Intelligence Artificielle et Jeux (LU3IN025)

Cours 3 :
1) La programmation linéaire comme outil de résolution
2) Manipulabilité

Nawal Benabbou

Licence Informatique - Sorbonne Université

2024-2025

SCIENCES SORBONNE UNIVERSITÉ

Un autre outil de résolution : la programmation linéaire

Dans le cours précédent

Nous avons vu que la théorie des graphes pouvait être utilisée pour déterminer des mariages efficaces/équitables.

Dans ce cours

Nous allons voir comment utiliser la programmation linéaire pour résoudre ces problèmes. Utiliser la programmation linéaire permettra d'ajouter d'autres types de contraintes au problème, sans avoir à changer la méthode de résolution.

Programmation linéaire

La programmation linéaire, aussi appelée optimisation linéaire, est la discipline qui étudie des problèmes d'optimisation dans lesquels la fonction à optimiser ainsi que les contraintes sont décrites par des fonctions linéaires en les variables du problème.

Un autre outil de résolution : la programmation linéaire

Un petit exemple introductif

Un étudiant à 2 examens à passer. Pour obtenir le diplôme, il doit avoir au moins 8/20 à chaque examen, et une moyenne supérieure ou égale à 10/20. Selon son programme de révision, il obtient les notes suivantes :

- 5/20 au premier examen sans révision, puis $+2$ points par heure de révision, avec un maximum de 5h de révision.
- 6/20 au second examen sans révision, puis $+1$ points par heure de révision, avec un maximum de 8h de révision.

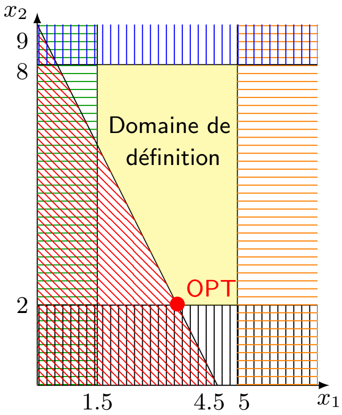
Objectif : minimiser le temps de révision.

Modélisation du problème sous la forme d'un programme linéaire

Variables : x_i le nombre d'heures à réviser le $i^{\text{ème}}$ examen, avec $i \in \{1, 2\}$.
Fonction objectif : $\min \quad x_1 + x_2$
Contraintes :

$$\begin{aligned} 5 + 2x_1 &\geq 8 && \text{(au moins 8/20 au 1er examen)} \\ 6 + x_2 &\geq 8 && \text{(au moins 8/20 au 2ème examen)} \\ (5 + 2x_1 + 6 + x_2)/2 &\geq 10 && \text{(au moins 10/20 de moyenne)} \\ 0 \leq x_1 \leq 5,\ \ 0 \leq x_2 &\leq 8 && \text{(le nombre d'heures autorisé)} \end{aligned}$$

Résolution graphique (cas avec seulement deux variables)



Contraintes :

$$
\begin{aligned}
5 + 2x_1 &\geq 8 \\
6 + x_2 &\geq 8 \\
(5 + 2x_1 + 6 + x_2)/2 &\geq 10 \\
0 \leq x_1 &\leq 5 \\
0 \leq x_2 &\leq 8
\end{aligned}
$$

Théorème

Si le programme linéaire admet une solution optimale, alors un des sommets du domaine de définition est optimal.

Détermination d'une solution optimale pour un programme linéaire

On peut calculer la valeur de la fonction objectif de tous les sommets du domaine, et on retourne le sommet qui a la meilleure valeur.

$\rightarrow$ Ici, le sommet $(3.5, 2)$ à l'intersection des droites $(5 + 2x_1 + 6 + x_2)/2 = 10$ et $6 + x_2 = 8$ est une solution optimale.

Modélisation pour les mariages efficaces

Programme linéaire : formulation générale

Exprimer le problème sous la forme :

$$\max \text{ (ou } \min) \quad f(x)$$
$$\text{sous contraintes : } \quad x \in D$$

où $x = (x_1, \ldots x_m)$ sont des variables réelles, D est l'ensemble de définition caractérisé par des contraintes linéaires et $f(x)$ est la fonction objectif (linéaire).

Modélisation de la recherche du mariage maximisant l'utilité totale (efficacité)

Variables : Pour tout $i \in \{1, \ldots, n\}$ et $j \in \{1, \ldots, n\}$, on définit la variable x_{ij} telle que $x_{ij} = 1$ si l'homme h_i est marié avec la femme f_j et $x_{ij} = 0$ sinon.

