

## Gestion des approvisionnements

Ce chapitre est consacré à l'étude de la gestion du stock d'un produit, une pièce de rechange automobile par exemple, sur plusieurs périodes de temps. Au début de chaque période, le gestionnaire du stock effectue une commande auprès de son fournisseur ; pendant la période, la quantité commandée est livrée et des clients formulent des demandes que le gestionnaire peut ou non satisfaire suivant le stock dont il dispose. On appelle stratégie du gestionnaire la manière dont il décide de la quantité commandée en fonction du passé. L'objectif du chapitre est de caractériser les stratégies optimales en termes de minimisation du coût total qui s'exprime comme la somme du coût d'achat du produit auprès du fournisseur, du coût de stockage et du coût associé aux demandes de clients qui n'ont pu être honorées faute de stock. Bien entendu, minimiser la somme du coût d'achat et du coût de stockage et minimiser le coût associé aux demandes non satisfaites sont des objectifs antagonistes.

Dans le premier paragraphe, nous commençons par résoudre le problème d'optimisation sur une seule période de temps avant de décrire précisément le problème dynamique de gestion de stock sur plusieurs périodes de temps. Le second paragraphe est consacré à une introduction au contrôle de chaînes de Markov, dans un cadre qui englobe le problème dynamique de gestion de stock. L'objectif est de montrer comment le principe de la programmation dynamique introduit par Bellman [1] à la fin des années 1950 permet de ramener l'étude d'un problème d'optimisation à  $N$  périodes de temps à celle de  $N$  problèmes à une seule période de temps. Nous illustrons sur le problème dit de la secrétaire comment ce principe permet d'explicitier la stratégie optimale d'un recruteur. On suppose que le recruteur sait classer les  $N$  candidats à un poste qui se présentent successivement pour passer un entretien avec lui et que tout candidat qui ne reçoit pas de réponse positive lors de son entretien trouve un emploi ailleurs. Pour maximiser la probabilité de choisir le meilleur des  $N$  candidats lors de cette procédure, le recruteur doit d'abord observer une proportion proche de  $1/e$  de candidats sans les recruter, puis choisir ensuite tout candidat meilleur que ceux qu'il observés (si aucun candidat meilleur ne se présente ensuite, il recrute le dernier candidat qu'il reçoit).

Enfin, dans le troisième paragraphe, nous utilisons les résultats du second paragraphe pour déterminer les stratégies optimales du gestionnaire de stock dans le modèle dynamique : à chaque instant, lorsque le stock est égal à  $x$ , le gestionnaire doit commander la quantité  $\mathbf{1}_{\{x \leq s\}}(S - x)^+$  qui permet de se ramener au stock objectif  $S$  si le stock est inférieur au seuil  $s$ , où le couple  $(s, S)$  peut dépendre ou non du temps.

### 3.1 Le modèle probabiliste de gestion de stock

#### 3.1.1 Le modèle à une période de temps

Ce modèle comporte un seul produit et une seule période de temps. Au début de la période, pour satisfaire les besoins de sa clientèle, un gestionnaire (par exemple un vendeur de journaux) commande une quantité  $q$  de produit qui lui est facturée au coût unitaire  $c > 0$  par le fournisseur. Il est souvent naturel de supposer que le produit se présente sous forme d'unités (cas des journaux par exemple), auquel cas  $q \in \mathbb{N}$ . Mais on peut aussi se placer dans un cadre continu (cas d'un liquide comme l'essence par exemple) et supposer  $q \in \mathbb{R}_+$ , ce qui simplifie parfois le problème d'optimisation. Comme la demande des clients sur la période de temps n'est pas connue du gestionnaire au moment où il effectue sa commande, il est naturel de la modéliser par une variable aléatoire  $D$  à valeurs dans  $\mathbb{N}$  ou  $\mathbb{R}_+$ . On suppose que  $D$  est d'espérance finie  $\mathbb{E}[D] = \mu$  et que l'on connaît sa loi au travers de sa fonction de répartition  $F(x) = \mathbb{P}(D \leq x)$ .

À l'issue de la période trois situations sont possibles :

- $q = D$  : le gestionnaire a visé juste.
- $q > D$  i.e. il y a  $(q - D)$  unités de produit en surplus. On associe à ce surplus un coût unitaire  $c_S$  qui correspond par exemple au coût de stockage sur la période. Notons que l'on peut supposer  $c_S$  négatif pour rendre compte de la possibilité de retourner le surplus au fournisseur (c'est le cas pour les journaux qui sont repris par les Nouvelles Messageries de la Presse Parisienne). Dans ce cas, il est naturel de supposer que  $c > -c_S$  i.e. que le prix auquel le fournisseur reprend le surplus est plus petit que le coût unitaire  $c$  auquel le gestionnaire se fournit. Nous supposons donc désormais que  $\boxed{c > 0 \text{ et } c > -c_S}$ .
- $q < D$  : on appelle manquants les demandes que le gestionnaire n'a pu satisfaire. On associe aux  $D - q$  manquants un coût unitaire  $\boxed{c_M \geq 0}$  qui rend compte de la détérioration de l'image du gestionnaire auprès des clients non servis.

L'objectif pour le gestionnaire est de trouver  $q \geq 0$  qui minimise l'espérance du coût total

$$g(q) = cq + c_S \mathbb{E}[(q - D)^+] + c_M \mathbb{E}[(D - q)^+] \quad (3.1)$$

où pour  $y \in \mathbb{R}$ ,  $y^+ = \max(y, 0)$  désigne la partie positive de  $y$ . Même si la quantité commandée est positive, nous allons supposer que la fonction  $g$  est définie sur  $\mathbb{R}$  car cela sera utile lorsque nous introduirons un coût fixe d'approvisionnement et un stock initial. Vérifions que  $g$  est continue. Comme la demande  $D$  est positive, pour  $q \leq 0$ ,  $(D-q)^+ = D-q$  et  $\mathbb{E}[(D-q)^+] = \mu - q$ . Donc  $q \rightarrow \mathbb{E}[(D-q)^+]$  est continue sur  $\mathbb{R}_-$ . Pour  $q \geq 0$ ,  $0 \leq (D-q)^+ \leq D$ . Comme  $q \rightarrow (D-q)^+$  est continue, on en déduit par convergence dominée que  $q \rightarrow \mathbb{E}[(D-q)^+]$  est continue sur  $\mathbb{R}_+$  et donc sur  $\mathbb{R}$ . L'égalité  $y^+ = y + (-y)^+$  entraîne que

$$\forall q \in \mathbb{R}, \mathbb{E}[(q-D)^+] = q - \mu + \mathbb{E}[(D-q)^+]. \quad (3.2)$$

On en déduit la continuité de  $q \rightarrow \mathbb{E}[(q-D)^+]$  puis celle de  $g$ .

Pour assurer que  $\inf_{q \geq 0} g(q)$  est atteint, il suffit maintenant de vérifier que  $\lim_{q \rightarrow +\infty} g(q) = +\infty$ . Pour montrer ce résultat, on distingue deux cas dans lesquels on utilise respectivement les inégalités  $c > 0$  et  $c > -c_S$  :

- si  $c_S \geq 0$ , alors  $g(q) \geq cq$ ,
- si  $c_S < 0$ , pour  $q \geq 0$ ,  $(q-D)^+ \leq q$  et donc  $\mathbb{E}[(q-D)^+] \leq q$  ce qui implique  $g(q) \geq (c + c_S)q$ .

Ainsi  $\inf_{q \geq 0} g(q)$  est atteint. La proposition suivante indique quelle quantité le gestionnaire doit commander pour minimiser le coût.

**Proposition 3.1.1.**

- Si  $c_M \leq c$  alors la fonction  $g$  est croissante et ne rien commander est optimal.
- Sinon,  $(c_M - c)/(c_M + c_S) \in ]0, 1[$  et si on pose

$$S = \inf\{z \in \mathbb{R} : F(z) \geq (c_M - c)/(c_M + c_S)\} \quad (3.3)$$

alors  $S \in \mathbb{R}_+$ . En outre, si  $D$  est une variable aléatoire entière i.e.  $\mathbb{P}(D \in \mathbb{N}) = 1$ , alors  $S \in \mathbb{N}$ . Enfin,  $g$  est décroissante sur  $] - \infty, S]$  et croissante sur  $[S, +\infty[$ , ce qui implique que commander  $S$  est optimal.

**Remarque 3.1.2.** L'hypothèse  $c_M > c$  qui rend le problème d'optimisation intéressant peut se justifier par des considérations économiques. En effet, il est naturel de supposer que le prix de vente unitaire du produit par le gestionnaire à ses clients est supérieur au coût  $c$  auquel il s'approvisionne. Comme la perte  $c_M$  correspondant à une unité de manquants est égale à la somme de ce prix de vente unitaire et du coût en termes d'image, on a alors  $c_M > c$ .  $\diamond$

*Démonstration.* D'après (3.1) et (3.2),

$$g(q) = c_M \mu + (c - c_M)q + (c_M + c_S)\mathbb{E}[(q-D)^+]. \quad (3.4)$$

Pour  $q \geq 0$ ,  $(q-D)^+ = \int_0^q \mathbf{1}_{\{z \geq D\}} dz$  ce qui implique en utilisant le théorème de Fubini,

$$\mathbb{E}[(q-D)^+] = \mathbb{E} \left[ \int_0^q \mathbf{1}_{\{z \geq D\}} dz \right] = \int_0^q \mathbb{E} [\mathbf{1}_{\{D \leq z\}}] dz = \int_0^q F(z) dz.$$

Comme  $D$  est une variable aléatoire positive, sa fonction de répartition  $F$  est nulle sur  $] -\infty, 0[$  et l'égalité précédente reste vraie pour  $q < 0$ . Donc

$$\forall q \in \mathbb{R}, g(q) = c_M \mu + \int_0^q ((c - c_M) + (c_M + c_S)F(z)) dz. \quad (3.5)$$

- Si  $c_M \leq c$  alors comme  $F$  est à valeurs dans  $[0, 1]$ , l'intégrande dans le membre de droite est minoré par  $c - c_M + \min(0, c_M + c_S) = \min(c - c_M, c + c_S)$ . Ce minorant est positif si bien que  $g$  est une fonction croissante et ne rien commander est optimal.
- Si  $c_M > c$ , comme  $c > -c_S$ ,  $0 < c_M - c < c_M + c_S$ . D'où

$$0 < \frac{c_M - c}{c_M + c_S} < 1.$$

Comme  $F$  est nulle sur  $] -\infty, 0[$ , l'ensemble  $\{z \in \mathbb{R} : F(z) \geq (c_M - c)/(c_M + c_S)\}$  est inclus dans  $\mathbb{R}_+$ . Puisque  $\lim_{z \rightarrow +\infty} F(z) = 1$ , il est non vide. Donc sa borne inférieure  $S$  définie par (3.3) est dans  $\mathbb{R}_+$ .

Lorsque la demande  $D$  est une variable entière, la fonction de répartition  $F$  est constante sur les intervalles  $[n, n+1[$ ,  $n \in \mathbb{N}$  et présente des sauts égaux à  $\mathbb{P}(D = n)$  aux points  $n \in \mathbb{N}$ . On en déduit que  $S = \min\{n \in \mathbb{N}, \sum_{k=0}^n \mathbb{P}(D = k) \geq (c_M - c)/(c_M + c_S)\}$  et que  $S \in \mathbb{N}$ .

L'intégrande dans le membre de droite de (3.5) est croissant, négatif pour  $z < S$  et positif pour  $z \geq S$ . Donc  $g$  est décroissante sur  $] -\infty, S]$ , croissante sur  $[S, +\infty[$  et commander  $S$  est optimal.

□

**Remarque 3.1.3.** – Lorsque  $c_M > c$ , si la fonction de répartition  $F$  de la demande  $D$  est continue (c'est le cas par exemple si la variable aléatoire  $D$  possède une densité), alors  $F(S) = (c_M - c)/(c_M + c_S)$ . Cette égalité peut se voir comme la condition d'optimalité du premier ordre  $g'(S) = 0$  puisque  $g'(q) = (c - c_M) + (c_M + c_S)F(q)$ . On peut la récrire

$$c + c_S \mathbb{P}(D \leq S) = c_M \mathbb{P}(D > S).$$

Elle a donc l'interprétation économique suivante : le surcoût moyen  $c + c_S \mathbb{P}(D \leq S)$  lié à la commande d'une unité supplémentaire est compensé par l'économie moyenne  $c_M \mathbb{P}(D > S)$  réalisée grâce à cette unité supplémentaire.

- La fonction  $q \rightarrow (q - D)^+$  est convexe sur  $\mathbb{R}$ . Lorsque  $c_M + c_S \geq 0$ , on en déduit en multipliant par  $c_M + c_S$  et en prenant l'espérance que  $q \rightarrow (c_M + c_S)\mathbb{E}[(q - D)^+]$  est convexe sur  $\mathbb{R}$ . Avec (3.4), on conclut que  $g$  est alors une fonction convexe sur  $\mathbb{R}$ .

◇

### Choix de la loi de la demande $D$

Il est souvent raisonnable de considérer que la demande  $D$  provient d'un grand nombre  $n$  de clients indépendants qui ont chacun une probabilité  $p$  de

commander une unité du produit. C'est par exemple le cas pour une pièce de rechange automobile : les  $n$  clients potentiels sont les détenteurs de la voiture pour laquelle la pièce est conçue. Dans ces conditions, la demande suit la loi binomiale de paramètre  $(n, p)$  i.e. pour  $0 \leq k \leq n$ , la probabilité qu'elle vaille  $k$  est donnée par  $\binom{n}{k} p^k (1-p)^{n-k}$ . En particulier  $\mu = \mathbb{E}[D] = np$ . La loi binomiale n'étant pas d'une manipulation très agréable, on pourra préférer les deux lois obtenues dans les passages à la limite suivants :

- $n$  grand ( $n \rightarrow +\infty$ ) et  $p$  petit ( $p \rightarrow 0$ ) avec  $np \rightarrow \mu > 0$ . Dans cette asymptotique, d'après l'exemple ??, la loi binomiale de paramètres  $(n, p)$  converge étroitement vers la loi de Poisson de paramètre  $\mu$ . On peut donc modéliser la demande comme une variable de Poisson de paramètre  $\mu$ .
- $n$  grand ( $n \rightarrow +\infty$ ) avec  $p > 0$  fixé, alors le théorème central limite ?? justifie l'utilisation de la loi gaussienne  $\mathcal{N}(\mu, \sigma^2)$  (avec  $\sigma^2 = \mu(1 - \mu/n)$ ) de densité  $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$  comme loi pour  $D$ . Une variable aléatoire gaussienne de variance  $\sigma^2 > 0$  a toujours une probabilité strictement positive de prendre des valeurs négatives. Mais dans les conditions d'application du théorème central limite, cette probabilité est très faible pour la loi limite  $\mathcal{N}(\mu, \sigma^2)$  et la commande de la quantité  $S$  donnée par la proposition 3.1.1 est une bonne stratégie.

### Taux de manquants

On suppose  $c_M > c$ . Du point de vue du gestionnaire, le taux de manquants i.e. le taux de demandes de clients non satisfaites est un indicateur important. Lorsqu'il a commandé  $S$ , ce taux est égal à  $\frac{(D-S)^+}{D}$ . On a

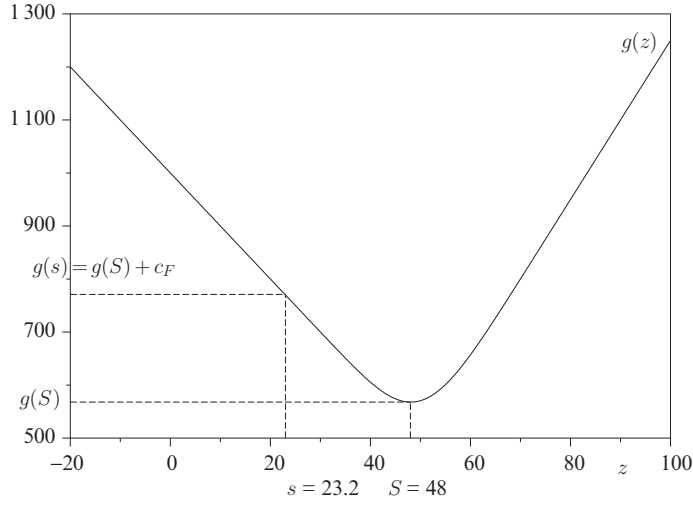
$$\mathbb{E} \left[ \frac{(D-S)^+}{D} \right] \leq \mathbb{E} [\mathbf{1}_{\{D>S\}}] = \mathbb{P}(D > S) = 1 - F(S) \leq \frac{c + c_S}{c_M + c_S},$$

la dernière inégalité étant une égalité si  $F$  est continue en  $S$ . En général, l'espérance du taux de manquants est même significativement plus petite que la probabilité pour qu'il y ait des manquants.

**Exemple 3.1.4.** Dans le cas particulier où  $D$  suit la loi de Poisson de paramètre 50,  $c = 10$ ,  $c_M = 20$  et  $c_S = 5$ , on vérifie numériquement que le stock objectif vaut  $S = 48$ . Le rapport  $\frac{c + c_S}{c_M + c_S}$  qui vaut 0.6 est légèrement supérieur à  $\mathbb{P}(D > S) \simeq 0.575$  mais très sensiblement supérieur au taux de manquants qui est égal à 6.8 %. La fonction  $g$  correspondant à ce cas particulier est représentée sur la Fig. 3.1.  $\diamond$

### Résolution avec stock initial et coût fixe d'approvisionnement

On suppose maintenant que si le gestionnaire commande une quantité  $q$  non nulle auprès de son fournisseur, il doit payer un coût fixe  $c_F$  positif en plus



**Fig. 3.1.** Représentation de  $g(z)$ ,  $s$  et  $S$  ( $c = 10$ ,  $c_M = 20$ ,  $c_S = 5$ ,  $c_F = 200$ ,  $D$  distribuée suivant la loi de Poisson de paramètre 50)

du coût unitaire  $c$ . On suppose également que le stock initial  $x$  n'est pas nécessairement nul. On autorise même la situation où  $x$  est négatif qui traduit le fait qu'une quantité égale à  $-x$  de demandes formulées par des clients avant le début de la période n'ont pu être satisfaites et sont maintenues par ces clients. Le problème est maintenant de trouver la quantité commandée  $q \geq 0$  qui minimise

$$c_F \mathbf{1}_{\{q > 0\}} + cq + c_S \mathbb{E}[(x + q - D)^+] + c_M \mathbb{E}[(D - x - q)^+]. \quad (3.6)$$

**Proposition 3.1.5.** *On suppose  $c_M > c$ . Alors l'ensemble*

$$\{z \in ]-\infty, S], g(z) \geq c_F + g(S)\}$$

*est non vide. Si on note  $s$  sa borne supérieure, la stratégie  $(s, S)$  qui consiste à commander la quantité  $q = (S - x)$  permettant d'atteindre le stock objectif  $S$  lorsque  $x$  est inférieur au stock seuil  $s$  et à ne rien faire ( $q = 0$ ) sinon est optimale au sens où elle minimise (3.6).*

La figure 3.1 représente la fonction  $g$  et fournit une interprétation graphique du couple  $(s, S)$  dans le cas particulier où  $D$  suit la loi de Poisson de paramètre 50,  $c = 10$ ,  $c_M = 20$ ,  $c_S = 5$  et  $c_F = 200$ .

*Démonstration.* Lorsque  $z \leq 0$ , on a  $(z - D)^+ = 0$  et  $(D - z)^+ = D - z$ , d'où pour la fonction  $g$  définie par (3.1),

$$\forall z \leq 0, g(z) = cz + c_M(\mathbb{E}[D] - z) = (c - c_M)z + c_M \mu.$$

Comme  $c_M > c$ , cela implique que  $\lim_{z \rightarrow -\infty} g(z) = +\infty$ . Ainsi l'ensemble  $\{z \in ]-\infty, S], g(z) \geq c_F + g(S)\}$  est non vide et sa borne supérieure  $s$  est dans  $] -\infty, S]$ .

On ne change rien à la quantité  $q$  optimale en ajoutant  $cx$  au coût (3.6). Le critère à minimiser devient  $c_F \mathbf{1}_{\{q>0\}} + g(x+q)$ . Comme d'après la proposition 3.1.1,  $g$  est croissante sur  $[S, +\infty[$ , il est clairement optimal de ne rien commander si  $x \geq S$ .

Pour  $x < S$ , comme  $g$  atteint son minimum en  $S$ , il est optimal de commander  $S - x$  si  $g(x) \geq c_F + g(S)$  et de ne rien commander sinon. Notons que lorsque le produit se présente sous forme d'unités, le stock initial  $x$  est entier tout comme le stock objectif  $S$  (voir la proposition 3.1.1), ce qui assure que le gestionnaire peut bien commander la quantité  $S - x$  qui est entière. La décroissance et la continuité de  $g$  sur  $] -\infty, S]$  entraînent que  $\{x \in ] -\infty, S], g(x) \geq c_F + g(S)\} = ] -\infty, s]$ , ce qui achève la démonstration.  $\square$

**Remarque 3.1.6.**

- Si  $c_F = 0$ , alors  $s = S$ . Comme la stratégie  $(S, S)$  consiste à commander  $(S - x)$  lorsque  $x \leq S$ , on retrouve bien le résultat de la proposition 3.1.1.
- Le coût minimal associé à la stratégie  $(s, S)$  est

$$\begin{aligned} u(x) &= -cx + \inf_{q \geq 0} (c_F \mathbf{1}_{\{q>0\}} + g(x+q)) \\ &= -cx + (g(S) + c_F) \mathbf{1}_{\{x \leq s\}} + g(x) \mathbf{1}_{\{x > s\}}. \end{aligned}$$

Dans la définition ??, nous introduirons une notion de convexité spécifique à la gestion des approvisionnements. Le coût minimal  $u(x)$  fournira un exemple typique de fonction satisfaisant cette propriété de convexité.  $\diamond$

### 3.1.2 Le modèle dynamique de gestion de stock

Le modèle comporte toujours un seul produit mais  $N$  périodes de temps (typiquement des journées, des semaines ou des mois). Pour  $t \in \{0, \dots, N-1\}$ , une demande  $D_{t+1}$  positive est formulée par les clients sur la période  $[t, t+1]$ . Les variables aléatoires  $D_1, D_2, \dots, D_N$  sont supposées indépendantes et identiquement distribuées de fonction de répartition  $F$  et d'espérance finie i.e.  $\mathbb{E}[D_1] = \mu < +\infty$ .

