Vers l'analyse de problèmes transitoires pour les grands modèles Markoviens

Laurent Truffet
IRCyN (UMR-6597) - EMN

Plan de l'exposé

- Contexte/Motivations
- Modèles stochastiques / Modèles Markoviens
- Problématique de base
- Notion centrale: les cônes et leurs propriétés
  - Outil pour comparer
  - Invariance et agrégation faible des chaînes
- Résultat général : un nouveau type de couplage
- Synthèse

Deux grandes classes de problèmes

I. Les problèmes par nature transitoires :

- Fiabilité des systèmes

- Gestion des alarmes dans les sytèmes de production

- Détection de congestion dans les réseaux

II. Les problèmes d'optimisation :

Il faut explorer un ensemble de modèles paramétrisés

$\{$ paramètres caractérisant le modèle $\} =$ configuration

Il faut décider (rapidement) quelle configuration est (non) pertinente

Exemple: Dimensionnement de réseaux

Modèles aléatoires, modèles markoviens

- Analyse Quantitative de S.E.D
- Modèles Analytiques

  $\{$ Var. d'Etat $\} + \{$ Relations $\}$
- On étudie une quantité $\Theta(n)$ au cours du temps telle que

$$\forall n \; \Theta(n) = f((\chi(0), \ldots, \chi(n))) \tag{1}$$

- Modèles Stochastiques : $\chi(n)$ et $\Theta(n)$ sont des variables aléatoires

Modèles markoviens (homogènes)

On considère un processus $\{\chi(n), n \in I\!N\}$ à espace d'état $E = \{1, \ldots, \eta\}$ tel que :

(i) $\forall n,\ \chi(n+1) = \phi(\chi(n), U(n))$

(ii) $(U(n))_{n \in I\!N}$ est une suite de v.a i.i.d

(iii) $\forall n$, $\chi(n)$ et $U(n)$ sont indépendantes

<u>Autre représentation (par rapport au distributions)</u> :

$x(n)$:= distribution de $\chi(n)$, i.e., $x_k(n) = \ \mathbf{P}(\chi(n) = k)$

A matrice des probabilités de transition entre état, i.e.

$$\begin{aligned} A(i,j) = &\quad \mathbf{P}(\chi(n+1) = j | \chi(n) = i) \\ = &\quad \mathbf{P}(\phi(i, U(n)) = j). \end{aligned}$$

D'où le système caractérisant l'évolution du système :

$$\left\{ \begin{array}{l} x(0) \\ x(n+1) = x(n)\ A. \end{array} \right. \tag{2}$$

Un exemple pour fixer les idées

$\chi(n)$ nombre de clients dans la file d'attente suivante:

1) $(A(n))_{n\geq 0}$ arrivées i.i.d, $\forall n$, $0 \leq A(n) \leq a$
$b_i = \mathbf{P}(A(n) = i)$, $i = 0, \ldots, a$.

2) Un seul serveur, de temps de service 1, discipline FIFO.

3) Capacité limitée à M.

4) à 0, la file est vide.

$$\begin{array}{ll} \chi(0) = & 0 \text{ (p.s)} \\ \chi(n+1) = & \min(\ M, \chi(n) + (A(n)-1)_+\ ) \end{array}$$

ou bien

$$x(0) = (1, 0, \ldots, 0) \text{ et } A = \begin{pmatrix} b_0 & b_1 & \ldots & b_a & 0 \\ b_0 & b_1 & \ldots & b_a & 0 \\ 0 & b_0 & b_1 & \ldots & b_a \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & b_0 & \sum_{i=1}^{a} b_i \end{pmatrix}.$$

$x(n+1) = x(n)\ A$.