Fonction objectif : $\max \quad \sum_{i=1}^{n} \sum_{j=1}^{n} \; x_{ij}(u_{h_i}(f_j) + u_{f_j}(h_i))$

Contraintes :

$\sum_{j=1}^{n} x_{ij} = 1, \; \forall i \in \{1, \ldots, n\}$ (chaque homme est marié à une seule femme)

$\sum_{i=1}^{n} x_{ij} = 1, \; \forall j \in \{1, \ldots, n\}$ (chaque femme est mariée à un seul homme)

$x_{ij} \in \{0, 1\}, \; \forall i, j \in \{1, \ldots, n\}$ (domaine des variables)

Remarque : tout est linéaire, sauf les contraintes "$x_{ij} \in \{0, 1\}$" (contraintes d'intégrité). Ce programme est un **programme linéaire en nombres entiers**.

Retour sur l'exemple

$u_h(f) + u_f(h)$:

	A	B	C
X	4	0	2
Y	1	4	0
Z	2	3	2

Programme linéaire en nombre entiers correspondant :

$$
\begin{array}{rl}
\max & 4x_{XA} + 2x_{XC} + x_{YA} + 4x_{YB} + 2x_{ZA} + 3x_{ZB} + 2x_{ZC} \\
\text{s.c.} & x_{XA} + x_{XB} + x_{XC} = 1 \\
& x_{YA} + x_{YB} + x_{YC} = 1 \\
& x_{ZA} + x_{ZB} + x_{ZC} = 1 \\
& x_{XA} + x_{YA} + x_{ZA} = 1 \\
& x_{XB} + x_{YB} + x_{ZB} = 1 \\
& x_{XC} + x_{YC} + x_{ZC} = 1 \\
& x_{XA}, x_{XB}, x_{XC}, x_{YA}, x_{YB}, x_{YC}, x_{ZA}, x_{ZB}, x_{ZC} \in \{0, 1\}
\end{array}
$$

Modélisation pour les mariages équitables

On souhaite maintenant modéliser le problème de mariage équitable sous forme de programme linéaire (en nombres entiers).

Observation

En utilisant la modélisation précédente, chercher un mariage équitable revient à remplacer la fonction objectif précédente par la suivante :

$$\max \min\{\min_{i=1\dots n} \sum_{j=1}^{n} x_{ij} u_{h_i}(f_j), \min_{j=1\dots n} \sum_{i=1}^{n} x_{ij} u_{f_j}(h_i)\}$$

$\Rightarrow$ Cette fonction objectif n'est pas linéaire en les variables.

Dans le cours précédent, nous avons vu que, lorsque les utilités sont des scores de Borda, on peut déterminer un mariage équitable par le biais des problèmes intermédiaires suivants : pour un k fixé, est-il possible de marier tout le monde avec une personne dans ses k premiers choix ?

Question :

Pour un k fixé, est-il possible de modéliser le problème intermédiaire sous forme de programme linéaire ?

Modélisation pour les mariages équitables

Modélisation du problème intermédiaire avec k fixé

Soit E l'ensemble des paires (i, j) telles que h_i est dans les k premiers choix de f_j et f_j est dans les k premiers choix de h_i.

Variables : Pour tout $(i, j) \in E$, on définit la variable x_{ij} telle que $x_{ij} = 1$ si l'homme h_i est marié avec la femme f_j et $x_{ij} = 0$ sinon.

Fonction objectif : $\max \ \sum_{(i,j)\in E} \ x_{ij}$

Contraintes :

$\sum_{(i,j)\in E} x_{ij} \leq 1, \forall i \in \{1, \ldots, n\}$ (un homme est avec au plus une femme)

$\sum_{(i,j)\in E}^{n} x_{ij} \leq 1, \forall j \in \{1, \ldots, n\}$ (une femme est avec au plus un homme)

$x_{ij} \in \{0, 1\}, \forall (i, j) \in E$ (domaine des variables)

Analyse du résultat

Si la valeur de la fonction objectif à l'optimum est égale à n, alors il existe un mariage parfait tel que chaque personne est mariée avec quelqu'un dans ses k premiers choix. Si en revanche la valeur optimale est strictement inférieure à n, alors il n'existe pas de tel mariage parfait.

