

CHAPITRE IV

***UTILISATION DES METHODES DE
PERTURBATIONS SINGULIERES ET DU
DECOUPLAGE EN REGIME PERMANENT POUR
LA SIMPLIFICATION DES CHAINES DE
MARKOV A COMMANDE***

I. Introduction

Lors de la prise des décisions, nous nous trouvons souvent confrontés à des problèmes ayant une ou plusieurs des caractéristiques suivantes :

- les objectifs sont multiples, difficiles à mesurer et parfois même contradictoires,
- les décisions doivent être prises dans un environnement de risque et d'incertitude,
- il est nécessaire de contrôler l'impact des décisions sur l'évolution future du système.

La prise en compte de toutes ces contraintes dans les systèmes étudiés nécessite une modélisation adéquate, avec l'association des critères correspondant aux performances exigées. L'évaluation complète des performances d'un système qui évolue dans un environnement incertain impose l'utilisation des modèles aléatoires. Dans ce contexte, les chaînes de Markov à commande représentent un outil mathématique puissant. La programmation dynamique et les techniques de commande optimale associées à ce type de modèles permettent une analyse rigoureuse du problème.

Suite au grand nombre d'états à prendre en compte, dans la majorité des cas pratiques les algorithmes de calcul de la commande d'une chaîne de Markov deviennent lourds et parfois peuvent diverger. Pour éviter cet inconvénient, leur résolution demande l'introduction des méthodes de simplification.

Dans ce chapitre, nous proposons l'application des résultats des chapitres précédents pour simplifier le calcul de la commande optimale des chaînes de Markov ergodiques.

II. Généralités sur les chaînes de Markov à commande

Les chaînes de Markov à commande sont des modèles décrivant l'évolution des systèmes dynamiques de nature stochastique, contrôlés par des séquences de décisions ou actions. Nous considérons des systèmes observés périodiquement et influencés au moment de l'observation par la prise d'une décision parmi un ensemble de possibilités.

II.1. Eléments de base des chaînes de Markov à commande [BEL, 57], [PAL, 66], [BOU, 69] [DER, 70]

Soit un système stochastique, observé à des instants discrets $n = 0, 1, 2, \dots$. Le résultat ξ_n de cette observation est une variable aléatoire discrète identifiée à un état de l'ensemble fini d'états possibles du système :

$$E = \{e_1, e_2, \dots, e_r\}. \quad (IV.1)$$

Après chaque observation ξ_n , une décision ϕ_n est prise pour faire évoluer le système. Cette décision est modélisée par un *vecteur de commande* noté par $U = [u_1 \ u_2 \ \dots \ u_K]^T$. Chaque composante scalaire u_k ($k = 1, \dots, K$) de ce vecteur prend des valeurs continues dans un intervalle de commande $[u_{k \min}, u_{k \max}]$ donné :

$$u_k \in \hat{e}, \quad u_{k \min} \leq u_k \leq u_{k \max}, \quad \forall k = 1, 2, \dots, K \quad (IV.2)$$

Dans ce qui suit, nous allons noter par Ω le domaine de variation de la commande U :

$$\Omega = [U_{\min}, U_{\max}] \quad \text{où} \quad U_{\min} = \begin{bmatrix} u_{1 \min} \\ u_{2 \min} \\ \dots \\ u_{K \min} \end{bmatrix}, \quad U_{\max} = \begin{bmatrix} u_{1 \max} \\ u_{2 \max} \\ \dots \\ u_{K \max} \end{bmatrix} \quad (IV.3)$$

Notre étude porte sur des chaînes de Markov homogènes; par conséquent, les probabilités de transition entre les différents états se définissent comme étant :

$$p_{ij}(U) : \Omega \rightarrow [0, 1] \quad \text{tel que} \quad p_{ij}(U) = \text{Prob}\{\xi_{n+1} = e_j / \xi_n = e_i, \phi_n = U\} \quad (IV.4)$$

$$i, j = 1, \dots, r \quad \text{et} \quad n = 0, 1, 2, \dots$$

En utilisant les éléments que nous avons définis, l'équation fondamentale d'une chaîne de Markov homogène finie à commande peut s'écrire sous la forme suivante :

$$P(n+1) = P(n) \dagger(U) \quad (IV.5)$$

où $P(n) = [P_1(n) \ \dots \ P_r(n)]$ représente le vecteur des probabilités d'état à l'instant n :

$$\text{avec} \quad P_i(n) \in [0, 1], \quad \forall i = 1, \dots, r, \quad \forall n = 0, 1, 2, \dots,$$

$$\sum_{i=1}^r P_i(n) = 1, \quad \forall n = 0, 1, 2, \dots,$$

$\dagger(U)$ représente la matrice de transition de la chaîne d'éléments $[p_{ij}(U)]_{i,j=1 \dots r}$,

et $U = \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_K \end{bmatrix} \in \Omega$ le vecteur de commande appliqué.

Les probabilités de transition du modèle vérifient la propriété de stochasticité :

$$p_{ij}(U) \in [0, 1], \quad \forall i, j = 1, \dots, r, \quad \forall U \in \Omega$$

$$\sum_{j=1}^r p_{ij}(U) = 1, \forall i = 1, \dots, r, \forall U \in \Omega. \quad (IV.6)$$

Lorsqu'une décision est prise, le vecteur de commande U prend une valeur précise dans Ω et fixe la configuration des probabilités de transition.

Pour évaluer les performances du système étudié, nous faisons correspondre à chaque transition d'un état e_i à un autre état e_j , un *revenu élémentaire* r_{ij} . Il nous vient alors une *matrice de revenus* \mathbf{R} associée à notre système :

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1r} \\ r_{21} & r_{22} & \dots & r_{2r} \\ \dots & \dots & \dots & \dots \\ r_{r1} & r_{r2} & \dots & r_{rr} \end{pmatrix} \quad (IV.7)$$

Nous appelons *loi de commande* ou *stratégie de commande*, toute application

$$g : \mathbb{N} \times E \rightarrow \Omega \text{ telle que } g(n, e_i) = U$$

qui permet de déterminer une commande en fonction de l'instant n et de l'état e_i du système. Cette commande est optimale lorsqu'elle permet d'atteindre un *objectif* donné en optimisant un *critère de revenu*. Vu l'évolution aléatoire du processus, nous recherchons la loi de commande assurant l'optimum de *l'espérance mathématique du revenu total*, qui dans le cas général des modèles aléatoires prend la forme suivante [WHI, 88] :

$$\Psi_i \{ p_{ij}(g(n, e_i)), r_{ij} \} \quad (IV.8)$$

avec Ψ_i l'opérateur d'espérance mathématique. Pour les chaînes de Markov, le critère (IV.8) présente la forme simplifiée suivante [BOU, 69] :

$$\sum_n [p_{ij}(g(n, e_i)) \times r_{ij}] \quad (IV.9)$$

II.2. Classification des problèmes de commande [PAL, 66], [BOU, 69], [WHI, 88]

Les problèmes de commande se classifient en fonction de l'objectif imposé. Deux grandes catégories peuvent être mises en évidence : les problèmes terminaux et les problèmes avec cible.

II.2.1. Problèmes terminaux

Nous appelons *problèmes terminaux*, les problèmes pour lesquels nous cherchons à atteindre un revenu total optimal durant une évolution limitée dans le temps. Un cas particulier est représenté par les problèmes pour lesquels les probabilités d'être dans un état donné à un instant donné doit être optimisée.

L'instant final de cette évolution porte le nom d'*horizon*. La nature de ce dernier nous donne la sous classification suivante :

- problèmes terminaux à horizon fixé ;
- problèmes terminaux à horizon aléatoire.

a) Problèmes terminaux à horizon fixé

La classe des problèmes terminaux à horizon fixé concerne les systèmes pour lesquels la politique de commande doit assurer l'optimisation d'un critère de revenu pour une durée d'évolution déterminée à l'avance. L'horizon N donné par l'instant final peut être fini ou infini.

a.1. Horizon fini

Soit $w_i^O(n)$ l'espérance de revenu total entre l'instant n et l'horizon fini N , sachant qu'à l'instant n le système se trouve dans l'état e_i et qu'on utilise la politique optimale.

L'utilisation de la commande U donne la probabilité $p_{ij}(U)$ d'être dans l'état e_j à l'instant $n+1$. L'espérance de gain correspondant à cette transition est une fonction du type (IV.9) exprimée par le produit entre la probabilité de transition $p_{ij}(U)$ et le revenu élémentaire $r_{ij}(U)$. Une fois que le système se trouve dans l'état suivant e_j , on aura l'espérance $w_j^O(n+1)$ de revenu total jusqu'à l'instant final N en utilisant la politique optimale. Nous obtenons ainsi le critère d'optimalité suivant :

$$w_i^O(n) = \underset{U \in \Omega}{\text{opt}} \left\{ \sum_{j=1}^r \left(p_{ij}(U) r_{ij} + p_{ij}(U) w_j^O(n+1) \right) \right\} \quad (\text{IV.10})$$

Cette formule nous indique que l'espérance de revenu total optimal à l'instant n se décompose en une espérance de revenu acquis lors de la première transition ($n \rightarrow n+1$) et d'une espérance de revenu total obtenu à partir de l'instant suivant ($n+1 \rightarrow N$). Pour $i = 1, \dots, r$, nous obtenons $\mathbf{w}^O(n)$ le vecteur des espérances de revenus totaux optimaux perçus entre l'instant n et l'instant final N :

$$\mathbf{w}^O(n) = \begin{bmatrix} w_1^O(n) \\ w_2^O(n) \\ \dots \\ w_r^O(n) \end{bmatrix} \quad (\text{IV.11})$$

En notant l'espérance mathématique de revenu immédiat perçu au cours d'une transition si le système se trouve dans l'état e_i et si la commande appliquée est U par q_i :

$$q_i(U) = \sum_{j=1}^r \left(p_{ij}(U) r_{ij} \right) \quad (\text{IV.12})$$

nous obtenons le vecteur des espérances de revenus immédiats $\mathbf{q}(U)$ tel que :

$$\mathbf{q}(\mathbf{U}) = \begin{bmatrix} q_1(\mathbf{U}) \\ q_2(\mathbf{U}) \\ \dots \\ q_r(\mathbf{U}) \end{bmatrix} = \text{diag} \left(\dagger(\mathbf{U}) \times \mathcal{R}^T \right) \quad (\text{IV.13})$$

Avec les notations (IV.11) et (IV.13), la relation (IV.10) s'écrit sous la forme vectorielle suivante :

$$\mathbf{w}(\mathbf{n}) = \underset{\mathbf{U} \in \Omega}{\text{opt}} \left\{ \mathbf{q}(\mathbf{U}) + \dagger(\mathbf{U}) \mathbf{w}(\mathbf{n}+1) \right\}, \quad \text{À } n \in \{0, 1, 2, 3, \dots, N-1\}. \quad (\text{IV.14})$$

Cette relation permet de calculer $\mathbf{w}(\mathbf{n})$ pour la commande optimale correspondante. Le calcul de celle-ci se fait à partir de l'instant final N vers l'instant présent n (donc dans le sens chronologiquement inverse). Dans ce calcul, nous supposons $\mathbf{w}(N)$ connu :

$$\mathbf{w}(N) = \begin{bmatrix} w_1(N) \\ w_2(N) \\ \dots \\ w_r(N) \end{bmatrix} \quad (\text{IV.15})$$

avec $w_i(N)$ le revenu final perçu si à l'instant final N le système se trouve dans l'état e_i .

En optimisant (minimisant ou maximisant - selon le type de problème) le deuxième membre de la relation (IV.14), nous obtenons une commande optimale $\mathbf{U}^*(n, i) \in \Omega$ correspondant à chaque instant $n = 0, 1, 2, \dots, N-1$ et à chaque état e_i ($i = 1, 2, \dots, r$). Nous pouvons représenter la stratégie de commande optimale par une fonction $\mathbf{G}(n, i)$ donnée par la matrice de dimension $(K \times r)$ suivante :

$$\mathbf{G}(n) = [\hat{\mathbf{U}}(n,1) \ \hat{\mathbf{U}}(n,2) \ \dots \ \hat{\mathbf{U}}(n,r)] , \quad \text{À } n = 0, 1, \dots, N-1. \quad (\text{IV.16})$$

Ceci correspond à une politique optimale en boucle fermée [BOU, 69].

a.2. Horizon infini

La commande optimale avec horizon infini correspond à la recherche d'un revenu total optimal en effectuant un nombre infini de transitions ($N = \infty$).

Une période d'évolution très importante impose des précautions lors de la modélisation. En effet, un modèle valable au début de l'évolution peut devenir imprécis au fur et à mesure des transitions. Dans le cas de la commande optimale avec horizon infini, la prise en compte de ce phénomène se fait par l'introduction d'un facteur de perte $\gamma \in [0, 1]$. Plus la valeur de γ est proche de 1, plus les pertes d'informations sont réduites, donc le modèle initial reste fiable à l'infini. L'équation d'optimalité prend dans ce cas la forme suivante [WHI, 88] :

$$w_n = \underset{U \in \Omega}{\text{opt}} \left\{ q(U) + \gamma \sum_{i \in I} p_{ij}(U) w_{n+1} \right\}, \quad n = 0, 1, 2, 3, \dots \quad (\text{IV.17})$$

Dans le cas des chaînes de Markov avec $\gamma < 1$ (modèle imprécis à long terme), l'espérance de revenus w_n converge vers une valeur finie. Le calcul de la commande optimale se fait alors de la même façon que dans le cas de l'horizon fini, en prenant en compte le *vecteur des revenus finaux* w_N considéré comme étant donné pour le cas limite $N \rightarrow \infty$.

Dans le cas des chaînes de Markov avec $\gamma = 1$ (précision du modèle durant l'évolution à long terme), le critère donné par l'équation (IV.17) est divergent. Par conséquent, le calcul de la commande optimale se fait en prenant en compte une espérance moyenne par période de décision qui aura comme équation d'optimalité [WHI, 88] :

$$w_n = \frac{\underset{U \in \Omega}{\text{opt}} \left\{ q(U) + \sum_{i \in I} p_{ij}(U) w_{n+1} \right\}}{n}, \quad n = 1, 2, 3, \dots \quad (\text{IV.18})$$

Dans la pratique industrielle on a souvent affaire à des situations dans lesquelles l'horizon est de nature aléatoire. Dans cette situation, l'horizon d'évolution du système n'est pas connu a priori, mais uniquement par l'intermédiaire d'une probabilité. Nous retrouvons ainsi une deuxième grande catégorie de problèmes terminaux : les problèmes avec horizon aléatoire.

b) Problèmes terminaux à horizon aléatoire

Dans cette classe de problèmes, l'instant final N est une variable aléatoire. Sa valeur est donnée par une densité de probabilité connue. De cette façon, chaque instant de l'évolution a une certaine probabilité connue d'être l'instant final.

A titre d'exemple, pour un horizon aléatoire correspondant à un processus de Poisson discret, on a une probabilité constante ζ d'avoir l'instant suivant $(n+1)$ comme instant final, caractérisé par un vecteur d'espérances de revenu w_{final} . Dans ce cas, l'utilisation du principe d'optimalité permet d'écrire :

$$w_n = \underset{U \in \Omega}{\text{opt}} \left\{ q(U) + (1 - \zeta) \sum_{i \in I} p_{ij}(U) w_{n+1} + \zeta \sum_{i \in I} p_{ij}(U) w_{\text{final}} \right\}, \quad n = 0, 1, 2, \dots \quad (\text{IV.19})$$

Le premier terme de l'équation d'optimalité correspond à l'espérance mathématique de revenu immédiat. Le deuxième est associé à l'espérance mathématique de revenu futur dans le cas où l'horizon n'est pas encore atteint. Le dernier terme de l'équation d'optimalité correspond à l'espérance mathématique de revenu si l'horizon est déjà atteint.