Au début de chaque période  $[t, t+1]$ , le gestionnaire décide de commander une quantité  $Q_t \geq 0$  qui lui est livrée pendant la période. Plutôt que le stock physique qui est toujours positif ou nul la grandeur intéressante à considérer est le **stock système** qui peut prendre des valeurs négatives. Le stock système est défini comme le stock physique si celui-ci est strictement positif et comme moins la quantité de manquants (i.e. les demandes de clients qui n'ont pu être honorées faute de stock) sinon. On note  $X_t$  le stock système à l'instant  $t \in \{0, \dots, N\}$ .

On suppose que

- le gestionnaire ne renvoie pas de produit à son fournisseur,
- les demandes correspondant aux manquants sont maintenues par les clients jusqu'à ce que le gestionnaire puisse les servir.

Ainsi la dynamique du stock système est donnée par

$$X_{t+1} = X_t + Q_t - D_{t+1}. \quad (3.7)$$

Au cours de la période  $[t, t+1]$ , le coût est donné par une formule analogue à celle du modèle à une seule période :  $c_F \mathbf{1}_{\{Q_t > 0\}} + cQ_t + c_S(X_{t+1})^+ + c_M(X_{t+1})^-$  où  $y^- = \max(-y, 0)$ . On associe au stock système final  $X_N$  un coût  $u_N(X_N)$  (par exemple, estimer que ce stock final a une valeur unitaire égale à  $c$  consiste à poser  $u_N(x) = -cx$ ). Enfin pour traduire la préférence du gestionnaire pour 1 Euro à la date  $t$  par rapport à 1 Euro à la date  $t+1$ , on utilise un facteur d'actualisation  $\alpha \in [0, 1]$ . L'espérance du coût total actualisé est donnée par

$$\mathbb{E} \left[ \sum_{t=0}^{N-1} \alpha^t (c_F \mathbf{1}_{\{Q_t > 0\}} + cQ_t + c_S(X_{t+1})^+ + c_M(X_{t+1})^-) + \alpha^N u_N(X_N) \right]. \quad (3.8)$$

L'objectif du gestionnaire est de choisir les quantités  $Q_t$  au vu du passé jusqu'à l'instant  $t$  de manière à minimiser l'espérance du coût total. Le but du paragraphe suivant est de montrer que la résolution d'un tel problème d'optimisation à  $N$  périodes de temps peut se ramener à la résolution de  $N$  problèmes à une période de temps définis par récurrence descendante.

## 3.2 Éléments de contrôle de chaînes de Markov

Nous allons dans ce paragraphe décrire puis résoudre un problème de contrôle de chaînes de Markov qui englobe le problème dynamique de gestion de stock que nous venons de présenter.

### 3.2.1 Description du modèle

On considère un modèle d'évolution d'un système commandé par un gestionnaire sur  $N$  périodes de temps. On note  $\mathcal{E}$  l'ensemble des états possibles du système et  $\mathcal{A}$  l'ensemble des actions possibles du gestionnaire. Ces deux ensembles sont supposés discrets. L'état du système à l'instant  $t \in \{0, \dots, N\}$  est noté  $X_t$  tandis que l'action choisie par le gestionnaire à l'instant  $t \in \{0, \dots, N-1\}$  est notée  $A_t$ . Pour  $t \in \{0, \dots, N-1\}$ , le gestionnaire choisit l'action  $A_t$  au vu de l'histoire  $H_t = (X_0, A_0, X_1, A_1, \dots, X_{t-1}, A_{t-1}, X_t) \in (\mathcal{E} \times \mathcal{A})^t \times \mathcal{E}$  jusqu'à l'instant  $t$ . On suppose que l'état  $X_{t+1}$  du système à l'instant  $t+1$  ne dépend de l'histoire  $H_t$  et de l'action  $A_t$  qu'au travers du couple  $(X_t, A_t) : \forall t \in \{0, \dots, N-1\}$ ,



$$\begin{aligned}
\forall h_t = (x_0, a_0, x_1, a_1, \dots, x_t) \in (\mathcal{E} \times \mathcal{A})^t \times \mathcal{E}, \forall a_t \in \mathcal{A}, \forall x_{t+1} \in \mathcal{E}, \\
\mathbb{P}(X_{t+1} = x_{t+1} | H_t = h_t, A_t = a_t) = \mathbb{P}(X_{t+1} = x_{t+1} | X_t = x_t, A_t = a_t) \\
= p_t((x_t, a_t), x_{t+1}), \tag{3.9}
\end{aligned}$$

où la matrice  $p_t : (\mathcal{E} \times \mathcal{A}) \times \mathcal{E} \rightarrow [0, 1]$  vérifie  $\sum_{y \in \mathcal{E}} p_t((x, a), y) = 1$  pour tout  $(x, a) \in \mathcal{E} \times \mathcal{A}$ .

Pour  $t \in \{0, \dots, N-1\}$ , le choix de l'action  $A_t$  induit un coût égal à  $\varphi_t(X_t, A_t, X_{t+1})$  sur la période  $[t, t+1]$  où  $\varphi_t : \mathcal{E} \times \mathcal{A} \times \mathcal{E} \rightarrow \mathbb{R}$ . A l'instant final, un coût  $u_N(X_N)$  est associé à l'état final  $X_N$  avec  $u_N : \mathcal{E} \rightarrow \mathbb{R}$ . L'espérance du coût total actualisé avec le facteur d'actualisation  $\alpha \in [0, 1]$  est donnée par

$$\mathbb{E} \left[ \sum_{n=0}^{N-1} \alpha^n \varphi_n(X_n, A_n, X_{n+1}) + \alpha^N u_N(X_N) \right].$$

**Exemple 3.2.1.** Le problème dynamique de gestion de stock décrit au paragraphe 3.1.2 entre dans ce cadre lorsque le produit se présente sous forme d'unités. L'espace d'état est  $\mathcal{E} = \mathbb{Z}$  pour le stock système  $X_t$  et l'ensemble d'actions  $\mathcal{A} = \mathbb{N}$  pour la quantité  $Q_t$  de produit commandée. On a

$$X_{t+1} = X_t + Q_t - D_{t+1}$$

où les demandes  $(D_1, \dots, D_N)$  sont des variables aléatoires positives indépendantes et identiquement distribuées que l'on suppose à valeurs dans  $\mathbb{N}$ . Donc pour  $t \in \{0, \dots, N-1\}$  et  $h_t = (x_0, q_0, x_1, q_1, \dots, x_t) \in (\mathbb{Z} \times \mathbb{N})^t \times \mathbb{Z}$ , on a

$$\begin{aligned}
\mathbb{P}(X_{t+1} = x_{t+1} | H_t = h_t, Q_t = q_t) \\
= \mathbb{P}(X_t + Q_t - D_{t+1} = x_{t+1} | H_t = h_t, Q_t = q_t) \\
= \frac{\mathbb{P}(D_{t+1} = x_t + q_t - x_{t+1}, H_t = h_t, Q_t = q_t)}{\mathbb{P}(H_t = h_t, Q_t = q_t)} \\
= \mathbb{P}(D_{t+1} = x_t + q_t - x_{t+1}) = \mathbb{P}(D_1 = x_t + q_t - x_{t+1}),
\end{aligned}$$

car les variables  $H_t$  et  $Q_t$  ne dépendent que de  $D_1, D_2, \dots, D_t$  et sont indépendantes de  $D_{t+1}$ . Ainsi (3.9) est vérifié pour  $p_t((x, q), y) = \mathbb{P}(D_1 = x + q - y)$ . Enfin la fonction  $\varphi_t$  qui intervient dans l'expression de l'espérance du coût total actualisé est donnée par :  $\forall (t, x, q, y) \in \{0, \dots, N-1\} \times \mathbb{Z} \times \mathbb{N} \times \mathbb{Z}$ ,

$$\varphi_t(x, q, y) = c_F \mathbf{1}_{\{q > 0\}} + cq + c_S y^+ + c_M y^-.$$

Elle ne dépend que des deux dernières variables.  $\diamond$

Le fait que le gestionnaire décide de l'action  $A_t$  au vu de l'histoire  $H_t$  se traduit par l'existence d'une application  $d_t$  de l'ensemble  $(\mathcal{E} \times \mathcal{A})^t \times \mathcal{E}$  des histoires possibles jusqu'à l'instant  $t$  dans l'ensemble  $\mathcal{A}$  des actions telle que  $A_t = d_t(H_t)$ . Cette application  $d_t$  est appelée règle de décision du gestionnaire à l'instant  $t$ . La stratégie  $\pi = (d_0, \dots, d_{N-1})$  du gestionnaire est constituée

de l'ensemble de ses règles de décision. Bien sûr, la dynamique du système  $X_t$ ,  $t \in \{0, \dots, N\}$  dépend de la stratégie du gestionnaire. Pour expliciter cette dépendance on note désormais  $X_t^\pi$  et  $H_t^\pi$  l'état du système et l'histoire à l'instant  $t \in \{0, \dots, N\}$  qui correspondent à la stratégie  $\pi$ . Sachant que l'état initial du système est  $x_0$ , le coût moyen associé à la stratégie  $\pi$  est donné par :

$$v_0^\pi(x_0) = \mathbb{E} \left[ \sum_{n=0}^{N-1} \alpha^n \varphi_n(X_n^\pi, d_n(H_n^\pi), X_{n+1}^\pi) + \alpha^N u_N(X_N^\pi) \middle| X_0^\pi = x_0 \right].$$

On note

$$v_0^*(x_0) = \inf_{\pi} v_0^\pi(x_0).$$

L'objectif du gestionnaire est de trouver à l'instant initial une stratégie optimale  $\pi^*$  qui réalise l'infimum lorsque celui-ci est atteint ou une stratégie  $\varepsilon$ -optimale  $\pi^\varepsilon$  vérifiant  $\forall x_0 \in \mathcal{E}$ ,  $v_0^{\pi^\varepsilon}(x_0) \leq v_0^*(x_0) + \varepsilon$  avec  $\varepsilon > 0$  arbitraire lorsque l'infimum n'est pas atteint.

Dans le paragraphe 3.2.2 nous allons montrer comment évaluer le coût associé à une stratégie par récurrence descendante. Puis dans le paragraphe 3.2.3, nous en déduirons les équations d'optimalité et nous montrerons que leur résolution permet de construire des stratégies optimales ou  $\varepsilon$ -optimales. Enfin, dans le paragraphe ??, nous appliquerons la théorie développée pour déterminer la stratégie optimale d'un recruteur dans « le problème de la secrétaire ».

### 3.2.2 Évaluation du coût associé à une stratégie

Nous allons montrer que le coût moyen associé à la stratégie  $\pi$  peut être évalué par récurrence descendante. On introduit pour cela le coût à venir à l'instant  $t \in \{1, \dots, N\}$  sachant que l'histoire est  $h_t \in (\mathcal{E} \times \mathcal{A})^t \times \mathcal{E}$  :

$$v_t^\pi(h_t) = \mathbb{E} \left[ \sum_{n=t}^{N-1} \alpha^{n-t} \varphi_n(X_n^\pi, d_n(H_n^\pi), X_{n+1}^\pi) + \alpha^{N-t} u_N(X_N^\pi) \middle| H_t^\pi = h_t \right].$$

Cette notation est bien compatible avec la définition de  $v_0^\pi$  au paragraphe précédent. La proposition suivante indique comment exprimer  $v_t^\pi$  en fonction de  $v_{t+1}^\pi$  :

#### Proposition 3.2.2.

$$\forall h_N = (x_0, a_0, x_1, a_1, \dots, x_N) \in (\mathcal{E} \times \mathcal{A})^N \times \mathcal{E}, \quad v_N^\pi(h_N) = u_N(x_N).$$

En outre, pour  $t \in \{0, \dots, N-1\}$ ,  $\forall h_t = (x_0, a_0, x_1, a_1, \dots, x_t) \in (\mathcal{E} \times \mathcal{A})^t \times \mathcal{E}$ ,

$$v_t^\pi(h_t) = \sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, d_t(h_t)), x_{t+1})) \left[ \varphi_t(x_t, d_t(h_t), x_{t+1}) + \alpha v_{t+1}^\pi((h_t, d_t(h_t), x_{t+1})) \right].$$

La formule précédente peut s'interpréter de la façon suivante : lorsque l'histoire jusqu'à l'instant  $t$  est  $h_t \in (\mathcal{E} \times \mathcal{A})^t \times \mathcal{E}$ , sous la stratégie  $\pi$ , l'action du gestionnaire en  $t$  est  $d_t(h_t)$ . Pour tout  $x_{t+1} \in \mathcal{E}$ , l'état  $X_{t+1}^\pi$  du système à l'instant  $t+1$  et l'histoire  $H_{t+1}^\pi$  jusqu'à l'instant  $t+1$  sont donc respectivement égaux à  $x_{t+1}$  et à  $(h_t, d_t(h_t), x_{t+1}) \in (\mathcal{E} \times \mathcal{A})^{t+1} \times \mathcal{E}$  avec probabilité  $p_t((x_t, d_t(h_t)), x_{t+1})$ . Le coût à venir à l'instant  $t$  sachant que l'histoire est  $h_t$  se décompose donc comme la somme sur les états possibles  $x_{t+1} \in \mathcal{E}$  du système à l'instant  $t+1$  pondérée par  $p_t((x_t, d_t(h_t)), x_{t+1})$  du coût  $\varphi_t(x_t, d_t(h_t), x_{t+1})$  sur la période  $[t, t+1]$  plus le coût actualisé  $\alpha v_{t+1}^\pi((h_t, d_t(h_t), x_{t+1}))$  à venir à l'instant  $t+1$  sachant que l'histoire est  $(h_t, d_t(h_t), x_{t+1})$ .

*Démonstration.* Par définition,

$$v_N^\pi(h_N) = \mathbb{E}[u_N(X_N^\pi) | H_N^\pi = h_N] = \frac{\mathbb{E}[u_N(X_N^\pi) \mathbf{1}_{\{H_N^\pi = h_N\}}]}{\mathbb{P}(H_N^\pi = h_N)}.$$

Comme  $H_N^\pi = h_N = (x_0, a_0, x_1, a_1, \dots, x_N)$  implique  $X_N^\pi = x_N$ , on a

$$\mathbb{E}[u_N(X_N^\pi) \mathbf{1}_{\{H_N^\pi = h_N\}}] = \mathbb{E}[u_N(x_N) \mathbf{1}_{\{H_N^\pi = h_N\}}] = u_N(x_N) \mathbb{P}(H_N^\pi = h_N),$$

et on conclut que  $v_N^\pi(h_N) = u_N(x_N)$ .

Soit  $t \in \{0, \dots, N-1\}$ . Par linéarité de l'espérance,

$$\begin{aligned} \mathbb{P}(H_t^\pi = h_t) v_t^\pi(h_t) &= \mathbb{E}\left[\mathbf{1}_{\{H_t^\pi = h_t\}} \left( \sum_{n=t}^{N-1} \alpha^{n-t} \varphi_n(X_n^\pi, d_n(H_n^\pi), X_{n+1}^\pi) \right. \right. \\ &\quad \left. \left. + \alpha^{N-t} u_N(X_N^\pi) \right) \right] \\ &= \sum_{x_{t+1} \in \mathcal{E}} \mathbb{E}\left[\mathbf{1}_{\{X_{t+1}^\pi = x_{t+1}, H_t^\pi = h_t\}} \left( \sum_{n=t}^{N-1} \alpha^{n-t} \varphi_n(X_n^\pi, d_n(H_n^\pi), X_{n+1}^\pi) \right. \right. \\ &\quad \left. \left. + \alpha^{N-t} u_N(X_N^\pi) \right) \right] \end{aligned}$$

Comme sous la stratégie  $\pi$ ,  $H_t^\pi = h_t$  implique  $A_t = d_t(h_t)$ ,  $(H_t^\pi = h_t, X_{t+1}^\pi = x_{t+1})$  implique  $H_{t+1}^\pi = (h_t, d_t(h_t), x_{t+1})$ . Donc

$$\begin{aligned} \mathbb{P}(H_t^\pi = h_t) v_t^\pi(h_t) &= \sum_{x_{t+1} \in \mathcal{E}} \mathbb{E}\left[\mathbf{1}_{\{H_{t+1}^\pi = (h_t, d_t(h_t), x_{t+1})\}} \left( \varphi_t(x_t, d_t(h_t), x_{t+1}) \right. \right. \\ &\quad \left. \left. + \alpha \left\{ \sum_{n=t+1}^{N-1} \alpha^{n-t-1} \varphi_n(X_n^\pi, d_n(H_n^\pi), X_{n+1}^\pi) + \alpha^{N-t-1} u_N(X_N^\pi) \right\} \right) \right] \\ &= \sum_{x_{t+1} \in \mathcal{E}} \mathbb{P}(H_{t+1}^\pi = (h_t, d_t(h_t), x_{t+1})) \left[ \varphi_t(h_t, d_t(h_t), x_{t+1}) \right. \\ &\quad \left. + \alpha v_{t+1}^\pi((h_t, d_t(h_t), x_{t+1})) \right]. \end{aligned}$$

Il suffit de remarquer que

$$\begin{aligned} \frac{\mathbb{P}(H_{t+1}^\pi = (h_t, d_t(h_t), x_{t+1}))}{\mathbb{P}(H_t^\pi = h_t)} &= \frac{\mathbb{P}(X_{t+1}^\pi = x_{t+1}, H_t^\pi = h_t, A_t = d_t(h_t))}{\mathbb{P}(H_t^\pi = h_t, A_t = d_t(h_t))} \\ &= \mathbb{P}(X_{t+1}^\pi = x_{t+1} | H_t^\pi = h_t, A_t = d_t(h_t)) \\ &= p_t((x_t, d_t(h_t)), x_{t+1}) \end{aligned}$$

d'après (3.9) pour conclure la démonstration de la formule de récurrence.  $\square$

### 3.2.3 Équations d'optimalité

On définit  $v_t^*(h_t) = \inf_\pi v_t^\pi(h_t)$ . Le théorème suivant explique comment évaluer  $v_t^*$  par récurrence descendante :

**Théorème 3.2.3.** *Pour tout  $t \in \{0, \dots, N\}$ ,*

$$\forall h_t = (x_0, a_0, x_1, a_1, \dots, x_t) \in (\mathcal{E} \times \mathcal{A})^t \times \mathcal{E}, \quad v_t^*(h_t) = u_t(x_t)$$

où les fonctions  $u_t$  sont définies par récurrence descendante à partir du coût terminal  $u_N$  par les **équations d'optimalité** : pour  $t \in \{0, \dots, N-1\}$ ,

$$\forall x_t \in \mathcal{E}, \quad u_t(x_t) = \inf_{a \in \mathcal{A}} \sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, a), x_{t+1}) [\varphi_t(x_t, a, x_{t+1}) + \alpha u_{t+1}(x_{t+1})]. \quad (3.10)$$

*Démonstration.* D'après la proposition 3.2.2,  $\forall \pi, v_N^\pi(h_N) = u_N(x_N)$ . Donc  $v_N^*(h_N) = u_N(x_N)$ , ce qui permet d'initialiser la démonstration par récurrence descendante de l'assertion  $\forall t \in \{0, \dots, N\}, v_t^*(h_t) \geq u_t(x_t)$ . Supposons que l'hypothèse de récurrence est vraie à l'instant  $t+1$ .

Soit  $\pi$  une stratégie. On a  $v_{t+1}^\pi(h_{t+1}) \geq v_{t+1}^*(h_{t+1}) \geq u_{t+1}(x_{t+1})$ . En insérant cette inégalité dans la formule de récurrence de la proposition 3.2.2, on obtient

$$\begin{aligned} v_t^\pi(h_t) &\geq \sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, d_t(h_t)), x_{t+1}) [\varphi_t(x_t, d_t(h_t), x_{t+1}) + \alpha u_{t+1}(x_{t+1})] \\ &\geq \inf_{a \in \mathcal{A}} \sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, a), x_{t+1}) [\varphi_t(x_t, a, x_{t+1}) + \alpha u_{t+1}(x_{t+1})] = u_t(x_t). \end{aligned}$$

Comme la stratégie  $\pi$  est arbitraire, on conclut que  $v_t^*(h_t) = \inf_\pi v_t^\pi(h_t) \geq u_t(x_t)$  i.e. que l'hypothèse de récurrence est vérifiée au rang  $t$ .

Pour montrer l'inégalité inverse, on fixe  $\gamma > 0$ . Pour  $t \in \{0, \dots, N-1\}$  et  $x_t \in \mathcal{E}$ , d'après (3.10), il existe  $\delta_t(x_t) \in \mathcal{A}$  t.q.

$$\sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, \delta_t(x_t)), x_{t+1}) [\varphi_t(x_t, \delta_t(x_t), x_{t+1}) + \alpha u_{t+1}(x_{t+1})] \leq u_t(x_t) + \gamma.$$

Notons  $\pi^\gamma$  la stratégie telle que la règle de décision à chaque instant  $t$  dans  $\{0, \dots, N-1\}$  est  $d_t(h_t) = \delta_t(x_t)$  et montrons par récurrence descendante que

$$\forall t \in \{0, \dots, N\}, \forall h_t = (x_0, a_0, x_1, a_1, \dots, x_t), v_t^{\pi^\gamma}(h_t) \leq u_t(x_t) + (N-t)\gamma.$$

On a  $v_N^{\pi^\gamma}(h_N) = u_N(x_N)$ . Supposons que l'hypothèse est vérifiée pour  $n$  dans  $\{t+1, \dots, N\}$ . En insérant l'inégalité  $v_{t+1}^{\pi^\gamma}(h_{t+1}) \leq u_{t+1}(x_{t+1}) + (N-t-1)\gamma$  dans la formule de récurrence de la proposition 3.2.2, on obtient

$$\begin{aligned} v_t^{\pi^\gamma}(h_t) &\leq \sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, d_t(h_t)), x_{t+1}) \\ &\quad [\varphi_t(x_t, d_t(h_t), x_{t+1}) + \alpha (u_{t+1}(x_{t+1}) + (N-t-1)\gamma)] \end{aligned}$$

En utilisant successivement  $\sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, d_t(h_t)), x_{t+1}) = 1$ , la définition de  $d_t$  et  $\alpha \leq 1$ , on en déduit que

$$\begin{aligned} v_t^{\pi^\gamma}(h_t) &\leq \sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, \delta_t(x_t)), x_{t+1}) [\varphi_t(x_t, \delta_t(x_t), x_{t+1}) + \alpha u_{t+1}(x_{t+1})] \\ &\quad + \alpha(N-t-1)\gamma \\ &\leq u_t(x_t) + \gamma + (N-t-1)\gamma. \end{aligned}$$

Donc l'hypothèse de récurrence est vérifiée au rang  $t$ .