Pourquoi et comment calculer des bornes

$$\Theta(n) = f(\chi(n))$$

Trouver $Y_1(n)$, $Y_2(n)$, $\prec_i, i = 1, 2$ tels que :

$$f(Y_1(n)) \prec_1 f(\chi(n)) \prec_2 f(Y_2(n))$$

$f(Y_i(n))$ plus facile à calculer

$+$

Borne Erreur(n) = $f(Y_2(n)) - f(Y_1(n))$

Le borneur évolue dans un univers hostile

- " une proba. comprise à coup sûr entre -3 et $+50$: bien ...Laurent...bien"

- "Euh...$[0,1]$ pour une proba. ...c'est de la borne non?"

- beaucoup de réconfort : " Essaie, mais ça ne marchera jamais... "

- des phrases qui lassent : "salut, ça borne ?"

Problèmatique de base

Dynamique du système : modélisée par $\chi(.)$, à espace d'état $E = \{1, \ldots, \eta\}$

On considère que la caractéristique du sytème, $\Theta(.)$, est une fonction monotone de groupes d'états de la chaîne $\chi = (x(0), A)$.

On formalise de la façon suivante :

$$\forall n,\ \Theta(n) = g(\epsilon(\chi(0), \ldots, \epsilon(\chi(n))) \tag{3}$$

avec:

$g\ :\ E^{I\!N} \longrightarrow I\!R$, croissante au sens de l'ordre composantes par composantes,

$\epsilon\ :\ E \longrightarrow \Sigma = \{1, \ldots, N\}$, avec $N < \eta$, ϵ croissante surjective et telle que $(\epsilon^{-1}(I))_{\{I \in \Sigma\}}$ est une partition de E

En général $(\epsilon(\chi)(n))_{n \geq 0}$ n'est pas une chaîne de Markov homogène.

Idée générale: Trouver des chaînes $Y_1(.)$, $Y_2(.)$, un ordre stochastique $\prec$ tels que

$Y_1(.) \prec \epsilon(\chi)(.) \prec Y_2(.)$ implique $\Theta_1(.) \prec \Theta(.) \prec \Theta_2(.)$

Intérêt : $Y_i(.)$ à valeurs dans Σ !

Les cônes : une notion centrale

Les cônes et les comparaisons de distributions

$\mathcal{A}(\eta)$={ vecteurs de probabilités }, $\mathcal{A}(\eta) \subset I\!R^{\eta}$

1) Ordre partiel sur $\mathcal{A}(\eta)$

$\forall x, y, z \ \in \mathcal{A}(\eta)$

(R) $x \leq_{ord} x$
(T) $x \leq_{ord} y$ et $y \leq_{ord} z \Longrightarrow x \leq_{ord} z$
(AS) $x \leq_{ord} y$ et $y \leq_{ord} x \Longrightarrow x = y$.

2) La construction d'un ordre entre les distributions (donc entre les variables aléatoires)

Cône polyhédral: $\mathcal{K} = I\!R^{\eta}_{+} K$, K $(\eta \times l)$-matrice

Ordre: $y \geq_{\mathcal{K}} x \iff y - x \in \mathcal{K}$

Cône dual : $\mathcal{K}^* = \{y \in I\!R^{\eta} | y \ K \geq 0\}$

Ordre* : $y \geq_{\mathcal{K}^*} x \iff (y - x) \ K \geq 0$

Remarques:
1) On prend des cônes propres pour engendrés les ordres

2) En général ces cônes correspondent à des familles de fonctions, à des familles d'ensembles.

3) Pour des raisons numériques on s'intéresse aux ordres*

Les cônes et les comparaisons de chaînes

Soit $\mathcal{K}^*$ un cône propre comme dual d'un cône propre $\mathcal{K}$, i.e.:

$\mathcal{K}$ cône convexe fermé (par ex. polyhédral), interieur non vide, $\mathcal{K} \cap (-\mathcal{K}) = 0$.