II.2.2. Problèmes avec cible

Nous appelons *problèmes avec cible*, les problèmes dans lesquels on cherche à atteindre un ensemble d'états du système en optimisant un critère de revenu.

Considérons une chaîne de Markov finie homogène et ergodique, ayant un nombre r d'états. Dans les états de la chaîne nous distinguons deux sous-ensembles, en général imposés d'avance :

- Le sous-ensemble I des états initiaux ($\text{card}(I) = M$),
- Le sous-ensemble C des états cibles ($\text{card}(C) = m$) avec $r = M + m$.

Le problème consiste à aller d'un état initial donné à l'un quelconque des états cibles en optimisant (en général - minimisant) les revenus (coûts) totaux.

Vu la partition de l'espace d'état, nous pouvons écrire la matrice de transition de la chaîne de Markov sous la forme suivante :

$$\mathbf{\dagger}(\mathbf{U}) = \begin{pmatrix} \dagger_I(\mathbf{U}) & \dagger_{IC}(\mathbf{U}) \\ \dagger_{CI}(\mathbf{U}) & \dagger_C(\mathbf{U}) \end{pmatrix} \quad (\text{IV.20})$$

avec $\dagger_I(\mathbf{U}) \in \mathbb{R}^{M \times M}$ la matrice de transition entre les états initiaux,
 $\dagger_C(\mathbf{U}) \in \mathbb{R}^{m \times m}$ la matrice de transition entre les états cibles,
 $\dagger_{IC}(\mathbf{U}) \in \mathbb{R}^{M \times m}$ et $\dagger_{CI}(\mathbf{U}) \in \mathbb{R}^{m \times M}$ les matrices de transition entre les états initiaux et les états cibles.

De même, la matrice des revenus s'écrit sous la forme partitionnée :

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_I & \mathbf{R}_{IC} \\ \mathbf{R}_{CI} & \mathbf{R}_C \end{pmatrix} \quad (\text{IV.21})$$

où $\mathbf{R}_I \in \mathbb{R}^{M \times M}$ représente les revenus de passage entre les états initiaux,
 $\mathbf{R}_C \in \mathbb{R}^{m \times m}$ les revenus de passage entre les états cibles,
 $\mathbf{R}_{IC} \in \mathbb{R}^{M \times m}$ et $\mathbf{R}_{CI} \in \mathbb{R}^{m \times M}$ les revenus correspondant aux transitions entre les états initiaux et les états cibles.

Soit $w_{ij}^*(n)$ l'espérance de revenu optimal à partir de l'instant n jusqu'à l'arrivée dans un état cible, sachant qu'à l'instant n le système se trouve dans l'état initial e_i ($e_i \in I$). Cette espérance mathématique est égale à la somme de deux éléments :

- l'espérance mathématique de gain immédiat r_{ij} si la transition vers l'état e_j a eu lieu ;
- l'espérance mathématique de gain relatif à l'état suivant $w_{ij}^*(n+1)$, si la transition vers e_j a eu lieu et si e_j n'est pas état cible ($e_j \notin C$).

Comme la probabilité d'apparition de la transition $e_i \rightarrow e_j$ est $p_{ij}(\mathbf{U})$, nous pouvons écrire :

$$\mathbf{w}_i^{\hat{O}}(n) = \underset{\mathbf{U} \in \Omega}{\text{opt}} \left(\sum_{j=1}^r p_{ij}(\mathbf{U}) r_{ij} + \sum_{j=1}^M p_{ij}(\mathbf{U}) \mathbf{w}_j^{\hat{O}}(n+1) \right) \quad \text{À } i = 1, 2, \dots, M. \quad (\text{IV.22})$$

En posant :

$$\mathbf{w}^{\hat{O}}(n) = \begin{bmatrix} \mathbf{w}_1^{\hat{O}}(n) \\ \mathbf{w}_2^{\hat{O}}(n) \\ \dots \\ \mathbf{w}_M^{\hat{O}}(n) \end{bmatrix} \quad (\text{IV.23})$$

pour le vecteur des espérances de revenus totaux optimaux perçus à partir de l'instant n jusqu'à l'arrivée dans un état cible, le critère d'optimalité prend la forme suivante :

$$\mathbf{w}^{\hat{O}}(n) = \underset{\mathbf{U} \in \Omega}{\text{opt}} \left\{ \mathbf{q}'(\mathbf{U}) + \dagger_I(\mathbf{U}) \mathbf{w}^{\hat{O}}(n+1) \right\} \quad n = 0, 1, 2, 3, \dots \quad (\text{IV.24})$$

où $\mathbf{q}'(\mathbf{U})$ représente le vecteur des espérances de revenu (coût) immédiat en partant des états initiaux et en utilisant la commande \mathbf{U} :

$$\mathbf{q}'(\mathbf{U}) = \begin{bmatrix} q_1(\mathbf{U}) \\ q_2(\mathbf{U}) \\ \vdots \\ q_M(\mathbf{U}) \end{bmatrix} = \text{diag} \left(\left[\dagger_I(\mathbf{U}) \quad \dagger_{IC}(\mathbf{U}) \right] \times \left[\mathbf{R}_I \quad \mathbf{R}_{IC} \right]^T \right) \quad (\text{IV.25})$$

avec :

$$q_i(\mathbf{U}) = \sum_{j=1}^r (p_{ij}(\mathbf{U}) \times r_{ij}) \quad \text{À } i = 1, \dots, M. \quad (\text{IV.26})$$

l'espérance mathématique de revenu immédiat perçu au cours d'une transition si le système se trouve dans l'état e_i ($i = 1, 2, \dots, M$) et si la commande appliquée est \mathbf{U} .

L'équation d'optimalité (IV.24) correspondant au problème avec cible est une équation implicite. Elle se résout par itération, à partir d'une solution initiale quelconque $\mathbf{w}^{\hat{O}}$. Si la condition :

$$\max_{i=1, \dots, M} \max_{\mathbf{U} \in \Omega} \left\{ \sum_{j=1}^M p_{ij}(\mathbf{U}) \right\} < 1 \quad (\text{IV.27})$$

correspondant à la contraction de l'opérateur $\text{opt}[\cdot]$ de l'équation (IV.24) est respectée [BOU, 69], l'algorithme converge vers une valeur unique de l'espérance de revenu et permet le calcul de la commande optimale $\mathbf{U}_i^{\hat{O}}(n)$. Par conséquent, la stratégie de commande optimale sera donnée

par une matrice de dimension $(K \times M)$:

$$\mathbf{G}(n) = [\hat{\mathbf{U}}(n,1) \ \hat{\mathbf{U}}(n,2) \ \dots \ \hat{\mathbf{U}}(n,M)] , \ \text{À } n = 0, 1, \dots \quad (\text{IV.28})$$

où \mathbf{g} est la fonction de commande.

Nous observons que la condition (IV.27) est équivalente à :

$$\max_{\mathbf{U} \in \Omega} \left\| \dot{\mathbf{T}}_1(\mathbf{U}) \right\| < 1 \quad (\text{IV.29})$$

donc il doit y avoir au moins une transition directe de chaque état initial vers un des états cibles, quelle que soit la commande appliquée. Cette condition est vérifiée dans le cas des chaînes de Markov irréductibles. Par conséquent, ce type de systèmes se caractérise par la convergence du critère d'optimalité (IV.24).

Les équations d'optimalité des chaînes de Markov sont des équations vectorielles récurrentes qui demandent l'emploi de moyens informatiques considérables. Pour des systèmes de grande taille, comme c'est le cas pour la majorité des chaînes de Markov à commande, ces équations récurrentes deviennent vite intraitables. Par conséquent, la résolution des problèmes réels de commande optimale des chaînes de Markov impose l'utilisation de méthodes de simplification.

Dans les chapitres précédents nous avons étudié le découplage des chaînes de Markov non commandées à double échelle de temps respectivement double échelle de pondération. Le sens des propriétés de double échelle de temps et double échelle de pondération reste le même dans le cas des chaînes de Markov à commande. L'apparition du vecteur de commande dans les chaînes de Markov, introduit une non-linéarité dans le modèle et exige le développement de techniques de simplification plus élaborées.

III. Application des perturbations singulières pour la simplification des chaînes de Markov à commande

III.1. Systèmes étudiés

Nous nous proposons d'étudier les chaînes de Markov homogènes ergodiques à commande, caractérisées par l'équation fondamentale suivante :

$$\mathbf{P}(n+1) = \mathbf{P}(n) \dot{\mathbf{T}}(\mathbf{U}) \quad (\text{IV.30})$$

avec $\mathbf{P}(n) \in \mathbb{R}^{1 \times r}$ la distribution des probabilités à l'instant n ($n = 0, 1, 2, \dots$),

$\dot{\mathbf{T}}(\mathbf{U}) = [\dot{p}_{ij}(\mathbf{U})]$, la matrice de transition de la forme :

$$\dot{\mathbf{T}}(\mathbf{U}) = \dot{\mathbf{T}}_0 + \dot{\mathbf{T}}_1 \cdot u_1 + \dot{\mathbf{T}}_2 \cdot u_2 + \dots + \dot{\mathbf{T}}_K \cdot u_K \quad (\text{IV.31})$$

et la commande $\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_K \end{bmatrix} \in \Omega$ borné,

Les matrices \dagger_k sont de matrices constantes de type :

$$\dagger_k = \begin{pmatrix} p_{11}^k & p_{12}^k & \dots & p_{1r}^k \\ p_{21}^k & p_{22}^k & \dots & p_{2r}^k \\ \dots & \dots & \dots & \dots \\ p_{r1}^k & p_{r2}^k & \dots & p_{rr}^k \end{pmatrix}, \quad k = 0, 1, \dots, K,$$

avec $p_{ij}(\mathbf{U}) = p_{ij}^0 + p_{ij}^1 u_1 + p_{ij}^2 u_2 + \dots + p_{ij}^K u_K$, $\forall i, j = 1, 2, \dots, r$, $\forall \mathbf{U} = \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_K \end{pmatrix} \in \Omega$

Dans l'expression (IV.31) de la matrice de transition $\dagger(\mathbf{U})$, la matrice \dagger_0 apparaît comme une composante constante, indépendante de la commande.

La stochasticité de la matrice de transition $\dagger(\mathbf{U})$, induit les propriétés suivantes sur les matrices \dagger_k :

$$\sum_{j=1}^r p_{ij}^0 = 1, \quad \forall i = 1, 2, \dots, r,$$

$$\sum_{j=1}^r p_{ij}^k = 0, \quad \forall i = 1, 2, \dots, r \text{ et } \forall k = 1, 2, \dots, K.$$

Remarque

La matrice de transition (IV.31) peut être écrite sous la forme contractée suivante :

$$\dagger(\mathbf{U}) = \dagger_0 + \sum_{k=1}^K \dagger_k \cdot u_k$$

qui introduite dans la formule fondamentale (IV.30) donne :

$$\mathbf{P}(n+1) = \mathbf{P}(n) \left(\dagger_0 + \sum_{k=1}^K \dagger_k \cdot u_k \right) \quad (\text{IV.32})$$

Nous reconnaissons ici une forme bilinéaire (affines en l'état et en la commande). Il s'agit plus précisément de la classe des systèmes bilinéaires *réguliers* [BRU, 74].

□

Les systèmes bilinéaires représentent une classe de systèmes permettant l'extension de certaines méthodes propres aux systèmes linéaires, tout en restant des approches viables des systèmes non-linéaires. Ces systèmes ont la propriété d'être suffisamment adaptatifs, grâce à la présence de produits de variables d'état par des variables d'entrée. Ceci permet de modéliser avec

précision un nombre important de systèmes réels.

Plusieurs auteurs, comme Bruni, Dippilo, Koch [BRU, 74] et Mohler [MOH, 73] [MOH, 80] ont étudié les systèmes bilinéaires déterministes. Des méthodes de simplification de ces systèmes ont ainsi été développées dans le cas continu [ROT, 76] et une définition de la double échelle de temps a été donnée.

Nous proposons une extension de ces travaux dans le cas des chaînes de Markov bilinéaires à commande.

III.2. Définition de la double échelle de temps des chaînes de Markov à commande

Définition

Une chaîne de Markov homogène ergodique à commande, d'équation fondamentale :

$$\mathbf{P}(n+1) = \mathbf{P}(n) \dot{\mathbf{T}}(\mathbf{U}), \quad n = 0, 1, 2, 3, \dots,$$

$$\mathbf{P}(n) \text{ ' } \hat{\mathbf{e}}^{1 \times r} \text{ de conditions initiales } \mathbf{P}(0),$$

$$\text{et } \mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_K \end{bmatrix} \text{ ' } \Omega \text{ borné,}$$

possède la propriété de double échelle de temps sur Ω si le spectre de la matrice $\dot{\mathbf{T}}(\mathbf{U})$ peut être partitionné en deux sous-ensembles disjoints $\mathcal{L}(\mathbf{U})$ et $\mathcal{R}(\mathbf{U})$ tels que :

$$\forall \mathbf{U} \text{ ' } \Omega, \quad \min_{\lambda_i(\mathbf{U}) \text{ ' } \mathcal{L}(\mathbf{U})} |\lambda_i(\mathbf{U})| = \dot{\lambda}(\mathbf{U}) \gg r\dot{\lambda}(\mathbf{U}) = \max_{\lambda_j(\mathbf{U}) \text{ ' } \mathcal{R}(\mathbf{U})} |\lambda_j(\mathbf{U})|$$

□

Une indication de la séparation des échelles de temps peut être donnée par un coefficient μ tel que :

$$\mu = \max_{\mathbf{U} \text{ ' } \Omega} \left(\frac{r\dot{\lambda}(\mathbf{U})}{\dot{\lambda}(\mathbf{U})} \right) \ll 1$$

Dans le cas de l'existence de la propriété de double échelle de temps, le système peut être écrit sous la forme partitionnée suivante :

$$[\mathbf{L}(n+1) \quad \mathbf{R}(n+1)] = [\mathbf{L}(n) \quad \mathbf{R}(n)] \begin{pmatrix} \dot{\mathbf{T}}_{11}(\mathbf{U}) & \dot{\mathbf{T}}_{12}(\mathbf{U}) \\ \dot{\mathbf{T}}_{21}(\mathbf{U}) & \dot{\mathbf{T}}_{22}(\mathbf{U}) \end{pmatrix}, \quad \forall \mathbf{U} \text{ ' } \Omega \quad (\text{IV.33})$$

avec :

$\mathbf{P}(n) = [\mathbf{P}_1(n) \quad \mathbf{P}_2(n) \quad \mathbf{P}_3(n) \dots \mathbf{P}_r(n)] = [\mathbf{L}(n) \quad \mathbf{R}(n)]$ les probabilités d'état du système initial,

$$\dot{\mathbf{T}}_{11}(\mathbf{U}) \text{ ' } \hat{\mathbf{e}}^{r_1 \times r_1}, \dot{\mathbf{T}}_{12}(\mathbf{U}) \text{ ' } \hat{\mathbf{e}}^{r_1 \times r_2}, \dot{\mathbf{T}}_{21}(\mathbf{U}) \text{ ' } \hat{\mathbf{e}}^{r_2 \times r_1}, \dot{\mathbf{T}}_{22}(\mathbf{U}) \text{ ' } \hat{\mathbf{e}}^{r_2 \times r_2}, \quad r_1 + r_2 = r,$$

et :

$\mathbf{L}(n) = [P_1(n) \ P_2(n) \ \dots \ P_r(n)]$ le vecteur de probabilités d'état de la partie lente

$\mathbf{R}(n) = [P_{r+1}(n) \ P_{r+2}(n) \ \dots \ P_r(n)]$ la distribution des probabilités de la partie rapide.