Pour  $t \in \{0, \dots, N\}$  et  $h_t \in (\mathcal{E} \times \mathcal{A})^t \times \mathcal{E}$ , on conclut que

$$v_t^*(h_t) \leq \inf_{\gamma > 0} v_t^{\pi^\gamma}(h_t) \leq \inf_{\gamma > 0} (u_t(x_t) + (N-t)\gamma) = u_t(x_t).$$

□

#### Définition 3.2.4.

1. Une stratégie  $\pi^*$  est dite optimale si

$$\forall t \in \{0, \dots, N\}, \forall h_t \in (\mathcal{E} \times \mathcal{A})^t \times \mathcal{E}, v_t^{\pi^*}(h_t) = \inf_{\pi} v_t^{\pi}(h_t) = v_t^*(h_t).$$

2. Une stratégie  $\pi^\varepsilon$  est dite  $\varepsilon$ -optimale si

$$\forall t \in \{0, \dots, N\}, \forall h_t \in (\mathcal{E} \times \mathcal{A})^t \times \mathcal{E}, v_t^{\pi^\varepsilon}(h_t) \leq v_t^*(h_t) + \varepsilon.$$

3. Une stratégie  $\pi_M = (d_0, \dots, d_{N-1})$  est dite markovienne si pour tout  $t$  dans  $\{0, \dots, N-1\}$ , la règle de décision à l'instant  $t$  ne dépend du passé qu'au travers de l'état du système à l'instant  $t$  i.e. il existe  $\delta_t : \mathcal{E} \rightarrow \mathcal{A}$  t.q.  $\forall h_t = (x_0, a_0, x_1, a_1, \dots, x_t), d_t(h_t) = \delta_t(x_t)$ .

**Remarque 3.2.5.**

- Les équations d’optimalité (3.10) expriment que pour qu’une stratégie soit optimale sur la période  $[t, N]$ , il faut que la décision prise en  $t$  soit optimale mais aussi que toutes les décisions ultérieures le soient. Ainsi, lorsqu’une stratégie  $\pi$  est optimale au sens introduit à la fin du paragraphe 3.2.1, à savoir  $v_0^\pi(x_0) = v_0^*(x_0)$ , elle est aussi optimale au sens du point 1 de la définition précédente (qui peut sembler plus exigeant en première lecture).
- Les stratégies markoviennes sont particulièrement intéressantes car il n’est pas nécessaire de garder en mémoire tout le passé pour les appliquer. La terminologie markovienne provient de ce que sous une telle stratégie  $\pi_M = (\delta_0, \dots, \delta_{N-1})$ ,  $H_t^{\pi_M} = h_t$  entraîne  $A_t = \delta_t(x_t)$  et donc

$$\begin{aligned} \mathbb{P}(X_{t+1}^{\pi_M} = x_{t+1} | H_t^{\pi_M} = h_t) \\ = \mathbb{P}(X_{t+1}^{\pi_M} = x_{t+1} | H_t^{\pi_M} = h_t, A_t = \delta_t(x_t)) = p_t((x_t, \delta_t(x_t)), x_{t+1}) \end{aligned}$$

d’après (3.9). Ainsi l’état  $X_{t+1}^{\pi_M}$  du système à l’instant  $t + 1$  ne dépend du passé  $H_t^{\pi_M}$  qu’au travers de l’état  $X_t^{\pi_M}$  du système à l’instant  $t$  i.e. la suite  $(X_t^{\pi_M})_{t \in \{0, \dots, N\}}$  est une chaîne de Markov.

◇

Il faut noter que pour  $\gamma = \varepsilon/N$ , la stratégie  $\pi^\gamma$  qui a été construite dans la démonstration du théorème 3.2.3 est  $\varepsilon$ -optimale et markovienne, ce qui fournit la première assertion du corollaire suivant :

**Corollaire 3.2.6.**

1. Pour tout  $\varepsilon > 0$ , il existe une stratégie markovienne  $\varepsilon$ -optimale.
2. Si  $\mathcal{A}$  est fini ou bien s’il existe une stratégie  $\pi^* = (d_0, \dots, d_{N-1})$  optimale, alors il existe une stratégie markovienne optimale.

*Démonstration.* Si  $\mathcal{A}$  est fini, alors pour tout  $t \in \{0, \dots, N-1\}$  et tout  $x_t \in \mathcal{E}$ , l’infimum de l’application

$$a \in \mathcal{A} \rightarrow \sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, a), x_{t+1}) [\varphi_t(x_t, a, x_{t+1}) + \alpha u_{t+1}(x_{t+1})] \quad (3.11)$$

est un minimum i.e. il est atteint. Vérifions maintenant que cette propriété reste vraie s’il existe une stratégie  $\pi^* = (d_0, \dots, d_{N-1})$  optimale. Pour  $h_t = (x_0, a_0, x_1, a_1, \dots, x_t)$ , d’après le théorème 3.2.3,  $u_t(x_t) = v_t^*(h_t) = v_t^{\pi^*}(h_t)$ . En utilisant l’équation d’optimalité (3.10) et la proposition 3.2.2, on en déduit que

$$\begin{aligned} \inf_{a \in \mathcal{A}} \sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, a), x_{t+1}) [\varphi_t(x_t, a, x_{t+1}) + \alpha u_{t+1}(x_{t+1})] &= u_t(x_t) \\ &= v_t^{\pi^*}(h_t) = \sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, d_t(h_t)), x_{t+1}) \left[ \varphi_t(x_t, d_t(h_t), x_{t+1}) \right. \\ &\quad \left. + \alpha v_{t+1}^{\pi^*}((h_t, d_t(h_t), x_{t+1})) \right]. \end{aligned}$$

Comme  $v_{t+1}^*((h_t, d_t(h_t), x_{t+1})) = v_{t+1}^*((h_t, d_t(h_t), x_{t+1})) = u_{t+1}(x_{t+1})$ , on en déduit que

$$\begin{aligned} & \inf_{a \in \mathcal{A}} \sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, a), x_{t+1}) [\varphi_t(x_t, a, x_{t+1}) + \alpha u_{t+1}(x_{t+1})] \\ &= \sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, d_t(h_t)), x_{t+1}) [\varphi_t(x_t, d_t(h_t), x_{t+1}) + \alpha u_{t+1}(x_{t+1})]. \end{aligned}$$

Puisque  $d_t(h_t) \in \mathcal{A}$ , l'infimum de (3.11) est atteint pour  $a = d_t(h_t)$ .

La quantité minimisée (3.11) ne dépend que de  $x_t$  et  $a$ . Donc il existe  $\delta_t^* : \mathcal{E} \rightarrow \mathcal{A}$  t.q. l'infimum est atteint pour  $a = \delta_t^*(x_t)$ . On en déduit que

$$u_t(x_t) = \sum_{x_{t+1} \in \mathcal{E}} p_t((x_t, \delta_t^*(x_t)), x_{t+1}) [\varphi_t(x_t, \delta_t^*(x_t), x_{t+1}) + \alpha u_{t+1}(x_{t+1})].$$

Il suffit de reprendre l'argument de récurrence descendante de la fin de la démonstration du théorème 3.2.3 avec  $\gamma = 0$  pour voir que la stratégie markovienne  $\pi_M^* = (\delta_0^*, \dots, \delta_{N-1}^*)$  est optimale.  $\square$

Les équations d'optimalité (3.10) sont aussi appelées équations de Bellman ou équations de la programmation dynamique. Elles permettent de ramener la résolution du problème d'optimisation à  $N$  périodes de temps à celle plus simple de  $N$  problèmes d'optimisation à une période. Leur résolution par récurrence descendante fournit une procédure effective qui porte le nom de programmation dynamique pour calculer le coût minimal et déterminer les stratégies markoviennes optimales (ou  $\varepsilon$ -optimales).

La théorie développée dans ce paragraphe peut être étendue à un cadre beaucoup plus général. Par exemple, on peut considérer des espaces d'états  $\mathcal{E}$  et d'actions  $\mathcal{A}$  non discrets mais pour cela il faut utiliser la notion d'espérance conditionnelle sachant une tribu qui dépasse le cadre de ce livre. Néanmoins, les conclusions sont analogues à celles que nous avons obtenues.

Il est également intéressant de considérer le problème à horizon infini dans le cas où les fonctions  $p_t$  et  $\varphi_t$  ne dépendent pas du temps et sont notées respectivement  $p$  et  $\varphi$ . On suppose que le facteur d'actualisation  $\alpha$  est strictement plus petit que 1 pour pouvoir définir le coût total. Le passage à la limite formel  $t \rightarrow +\infty$  dans (3.10) explique pourquoi on travaille alors avec l'équation d'optimalité :

$$u(x) = \inf_{a \in \mathcal{A}} \sum_{y \in \mathcal{E}} p((x, a), y) [\varphi(x, a, y) + \alpha u(y)].$$

Nous renvoyons par exemple aux livres de Puterman [8], de Bertsekas [2, 3], de Bertsekas et Shreve [4], de White [10] et de Whittle [11] pour plus de détails concernant les généralisations possibles.

### 3.3 Résolution du problème dynamique de gestion de stock

D'après l'exemple 3.2.1, lorsque le produit se présente sous forme d'unités, le problème dynamique de gestion de stock décrit au paragraphe 3.1.2 entre dans le cadre du paragraphe précédent consacré au contrôle de chaînes de Markov. L'espace d'état est  $\mathcal{E} = \mathbb{Z}$  pour le stock système  $X_t$  et l'ensemble d'actions  $\mathcal{A} = \mathbb{N}$  pour la quantité  $Q_t$  de produit commandée. En outre, pour tout  $(t, x, q, y)$  dans  $\{0, \dots, N-1\} \times \mathbb{Z} \times \mathbb{N} \times \mathbb{Z}$ ,

$$\begin{cases} p_t((x, q), y) = \mathbb{P}(D_1 = x + q - y), \\ \varphi_t(x, q, y) = c_F \mathbf{1}_{\{q > 0\}} + cq + c_S y^+ + c_M y^-. \end{cases}$$

Les équations d'optimalité (3.10) s'écrivent : pour  $t \in \{0, \dots, N-1\}$ ,  $x \in \mathbb{Z}$

$$\begin{aligned} u_t(x) &= \inf_{q \in \mathbb{N}} \sum_{y \in \mathbb{Z}} p_t((x, q), y) [\varphi_t(x, q, y) + \alpha u_{t+1}(y)] \\ &= \inf_{q \in \mathbb{N}} \sum_{y \in \mathbb{Z}} \mathbb{P}(D_1 = x + q - y) [c_F \mathbf{1}_{\{q > 0\}} + cq + c_S y^+ + c_M y^- + \alpha u_{t+1}(y)] \\ &= \inf_{q \in \mathbb{N}} \sum_{z \in \mathbb{Z}} \mathbb{P}(D_1 = z) [c_F \mathbf{1}_{\{q > 0\}} + cq + c_S(x + q - z)^+ + c_M(x + q - z)^- \\ &\quad + \alpha u_{t+1}(x + q - z)] \text{ en posant } z = x + q - y, \\ &= \inf_{q \in \mathbb{N}} \mathbb{E}[c_F \mathbf{1}_{\{q > 0\}} + cq + c_S(x + q - D_1)^+ + c_M(x + q - D_1)^- \\ &\quad + \alpha u_{t+1}(x + q - D_1)]. \end{aligned}$$

Nous allons étudier ces équations dans le cadre continu où le stock système est réel, les demandes sont des variables aléatoires positives non nécessairement entières et les quantités commandées sont des réels positifs, car leur résolution est plus simple que dans le cadre discret où nous les avons établies. Elles s'écrivent alors : pour  $t \in \{0, \dots, N-1\}$ ,

$$\begin{aligned} u_t(x) &= \inf_{q \geq 0} \mathbb{E}[c_F \mathbf{1}_{\{q > 0\}} + cq + c_S(x + q - D_1)^+ \\ &\quad + c_M(x + q - D_1)^- + \alpha u_{t+1}(x + q - D_1)]. \end{aligned} \quad (3.12)$$

#### 3.3.1 Gestion sans coût fixe d'approvisionnement

En plus de la nullité du coût fixe ( $c_F = 0$ ), nous supposons dans ce paragraphe que la fonction de coût terminale est  $u_N(x) = -cx$ , ce qui revient à associer une valeur unitaire  $c$  au stock système terminal  $X_N$ . Cette hypothèse simplificatrice qui va nous permettre d'obtenir une stratégie optimale stationnaire (i.e. telle que la règle de décision à l'instant  $t$  ne dépend pas de  $t$ ) est



discutable : si  $X_N \leq 0$ , il y a  $-X_N$  manquants et il est naturel de leur associer le coût  $-cX_N$  puisque c'est le prix à payer au fournisseur pour obtenir la quantité  $-X_N$  ; en revanche affecter le coût  $-cX_N$  et donc la valeur  $cX_N$  au stock physique résiduel  $X_N \geq 0$  est moins naturel.

Commençons par déterminer la quantité de produit commandée optimale à l'instant  $N - 1$ . Comme  $u_N(x) = -cx$ ,

$$\begin{aligned} u_{N-1}(x) &= \alpha c \mathbb{E}[D_1] - cx + \inf_{q \geq 0} \left( (1 - \alpha)c(x + q) + c_S \mathbb{E}[(x + q - D_1)^+] \right. \\ &\quad \left. + c_M \mathbb{E}[(D_1 - x - q)^+] \right) \\ &= \alpha c \mathbb{E}[D_1] - cx + \inf_{q \geq 0} g_\alpha(x + q) \end{aligned} \quad (3.13)$$

où la fonction  $g_\alpha(y) = (1 - \alpha)cy + c_S \mathbb{E}[(y - D_1)^+] + c_M \mathbb{E}[(D_1 - y)^+]$  est définie comme la fonction de coût moyen  $g$  du modèle étudié au paragraphe 3.1.1 (voir équation (3.1)) à ceci près que le coût unitaire  $c$  est remplacé par  $(1 - \alpha)c$ . On se place dans le cas intéressant où  $c_M > (1 - \alpha)c > -c_S$ . D'après l'analyse menée au paragraphe 3.1.1, la fonction  $g_\alpha$  est continue, décroissante sur  $]-\infty, S_\alpha]$  et croissante sur  $[S_\alpha, +\infty[$  avec

$$S_\alpha = \inf\{z \geq 0 : F(z) \geq (c_M - (1 - \alpha)c)/(c_M + c_S)\} \in \mathbb{R}_+,$$

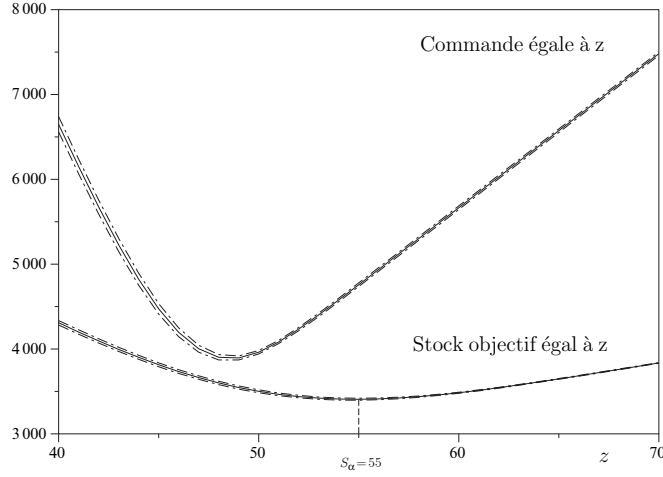
où  $F$  est la fonction de répartition commune des demandes. La décision optimale à l'instant  $N - 1$  consiste donc à commander la quantité  $(S_\alpha - x)^+$  (i.e.  $(S_\alpha - x)$  si  $x \leq S_\alpha$  et rien du tout sinon) si le stock système est  $x$ . Nous allons vérifier que cela reste vrai à tout instant  $t$  dans  $\{0, \dots, N - 2\}$ .

**Théorème 3.3.1.** *On suppose  $c_F = 0$ ,  $c_M > (1 - \alpha)c > -c_S$  et  $u_N(x) = -cx$ . La stratégie avec stock objectif*

$$S_\alpha = \inf\{z \geq 0 : F(z) \geq (c_M - (1 - \alpha)c)/(c_M + c_S)\},$$

*qui consiste pour tout  $t \in \{0, \dots, N - 1\}$  à commander la quantité  $(S_\alpha - x)^+$  lorsque le stock système vaut  $x$  est optimale.*

La figure 3.2 illustre l'optimalité de la stratégie de stock objectif  $S_\alpha = 55$  dans le cas particulier où les demandes sont distribuées suivant la loi de Poisson de paramètre 50,  $N = 10$ ,  $\alpha = 0.9$ ,  $c = 10$ ,  $c_M = 20$ ,  $c_S = 5$ ,  $c_F = 0$ ,  $u_N(x) = -cx$  et le stock initial est  $X_0 = 20$ . Pour  $z \in \{40, 41, \dots, 70\}$  les coûts moyens (3.8) associés à la stratégie qui consiste à commander  $z$  à chaque instant d'une part et à la stratégie de stock objectif  $z$  d'autre part ont été évalués en effectuant la moyenne empirique des coûts sur 1 000 réalisations indépendantes  $(D_1^i, \dots, D_{10}^i)_{1 \leq i \leq 1000}$  des demandes. Les mêmes réalisations de ces variables ont été utilisées pour chacune des stratégies. Plus précisément, le coût moyen est approché par



**Fig. 3.2.** Comparaison des coûts entre la stratégie de stock objectif  $z$  et la stratégie de commande constante égale à  $z$  ( $N = 10$ , stock initial  $X_0 = 20$ ,  $\alpha = 0.9$ ,  $c = 10$ ,  $c_M = 20$ ,  $c_S = 5$ ,  $c_F = 0$ ,  $u_N(x) = -cx$ , demandes distribuées suivant la loi de Poisson de paramètre 50)

$$\frac{1}{1\,000} \sum_{i=1}^{1\,000} \left[ \sum_{t=0}^9 (0.9)^t (10Q_t^i + 5(X_{t+1}^i)^+ + 20(X_{t+1}^i)^-) - 10(0.9)^{10} X_N^i \right]$$

où pour  $i \in \{1, \dots, 1\,000\}$ ,  $X_0^i = 20$  et pour  $t \in \{0, \dots, 9\}$ ,  $Q_t^i = z$  et  $X_{t+1}^i = X_t^i + z - D_{t+1}^i$  dans la stratégie qui consiste à commander  $z$  à chaque instant et  $Q_t^i = (z - X_t^i)^+$  et  $X_{t+1}^i = \max(X_t^i, z) - D_{t+1}^i$  dans la stratégie de stock objectif  $z$ . Les courbes en pointillés représentent les bornes des intervalles de confiance à 95 % obtenus pour chacun des coûts moyens. On observe bien que parmi les stratégies considérées, le coût moyen minimal est obtenu pour la stratégie de stock objectif  $S_\alpha = 55$ .

*Démonstration.* Comme la quantité  $q$  optimale dans (3.13) est  $(S_\alpha - x)^+$ , on a

$$u_{N-1}(x) = \begin{cases} \alpha c \mathbb{E}[D_1] - cx + g_\alpha(S_\alpha) & \text{si } x \leq S_\alpha \\ \alpha c \mathbb{E}[D_1] - cx + g_\alpha(x) & \text{sinon.} \end{cases}$$

Ainsi  $u_{N-1}(x) = K_{N-1} - cx + w_{N-1}(x)$  où  $K_{N-1} = \alpha c \mathbb{E}[D_1] + g_\alpha(S_\alpha)$  est une constante et  $w_{N-1}(x) = \mathbf{1}_{\{x \geq S_\alpha\}}(g_\alpha(x) - g_\alpha(S_\alpha))$  est une fonction croissante et nulle sur  $] -\infty, S_\alpha]$ .

Nous allons vérifier que cette écriture se généralise aux instants antérieurs en démontrant par récurrence descendante que pour tout  $t$  dans  $\{0, \dots, N-1\}$ , la décision optimale à l'instant  $t$  consiste à commander  $(S_\alpha - x)^+$  si le stock système est  $x$  et que  $u_t(x) = K_t - cx + w_t(x)$  où  $K_t$  est une constante et  $w_t$  est une fonction croissante et nulle sur  $] -\infty, S_\alpha]$ .

Supposons que l'hypothèse de récurrence est vérifiée au rang  $t + 1$ . En insérant l'égalité  $u_{t+1}(x) = K_{t+1} - cx + w_{t+1}(x)$  dans l'équation (3.12) avec  $c_F = 0$ , on obtient

$$\begin{aligned} u_t(x) &= \inf_{q \geq 0} \mathbb{E} \left[ cq + c_S(x + q - D_1)^+ + c_M(x + q - D_1)^- \right. \\ &\quad \left. + \alpha(K_{t+1} - c(x + q - D_1) + w_{t+1}(x + q - D_1)) \right] \\ &= \alpha(K_{t+1} + c\mathbb{E}[D_1]) - cx + \inf_{q \geq 0} (g_\alpha(x + q) + \alpha\mathbb{E}[w_{t+1}(x + q - D_1)]). \end{aligned}$$

Pour analyser la minimisation du dernier terme, on remarque que

- la fonction  $g_\alpha(y)$  est croissante sur  $[S_\alpha, +\infty[$  et atteint son minimum pour  $y = S_\alpha$  ;
- la fonction  $y \rightarrow \mathbb{E}[w_{t+1}(y - D_1)]$  est croissante par croissance de  $w_{t+1}$  ;
- cette fonction est nulle pour  $y \leq S_\alpha$  puisque comme  $D_1 \geq 0$ ,  $y - D_1$  est alors dans l'ensemble  $] - \infty, S_\alpha]$  où  $w_{t+1}$  s'annule.

