Sigma_{43}(t + 1) = \Sigma_{34}^T(t + 1) \tag{24}$$

$$\begin{aligned} \Sigma_{13}(t + 1) &= \Sigma_{14}(t + 1) = \Sigma_{23}(t + 1) = \Sigma_{24}(t + 1) = \Sigma_{31}(t + 1) \\ &= \Sigma_{41}(t + 1) = \Sigma_{32}(t + 1) = \Sigma_{42}(t + 1) = 0 \end{aligned} \tag{25}$$

où \mathbf{U} est la matrice dont tous les éléments valent 1, \odot représente le produit de Hadamard (c'est-à-dire le produit élément par élément; voir, par exemple, Horn et Johnson, 1992), $\mathbf{R}_f, \mathbf{R}_m, \mathbf{D}_f$ et \mathbf{D}_m désignent respectivement les matrices de recombinaison et les matrices diagonales de mutation chez les femelles (lorsque l'indice est f) et chez les mâles (lorsque l'indice est m). Remarquons que l'équation [17] peut s'écrire sous la forme

$$\Sigma_{11}(t+1) = (\mathbf{U} - 2\mathbf{R}_f) \odot \left(\frac{\Sigma_{11}^w(t) + \Sigma_{22}^w(t) - \Sigma_{12}^w(t) - \Sigma_{21}^w(t)}{4} \right) + \left(\frac{\Sigma_{11}^w(t) + \Sigma_{22}^w(t) + \Sigma_{12}^w(t) + \Sigma_{21}^w(t)}{4} \right) + \mathbf{D}_f \quad [26]$$

Puisque $(\mathbf{U} - 2\mathbf{R}_f)$ est une matrice symétrique définie semi-positive (lemme 1), que

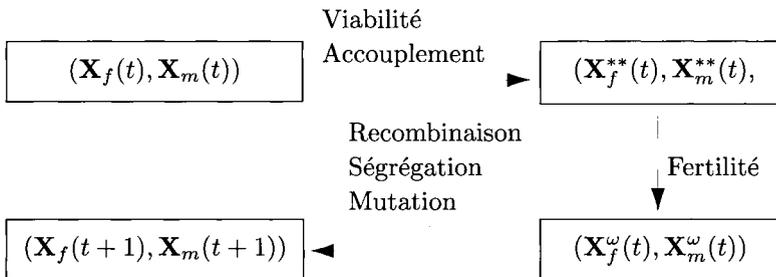
$$\mathbf{X}^T \left(\frac{\Sigma_{11}^w(t) + \Sigma_{22}^w(t) - \Sigma_{12}^w(t) - \Sigma_{21}^w(t)}{4} \right) \mathbf{X} = (\mathbf{X}, -\mathbf{X}, \mathbf{0}, \mathbf{0})^T \Sigma^w(t) (\mathbf{X}, -\mathbf{X}, \mathbf{0}, \mathbf{0}) \geq 0 \quad [27]$$

et

$$\mathbf{X}^T \left(\frac{\Sigma_{11}^w(t) + \Sigma_{22}^w(t) + \Sigma_{12}^w(t) + \Sigma_{21}^w(t)}{4} \right) \mathbf{X} = (\mathbf{X}, \mathbf{X}, \mathbf{0}, \mathbf{0})^T \Sigma^w(t) (\mathbf{X}, \mathbf{X}, \mathbf{0}, \mathbf{0}) \geq 0 \quad [28]$$

que le produit de Hadamard de 2 matrices symétriques définies semi-positives est une matrice symétrique définie semi-positive (voir, par exemple, Horn et Johnson, 1992), et que la somme de matrices symétriques définies semi-positives est une matrice symétrique définie semi-positive, alors on voit que $\Sigma_{11}(t+1)$ est bien une matrice symétrique définie semi-positive, comme d'ailleurs doit l'être une matrice de covariance.

Le cycle complet du modèle est illustré par le schéma suivant :



CONDITION DE CONVERGENCE

Rappelons les définitions suivantes concernant les matrices symétriques :

Définition 1 : On dit qu'une matrice symétrique \mathbf{A} est définie positive (respectivement, définie semi-positive) si pour tout vecteur \mathbf{X} non nul on a $\mathbf{X}^T \mathbf{A} \mathbf{X} > 0$ (respectivement, ≥ 0).

Définition 2 : Soit \mathbf{A} et \mathbf{B} 2 matrices symétriques définies semi-positives. On dit qu' \mathbf{A} est plus grand que \mathbf{B} et on note $\mathbf{A} \geq \mathbf{B}$ (respectivement, \mathbf{A} strictement plus grand que \mathbf{B} et on note $\mathbf{A} > \mathbf{B}$) si la matrice $\mathbf{A} - \mathbf{B}$ est définie semi-positive (respectivement, positive).

La convergence des itérés d'une transformation de matrices symétriques définies semi-positives qui admet un équilibre unique et qui est croissante et concave par rapport à la relation d'ordre \geq (voir les définitions exactes ci-dessous) découle du théorème suivant :

Théorème 1 : Soit \mathbf{T} une transformation continue définie sur l'ensemble des matrices symétriques définies semi-positives d'une dimension donnée, et vérifiant les propriétés suivantes :

- i) \mathbf{T} est croissante, c'est-à-dire que $\mathbf{T}(\mathbf{C}_1) \geq \mathbf{T}(\mathbf{C}_2)$ dès que $\mathbf{C}_1 \geq \mathbf{C}_2$ où \mathbf{C}_1 et \mathbf{C}_2 sont des matrices symétriques définies semi-positives ;
- ii) \mathbf{T} est concave, c'est-à-dire que, pour toutes matrices symétriques définies semi-positives \mathbf{C}_1 et \mathbf{C}_2 , on a :

$$\mathbf{T}\left(\frac{\mathbf{C}_1 + \mathbf{C}_2}{2}\right) \geq \frac{\mathbf{T}(\mathbf{C}_1) + \mathbf{T}(\mathbf{C}_2)}{2}$$

Si \mathbf{T} possède un point fixe unique qui est une matrice symétrique définie positive, c'est-à-dire $\mathbf{T}(\widehat{\mathbf{C}}) = \widehat{\mathbf{C}} > 0$, alors la suite des itérés de \mathbf{T} , et ce à partir de n'importe quelle matrice symétrique définie positive \mathbf{C} , converge vers $\widehat{\mathbf{C}}$, c'est-à-dire

$$\lim_{n \rightarrow \infty} \mathbf{T}^{(n)}(\mathbf{C}) = \widehat{\mathbf{C}}$$

Nous présentons une démonstration directe de ce théorème à l'annexe A.2. Il existe sans doute une version plus générale de ce théorème dans la théorie des opérateurs (voir, par exemple, Krasnosel'skii, 1964). Par ailleurs, la condition de croissance s'avère redondante avec les autres hypothèses du théorème.

Théorème 2 : Soit \mathbf{T} est une transformation continue définie sur l'ensemble des matrices symétriques définies semi-positives. Si la transformation \mathbf{T} est concave, alors elle est croissante.

Ce résultat, démontré à l'annexe A.3, simplifie grandement la tâche dans les preuves de la stabilité globale des points fixes uniques de transformations de matrices symétriques définies semi-positives.

CONVERGENCE DE LA VARIABILITÉ

Matrice de covariance

En vertu de ce qui précède, pour établir la convergence de la matrice de covariance, il suffit de démontrer la concavité de la transformation de la matrice de covariance des effets alléliques d'une génération à la suivante. À remarquer que cette matrice est nécessairement une matrice symétrique définie semi-positive.

