Jeux matriciels

Bruno Bouzy

7 décembre 2005

Introduction

Ce document décrit ma compréhension des jeux matriciels. J'ai constaté que les jeux matriciels constituaient un formalisme utilisé pour le RLMAS. Je me suis donc intéressé aux jeux matriciels au travers des 5 articles [Littman 1994, 2001], [Hu & Wellmann 1998], [Bowling 2000], [Tesauro 2003].

Définition

Un *jeu matriciel* oppose n joueurs ou agents. A chaque unité de temps, un agent A^i choisit une action j. Pour simplifier, si #agent=2 et #action=2. Le jeu matriciel est défini par deux matrices de récompenses R_1 et R_2. $R^i_{j,k}$ est la récompense de l'agent k (k=1,2) s'il choisit l'action j (j=1,2) et que l'autre agent choisit l'action i. En général chaque valeur de j correspond à une ligne et chaque valeur de i correspond à une colonne. Si une matrice est l'opposée de l'autre on dit que le jeu est à *somme nulle*. Chaque joueur a pour but de maximiser sa récompense cumulée.

Exemples

La bataille des sexes

Un homme et une femme sont mariés ensemble et chaque soir ils choisissent d'aller à l'opéra ou au foot. La matrice du jeu est:

	Opéra	***Foot***
Opéra	2, 1	0, 0
Foot	0, 0	1, 2

Les deux matrices de récompenses ont été réunies dans le même tableau. Une cellule du tableau correspond donc aux récompenses respectives données à la femme (chiffre de gauche) et à l'homme (chiffre de droite) sachant que la femme a choisi l'action correspondant à la ligne de la cellule, et l'homme celle correspondant à la colonne. Dans cet exemple, si la femme et l'homme ont choisi tous les deux l'opéra, alors la femme aura un récompense 2 et l'homme une récompense 1. Cela exprime le fait que la femme aime plus l'opéra que l'homme n'aime l'opéra. Si la femme et l'homme ont choisi tous les deux le foot, alors la femme aura une récompense 1 et l'homme une récompense 2, exprimant que le fait que l'homme aime plus le foot que sa femme n'aime le foot. Si l'homme et la femme choisissent l'un le foot et l'autre l'opéra, ils ont tous les deux une récompense nulle, exprimant le fait qu'ils préfèrent passer la soirée ensemble plutôt que séparés. On voit que ce jeu est un jeu de coordination.

La poule mouillée

Les deux joueurs sont chacun assis dans leur voiture roulant en sens inverse l'une de l'autre sur la même route à une voie. Les deux voitures se rapprochent. Chaque joueur choisi une action: rester sur la même voie ou sortir c'est-à-dire craquer. La matrice du jeu est:

	rester	***craquer***
rester	-10, -10	+1, -1
craquer	-1, +1	0, 0

Une cellule du tableau contient deux valeurs : celle de gauche donnée au joueur choisissant une ligne du tableau et celle de droite au joueur choisissant une colonne. Si les deux joueurs choisissent de rester sur la route, ils se percutent et ont une « récompense » -10 chacun. Si un joueur reste et l'autre craque, celui qui est resté gagne +1 et celui qui a craqué, la poule mouillée, reçoit -1. Si les deux joueurs sortent, ils reçoivent 0.

La conduite

Les deux joueurs sont chacun assis dans leur voiture roulant en sens inverse l'une de l'autre sur la même route à *deux* voies. Les deux voitures se rapprochent. Chaque joueur choisi une voie: gauche ou droite. La matrice du jeu est:

	gauche	***droite***
gauche	0, 0	-10, -10
droite	-10, -10	0, 0

Une cellule du tableau contient deux valeurs : celle de gauche donnée au joueur choisissant une ligne du tableau et celle de droite au joueur choisissant une colonne. Si les deux joueurs choisissent tous les deux de rouler à gauche, ils se croisent et n'ont rien. Idem s'ils choisissent tous les deux de rouler à droite. En revanche, si un joueur choisi de rouler à gauche et l'autre à droite, ils se percutent et reçoivent -10 chacun.

Le dilemme du prisonnier

Les deux joueurs sont prisonniers et ils peuvent choisir une action: « coopérer » ou « trahir ». La matrice du jeu est:

	coop	***trahir***
coop	4, 4	0, 5
trahir	5, 0	3, 3

Si les deux joueurs choisissent tous les deux de coopérer, ils obtiennent 4. Ils reçoivent 3 s'ils choisissent de trahir. Si un joueur trahit et l'autre « coopère », le traître reçoit 5 et le coopérateur ne reçoit rien.

Pierre Papier Ciseaux

C'est un jeu à somme nulle. Les 2 joueurs peuvent choisir une action: «pierre», «papier» ou «ciseaux». La matrice du jeu est:

	pierre	***papier***	***ciseaux***
pierre	0	-1	+1
papier	+1	0	-1
ciseaux	-1	+1	0

La pierre bat les ciseaux qui battent le papier qui bat la pierre.