Exemple

Listes des hommes :

	1	2	3
X	A	C	B
Y	B	A	C
Z	B	A	C

Listes des femmes :

	1	2	3
A	X	Z	Y
B	Y	Z	X
C	Z	X	Y

Pour $k = 2$, on a $E = \{(X, A), (X, C), (Y, B), (Z, A), (Z, B)\}$, ce qui donne le programme linéaire en nombres entiers suivant :

$$
\begin{aligned}
\max \quad & x_{XA} + x_{XC} + x_{YB} + x_{ZA} + x_{ZB} \\
\text{s.c.} \quad & x_{XA} + x_{XC} \leq 1 \\
& x_{YB} \leq 1 \\
& x_{ZA} + x_{ZB} \leq 1 \\
& x_{XA} + x_{ZA} \leq 1 \\
& x_{YB} + x_{ZB} \leq 1 \\
& x_{XC} \leq 1 \\
& x_{XA}, x_{XC}, x_{YB}, x_{ZA}, x_{ZB} \in \{0, 1\}
\end{aligned}
$$

Modélisation plus générale pour les mariages équitables

Peut-on résoudre le problème sans passer par le problème intermédiaire ?

Réponse : OUI, en utilisant une technique classique de linéarisation du $\min$.

Variables : Pour tout $i \in \{1, \ldots, n\}$ et $j \in \{1, \ldots, n\}$, on définit la variable x_{ij} telle que $x_{ij} = 1$ si l'homme h_i est marié avec la femme f_j et $x_{ij} = 0$ sinon. On ajoute une variable z représentant le $\min$ des utilités.

Fonction objectif : $\max \ z$

Contraintes :

$z \leq \sum_{j=1}^{n} u_{h_i}(f_j) x_{ij}, \forall i \in \{1, \ldots, n\}$ (chaque homme a une utilité d'au moins z)

$z \leq \sum_{i=1}^{n} u_{f_j}(h_i) x_{ij}, \forall j \in \{1, \ldots, n\}$ (chaque femme a une utilité d'au moins z)

$\sum_{j=1}^{n} x_{ij} = 1, \ \forall i \in \{1, \ldots, n\}$ (chaque homme est marié à une seule femme)

$\sum_{i=1}^{n} x_{ij} = 1, \ \forall j \in \{1, \ldots, n\}$ (chaque femme est mariée à un seul homme)

$z \geq 0, \ x_{ij} \in \{0, 1\}, \ \forall i, j \in \{1, \ldots, n\}$ (domaine des variables)

Pourquoi cela fonctionne ?

Intuition : on maximise z sous contrainte que z soit inférieure à l'utilité de chaque personne (z sera donc égale à la plus petite de ces utilités à l'optimum).

Cette méthode fonctionne avec les scores de Borda, ou toute autre fonction d'utilité (contrairement à la méthode précédente).

Retour sur l'exemple

Listes des hommes :

	1	2	3
X	A	C	B
Y	B	A	C
Z	B	A	C

Listes des femmes :

	1	2	3
A	X	Z	Y
B	Y	Z	X
C	Z	X	Y

$$
\begin{array}{rl}
\max & z \\
\text{s.c.} & z \leq 2x_{XA} + x_{XC} \\
& z \leq x_{YA} + 2x_{YB} \\
& z \leq x_{ZA} + 2x_{ZB} \\
& z \leq 2x_{XA} + x_{ZA} \\
& z \leq 2x_{YB} + x_{ZB} \\
& z \leq x_{XC} + 2x_{ZC} \\
& x_{XA} + x_{XB} + x_{XC} = 1 \\
& x_{YA} + x_{YB} + x_{YC} = 1 \\
& x_{ZA} + x_{ZB} + x_{ZC} = 1 \\
& x_{XA} + x_{YA} + x_{ZA} = 1 \\
& x_{XB} + x_{YB} + x_{ZB} = 1 \\
& x_{XC} + x_{YC} + x_{ZC} = 1 \\
& z \geq 0 \\
& x_{XA}, x_{XB}, x_{XC}, x_{YA}, x_{YB}, x_{YC}, x_{ZA}, x_{ZB}, x_{ZC} \in \{0, 1\}
\end{array}
$$

Et pour les colocataires ? les internes-hôpitaux ?

Modélisation pour le problème des colocataires

Pas de difficulté particulière, on définit une variable booléenne x_{ij} pour toutes les paires d'individus (i, j), avec $i, j \in \{1, \ldots, 2n\}$, et tels que $i < j$ (pour éviter les redondances).