Les étapes principales de la dynamique des variances et des covariances des effets alléliques sont décrites par les transformations matricielles \mathbf{T}_1 et \mathbf{T}_2 définies comme suit :

$$\begin{aligned}\mathbf{T}_1 : \Sigma(t) &\mapsto \Sigma^\omega(t) \\ \mathbf{T}_2 : \Sigma^\omega(t) &\mapsto \Sigma(t+1)\end{aligned}$$

La transformation \mathbf{T}_1 , décrite par l'équation [12], représente les changements dus à la sélection et la transformation \mathbf{T}_2 , décrite par les équations [17] à [25], ceux dus à la transmission. Les équations de récurrence des variances et des covariances sont donc résumées par la transformation composée $\mathbf{T}_2 \circ \mathbf{T}_1$.

Ces transformations sont définies sur l'ensemble des matrices symétriques définies semi-positives (car elles sont des transformations de matrices de covariance). La transformation \mathbf{T}_1 est continue et concave en vertu du lemme suivant :

Lemme 2 : Si Ω est une matrice symétrique définie semi-positive, alors la transformation

$$\mathbf{T}(\Sigma) = (\Sigma^{-1} + \Omega)^{-1} \quad [29]$$

qui peut être définie pour toute matrice symétrique définie semi-positive Σ de même dimension que Ω , est concave, en plus d'être continue.

Le lemme 2 est démontré par Karlin (1979), dans le cas où Ω est inversible, par une analyse locale avec référence à la théorie générale des opérateurs (Krasnosel'skii, 1964). Pour ce faire, il considère une matrice symétrique \mathbf{A} de petite norme ayant la même dimension que Σ et il fait le développement de Taylor suivant :

$$\mathbf{T}(\Sigma + \mathbf{A}) - \mathbf{T}(\Sigma) = \mathbf{T}'(\mathbf{A}) + \frac{1}{2}\mathbf{T}''(\mathbf{A}) + \mathbf{O}(\|\mathbf{A}\|^3) \quad [30]$$

où $\mathbf{T}'(\mathbf{A})$ est une transformation linéaire, $\mathbf{T}''(\mathbf{A})$ est une transformation quadratique et $\mathbf{O}(\|\mathbf{A}\|^3)$ est une fonction analytique de \mathbf{A} d'ordre $\|\mathbf{A}\|^3$. Il établit la croissance de \mathbf{T} en montrant que $\mathbf{T}'(\mathbf{A}) > 0$ et la concavité de \mathbf{T} en montrant que $\mathbf{T}''(\mathbf{A}) < 0$. Nous présentons à l'annexe A.4 une démonstration directe de ce résultat en utilisant la décomposition spectrale des matrices définies semi-positives, qui inclut aussi le cas où Ω est singulière.

En outre, il est aisé de voir que la transformation \mathbf{T}_2 est *affine* dans le sens que

$$\mathbf{T}_2 \left(\frac{\Sigma_1 + \Sigma_2}{2} \right) = \frac{\mathbf{T}_2(\Sigma_1) + \mathbf{T}_2(\Sigma_2)}{2} \quad [31]$$

pour toutes matrices symétriques définies semi-positives Σ_1 et Σ_2 . En particulier, \mathbf{T}_2 est concave, en plus d'être continue.

La concavité de la transformation globale $\mathbf{T}_2 \circ \mathbf{T}_1$ est alors assurée par le lemme suivant démontré à l'annexe A.5 :

Lemme 3 : Soit \mathbf{T}_1 et \mathbf{T}_2 2 transformations continues, définies sur l'ensemble des matrices symétriques définies semi-positives. Si \mathbf{T}_1 et \mathbf{T}_2 sont concaves, alors leur composée, $\mathbf{T}_2 \circ \mathbf{T}_1$, est aussi concave, en plus d'être continue.

Finalement, en utilisant le théorème 2, on aboutit au résultat fondamental suivant :

Résultat 1 : Si un point d'équilibre unique (matrice symétrique définie positive) existe pour la matrice de covariance, alors la suite des itérés de la matrice de covariance converge vers cet équilibre, et ce à partir de n'importe quelle matrice symétrique définie positive de départ.

Dans les modèles proposés par Lande, l'existence d'une matrice d'équilibre unique définie positive est établie. La stabilité globale de cet équilibre est une conséquence directe du résultat ci-dessus.

Vecteur moyenne

En regroupant les équations [13] et [15], la dynamique du vecteur moyenne $\bar{\mathbf{x}}(t) = (\bar{\mathbf{x}}_f(t), \bar{\mathbf{x}}_m(t))$ est donnée par l'équation de récurrence

$$\bar{\mathbf{x}}(t + 1) = \mathbf{P}(t)\bar{\mathbf{x}}(t) + \boldsymbol{\epsilon}(t) \tag{32}$$

où
$$\mathbf{P}(t) = \mathbf{J}(\boldsymbol{\Sigma}^{-1}(t) + \boldsymbol{\Omega})^{-1}\boldsymbol{\Sigma}^{-1}(t)$$

$$\boldsymbol{\epsilon}(t) = \mathbf{P}(t)\boldsymbol{\Sigma}(t)\boldsymbol{\Omega}\boldsymbol{\theta}$$

L'équation [32] est identique à l'équation [4.5] donnée dans Karlin (1979), car la recombinaison n'intervient pas dans la dynamique du vecteur moyenne. La convergence du vecteur moyenne découle de celle de la matrice de covariance. Le résultat de convergence du vecteur moyenne est démontré dans Karlin (1979) lorsque le rayon spectral de la matrice $\mathbf{J}\mathbf{J}^T$ est plus petit ou égal à 1 et que la matrice symétrique $\boldsymbol{\Omega}$ est définie positive. Dans notre cas, la première condition est vérifiée, car la matrice \mathbf{J} est doublement stochastique. On a donc *convergence du vecteur moyenne au moins dans le cas $\boldsymbol{\Omega} > 0$* .

RÉSULTATS NUMÉRIQUES

Pour avoir un aperçu sur la convergence et la vitesse de convergence de la suite $\boldsymbol{\Sigma}(t)$ pour la matrice de covariance, nous avons étudié numériquement la variation de la distance entre les itérés consécutifs de cette matrice en fonction des différents facteurs d'évolution. Nous donnons dans les figures 1 à 6 quelques exemples de graphiques résultant de cette étude. La distance est exprimée par la norme de la matrice des différences consécutives définie par $\max_{i,j} |\boldsymbol{\Sigma}(t)(i, j) - \boldsymbol{\Sigma}(t - 1)(i, j)|$, pour toute génération t . Les forces de mutation et de sélection sont mesurées par la norme spectrale, c'est-à-dire, la plus grande valeur propre en module, des matrices associées.

Dans les différents cas analysés, avec ou sans distinction de sexes, nous avons observé une décroissance assez rapide de la distance entre les matrices de covariance

consécutives, en moyenne en dessous de 10^{-3} après 100 itérations. De plus, l'étude a montré que les influences du nombre de locus et des taux de recombinaison ne sont pas significatives (voir, par exemple, les figures 1 et 2). Seules les forces de mutation et de sélection affectent la vitesse de convergence. On remarque que l'on se rapproche de plus en plus vite de la matrice de covariance d'équilibre au fur à mesure que la norme de la matrice de mutation augmente (voir, par exemple, la figure 3). D'autre part, lorsque la norme de la matrice associée à la mutation est faible relativement à celle de la sélection, nous remarquons que la vitesse de convergence est plus grande pour les grandes valeurs de la norme de la matrice de sélection (voir la figure 4).