Equilibre de Nash

Il y a deux sortes d'équilibre de Nash: équilibre pur dans lequel chaque agent choisi son action de manière déterministe et un équilibre mixte dans lequel l'action de l'agent est choisie suivant une « stratégie ».

Equilibre pur

Dans le cas à N agents, un équilibre de Nash est un N-uple d'actions $(A_1^*, A_2^*, ..., A_N^*)$ tel que pour chaque agent i et chaque action A_i choisie par l'agent i, on a:

$$R_i((A_1^*, A_2^*, ..., A_{i-1}^*, A_i^*, A_{i+1}^*, ..., A_N^*) \geq R_i((A_1^*, A_2^*, ..., A_{i-1}^*, \mathbf{A_i}, A_{i+1}^*, ..., A_N^*)$$

$(A_1^*, A_2^*, ..., A_N^*)$ s'appelle un équilibre pur. Cela signifie que, au point d'équilibre, si tous les agents gardent la même action sauf un agent, alors celui-ci voit sa récompense diminuer. Donc à l'équilibre, tous les agents ont intérêt à garder la même action [Hu & Wellmann 1998], [Nash 1950].

Equilibre mixte

Ici un agent choisit, non pas une action déterministe, mais une « stratégie ». Une stratégie est un N-uple de probabilités Pj = (Pij) fixé, tel que Pij est la probabilité que l'agent i choisisse l'action j.

Dans le cas à N agents, un équilibre de Nash est un N-uple d'actions $(P_1^*, P_2^*, ..., P_N^*)$ tel que pour chaque agent i et chaque stratégie P_i suivie par l'agent i, on a:

$$R_i((P_1^*, P_2^*, ..., P_{i-1}^*, P_i^*, P_{i+1}^*, ..., P_N^*) \geq R_i((P_1^*, P_2^*, ..., P_{i-1}^*, \mathbf{P_i}, P_{i+1}^*, ..., P_N^*)$$

$(P_1^*, P_2^*, ..., P_N^*)$ s'appelle un équilibre mixte. Cela signifie que, au point d'équilibre, si tous les agents gardent la même stratégie sauf un agent, alors celui-ci voit sa récompense diminuer en moyenne. Donc à l'équilibre, tous les agents ont intérêt à garder la même stratégie.

Equilibre inimical

[Littman 2001] Dans le cas à N agents, un équilibre de Nash *inimical* est un N-uple d'actions $(P_1^*, P_2^*, ..., P_N^*)$ tel que pour chaque agent i et chaque stratégie P_i suivie par l'agent i, on a:

$$R_i((P_1^*, P_2^*, ..., P_{i-1}^*, P_i^*, P_{i+1}^*, ..., P_N^*) \leq R_i((P_1, P_2, ..., P_{i-1}, \mathbf{P_i^*}, P_{i+1}, ..., P_N)$$

Un jeu matriciel n'a pas nécessairement un équilibre inimical. Dans un jeu à somme nulle, tout équilibre est inimical.

Equilibre de coordination

Dans le cas à N agents, un équilibre de Nash *de coordination* est un N-uple d'actions (P_1*, P_2*, ..., P_N*) tel que pour chaque agent i et chaque stratégie P_i suivie par l'agent i, on a:

$$R_i((P_1^*, P_2^*, ..., P_{i-1}^*, P_i^*, P_{i+1}^*, ..., P_N^*) = \max_{aj} R_i((a_1, a_2, ..., a_N)$$

Un tel équilibre n'existe pas toujours.

Exercices

La bataille des sexes
Montrer que (opéra, opéra) et (foot, foot) sont des équilibres purs.
Montrer que ((2/3, 1/3), (1/3, 2/3)) est un équilibre mixte.

La poule mouillée
Montrer que (rester, craquer) et (craquer, rester) sont des équilibres purs.
Montrer que ((1/10, 9/10), (1/10, 9/10)) est un équilibre mixte.

La conduite
Montrer que (gauche, gauche) et (droite, droite) sont des équilibres purs.
Montrer que ((1/2, 1/2), (1/2, 1/2)) est un équilibre mixte.

Le dilemme du prisonnier
Montrer que (trahir, trahir) est un équilibre pur.
Montrer que ((2/3, 1/3), (2/3, 1/3)) est un équilibre mixte.

Pierre papier ciseaux
Montrer qu'il n'y a pas d'équilibre pur.
Montrer que ((1/3, 1/3, 1/3), (1/3, 1/3, 1/3)) est un équilibre mixte.

Références

[Hu & Wellmann 1998], Junling Hu, Michael Wellman, "Multiagent reinforcement learning: Theoretical Framework and an algorithm", ICML, pages 242-250, 1998.

[Littman 2001], Michael Littman, "Friend-or-Foe Q-learning in General Sum Games", ICML, pages ?-?, 2001.

[Nash 1950], John Nash, "Non cooperative games", *Annals of Mathematics*, 51, pages 286-295, 1950.

[Nash 1950], John Nash, "Equilibrium points in n-person games", *PNAS*, 36, 48, 49, Reprinted in H. W. Kuhn (ed.), 1997, Classics in game theory, Princeton, NJ, Princeton University Press.