Les conditions initiales du système partitionné sont : $\mathbf{P}(0) = [\mathbf{L}(0) \ \mathbf{R}(0)]$.

Les sous-matrices de (IV.33) s'écrivent :

$$\begin{aligned} \dot{\mathbf{T}}_{11}(\mathbf{U}) &= \dot{\mathbf{T}}_{011} + \sum_{k=1}^K \dot{\mathbf{T}}_{k11} u_k, \quad \dot{\mathbf{T}}_{12}(\mathbf{U}) = \dot{\mathbf{T}}_{012} + \sum_{k=1}^K \dot{\mathbf{T}}_{k12} u_k, \\ \dot{\mathbf{T}}_{21}(\mathbf{U}) &= \dot{\mathbf{T}}_{021} + \sum_{k=1}^K \dot{\mathbf{T}}_{k21} u_k, \quad \dot{\mathbf{T}}_{22}(\mathbf{U}) = \dot{\mathbf{T}}_{022} + \sum_{k=1}^K \dot{\mathbf{T}}_{k22} u_k \end{aligned}$$

où :

$$\dot{\mathbf{T}}_k = \begin{pmatrix} \dot{\mathbf{T}}_{k11} & \dot{\mathbf{T}}_{k12} \\ \dot{\mathbf{T}}_{k21} & \dot{\mathbf{T}}_{k22} \end{pmatrix} \quad \text{pour } k = 0, 1, \dots, K.$$

avec $\dot{\mathbf{T}}_{k11} \sim \hat{\mathbf{e}}^{r_1 x_1}, \dot{\mathbf{T}}_{k12} \sim \hat{\mathbf{e}}^{r_1 x_2}, \dot{\mathbf{T}}_{k21} \sim \hat{\mathbf{e}}^{r_2 x_1}, \dot{\mathbf{T}}_{k22} \sim \hat{\mathbf{e}}^{r_2 x_2}$, À $k = 0, 1, \dots, K$

Parmi les méthodes de perturbation singulières développées dans le cas linéaire, le découplage de Bennis s'avère intéressant. Il permet d'avoir un sous-système lent découplé stochastique. Nous nous proposons de généraliser cette technique pour le découplage des chaînes de Markov à commande.

III.3. Découplage des dynamiques

Soit la chaîne de Markov à commande sous la forme partitionnée (IV.33) :

$$\begin{cases} \mathbf{L}(n+1) = \mathbf{L}(n) \dot{\mathbf{T}}_{11}(\mathbf{U}) + \mathbf{R}(n) \dot{\mathbf{T}}_{21}(\mathbf{U}) \\ \mathbf{R}(n+1) = \mathbf{L}(n) \dot{\mathbf{T}}_{12}(\mathbf{U}) + \mathbf{R}(n) \dot{\mathbf{T}}_{22}(\mathbf{U}) \end{cases} \quad (\text{IV.34})$$

- partie lente -

Le principe du découplage consiste à considérer que les probabilités des états de la partie rapide de la chaîne de Markov tendent rapidement vers leur régime permanent. Dans ce cas, nous pouvons écrire :

$$\mathbf{R}(n+1) = \mathbf{R}(n) = \mathbf{R}_1(n),$$

Lorsque nous reportons cette relation dans (IV.34) nous retrouvons le système lent découplé suivant :

$$\mathbf{L}_1(n+1) = \mathbf{L}_1(n) \left(\dot{\mathbf{T}}_{11}(\mathbf{U}_1) + \dot{\mathbf{T}}_{12}(\mathbf{U}_1) \left(\mathbf{I} - \dot{\mathbf{T}}_{22}(\mathbf{U}_1) \right)^{-1} \dot{\mathbf{T}}_{21}(\mathbf{U}_1) \right), \mathbf{U}_1 \sim \Omega \quad (\text{IV.35})$$

Nous notons cette matrice d'évolution par $\dot{\mathbf{T}}_1$.

La composante lente de la distribution rapide $\mathbf{R}_1(n)$ peut être calculée par l'expression

ci-dessous :

$$\mathbf{R}_1(n) = \mathbf{L}_1(n) \mathbin{\dot{\nabla}}_{12}(\mathbf{U}_1) \left(\mathbf{I} - \mathbin{\dot{\nabla}}_{22}(\mathbf{U}_1) \right)^{-1} \quad (\text{IV.36})$$

dont les conditions initiales sont données par les relations :

$$\mathbf{L}_1(0) = \mathbf{L}(0) \text{ respectivement } \mathbf{R}_1(0) = \mathbf{L}(0) \mathbin{\dot{\nabla}}_{12}(\mathbf{U}_1) \left(\mathbf{I} - \mathbin{\dot{\nabla}}_{22}(\mathbf{U}_1) \right)^{-1}$$

- partie rapide -

L'équation qui décrit le sous-système rapide est obtenu en définissant la probabilité d'état $\mathbf{R}_r(n)$ de la façon suivante :

$$\mathbf{R}_r(n) = \mathbf{R}(n) - \mathbf{R}_1(n) \quad (\text{IV.37})$$

avec $\mathbf{R}_1(n)$ donnée par la relation (IV.36).

En introduisant (IV.37) dans le système initial (IV.34), nous obtenons le système rapide découplé suivant :

$$\mathbf{R}_r(n+1) = \mathbf{R}_r(n) \mathbin{\dot{\nabla}}_{22}(\mathbf{U}_r), \mathbf{U}_r \in \Omega \quad (\text{IV.38})$$

La matrice d'évolution est notée $\mathbin{\dot{\nabla}}_r$ et les conditions initiales sont :

$$\mathbf{R}_r(0) = \mathbf{R}(0) - \mathbf{L}(0) \mathbin{\dot{\nabla}}_{12}(\mathbf{U}_r) \left(\mathbf{I} - \mathbin{\dot{\nabla}}_{22}(\mathbf{U}_r) \right)^{-1}$$

III.4. Propriétés des matrices $\mathbin{\dot{\nabla}}_1$ et $\mathbin{\dot{\nabla}}_r$

a) Matrice $\mathbin{\dot{\nabla}}_1$:

1°. La matrice $\mathbin{\dot{\nabla}}_1$ relative au système lent découplé est une matrice stochastique.

- Démonstration -

Nous notons par $\mathbf{1}_{r_1}$ (respectivement $\mathbf{1}_{r_2}$) le vecteur à éléments égaux à 1, de dimension r_1 (respectivement r_2) :

$$\mathbf{1}_{r_1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{r_1 \times 1} \quad \text{et} \quad \mathbf{1}_{r_2} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{r_2 \times 1} \quad (\text{IV.39})$$

avec r_1 et r_2 le nombre d'états lents respectivement rapides de la chaîne de Markov.

- La somme de chaque ligne de la matrice $\mathbin{\dot{\nabla}}_1$ est égale à 1

La matrice de transition initiale $\mathbin{\dot{\nabla}}(\mathbf{U})$ étant stochastique, nous pouvons écrire, en gardant les notations (IV.33), la relation suivante :

$$\mathbf{1}_{r_2} - \ddagger_{22}(\mathbf{U}) \mathbf{1}_{r_2} = \ddagger_{21}(\mathbf{U}) \mathbf{1}_{r_1}, \quad \forall \mathbf{U} \in \Omega$$

ou sous une autre forme :

$$\mathbf{1}_{r_2} = \left(\mathbf{I} - \ddagger_{22}(\mathbf{U}) \right)^{-1} \ddagger_{21}(\mathbf{U}) \mathbf{1}_{r_1}, \quad \forall \mathbf{U} \in \Omega$$

En multipliant à gauche par $\ddagger_{12}(\mathbf{U})$ il nous vient :

$$\ddagger_{12}(\mathbf{U}) \mathbf{1}_{r_2} = \ddagger_{12}(\mathbf{U}) \left(\mathbf{I} - \ddagger_{22}(\mathbf{U}) \right)^{-1} \ddagger_{21}(\mathbf{U}) \mathbf{1}_{r_1}, \quad \forall \mathbf{U} \in \Omega$$

En rajoutant $\mathbf{1}_{r_1}$ aux deux membres de l'égalité, nous obtenons :

$$\mathbf{1}_{r_1} - \ddagger_{12}(\mathbf{U}) \mathbf{1}_{r_2} + \ddagger_{12}(\mathbf{U}) \left(\mathbf{I} - \ddagger_{22}(\mathbf{U}) \right)^{-1} \ddagger_{21}(\mathbf{U}) \mathbf{1}_{r_1} = \mathbf{1}_{r_1}, \quad \forall \mathbf{U} \in \Omega$$

La propriété de stochasticité de la matrice initiale nous permet d'écrire :

$$\ddagger_{11}(\mathbf{U}) \mathbf{1}_{r_1} + \ddagger_{12}(\mathbf{U}) \left(\mathbf{I} - \ddagger_{22}(\mathbf{U}) \right)^{-1} \ddagger_{21}(\mathbf{U}) \mathbf{1}_{r_1} = \mathbf{1}_{r_1}, \quad \forall \mathbf{U} \in \Omega$$

donc :

$$\ddagger_{11}(\mathbf{U}) \mathbf{1}_{r_1} = \mathbf{1}_{r_1}, \quad \forall \mathbf{U} \in \Omega$$

Vu que $\mathbf{U}_1 \in \Omega$, nous avons :

$$\ddagger_{11}(\mathbf{U}_1) \mathbf{1}_{r_1} = \mathbf{1}_{r_1}, \quad \forall \mathbf{U}_1 \in \Omega \quad (\text{IV.40})$$

La somme des éléments de chacune des lignes de $\ddagger_1(\mathbf{U}_1)$ est égale à 1.

- Tous les éléments de la matrice \ddagger_1 sont non-négatifs.

La norme de $\ddagger_{22}(\mathbf{U})$ est plus petite que 1 quel que soit \mathbf{U} dans Ω . Cette propriété nous permet d'effectuer un développement en série de la matrice $(\mathbf{I} - \ddagger_{22}(\mathbf{U}))^{-1}$. Nous obtenons par conséquent :

$$\left(\mathbf{I} - \ddagger_{22}(\mathbf{U}) \right)^{-1} = \mathbf{I} + \ddagger_{22}(\mathbf{U}) + \left[\ddagger_{22}(\mathbf{U}) \right]^2 + \left[\ddagger_{22}(\mathbf{U}) \right]^3 + \dots = \sum_{p=0} \left[\ddagger_{22}(\mathbf{U}) \right]^p, \quad \forall \mathbf{U} \in \Omega$$

Comme $\ddagger_{22}(\mathbf{U})$ est une matrice non-négative
Erreur !

$$\left(\mathbf{I} - \ddagger_{22}(\mathbf{U}) \right)^{-1} \geq 0, \quad \forall \mathbf{U} \in \Omega$$

Vu que les matrices $\ddagger_{11}(\mathbf{U})$, $\ddagger_{12}(\mathbf{U})$ et $\ddagger_{21}(\mathbf{U})$ sont non-négatives quel que soit \mathbf{U} dans Ω , il en sera de même pour $\ddagger_1(\mathbf{U})$:

$$\dot{\mathbf{x}}_1(\mathbf{U}) = \dot{\mathbf{x}}_{11}(\mathbf{U}) + \dot{\mathbf{x}}_{12}(\mathbf{U}) \left(\mathbf{I} - \dot{\mathbf{x}}_{22}(\mathbf{U}) \right)^{-1} \dot{\mathbf{x}}_{21}(\mathbf{U}) \quad 0, \quad \forall \mathbf{U} \in \Omega$$

et comme $\mathbf{U}_1 \in \Omega$, nous obtenons :

$$\dot{\mathbf{x}}_1(\mathbf{U}_1) = \dot{\mathbf{x}}_{11}(\mathbf{U}_1) + \dot{\mathbf{x}}_{12}(\mathbf{U}_1) \left(\mathbf{I} - \dot{\mathbf{x}}_{22}(\mathbf{U}_1) \right)^{-1} \dot{\mathbf{x}}_{21}(\mathbf{U}_1) \quad 0, \quad \forall \mathbf{U}_1 \in \Omega$$

En conclusion, le système lent discret obtenu par application du découplage Bennis sur une chaîne de Markov à commande représente une chaîne de Markov d'ordre réduit (r_1 états).

□

2°. Remarque

La matrice de transition de la partie lente ne conserve pas la forme bilinéaire du système initial :

$$\dot{\mathbf{x}}_1(\mathbf{U}) = \dot{\mathbf{x}}_{011} + \sum_{k=1}^K \left(\dot{\mathbf{x}}_{k11} u_k \right) + \left[\dot{\mathbf{x}}_{012} + \sum_{k=1}^K \left(\dot{\mathbf{x}}_{k12} u_k \right) \right] \left[\mathbf{I} - \left(\dot{\mathbf{x}}_{022} + \sum_{k=1}^K \left(\dot{\mathbf{x}}_{k22} u_k \right) \right) \right]^{-1} \left[\dot{\mathbf{x}}_{021} + \sum_{k=1}^K \left(\dot{\mathbf{x}}_{k21} u_k \right) \right]$$

b) Matrice :

1°. La matrice $\dot{\mathbf{x}}_r$ n'est pas une matrice stochastique

2°. La matrice de transition de la partie rapide est affine en la commande

$$\dot{\mathbf{x}}_r(\mathbf{U}) = \dot{\mathbf{x}}_{022} + \sum_{k=1}^K \left(\dot{\mathbf{x}}_{k22} u_k \right)$$

III.5. Bilinearisation du sous-système lent

Nous introduisons une matrice constante \mathbf{G} vérifiant la relation (IV.41) $\forall \mathbf{U} \in \Omega$:

$$\mathbf{G} = \left[\mathbf{I} - \left(\dot{\mathbf{x}}_{022} + \sum_{k=1}^K \left(\dot{\mathbf{x}}_{k22} u_k \right) \right) \right]^{-1} \left[\dot{\mathbf{x}}_{021} + \sum_{k=1}^K \left(\dot{\mathbf{x}}_{k21} u_k \right) \right], \quad \mathbf{G} \in \mathbb{R}^{r_2 \times r_1} \quad (\text{IV.41})$$

Comme l'égalité (IV.41) doit être vérifiée quelle que soit la commande dans Ω , la matrice \mathbf{G} représente la solution de l'équation matricielle :

$$\begin{bmatrix} \dot{\dagger}_{0_{21}} \\ \dot{\dagger}_{1_{21}} \\ \dots \\ \dot{\dagger}_{K_{21}} \end{bmatrix} = \begin{bmatrix} \left(I - \dot{\dagger}_{0_{22}} \right) \\ - \dot{\dagger}_{1_{22}} \\ \dots \\ - \dot{\dagger}_{K_{22}} \end{bmatrix} \times \mathbf{G}$$