On en déduit que l'infimum est atteint pour  $q = (S_\alpha - x)^+$  et que  $u_t(x) = K_t - cx + w_t(x)$  avec

$$\begin{cases} K_t = \alpha(K_{t+1} + c\mathbb{E}[D_1]) + g_\alpha(S_\alpha) \\ w_t(x) = \mathbf{1}_{\{x \geq S_\alpha\}}(g_\alpha(x) - g_\alpha(S_\alpha) + \alpha\mathbb{E}[w_{t+1}(x - D_1)]). \end{cases}$$

La fonction  $w_t$  est clairement croissante et nulle sur  $] - \infty, S_\alpha]$ . □

**Remarque 3.3.2.** Notons que si le produit se présente sous forme d'unités, les demandes des clients sont entières et on peut vérifier comme dans la démonstration de la proposition 3.1.1 que le stock objectif  $S_\alpha$  est entier. À l'instant  $t$ , le stock système  $X_t$  est entier et il est possible pour le gestionnaire de commander la quantité  $(S_\alpha - X_t)^+$  qui est un entier positif. Cette stratégie est optimale car elle l'est pour le modèle où le produit se présente sous forme continue qui offre plus d'opportunités en termes de quantité commandée. ◇

### 3.4 Conclusion

Dans ce chapitre nous avons mis en évidence l'intérêt des règles de décision de type  $(s, S)$  pour la gestion dynamique du stock d'un produit (pièce de rechange par exemple). En effet, nous avons montré l'optimalité d'une stratégie composée de règles de décision markoviennes de ce type à chaque période parmi toutes les stratégies composées de règles de décision fonctions de tout le passé. Ce résultat reste valable pour des modèles plus généraux que celui que nous avons étudié : fonctions de coût de surplus et de coût de manquants convexes (elles sont supposées linéaires ici), demandes indépendantes mais non identiquement distribuées, fonctions de coût dépendant du temps, coûts fixes dépendant du temps sous réserve que  $\alpha c_F(t+1) \leq c_F(t)$ , etc. Nous renvoyons

à la présentation de Porteus [7] ou au livre de Liu et Esogbue [6] pour la description de ces modèles généraux. L'optimalité d'une stratégie  $(s, S)$  reste également valable pour le problème de gestion de stock à horizon temporel infini, auquel est consacré le Chap. 13 du livre de Whittle [11].

D'un point de vue pratique, on doit identifier la loi de la demande. À cet effet, il est par exemple possible d'effectuer un traitement statistique des demandes passées. On peut ensuite déterminer les  $(s_t, S_t)$  en résolvant par récurrence descendante les équations d'optimalité.

## Références

1. R. Bellman. *Dynamic programming*. Princeton University Press, Princeton, N.J., 1957.
2. D.P. Bertsekas. *Dynamic programming and stochastic control*. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1976. Mathematics in Science and Engineering, 125.
3. D.P. Bertsekas. *Dynamic Programming and Optimal Control, Vol. 1 et 2*. Athena Scientific, 1995.
4. D.P. Bertsekas et S.E. Shreve. *Stochastic optimal control*, volume 139 de *Mathematics in Science and Engineering*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1978. The discrete time case.
5. D. Lamberton et B. Lapeyre. *Introduction au calcul stochastique appliqué à la finance*. Ellipses Édition Marketing, Paris, seconde édition, 1997.
6. B. Liu et A.O. Esogbue. *Decision criteria and optimal inventory processes*. International Series in Operations Research & Management Science, 20. Kluwer Academic Publishers, Boston, MA, 1999.
7. E.L. Porteus. Stochastic inventory theory. In *Stochastic models*, volume 2 de *Handbooks Oper. Res. Management Sci.*, pages 605–652. North-Holland, Amsterdam, 1990.
8. M.L. Puterman. *Markov decision processes : discrete stochastic dynamic programming*. Wiley Series in Probability and Mathematical Statistics : Applied Probability and Statistics. John Wiley & Sons Inc., New York, 1994.
9. H. Scarf. The optimality of  $(s, S)$  policies in the dynamic inventory problem. In S.U. Press, editor, *Proceeding of the 1959 Stanford Symposium on Mathematical Methods in the Social Sciences*, pages 196–202, 1960.
10. D.J. White. *Markov decision processes*. John Wiley & Sons Ltd., Chichester, 1993.
11. P. Whittle. *Optimization over time. Vol. I*. Wiley Series in Probability and Mathematical Statistics : Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1982. Dynamic programming and stochastic control.

---

## Recherche de zones homogènes dans l'ADN

Le bactériophage lambda est un parasite de la bactérie *Escherichia coli*. Son ADN (acide désoxyribonucléique) circulaire comporte  $N_0 = 48\,502$  paires de nucléotides (voir [15]), et il est essentiellement constitué de régions codantes, i.e. de régions lues et traduites en protéines. La transcription, c'est-à-dire la lecture de l'ADN, s'effectue sur des parties de chacun des deux brins qui forment la double hélice de l'ADN. Ainsi sur la séquence d'ADN d'un seul brin on peut distinguer deux types de zones : celles où la transcription a lieu sur le brin et celles où la transcription a lieu sur le brin apparié. On observe sur les parties codantes une certaine fréquence d'apparition des différents nucléotides Adénine (A), Cytosine (C), Guanine (G) et Thymine (T). Le nucléotide A (resp. C) d'un brin est apparié avec le nucléotide T (resp. G) du brin apparié et vice versa. Les deux types de zones d'un brin décrites plus haut correspondent en fait à des fréquences d'apparitions différentes des quatre nucléotides. Les biologistes ont d'abord analysé l'ADN du bactériophage lambda en identifiant les gènes de l'ADN, c'est-à-dire les parties codantes de l'ADN, et les protéines correspondantes. Et ils ont ainsi constaté que les deux brins de l'ADN comportaient des parties codantes. Il est naturel de vouloir détecter a priori les parties codantes, ou susceptibles d'être codantes, à partir d'une analyse statistique de l'ADN. Cela peut permettre aux biologistes d'identifier plus rapidement les parties codantes pour les organismes dont la séquence d'ADN est connue.

Les paragraphes qui suivent montrent comment, en modélisant la séquence d'ADN comme une réalisation partielle d'une chaîne de Markov, on peut détecter les zones où les fréquences d'apparitions des quatre nucléotides sont significativement différentes. L'algorithme EM (Espérance Maximisation) que nous présentons et son utilisation pour l'analyse de l'ADN ont été étudiés en détail et dans un cadre plus général par Muri [12]. De nombreux travaux récents permettent d'améliorer l'algorithme EM pour la détection de zones intéressantes de l'ADN en tenant compte d'informations biologiques connues a priori, voir par exemple les travaux du Laboratoire Statistique et Génome (<http://stat.genopole.cnrs.fr>).

Les algorithmes EM ont été initialement introduits en 1977 par Dempster, Laid et Rubin [8]. Ils sont utilisés pour l'estimation de paramètres dans des modèles où des variables sont cachées, c'est-à-dire non observées (voir par exemple [5] p. 213). Dans l'exemple ci-dessus, avec l'interprétation biologique que l'on espère retrouver, on ne sait pas si le  $k$ -ième nucléotide observé appartient à une zone transcrite ou à une zone appariée à une zone transcrite. Le brin transcrit au niveau du  $k$ -ième nucléotide est donc une variable cachée que l'on désire retrouver. Il existe de nombreuses applications des algorithmes EM, voir par exemple [10]. Signalons, sans être exhaustif, que ces algorithmes sont utilisés dans les domaines suivants :

- Classification ou étude de données mélangées dont les sources sont inconnues : pour les données mélangées voir par exemple [11], pour l'analyse d'image voir par exemple [7 et 9], voir aussi l'exemple du paragraphe 5.5.1.
- Analyse de données censurées ou tronquées, voir le problème du paragraphe ??.
- Estimation de matrice de covariance avec des données incomplètes, etc.

Nous présentons brièvement le modèle mathématique pour la séquence d'un brin d'ADN,  $y_1 \dots y_{N_0}$  du bactériophage lambda. À la séquence d'ADN, on peut associer la séquence non observée, dite séquence cachée,  $s_1 \dots s_{N_0}$ , où si  $s_k = +1$ , alors  $y_k$  est la réalisation d'une variable aléatoire,  $Y_k$ , de loi  $p_+$  sur  $\mathcal{X} = \{\text{A, C, G, T}\}$ , et si  $s_k = -1$  alors la loi de  $Y_k$  est  $p_-$ . Les probabilités  $p_+$  et  $p_-$  sont distinctes mais inconnues. On modélise la suite  $s_1, \dots, s_{N_0}$  comme la réalisation d'une chaîne de Markov,  $(S_n, n \geq 1)$ , sur  $\mathcal{I} = \{+1, -1\}$  de matrice de transition,  $a$ , également inconnue.

**Remarque.** La matrice de transition  $a$  est de la forme

$$a = \begin{pmatrix} 1 - \varepsilon & \varepsilon \\ \varepsilon' & 1 - \varepsilon' \end{pmatrix},$$

où  $\varepsilon$  et  $\varepsilon'$  sont intuitivement inversement proportionnels à la longueur moyenne des zones homogènes où les fréquences d'apparitions des quatre nucléotides sont constantes. En effet, si par exemple  $S_1 = +1$ , la loi du premier instant où la chaîne de Markov change d'état,  $T = \inf\{k \geq 1; S_{k+1} \neq +1\}$ , suit une loi géométrique de paramètre  $\varepsilon$  car

$$\mathbb{P}(T = k | S_1 = +1) = \mathbb{P}(S_{k+1} = -1, S_k = 1, \dots, S_2 = 1 | S_1 = 1) = (1 - \varepsilon)^{k-1} \varepsilon,$$

et son espérance vaut  $1/\varepsilon$ .  $\diamond$

Le modèle utilisé comporte une chaîne de Markov,  $S$ , qui n'est pas directement observée. Ce type de modèle, dit modèle de chaînes de Markov cachées, est présenté de manière détaillée au paragraphe 5.1. Pour identifier les zones homogènes, il faut estimer les paramètres inconnus  $a$ ,  $p_+$  et  $p_-$ . Pour cela, on utilisera les estimateurs du maximum de vraisemblance (EMV) qui possèdent de bonnes propriétés. La construction de ces estimateurs ainsi que leurs propriétés asymptotiques sont présentées dans le paragraphe 5.2 au

travers d'un exemple élémentaire et dans un cadre simple. La convergence de l'EMV vers les paramètres inconnus du modèle dans le cadre des chaînes de Markov cachées est plus complexe à établir. Ce résultat et sa démonstration technique sont reportés au paragraphe ?? . Nous verrons au paragraphe 5.3, qu'il est impossible de calculer explicitement l'EMV dans le cas particulier des chaînes de Markov cachées. Mais nous exhiberons une méthode, l'algorithme EM, pour en donner une bonne approximation. Le paragraphe 5.4, qui est le cœur de ce chapitre, présente la mise en œuvre explicite de l'algorithme EM. En particulier, on calcule la loi des états cachés  $S_1, \dots, S_{N_0}$  sachant les observations  $y_1, \dots, y_{N_0}$ , voir la Fig. 5.7 pour les valeurs de  $\mathbb{P}(S_n = +1 | y_1, \dots, y_{N_0})$  concernant l'ADN du bactériophage lambda. Dans les modèles de mélanges ou de données censurées, l'algorithme EM s'exprime simplement. Ces applications importantes, en marge des modèles de chaînes de Markov, sont abordées au paragraphe 5.5. Pour les modèles de mélanges, on utilise les données historiques des crabes de Weldon, analysées par Pearson en 1894, première approche statistique d'un modèle de mélange. Les modèles de données censurées sont évoqués au travers d'un problème. Enfin, dans la conclusion, paragraphe 5.6, nous présentons les résultats numériques obtenus pour le bactériophage lambda ainsi que quelques commentaires sur la méthode utilisée.

## 5.1 Chaînes de Markov cachées

On rappelle la notation condensée suivante  $x_m^n$  pour le vecteur  $(x_m, \dots, x_n)$  avec  $m \leq n \in \mathbb{Z}$ . On considère  $S = (S_n, n \geq 1)$  une chaîne de Markov à valeurs dans  $\mathcal{I}$ , un espace fini non réduit à un élément, de matrice de transition  $a$  et de loi initiale  $\pi_0$ . Soit  $(Y_n, n \geq 1)$  une suite de variables à valeurs dans  $\mathcal{X}$ , un espace d'état fini, telle que conditionnellement à  $S$  les variables aléatoires  $(Y_n, n \geq 1)$  sont indépendantes et la loi de  $Y_k$  sachant  $S$  ne dépend que de la valeur de  $S_k$ . Plus précisément, pour tout  $N \geq 1$ , conditionnellement à  $S_1^N$ , les variables aléatoires  $Y_1^N$  sont indépendantes : pour tous  $N \geq 1$ ,  $y_1^N \in \mathcal{X}^N$  et  $s_1^N \in \mathcal{I}^N$ , on a

$$\mathbb{P}(Y_1^N = y_1^N | S_1^N = s_1^N) = \prod_{n=1}^N \mathbb{P}(Y_n = y_n | S_1^N = s_1^N), \quad (5.1)$$

de plus il existe une matrice  $b = (b(i, x); i \in \mathcal{I}, x \in \mathcal{X})$ , telle que

$$\mathbb{P}(Y_n = y_n | S_1^N = s_1^N) = \mathbb{P}(Y_n = y_n | S_n = s_n) = b(s_n, y_n). \quad (5.2)$$

**Lemme 5.1.1.** *La suite  $((S_n, Y_n), n \geq 1)$  est une chaîne de Markov. On a pour tous  $n \geq 2$ ,  $s_1^n \in \mathcal{I}^n$  et  $y_1^n \in \mathcal{X}^n$ ,*

$$\mathbb{P}(S_n = s_n, Y_n = y_n | S_1^{n-1} = s_1^{n-1}, Y_1^{n-1} = y_1^{n-1}) = a(s_{n-1}, s_n) b(s_n, y_n),$$

et

$$\mathbb{P}(S_n = s_n | S_1^{n-1} = s_1^{n-1}, Y_1^{n-1} = y_1^{n-1}) = a(s_{n-1}, s_n).$$

*Démonstration.* En utilisant les égalités (5.1) et (5.2) ainsi que la propriété de Markov pour  $(S_n, n \geq 1)$ , on a

$$\begin{aligned}
\mathbb{P}(S_1^n = s_1^n, Y_1^n = y_1^n) &= \mathbb{P}(Y_1^n = y_1^n | S_1^n = s_1^n) \mathbb{P}(S_1^n = s_1^n) \\
&= \left( \prod_{k=1}^n b(s_k, y_k) \right) \mathbb{P}(S_n = s_n | S_1^{n-1} = s_1^{n-1}) \mathbb{P}(S_1^{n-1} = s_1^{n-1}) \\
&= \left( \prod_{k=1}^n b(s_k, y_k) \right) \mathbb{P}(S_n = s_n | S_{n-1} = s_{n-1}) \mathbb{P}(S_1^{n-1} = s_1^{n-1}) \\
&= \left( \prod_{k=1}^n b(s_k, y_k) \right) a(s_{n-1}, s_n) \mathbb{P}(S_1^{n-1} = s_1^{n-1}).
\end{aligned}$$

D'autre part, en sommant sur  $y_n \in \mathcal{X}$  et en utilisant  $\sum_{x \in \mathcal{X}} b(s_n, x) = 1$ , puis en sommant sur  $s_n \in \mathcal{I}$  et en utilisant  $\sum_{s_n \in \mathcal{I}} a(s_{n-1}, s_n) = 1$ , il vient

$$\mathbb{P}(S_1^{n-1} = s_1^{n-1}, Y_1^{n-1} = y_1^{n-1}) = \left( \prod_{k=1}^{n-1} b(s_k, y_k) \right) \mathbb{P}(S_1^{n-1} = s_1^{n-1}).$$

On en déduit donc la première égalité du lemme. Ainsi la suite  $((S_n, Y_n), n \geq 1)$  est une chaîne de Markov.

La deuxième égalité du lemme se déduit de la première en sommant sur  $y_n \in \mathcal{X}$ .  $\square$

Dans le modèle de chaîne de Markov cachée, lors d'une réalisation, on observe simplement  $y_1^N$ , une réalisation de  $Y_1^N$ . Les variables  $S_1^N$  sont appelées variables cachées, et leur valeur prise lors d'une réalisation, les valeurs cachées. Dans ce modèle, on cherche à estimer, à partir de l'observation  $y_1^N$ , le paramètre  $\theta = (a, b, \pi_0)$  puis à calculer, pour  $i \in \mathcal{I}$ , les probabilités  $\mathbb{P}(S_n = i | Y_1^N = y_1^N)$ . L'ensemble des paramètres possibles forme un compact  $\Theta$  de  $[0, 1]^{\mathcal{I}^2} \times [0, 1]^{\mathcal{I} \times \mathcal{X}} \times [0, 1]^{\mathcal{I}}$ .

Remarquons que la loi du vecteur des observations  $Y_1^N$  ne détermine pas complètement le paramètre  $\theta = (a, b, \pi_0)$ . En effet, soit  $\sigma$  une permutation de  $\mathcal{I}$ . On note  $\theta_\sigma = (a(\sigma_i, \sigma_j), i \in \mathcal{I}, j \in \mathcal{I}), (b(\sigma_i, x), i \in \mathcal{I}, x \in \mathcal{X}), (\pi_0(\sigma_i), i \in \mathcal{I}))$ . Les paramètres  $\theta$  et  $\theta_\sigma$  génèrent la même loi pour le processus des observations  $Y_1^N$  (mais pas pour  $(S_1^N, Y_1^N)$  en général). On ne peut pas espérer distinguer  $\theta$  de  $\theta_\sigma$  à la seule vue des observations. On dit que le modèle n'est pas identifiable.

Remarquons également que s'il existe une probabilité  $p$  sur  $\mathcal{X}$  telle que pour tous  $i \in \mathcal{I}$ ,  $x \in \mathcal{X}$ ,  $b(i, x) = p(x)$  alors la suite  $Y$  est une suite de variables aléatoires indépendantes et de même loi  $p$ . En particulier la loi de  $Y_1^N$  ne dépend pas des valeurs de  $a$  et  $\pi_0$ .



**Définition 5.1.2.** Soit  $X$  une variable aléatoire dont on peut observer les réalisations. On suppose que la loi, a priori inconnue, de  $X$  appartient à une famille de lois indicée par un paramètre :  $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ , où  $\Theta$  est un ensemble de paramètres. Il s'agit d'un **modèle paramétrique**. On dit que le modèle est **identifiable** si pour  $\theta \neq \theta' \in \Theta$ , on a  $P_\theta \neq P_{\theta'}$ .

Nous reviendrons sur la notion de modèle identifiable dans le paragraphe suivant lors de la construction d'un estimateur de  $\theta$ , paramètre a priori inconnu de la loi de  $X$ .

Pour que le modèle de chaîne de Markov cachée soit identifiable, il suffit de restreindre l'ensemble des paramètres possibles à un sous-ensemble  $\Theta'$  de  $\Theta$ . Pour cela, on vérifie qu'il est possible de choisir  $\Theta'$  un ouvert de  $\Theta$ , tel que si  $\theta = (a, b, \pi_0) \in \Theta$ , alors

- soit il existe  $i \neq i'$  et pour tout  $x \in \mathcal{X}$ ,  $b(i, x) = b(i', x)$ , alors on a  $\theta \notin \Theta'$ ,
- soit pour tous  $i \neq i'$ , il existe  $x \in \mathcal{X}$  tel que  $b(i, x) \neq b(i', x)$ , et alors il existe  $\sigma$  une permutation unique de  $\mathcal{I}$  telle que  $\theta_\sigma \in \Theta'$ .

Ainsi si  $\theta$  et  $\theta'$  appartiennent à  $\Theta'$  et sont distincts, alors les lois des observations  $Y_1^N$  sont différentes. Le modèle, où l'ensemble des paramètres est  $\Theta'$ , est alors identifiable.

## 5.2 L'estimateur du maximum de vraisemblance (EMV)

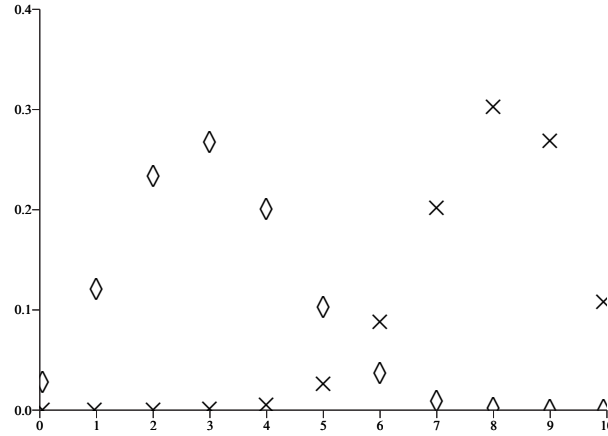
### 5.2.1 Définitions et exemples

Nous illustrons l'estimation par maximum de vraisemblance dans l'exemple qui suit.

**Exemple 5.2.1.** Vous désirez jouer à un jeu de pile ou face avec un adversaire. Vous savez qu'il dispose en fait de deux pièces biaisées. On note  $\theta_i$  la probabilité d'obtenir pile pour la pièce  $i \in \{1, 2\}$ , avec  $\theta_1 = 0.3$  et  $\theta_2 = 0.8$ . La pièce avec laquelle votre adversaire vous propose de jouer, a déjà été utilisée dans le jeu de pile ou face précédent, où vous avez observé  $k_0 = 4$  piles sur  $n = 10$  lancers.