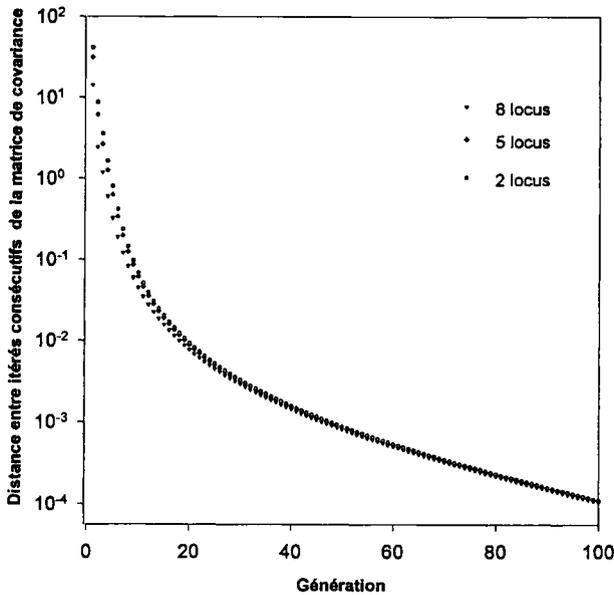


Fig 1. Modèle sans distinction de sexe avec plusieurs locus. (La recombinaison est libre et les normes des matrices de sélection et de mutation sont respectivement 10^{-2} et 10^{-5} .)

Dans le cas où la sélection est différente selon le sexe et qu'elle est suffisamment forte, nous constatons non seulement qu'il y a convergence, que la recombinaison soit libre ou non, mais aussi que cette convergence est rapide, en fait d'autant plus rapide que la norme de la matrice de mutation est petite (voir, par exemple, la figure 5). Mais si la sélection est très faible, la convergence est très lente après quelques générations (voir pour illustration la figure 6).

DISCUSSION

Le modèle polygénique gaussien que nous avons considéré pour un caractère quantitatif s'inspire à la fois du modèle phénotypique multifactoriel de Karlin (1979) et du modèle polygénique de Lande (1976). Il emprunte à Lande la détermination

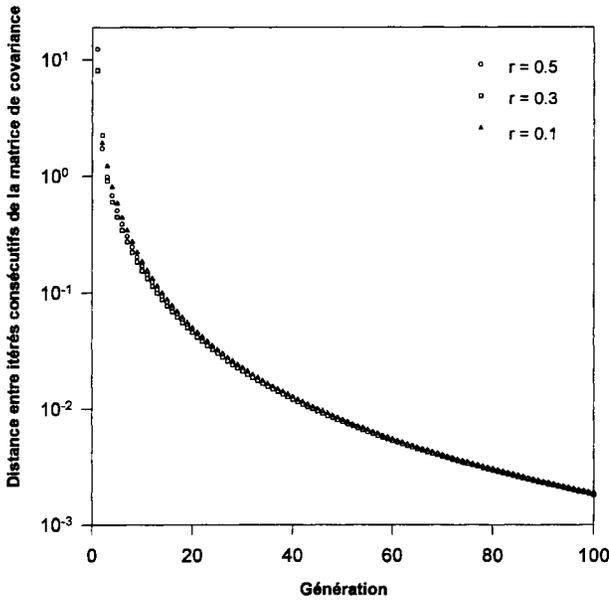


Fig 2. Modèle sans distinction de sexes avec plusieurs niveaux de recombinaison. (Le nombre de locus est 2 et les normes des matrices de sélection et de mutation sont respectivement 10^{-4} et 10^{-5} .)

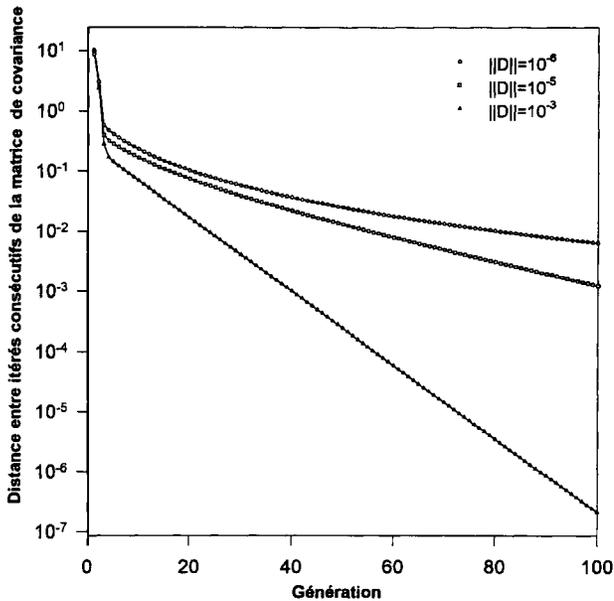


Fig 3. Modèle sans distinction de sexes avec divers niveaux de mutation. (La recombinaison est libre, le nombre de locus est 2 et la norme de la matrice de sélection est fixée à 10^{-3} .)

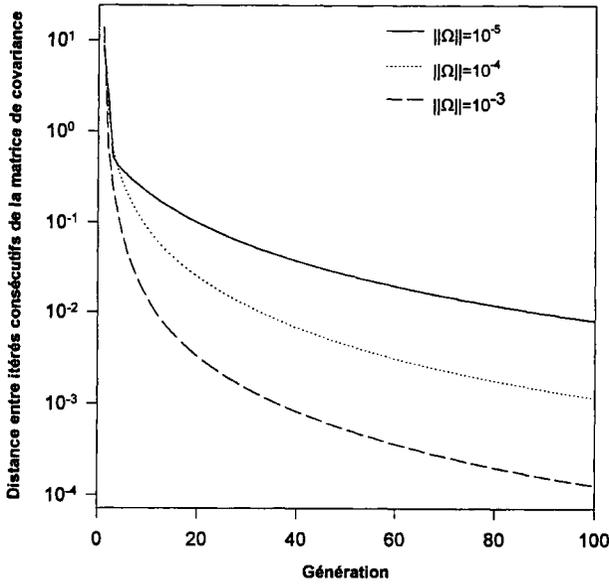


Fig 4. Modèle sans distinction de sexes avec 3 niveaux de sélection. (Le nombre de locus est 2, la recombinaison est libre et la norme de la matrice de mutation est 10^{-6} .)

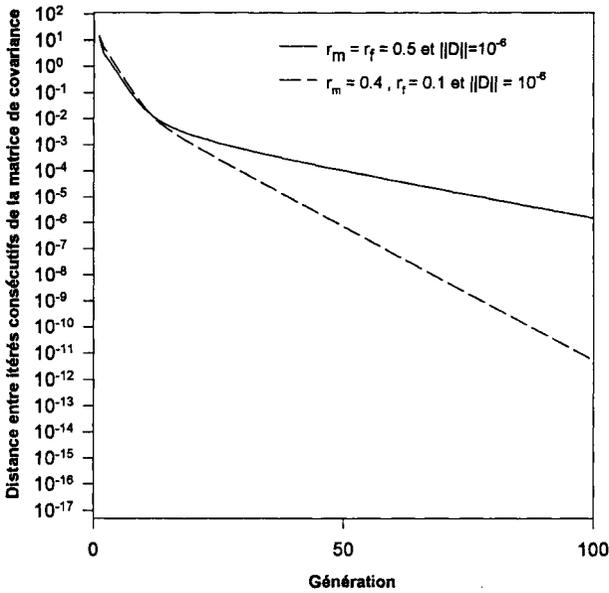


Fig 5. Modèle avec distinction de sexe et sélection forte. (Les normes des matrices de sélection chez les mâles et chez les femelles sont respectivement 10 et 10^4 .)