La solution de cette équation peut être obtenue en calculant la pseudo-inverse au sens de Moore-Penrose de la matrice de dimension $(K r_2 \times r_2)$ qui multiplie l'inconnue \mathbf{G} . Ce calcul nous donne le résultat suivant :

$$\mathbf{G} = \left[\left(I - \dot{\dagger}_{0_{22}} \right)^T \left(I - \dot{\dagger}_{0_{22}} \right) + \sum_{k=1}^K \dot{\dagger}_{k_{22}}^T \dot{\dagger}_{k_{22}} \right]^{-1} \left[\left(I - \dot{\dagger}_{0_{22}} \right)^T \dot{\dagger}_{0_{21}} - \sum_{k=1}^K \dot{\dagger}_{k_{22}}^T \dot{\dagger}_{k_{21}} \right]$$

La forme découplée de l'équation fondamentale de la chaîne ergodique devient :

$$\begin{cases} \begin{bmatrix} \mathbf{L}_1(n+1) & \mathbf{R}_r(n+1) \end{bmatrix} = \begin{bmatrix} \mathbf{L}_1(n) & \mathbf{R}_r(n) \end{bmatrix} \begin{pmatrix} \dot{\dagger}_1(\mathbf{U}_1) & 0 \\ 0 & \dot{\dagger}_r(\mathbf{U}_r) \end{pmatrix} \\ \mathbf{R}_l(n) = \mathbf{L}_1(n) \dot{\dagger}_{\text{infl}}(\mathbf{U}_1) \end{cases}$$

avec :

$$\dot{\dagger}_1(\mathbf{U}_1) = \dot{\dagger}_{11}(\mathbf{U}_1) + \dot{\dagger}_{12}(\mathbf{U}_1) \mathbf{G}$$

$$\dot{\dagger}_r(\mathbf{U}_r) = \dot{\dagger}_{22}(\mathbf{U}_r)$$

$$\dot{\dagger}_{\text{infl}}(\mathbf{U}_1) = \dot{\dagger}_{12}(\mathbf{U}_1) \left(I - \dot{\dagger}_{22}(\mathbf{U}_1) \right)^{-1}$$

De cette façon, les matrices de transition des sous-systèmes découplés sont affines en la commande. Les équations des systèmes lent et rapide seront ainsi bilinéaires. En plus, vu que la matrice constante \mathbf{G} vérifie :

$$\mathbf{G} = \left(I - \dot{\dagger}_{22}(\mathbf{U}_1) \right)^{-1} \dot{\dagger}_{21}(\mathbf{U}_1), \text{ À } \mathbf{U}_1 \in \Omega,$$

la matrice $\dot{\dagger}_1$ du système lent reste stochastique.

Une étape importante dans la simplification des chaînes de Markov par les perturbations singulières concerne la mise en évidence des dynamiques.

III.6. Aspects de la mise en évidence des dynamiques des chaînes de Markov bilinéaires

La mise en évidence des dynamiques des systèmes bilinéaires peut être effectuée par des méthodes analytiques et des méthodes géométriques. Nous développons ici la méthode

géométrique basée sur le tracé des disques de Gershgorine.

Dans le cas des chaînes de Markov bilinéaires, la localisation des valeurs propres à l'aide des méthodes géométriques peut être obtenue de deux façons différentes :

- soit au niveau global, pour détecter des groupements de probabilités d'état lentes et de probabilités d'état rapides, valables sur tout l'intervalle de variation de la commande.
- soit au niveau local, pour déterminer les variations admissibles sur les variables de commande assurant une certaine séparation des dynamiques des probabilités d'état.

III.6.1. Niveau global

La mise en évidence globale des dynamiques d'une chaîne de Markov à commande de type bilinéaire constitue une application des techniques de détection des dynamiques sur des matrices constantes. Dans ce sens, en généralisant les résultats du continu [ROT, 85], nous pouvons énoncer le théorème analogue suivant :

Théorème 5

Soit la chaîne de Markov finie homogène ergodique à commande, caractérisée par l'équation fondamentale suivante :

$$\mathbf{P}(n+1) = \mathbf{P}(n) \left(\dagger_0 + \sum_{k=1}^K \dagger_k u_k \right) \quad (\text{IV.42})$$

avec $\mathbf{P}(n) \in \mathbb{R}^{1 \times r}$ à $n = 0, 1, 2, \dots$; $\dagger_k \in \mathbb{R}^{r \times r}$ à $k = 1, \dots, K$; $\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_K \end{bmatrix} \in \Omega \subset \mathbb{R}^K$.

Si pour toutes les matrices \dagger_k le tracé des disques de Gershgorine indique une séparation en deux dynamiques et que les variables lentes et les variables rapides sont dans tous les cas les mêmes, alors la chaîne de Markov à commande est à double échelle de temps sur tout l'espace de commande Ω .

- Démonstration -

La démonstration basée sur une simple addition des relations données par la séparation des disques de Gershgorine correspondant à chaque matrice constante \dagger_k (voir l'annexe 4).

□

III.6.2. Niveau local

a) Disques de Gershgorine dynamiques

L'utilisation du théorème de Gershgorine sur le système (IV.33) donne une localisation des valeurs propres pour une commande $\mathbf{U} \in \Omega$ fixée. Une étude complète de la séparation des dynamiques de la chaîne de Markov impose l'utilisation des disques de Gershgorine de façon dynamique. Nous obtenons ainsi une localisation des valeurs propres $\lambda_i(\mathbf{U})$ à l'intérieur d'un ensemble de r disques définis par :

$$|x - p_{ii}(U)| \leq (R_i(U))^\alpha (R'_i(U))^{1-\alpha}, \quad i = 1, \dots, r, \quad \alpha \in [0, 1] \text{ et } U \in \Omega \quad (\text{IV.43})$$

avec :

$$R_i(U) = \sum_{j=1, j \neq i}^r p_{ij}(U) \quad (\text{IV.44})$$

$$R'_i(U) = \sum_{j=1, j \neq i}^r p_{ji}(U) \quad (\text{IV.45})$$

et nous obtenons le théorème qui suit :

Théorème 6

Si les disques de Gershgorine associés à $\dot{T}(U)$ vérifient l'inégalité suivante :

$$\begin{aligned} &\exists I_i, I_j, \quad I_i \cap I_j = \emptyset, \quad I_i \cup I_j = \{1, \dots, r\} \\ &\exists \alpha \in [0, 1] \end{aligned}$$

tels que

$$|p_{ii}(U) - p_{jj}(U)| \gg \frac{(R_i(U))^\alpha (R'_i(U))^{1-\alpha} + (R_j(U))^\alpha (R'_j(U))^{1-\alpha}}{\forall i \in I_i, \forall j \in I_j, \forall U \in \Omega} \quad (\text{IV.46})$$

alors dans Ω , la chaîne de Markov à commande possède la double échelle de temps.

Une indication sur la séparation des dynamiques est alors donnée par le paramètre :

$$\mu = \sup_{U \in \Omega} \min_{i \in I_i, j \in I_j} \frac{(R_i(U))^\alpha (R'_i(U))^{1-\alpha} + (R_j(U))^\alpha (R'_j(U))^{1-\alpha}}{|p_{ii}(U) - p_{jj}(U)|}$$

qui est inférieur à 1 s'il y a deux domaines distincts de localisation de valeurs propres.

□

Pour des raisons de volume de calcul, la vérification pratique de l'inégalité (IV.46) demande l'introduction d'un échantillonnage de l'espace de commande. Dans ce sens, nous pouvons utiliser un pas δU inversement proportionnel au conditionnement de la matrice $\dot{T}(U)$:

$$\delta U = \frac{1}{\text{cond}\{\dot{T}(U)\}} = \frac{1}{\|\dot{T}(U)\| \cdot \|\dot{T}^{-1}(U)\|} \quad (\text{IV.47})$$

En effet, le conditionnement d'une matrice exprime la sensibilité du spectre du système en fonction des variations de ses éléments [STO, 80]. Comme dans notre cas les probabilités de transition varient en fonction de la commande, le conditionnement de la matrice exprime la correspondance entre la variation de la commande et la variation induite au niveau des valeurs propres du système.

b) Méthode de conditionnement [CIA, 82]

La méthode des disques de Gershgorine dynamiques fait appel à la comparaison entre des disques dont les centres et les rayons sont fonctions affines de la commande. Ceci peut être contourné en utilisant une méthode de conditionnement qui permet de fixer les centres des disques de Gershgorine [CIA, 82].

Théorème 7 [CIA, 82]

Soit $\dot{\mathbf{T}}$ une matrice carrée $r \times r$, \mathbf{J} la forme Jordan associée et \mathbf{P} une matrice régulière telles que :

$$\mathbf{P}^{-1} \dot{\mathbf{T}} \mathbf{P} = \mathbf{J} = \mathbf{D} + \mathbf{H} \text{ où } \mathbf{D} = \text{Erreur !} , \lambda_i \in \text{Sp}(\dot{\mathbf{T}})$$

Soit $\|\cdot\|$ une norme matricielle vérifiant la condition :

$$\|\text{diag}\{d_i\}\| = \max_i |d_i|$$

Alors pour toute matrice $\Delta \dot{\mathbf{T}}$,

$$\text{Sp}(\dot{\mathbf{T}} + \Delta \dot{\mathbf{T}}) \subset \bigcup_{i=1}^r D_i \text{ où } D_i = \{z \in \mathbb{C} ; |z - \lambda_i| < \varepsilon + \text{cond}(\mathbf{P}) \|\Delta \dot{\mathbf{T}}\|\}$$

avec $\text{cond}(\mathbf{P}) = \|\mathbf{P}\| \cdot \|\mathbf{P}^{-1}\|$ et $\varepsilon = 0$ si $\dot{\mathbf{T}}$ est diagonalisable, $\varepsilon = 1$ sinon.

□

L'application de la méthode de conditionnement dans le cas des chaînes de Markov bilinéaires, présente les étapes suivantes :

i) sélectionner parmi les matrices constantes déterminant le système bilinéaire (IV.42), une matrice particulière $\dot{\mathbf{T}}_k$, ayant des valeurs propres très différentes ;

ii) déterminer l'ensemble des valeurs et vecteurs propres donc $\text{Sp}(\dot{\mathbf{T}}_k)$ et \mathbf{P}_k tels que

$$\mathbf{P}_k^{-1} \dot{\mathbf{T}}_k \mathbf{P}_k = \mathbf{J}_k$$

iii) appliquer le théorème 8 sur $\dot{\mathbf{T}}_k$ avec la matrice de perturbation :

$$\Delta \dot{\mathbf{T}}_k = \mathbf{u}_k^{-1} \sum_{i=0, i \neq k}^K \dot{\mathbf{T}}_i \mathbf{u}_i \text{ avec } \mathbf{u}_0 = 1.$$

Le conditionnement proposé confère à la nouvelle matrice, une diagonale principale constante (égale à celle de $\dot{\mathbf{T}}_k$). Seuls les rayons des disques de Gershgorine correspondants seront variables en fonction de la commande.

c) Utilisation de la matrice majorante

Dans le cas des chaînes de Markov à commande, il est intéressant d'appliquer la relation (IV.46) pour $\alpha = 0$, correspondant aux disques de Gershgorine sur les colonnes. Pour éviter de construire les disques de Gershgorine dynamiques dans cette situation, le théorème suivant nous permet de travailler sur des éléments constants :

Théorème 8

Soit la chaîne de Markov homogène finie à commande, présentant l'équation fondamentale suivante

$$\mathbf{P}(n+1) = \mathbf{P}(n) \mathbf{\ddagger}(\mathbf{U})$$

$$\text{avec } \mathbf{P}(n) \in \mathbb{R}^{1 \times r} \text{ } \forall n = 0, 1, 2, \dots; \mathbf{\ddagger}(\mathbf{U}) = [p_{ij}] \in \mathbb{R}^{r \times r} \text{ } \forall i, j = 1, \dots, r, \text{ } \forall \mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_K \end{bmatrix} \in \Omega.$$

Si les probabilités de transition vérifient l'inégalité :

$$\max_{\mathbf{U} \in \Omega} \{p_{ii}(\mathbf{U})\} + \sum_{k=1, k \neq i}^r \max_{\mathbf{U} \in \Omega} \{p_{ki}(\mathbf{U})\} << \min_{\mathbf{U} \in \Omega} \{p_{jj}(\mathbf{U})\} - \sum_{k=1, k \neq j}^r \max_{\mathbf{U} \in \Omega} \{p_{kj}(\mathbf{U})\} \quad (\text{IV.48})$$

$$\forall i \in I_i, \forall j \in I_j, \forall \mathbf{U} \in \Omega, \text{ avec } I_i \cap I_j = \emptyset, I_i \cup I_j = \{1, \dots, r\},$$

alors la chaîne de Markov à commande présente la propriété de double échelle de temps sur Ω .

- Démonstration -

En majorant les éléments de la colonne $i \in I_i$ de la matrice de transition $\mathbf{\ddagger}(\mathbf{U})$, nous obtenons l'inégalité suivante :

$$p_{ii}(\mathbf{U}) + \sum_{k=1, k \neq i}^r p_{ki}(\mathbf{U}) \leq \max_{\mathbf{U} \in \Omega} \{p_{ii}(\mathbf{U})\} + \sum_{k=1, k \neq i}^r \max_{\mathbf{U} \in \Omega} \{p_{ki}(\mathbf{U})\}$$

En utilisant par la suite l'inégalité (IV.48), nous obtenons :

$$p_{ii}(\mathbf{U}) + \sum_{k=1, k \neq i}^r p_{ki}(\mathbf{U}) << \min_{\mathbf{U} \in \Omega} \{p_{jj}(\mathbf{U})\} - \sum_{k=1, k \neq j}^r \max_{\mathbf{U} \in \Omega} \{p_{kj}(\mathbf{U})\} \quad (\text{IV.49})$$

et par la suite, en continuant la majoration :

$$\min_{\mathbf{U} \in \Omega} \{p_{jj}(\mathbf{U})\} - \sum_{k=1, k \neq j}^r \max_{\mathbf{U} \in \Omega} \{p_{kj}(\mathbf{U})\} \leq p_{jj}(\mathbf{U}) - \sum_{k=1, k \neq j}^r p_{kj}(\mathbf{U}) \quad (\text{IV.50})$$

Les inégalités (IV.49) et (IV.50) nous permet d'écrire :

$$p_{ii}(\mathbf{U}) + \sum_{k=1, k \neq i}^r p_{ki}(\mathbf{U}) << p_{jj}(\mathbf{U}) - \sum_{k=1, k \neq j}^r p_{kj}(\mathbf{U})$$

$$\forall i \in I_i, \forall j \in I_j, \forall \mathbf{U} \in \Omega \text{ avec } I_i \cap I_j = \emptyset, I_i \cup I_j = \{1, \dots, r\},$$

prouvant, selon le théorème de Gershgorine, l'existence de la double échelle de temps de la chaîne de Markov.

□

Après la mise en évidence des dynamiques et la simplification de la chaîne de Markov à commande, nous présentons le principe de calcul de la commande optimale des sous-systèmes découplés.