Le nombre de pile obtenu lors de  $n$  lancers suit une loi binomiale de paramètre  $(n, \theta)$ ,  $\theta$  étant la probabilité d'obtenir pile. On note  $p(\theta, k) = \binom{n}{k} \theta^k (1-\theta)^{n-k}$  la probabilité qu'une variable de loi binomiale de paramètre  $(n, \theta)$  prenne la valeur  $k$ . On visualise ces probabilités pour  $n = 10$ , et  $\theta \in \{\theta_1, \theta_2\}$  sur la Fig. 5.1. Il est raisonnable de supposer que la pièce utilisée est la pièce 1 car la probabilité d'observer  $k_0 = 4$  est plus grande pour  $\theta = \theta_1$ ,  $p(\theta_1; k_0) \simeq 0.2$ , que pour  $\theta = \theta_2$ ,  $p(\theta_2; k_0) \simeq 0.006$ . On choisit ainsi le paramètre  $\theta \in \{\theta_1, \theta_2\}$  qui maximise la fonction

$$\theta \rightarrow p(\theta; k_0) \quad \text{où } \theta \in \Theta = \{\theta_1, \theta_2\}.$$



**Fig. 5.1.** Lois binomiales de paramètres  $(10, \theta)$  avec  $\theta = 0.3$  (losanges) et  $\theta = 0.8$  (croix)

La fonction ci-dessus s'appelle la vraisemblance, et le paramètre qui la maximise s'appelle l'estimateur du maximum de vraisemblance.  $\diamond$

**Définition 5.2.2.** On considère un modèle paramétrique :  $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$  une famille de lois (resp. de lois à densité) sur un espace  $E$  discret (resp. sur  $E = \mathbb{R}^n$ ), où  $\Theta$  est un ensemble de paramètres. On note  $p(\theta; x)$  la probabilité qu'une variable de loi  $P_\theta$  prenne la valeur  $x \in E$  (resp. la densité en  $x \in E$  d'une variable de loi  $P_\theta$ ). La fonction définie sur  $\Theta$ , à  $x \in E$  fixé, par  $\theta \rightarrow p(\theta; x)$  s'appelle la **vraisemblance**.

Supposons que pour tout  $x \in E$ , il existe une unique valeur de  $\theta$ , notée  $\hat{\theta}(x)$ , telle que la vraisemblance soit maximale en  $\hat{\theta}(x)$  :  $p(\hat{\theta}(x); x) > p(\theta, x)$  pour tous  $x \in E$ ,  $\theta \in \Theta$  et  $\theta \neq \hat{\theta}(x)$ . La fonction  $x \rightarrow \hat{\theta}(x)$ , pour  $x \in E$ , s'appelle l'**Estimateur du Maximum de Vraisemblance (EMV)** de  $\theta$ .

Comme la fonction log est strictement croissante, on peut aussi rechercher l'EMV comme la valeur de  $\theta$  qui maximise la **log-vraisemblance**  $\theta \rightarrow \log p(\theta; x)$ . Cette approche est souvent techniquement plus simple. Dans le cas où la vraisemblance atteint son maximum en plusieurs points, l'EMV est mal défini, voir la remarque 5.2.10 à ce sujet.

Soit  $X$  est une variable aléatoire de loi  $P_{\theta_0}$ , avec  $\theta_0 \in \Theta$  inconnu.

**Définition 5.2.3.** La variable aléatoire  $\hat{\theta} = \hat{\theta}(X)$  est également appelée l'EMV de  $\theta$ .

**Remarque 5.2.4.** Si  $g$  est une bijection définie sur  $\Theta$ , alors on vérifie facilement que l'EMV de  $r = g(\theta)$ , qui est l'EMV associé au modèle paramétrique

$\mathcal{Q} = \{Q_r = P_{g^{-1}(r)}, r \in g(\Theta)\}$  est  $g(\hat{\theta})$ , où  $\hat{\theta}$  est l'EMV de  $\theta$ . Par convention si  $g$  est une fonction définie sur  $\Theta$ , alors l'EMV de  $g(\theta)$  est  $g(\hat{\theta})$ .  $\diamond$

L'estimateur du maximum de vraisemblance possède de bonnes propriétés statistiques : convergence, normalité asymptotique. En revanche, cet estimateur est souvent biaisé. Ce handicap est généralement compensé par le fait qu'il permet de construire un intervalle de confiance étroit comparativement à d'autres estimateurs. Après avoir donné une définition précise aux termes précédents, nous illustrerons ces propriétés sur l'estimation de la probabilité d'avoir un garçon à la naissance.

**Définition 5.2.5.** On considère un modèle paramétrique. Soit  $X = (X_n, n \geq 1)$  une suite de variables aléatoires à valeurs dans  $E$ , dont la loi  $P_{\theta_0}$  appartient à une famille de lois  $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ , où  $\Theta$  est un ensemble de paramètres. Le paramètre  $\theta_0$  est inconnu a priori.

Un **estimateur** de  $\theta_0$  construit à partir de  $X_1^n$  est une fonction explicite de  $X_1^n$ . En particulier, elle ne fait pas intervenir  $\theta_0$ .

Un estimateur  $h(X_1^n)$  de  $\theta_0$  est dit **sans biais** s'il est intégrable et si  $\mathbb{E}[h(X_1^n)] = \theta_0$  pour tout  $\theta_0 \in \Theta$ . Sinon, on dit que l'estimateur est **biaisé**.

Soit  $(\delta_n, n \geq 1)$  une suite d'estimateurs de  $\theta_0$ , où  $\delta_n$  est construit à partir de  $X_1^n$ . On dit que la suite  $(\delta_n, n \geq 1)$  est un estimateur **convergent** (on dit aussi **fortement convergent**) de  $\theta_0$ , si pour tout  $\theta_0 \in \Theta$ , on a p.s.

$$\lim_{n \rightarrow \infty} \delta_n = \theta_0.$$

Si la convergence a lieu en probabilité seulement, on parle d'estimateur **faiblement convergent**.

On dit que la suite  $(\delta_n, n \geq 1)$  est un estimateur **asymptotiquement normal** de  $\theta_0$ , si pour tout  $\theta_0 \in \Theta$ , on a

$$\sqrt{n}(\delta_n - \theta_0) \xrightarrow[n \rightarrow \infty]{\text{Loi}} \mathcal{N}(0, \Sigma(\theta_0)),$$

où  $\Sigma(\theta_0)$  est la **variance asymptotique** (ou **matrice de covariance asymptotique** si le paramètre  $\theta$  est multidimensionnel).

**Exemple 5.2.6.** On suppose qu'à la naissance chaque bébé a une probabilité  $\theta_0$  d'être un garçon et une probabilité  $1 - \theta_0$  d'être une fille. On considère une population de  $n$  bébés et on désire donner une estimation de  $\theta_0$  afin de savoir si à la naissance il naît significativement plus de garçons que de filles ou plus de filles que de garçons ou autant de garçons que de filles.

On précise le modèle paramétrique. On note  $X_i = 1$  si le  $i^{\text{ème}}$  bébé est un garçon et  $X_i = 0$  sinon. Il est naturel de supposer que les variables aléatoires  $X = (X_i, i \geq 1)$  sont indépendantes et de même loi de Bernoulli de paramètre  $\theta_0 \in \Theta = [0, 1]$ .

Pour calculer l'EMV  $\hat{\theta}_n$ , à partir de l'échantillon  $X_1^n$ , on remarque que  $\mathbb{P}(X_i = x_i) = \theta_0^{x_i}(1 - \theta_0)^{1-x_i}$ . Par indépendance, on en déduit que la vraisemblance est la fonction de  $\theta$  définie sur  $\Theta$ , pour  $x = x_1^n \in \{0, 1\}^n$ , par

$$p_n(\theta; x) = \prod_{i=1}^n \theta^{x_i} (1 - \theta)^{1-x_i} = \theta^{n_1} (1 - \theta)^{n-n_1},$$

où  $n_1 = \sum_{i=1}^n x_i$  représente le nombre de garçons. La log-vraisemblance est définie par

$$L_n(\theta; x) = \log p_n(\theta; x) = n_1 \log(\theta) + (n - n_1) \log(1 - \theta),$$

avec la convention que  $0 \log 0 = 0$ . Dans un premier temps on cherche  $\theta$  qui annule sa dérivée :

$$\frac{\partial L_n}{\partial \theta}(\theta; x) = \frac{n_1}{\theta} - \frac{n - n_1}{1 - \theta} = 0,$$

soit  $\theta = n_1/n$ . Comme  $\partial L_n / \partial \theta$  est une fonction strictement décroissante, on en déduit dans un deuxième temps, que la log-vraisemblance est maximale

pour  $\theta = \frac{n_1}{n} = \frac{1}{n} \sum_{i=1}^n x_i$ . L'EMV est donc  $\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n X_i$ . On peut remar-

quer que dans ce cas précis l'EMV est un estimateur sans biais de  $\theta_0$ .

On déduit de la loi forte des grands nombres que  $(\hat{\theta}_n, n \geq 1)$  converge presque sûrement vers le vrai paramètre inconnu  $\theta_0$ , i.e. l'EMV est convergent. On déduit du théorème central limit que  $(\sqrt{n}(\hat{\theta}_n - \theta_0), n \geq 1)$  converge en loi vers une variable de loi gaussienne  $\mathcal{N}(0, \Sigma(\theta_0))$ , où  $\Sigma(\theta_0) = \text{Var}(X_1) = \theta_0(1 - \theta_0)$ . L'EMV est donc asymptotiquement normal de variance asymptotique  $\Sigma(\theta_0)$ .

On peut alors, en remplaçant  $\Sigma(\theta_0)$  par l'estimation  $\hat{\theta}_n(1 - \hat{\theta}_n)$ , en déduire un intervalle de confiance de  $\theta_0$  de niveau asymptotique 95 % (cf. le paragraphe ??). Par exemple, aux U.S.A. en 1996 on compte  $n_1 = 1\,990\,480$  naissances de garçons et  $n - n_1 = 1\,901\,014$  naissances de filles. On en déduit que  $\hat{\theta}_n = n_1/n \simeq 0.511495$ . On calcule l'intervalle de confiance de niveau asymptotique 95 % pour  $\theta$  :  $\left[ \hat{\theta}_n \pm 1.96 \sqrt{\hat{\theta}_n(1 - \hat{\theta}_n)/\sqrt{n}} \right] \simeq [0.511, 0.512]$ . Il naît significativement plus de garçons que de filles.  $\diamond$

**Exercice 5.2.7.** On considère le modèle paramétrique gaussien. Soit  $(X_n, n \geq 1)$  une suite de variables aléatoires indépendantes et de même loi gaussienne de moyenne  $\mu \in \mathbb{R}$  et de variance  $\sigma^2 > 0$ . Le paramètre est  $\theta = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times ]0, \infty[$ . La vraisemblance associée au vecteur  $X_1^n$  est sa densité.

1. Expliciter la vraisemblance et la log-vraisemblance du modèle gaussien.
2. Montrer que l'EMV,  $\hat{\theta}_n = (\hat{\mu}_n, \hat{\sigma}_n^2)$ , de  $\theta$  associé à  $X_1^n$  est défini par la moyenne empirique et la variance empirique :

$$\hat{\mu}_n = \frac{1}{n} \sum_{k=1}^n X_k, \quad \hat{\sigma}_n^2 = \frac{1}{n} \sum_{k=1}^n (X_k - \hat{\mu}_n)^2 = \frac{1}{n} \sum_{k=1}^n X_k^2 - \hat{\mu}_n^2. \quad (5.3)$$

3. Vérifier que la moyenne empirique est un estimateur sans biais de  $\mu$ , mais que la variance empirique est un estimateur biaisé de  $\sigma^2$ . Construire à partir de  $\hat{\sigma}_n^2$  un estimateur sans biais de  $\sigma^2$ .
4. Vérifier que l'EMV,  $\hat{\theta}_n$ , est un estimateur de  $\theta$  convergent et asymptotiquement normal.



Sous des hypothèses assez générales sur la vraisemblance, on peut montrer que la suite d'estimateurs  $(\hat{\theta}_n, n \geq 1)$ , où  $\hat{\theta}_n$  est l'EMV de  $\theta$  construit à partir de  $X_1^n$ , est convergente et asymptotiquement normale. On pourra consulter [4], paragraphe 16, pour une démonstration précise de ces résultats quand les variables  $(X_n, n \geq 1)$  sont indépendantes et de même loi.

Dans le paragraphe qui suit, nous démontrons la convergence de l'EMV, dans le cas où les variables  $(X_n, n \geq 1)$ , indépendantes et de même loi, dépendant d'un paramètre  $\theta \in \Theta$ , sont à valeurs dans un espace discret  $E$ . Des arguments similaires permettront de montrer la convergence de l'EMV pour les chaînes de Markov cachées (voir le paragraphe ??). Au chapitre ??, la démonstration du lemme ?? établit directement la convergence de l'EMV de la matrice de transition pour un modèle de chaîne de Markov à l'aide du théorème ergodique pour les chaînes de Markov et la normalité asymptotique à l'aide du TCL ergodique.

### 5.2.2 Convergence de l'EMV dans un modèle simple

On considère  $\mathcal{P} = \{P_\theta; \theta \in \Theta\}$ , une famille de lois sur un espace discret  $E$ , indexée par un paramètre  $\theta \in \Theta$ . Soit  $(X_n, n \geq 1)$  une suite de variables aléatoires indépendantes et de même loi  $P_{\theta_0}$ ,  $\theta_0$  étant inconnu. Pour  $x_1 \in E$  on pose  $p(\theta_0; x_1) = \mathbb{P}(X_1 = x_1)$ . La vraisemblance associée à  $X_1$  est donc  $\theta \mapsto p(\theta, x_1)$ . Par indépendance, la vraisemblance associée à l'échantillon  $X_1^n$  est, pour  $x_1^n \in E$ ,  $p_n(\theta; x_1^n) = \prod_{i=1}^n p(\theta; x_i)$ . La log-vraisemblance est

$$L_n(\theta; x_1^n) = \sum_{i=1}^n \log p(\theta; x_i). \quad (5.4)$$

On suppose que l'EMV de  $\theta$ ,  $\hat{\theta}_n$  est bien défini : la variable aléatoire  $\hat{\theta}_n$  est l'unique valeur de  $\Theta$  en laquelle  $L_n(\theta; X_1^n)$  et donc  $\frac{1}{n} L_n(\theta; X_1^n)$  sont maximaux :

$$\hat{\theta}_n = \operatorname{argmax}_{\theta \in \Theta} \frac{1}{n} \sum_{i=1}^n \log p(\theta; X_i). \quad (5.5)$$

On considère le lemme suivant dont la démonstration est reportée à la fin de ce paragraphe.

**Lemme 5.2.8.** *Soit un espace  $E$  discret,  $\mathcal{P}_E$  l'ensemble des probabilités sur  $E$ , et  $p = (p(x), x \in E) \in \mathcal{P}_E$ . On considère la fonction  $\mathcal{H}_p$  à valeurs dans  $[-\infty, 0]$  définie sur  $\mathcal{P}_E$  par*

$$\mathcal{H}_p : p' \rightarrow \mathcal{H}_p(p') = \sum_{x \in E} p(x) \log p'(x),$$

avec la convention  $0 \log 0 = 0$ . On suppose que  $\mathcal{H}_p(p) > -\infty$ . Alors la fonction  $\mathcal{H}_p$  atteint son unique maximum pour  $p' = p$ .

Remarquons que la quantité  $\mathcal{H}_p(p)$  est au signe près l'**entropie** de  $p$ .

Par simplicité d'écriture, on note pour  $\theta, \theta_0 \in \Theta$ ,  $\mathcal{H}_{\theta_0}(\theta) = \mathcal{H}_{p_0}(p)$ , où  $p = p(\theta; \cdot)$  et  $p_0 = p(\theta_0; \cdot)$ . Comme  $p(\theta; X_1) \leq 1$ , on a

$$\mathcal{H}_{\theta_0}(\theta) = \sum_{x \in E} p(\theta_0; x) \log p(\theta, x) = \mathbb{E}[\log p(\theta; X_1)] \in [-\infty, 0],$$

et par la loi forte des grands nombres, cf. corollaire ??,

$$\frac{1}{n} L_n(\theta; X_1^n) = \frac{1}{n} \sum_{i=1}^n \log p(\theta; X_i) \xrightarrow[n \rightarrow \infty]{\text{p.s.}} \mathcal{H}_{\theta_0}(\theta). \quad (5.6)$$

Ceci suggère que  $\hat{\theta}_n = \operatorname{argmax}_{\theta \in \Theta} \frac{1}{n} L_n(\theta; X_1^n)$  converge presque sûrement vers  $\operatorname{argmax}_{\theta \in \Theta} \mathcal{H}_{\theta_0}(\theta)$  (i.e. vers  $\theta_0$  d'après le lemme 5.2.8) quand  $n$  tend vers l'infini; et donc que l'EMV est convergent. Plus précisément, on a le théorème suivant.

**Théorème 5.2.9.** *On suppose les conditions suivantes :*

1.  $\Theta$  est compact.
2. Le modèle est identifiable.
3. La vraisemblance définie sur  $\Theta$ ,  $\theta \rightarrow p(\theta; x)$ , est continue pour tout  $x \in E$ .
4. P.s. pour  $n$  assez grand, (5.5) définit uniquement l'EMV  $\hat{\theta}_n$ .
5. La quantité  $\mathcal{H}_{\theta_0}(\theta)$  est finie pour tout  $\theta \in \Theta$ .

Alors l'EMV de  $\theta$ , défini par (5.5), est un estimateur convergent.

*Démonstration.* On pose pour tout  $x \in E$ ,  $f_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i=x\}}$ , et on remarque que

$$\frac{1}{n} L_n(\theta; X_1^n) = \sum_{x \in E} \log p(\theta; x) f_n(x).$$

La loi forte des grands nombres assure que p.s. pour tout  $x \in E$ ,

$$f_n(x) \xrightarrow[n \rightarrow \infty]{\text{p.s.}} p(\theta_0; x).$$

Comme  $\Theta$  est compact, et que p.s. l'EMV est bien défini pour  $n$  assez grand, la suite des EMV admet au moins un point d'accumulation  $\theta_* \in \Theta$ . Et il existe p.s. une fonction strictement croissante (aléatoire),  $\sigma$ , de  $\mathbb{N}^*$  dans  $\mathbb{N}^*$ , telle que

la suite  $(\hat{\theta}_{\sigma(n)}, n \geq 1)$  converge vers  $\theta_*$ . Par continuité de la vraisemblance, on a pour tout  $x \in E$ ,

$$\log p(\hat{\theta}_{\sigma(n)}; x) \xrightarrow[n \rightarrow \infty]{\text{p.s.}} \log p(\theta_*; x).$$

Comme les fonctions  $f_{\sigma(n)}$  et  $-\log p(\hat{\theta}_{\sigma(n)}; \cdot)$  sont positives, on déduit du lemme de Fatou que p.s.

$$\begin{aligned} \liminf_{n \rightarrow \infty} -\frac{1}{\sigma(n)} L_{\sigma(n)}(\hat{\theta}_{\sigma(n)}; X_1^{\sigma(n)}) &= \liminf_{n \rightarrow \infty} \sum_{x \in E} -\log p(\hat{\theta}_{\sigma(n)}; x) f_{\sigma(n)}(x) \\ &\geq \sum_{x \in E} \liminf_{n \rightarrow \infty} -\log p(\hat{\theta}_{\sigma(n)}; x) f_{\sigma(n)}(x) \\ &= -\mathcal{H}_{\theta_0}(\theta_*). \end{aligned}$$

Comme  $\hat{\theta}_n$  est l'EMV, on a  $L_n(\theta_0; X_1^n) \leq L_n(\hat{\theta}_n; X_1^n)$ , et grâce à (5.6), p.s.

$$\begin{aligned} \liminf_{n \rightarrow \infty} -\frac{1}{\sigma(n)} L_{\sigma(n)}(\hat{\theta}_{\sigma(n)}; X_1^{\sigma(n)}) &\leq \liminf_{n \rightarrow \infty} -\frac{1}{\sigma(n)} L_{\sigma(n)}(\theta_0; X_1^{\sigma(n)}) \\ &= -\mathcal{H}_{\theta_0}(\theta_0). \end{aligned}$$

On déduit de ces inégalités, que p.s.  $\mathcal{H}_{\theta_0}(\theta_*) \geq \mathcal{H}_{\theta_0}(\theta_0)$ . Le modèle étant identifiable et  $\mathcal{H}_{\theta_0}(\theta_0)$  fini, on déduit du lemme 5.2.8 que p.s.  $\theta_* = \theta_0$ . Ceci implique que la suite des EMV admet p.s. un seul point d'accumulation  $\theta_0$ . Elle est donc p.s. convergente et sa limite est  $\theta_0$ . Ceci démontre donc le théorème.  $\square$

**Remarque 5.2.10.** Si la vraisemblance  $p_n(\cdot, X_1^n)$  atteint son maximum en plusieurs points l'EMV n'est pas défini (condition 4 du théorème 5.2.9 non vérifiée). Les arguments de la démonstration ci-dessus assurent en fait que toute suite  $(\hat{\theta}_n, n \geq 1)$ , telle que la vraisemblance  $p_n(\cdot, X_1^n)$  atteint son maximum en  $\hat{\theta}_n$ , admet p.s. un seul point d'accumulation qui est  $\theta_0$ . La suite converge donc p.s. vers la vraie valeur  $\theta_0$ . Ainsi, on peut étendre la définition de l'EMV à tout point de  $\Theta$  tel que la vraisemblance soit maximale en ce point, et conserver les propriétés de convergence.  $\diamond$

*Démonstration du lemme 5.2.8.* Remarquons que pour  $r \geq 0$ , on a  $\log r \leq r - 1$  avec égalité si et seulement si  $r = 1$ . Pour  $y > 0$ ,  $z \geq 0$ , on a  $y \log z - y \log y = y \log(z/y) \leq y \left( \frac{z}{y} - 1 \right) = z - y$ , avec égalité si et seulement si  $z = y$ . Avec la convention  $0 \log 0 = 0$ , on obtient que pour  $y \geq 0$ ,  $z \geq 0$

$$y \log z - y \log y \leq z - y, \quad (5.7)$$

avec égalité si et seulement si  $y = z$ . On a

$$\begin{aligned}\mathcal{H}_p(p') - \mathcal{H}_p(p) &= \sum_{x \in E} p(x) \log p'(x) - \sum_{x \in E} p(x) \log p(x) \\ &= \sum_{x \in E} p(x) [\log p'(x) - \log p(x)] \\ &\leq \sum_{x \in E} p'(x) - p(x) \\ &= 0,\end{aligned}$$

où l'on a utilisé le fait que  $\sum_{x \in E} p(x) |\log p(x)| < \infty$  pour la deuxième égalité et (5.7) pour l'inégalité. Ainsi on a  $\mathcal{H}_p(p') \leq \mathcal{H}_p(p)$ . Enfin comme (5.7) n'est une égalité que si  $y = z$ , on en déduit que  $\mathcal{H}_p(p') = \mathcal{H}_p(p)$  si et seulement si  $p' = p$ .  $\square$

### 5.3 Présentation générale de l'algorithme EM

On écrit  $\mathbb{P}_\theta$  et  $\mathbb{E}_\theta$  pour les probabilités et espérances calculées quand le vrai paramètre de la chaîne de Markov  $((S_n, Y_n), n \geq 1)$  est  $\theta = (a, b, \pi_0)$ . Pour abréger les notations, on notera  $S = S_1^N$ ,  $s = s_1^N \in \mathcal{I}^N$ ,  $Y = Y_1^N$  et  $y = y_1^N \in \mathcal{X}$ . La vraisemblance du modèle incomplet est définie par  $p_N(\theta; y) = \mathbb{P}_\theta(Y = y)$ . On a, en utilisant (5.1) et (5.2),

$$\begin{aligned}p_N(\theta; y) &= \sum_{s \in \mathcal{I}^N} \mathbb{P}_\theta(Y = y | S = s) \mathbb{P}_\theta(S = s) \\ &= \sum_{s \in \mathcal{I}^N} \left( \prod_{n=1}^N b(s_n, y_n) \right) \pi_0(s_1) \prod_{n=2}^N a(s_{n-1}, s_n).\end{aligned}\quad (5.8)$$

La log-vraisemblance du modèle incomplet est  $L_N(\theta; y) = \log p_N(\theta; y)$ . Pour déterminer l'EMV de  $\theta$ , il faut maximiser  $p_N(\cdot; y)$  en  $\theta = (a, b, \pi)$ . Bien sûr, il faut tenir compte des contraintes suivantes :  $\sum_{j \in \mathcal{I}} a(i, j) = 1$  pour tout  $i \in \mathcal{I}$  ( $a$  est la matrice de transition d'une chaîne de Markov),  $\sum_{x \in \mathcal{X}} b(i, x) = 1$  pour tout  $i \in \mathcal{I}$  ( $b(i, \cdot)$  est une probabilité) et  $\sum_{i \in \mathcal{I}} \pi_0(i) = 1$  ( $\pi_0$  est la loi de  $S_1$ ). On choisit  $\Theta'$ , défini à la fin du paragraphe 5.1, pour l'ensemble des paramètres possibles de sorte que le modèle paramétrique soit identifiable.