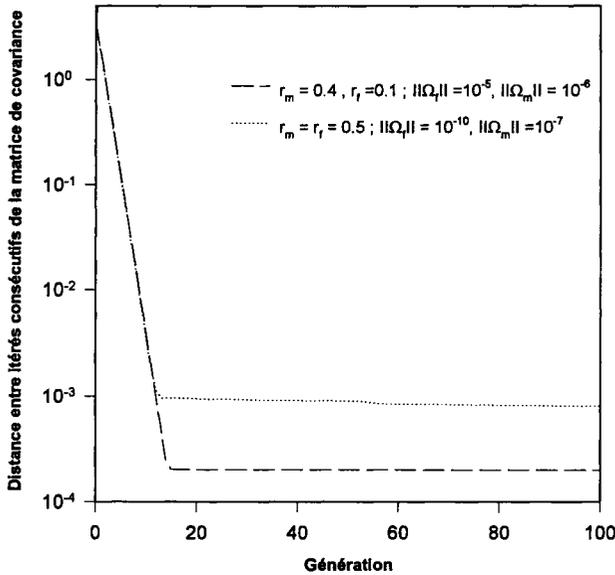


Fig 6. Modèle avec distinction de sexes et sélection faible. (Le nombre de locus est 2 et la norme de la matrice de mutation est 10^{-4} chez les 2 sexes.)

du caractère par des effets alléliques à plusieurs locus liés qui sont sujets à la transmission mendélienne avec mutation, ségrégation et recombinaison, et à Karlin la formulation générale des facteurs de sélection (viabilité, accouplement préférentiel, fertilité). Les équations de récurrence que nous obtenons pour la matrice de covariance des effets alléliques font donc intervenir une matrice de recombinaison qui est absente dans les équations de Karlin. D'autre part, ces équations sont plus générales que celles de Lande, car les facteurs de sélection agissent de façon quelconque sur les effets alléliques des individus, et non pas seulement sur une somme de ces effets perturbée par une déviation environnementale indépendante et centrée comme dans le modèle de Lande. Dans ce cas, les matrices d'interaction des effets alléliques associées aux différents facteurs de sélection sont des matrices dont tous les blocs correspondant à un même individu ou à une paire d'individus sont des constantes fois la matrice \mathbf{U} , dont tous les éléments sont des 1.

Le résultat de convergence de la matrice de covariance des effets alléliques s'applique donc en particulier au modèle panmictique de Lande (1976), sans distinction de sexes et sans différences de fertilité, ce qui confirme les résultats des itérations numériques effectuées jusqu'à maintenant. En revanche, la convergence de la moyenne des effets alléliques dans ce cas ne découle pas du résultat énoncé à ce sujet dans le présent article, la matrice d'interaction Ω associée à la sélection étant alors singulière, bien que la convergence puisse encore être démontrée (Mahdi, 1994). Ce cas correspond à un continuum d'états d'équilibre pour la moyenne du vecteur des effets alléliques. En fait, dans le modèle de Lande, la condition d'équilibre porte sur la somme des composantes de ce vecteur pour un individu, qui doit être à l'optimum de sélection. Mais en général, dans le cas où la matrice symétrique Ω est

définie positive; l'optimum de sélection n'est pas atteint par la moyenne du vecteur des effets alléliques. D'après Karlin (1979), il suffirait pour cela que la matrice de transmission \mathbf{J} , qui modélise la ségrégation mendélienne, commute avec la matrice $\mathbf{\Omega}^{-1}$, l'inverse de la matrice $\mathbf{\Omega}$. Cela ne peut survenir que pour des cas particuliers. Il en est sans doute de même pour la croissance de la valeur sélective moyenne ou du produit des valeurs sélectives moyennes chez les mâles et chez les femelles dans les populations avec distinction de sexes, mais avec accouplement au hasard et sans différences de fertilité, même lorsque l'on suppose que la structure de covariance a atteint un état d'équilibre (voir, par exemple, Lande, 1979, 1980b, pour les modèles polygéniques additifs avec transmission mendélienne, et Lessard, 1991, pour les modèles phénotypiques avec mélange, équitable et sans interaction, des caractères parentaux). Le «théorème fondamental de la sélection naturelle» (Fisher, 1930) n'est pas de validité générale ici comme dans les modèles génétiques exacts.

Seuls les écarts à la panmixie qui peuvent s'exprimer à l'aide d'une fonction de sélection sur les couples formés au hasard ont été considérés dans le présent article. Une telle fonction de sélection a les mêmes effets que des différences de fertilité. Il est à remarquer que Lande (1977) modélise les écarts à la panmixie de façon différente. Il traite la consanguinité en supposant que les effets alléliques au même locus chez un individu ont un coefficient de corrélation constant, et l'homogamie en supposant que les caractères des individus qui s'accouplent ont un coefficient de corrélation constant. De la consanguinité peut aussi être créée par dérive génétique dans une population d'effectif limité, comme cela a été traité par Chevalet (1988), puis Perrin (1993), pour des modèles polygéniques additifs avec pondération des locus. Ces cas ont aussi été analysés et des conclusions semblables à celles présentées ci-dessus ont été obtenues.

Le cas de caractères pléiotropiques avec des effets alléliques multidimensionnels comme dans Lande (1980a) est inclus dans la présente analyse. En effet, il suffit de regrouper les effets alléliques en blocs indissociables à l'intérieur desquels la recombinaison est nulle. (Voir aussi Mahdi, 1994, pour une analyse directe de ce cas.)

La condition de convergence des itérés d'une transformation d'une matrice de covariance, à savoir la concavité, pourrait avoir des applications à bien d'autres modèles gaussiens que le modèle phénotypique de Karlin (1979) et que le présent modèle polygénique. La condition d'existence et d'unicité d'un état d'équilibre reste cependant à être vérifiée dans chaque cas. Dans le modèle de Karlin, seule l'existence est à vérifier, puisque l'unicité est garantie par une concavité stricte de la transformation. Dans le modèle polygénique additif de Lande (1976), un état d'équilibre unique est déterminé explicitement. Cela reste à faire pour d'autres modèles polygéniques et pour d'autres modèles éventuels. Le résultat de convergence assure alors que l'état d'équilibre sera atteint et maintenu, ce qui exclut la possibilité de cycles limites ou de chaos. Pour une discussion sur l'état d'équilibre proprement dit dans le modèle polygénique additif, nous renvoyons le lecteur à Lande (1976) et Bulmer (1980).

Des restrictions importantes limitent encore la généralité d'application du résultat de convergence aux modèles polygéniques : pour que la transformation de la matrice de covariance soit fermée, on doit supposer ou bien que les forces de sélection ne différencient pas les sexes, ou bien que les taux de recombinaison sont tous

égaux à $1/2$. Sinon, la transformation fait intervenir les moyennes des effets alléliques. La première condition laisse ouvert le problème de la convergence dans le cas d'effets alléliques à des locus liés avec distinction de sexes et remet en question la croissance du produit des valeurs sélectives moyennes chez les mâles et chez les femelles qui en découle (Lande, 1980b). La seconde condition, qui stipule que la ségrégation se fait indépendamment à chaque locus, donc que la recombinaison est libre, rapproche le présent modèle du modèle infinitésimal (voir, par exemple, Bulmer, 1980), qui suppose en plus un très grand nombre de locus et une petite contribution additive au caractère de la part de chacun d'eux. La présente étude est reliée de cette façon aux travaux de Langlois (1990) sur les différentes composantes de la variance dans le cas de l'homogamie et aux questions sur la convergence de la structure de covariance soulevées par Feldman et Cavalli Sforza (1979) dans un contexte plus général de transmission de caractères.

Les itérations numériques ont confirmé la convergence de la matrice de covariance, même dans les cas où la sélection est différente chez les 2 sexes et la recombinaison n'est pas libre. De plus, la convergence est pratiquement complétée en quelques dizaines de générations seulement, ce qui autorise à utiliser la matrice de covariance d'équilibre dans l'équation de récurrence pour le vecteur moyenne, qui converge habituellement moins rapidement.