III.7. Utilisation des perturbations singulières pour la simplification du calcul de la commande optimale

Dans le cas général, le découplage d'un système en deux sous systèmes représentant l'un la partie rapide et l'autre la partie lente, permet de concevoir un régulateur du système global en commandant chacun des deux sous-systèmes. Cette méthode a été souvent utilisée dans le cas des systèmes continus singulièrement perturbés [CHO, 76] [DAU, 83] et généralisée dans le cas des systèmes échantillonnés singulièrement perturbés [ELM, 85].

Pour les chaînes de Markov, la commande est souvent de nature qualitative et ne peut pas toujours être le résultat d'action d'un régulateur ou d'un quelconque dispositif automatisé. En plus, le critère d'optimalité associé aux problèmes de commande des chaînes de Markov est un critère vectoriel. Cela implique que la commande optimale doit être calculée à chaque instant et pour chaque état du système stochastique. Dans ce cas, l'application des perturbations singulières pour la simplification du calcul de la commande optimale revient à déterminer une politique de commande qui optimise l'évolution de chaque sous-système. L'ensemble des commandes ainsi obtenues constitue une commande quasi-optimale pour le système initial.

III.7.1. Position du problème

Soit la chaîne de Markov :

$$\begin{cases} \mathbf{L}(n+1) = \mathbf{L}(n) \mathbf{T}_{11}(\mathbf{U}) + \mathbf{R}(n) \mathbf{T}_{21}(\mathbf{U}) \\ \mathbf{R}(n+1) = \mathbf{L}(n) \mathbf{T}_{12}(\mathbf{U}) + \mathbf{R}(n) \mathbf{T}_{22}(\mathbf{U}) \end{cases} \quad (\text{IV.51})$$

$$\mathbf{U} \in \Omega, \mathbf{L}(n) \in \mathbb{R}^{1 \times r_1}, \mathbf{R}(n) \in \mathbb{R}^{1 \times r_2}, \mathbf{r} = \mathbf{r}_1 + \mathbf{r}_2$$

avec les conditions initiales $\mathbf{L}(0)$, $\mathbf{R}(0)$.

Suivant la partition de la matrice de transition de la chaîne de Markov à commande, nous introduisons une partition similaire de la matrice des revenus élémentaires :

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix}$$

avec : $\mathcal{R}_{11} \in \mathbb{R}^{r_1 \times r_1}$ la matrice des revenus élémentaires de passage entre les états lents,

$\mathcal{R}_{12} \in \mathbb{R}^{r_1 \times r_2}$, $\mathcal{R}_{21} \in \mathbb{R}^{r_2 \times r_1}$ les matrices des revenus élémentaires de passage entre les états rapides et les lents,

$\mathcal{R}_{22} \in \mathbb{R}^{r_2 \times r_2}$ la matrice des revenus élémentaires de passage entre les états rapides.

Nous présentons la démarche de calcul relative aux problèmes terminaux, en spécifiant

que le cas des problèmes avec cible présente une résolution analogue.

L'équation d'optimalité du système initial, correspondant aux problèmes terminaux, présente la forme suivante :

$$\mathbf{w}(\mathbf{0}) = \underset{\mathbf{U} \in \mathcal{Z}}{\text{opt}} \left\{ \text{diag} \left[\mathbf{\ddagger}(\mathbf{U}) \times \mathbf{R}^T \right] + \mathbf{\ddagger}(\mathbf{U}) \mathbf{w}(\mathbf{0}+1) \right\} \quad (\text{IV.52})$$

avec le vecteur des revenus finaux :

$$\mathbf{w}(\mathbf{N}) = \begin{bmatrix} w_1(\mathbf{N}) \\ w_2(\mathbf{N}) \\ \dots \\ w_r(\mathbf{N}) \end{bmatrix}$$

où N représente l'horizon fini fixé.

III.7.2. Décomposition du système

L'utilisation de la méthode des perturbations singulières sur le système (IV.52) permet de le découpler en une partie lente et une rapide.

a) Partie lente

En utilisant le découplage de la chaîne de Markov à double échelle de temps, l'équation d'optimalité de la partie lente est directement exploitable, suite à la stochasticité de la matrice de transition $\mathbf{\ddagger}_1$. Cette équation s'écrit sous la forme suivante :

$$\mathbf{x}(\mathbf{0}) = \underset{\mathbf{U}_1 \in \mathcal{Z}}{\text{opt}} \left\{ \text{diag} \left[\mathbf{\ddagger}_1(\mathbf{U}_1) \times \mathbf{R}_1^T \right] + \mathbf{\ddagger}_1(\mathbf{U}_1) \mathbf{x}(\mathbf{0}+1) \right\} \quad (\text{IV.53})$$

avec la partition $\mathbf{w}(\mathbf{0}) = \begin{bmatrix} \mathbf{x}(\mathbf{0}) \\ \mathbf{z}(\mathbf{0}) \end{bmatrix}$, où $\mathbf{x}(\mathbf{0}) \in \mathbb{R}^{r_1}$ correspond à la partie lente et $\mathbf{z}(\mathbf{0}) \in \mathbb{R}^{r_2}$ à la partie rapide.

La matrice de revenus est indépendante de la matrice de transition de la chaîne de Markov. Le seul lien entre ces deux matrices est représenté par les états du système. En effet, une probabilité de transition p_{ij} et le revenu élémentaire r_{ij} concernent la même transition entre les états e_i et e_j . En découplant la matrice $\mathbf{\ddagger}$, les probabilités de transition ont changé mais les états sont les mêmes. Par conséquent, la matrice \mathcal{R}_l est égale à \mathcal{R}_1 . L'équation (IV.53) d'optimalité de la partie lente prend la forme suivante :

$$\mathbf{x}(\mathbf{0}) = \underset{\mathbf{U}_1 \in \mathcal{Z}}{\text{opt}} \left\{ \text{diag} \left[\mathbf{\ddagger}_1(\mathbf{U}_1) \times \mathbf{R}_{11}^T \right] + \mathbf{\ddagger}_1(\mathbf{U}_1) \mathbf{x}(\mathbf{0}+1) \right\} \quad (\text{IV.54})$$

avec les revenus finaux : $\mathbf{x}^O(N) = \begin{bmatrix} w_1(N) \\ w_2(N) \\ \dots \\ w_{r_1}(N) \end{bmatrix}$

La commande optimale de la partie lente sera représentée par la matrice de dimension $(K \times r_1)$ suivante :

$$\mathbf{G}_1(n) = [g_1(n, 1) \ g_1(n, 2) \ \dots \ g_1(n, r_1)]$$

avec

$$g_1(n, i) = \hat{U}_1(n, i), \quad \forall i = 1, 2, \dots, r_1 \quad \text{et} \quad \forall n = 0, 1, \dots, N-1.$$

b) Partie rapide

Pour calculer la commande optimale de la partie rapide de la chaîne de Markov à commande, nous utilisons une réinjection des espérances $\hat{\mathbf{x}}_1$ de la partie lente dans l'équation d'optimalité initiale. Cette fois, nous avons affaire à la commande initiale \mathbf{U} .

La forme partitionnée du critère d'optimalité (IV.52) du système initial présente la forme suivante :

$$\begin{pmatrix} \mathbf{x}^O \\ \mathbf{z}^O \end{pmatrix} = \underset{\mathbf{U} \in \mathcal{Z}}{\text{opt}} \left\{ \text{diag} \left(\begin{bmatrix} \dagger_{11}(\mathbf{U}) & \dagger_{12}(\mathbf{U}) \\ \dagger_{21}(\mathbf{U}) & \dagger_{22}(\mathbf{U}) \end{bmatrix} \times \begin{bmatrix} \mathbf{R}_{11}^T & \mathbf{R}_{21}^T \\ \mathbf{R}_{12}^T & \mathbf{R}_{22}^T \end{bmatrix} \right) + \begin{bmatrix} \dagger_{11}(\mathbf{U}) & \dagger_{12}(\mathbf{U}) \\ \dagger_{21}(\mathbf{U}) & \dagger_{22}(\mathbf{U}) \end{bmatrix} \times \begin{pmatrix} \mathbf{x}^{O+1} \\ \mathbf{z}^{O+1} \end{pmatrix} \right\}$$

ou sous une autre forme :

$$\begin{pmatrix} \mathbf{x}^O \\ \mathbf{z}^O \end{pmatrix} = \underset{\mathbf{U} \in \mathcal{Z}}{\text{opt}} \left\{ \text{diag} \left(\begin{bmatrix} \dagger_{11}(\mathbf{U}) \mathbf{R}_{11}^T + \dagger_{12}(\mathbf{U}) \mathbf{R}_{12}^T \\ \dagger_{21}(\mathbf{U}) \mathbf{R}_{21}^T + \dagger_{22}(\mathbf{U}) \mathbf{R}_{22}^T \end{bmatrix} \right) + \begin{bmatrix} \dagger_{11}(\mathbf{U}) & \dagger_{12}(\mathbf{U}) \\ \dagger_{21}(\mathbf{U}) & \dagger_{22}(\mathbf{U}) \end{bmatrix} \times \begin{pmatrix} \mathbf{x}^{O+1} \\ \mathbf{z}^{O+1} \end{pmatrix} \right\}$$

et en écrivant uniquement l'équation de la partie rapide :

$$\mathbf{z}^O = \underset{\mathbf{U} \in \mathcal{Z}}{\text{opt}} \left\{ \text{diag} \left(\dagger_{21}(\mathbf{U}) \mathbf{R}_{21}^T + \dagger_{22}(\mathbf{U}) \mathbf{R}_{22}^T \right) + \dagger_{21}(\mathbf{U}) \mathbf{x}^{O+1} + \dagger_{22}(\mathbf{U}) \mathbf{z}^{O+1} \right\} \quad (\text{IV.55})$$

Lorsque nous remplaçons l'espérance de revenu du système lent $\hat{\mathbf{x}}(n)$ par les valeurs $\hat{\mathbf{x}}_1(n)$ (calculé dans (IV.54)) dans la relation (IV.55), le critère d'optimalité de la partie rapide se réécrit :

$$\mathbf{z}^O = \underset{\mathbf{U} \in \mathcal{Z}}{\text{opt}} \left\{ \text{diag} \left(\dagger_{21}(\mathbf{U}) \mathbf{R}_{21}^T + \dagger_{22}(\mathbf{U}) \mathbf{R}_{22}^T \right) + \dagger_{21}(\mathbf{U}) \mathbf{x}_1^{O(n+1)} + \dagger_{22}(\mathbf{U}) \mathbf{z}^{O+1} \right\} \quad (\text{IV.56})$$

avec les conditions finales : $\mathbf{z}^{\circ}(N) = \begin{bmatrix} w_{r_1+1}(N) \\ w_{r_1+2}(N) \\ \dots \\ w_r(N) \end{bmatrix}$

La commande optimale des états rapides sera représentée par la matrice de dimension $(K \times r_2)$ suivante :

$$\mathbf{G}_r(n) = \begin{bmatrix} g_r(n, r_1+1) & g_r(n, r_1+2) & \dots & g_r(n, r) \end{bmatrix}$$

avec $g_r(n, i) = \hat{U}(n, i)$, $\forall i = r_1+1, r_1+2, \dots, r$ et $\forall n = 0, 1, \dots, N-1$.

III.7.3. Détermination de la commande quasi-optimale du système initial

La commande quasi-optimale du système initial sera obtenue à partir des commandes optimales des sous-systèmes lent et rapide :

$$\mathbf{G}(n) = \begin{bmatrix} \mathbf{G}_l(n) & \mathbf{G}_r(n) \end{bmatrix}$$

et l'évolution des revenus est approchée par les évolutions partielles :

$$\mathbf{w}^{\circ}(n) = \hat{\mathbf{Y}} \begin{bmatrix} \mathbf{x}(n) \\ \mathbf{z}(n) \end{bmatrix}$$

Nous remarquons que la commande optimale de la partie lente est approchée par \mathbf{U}_l . Pour la partie rapide, la commande calculée par réinjection est directement une commande composite $\mathbf{U} = \mathbf{U}_l + \mathbf{U}_r$.

La technique proposée s'applique à la simplification de l'équation d'optimalité dans tous les cas de problèmes terminaux (horizon fini, horizon infini et horizon aléatoire), ainsi que dans le cas des problèmes avec cible.

Nous présentons par la suite un exemple numérique illustrant l'efficacité des techniques développées.

III.8. Exemple numérique

Soit une chaîne de Markov à commande, formée par 3 états et représentée par la matrice de transition affine en la commande ($K = 2$) :

$$\dot{\mathbf{x}}(\mathbf{U}) = \begin{pmatrix} 0.94 & 0.03 & 0.03 \\ 0.04 & 0.9 & 0.06 \\ 0.4 & 0.4 & 0.2 \end{pmatrix} + \begin{pmatrix} 0.0015 & -0.0005 & -0.001 \\ -0.0005 & 0.0015 & -0.001 \\ 0.002 & 0.002 & -0.004 \end{pmatrix} \mathbf{u}_1 + \begin{pmatrix} 0.002 & -0.001 & -0.001 \\ -0.0005 & 0.001 & -0.0005 \\ 0.002 & 0.002 & -0.004 \end{pmatrix} \mathbf{u}_2$$

avec les notations : $\dot{\mathbf{x}}(\mathbf{U}) = \dot{\mathbf{x}}_0 + \dot{\mathbf{x}}_1 \mathbf{u}_1 + \dot{\mathbf{x}}_2 \mathbf{u}_2$,

et soit aussi la matrice de revenus :

$$\mathbf{R} = \begin{pmatrix} 0.01202 & 0.012 & 0.012 \\ 0.01202 & 0.01201 & 0.01198 \\ 0.012 & 0.01201 & 0.0121 \end{pmatrix}$$

Nous considérons une commande $\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ comprise dans l'espace de commande Ω donné par les intervalles :

$$u_1 \in [0, 2] \text{ et } u_2 \in [0, 4]$$

Nous traitons le problème de commande avec horizon fini de 15 transitions ($N = 15$), dans lequel on cherche à atteindre un revenu maximum. L'équation d'optimalité associé à ce type de problème, présente la forme suivante :

$$\mathbf{w}(0) = \max_{\mathbf{U} \in \mathcal{Z}} \left\{ \text{diag} \left[\mathbf{I}^T (\mathbf{U}) \times \mathbf{R}^T \right] + \mathbf{I}^T (\mathbf{U}) \mathbf{w}(1) \right\}$$

En considérant une espérance nulle de revenu à la fin des 15 transitions :

$$\mathbf{w}(N) = \mathbf{w}(15) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

l'équation d'optimalité de la chaîne nous donne l'évolution suivante du vecteur des espérances de revenus totaux :

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
\mathbf{e}_1	0.180 3	0.168 2	0.156 2	0.144 2	0.132 2	0.120 2	0.108 2	0.096 1	0.084 1	0.072 1	0.060 1	0.048 1	0.036 1	0.024 1	0.012 1	0
\mathbf{e}_2	0.180 2	0.168 2	0.156 2	0.144 1	0.132 1	0.120 1	0.108 1	0.096 1	0.084 1	0.072 1	0.060 1	0.048 1	0.036 1	0.024 1	0.012 1	0
\mathbf{e}_3	0.180 2	0.168 2	0.156 2	0.144 2	0.132 2	0.120 2	0.108 1	0.096 1	0.084 1	0.072 1	0.060 1	0.048 1	0.036 1	0.024 1	0.012 1	0

La commande optimale ainsi obtenue est :