L'existence et la convergence de l'EMV sont présentées au paragraphe ??, voir le théorème ??. Pour calculer numériquement l'EMV, remarquons qu'il faut maximiser  $p_N(\theta; y)$ , un polynôme de degré  $2N$  à Card  $(\mathcal{I}^2 \times (\mathcal{I} \times \mathcal{X}) \times \mathcal{I})$  variables sous  $2\text{Card}(\mathcal{I}) + 1$  contraintes linéaires libres. Pour des applications courantes, on ne peut pas espérer calculer numériquement l'EMV par des algorithmes classiques d'optimisation. On peut, en revanche, utiliser des algorithmes de recuit simulé (cf. [16] pour la détection de zones homogènes de l'ADN).



Pour une autre approche, on considère la vraisemblance du modèle complet définie par  $p_N^{complet}(\theta; s, y) = \mathbb{P}_\theta(S = s, Y = y)$ . On a

$$p_N^{complet}(\theta; s, y) = \pi_0(s_1)b(s_1, y_1) \prod_{n=2}^N a(s_{n-1}, s_n)b(s_n, y_n).$$

Nous verrons au paragraphe 5.4.2 que le calcul de l'EMV pour le modèle complet est élémentaire.

La loi conditionnelle de  $S$  sachant  $Y$  est donnée par

$$\pi_N(\theta; s|y) = \mathbb{P}_\theta(S = s|Y = y) = \frac{\mathbb{P}_\theta(S = s, Y = y)}{\mathbb{P}_\theta(Y = y)} = \frac{p_N^{complet}(\theta; s, y)}{p_N(\theta; y)}. \quad (5.9)$$

On écrit artificiellement la log-vraisemblance du modèle incomplet, calculée pour  $\theta$ , avec la log-vraisemblance du modèle complet calculée pour  $\theta'$  distinct de  $\theta$  a priori. Comme  $\sum_{s \in \mathcal{I}^N} \pi_N(\theta'; s|y) = 1$ , on a

$$\begin{aligned} L_N(\theta; y) &= \log p_N(\theta; y) \\ &= \sum_{s \in \mathcal{I}^N} \pi_N(\theta'; s|y) \log p_N(\theta; y) \\ &= \sum_{s \in \mathcal{I}^N} \pi_N(\theta'; s|y) \log p_N^{complet}(\theta; s, y) - \sum_{s \in \mathcal{I}^N} \pi_N(\theta'; s|y) \log \pi_N(\theta; s|y), \end{aligned}$$

où l'on a utilisé la définition (5.9) de  $\pi_N(\theta; s|y)$  pour la dernière égalité. On pose, pour  $y \in \mathcal{X}^N$ ,

$$Q(\theta, \theta') = \sum_{s \in \mathcal{I}^N} \pi_N(\theta'; s|y) \log p_N^{complet}(\theta; s, y),$$

et

$$\mathcal{H}_{\theta'}(\theta) = \sum_{s \in \mathcal{I}^N} \pi_N(\theta'; s|y) \log \pi_N(\theta; s|y).$$

On a donc

$$L_N(\theta; y) = Q(\theta, \theta') - \mathcal{H}_{\theta'}(\theta).$$

On établit le lemme suivant.

**Lemme 5.3.1.** *Soit  $\theta'$  fixé. Soit  $\theta^*$  le (ou un) paramètre qui maximise la fonction  $\theta \rightarrow Q(\theta, \theta')$ . Alors  $L_N(\theta^*; y) \geq L_N(\theta'; y)$ .*

*Démonstration.* On déduit du lemme 5.2.8 que  $\mathcal{H}_{\theta'}(\theta^*) \leq \mathcal{H}_{\theta'}(\theta')$ . Comme  $Q(\theta^*, \theta') \geq Q(\theta', \theta')$ , cela implique que  $L_N(\theta^*; y) \geq L_N(\theta'; y)$ .  $\square$

L'algorithme EM (Espérance Maximisation) consiste à construire par récurrence une suite de paramètres  $(\theta^{(r)}, r \in \mathbb{N})$  de la manière suivante.  $\theta^{(0)} \in \Theta'$  est choisi de manière quelconque. On suppose  $\theta^{(r)}$  construit. On calcule  $Q(\theta, \theta^{(r)})$ . Il s'agit d'un calcul d'espérance (étape E). Puis, on choisit  $\theta^{(r+1)}$  tel que la fonction  $\theta \rightarrow Q(\theta, \theta^{(r)})$  atteigne son maximum en la valeur  $\theta^{(r+1)}$ . Il s'agit d'une maximisation (étape M). D'après le lemme précédent, la suite  $(L_N(\theta^{(r)}; y), r \in \mathbb{N})$  est donc croissante.

Soit  $\delta > 0$ . On considère l'hypothèse suivante notée  $(H_\delta)$  : *Pour tous  $i, j \in \mathcal{I}$  et  $x \in \mathcal{X}$ , on a  $a(i, j) > \delta$ ,  $b(i, x) > \delta$  et  $\pi_0(i) > \delta$ .*

Soit  $\Theta_\delta$  l'ensemble des paramètres  $\theta \in \Theta'$  qui vérifient la condition  $(H_\delta)$ . On suppose que l'on peut choisir  $\delta$  assez petit pour que le vrai paramètre  $\theta_0$  soit dans  $\Theta_\delta$ . On admet le théorème suivant qui découle des résultats de [14].

**Théorème 5.3.2.** *La suite construite par l'algorithme EM dans  $\Theta_\delta$ ,  $(\theta^{(r)}, r \in \mathbb{N})$ , converge vers l'EMV de  $\theta$ ,  $\hat{\theta}_N$ , dès que  $\theta^{(0)}$  est assez proche de  $\hat{\theta}_N$ .*

Comme le souligne le théorème, la difficulté de l'algorithme EM réside dans le choix du point d'initialisation  $\theta^{(0)}$ . On peut démontrer, sous des hypothèses assez générales, que la suite générée par l'algorithme EM converge vers un point en lequel la dérivée de la log-vraisemblance s'annule. Il peut très bien s'agir d'un point selle ou d'un maximum local et non du maximum global  $\hat{\theta}_N$ . De plus l'algorithme EM converge mal si le point initial se trouve dans une région où la log-vraisemblance ne varie pas beaucoup. Il existe des procédures pour introduire de l'aléatoire dans les premières itérations de l'algorithme (variante stochastique (SEM) de l'algorithme EM) afin de s'affranchir de ces problèmes. On peut aussi utiliser l'algorithme EM avec plusieurs points d'initialisation. On pourra consulter [10] pour des résultats précis concernant ces questions.

**Exemple 5.3.3.** Les calculs explicites du paragraphe suivant permettent d'implémenter facilement l'algorithme EM pour l'estimation des paramètres et des variables cachées pour les chaînes de Markov cachées. On peut ainsi vérifier la pertinence de cet algorithme sur des simulations. On choisit un exemple avec des paramètres proches de ceux estimés dans l'exemple de la séquence d'ADN du bactériophage lambda (voir le paragraphe 5.6 pour les résultats numériques et plus particulièrement (5.14) pour la valeur des paramètres estimés). On considère une simulation  $(s_1^{N_0}, y_1^{N_0})$  de la chaîne de Markov cachée  $(S_1^{N_0}, Y_1^{N_0})$ , avec  $N_0 = 48\,502$ ,  $\mathcal{I} = \{+1, -1\}$ ,  $\mathcal{X} = \{A, C, G, T\}$ , et les paramètres suivants :

$$a = \begin{pmatrix} 0.9999 & 0.0001 \\ 0.0002 & 0.9998 \end{pmatrix}, \quad b = \begin{pmatrix} 0.246 & 0.248 & 0.298 & 0.208 \\ 0.270 & 0.208 & 0.198 & 0.324 \end{pmatrix} \quad \text{et} \quad \pi_0 = (1, 0).$$

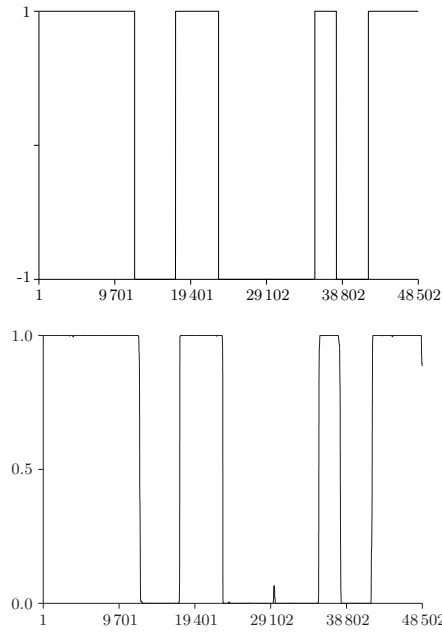
Après 1 000 itérations de l'algorithme EM, initialisé avec

$$a^{(0)} = \begin{pmatrix} 0.28 & 0.72 \\ 0.19 & 0.81 \end{pmatrix}, \quad b^{(0)} = \begin{pmatrix} 0.21 & 0.36 & 0.37 & 0.06 \\ 0.27 & 0.27 & 0.26 & 0.20 \end{pmatrix} \quad \text{et} \quad \pi_0^{(0)} = (0.5, 0.5), \quad (5.10)$$

on obtient l'estimation suivante des paramètres  $a$  et  $b$  :

$$a \simeq \begin{pmatrix} 0.99988 & 0.00012 \\ 0.00015 & 0.99985 \end{pmatrix}, \quad b \simeq \begin{pmatrix} 0.2456 & 0.2505 & 0.2946 & 0.2096 \\ 0.2723 & 0.2081 & 0.1952 & 0.3244 \end{pmatrix}.$$

Cette estimation dépend assez peu du point de départ pourvu que les termes diagonaux de  $a^{(0)}$  ne soient pas simultanément trop petits. Dans la Fig. 5.2, on présente les valeurs de la simulation  $s_1^{N_0}$  et les valeurs restaurées ( $\mathbb{P}(S_n = +1 | Y_1^{N_0} = y_1^{N_0}), n \in \{1, \dots, N_0\}$ ). La figure 5.3 (resp. 5.4) présente l'évolution au cours des itérations de l'algorithme EM, des coefficients diagonaux de  $a$  (resp. des coefficients de  $b$ ). On constate que les estimations sont constantes après un petit nombre (devant  $N_0$ ) d'itérations, et que les valeurs numériques estimées sont proches des vraies valeurs des paramètres.  $\diamond$

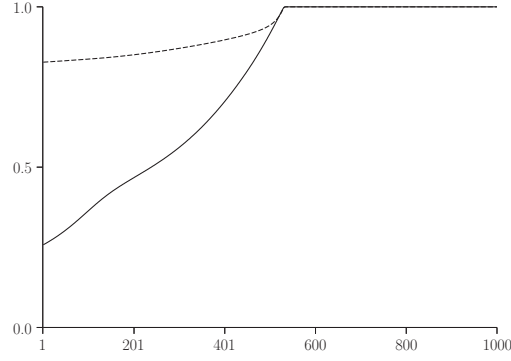


**Fig. 5.2.** Valeur des états cachés simulés ( $n \rightarrow S_n$ ) en haut et valeur des probabilités estimées de l'état caché  $+1$  ( $n \rightarrow \mathbb{P}(S_n = +1 | Y_1^{N_0} = y_1^{N_0})$ ) en bas

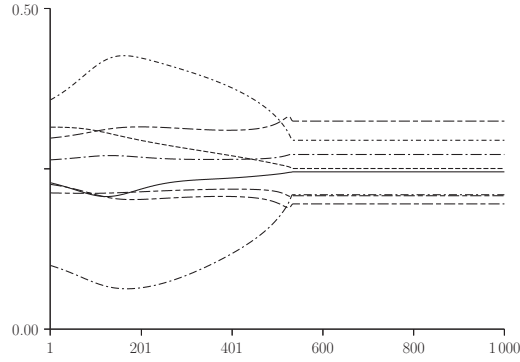
## 5.4 Mise en œuvre de l'algorithme EM

### 5.4.1 L'étape espérance : étape E

On suppose construit  $\theta^{(r)}$ . On désire construire  $\theta^{(r+1)}$ . Pour cela, il faut calculer  $Q(\theta; \theta^{(r)})$ . Soit  $\theta, \theta' \in \Theta_\delta$ . On note  $\theta = (a, b, \pi_0)$  et  $\theta' = (a', b', \pi'_0)$ . On rappelle que  $\mathbb{P}_{\theta'}$  et  $\mathbb{E}_{\theta'}$  désignent les probabilités et espérances quand le vrai



**Fig. 5.3.** Évolution de l'estimation des termes diagonaux de la matrice de transition des états cachés en fonction des itérations, obtenue pour 1 000 itérations



**Fig. 5.4.** Évolution de l'estimation des termes de la matrice  $b$  en fonction des itérations, obtenue pour 1 000 itérations

paramètre de la chaîne de Markov  $((S_n, Y_n), n \geq 1)$  est  $\theta'$ . Comme  $\theta' \in \Theta_\delta$ , remarquons que toutes les probabilités de transition sont strictement positives. En particulier, pour tous  $n \geq 1$ ,  $s_1^n \in \mathcal{I}^n$ ,  $y_1^n \in \mathcal{X}^n$ ,  $\mathbb{P}_{\theta'}(S_1^n = s_1^n, Y_1^n = y_1^n) > 0$ .

Avec les notations du paragraphe précédent, on calcule  $Q(\theta, \theta')$  pour  $y \in \mathcal{X}^N$  :

$$\begin{aligned}
 Q(\theta, \theta') &= \sum_{s \in \mathcal{I}^N} \pi_N(\theta'; s|y) \log p_N^{\text{complet}}(\theta; s, y) \\
 &= \mathbb{E}_{\theta'} [\log p_N^{\text{complet}}(\theta; S, y) | Y = y] \\
 &= \mathbb{E}_{\theta'} \left[ \log \left( \pi_0(S_1) b(S_1, y_1) \prod_{n=2}^N a(S_{n-1}, S_n) b(S_n, y_n) \right) \middle| Y = y \right]
 \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}_{\theta'}[\log \pi_0(S_1)|Y = y] + \mathbb{E}_{\theta'}[\log b(S_1, y_1)|Y = y] \\
&\quad + \sum_{n=2}^N \mathbb{E}_{\theta'}[\log a(S_{n-1}, S_n)|Y = y] + \sum_{n=2}^N \mathbb{E}_{\theta'}[\log b(S_n, y_n)|Y = y] \\
&= \sum_{i \in \mathcal{I}} \mathbb{P}_{\theta'}(S_1 = i|Y = y) \log \pi_0(i) \\
&\quad + \sum_{n=1}^N \sum_{i \in \mathcal{I}} \mathbb{P}_{\theta'}(S_n = i|Y = y) \log b(i, y_n) \\
&\quad + \sum_{n=2}^N \sum_{i, j \in \mathcal{I}} \mathbb{P}_{\theta'}(S_{n-1} = i, S_n = j|Y = y) \log a(i, j).
\end{aligned}$$

Nous devons donc calculer, pour la chaîne de Markov de paramètre  $\theta'$ , les probabilités  $\mathbb{P}_{\theta'}(S_{n-1} = i, S_n = j|Y = y)$  pour  $2 \leq n \leq N$  et  $\mathbb{P}_{\theta'}(S_n = i|Y = y)$  pour  $1 \leq n \leq N$ . Pour résoudre ce problème, appelé problème de **filtrage**, on effectue les étapes suivantes :

1. Prédire la valeur de  $S_n$  connaissant les observations partielles jusqu'à l'instant  $n - 1$ . Il s'agit de la prévision.
2. Estimer la valeur de  $S_n$  connaissant les observations partielles jusqu'à l'instant  $n$ . Il s'agit du filtrage.
3. Estimer la valeur de  $S_n$  connaissant les observations partielles jusqu'à l'instant final  $N$ . Il s'agit du lissage.

**Lemme 5.4.1 (Prévision).** On a, pour  $n \geq 2$ ,  $y_1^{n-1} \in \mathcal{X}^{n-1}$ ,

$$\mathbb{P}_{\theta'}(S_n = i|Y_1^{n-1} = y_1^{n-1}) = \sum_{j \in \mathcal{I}} a'(j, i) \mathbb{P}_{\theta'}(S_{n-1} = j|Y_1^{n-1} = y_1^{n-1}).$$

*Démonstration.* On décompose suivant les valeurs de  $S_1^{n-1}$  :

$$\begin{aligned}
&\mathbb{P}_{\theta'}(S_n = i|Y_1^{n-1} = y_1^{n-1}) \\
&= \sum_{s_1^{n-1} \in \mathcal{I}^{n-1}} \mathbb{P}_{\theta'}(S_n = i, S_1^{n-1} = s_1^{n-1}|Y_1^{n-1} = y_1^{n-1}) \\
&= \sum_{s_1^{n-1} \in \mathcal{I}^{n-1}} \mathbb{P}_{\theta'}(S_n = i|S_1^{n-1} = s_1^{n-1}, Y_1^{n-1} = y_1^{n-1}) \\
&\quad \mathbb{P}_{\theta'}(S_1^{n-1} = s_1^{n-1}|Y_1^{n-1} = y_1^{n-1}) \\
&= \sum_{s_1^{n-1} \in \mathcal{I}^{n-1}} a'(s_{n-1}, i) \mathbb{P}_{\theta'}(S_1^{n-1} = s_1^{n-1}|Y_1^{n-1} = y_1^{n-1}) \\
&= \sum_{s_{n-1} \in \mathcal{I}} a'(s_{n-1}, i) \mathbb{P}_{\theta'}(S_{n-1} = s_{n-1}|Y_1^{n-1} = y_1^{n-1}),
\end{aligned}$$

où l'on a utilisé le lemme 5.1.1 pour la troisième égalité.  $\square$

**Lemme 5.4.2 (Filtrage).** On a, pour  $n \geq 1$ ,  $y_1^n \in \mathcal{X}^n$ ,

$$\mathbb{P}_{\theta'}(S_n = i | Y_1^n = y_1^n) = \frac{b'(i, y_n) \mathbb{P}_{\theta'}(S_n = i | Y_1^{n-1} = y_1^{n-1})}{\sum_{j \in \mathcal{I}} b'(j, y_n) \mathbb{P}_{\theta'}(S_n = j | Y_1^{n-1} = y_1^{n-1})}.$$

Remarquons que les termes de prévision à l'instant  $n$  s'écrivent en fonction des termes de filtrage à l'instant  $n-1$ . Ces derniers s'écrivent en fonction des termes de prévision à l'instant  $n-1$ . On en déduit que l'on peut donc calculer les termes de prévision et de filtrage à l'instant  $n$  en fonction de  $a'$ ,  $b'$  et  $\mathbb{P}_{\theta'}(S_1 = i | Y_1 = y_1)$ . Or d'après la formule de Bayes, on a

$$\mathbb{P}_{\theta'}(S_1 = i | Y_1 = y_1) = \frac{\mathbb{P}_{\theta'}(S_1 = i, Y_1 = y_1)}{\sum_{j \in \mathcal{I}} \mathbb{P}_{\theta'}(S_1 = j, Y_1 = y_1)} = \frac{b'(i, y_1) \pi'_0(i)}{\sum_{j \in \mathcal{I}} b'(j, y_1) \pi'_0(j)}.$$

On en déduit donc que l'on peut exprimer les termes de prévision et de filtrage en fonction de  $\theta' = (a', b', \pi'_0)$ .