Finalement, on ne peut pas passer sous silence l'hypothèse la plus forte qui est omniprésente dans cette étude : la normalité. Tous les facteurs d'évolution (sélection, mutation) sont exprimés par des fonctions unimodales gaussiennes ; les effets alléliques qui déterminent le caractère forment un vecteur de loi multinormale ; et la normalité est préservée d'une génération à la suivante. Or, dans la réalité, des fonctions de sélection multimodales (pour des populations naturelles occupant un vaste territoire aux conditions locales fort diverses) ou des fonctions de sélection tronquées (pour des caractères quantitatifs sous-jacents à certaines maladies létales) sont fréquentes. En particulier, une fonction de sélection tronquée dans une population d'effectif limité, comme cela est courant en sélection artificielle, peut entraîner une réponse qui s'écarte substantiellement à long terme de prévisions fondées sur des approximations gaussiennes (voir, par exemple, Bulmer, 1980 ; Robertson et Hill, 1983 ; Chevalet, 1988 ; Verrier *et al.*, 1991 ; Hospital et Chevalet, 1993), même en supposant un modèle additif sans dominance ni épistasie. De plus, toutes les critiques formulées à l'endroit du modèle additif gaussien de Lande (1976), fondé sur l'hypothèse de multinormalité des effets alléliques, s'appliquent au modèle plus général du présent article, en particulier celles de Turelli et Barton (1990) à partir d'approximations plus précises qui prennent en compte les récurrences pour les moments d'ordre supérieur à 2. Enfin, même en supposant la multinormalité des effets alléliques à une génération donnée, la recombinaison (comme la mutation d'ailleurs) brise cette multinormalité et crée un mélange de lois multinormales qui sera approximativement de loi multinormale seulement si les paramètres des lois multinormales du mélange ayant un poids non négligeable ne sont pas trop différents, ce qui sera généralement le cas si la sélection et la mutation sont des forces faibles par rapport à la recombinaison (Kimura, 1965 ; Felsenstein, 1977 ; Fleming, 1979).

Cependant, les modèles gaussiens restent approximatifs et, les erreurs s'accumulant de génération en génération, ils peuvent difficilement servir à faire des prévisions à long terme. Donc, même si les résultats du présent article concernent la

convergence globale de matrices de covariance satisfaisant des équations de récurrence précises, ils peuvent suggérer en réalité une stabilité locale plutôt que globale des états d'équilibres, car les équations de récurrence risquent fort d'être valides seulement près de l'équilibre.

REMERCIEMENTS

Travail partiellement subventionné par des fonds CRSNG du gouvernement du Canada et des fonds FCAR du gouvernement du Québec.

RÉFÉRENCES

- Bartlett MS (1951) An inverse matrix adjustment arising in discriminant analysis. *Ann Math Stat* 22, 107-111
- Bulmer MG (1980) *The Mathematical Theory of Quantitative Genetics*. Clarendon Press, Oxford
- Chevalet C (1988) Control of genetic drift in selected populations. In : *Proceedings of the Second International Conference on Quantitative Genetics* (BS Weir, EJ Eisen, MM Goodman, G Namkoong, eds), Sinauer, Sunderland, 379-394
- Feldman MW, Cavalli-Sforza LL (1979) Aspects of variance and covariance analysis with cultural inheritance. *Theor Popul Biol* 15, 276-307
- Felsenstein J (1977) Multivariate normal genetic models with a finite number of loci. In : *Proceedings of the International Conference on Quantitative Genetics* (E Pollak, O Kempthorne, TB Bailey, Jr, eds), Iowa State Univ Press, Ames, IA, 227-246
- Fisher RA (1918) The correlation between relatives on the supposition of Mendelian inheritance. *Trans Roy Soc Edinburgh* 52, 399-433
- Fisher RA (1930) *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford
- Fleming WH (1979) Equilibrium distributions of continuous polygenic traits. *SIAM J Appl Math* 36, 148-168
- Galton F (1887) *Natural Inheritance*. Macmillan, Londres
- Horn RA, Johnson CR (1992) *Matrix Analysis*. Cambridge University Press, New York
- Hospital F, Chevalet C (1993) Effects of population size and linkage on optimal selection intensity. *Theor Appl Genet* 86, 775-780
- Karlin S (1979) Models of multifactorial inheritance. I. Multivariate formulations and basic convergence results. *Theor Pop Biol* 15, 308-355
- Kimura M (1965) A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc Natl Acad Sci USA* 54, 731-736
- Krasnosel'skii DO (1964) *Positive Solutions of Operator Equations*. Noordhoff, Groningen
- Lande R (1976) The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet Res Cambridge* 26, 221-235
- Lande R (1977) The influence of the mating system on the maintenance of genetic variability in polygenic characters. *Genetics* 86, 485-498
- Lande R (1979) Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* 33, 402-416
- Lande R (1980a) The genetic covariance between traits maintained by linked pleiotropic mutation. *Genetics* 91, 203-215
- Lande R (1980b) Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution* 34, 292-303

- Langlois (1990) Incidence de la sélection et de l'homogamie sur les paramètres du modèle génétique additif. *Genet Sel Evol* 22, 119-132
- Lessard S (1991) Statistical models for sex-ratio evolution. In : *Handbook of Statistics 8* (CR Rao, R Chakraborty, eds), North-Holland, Amsterdam, 347-372
- Mahdi S (1994) Convergence de la variabilité de caractères génétiques quantitatifs. Thèse de doctorat, université de Montréal, Montréal
- Pearson K (1903) Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. *Phil Trans Roy Soc London*, Ser A 200, 1-66
- Perrin A (1993) Dynamique des systèmes multi-loci polygéniques. D'un modèle quantitatif vers une nouvelle méthodologie. Thèse de doctorat, ENGREF, Paris
- Robertson A, Hill WG (1983) Population and quantitative genetics of many linked loci in finite populations. *Proc R Soc Lond B* 219, 253-264
- Turelli M, Barton NH (1990) Dynamics of polygenic characters under selection. *Theor Pop Biol* 38, 1-57
- Verrier E, Colleau JJ, Foulley JL (1991) Methods for predicting response to selection in small populations under additive genetic models: a review. *Livest Prod Sci* 29, 93-114

ANNEXE

A.1 Démonstration du lemme 1

On considère un modèle diploïde à n locus et on désigne par $i_1^{(0)}, \dots, i_n^{(0)}$ le gamète d'origine maternelle et $i_1^{(1)}, \dots, i_n^{(1)}$ le gamète d'origine paternelle chez un individu. Cet individu transmet le gamète $i_1^{(\varepsilon_1)}, \dots, i_n^{(\varepsilon_n)}$ avec la probabilité $R(\varepsilon)$ où

$$\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$$

avec

$$\varepsilon_j = \begin{cases} 0 & \text{si au locus } j \text{ le gène est d'origine maternelle} \\ 1 & \text{si au locus } j \text{ le gène est d'origine paternelle} \end{cases}$$

pour tout $j = 1, \dots, n$.