- commande u_1

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
\mathbf{e}_1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
\mathbf{e}_2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2
\mathbf{e}_3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

- commande u_2

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
\mathbf{e}_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
\mathbf{e}_2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0
\mathbf{e}_3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

Quelle que soit la commande $\mathbf{U} \in \Omega$ adoptée, la partie lente contiendra toujours les états : \mathbf{e}_1 , \mathbf{e}_2 et la partie rapide l'état \mathbf{e}_3 . Ceci est mis en évidence par la méthode globale qui donne pour les matrices constantes \mathbf{I}_1 , \mathbf{I}_2 et \mathbf{I}_3 le tracé suivant des disques de Gershgorine :

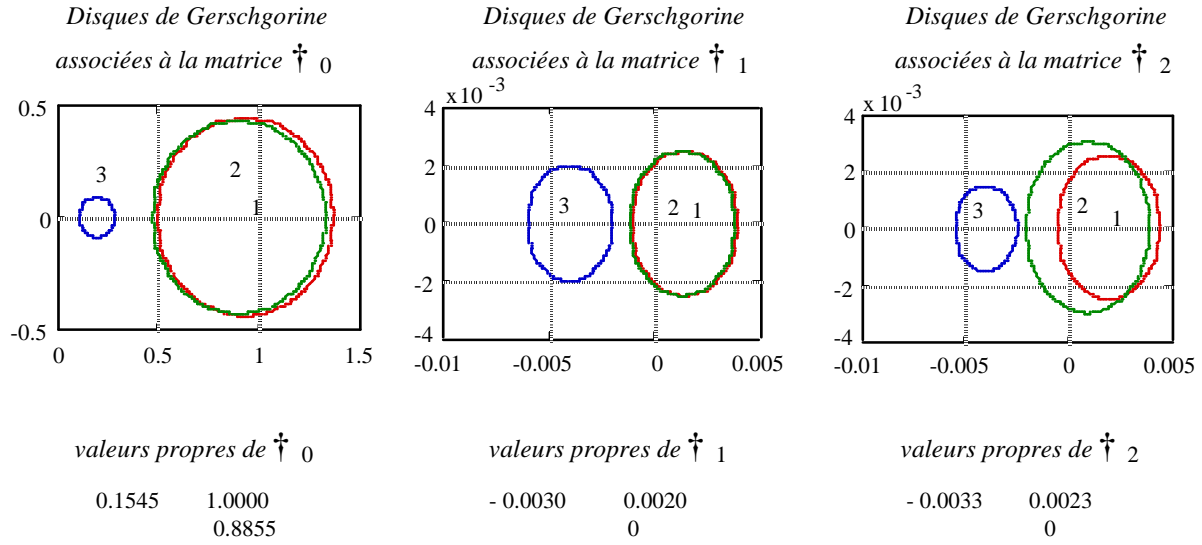


Fig.IV.1. Application de la méthode globale pour la détection des dynamiques.

L'existence des dynamiques différentes est prouvée aussi par application du théorème 8 :

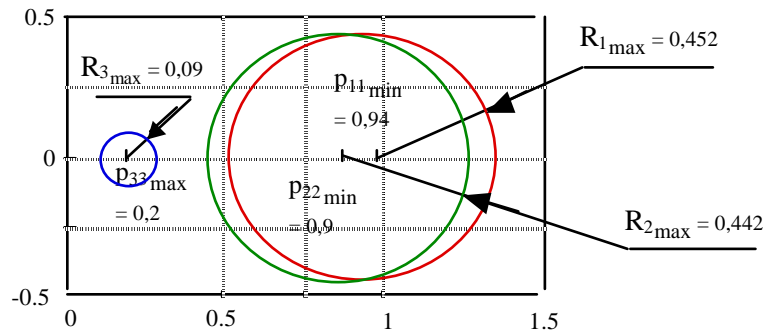


Fig.IV.2. Détection des dynamiques en utilisant la matrice majorante.

ainsi que par l'utilisation des disques de Gershgorine dynamiques sur les colonnes :

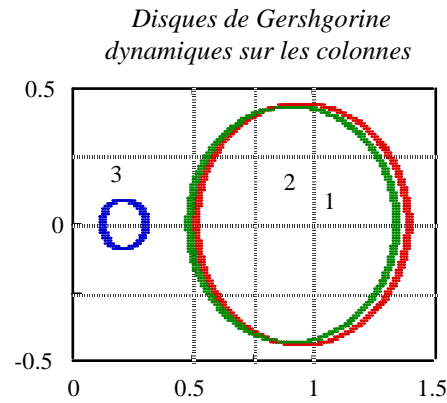


Fig.IV.3. Application de la méthode des disques de Gershgorine dynamiques.

L'évolution du coefficient de découplage des dynamiques en fonction de la commande est représentée dans la figure qui suit :

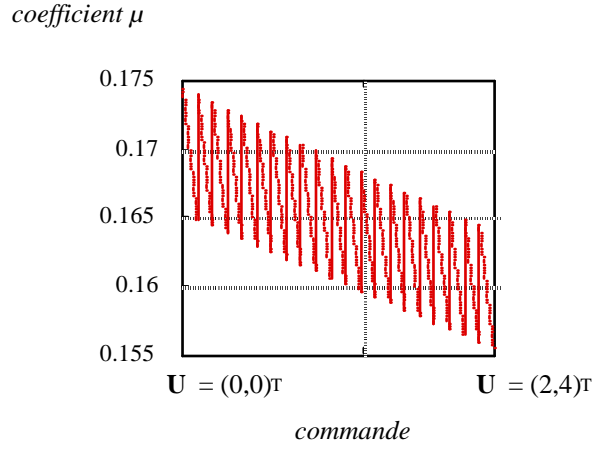


Fig.IV.4. Evolution du coefficient de découplage en fonction de la commande.

Nous remarquons que μ varie entre 0,155 et 0,175 sur tout l'ensemble de commande. Le système est caractérisé par une forte séparation des dynamiques. L'application des perturbations singulières sur ce système nous permet d'obtenir le système découplé avec les matrices de transition suivantes :

$$\dot{\dagger}_1(\mathbf{U}) = \begin{pmatrix} 0.955 & 0.045 \\ 0.07 & 0.93 \end{pmatrix} + \begin{pmatrix} 0.001 & -0.001 \\ -0.001 & 0.001 \end{pmatrix} \mathbf{u}_1 + \begin{pmatrix} 0.0015 & -0.0015 \\ -0.0008 & 0.0008 \end{pmatrix} \mathbf{u}_2$$

et

$$\dot{\dagger}_r(\mathbf{U}) = 0.2 - 0.004 \mathbf{u}_1 - 0.004 \mathbf{u}_2$$

Avec la matrice de bilinéarisation du système lent \mathbf{G} de la forme :

$$\mathbf{G} = \begin{pmatrix} 0.5 & 0.5 \end{pmatrix}$$

obtenue par remplacement des sous matrices :

$$\begin{aligned} \dot{\dagger}_0 &= \begin{pmatrix} \dot{\dagger}_{01} & \dot{\dagger}_{02} \\ \dot{\dagger}_{11} & \dot{\dagger}_{12} \\ \dot{\dagger}_{21} & \dot{\dagger}_{22} \end{pmatrix} = \begin{pmatrix} 0.94 & 0.03 & 0.03 \\ 0.04 & 0.9 & 0.06 \\ 0.4 & 0.4 & 0.2 \end{pmatrix}, \quad \dot{\dagger}_1 = \begin{pmatrix} \dot{\dagger}_{11} & \dot{\dagger}_{12} \\ \dot{\dagger}_{21} & \dot{\dagger}_{22} \end{pmatrix} = \begin{pmatrix} 0.0015 & -0.0005 & -0.001 \\ -0.0005 & 0.0015 & -0.001 \\ 0.002 & 0.002 & -0.004 \end{pmatrix}, \\ \dot{\dagger}_2 &= \begin{pmatrix} \dot{\dagger}_{21} & \dot{\dagger}_{22} \\ \dot{\dagger}_{21} & \dot{\dagger}_{22} \end{pmatrix} = \begin{pmatrix} 0.002 & -0.001 & -0.001 \\ -0.0005 & 0.001 & -0.0005 \\ 0.002 & 0.002 & -0.004 \end{pmatrix} \end{aligned}$$

dans la formule :

$$\mathbf{G} = \left[\left(\mathbf{I} - \dot{\dagger}_{022} \right)^T \left(\mathbf{I} - \dot{\dagger}_{022} \right) + \dot{\dagger}_{122}^T \dot{\dagger}_{122} + \dot{\dagger}_{222}^T \dot{\dagger}_{222} \right]^{-1} \left[\left(\mathbf{I} - \dot{\dagger}_{022} \right)^T \dot{\dagger}_{021} - \dot{\dagger}_{122}^T \dot{\dagger}_{121} - \dot{\dagger}_{222}^T \dot{\dagger}_{221} \right]$$

La commande optimale de la partie lente se fait selon l'équation d'optimalité :

$$\mathbf{x}(\mathbf{n}) = \max_{\mathbf{U}_1 \in \mathcal{Z}} \left\{ \text{diag} \left[\dagger_1(\mathbf{U}_1) \times \mathbf{R}_{11}^T \right] + \dagger_1(\mathbf{U}_1) \mathbf{x}(\mathbf{n}+1) \right\}$$

avec :

$$\mathbf{R}_{11} = \begin{pmatrix} 0.01202 & 0.012 \\ 0.01202 & 0.01201 \end{pmatrix}$$

Après avoir effectué les calculs, l'évolution de l'espérance de revenu de la partie lente est

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
\mathbf{e}_1	0.180 3	0.168 3	0.156 2	0.144 2	0.132 2	0.120 2	0.108 2	0.096 1	0.084 1	0.072 1	0.060 1	0.048 1	0.036 1	0.024 1	0.012 0	0
\mathbf{e}_2	0.180 2	0.168 2	0.156 2	0.144 2	0.132 1	0.120 1	0.108 1	0.096 1	0.084 1	0.072 1	0.060 1	0.048 0	0.036 0	0.024 0	0.012 0	0

avec la politique optimale suivante :

- commande \mathbf{u}_1

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
\mathbf{e}_1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
\mathbf{e}_2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

- commande \mathbf{u}_2

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
\mathbf{e}_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
\mathbf{e}_2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

Après la réinjection des résultats du système lent dans l'équation d'optimalité initiale, nous obtenons l'équation d'optimalité suivante :

$$\mathbf{z}(\mathbf{n}) = \max_{\mathbf{U} \in \mathcal{Z}} \left\{ \text{diag} \left(\left[\dagger_{21}(\mathbf{U}) \dagger_{22}(\mathbf{U}) \times \begin{bmatrix} \mathbf{R}_{21}^T \\ \mathbf{R}_{22}^T \end{bmatrix} \right] + \left[\dagger_{21}(\mathbf{U}) \dagger_{22}(\mathbf{U}) \right] \times \begin{pmatrix} \mathbf{x}(\mathbf{n}+1) \\ \mathbf{z}(\mathbf{n}+1) \end{pmatrix} \right) \right\}$$

avec :

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix} = \begin{pmatrix} 0.01202 & 0.012 & 0.012 \\ 0.01202 & 0.01201 & 0.01198 \\ 0.012 & 0.01201 & 0.0121 \end{pmatrix}$$

qui donne les espérances optimales la partie rapide :

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
\mathbf{e}_3	0.180 2	0.168 2	0.156 2	0.144 2	0.132 2	0.120 2	0.108 1	0.096 1	0.084 1	0.072 1	0.060 1	0.048 1	0.036 1	0.024 1	0.012 0	0

et la commande optimale associée :

- commande \mathbf{u}_1

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
\mathbf{e}_3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

- commande \mathbf{u}_2

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
e_3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

Les erreurs maximales de l'espérance de revenu sont les suivantes (valeurs constatées au moment initial) :

0,002 % pour l'état 1
 0,0083 % pour l'état 2
 0,0048 % pour l'état 3

La précision d'approximation de l'évolution de l'espérance de revenu du système initial par les espérances de revenu des systèmes découplés, nous donne une image de l'efficacité des perturbations singulières. En travaillant sur des équations récurrentes d'ordre réduit, nous obtenons la commande quasi-optimale du système global tout en ayant une bonne approximation de ses performances en terme d'espérance de revenu.

La prise en compte des pondérations des états dans le calcul de la commande optimale d'une chaîne de Markov, offre l'avantage d'une commande optimale effectuée à partir de deux équations d'optimalité d'ordre réduit qui peuvent être traitées indépendamment. Par conséquent, dans ce qui suit nous allons appliquer la méthode de découplage en régime permanent sur la classe des chaînes de Markov bilinéaires à commande.

IV. Simplification de la commande des chaînes de Markov à double échelle de pondération

Comme pour l'utilisation de la notion de double échelle de temps, la propriété de double échelle de pondération d'une chaîne de Markov à commande permet la séparation du problème de commande en deux problèmes de commande optimale : l'un concerne l'évolution des états à probabilités fortes et l'autre les états à faibles probabilités.

Nous commençons notre étude par donner la définition de la propriété de double échelle de pondération des chaînes de Markov à commande.

IV.1. Double échelle de pondération dans les chaînes de Markov à commande

Définition

Une chaîne de Markov homogène, finie et ergodique :

$$\mathbf{P}(n+1) = \mathbf{P}(n) \mathbf{\ddagger}(U), \quad n = 0, 1, 2, 3, \dots, \quad U \in \Omega \text{ borné}$$

$$\mathbf{P}(n) \xrightarrow{1 \times r} \text{de conditions initiales } \mathbf{P}(0),$$

possède la propriété de double échelle de pondération sur Ω si il existe une permutation indépendante de U :

$$J : \{1, 2, \dots, r\} \rightarrow \{1, 2, \dots, r\}$$

qui appliquée à la forme propre correspondant à la valeur propre 1 de la matrice de transition $\dot{\mathbf{T}}(\mathbf{U})$

$$\mathbf{f}_1(\mathbf{U}) = [f_1^1(\mathbf{U}) \ f_1^2(\mathbf{U}) \dots f_1^r(\mathbf{U})]$$

permet de la partitionner en deux sous-ensembles $\mathbf{f}_{1(s)}(\mathbf{U})$ et $\mathbf{f}_{1(w)}(\mathbf{U})$:

$$\mathbf{f}_1^J(\mathbf{U}) = [\mathbf{f}_{1(s)}(\mathbf{U}) \ \mathbf{f}_{1(w)}(\mathbf{U})] \quad (\text{IV.57})$$

de telle façon que l'inégalité :

$$\min_{f_1^j(\mathbf{U}) \in f_{1(s)}(\mathbf{U})} |f_1^j(\mathbf{U})| = \tilde{\mathbf{s}}(\mathbf{U}) \gg \tilde{\mathbf{w}}(\mathbf{U}) = \max_{f_1^j(\mathbf{U}) \in f_{1(w)}(\mathbf{U})} |f_1^j(\mathbf{U})|$$

soit satisfaite quel que soit $\mathbf{U} \in \Omega$.

Dans la forme (IV.57), le vecteur $\mathbf{f}_{1(s)}$ représente les états forts, et $\mathbf{f}_{1(w)}$ les états faibles.