Modélisation pour le problème des internes-hôpitaux

Pas de difficulté particulière, on définit une variable booléenne x_{ij} pour chaque paire (interne, hôpital), c'est à dire pour chaque paire (i, j) avec $i \in \{1 \ldots, n\}$ et $j \in \{1 \ldots, m\}$. De plus, il faut ajouter une contrainte pour chaque hôpital pour ne pas dépasser sa capacité :

$$\sum_{i=1}^{n} x_{ij} \leq C_j, \forall j \in \{1, \ldots, m\}$$

Et pour les mariages stables ?

Observation

Pour qu'une paire (h_p, f_q) ne soit pas instable, il faudrait imposer que :

$$x_{pq} = 1 \text{ ou } \left(\sum_{j=1}^{n} u_{h_p}(f_j)x_{pj} > u_{h_p}(f_q) \text{ ou } \sum_{i=1}^{n} u_{f_q}(h_i)x_{iq} > u_{f_q}(h_p)\right)$$

$\Rightarrow$ Cette contrainte n'est pas linéaire (à cause des "ou logique").

Linéarisation de la contrainte de stabilité

Pour garantir la stabilité d'une solution, il suffit d'imposer la contrainte linéaire suivante pour chaque paire (h_p, f_q) :

$$x_{pq} + \sum_{f_j \in F_p(q)} x_{pj} + \sum_{h_i \in H_q(p)} x_{iq} \geq 1$$

où $F_p(q)$ est l'ensemble des femmes que l'homme h_p préfère à la femme f_q, et $H_q(p)$ est l'ensemble des hommes que la femme f_q préfère à l'homme h_p.

Pourquoi cela fonctionne ?

Si $x_{pq} = 1$, alors la contrainte est vraie (OK car en couple donc pas instable). Si $x_{pq} = 0$, alors la contrainte est vraie ssi $\sum_{f_j \in F_p(q)} x_{pj} + \sum_{h_i \in H_q(p)} x_{iq} \geq 1$, c-à-d ssi au moins l'un des deux est en couple avec quelqu'un qu'il préfère à l'autre (c-à-d quand la paire n'est pas instable).

Et pour les mariages stables ?

Question

Pourquoi utiliser la programmation linéaire pour les mariages stables alors que, dans le cours 1, nous avons vu que l'algorithme de Gale-Shapley permet de déterminer rapidement un mariage stable ?

Une réponse possible

En TD, nous avons vu que le nombre de mariages stables peut être très grand (parfois même exponentiel...). Avec la programmation linéaire, on peut par exemple déterminer celui qui maximise la somme des utilités ou celui qui maximise le min des utilités. En effet, il suffit de choisir la fonction objectif qui nous intéresse, et d'imposer la stabilité via l'ajout de contraintes linéaires.

$\Rightarrow$ De manière générale, la programmation linéaire est un outil puissant, permettant de résoudre de nombreux problèmes avec des contraintes diverses (pourvu que l'on arrive à modéliser le problème avec des contraintes et une fonction objectif linéaires).

Résolution d'un programme linéaire

Une fois le problème modéliser sous forme de programme linéaire, il faut le résoudre pour trouver la solution optimale.

Résolution d'un programme linéaire avec des variables réelles

Pour les programmes linéaires avec seulement deux variables, on peut utiliser la méthode de résolution graphique (voir transparent 4). Pour les cas avec au moins trois variables, il existe des méthodes de résolution très efficaces (comme le "simplexe"). Ces méthodes sont implémentées par des solveurs existants (comme Gurobi, GLPK ou encore CPLEX).

Résolution d'un programme linéaire en nombres entiers

Quand certaines variables sont entières, la méthode graphique ne fonctionne pas toujours (elle fonctionne uniquement si le meilleur sommet est entier). De manière générale, le problème devient beaucoup plus compliqué avec des variables entières, mais d'autres méthodes de résolution ont été développées et implémentées par des solveurs existants.

→ En TP, vous apprendrez à utiliser l'un de ces solveurs (Gurobi).

→ En master ANDROIDE, les méthodes de résolution classiques seront vues en détail (exemple : le "simplexe", "branch and bound", etc).

Véracité et Manipulation

Un individu peut-il avoir intérêt à mentir dans l'algorithme de Gale-Shapley ?