Ainsi, la probabilité de ne pas avoir de recombinaison entre les locus k et l est donnée par

$$1 - R_{kl} = \sum_{\varepsilon|\varepsilon_k=\varepsilon_l=0} R(\varepsilon) + \sum_{\varepsilon|\varepsilon_k=\varepsilon_l=1} R(\varepsilon)$$

Maintenant si on multiplie à droite et à gauche la matrice $\mathbf{U} - \mathbf{R}$ par un vecteur \mathbf{X} non nul et par le vecteur transposé \mathbf{X}^T , on arrive à

$$\begin{aligned}
\mathbf{X}^T(\mathbf{U} - \mathbf{R})\mathbf{X} &= \sum_{k,l} x_k x_l (1 - R_{kl}) \\
&= \sum_{\varepsilon} \sum_{k,l|\varepsilon_k=\varepsilon_l=0} x_k x_l R(\varepsilon) + \sum_{\varepsilon} \sum_{k,l|\varepsilon_k=\varepsilon_l=1} x_k x_l R(\varepsilon) \\
&= \sum_{\varepsilon} \left[\left(\sum_{k|\varepsilon_k=0} x_k \right)^2 + \left(\sum_{k|\varepsilon_k=1} x_k \right)^2 \right] R(\varepsilon) \geq 0 \quad [33]
\end{aligned}$$

d'où la matrice symétrique $\mathbf{U} - \mathbf{R}$ est définie semi-positive. Si on fait de même avec la matrice $\mathbf{U} - 2\mathbf{R}$, on aboutit à

$$\begin{aligned}
\mathbf{X}^T(\mathbf{U} - 2\mathbf{R})\mathbf{X} &= \sum_{k,l} x_k x_l (1 - 2R_{kl}) \\
&= -\sum_{k,l} x_k x_l + 2 \sum_{k,l} x_k x_l (1 - R_{kl}) \\
&= \sum_{\varepsilon} \left[-\left(\sum_{k,l} x_k x_l \right) + 2 \left(\sum_{k|\varepsilon_k=0} x_k \right)^2 + 2 \left(\sum_{k|\varepsilon_k=1} x_k \right)^2 \right] R(\varepsilon) \\
&= \sum_{\varepsilon} \left[-\left(\sum_{k|\varepsilon_k=0} x_k + \sum_{k|\varepsilon_k=1} x_k \right)^2 + 2 \left(\sum_{k|\varepsilon_k=0} x_k \right)^2 + 2 \left(\sum_{k|\varepsilon_k=1} x_k \right)^2 \right] R(\varepsilon) \\
&= \sum_{\varepsilon} \left[\sum_{k|\varepsilon_k=0} x_k - \sum_{k|\varepsilon_k=1} x_k \right]^2 R(\varepsilon) \geq 0 \quad [34]
\end{aligned}$$

d'où la matrice symétrique $\mathbf{U} - 2\mathbf{R}$ est aussi définie semi-positive.

A.2 Démonstration du théorème 1

Montrons d'abord ce qui suit :

La suite des itérés de la matrice identiquement nulle, notée 0, converge vers le point fixe $\widehat{\mathbf{C}}$, c'est-à-dire

$$\lim_{n \rightarrow \infty} \mathbf{T}^{(n)}(0) = \widehat{\mathbf{C}}$$

En utilisant la croissance de \mathbf{T} et l'hypothèse $\mathbf{T}(\widehat{\mathbf{C}}) = \widehat{\mathbf{C}} > 0$, on peut écrire :

$$\widehat{\mathbf{C}} = \mathbf{T}(\widehat{\mathbf{C}}) \geq \mathbf{T}(0) \geq 0 \quad [35]$$

Ainsi, en répétant n fois ce procédé, on aboutit aux inégalités suivantes :

$$\widehat{\mathbf{C}} \geq \mathbf{T}^{(n)}(0) \geq \mathbf{T}^{(n-1)}(0) \geq 0 \quad [36]$$

On peut écrire alors, pour tout vecteur non nul \mathbf{X} les inégalités suivantes :

$$\mathbf{X}^T \widehat{\mathbf{C}} \mathbf{X} \geq \mathbf{X}^T \mathbf{T}^{(n)}(0) \mathbf{X} \geq \mathbf{X}^T \mathbf{T}^{(n-1)}(0) \mathbf{X} \geq 0 \quad [37]$$

La suite $\mathbf{X}^T \mathbf{T}^{(n)}(0) \mathbf{X}$ est croissante et bornée, elle est donc convergente. En particulier, en prenant $\mathbf{X} = \mathbf{e}_i$ où \mathbf{e}_i est le i^e vecteur de la base canonique de R^n , on définit pour tout i

$$C_{ii} = \lim_{n \rightarrow \infty} \mathbf{e}_i^T \mathbf{T}^{(n)}(0) \mathbf{e}_i$$

De même, pour tout $i \neq j$, on définit

$$a_{ij} = \lim_{n \rightarrow \infty} (\mathbf{e}_i + \mathbf{e}_j)^T \mathbf{T}^{(n)}(0) (\mathbf{e}_i + \mathbf{e}_j)$$

d'où

$$C_{ij} = \lim_{n \rightarrow \infty} \mathbf{e}_i^T \mathbf{T}^{(n)}(0) \mathbf{e}_j = \frac{a_{ij} - C_{ii} - C_{jj}}{2} \quad [38]$$

La suite $\mathbf{T}^{(n)}(0)$ converge donc vers une matrice \mathbf{C} dont les éléments sont les valeurs C_{ij} .

Montrons maintenant que $\mathbf{C} = \widehat{\mathbf{C}}$. En utilisant l'égalité $\mathbf{T}^n(0) = \mathbf{T}(\mathbf{T}^{n-1}(0))$ et la continuité de l'opérateur \mathbf{T} , on peut écrire :

$$\mathbf{C} = \lim_{n \rightarrow \infty} \mathbf{T}^{(n)}(0) = \lim_{n \rightarrow \infty} \mathbf{T}(\mathbf{T}^{(n-1)}(0)) = \mathbf{T}(\lim_{n \rightarrow \infty} \mathbf{T}^{(n-1)}(0)) = \mathbf{T}(\mathbf{C}) \quad [39]$$

L'unicité du point fixe $\widehat{\mathbf{C}}$ entraîne que $\mathbf{C} = \widehat{\mathbf{C}}$.

Étudions maintenant la convergence des itérés de \mathbf{T} pour une matrice symétrique définie positive quelconque. Pour cela, commençons d'abord par démontrer le résultat suivant :

Pour tout entier naturel k , la suite des itérés de $2^k \widehat{\mathbf{C}}$ par \mathbf{T} converge vers le point fixe $\widehat{\mathbf{C}}$.

La preuve de ce résultat repose sur les propriétés de continuité et de concavité de \mathbf{T} . En vertu de ces propriétés, on a

$$\mathbf{T}(\widehat{\mathbf{C}}) = \mathbf{T}\left(\frac{0 + 2\widehat{\mathbf{C}}}{2}\right) \geq \frac{\mathbf{T}(0) + \mathbf{T}(2\widehat{\mathbf{C}})}{2} \geq \frac{\mathbf{T}(2\widehat{\mathbf{C}})}{2} \quad [40]$$

qui s'écrit aussi sous la forme

$$\mathbf{T}(2\widehat{\mathbf{C}}) \leq 2\mathbf{T}(\widehat{\mathbf{C}}) = 2\widehat{\mathbf{C}}$$

En appliquant k fois ce procédé, on trouve

$$\mathbf{T}(2^k \widehat{\mathbf{C}}) \leq 2^k \mathbf{T}(\widehat{\mathbf{C}}) = 2^k \widehat{\mathbf{C}} \quad [41]$$

Puisque $\widehat{\mathbf{C}} \leq 2^k \widehat{\mathbf{C}}$ et la transformation \mathbf{T} est croissante, on peut écrire

$$\widehat{\mathbf{C}} = \mathbf{T}(\widehat{\mathbf{C}}) \leq \mathbf{T}(2^k \widehat{\mathbf{C}}) \leq 2^k \widehat{\mathbf{C}}$$