Avant de détailler la démonstration du lemme 5.4.2, nous démontrons un résultat technique intermédiaire.

**Lemme 5.4.3.** Soit  $\theta \in \Theta$ . Soit  $m \geq 0$ ,  $n \geq 2$ ,  $y_n^{n+m+1} \in \mathcal{X}^{m+1}$ ,  $s_k \in \mathcal{I}$  et  $J_n = \{(S_1^{n-1}, Y_1^{n-1}) \in B\}$ , où  $B \subset (\mathcal{I} \times \mathcal{X})^{n-1}$ . Si  $\mathbb{P}_\theta(S_n = s_n, J_n) > 0$ , alors on a

$$\mathbb{P}_\theta(Y_n^{n+m} = y_n^{n+m} | S_n = s_n, J_n) = \mathbb{P}_\theta(Y_n^{n+m} = y_n^{n+m} | S_n = s_n),$$

et si  $\mathbb{P}_\theta(Y_n^{n+m} = y_n^{n+m}, S_n = s_n, J_n) > 0$

$$\begin{aligned} \mathbb{P}_\theta(Y_{n+m+1} = y_{n+m+1} | Y_n^{n+m} = y_n^{n+m}, S_n = s_n, J_n) \\ = \mathbb{P}_\theta(Y_{n+m+1} = y_{n+m+1} | Y_n^{n+m} = y_n^{n+m}, S_n = s_n). \end{aligned}$$

*Démonstration.* On calcule dans un premier temps  $\mathbb{P}_\theta(Y_n = y_n | S_n = s_n, J_n)$ . En décomposant suivant les valeurs possibles de  $(S_1^{n-1}, Y_1^{n-1})$ , il vient

$$\begin{aligned} \mathbb{P}_\theta(Y_n = y_n, S_n = s_n, J_n) &= \sum_{(s_1^{n-1}, y_1^{n-1}) \in B} \mathbb{P}_\theta(S_1^n = s_1^n, Y_1^n = y_1^n) \\ &= \sum_{(s_1^{n-1}, y_1^{n-1}) \in B} \pi_0(s_1) b(s_1, y_1) \prod_{k=2}^n a(s_{k-1}, s_k) b(s_k, y_k), \end{aligned}$$

où l'on a utilisé pour la dernière égalité le fait que  $((S_k, Y_k), k \geq 1)$  est une chaîne de Markov, cf. le lemme 5.1.1, de loi initiale  $\mathbb{P}_\theta(S_1 = s_1, Y_1 = y_1) = \mathbb{P}_\theta(Y_1 = y_1 | S_1 = s_1) \mathbb{P}_\theta(S_1 = s_1) = \pi_0(s_1) b(s_1, y_1)$ . En sommant sur  $y_n \in \mathcal{X}$ , on en déduit que

$$\begin{aligned} \mathbb{P}_\theta(S_n = s_n, J_n) \\ = \sum_{(s_1^{n-1}, y_1^{n-1}) \in B} \pi_0(s_1) b(s_1, y_1) \left[ \prod_{k=2}^{n-1} a(s_{k-1}, s_k) b(s_k, y_k) \right] a(s_{n-1}, s_n), \end{aligned}$$

et donc si  $\mathbb{P}_\theta(S_n = s_n, J_n) > 0$ ,

$$\mathbb{P}_\theta(Y_n = y_n | S_n = s_n, J_n) = b(s_n, y_n) = \mathbb{P}_\theta(Y_n = y_n | S_n = s_n).$$

On suppose  $m \geq 1$ . On a alors

$$\begin{aligned} \mathbb{P}_\theta(Y_n^{n+m} = y_n^{n+m} | S_n = s_n, J_n) &= \mathbb{P}_\theta(Y_{n+1}^{n+m} = y_{n+1}^{n+m} | S_n = s_n, Y_n = y_n, J_n) \mathbb{P}_\theta(Y_n = y_n | S_n = s_n, J_n) \\ &= \mathbb{P}_\theta(Y_{n+1}^{n+m} = y_{n+1}^{n+m} | S_n = s_n, Y_n = y_n) \mathbb{P}_\theta(Y_n = y_n | S_n = s_n, J_n) \\ &= \mathbb{P}_\theta(Y_{n+1}^{n+m} = y_{n+1}^{n+m} | S_n = s_n, Y_n = y_n) \mathbb{P}_\theta(Y_n = y_n | S_n = s_n) \\ &= \mathbb{P}_\theta(Y_n^{n+m} = y_n^{n+m} | S_n = s_n), \end{aligned}$$

où on utilise la formule de décomposition des probabilités conditionnelles pour les première et dernière égalités, et la proposition ?? pour la chaîne de Markov  $((S_k, Y_k), k \geq 1)$  avec  $I_n = \{(S_{n+1}^{n+m}, Y_{n+1}^{n+m}) \in \mathcal{I}^m \times \{y_{n+1}^{n+m}\}\}$  pour la deuxième égalité. Ceci démontre la première égalité du lemme.

Remarquons, que grâce à la définition des probabilités conditionnelles, on a si  $\mathbb{P}_\theta(Y_n^{n+m} = y_n^{n+m}, S_n = s_n, J_n) > 0$ ,

$$\begin{aligned} \mathbb{P}_\theta(Y_{n+m+1} = y_{n+m+1} | Y_n^{n+m} = y_n^{n+m}, S_n = s_n, J_n) &= \frac{\mathbb{P}_\theta(Y_{n+m+1} = y_{n+m+1} | S_n = s_n, J_n)}{\mathbb{P}_\theta(Y_n^{n+m} = y_n^{n+m} | S_n = s_n, J_n)} \\ &= \frac{\mathbb{P}_\theta(Y_{n+m+1} = y_{n+m+1} | S_n = s_n)}{\mathbb{P}_\theta(Y_n^{n+m} = y_n^{n+m} | S_n = s_n)} \\ &= \mathbb{P}_\theta(Y_{n+m+1} = y_{n+m+1} | Y_n^{n+m} = y_n^{n+m}, S_n = s_n), \end{aligned}$$

où l'on a utilisé la première égalité du lemme deux fois pour obtenir l'avant-dernière égalité. Ceci termine la démonstration du lemme.  $\square$

*Démonstration du lemme 5.4.2.* On a

$$\begin{aligned} \mathbb{P}_{\theta'}(S_n = i | Y_1^n = y_1^n) &= \frac{\mathbb{P}_{\theta'}(Y_1^{n-1} = y_1^{n-1}, S_n = i, Y_n = y_n)}{\mathbb{P}_{\theta'}(Y_1^{n-1} = y_1^{n-1}, Y_n = y_n)} \\ &= \frac{\mathbb{P}_{\theta'}(Y_1^{n-1} = y_1^{n-1}, S_n = i, Y_n = y_n)}{\sum_{j \in \mathcal{I}} \mathbb{P}_{\theta'}(Y_1^{n-1} = y_1^{n-1}, S_n = j, Y_n = y_n)} \\ &= \frac{\mathbb{P}_{\theta'}(Y_n = y_n | Y_1^{n-1} = y_1^{n-1}, S_n = i) \mathbb{P}_{\theta'}(Y_1^{n-1} = y_1^{n-1}, S_n = i)}{\sum_{j \in \mathcal{I}} \mathbb{P}_{\theta'}(Y_n = y_n | Y_1^{n-1} = y_1^{n-1}, S_n = j) \mathbb{P}_{\theta'}(Y_1^{n-1} = y_1^{n-1}, S_n = j)} \\ &= \frac{\mathbb{P}_{\theta'}(Y_n = y_n | S_n = i) \mathbb{P}_{\theta'}(S_n = i | Y_1^{n-1} = y_1^{n-1})}{\sum_{j \in \mathcal{I}} \mathbb{P}_{\theta'}(Y_n = y_n | S_n = j) \mathbb{P}_{\theta'}(S_n = j | Y_1^{n-1} = y_1^{n-1})} \\ &= \frac{b'(i, y_n) \mathbb{P}_{\theta'}(S_n = i | Y_1^{n-1} = y_1^{n-1})}{\sum_{j \in \mathcal{I}} b'(j, y_n) \mathbb{P}_{\theta'}(S_n = j | Y_1^{n-1} = y_1^{n-1})}, \end{aligned}$$

où l'on a utilisé la première égalité du lemme 5.4.3 (avec  $m = 0$ ) pour la quatrième égalité.  $\square$

**Lemme 5.4.4 (Lissage).** *On a, pour  $2 \leq n \leq N$ ,  $y_1^N \in \mathcal{X}^N$ ,*

$$\begin{aligned} \mathbb{P}_{\theta'}(S_{n-1} = i, S_n = j | Y_1^N = y_1^N) \\ = a'(i, j) \frac{\mathbb{P}_{\theta'}(S_{n-1} = i | Y_1^{n-1} = y_1^{n-1})}{\mathbb{P}_{\theta'}(S_n = j | Y_1^{n-1} = y_1^{n-1})} \mathbb{P}_{\theta'}(S_n = j | Y_1^N = y_1^N), \end{aligned}$$

et, pour  $1 \leq n \leq N-1$ ,  $y_1^N \in \mathcal{X}^N$ ,

$$\begin{aligned} \mathbb{P}_{\theta'}(S_n = j | Y_1^N = y_1^N) \\ = \sum_{l \in \mathcal{I}} a'(j, l) \frac{\mathbb{P}_{\theta'}(S_n = j | Y_1^n = y_1^n)}{\mathbb{P}_{\theta'}(S_{n+1} = l | Y_1^n = y_1^n)} \mathbb{P}_{\theta'}(S_{n+1} = l | Y_1^N = y_1^N). \end{aligned}$$

*Démonstration.* Soit  $2 \leq n \leq N$ . En utilisant la première égalité du lemme 5.4.3 avec  $n + m = N$ , il vient

$$\begin{aligned} \mathbb{P}_{\theta'}(S_{n-1} = i, S_n = j, Y_1^N = y_1^N) \\ = \mathbb{P}_{\theta'}(Y_n^N = y_n^N | S_{n-1} = i, S_n = j, Y_1^{n-1} = y_1^{n-1}) \\ \mathbb{P}_{\theta'}(S_{n-1} = i, S_n = j, Y_1^{n-1} = y_1^{n-1}) \\ = \mathbb{P}_{\theta'}(Y_n^N = y_n^N | S_n = j) \mathbb{P}_{\theta'}(S_{n-1} = i, S_n = j, Y_1^{n-1} = y_1^{n-1}). \end{aligned}$$

On a aussi

$$\mathbb{P}_{\theta'}(S_n = j, Y_1^N = y_1^N) = \mathbb{P}_{\theta'}(Y_n^N = y_n^N | S_n = j) \mathbb{P}_{\theta'}(S_n = j, Y_1^{n-1} = y_1^{n-1}).$$

En particulier, en faisant le rapport de ces deux égalités, on obtient

$$\begin{aligned} \mathbb{P}_{\theta'}(S_{n-1} = i | S_n = j, Y_1^N = y_1^N) \\ = \frac{\mathbb{P}_{\theta'}(S_n = j, S_{n-1} = i, Y_1^{n-1} = y_1^{n-1})}{\mathbb{P}_{\theta'}(S_n = j, Y_1^{n-1} = y_1^{n-1})} \\ = \frac{\mathbb{P}_{\theta'}(S_n = j | S_{n-1} = i, Y_1^{n-1} = y_1^{n-1}) \mathbb{P}_{\theta'}(S_{n-1} = i | Y_1^{n-1} = y_1^{n-1})}{\mathbb{P}_{\theta'}(S_n = j | Y_1^{n-1} = y_1^{n-1})}. \end{aligned}$$

En utilisant la proposition ?? pour la chaîne de Markov  $((S_n, Y_n), n \geq 1)$  et le lemme 5.1.1, il vient



$$\begin{aligned}
\mathbb{P}_{\theta'}(S_n = j | S_{n-1} = i, Y_1^{n-1} = y_1^{n-1}) \\
&= \sum_{x \in \mathcal{X}} \mathbb{P}_{\theta'}(S_n = j, Y_n = x | S_{n-1} = i, Y_1^{n-1} = y_1^{n-1}) \\
&= \sum_{x \in \mathcal{X}} \mathbb{P}_{\theta'}(S_n = j, Y_n = x | S_{n-1} = i, Y_{n-1} = y_{n-1}) \\
&= \sum_{x \in \mathcal{X}} a'(i, j) b'(j, x) \\
&= a'(i, j).
\end{aligned}$$

On en déduit

$$\mathbb{P}_{\theta'}(S_{n-1} = i | S_n = j, Y_1^N = y_1^N) = a'(i, j) \frac{\mathbb{P}_{\theta'}(S_{n-1} = i | Y_1^{n-1} = y_1^{n-1})}{\mathbb{P}_{\theta'}(S_n = j | Y_1^{n-1} = y_1^{n-1})}.$$

On calcule maintenant  $\mathbb{P}_{\theta'}(S_{n-1} = i, S_n = j | Y_1^N = y_1^N)$ . On déduit de l'égalité précédente que

$$\begin{aligned}
\mathbb{P}_{\theta'}(S_{n-1} = i, S_n = j | Y_1^N = y_1^N) \\
&= \mathbb{P}_{\theta'}(S_{n-1} = i | S_n = j, Y_1^N = y_1^N) \mathbb{P}_{\theta'}(S_n = j | Y_1^N = y_1^N) \\
&= a'(i, j) \frac{\mathbb{P}_{\theta'}(S_{n-1} = i | Y_1^{n-1} = y_1^{n-1})}{\mathbb{P}_{\theta'}(S_n = j | Y_1^{n-1} = y_1^{n-1})} \mathbb{P}_{\theta'}(S_n = j | Y_1^N = y_1^N).
\end{aligned}$$

Il reste à calculer  $\mathbb{P}_{\theta'}(S_n = j | Y_1^N = y_1^N)$ . On déduit de l'égalité précédente, en remplaçant  $n$  par  $n+1$ , que, pour  $1 \leq n \leq N-1$ ,

$$\begin{aligned}
\mathbb{P}_{\theta'}(S_n = j | Y_1^N = y_1^N) \\
&= \sum_{l \in \mathcal{I}} \mathbb{P}_{\theta'}(S_n = j, S_{n+1} = l | Y_1^N = y_1^N) \\
&= \sum_{l \in \mathcal{I}} a'(j, l) \frac{\mathbb{P}_{\theta'}(S_n = j | Y_1^n = y_1^n)}{\mathbb{P}_{\theta'}(S_{n+1} = l | Y_1^n = y_1^n)} \mathbb{P}_{\theta'}(S_{n+1} = l | Y_1^N = y_1^N).
\end{aligned}$$

□

Remarquons que le calcul de  $\mathbb{P}_{\theta'}(S_N = j | Y_1^N = y_1^N)$  provient des équations de filtrage et de prévision. Son calcul nécessite le parcours complet de la suite  $y = y_1^N$ . À partir de cette quantité, on déduit des équations de lissage que l'on peut calculer par récurrence descendante  $\mathbb{P}_{\theta'}(S_n = j | Y_1^N = y_1^N)$  (on part donc de  $n = N$ ). Et parallèlement, on peut calculer les quantités  $\mathbb{P}_{\theta'}(S_{n-1} = i, S_n = j | Y_1^N = y_1^N)$ . Ces calculs nécessitent le parcours complet de la suite  $y = y_1^N$  de 1 à  $N$  puis de  $N$  à 1. On fait référence à ces calculs sous le nom d'algorithme « forward-backward ». On a ainsi calculé les coefficients de  $Q(\theta, \theta')$  qui sont fonction de  $\theta' = (a', b', \pi_0')$ .

### 5.4.2 L'étape maximisation : étape M

On recherche  $\theta = (a, b, \pi_0)$  qui maximise  $Q(\theta, \theta')$  à  $y$  et  $\theta'$  fixés. On maximise la quantité  $Q(\theta, \theta')$  définie par

$$\begin{aligned} \sum_{i \in \mathcal{I}} \mathbb{P}_{\theta'}(S_1 = i | Y = y) \log \pi_0(i) + \sum_{n=1}^N \sum_{i \in \mathcal{I}} \mathbb{P}_{\theta'}(S_n = i | Y = y) \log b(i, y_n) \\ + \sum_{n=2}^N \sum_{i, j \in \mathcal{I}} \mathbb{P}_{\theta'}(S_{n-1} = i, S_n = j | Y = y) \log a(i, j), \end{aligned}$$

sous les contraintes que  $(\pi_0(j), j \in \mathcal{I})$  est une probabilité, de même que  $(b(i, x), x \in \mathcal{X})$  et  $(a(i, j), j \in \mathcal{I})$  pour tout  $i \in \mathcal{I}$ . On remarque que l'on peut maximiser séparément les sommes correspondant à chacune des contraintes précédentes. Considérons par exemple la somme intervenant dans  $Q(\theta, \theta')$  qui fait intervenir les termes  $b(i, x)$  pour  $x \in \mathcal{X}$  et  $i$  fixé :

$$\sum_{n=1}^N \mathbb{P}_{\theta'}(S_n = i | Y = y) \log b(i, y_n).$$

On peut récrire cette somme de la manière suivante  $\sum_{x \in \mathcal{X}} q(x) \log b(i, x)$  où

$$q(x) = \sum_{n=1}^N \mathbf{1}_{\{y_n=x\}} \mathbb{P}_{\theta'}(S_n = i | Y = y).$$

Remarquons que la probabilité  $(b(i, x), x \in \mathcal{X})$  maximise  $\sum_{x \in \mathcal{X}} q(x) \log b(i, x)$  si et seulement si elle maximise la somme  $\sum_{x \in \mathcal{X}} p(x) \log b(i, x)$ , où  $p(x) = q(x) / \sum_{z \in \mathcal{X}} q(z)$ . La suite  $p = (p(x), x \in \mathcal{X})$  définit une probabilité sur  $\mathcal{X}$ . On déduit du lemme 5.2.8 que la somme est maximale pour  $b(i, \cdot) = p$ . On peut utiliser des arguments similaires pour calculer  $a$  et  $\pi_0$ . En définitive, on en déduit que  $Q(\theta, \theta')$  est maximal pour  $\theta = (a, b, \pi_0)$  défini pour  $i, j \in \mathcal{I}, x \in \mathcal{X}$  par

$$\begin{aligned} b(i, x) &= \frac{\sum_{n=1}^N \mathbf{1}_{\{y_n=x\}} \mathbb{P}_{\theta'}(S_n = i | Y = y)}{\sum_{n=1}^N \mathbb{P}_{\theta'}(S_n = i | Y = y)}, \\ a(i, j) &= \frac{\sum_{n=2}^N \mathbb{P}_{\theta'}(S_{n-1} = i, S_n = j | Y = y)}{\sum_{l \in \mathcal{I}} \sum_{n=2}^N \mathbb{P}_{\theta'}(S_{n-1} = i, S_n = l | Y = y)} \\ &= \frac{\sum_{n=2}^N \mathbb{P}_{\theta'}(S_{n-1} = i, S_n = j | Y = y)}{\sum_{n=1}^{N-1} \mathbb{P}_{\theta'}(S_{n-1} = i | Y = y)}, \\ \pi_0(i) &= \mathbb{P}_{\theta'}(S_1 = i | Y = y). \end{aligned}$$

Ceci termine l'étape M.

## 5.5 Autres exemples d'application de l'algorithme EM

### 5.5.1 Le mélange

Un des premiers exemples d'étude de loi de mélange remonte à la fin du  $XIX^{\text{ème}}$  siècle. Il s'agit aujourd'hui d'une problématique courante, voir par exemple [11] ou Chap. 9 dans [5].

#### Les crabes de Weldon

À la fin du  $XIX^{\text{ème}}$  siècle, Weldon mesure le rapport entre la largeur du front et la longueur du corps de 1 000 crabes de la baie de Naples. Le tableau 5.1 donne le nombre d'individus observés sur 29 intervalles pour le rapport des deux mesures (les mesures sont faites avec une précision du dixième de millimètre, et la longueur moyenne d'un animal est de 35 millimètres).

Si l'on suppose un modèle gaussien pour les données de ratio, on calcule à partir de (5.3) la moyenne empirique  $\mu_0 \simeq 0.645$  et l'écart type empirique, racine carrée de la variance empirique,  $\sigma_0 \simeq 0.019$ .

L'asymétrie des données, voir les histogrammes de la figure 5.5, indique que les données ne proviennent pas de réalisations de variables gaussiennes indépendantes et de même loi. Effectivement, un test d'adéquation de loi ( $\chi^2$ , Shapiro-Wilk, ..., cf. [2]) permet de rejeter l'hypothèse de normalité pour les données. Ceci induit Weldon à postuler l'existence de deux sous-populations. À partir de ces données, Pearson [13] estime les paramètres d'un

**Tableau 5.1.** Nombre de crabes de la baie de Naples (sur un total de 1 000 crabes) dont le ratio de la largeur du front par la longueur du corps sont dans les intervalles (Weldon, 1893)

Intervalle	Nombre	Intervalle	Nombre
[0.580, 0.584[	1	[0.640, 0.644[	74
[0.584, 0.588[	3	[0.644, 0.648[	84
[0.588, 0.592[	5	[0.648, 0.652[	86
[0.592, 0.596[	2	[0.652, 0.656[	96
[0.596, 0.600[	7	[0.656, 0.660[	85
[0.600, 0.604[	10	[0.660, 0.664[	75
[0.604, 0.608[	13	[0.664, 0.668[	47
[0.608, 0.612[	19	[0.668, 0.672[	43
[0.612, 0.616[	20	[0.672, 0.676[	24
[0.616, 0.620[	25	[0.676, 0.680[	19
[0.620, 0.624[	40	[0.680, 0.684[	9
[0.624, 0.628[	31	[0.684, 0.688[	5
[0.628, 0.632[	60	[0.688, 0.692[	0
[0.632, 0.636[	62	[0.692, 0.696[	1
[0.636, 0.640[	54		

modèle à  $I = 2$  populations différentes. Ainsi un crabe pris au hasard a une probabilité  $\pi_i$  d'appartenir à la population  $i$ ,  $\pi_i$  étant proportionnel à la taille de la population  $i$ . Et, au sein de la population  $i$ , les mesures du ratio sont distribuées suivant une loi gaussienne réelle de moyenne  $\mu_i$ , de variance  $\sigma_i^2$  et de densité  $f_{\mu_i, \sigma_i}$ . On suppose de plus que les mesures sont des réalisations de variables indépendantes  $(Y_n, n \geq 1)$ . L'objectif est d'estimer les probabilités  $\pi_i$  et les paramètres  $(\mu_i, \sigma_i), i \in \mathcal{I}$ . Le groupe,  $Z_n$ , du  $n$ -ième crabe mesuré est une variable cachée, que l'on essaie également de restaurer.