Définition : Algorithme/procédure à véracité garantie (non manipulable)

Un algorithme est dit "à véracité garantie" s'il n'existe pas de situation où un agent à intérêt à mentir sur ses données, même en connaissant les données des autres personnes. Formellement, un algorithme est dit "à véracité garantie" si et seulement si, pour tout ensemble de m agents, dont les vraies données sont respectivement $(d_1^*, \ldots, d_m^*)$, il n'existe aucune paire agent-données (i, d_i) telle que l'agent i préfère la solution retournée par l'algorithme avec $(d_1^*, \ldots, d_{i-1}^*, d_i, d_{i+1}^*, \ldots, d_m^*)$ à la solution retournée avec les vraies données $(d_1^*, \ldots, d_i^*, \ldots, d_m^*)$.

Un exemple : véracité garantie et élection présidentielle

Question :

On considère le mécanisme d'élection à la présidence de la république française. Ce mécanisme est-il à véracité garantie ? NON. Au second tour, il n'y a que deux candidats, donc on a évidemment envie de voter pour son candidat préféré. Par contre, au premier tour, il y a ce que l'on appelle le vote utile.

Exemple 1 : vos préférences sont A $>$ B $> \ldots >$ E $>$ F. Les sondages disent qu'il y a de fortes chances d'avoir "E face à F" au 2ème tour. De plus, B est en 3ème position dans les sondages, alors que A est dernier. Dans ce cas, vous pouvez avoir envie de voter pour B au lieu de A, pour essayer d'éviter le pire.

Exemple 2 : supposons maintenant que A est en tête dans les sondages, et que B et F sont au coude à coude pour la 2ème position. Dans ce cas, vous pouvez avoir envie de voter B au lieu de A pour éviter le pire au 2ème tour. Mais, si les sondages vous informent aussi que A perdrait face à B au 2ème tour, alors que A gagnerait face à F, vous pouvez avoir envie de voter pour F au 1er tour.

Pouvez-vous proposer un règle électorale qui ne soit pas manipulable ?

Théorème [Gibbart-Satterthwaite] (pour le cas avec au moins 3 candidats)

Considérons une règle de vote qui assure que chaque candidat peut être élu. Si la règle de vote est non-manipulable, alors c'est forcément une dictature.

Véracité et Mariage Stable

Question :

Est-ce que l'algorithme de Gale-Shapley est à véracité garantie ? Non.

Vraies préférences des hommes :

	1	2	3
Xavier	Bea	Amy	Claire
Yohan	Amy	Bea	Claire
Zach	Amy	Bea	Claire

Vraies préférences des femmes :

	1	2	3
Amy	Xavier	Zach	Yohan
Bea	Zach	Xavier	Yohan
Claire	Xavier	Yohan	Zach

Dans ce problème, il n'existe que deux mariages stables :

- le mariage homme-optimal : H-OPT = (X-B, Y-C, Z-A),
- et le mariage femme-optimal : F-OPT = (X-A, Y-C, Z-B).

Si l'algorithme retourne le mariage homme-optimal, alors en transmettant :

Amy : Xavier Yohan Zach

Amy améliore sa situation puisque F-OPT devient l'unique mariage stable.
Si l'algorithme retourne le mariage femme-optimal, alors en transmettant :

Xavier : Bea Claire Amy

Xavier améliore sa situation puisque H-OPT devient l'unique mariage stable.

Pouvez-vous donner un algorithme non manipulable pour les mariages stables ?

Véracité et Mariage Stable

Théorème [Roth 1982]

Pour le problème de mariage stable, il n'existe pas d'algorithme de recherche de mariage stable qui soit à véracité garantie.

Théorème [Roth 1982]

L'algorithme de Gale-Shapley en faveur des hommes est tel qu'aucun homme n'a intérêt à mentir sur ses préférences (véracité garantie "côté homme"). De même, quand ce sont les femmes qui font les propositions, l'algorithme est à véracité garantie "côté femme".

Remarque

Un résultat pas si négatif pour l'algorithme de Gale-Shapley car :

- La véracité garantie sur un seul côté est suffisant quand l'autre côté n'est pas stratégique. Par exemple, quand on affecte des tâches à des agents, les tâches ne vont pas se plaindre.
- La manipulation reste un exercice difficile, notamment quand le nombre de personnes impliquées est très grand.
- Comment manipuler sans connaître les données des autres ?