En appliquant n fois la transformation \mathbf{T} , on aboutit à

$$\widehat{\mathbf{C}} \leq \mathbf{T}^{(n)}(2^k \widehat{\mathbf{C}}) \leq \mathbf{T}^{(n-1)}(2^k \widehat{\mathbf{C}}) \leq \dots \leq \mathbf{T}(2^k \widehat{\mathbf{C}}) \leq 2^k \widehat{\mathbf{C}} \quad [42]$$

Maintenant, si on multiplie à droite et à gauche chaque membre de l'inégalité ci-dessus respectivement par un vecteur non nul \mathbf{X} et par le vecteur transposé \mathbf{X}^T , on obtient

$$\mathbf{X}^T \widehat{\mathbf{C}} \mathbf{X} \leq \mathbf{X}^T \mathbf{T}^{(n)}(2^k \widehat{\mathbf{C}}) \mathbf{X} \leq \mathbf{X}^T \mathbf{T}^{(n-1)}(2^k \widehat{\mathbf{C}}) \mathbf{X} \leq \dots \leq \mathbf{X}^T (2^k \widehat{\mathbf{C}}) \mathbf{X} \quad [43]$$

On en déduit donc que la suite $\mathbf{X}^T \mathbf{T}^{(n)}(2^k \widehat{\mathbf{C}}) \mathbf{X}$ est décroissante et bornée. Elle est donc convergente. En procédant comme précédemment, on trouve une limite \mathbf{L} à $\mathbf{T}^{(n)}(2^k \widehat{\mathbf{C}})$ lorsque $n \rightarrow \infty$, c'est-à-dire

$$\mathbf{L} = \lim_{n \rightarrow \infty} \mathbf{T}^{(n)}(2^k \widehat{\mathbf{C}})$$

On démontre aisément que $\mathbf{L} = \widehat{\mathbf{C}}$ en reprenant la technique utilisée pour montrer l'égalité $\mathbf{C} = \widehat{\mathbf{C}}$.

On est maintenant en mesure de montrer le résultat général, à savoir :

Pour toute matrice symétrique définie positive \mathbf{C} , on a :

$$\widehat{\mathbf{C}} = \lim_{n \rightarrow \infty} \mathbf{T}^{(n)}(\mathbf{C})$$

Pour déduire ce résultat, montrons que, pour tout $\mathbf{C} > 0$, il existe un nombre réel positif k , tel que : $0 < \mathbf{C} < 2^k \widehat{\mathbf{C}}$. Pour cela, montrons que, pour tout $\mathbf{C} > 0$, il existe un réel positif λ vérifiant $0 < \mathbf{C} < \lambda \widehat{\mathbf{C}}$, c'est-à-dire

$$0 < \mathbf{X}^T \mathbf{C} \mathbf{X} < \lambda \mathbf{X}^T \widehat{\mathbf{C}} \mathbf{X}$$

pour tout vecteur \mathbf{X} non nul. Définissons l'application

$$h(\mathbf{X}) = \frac{\mathbf{X}^T \mathbf{C} \mathbf{X}}{\mathbf{X}^T \widehat{\mathbf{C}} \mathbf{X}}$$

et considérons l'ensemble compact $\mathbf{E} = \{\mathbf{X} \text{ tel que } \|\mathbf{X}\| = 1\}$. La fonction h étant continue sur le compact \mathbf{E} , elle atteint ses bornes sur ce compact. Soit \mathbf{Z} dans \mathbf{E} tel que $\alpha = h(\mathbf{Z}) = \max\{h(\mathbf{X}) \text{ pour } \mathbf{X} \text{ dans } \mathbf{E}\}$. On a donc, pour tout \mathbf{X} appartenant à \mathbf{E} , $h(\mathbf{X}) \leq \alpha$, ce qui se traduit par :

$$\mathbf{X}^T \mathbf{C} \mathbf{X} \leq \alpha \mathbf{X}^T \widehat{\mathbf{C}} \mathbf{X} \quad [44]$$

Il existe donc un entier $k > 0$ tel que

$$\mathbf{X}^T \mathbf{C} \mathbf{X} < 2^k \mathbf{X}^T \widehat{\mathbf{C}} \mathbf{X}.$$

Cela se généralise aisément à tous les vecteurs sauf le vecteur nul car pour tout $\mathbf{Y} \neq \mathbf{0}$, il existe un réel $\lambda_{\mathbf{Y}} \neq 0$ et un \mathbf{X} appartenant à \mathbf{E} tel que $\mathbf{Y} = \lambda_{\mathbf{Y}} \mathbf{X}$. Ainsi donc, on a

$$\mathbf{Y}^T \mathbf{C} \mathbf{Y} = \lambda_{\mathbf{Y}}^2 \mathbf{X}^T \mathbf{C} \mathbf{X} < \lambda_{\mathbf{Y}}^2 2^k \mathbf{X}^T \widehat{\mathbf{C}} \mathbf{X} = 2^k \mathbf{Y}^T \widehat{\mathbf{C}} \mathbf{Y} \quad [45]$$

On a alors pour tout $\mathbf{C} > 0$

$$0 < \mathbf{C} < 2^k \widehat{\mathbf{C}}$$

Si on applique n fois l'opérateur \mathbf{T} , on aboutit à

$$\mathbf{T}^{(n)}(0) \leq \mathbf{T}^{(n)}(\mathbf{C}) \leq \mathbf{T}^{(n)}(2^k \widehat{\mathbf{C}}) \quad [46]$$

En utilisant la convergence des suites $\mathbf{T}^{(n)}(0)$ et $\mathbf{T}^{(n)}(2^k \widehat{\mathbf{C}})$ vers $\widehat{\mathbf{C}}$, on en déduit

$$\lim_{n \rightarrow \infty} \mathbf{T}^{(n)}(\mathbf{C}) = \widehat{\mathbf{C}}$$

A.3 Démonstration du théorème 2

On suppose qu'il existe 2 matrices symétriques définies positives, \mathbf{C}_1 et \mathbf{C}_2 , et un vecteur non nul \mathbf{X} vérifiant

$$\mathbf{C}_2 > \mathbf{C}_1 > 0 \quad \text{et} \quad \mathbf{X}^T \mathbf{T}(\mathbf{C}_1) \mathbf{X} > \mathbf{X}^T \mathbf{T}(\mathbf{C}_2) \mathbf{X} > 0$$

et on montre alors qu'il y a contradiction avec les hypothèses du théorème. Pour ce faire, construisons la suite des matrices symétriques définies positives \mathbf{C}_n de la façon suivante :

$$\mathbf{C}_n = \mathbf{C}_1 + 2^{n-2}(\mathbf{C}_2 - \mathbf{C}_1), \quad \text{pour } n \geq 3 \quad [47]$$

La suite s'écrit aussi sous la forme :

$$\mathbf{C}_n = 2\mathbf{C}_{n-1} - \mathbf{C}_1, \quad \text{pour } n \geq 3 \quad [48]$$

En appliquant la propriété de concavité de \mathbf{T} , on a

$$\mathbf{T}\left(\frac{\mathbf{C}_1 + \mathbf{C}_n}{2}\right) \geq \frac{\mathbf{T}(\mathbf{C}_1) + \mathbf{T}(\mathbf{C}_n)}{2} \quad [49]$$

En substituant \mathbf{C}_{n-1} par sa valeur dans l'équation ci-dessus, on trouve l'inégalité

$$\mathbf{T}(\mathbf{C}_n) \leq 2\mathbf{T}(\mathbf{C}_{n-1}) - \mathbf{T}(\mathbf{C}_1) \quad [50]$$

En itérant maintenant sur l'indice n , on aboutit à

$$\begin{aligned} \mathbf{T}(\mathbf{C}_n) &\leq 2^{n-2} \mathbf{T}(\mathbf{C}_2) - \left\{ \sum_{j=0}^{n-3} 2^j \mathbf{T}(\mathbf{C}_1) \right\} \\ &= 2^{n-2} \{ \mathbf{T}(\mathbf{C}_2) - \mathbf{T}(\mathbf{C}_1) \} + \mathbf{T}(\mathbf{C}_1) \end{aligned} \quad [51]$$