Dans ce qui suit, nous nous proposons d'estimer les paramètres avec leur EMV, que nous approchons à l'aide de l'algorithme EM. Plusieurs autres méthodes existent pour l'approximation de ces EMV, cf. [11, 5]. Historiquement, Pearson a estimé les paramètres de sorte que les cinq premiers moments de la loi de  $Y_n$  égalent les moments empiriques. Cette méthode conduit à rechercher les racines d'un polynôme de degré neuf.

### Le modèle de mélange.

Soit  $I \geq 2$  fixé. Le modèle complet est donné par une suite de variables aléatoires indépendantes de même loi,  $((Z_n, Y_n), n \geq 1)$ , où  $Z_n$  est à valeurs dans  $\mathcal{I} = \{1, \dots, I\}$ , de loi  $\pi = (\pi(i), i \in \mathcal{I})$ , et la loi de  $Y_n$  sachant  $Z_n = i$  a pour densité  $f_{\mu_i, \sigma_i}$ . On observe les réalisations des variables  $(Y_n, n \geq 1)$  et les variables  $(Z_n, n \geq 1)$  sont cachées. Il s'agit d'un **modèle de mélange** de lois gaussiennes. Le modèle est paramétrique, de paramètre inconnu  $\theta = (\pi, ((\mu_i, \sigma_i), i \in \mathcal{I})) \in \Theta = \mathcal{P}_{\mathcal{I}} \times (\mathbb{R} \times ]0, \infty])^I$ , où  $\mathcal{P}_{\mathcal{I}}$  est l'ensemble des probabilités sur  $\mathcal{I}$ .

Remarquons que le nombre  $I$  est fixé a priori. L'estimation du nombre  $I$  de populations différentes est un problème délicat en général, voir [11], Chap. 6.

Comme à la fin du paragraphe 5.1, il est facile de vérifier que le modèle est identifiable si l'on restreint l'ensemble des paramètres à  $\Theta' = \{\theta \in \Theta \text{ tels que si } i < j \in \mathcal{I}, \text{ alors soit } \mu_i < \mu_j \text{ soit } \mu_i = \mu_j \text{ et } \sigma_i < \sigma_j\}$ . On admet alors que l'EMV de  $\theta$ ,  $\hat{\theta}_n$ , construit à partir de  $(Z_1^n, Y_1^n)$ , est un estimateur convergent.

Pour déterminer la loi de  $Y_n$ , remarquons que pour tous  $a < b$ , on a, en utilisant la loi de  $Y_n$  sachant  $Z_n$ ,

$$\begin{aligned} \mathbb{P}(Y_n \in [a, b]) &= \sum_{i \in \mathcal{I}} \mathbb{P}(Y_n \in [a, b] | Z_n = i) \mathbb{P}(Z_n = i) \\ &= \sum_{i \in \mathcal{I}} \pi_i \int_{[a, b]} f_{\mu_i, \sigma_i}(y) dy = \int_{[a, b]} f_{\theta}(y) dy, \end{aligned}$$

avec  $f_{\theta} = \sum_{i \in \mathcal{I}} \pi_i f_{\mu_i, \sigma_i}$ . Ainsi,  $Y_n$  est une variable continue de densité  $f_{\theta}$ . Comme les variables  $(Y_n, n \geq 1)$  sont indépendantes, la vraisemblance du modèle associé à l'échantillon  $Y_1^N$  est pour  $y = y_1^N \in \mathbb{R}^N$  :

$$p_N(\theta; y) = \prod_{k=1}^N f_{\theta}(y_k),$$

et la log-vraisemblance

$$L_N(\theta; y) = \sum_{k=1}^N \log f_{\theta}(y_k).$$

La vraisemblance du modèle complet associé à l'échantillon  $(Z_1^N, Y_1^N)$  est pour  $z = z_1^N \in \mathcal{I}^N$ ,  $y = y_1^N \in \mathbb{R}^N$  :

$$p_N^{\text{complet}}(\theta; z, y) = \prod_{k=1}^N \pi_{z_k} f_{\mu_{z_k}, \sigma_{z_k}}(y_k).$$

Il est difficile de calculer numériquement l'EMV de  $\theta$  du modèle incomplet. L'algorithme EM, que nous explicitons, est rapide à mettre en œuvre dans ce cadre. Les mêmes arguments permettent de démontrer le lemme 5.3.1, avec  $Q$  défini ici par

$$Q(\theta, \theta') = \sum_{z \in \mathcal{I}^N} \pi_N(\theta'; z|y) \log p_N^{\text{complet}}(\theta; z, y),$$

où  $\theta, \theta' = (\pi', ((\mu'_i, \sigma'_i), i \in \mathcal{I})) \in \Theta'$ , et par définition

$$\pi_N(\theta'; z|y) = \frac{p_N^{\text{complet}}(\theta'; z, y)}{p_N(\theta'; y)} = \prod_{k=1}^N \rho'_{z_k, k},$$

où pour tous  $i \in \mathcal{I}, k \in \{1, \dots, N\}$

$$\rho'_{i, k} = \frac{\pi'_i f_{\mu'_i, \sigma'_i}(y_k)}{f_{\theta'}(y_k)}. \quad (5.11)$$

La quantité  $\rho'_{i, k}$  s'interprète comme la probabilité que  $Z_k = i$  sachant  $Y_k = y_k$ ,  $\theta'$  étant le paramètre du modèle. La quantité  $\pi_N(\theta'; z|y)$  s'interprète comme la loi conditionnelle des variables cachées  $Z_1^N$  sachant les variables observées  $Y_1^N$ .

### L'étape E

On explicite la fonction  $Q(\theta, \theta')$ . On remarque que

$$\log p_N^{\text{complet}}(\theta; z, y) = \sum_{k=1}^N \left[ \log(\pi_{z_k}) + \log(f_{\mu_{z_k}, \sigma_{z_k}}(y_k)) \right].$$

Comme pour tout  $l \in \{1, \dots, N\}$ , on a  $\sum_{j \in \mathcal{I}} \rho'_{j, l} = 1$ , il vient

$$\sum_{z \in \mathcal{I}^N; z_k = i} \pi_N(\theta'; z|y) = \rho'_{i, k}.$$

Cette égalité représente le calcul de la loi marginale de  $Z_k$  sachant  $Y_k$ . On en déduit donc que

$$Q(\theta, \theta') = \sum_{k=1}^N \sum_{i \in \mathcal{I}} \rho'_{i,k} [\log(\pi_i) + \log(f_{\mu_i, \sigma_i}(y_k))].$$

### L'étape M

On écrit  $Q(\theta, \theta') = NA_0 + \sum_{j \in \mathcal{I}} A_j$  avec  $A_0 = \sum_{i \in \mathcal{I}} \pi_i^* \log \pi_i$ ,

$$\pi_i^* = \frac{1}{N} \sum_{k=1}^N \rho'_{i,k}, \quad (5.12)$$

et  $A_j = \sum_{k=1}^N \rho'_{j,k} \log(f_{\mu_j, \sigma_j}(y_k))$  pour  $j \in \mathcal{I}$ . Remarquons que maximiser  $Q(\theta, \theta')$  en  $\theta \in \Theta'$  revient à maximiser séparément  $A_0$ , sous la contrainte que  $\pi \in \mathcal{P}_{\mathcal{I}}$ , et  $A_j$  pour  $j \in \mathcal{I}$ .

Comme  $\pi^* = (\pi_i^*, i \in \mathcal{I})$  est une probabilité sur  $\mathcal{I}$ , on déduit du lemme 5.2.8 que, sous la contrainte  $\pi \in \mathcal{P}_{\mathcal{I}}$ ,  $A_0$  est maximal pour  $\pi = \pi^*$ . On cherche ensuite les zéros des dérivées de  $A_j$  par rapport à  $\mu_j$  et  $\sigma_j$ . Comme  $\log f_{\mu, \sigma}(v) = -\frac{1}{2} \log(2\pi) - \log(\sigma) - \frac{(v - \mu)^2}{2\sigma^2}$ , on a

$$\frac{\partial A_j}{\partial \mu_j} = \sum_{k=1}^N \rho'_{j,k} \frac{(y_k - \mu_j)}{\sigma_j^2},$$

et

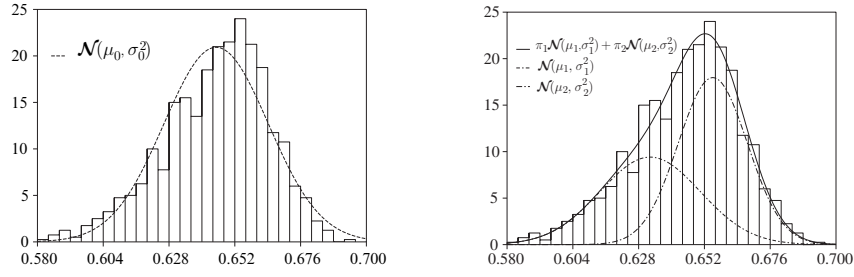
$$\frac{\partial A_j}{\partial \sigma_j} = - \sum_{k=1}^N \rho'_{j,k} \frac{1}{\sigma_j} \left[ 1 - \frac{(y_k - \mu_j)^2}{\sigma_j^2} \right].$$

Les deux dérivées ci-dessus s'annulent en

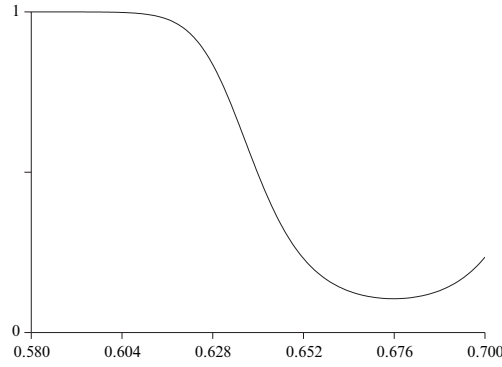
$$\mu_j^* = \frac{\sum_{k=1}^N \rho'_{j,k} y_k}{\sum_{k=1}^N \rho'_{j,k}} \quad \text{et} \quad (\sigma_j^*)^2 = \frac{\sum_{k=1}^N \rho'_{j,k} (y_k - \mu_j^*)^2}{\sum_{k=1}^N \rho'_{j,k}}. \quad (5.13)$$

On vérifie aisément que  $A_j$  possède un unique maximum pour  $(\mu_j, \sigma_j) \in \mathbb{R} \times ]0, \infty[$ , et qu'il est atteint en  $(\mu_j^*, \sigma_j^*)$ . On en déduit donc que  $\theta \rightarrow Q(\theta, \theta')$  atteint son unique maximum en  $\theta^* = (\pi^*, ((\mu_i^*, \sigma_i^*), i \in \mathcal{I}))$ .

L'algorithme EM consiste donc, à partir d'un point initial  $\theta^{(0)} \in \Theta'$ , à itérer les opérations définies par (5.11), (5.12) et (5.13). On pourra remarquer que les actualisations (5.12) et (5.13) s'interprètent comme le calcul de la moyenne empirique et de la variance empirique pondérées par  $\rho'_{i,\cdot}$ , qui est la probabilité, calculée avec les anciens paramètres, d'être dans la population  $i$ .



**Fig. 5.5.** Histogrammes des mesures pour les crabes de Weldon avec, à droite, la densité de la loi gaussienne  $\mathcal{N}(\mu_0, \sigma_0^2)$  et, à gauche, les densités des lois gaussiennes  $\mathcal{N}(\mu_1, \sigma_1^2)$ ,  $\mathcal{N}(\mu_2, \sigma_2^2)$  et la densité de la loi mélange  $f_\theta = \pi_1 \mathcal{N}(\mu_1, \sigma_1^2) + \pi_2 \mathcal{N}(\mu_2, \sigma_2^2)$ . On remarque une meilleure adéquation de la densité  $f_\theta$  (figure de droite) aux données par rapport à la densité gaussienne  $\mathcal{N}(\mu_0, \sigma_0^2)$  (figure de gauche)



**Fig. 5.6.** Restauration des données manquantes pour les crabes de Weldon : probabilité d'appartenir à la population 1, sachant la valeur du ratio

## Résultats

La limite et la convergence de l'algorithme dépendent peu du point de départ  $\theta^{(0)}$ . Dans l'exemple des crabes de Weldon, on obtient les valeurs numériques suivantes estimées par l'algorithme EM pour les paramètres du mélange :

$$\begin{aligned} \pi_1 &\simeq 0.434, & \mu_1 &\simeq 0.632, & \sigma_1 &\simeq 0.018, \\ \pi_2 &\simeq 0.566, & \mu_2 &\simeq 0.655, & \sigma_2 &\simeq 0.013. \end{aligned}$$

La figure 5.5 permet de visualiser l'adéquation des données à la densité pour les paramètres estimés. Enfin, dans la Fig. 5.6 on trace la probabilité d'appartenir à la population 1, sachant la valeur  $y$  du ratio :

$$y \rightarrow \pi(\theta; 1|y) = \frac{\pi_1 f_{\mu_1, \sigma_1}(y)}{\pi_1 f_{\mu_1, \sigma_1}(y) + \pi_2 f_{\mu_2, \sigma_2}(y)}.$$

Pour confirmer l'adéquation des données à la loi du mélange, on peut faire un test du  $\chi^2$ , voir [2]. On peut également proposer un modèle paramétrique à partir de la loi de Weibull qui est asymétrique. Cette approche fournit également une bonne adéquation aux données et donne une interprétation différente des observations.

## 5.6 Conclusion

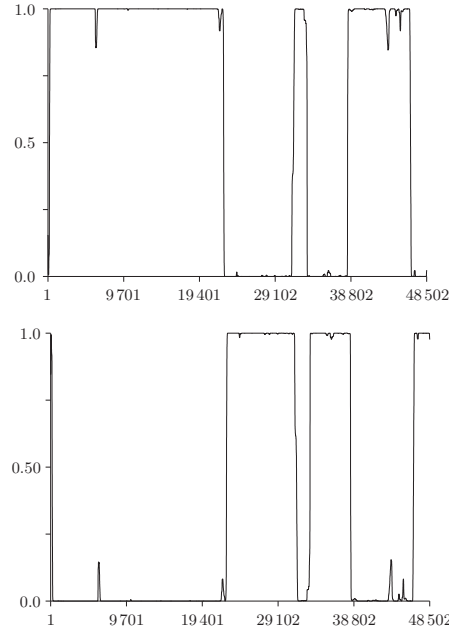
Pour le bactériophage lambda, on obtient après 1000 itérations de l'algorithme EM, initialisé avec (5.10), l'approximation suivante de l'EMV des paramètres  $a = (a(i, j); i, j \in \{-1, +1\})$  et  $b = (b(i, j); i \in \{-1, +1\}, j \in \{A, C, G, T\})$  :

$$a \simeq \begin{pmatrix} 0.99988 & 0.00012 \\ 0.00023 & 0.99977 \end{pmatrix}, \quad b \simeq \begin{pmatrix} 0.24635 & 0.24755 & 0.29830 & 0.20780 \\ 0.26974 & 0.20845 & 0.19834 & 0.32347 \end{pmatrix}. \quad (5.14)$$

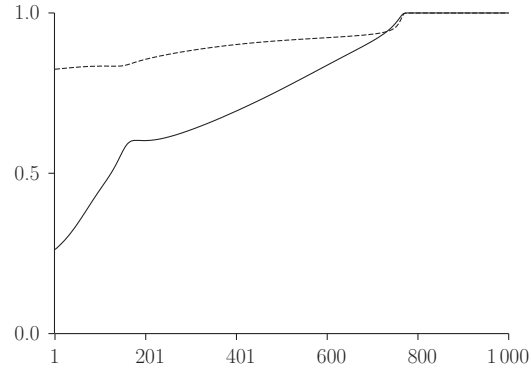
L'algorithme EM fournit également les valeurs de  $\mathbb{P}(S_n = i | Y = y)$  par les équations de lissage, calculées avec l'approximation (5.14) de l'EMV de  $\theta$ . La figure 5.7 met en évidence la présence de six grandes zones homogènes associées aux valeurs cachées +1 ou -1. Ces zones correspondent à des proportions différentes des quatre nucléotides. Ces proportions différentes pourraient provenir du fait que la transcription se fait sur le brin d'ADN analysé ou sur le brin apparié (voir [6]). Enfin les Figs. 5.8 et 5.9 représentent l'évolution des paramètres estimés, termes diagonaux de la matrice  $a$ , et termes de la matrice  $b$ , en fonction du nombre d'itérations de l'algorithme EM. On observe une convergence très rapide de l'algorithme. Toutefois, pour des initialisations éloignées des valeurs données dans (5.14), on observe une convergence de l'algorithme EM vers une valeur différente, correspondant à un maximum local de la log-vraisemblance.

Si on augmente le nombre de valeurs cachées possibles, certaines des zones précédentes se divisent, mais les résultats deviennent moins nets. On peut également choisir des modèles plus compliqués (et donc avec plus de paramètres) où la loi de  $Y_n$  peut dépendre de  $S_n$  et aussi de  $Y_{n-1}$  ; on peut aussi tenir compte dans les modèles du fait que trois nucléotides codent pour un acide aminé, etc. Nous renvoyons à [12] pour une étude très détaillée des différents modèles et des résultats obtenus pour chacun. On retrouve également dans ces modèles un découpage de l'ADN proche du cas présenté ici, où l'on se limite à deux états cachés. Les résultats sont donc robustes. Ils suggèrent donc vraiment l'existence de six zones homogènes de deux types différents.

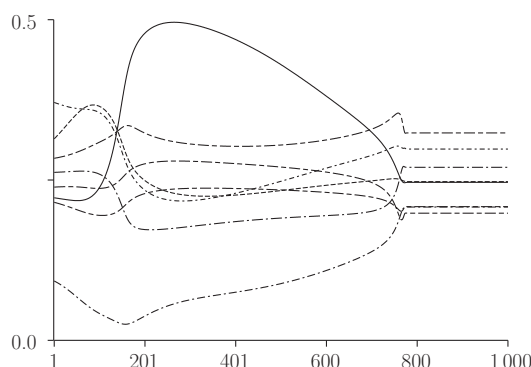




**Fig. 5.7.** Probabilité des états cachés pour la séquence d'ADN du bactériophage lambda dans un modèle à deux états cachés :  $\mathcal{I} = \{-1, 1\}$ , obtenu avec 1 000 itérations (de haut en bas :  $n \rightarrow \mathbb{P}(S_n = i | Y_1^{N_0} = y_1^{N_0})$ , pour  $i = -1, +1$ )



**Fig. 5.8.** Évolution de l'estimation des termes diagonaux de la matrice de transition des états cachés en fonction des itérations, obtenue pour 1 000 itérations



**Fig. 5.9.** Évolution de l'estimation des termes de la matrice  $b$  en fonction des itérations, obtenue pour 1 000 itérations

## Références

1. L. Baum et T. Petrie. Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Stat.*, 37 : 1554–1563, 1966.
2. P. Bickel et K. Doksum. *Mathematical statistics. Basic ideas and selected topics*. Holden-Day Series in Probability and Statistics. Holden-Day, San Francisco, 1977.
3. P. Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.
4. A.A. Borovkov. *Mathematical statistics*. Gordon and Breach Science Publishers, Amsterdam, 1998.
5. R. Casella et C. Robert. *Monte Carlo statistical methods*. Springer texts in statistics. Springer, 1999.
6. G. Churchill. Hidden Markov chains and the analysis of genome structure. *Comput. Chem.*, 16(2) : 105–115, 1992.
7. F. Dellaert. *Monte Carlo EM for data-association and its application in computer vision*. PhD thesis, Carnegie Mellon (Pittsburgh, U.S.A.), 2001.
8. A.P. Dempster, N.M. Laird et D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. of the Royal Stat. Soc. B*, 39 : 1–38, 1977.
9. D. Forsyth et J. Ponce. *Computer vision – A modern approach*. Prentice Hall, 2003.
10. G. McLachlan et T. Krishnan. *The EM algorithm and extensions*. Wiley Series in Probability and Mathematical Statistics. Wiley & Sons, 1997.
11. G. McLachlan et D. Peel. *Finite mixture models*. Wiley Series in Probability and Mathematical Statistics. Wiley & Sons, 2001.
12. F. Muri. *Comparaison d'algorithmes d'identification de chaînes de Markov cachées et application à la détection de régions homogènes dans les séquences d'ADN*. Thèse, Université René Descartes (Paris V), 1997.

13. K. Pearson. Contributions to the theory of mathematical evolution. *Phil. Trans. of the Royal Soc. of London A*, 185 : 71–110, 1894.
14. R. Redner et H. Walker. Mixture densities, maximum likelihood and the EM algorithm. *SIAM Rev.*, 26(2) : 195–239, 1984.
15. F. Sanger, A. Coulson, G. Hong, D. Hill et G. Petersen. Nucleotide sequence of bacteriophage lambda DNA. *J. Mol. Biol.*, 162(4) : 729–773, 1982.
16. P. Vandekerkhove. *Contribution à l'étude statistique des chaînes de Markov cachées*. Thèse, Université de Montpellier II, 1997.

Modèles aléatoires

Applications aux sciences de l'ingénieur et du vivant

Delmas, J.-F.; Jourdain, B.

2006, XXV, 431 p., Softcover

ISBN: 978-3-540-33282-4