□

Un indicateur de séparation des pondérations est donné par le rapport :

$$\mu_p = \max_{\mathbf{U} \in \Omega} \left(\frac{\mathbf{w} \ddot{\mathbf{Y}}(\mathbf{U})}{\mathbf{s} \ddot{\mathbf{Y}}(\mathbf{U})} \right) \ll 1$$

La mise en évidence des pondérations se fait en utilisant une technique locale. Cette technique consiste à étudier la forme propre \mathbf{f}_1 en fonction de la commande \mathbf{U} . Pour faciliter les calculs numériques, une technique similaire à (IV.47) peut être appliquée en utilisant le conditionnement de la matrice $\dot{\mathbf{T}}$.

Au cas où la chaîne de Markov à commande possède la propriété de double échelle de pondération, nous pouvons lui appliquer un découplage en régime permanent par extension des techniques développées dans le cas linéaire (sans commande).

IV.2. Découplage des pondérations différentes dans la chaîne de Markov à commande

L'équation fondamentale d'une chaîne de Markov bilinéaire présente la forme (IV.32) :

$$\mathbf{P}(n+1) = \mathbf{P}(n) \left(\dot{\mathbf{T}}_0 + \sum_{i=1}^K \dot{\mathbf{T}}_i \cdot u_i \right)$$

avec : $\mathbf{P}(n) \in \mathbb{R}^{1 \times r}$ désignant la distribution des probabilités à l'instant n ($n = 0, 1, 2, \dots$),

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_K \end{bmatrix} \in \Omega \subset \mathbb{R}^K \text{ désignant le vecteur de commande } (K \in \mathbb{N}^*)$$

et $\dot{\mathbf{T}}(\mathbf{U}) = [\dot{p}_{ij}(\mathbf{U})]_{i,j=1, \dots, r} \in \mathbb{R}^{r \times r}$ ($r \in \mathbb{N}^*$) la matrice de transition de la forme

$$\dot{\mathbf{T}}(\mathbf{U}) = (\dot{\mathbf{T}}_0 + \dot{\mathbf{T}}_1 \cdot u_1 + \dot{\mathbf{T}}_2 \cdot u_2 + \dots + \dot{\mathbf{T}}_K \cdot u_K)$$

où les matrices $\dot{\mathbf{T}}_k = [\mathbf{p}_{ij}^k(\mathbf{U})]_{i,j=1, \dots, r}$ ($k = 0, 1, 2, \dots, K$) sont des matrices constantes.

En arrangeant la matrice de transition $\dot{\mathbf{T}}(\mathbf{U})$ selon les pondérations des états, nous obtenons la forme partitionnée de l'équation fondamentale :

$$[\mathbf{S}(n+1) \ \mathbf{W}(n+1)] = [\mathbf{S}(n) \ \mathbf{W}(n)] \begin{pmatrix} \dot{\mathbf{T}}'_s(\mathbf{U}) & \dot{\mathbf{T}}'_{s-w}(\mathbf{U}) \\ \dot{\mathbf{T}}'_{w-s}(\mathbf{U}) & \dot{\mathbf{T}}'_w(\mathbf{U}) \end{pmatrix} \quad (\text{IV.58})$$

où on a noté par $\mathbf{P}(n) = [\mathbf{S}(n) \ \mathbf{W}(n)]$ la décomposition due la distribution du système initial en :

$\mathbf{S}' \hat{\mathbf{e}}^{1 \times r_1}$ la distribution des probabilités de la partie forte et

$\mathbf{W}' \hat{\mathbf{e}}^{1 \times r_2}$ ($r_1 + r_2 = r$) la distribution des probabilités de la partie faible.

Par application des hypothèses $\mathbf{S}(n+1) = \mathbf{S}(n) = \mathbf{S}_w(n)$ et $\mathbf{W}(n+1) = \mathbf{W}(n) = \mathbf{W}_s(n)$ caractéristiques au régime permanent des parties forte et faible, nous obtenons les formules des systèmes découplés :

$$[\mathbf{S}_s(n+1) \ \mathbf{W}_w(n+1)] = [\mathbf{S}_s(n) \ \mathbf{W}_w(n)] \begin{pmatrix} \dot{\mathbf{T}}'_s(\mathbf{U}) & 0 \\ 0 & \dot{\mathbf{T}}'_w(\mathbf{U}) \end{pmatrix} \quad (\text{IV.59})$$

avec les matrices stochastiques :

$$\dot{\mathbf{T}}'_s(\mathbf{U}) = \dot{\mathbf{T}}'_s(\mathbf{U}) + \dot{\mathbf{T}}'_{s-w}(\mathbf{U}) (\mathbf{I} - \dot{\mathbf{T}}'_s(\mathbf{U}))^{-1} \dot{\mathbf{T}}'_{w-s}(\mathbf{U}) \quad (\text{IV.60})$$

$$\dot{\mathbf{T}}'_w(\mathbf{U}) = \dot{\mathbf{T}}'_w(\mathbf{U}) + \dot{\mathbf{T}}'_{w-s}(\mathbf{U}) (\mathbf{I} - \dot{\mathbf{T}}'_s(\mathbf{U}))^{-1} \dot{\mathbf{T}}'_{s-w}(\mathbf{U}) \quad (\text{IV.61})$$

et les conditions initiales $\mathbf{S}_s(0)$ et $\mathbf{W}_w(0)$.

Remarque

Les matrices de transition des sous-systèmes fort et faible ne présentent plus une forme affine en la commande après le découplage :

$$\dot{\mathbf{T}}'_s(\mathbf{U}) = \dot{\mathbf{T}}'_{0_s} + \sum_{k=1}^K \left(\dot{\mathbf{T}}'_{k_s} u_k \right) + \left[\dot{\mathbf{T}}'_{0_{s-w}} + \sum_{k=1}^K \left(\dot{\mathbf{T}}'_{k_{s-w}} u_k \right) \right] \left[\mathbf{I} - \left(\dot{\mathbf{T}}'_{0_w} + \sum_{k=1}^K \left(\dot{\mathbf{T}}'_{k_w} u_k \right) \right) \right]^{-1} \left[\dot{\mathbf{T}}'_{0_{w-s}} + \sum_{k=1}^K \left(\dot{\mathbf{T}}'_{k_{w-s}} u_k \right) \right] \quad (\text{IV.62})$$

$$\dot{\mathbf{T}}'_w(\mathbf{U}) = \dot{\mathbf{T}}'_{0_w} + \sum_{k=1}^K \left(\dot{\mathbf{T}}'_{k_w} u_k \right) + \left[\dot{\mathbf{T}}'_{0_{w-s}} + \sum_{k=1}^K \left(\dot{\mathbf{T}}'_{k_{w-s}} u_k \right) \right] \left[\mathbf{I} - \left(\dot{\mathbf{T}}'_{0_s} + \sum_{k=1}^K \left(\dot{\mathbf{T}}'_{k_s} u_k \right) \right) \right]^{-1} \left[\dot{\mathbf{T}}'_{0_{s-w}} + \sum_{k=1}^K \left(\dot{\mathbf{T}}'_{k_{s-w}} u_k \right) \right] \quad (\text{IV.63})$$

Ceci nous empêche d'appliquer les méthodes classiques de calcul de la commande optimale directement sur les systèmes découplés. Par conséquent, nous procédons à un conditionnement des deux sous-systèmes.

□

IV.3. Bilinearisation des sous-systèmes

La technique de conditionnement est similaire à celle utilisée pour le sous-système lent dans le cas du découplage par perturbations singulières, et revient à calculer les matrices constantes \mathbf{V} et respectivement \mathbf{Y} , qui vérifient les relations suivantes :

$$\mathbf{V} = \left[\mathbf{I} - \left(\begin{array}{c} \ddagger' \\ 0_w \end{array} + \sum_{k=1}^K \left(\begin{array}{c} \ddagger' \\ k_w \end{array} u_k \right) \right) \right]^{-1} \left[\begin{array}{c} \ddagger' \\ 0_{w-s} \end{array} + \sum_{k=1}^K \left(\begin{array}{c} \ddagger' \\ k_{w-s} \end{array} u_k \right) \right], \text{ À } \mathbf{U}' \text{ ž} ; \mathbf{V}' \hat{e}^{r_2 \times r_1} \quad (\text{IV.64})$$

respectivement :

$$\mathbf{Y} = \left[\mathbf{I} - \left(\begin{array}{c} \ddagger' \\ 0_s \end{array} + \sum_{k=1}^K \left(\begin{array}{c} \ddagger' \\ k_s \end{array} u_k \right) \right) \right]^{-1} \left[\begin{array}{c} \ddagger' \\ 0_{s-w} \end{array} + \sum_{k=1}^K \left(\begin{array}{c} \ddagger' \\ k_{s-w} \end{array} u_k \right) \right], \text{ À } \mathbf{U}' \text{ ž} ; \mathbf{Y}' \hat{e}^{r_1 \times r_2} \quad (\text{IV.65})$$

Ces matrices constantes sont obtenues comme solutions des équations matricielles :

$$\begin{bmatrix} \ddagger' \\ 0_{w-s} \\ \ddagger' \\ 1_{w-s} \\ \dots \\ \ddagger' \\ K_{w-s} \end{bmatrix} = \begin{bmatrix} \left(\mathbf{I} - \begin{array}{c} \ddagger' \\ 0_w \end{array} \right) \\ -\ddagger' \\ 1_w \\ \dots \\ -\ddagger' \\ K_w \end{bmatrix} \times \mathbf{V} \quad (\text{IV.66})$$

respectivement

$$\begin{bmatrix} \ddagger' \\ 0_{s-w} \\ \ddagger' \\ 1_{s-w} \\ \dots \\ \ddagger' \\ K_{s-w} \end{bmatrix} = \begin{bmatrix} \left(\mathbf{I} - \begin{array}{c} \ddagger' \\ 0_s \end{array} \right) \\ -\ddagger' \\ 1_s \\ \dots \\ -\ddagger' \\ K_s \end{bmatrix} \times \mathbf{Y} \quad (\text{IV.67})$$

La résolution des équations (IV.66) et (IV.67) donne les résultats suivants :

$$\mathbf{V} = \left[\left(\mathbf{I} - \begin{array}{c} \ddagger' \\ 0_w \end{array} \right)^T \left(\mathbf{I} - \begin{array}{c} \ddagger' \\ 0_w \end{array} \right) + \sum_{k=0}^K \begin{array}{c} \ddagger' \\ k_w \end{array}^T \begin{array}{c} \ddagger' \\ k_w \end{array} \right]^{-1} \left[\left(\mathbf{I} - \begin{array}{c} \ddagger' \\ 0_w \end{array} \right)^T \begin{array}{c} \ddagger' \\ 0_{w-s} \end{array} - \sum_{k=1}^K \begin{array}{c} \ddagger' \\ k_w \end{array}^T \begin{array}{c} \ddagger' \\ k_{w-s} \end{array} \right] \quad (\text{IV.68})$$

respectivement :

$$\mathbf{Y} = \left[\left(\mathbf{I} - \begin{array}{c} \ddagger' \\ 0_s \end{array} \right)^T \left(\mathbf{I} - \begin{array}{c} \ddagger' \\ 0_s \end{array} \right) + \sum_{k=0}^K \begin{array}{c} \ddagger' \\ k_s \end{array}^T \begin{array}{c} \ddagger' \\ k_s \end{array} \right]^{-1} \left[\left(\mathbf{I} - \begin{array}{c} \ddagger' \\ 0_s \end{array} \right)^T \begin{array}{c} \ddagger' \\ 0_{s-w} \end{array} - \sum_{k=1}^K \begin{array}{c} \ddagger' \\ k_s \end{array}^T \begin{array}{c} \ddagger' \\ k_{s-w} \end{array} \right]$$

(IV.69)

Par conséquent, l'équation fondamentale découplée prend la forme :

$$[S_s(n+1) \ W_w(n+1)] = [S_s(n) \ W_w(n)] \begin{pmatrix} \dagger_s(U) & 0 \\ 0 & \dagger_w(U) \end{pmatrix} \quad (IV.70)$$

avec les matrices de transition :

$$\begin{aligned} \dagger_s(U) &= \dagger'_s(U) + \dagger'_{s-w}(U) \times \mathcal{V} \\ \dagger_w(U) &= \dagger'_w(U) + \dagger'_{w-s}(U) \times \mathcal{Y}, \end{aligned}$$

fonctions affines en la commande.

IV.4. Calcul de la commande optimale par découplage en régime permanent

IV.4.1. Position du problème

Le découplage en régime permanent de la chaîne de Markov à commande donne une approche du comportement asymptotique du système. Par conséquent, ce découplage est particulièrement adapté au calcul de la commande optimale dans le cas des problèmes terminaux à horizon infini. L'équation d'optimalité associée à un tel problème terminal présente la forme suivante :

$$\mathbf{w}(\infty) = \underset{U \in \mathcal{Z}}{\text{opt}} \left\{ \text{diag} \left[\dagger(U) \times \mathbf{R}^T \right] + \gamma \dagger(U) \mathbf{w}(\infty) \right\} \quad (IV.71)$$

avec le facteur de perte $\gamma \in [0,1]$ et le vecteur des revenus à l'infini :

$$\mathbf{w}(\infty) = \begin{bmatrix} w_1(\infty) \\ w_2(\infty) \\ \dots \\ w_r(\infty) \end{bmatrix}$$

IV.4.2. Commande optimale des sous-systèmes

En utilisant le découplage du système en régime permanent :

$$[S_s(n+1) \ W_w(n+1)] = [S_s(n) \ W_w(n)] \begin{pmatrix} \dagger_s(U) & 0 \\ 0 & \dagger_w(U) \end{pmatrix}$$

avec les matrices des sous-systèmes fort respectivement faible :

$$\begin{aligned} \dagger_s(U) &= \dagger'_s(U) + \dagger'_{s-w}(U) \times \mathcal{V} \\ \dagger_w(U) &= \dagger'_w(U) + \dagger'_{w-s}(U) \times \mathcal{Y}, \end{aligned}$$

les équations d'optimalité des parties forte et faible s'écrivent sous la forme suivante :

- pour la partie forte :

$$\mathbf{w}_s^{\hat{O}}(n) = \max_{\mathbf{U} \in \mathcal{Z}} \left\{ \text{diag} \left[\dot{\mathbf{r}}_s(\mathbf{U}) \mathbf{R}_{11}^T \right] + \gamma \dot{\mathbf{r}}_s(\mathbf{U}) \mathbf{w}_s^{\hat{O}}(n+1) \right\} \quad (\text{IV.72})$$

- pour la partie faible :

$$\mathbf{w}_w^{\hat{O}}(n) = \max_{\mathbf{U} \in \mathcal{Z}} \left\{ \text{diag} \left[\dot{\mathbf{r}}_w(\mathbf{U}) \mathbf{R}_{22}^T \right] + \gamma \dot{\mathbf{r}}_w(\mathbf{U}) \mathbf{w}_w^{\hat{O}}(n+1) \right\} \quad (\text{IV.73})$$

avec la partition $\begin{bmatrix} \mathbf{w}_s^{\hat{O}}(n) \\ \mathbf{w}_w^{\hat{O}}(n) \end{bmatrix}$ du vecteur $\mathbf{w}^{\hat{O}}(n)$ d'espérance de revenus totaux à partir de l'instant n , en $\mathbf{w}_s^{\hat{O}}(n) \in \mathbb{R}^{r_1}$ correspondant à la partie forte, et $\mathbf{w}_w^{\hat{O}}(n) \in \mathbb{R}^{r_2}$, à la partie faible.