En multipliant alors les 2 membres de [51] à gauche par \mathbf{X}^T et à droite par \mathbf{X} , on arrive à

$$\mathbf{X}^T \mathbf{T}(\mathbf{C}_n) \mathbf{X} \leq 2^{n-2} \mathbf{X}^T \{ \mathbf{T}(\mathbf{C}_2) - \mathbf{T}(\mathbf{C}_1) \} \mathbf{X} + \mathbf{X}^T \mathbf{T}(\mathbf{C}_1) \mathbf{X} \quad [52]$$

Si on pose $\mathbf{X}^T(\mathbf{T}(\mathbf{C}_2) - \mathbf{T}(\mathbf{C}_1))\mathbf{X} = a < 0$ et $\mathbf{X}^T\mathbf{T}(\mathbf{C}_1)\mathbf{X} = b > 0$, on montre aisément que

$$\mathbf{X}^T\mathbf{T}(\mathbf{C}_n)\mathbf{X} < 0, \text{ dès que } n \geq \frac{\ln(-b/a)}{\ln 2} + 2$$

ce qui contredit l'hypothèse que $\mathbf{T}(\mathbf{C}_n) \geq 0$ pour tout n .

Le cas $\mathbf{C}_2 \geq \mathbf{C}_1 \geq 0$ s'obtient facilement par continuité.

A.4 Démonstration du lemme 2

La matrice $\mathbf{\Omega}$ étant symétrique et définie semi-positive, il existe des constantes positives $\lambda_1, \dots, \lambda_k$ et des vecteurs non nuls $\mathbf{V}_1, \dots, \mathbf{V}_k$ tels que

$$\mathbf{\Omega} = \lambda_1 \mathbf{V}_1 \mathbf{V}_1^T + \dots + \lambda_k \mathbf{V}_k \mathbf{V}_k^T$$

Commençons par le cas où $\mathbf{\Omega}$ est de rang 1, qui correspond au cas $k = 1$. Dans ce cas, en vertu de la formule de Sherman-Morrison démontrée dans Bartlett (1951) et en posant $\lambda = \lambda_1$ et $\mathbf{V} = \mathbf{V}_1$, on a

$$(\mathbf{\Sigma}^{-1} + \lambda \mathbf{V} \mathbf{V}^T)^{-1} = \mathbf{\Sigma} - \frac{\lambda \mathbf{\Sigma} \mathbf{V} \mathbf{V}^T \mathbf{\Sigma}}{1 + \lambda \mathbf{V}^T \mathbf{\Sigma} \mathbf{V}} \quad [53]$$

car $1 + \lambda \mathbf{V}^T \mathbf{\Sigma} \mathbf{V} > 0$. Si on pose maintenant $\sqrt{\lambda} \mathbf{V} = \mathbf{W}$, la transformation ci-dessus s'écrit alors

$$\mathbf{T}(\mathbf{\Sigma}) = \mathbf{\Sigma} - \frac{\mathbf{\Sigma} \mathbf{W} \mathbf{W}^T \mathbf{\Sigma}}{1 + \mathbf{W}^T \mathbf{\Sigma} \mathbf{W}} \quad [54]$$

et elle est définie et continue pour toute matrice symétrique définie semi-positive $\mathbf{\Sigma}$. De plus, pour toutes matrices symétriques définies semi-positives $\mathbf{\Sigma}_1$ et $\mathbf{\Sigma}_2$, et pour tout vecteur non nul \mathbf{X} , on obtient

$$\begin{aligned} \mathbf{X}^T \left(\mathbf{T} \left(\frac{\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2}{2} \right) - \frac{\mathbf{T}(\mathbf{\Sigma}_1)}{2} - \frac{\mathbf{T}(\mathbf{\Sigma}_2)}{2} \right) \mathbf{X} \\ = \frac{1}{2(2 + a + b)} \left[y \sqrt{\frac{1+b}{1+a}} - z \sqrt{\frac{1+a}{1+b}} \right]^2 \geq 0 \end{aligned} \quad [55]$$

où

$$\begin{aligned} \mathbf{X}^T \mathbf{\Sigma}_1 \mathbf{W} &= y, & \mathbf{X}^T \mathbf{\Sigma}_2 \mathbf{W} &= z \\ \mathbf{W}^T \mathbf{\Sigma}_1 \mathbf{W} &= a \geq 0, & \mathbf{W}^T \mathbf{\Sigma}_2 \mathbf{W} &= b \geq 0 \end{aligned} \quad [56]$$

Cela montre la concavité de la transformation dans ce cas particulier. Pour montrer le résultat dans le cas général, on procède par induction sur le rang de $\mathbf{\Omega}$. On note $\mathbf{\Omega}_k$ cette matrice et \mathbf{T}_k la transformation associée lorsque le rang est k . Supposons que la continuité et la concavité sont vérifiées jusqu'à l'ordre $k - 1$ et montrons alors la continuité et la concavité pour l'ordre k . On a

$$\mathbf{\Omega}_k = \mathbf{\Omega}_{k-1} + \lambda_k \mathbf{V}_k \mathbf{V}_k^T$$

Si on pose

$$\mathbf{S}_{k-1} = (\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Omega}_{k-1})^{-1} = \mathbf{T}_{k-1}(\boldsymbol{\Sigma})$$

la transformation \mathbf{T}_k devient

$$\mathbf{T}_k(\boldsymbol{\Sigma}) = (\mathbf{S}_{k-1}^{-1} + \lambda_k \mathbf{V}_k \mathbf{V}_k^T)^{-1} = \mathbf{T}_1(\mathbf{T}_{k-1}(\boldsymbol{\Sigma})) \quad [57]$$

Puisque \mathbf{T}_1 et \mathbf{T}_{k-1} (d'après l'hypothèse de récurrence) sont 2 transformations continues et concaves sur l'ensemble des matrices symétriques définies semi-positives, il en est de même de la transformation composée $\mathbf{T}_1 \circ \mathbf{T}_{k-1}$ par le lemme 3.

A.5 Démonstration du lemme 3

Soit $\boldsymbol{\Sigma}_1$ et $\boldsymbol{\Sigma}_2$ 2 matrices symétriques et définies semi-positives, \mathbf{T}_1 et \mathbf{T}_2 2 transformations concaves et continues sur l'ensemble des matrices symétriques et définies semi-positives. En vertu du théorème 2 elles sont donc croissantes. On peut alors écrire

$$\begin{aligned} (\mathbf{T}_2 \circ \mathbf{T}_1) \left(\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2} \right) &= \mathbf{T}_2 \left(\mathbf{T}_1 \left(\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2} \right) \right) \\ &\geq \mathbf{T}_2 \left(\frac{\mathbf{T}_1(\boldsymbol{\Sigma}_1) + \mathbf{T}_1(\boldsymbol{\Sigma}_2)}{2} \right) \\ &\geq \frac{\mathbf{T}_2(\mathbf{T}_1(\boldsymbol{\Sigma}_1)) + \mathbf{T}_2(\mathbf{T}_1(\boldsymbol{\Sigma}_2))}{2} \\ &= \frac{(\mathbf{T}_2 \circ \mathbf{T}_1)(\boldsymbol{\Sigma}_1) + (\mathbf{T}_2 \circ \mathbf{T}_1)(\boldsymbol{\Sigma}_2)}{2} \end{aligned} \quad [58]$$

D'où la concavité de $\mathbf{T}_2 \circ \mathbf{T}_1$. De plus la composée de transformations continues est continue.