Direction extrêmale: x est direction extrêmale ssi

$$x \geq_{\mathcal{K}} y \geq_{\mathcal{K}} 0 \Longrightarrow \exists \lambda > 0,\ y = \lambda\, x$$

Opérateur monotone: A une matrice qui laisse $\mathcal{K}$ invariant (i.e. $\mathcal{K}\ A \subset \mathcal{K}$)

A monotone $\Longleftrightarrow$ ($\forall x,\ xK \geq 0 \Longrightarrow xAK \geq 0$)

Lorsque le cône est *simplicial* alors A est K-monotone ssi

$$K^{-1}\ A\ K \geq 0.$$

Autre façon de l'écrire:

A K-monotone ssi $pK \leq qK \implies pAK \leq qAK$.

Remarque : si A est la matrice de transition de $(\chi(n))_{n\geq 0}$:

$$\chi(l) \leq_K \chi(l+1) \Longrightarrow \forall n \geq l,\ \chi(n) \leq_K \chi(n+1)$$

L'ordre stochastique fort, $\leq_{st}$

L'ordre $\leq_{st}$ est engendré par les fonctions <u>croissantes</u> de $I\!R^\eta$ vers $I\!R$

$$X \leq_{st} Y \text{ ssi } \forall f \;\; Ef(X) \leq Ef(Y)$$

Cas particulier$^\heartsuit$ $E = \{1, \ldots, \eta\}$:

$$1_{\{x \geq k\}}, k \in E$$

$X : p = (p_1, \ldots, p_\eta)$ et $Y : q = (q_1, \ldots, q_\eta)$

$$K_{st} = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ 1 & 1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \tag{4}$$

$$X \leq_{st} Y \text{ ssi } p\, K_{st} \leq q\, K_{st}$$

Pour fixer les idées...

$$E = \{1, 2, 3\}$$

$$K_{st} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

$p = (0.8, 0.1, 0.1)$ et $q = (0.3, 0.2, 0.5)$

$p\ K_{st} = (1.0, 0.2, 0.1)$

$q\ K_{st} = (1.0, 0.7, 0.5)$

donc

$$p \leq_{st} q$$

Comparaison des chaînes, même espace d'état

$(X(n))_{n\geq 0}$ et $(Y(n))_{n\geq 0}$ deux chaînes de Markov à espace d'états $E = \{1, \ldots, \eta\}$ de matrices de transition P et Q.

SI

1. $X(0) \leq_{st} Y(0)$

2. $\forall i \in E \; P(i, .) \leq_{st} Q(i, .)$

3. P ou Q est $\leq_{st}$ $-$monotone, c'est à dire $K_{st}^{-1} P K_{st} \geq 0$ ou $K_{st}^{-1} Q K_{st} \geq 0$

ALORS (Dailey , Keister , ...)

$$\forall n \; X(n) \leq_{st} Y(n)$$

et même PIRE (couplage de CMTD) :

$$\forall n \; (X(0), \ldots, X(n)) \leq_{st} (Y(0), \ldots, Y(n))$$

Les cônes et l'agrégation faible (*Weak Lumpability*)

$\chi = (x(0), A)$ CMTDH à valeurs dans $E = \{1, \ldots, \eta\}$
$\chi_\epsilon = (\epsilon(\chi(n)))_{n\geq 0}$ processus agrégé

Problème: à quelle condition χ_ϵ est une CMTDH à valeurs dans $\Sigma = \{1, \ldots, N\}$?

Soit encore : à quelle condition a t-on

$\mathbf{P}(\epsilon(\chi(n+1) = J | \epsilon(\chi(n)) = I, \ldots, \epsilon(\chi(0)) = I_0)) = A^\epsilon(I, J)$?

On utilise un résultat dérivé des études sur l'agrégation faible.

(WL1) Existence d'un cône invariant i.e.

$\exists\, \mathcal{C} = \oplus_{I=1}^N \mathcal{C}_I$ tel que :

$$\mathcal{C}\ A \subset \mathcal{C}$$

(WL2) Condition initiale, $x(0)$, judicieusement choisie !