Ceci permet la résolution de la commande optimale sur chaque partie, pour tous les types de problèmes terminaux. La commande optimale de la partie forte sera représentée par la matrice de dimension $(K \times r_1)$ suivante :

$$\mathbf{G}_s(n) = [g_s(n, 1) \ g_s(n, 2) \ \dots \ g_s(n, r_1)]$$

avec $g_s(n, i) = \hat{\mathbf{U}}(n, i)$, $\forall i = 1, 2, \dots, r_1$ et $\forall n = 0, 1, \dots$

Pour la partie faible, nous obtenons une commande optimale représentée par une matrice $(K \times r_2)$

$$\mathbf{G}_w(n) = [g_w(n, r_1+1) \ g_w(n, r_1+2) \ \dots \ g_w(n, r)]$$

avec $g_w(n, i) = \hat{\mathbf{U}}(n, i)$, $\forall i = r_1+1, r_1+2, \dots, r$ et $\forall n = 0, 1, \dots$

IV.4.3. Détermination de la commande quasi-optimale du système initial

La commande quasi-optimale du système initial sera obtenue à partir des commandes optimales des sous-systèmes :

$$\mathbf{G}(n) = [\mathbf{G}_s(n) \ \mathbf{G}_w(n)]$$

et l'évolution des revenus est approchée par les évolutions partielles :

$$\mathbf{w}^{\hat{O}}(n) = \hat{\mathbf{Y}} \begin{bmatrix} \mathbf{w}_s^{\hat{O}}(n) \\ \mathbf{w}_w^{\hat{O}}(n) \end{bmatrix}$$

Le même principe de découplage peut être appliquée pour les problèmes terminaux à horizon aléatoire ou à horizon fini éloigné (N très grand), ainsi que pour les problèmes avec cible.

Dans ce qui suit, nous illustrons la démarche proposée sur un exemple numérique.

IV.5. Exemple numérique

Soit la chaîne de Markov à commande formée de 4 états et représentée par la matrice de transition suivante (forme bilinéaire avec $K = 2$) :

$$\dagger(\mathbf{U}) = \begin{pmatrix} 0.9 & 0.04 & 0.03 & 0.03 \\ 0.01 & 0.90 & 0.04 & 0.05 \\ 0.35 & 0.35 & 0.01 & 0.29 \\ 0.50 & 0.29 & 0.20 & 0.01 \end{pmatrix} + \begin{pmatrix} 0.007 & -0.003 & -0.003 & -0.001 \\ -0.001 & 0 & 0.001 & 0 \\ 0.005 & -0.005 & 0.001 & -0.001 \\ 0.004 & -0.006 & 0.001 & 0.001 \end{pmatrix} \mathbf{u}_1 + \begin{pmatrix} 0.001 & -0.005 & 0.003 & 0.001 \\ -0.001 & -0.003 & 0 & 0.004 \\ 0.007 & -0.004 & -0.002 & -0.001 \\ 0.006 & -0.005 & -0.002 & 0.001 \end{pmatrix} \mathbf{u}_2$$

où nous avons noté : $\dagger(\mathbf{U}) = \dagger_0 + \dagger_1 \mathbf{u}_1 + \dagger_2 \mathbf{u}_2$

et soit la matrice de revenus :

$$\mathbf{R} = \begin{pmatrix} 0.202 & 0.201 & 0.199 & 0.2 \\ 0.201 & 0.201 & 0.2 & 0.199 \\ 0.201 & 0.201 & 0.202 & 0.201 \\ 0.199 & 0.199 & 0.201 & 0.201 \end{pmatrix}$$

indépendante de la commande.

Nous considérons $\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ comprise dans l'espace de commande Ω donné par les intervalles :

$$u_1 \in [0.1, 0.9] \text{ et } u_2 \in [0.01, 1]$$

Nous traitons le problème de commande avec horizon infini, dans lequel on cherche à atteindre un revenu maximum. La perte d'information en régime permanent est exprimée par un coefficient $\gamma = 0.7$. Nous allons faire les calculs pour un nombre de 150 transitions, en considérant une espérance nulle de revenu à l'infini :

$$\mathbf{w}() = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

L'équation d'optimalité de la chaîne présente la forme suivante :

$$\mathbf{w}(\mathbf{U}) = \max_{\mathbf{U} \in \mathcal{U}} \left\{ \text{diag}[\dagger(\mathbf{U}) \mathbf{R}^T] + \gamma \dagger(\mathbf{U}) \mathbf{w}(\mathbf{U}+1) \right\}$$

Nous obtenons ainsi l'évolution de l'espérance de revenu du système initial, représentée dans le tableau suivant :

	0	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	∞
\mathbf{e}_1	0.6723	0.6723	0.6723	0.6723	0.6723	0.6723	0.6723	0.6723	0.6723	0.6723	0.6723	0.6723	0.6722	0.6719	0.6590	0.2018	0
\mathbf{e}_2	0.6703	0.6703	0.6703	0.6703	0.6703	0.6703	0.6703	0.6703	0.6703	0.6703	0.6703	0.6703	0.6702	0.6699	0.6570	0.2010	0
\mathbf{e}_3	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6701	0.6572	0.2010	0
\mathbf{e}_4	0.6693	0.6693	0.6693	0.6693	0.6693	0.6693	0.6693	0.6693	0.6693	0.6693	0.6693	0.6693	0.6693	0.6690	0.6561	0.1994	0

avec la commande optimale correspondante :

- commande u_1

	0	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
\mathbf{e}_1	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9

e_2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
e_3	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
e_4	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9

- commande u_2

	0	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
e_1	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
e_2	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
e_3	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02
e_4	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02

Suite à l'analyse de la forme propre associée à la valeur propre 1 de la matrice de transition, nous concluons que la partie forte contient les états e_1 et e_2 et la partie faible les états e_3 et e_4 , quelle que soit la commande appliquée. Le coefficient de découplage des pondérations varie entre 0,07 et 0,09. L'évolution de ce coefficient en fonction de la commande, est représentée dans la figure IV.5 :

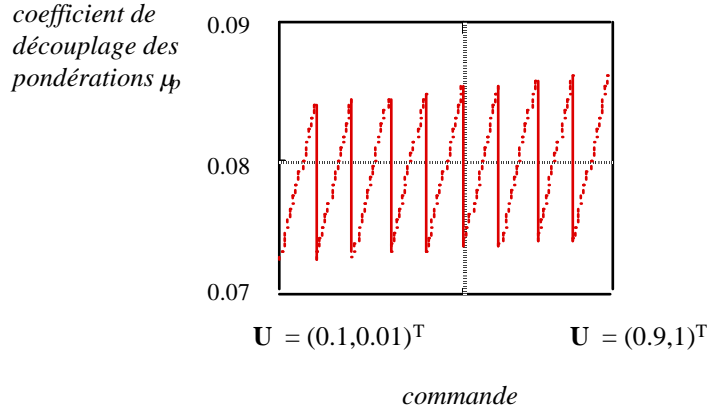


Fig.IV.5. Evolution du coefficient de découplage en fonction de la commande.

Pour l'étude de la partie forte, nous obtenons :

$$\mathbf{V} = \left[\begin{pmatrix} \mathbf{I} - \ddagger' & 0_w \end{pmatrix}^T \begin{pmatrix} \mathbf{I} - \ddagger' & 0_w \end{pmatrix} + \ddagger'_{1_w}^T \ddagger'_{1_w} + \ddagger'_{2_w}^T \ddagger'_{2_w} \right]^{-1} \left[\begin{pmatrix} \mathbf{I} - \ddagger' & 0_w \end{pmatrix}^T \ddagger'_{0_{w-s}} - \ddagger'_{1_w}^T \ddagger'_{1_{w-s}} - \ddagger'_{2_w}^T \ddagger'_{2_{w-s}} \right]$$

avec les notations suivantes :

$$\ddagger'_0 = \begin{pmatrix} \ddagger'_s & \ddagger'_{s-w} \\ \ddagger'_{w-s} & \ddagger'_w \end{pmatrix} = \begin{pmatrix} 0.9 & 0.04 & 0.03 & 0.03 \\ 0.01 & 0.90 & 0.04 & 0.05 \\ 0.35 & 0.35 & 0.01 & 0.29 \\ 0.50 & 0.29 & 0.20 & 0.01 \end{pmatrix}, \quad \ddagger'_1 = \begin{pmatrix} \ddagger'_s & \ddagger'_{s-w} \\ \ddagger'_{w-s} & \ddagger'_w \end{pmatrix} = \begin{pmatrix} 0.007 & -0.003 & -0.003 & -0.001 \\ -0.001 & 0 & 0.001 & 0 \\ 0.005 & -0.005 & 0.001 & -0.001 \\ 0.004 & -0.006 & 0.001 & 0.001 \end{pmatrix}$$

$$\dagger_2 = \begin{pmatrix} \dagger' & 2_s & \dagger' & 2_{s-w} \\ \dagger' & 2_{w-s} & \dagger' & 2_w \end{pmatrix} = \begin{pmatrix} 0.001 & -0.005 & 0.003 & 0.001 \\ -0.001 & -0.003 & 0 & 0.004 \\ 0.007 & -0.004 & -0.002 & -0.001 \\ 0.006 & -0.005 & 0.002 & 0.001 \end{pmatrix}$$

Ceci donne par remplacement :

$$\mathbf{V} = \begin{pmatrix} 0.5330 & 0.4670 \\ 0.6127 & 0.3873 \end{pmatrix}$$

donc

$$\dagger_s(\mathbf{U}) = \begin{pmatrix} 0.9344 & 0.0656 \\ 0.0915 & 0.9085 \end{pmatrix} + \begin{pmatrix} 0.0048 & -0.0048 \\ -0.0005 & 0.0005 \end{pmatrix} \mathbf{u}_1 + \begin{pmatrix} 0.0032 & -0.0032 \\ 0.0015 & -0.0015 \end{pmatrix} \mathbf{u}_2$$

En considérant les conditions finales suivantes :

$$\mathbf{w}_s() = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

et en utilisant l'équation d'optimalité de la partie forte :

$$\mathbf{w}_s^{\dagger}(\mathbf{U}) = \max_{\mathbf{U} \in \mathcal{Z}} \left\{ \text{diag} \left[\dagger_s(\mathbf{U}) \mathbf{R}_1^T \right] + \gamma \dagger_s(\mathbf{U}) \mathbf{w}_s^{\dagger}(\mathbf{U}+1) \right\}$$

où

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix} = \begin{pmatrix} 0.202 & 0.201 & 0.199 & 0.2 \\ 0.201 & 0.201 & 0.2 & 0.199 \\ 0.201 & 0.201 & 0.202 & 0.201 \\ 0.199 & 0.199 & 0.201 & 0.201 \end{pmatrix}$$

l'espérance de revenu de cette partie devient :

	0	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	∞
\mathbf{e}_1	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	0.6728	0.6724	0.6595	0.2019	0
\mathbf{e}_2	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6705	0.6701	0.6572	0.2010	0

avec la politique optimale suivante :

- commande \mathbf{u}_1

	0	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
\mathbf{e}_1	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
\mathbf{e}_2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.3

- commande \mathbf{u}_2

	0	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
\mathbf{e}_1	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
\mathbf{e}_2	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.05

Pour l'étude de la partie faible, nous obtenons :

$$\mathbf{Y} = \left[\left(\mathbf{I} - \dagger' \begin{matrix} 0_s \end{matrix} \right)^T \left(\mathbf{I} - \dagger' \begin{matrix} 0_s \end{matrix} \right) + \dagger' \begin{matrix} 1_s \end{matrix} \dagger' \begin{matrix} 1_s \end{matrix} + \dagger' \begin{matrix} 2_s \end{matrix} \dagger' \begin{matrix} 2_s \end{matrix} \right]^{-1} \left[\left(\mathbf{I} - \dagger' \begin{matrix} 0_s \end{matrix} \right)^T \dagger' \begin{matrix} 0_{s-w} \end{matrix} - \dagger' \begin{matrix} 1_s \end{matrix} \dagger' \begin{matrix} 1_{s-w} \end{matrix} - \dagger' \begin{matrix} 2_s \end{matrix} \dagger' \begin{matrix} 2_{s-w} \end{matrix} \right]$$

ce qui par remplacement donne :

$$\mathbf{Y} = \begin{pmatrix} 0.5001 & 0.4999 \\ 0.4998 & 0.5002 \end{pmatrix}$$

donc

$$\dagger_w(\mathbf{U}) = \begin{pmatrix} 0.3599 & 0.6401 \\ 0.595 & 0.405 \end{pmatrix} + \begin{pmatrix} 0.001 & -0.001 \\ 0.000002 & -0.000002 \end{pmatrix} \mathbf{u}_1 + \begin{pmatrix} -0.0005 & 0.0005 \\ -0.0015 & 0.0015 \end{pmatrix} \mathbf{u}_2$$

et par conséquent, en considérant

$$\mathbf{w}_w() = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

et en utilisant l'équation d'optimalité de la partie faible :

$$\mathbf{w}_w^Q(\mathbf{n}) = \max_{\mathbf{U} \in \mathcal{Z}} \left\{ \text{diag} \left[\dagger_w(\mathbf{U}) \mathbf{R}_{22}^T \right] + \gamma \dagger_w(\mathbf{U}) \mathbf{w}_w^Q(\mathbf{n}+1) \right\}$$

l'espérance de revenu optimal du système faible présente l'évolution suivante :

	0	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	∞
\mathbf{e}_3	0.6707	0.6707	0.6707	0.6707	0.6707	0.6707	0.6707	0.6707	0.6707	0.6707	0.6707	0.6707	0.6707	0.6704	0.6575	0.2014	0
\mathbf{e}_4	0.6704	0.6704	0.6704	0.6704	0.6704	0.6704	0.6704	0.6704	0.6704	0.6704	0.6704	0.6704	0.6704	0.6701	0.6572	0.2010	0

avec la commande optimale :

- commande \mathbf{u}_1

	0	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
\mathbf{e}_3	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
\mathbf{e}_4	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.1

- commande \mathbf{u}_2

	1	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150
\mathbf{e}_3	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
\mathbf{e}_4	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03

Les erreurs commises pour l'espérance de revenu sont :

0,0842 % pour l'état 1 ; 0,0367 % pour l'état 2 ; 0,0340 % pour l'état 3 ; 0,1624 % pour l'état 4

V. Conclusion

Dans ce chapitre, nous avons étudié la simplification du calcul de la commande des chaînes de Markov caractérisées par une formule fondamentale affine en la commande. Ce type de système fait partie des systèmes bilinéaires discrets, et présente un grand intérêt pratique.

Une première classe de méthodes développées est obtenue par application des techniques de perturbations singulières. Par la suite, le sous système lent est ramené sous une forme

bilinéaire et présente la propriété de stochasticté. De cette façon, le calcul de la commande optimale se fait uniquement sur cette partie du système. Les résultats de la partie lente sont réinjectés dans le système initial et donnent la commande optimale de la partie rapide.

La deuxième classe de méthodes étudiées dans ce chapitre est le découplage en régime permanent. Elle correspond à l'existence de la propriété de double échelle de pondération de la chaîne de Markov à commande. Les sous-systèmes stochastiques fort et faible sont ramenés sous une forme bilinéaire après découplage. Le calcul de la commande optimale s'applique sur chaque sous-système séparément. Cette technique de découplage donne de très bons résultats dans le cas des problèmes terminaux à horizon infini.