Résultat central (Borne supérieure)

Données : $\chi(.) = (x(0), A)$ CMTDH esp. d'état $E = \{1, \ldots, \eta\}$ $\epsilon : E \longrightarrow \Sigma = \{1, \ldots, N\}$

$V_\epsilon = diag(\mathbf{1}^*_{\eta_I})$ $(N \times \eta)$-matrice $V(I, j) = 1_{\{\epsilon(j)=I\}}$, $I \in \Sigma$, $j \in E$.

$\forall n,\ \Theta(n) = g(\epsilon(\chi(0), \ldots, \epsilon(\chi(n)))$

g, ϵ croissantes

On énonce alors le résultat suivant:

1) Si on peut construire une matrice B $\leq_{st}$-monotone telle que $\forall i, A(i, .) \leq_{st} B(i, .)$

2) Si on peut trouver un cône $\mathcal{C} = \oplus_{I=1}^{N} \mathcal{C}_I$ et prouver qu'il existe une matrice $\widehat{B}$ tels que :

(a) $\mathcal{C}\ \widehat{B} \subset \mathcal{C}$.
(b) $\forall i,\ B(i, .) \leq_{st} \widehat{B}(i, .)$

3) Si on peut trouver une condition initiale $\widehat{x}(0)$ judicieusement choisie et telle que $x(0) \leq_{st} \widehat{x}(0)$

Alors il existe une CMTDH $Y = (\widehat{x}(0)\ V_\epsilon, A^\epsilon)$ à valeurs dans Σ telle que $\chi_\epsilon(.) \leq_{st} Y(.)$

ET DONC $\Theta(.) \leq_{st} \widehat{\Theta}(.)$
avec :

$$\forall n,\ \widehat{\Theta}(n) = g(Y(0), \ldots, Y(n)).$$

Borne Markovienne sur un processus de sortie de file d'attente

$\chi(.)$ CMTDH nombre de clients dans file d'attente

$E = \{1,2,3,4\}$, $\epsilon(\chi)(n) = 1 + 1_{\{\chi(n)\geq 2\}}$, $\Sigma = \{1,2\}$

$V_\epsilon = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$.

$x(0) = (1,0,0,0)$ et la matrice de transition est

$$A = \left(\begin{array}{c|ccc} 0.35 & 0.05 & 0.4 & 0.2 \\ \hline 0.35 & 0.05 & 0.4 & 0.2 \\ 0.0 & 0.35 & 0.05 & 0.6 \\ 0.0 & 0.0 & 0.35 & 0.65 \end{array}\right).$$

1) $B = A$ car A est $\leq_{st}$-monotone

2) On prend $\mathcal{C} = I\!R^4_+ \; diag((1), (\frac{1}{10}\,(1,2,7)))$
On explicite des conditions sous lesquelles on peut prouver l'existence d'une matrice $\widehat{B}$ vérifiant (2a) et (2b).

3) $\widehat{x}(0) = (1\ (1),\ 0\ (\frac{1}{10}\,(1,2,7)))$

$Y = ((1,0), A^\epsilon)$ est une borne Markovienne du processus de sortie, avec:

$$A^\epsilon = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{90} & \frac{89}{90} \end{pmatrix} \tag{5}$$

Remarque: en étudiant seulement $Y(.)$ on peut obtenir des bornes sur $\chi(.)$ également.

Synthèse

Situation des résultats obtenus

Par rapport à la théorie

$\leq_{st}$-comparaison de CMTDH + agrégation forte = couplage classique

$\leq_{st}$-comparaison de CMTDH + agrégation faible = nouvelle méthode de couplage

Par rapport à des applications potentielles

Il manque le développement de tous les aspects numériques

Problèmes, travaux en cours et futurs

En cours: les cônes invariants sont des demi-droites

Futurs: généralisation (1) à d'autres cônes invariants, (2) pour d'autres ordres stochastiques, (3) Briser la monotonie !

Problèmes

politique de choix des cônes ?

critères ?