Université Côte d'Azur

Projet Processus markoviens

Année 2024

Secretary problem

Auteurs:
Man David
Mathis Chabaud
Kouame Gerard Kra
Seyedi Seyed Hossein

May 31, 2024

Sommaire

Introduction

1 Modélisation du problème

0.1 Historique du problème

Le problème de la secrétaire a apparemment été introduit en 1949 par Merrill M. Flood, qui l'a appelé le problème de la fiancée dans une conférence qu'il a donnée cette année-là. Il y fit référence à plusieurs reprises au cours des années 1950, par exemple lors d'une conférence à Purdue le 9 mai 1958, et il devint finalement largement connu dans le folklore bien que rien n'ait été publié à l'époque. En 1958, il envoie une lettre à Leonard Gillman, avec copie à une douzaine d'amis dont Samuel Karlin et J. Robbins, décrivant une preuve de la stratégie optimale, avec une annexe de R. Palermo qui prouve que toutes les stratégies sont dominées par une stratégie du formulaire "rejeter le premier p sans condition, puis accepter le candidat suivant qui est meilleur". La première publication fut apparemment celle de Martin Gardner dans Scientific American, en février 1960. Il en avait entendu parler par John H. Fox Jr. et L. Gerald Marnie, qui avaient posé indépendamment un problème équivalent en 1958 ; ils l'appelaient le "jeu de Googol". Fox et Marnie ne connaissaient pas la solution optimale ; Gardner a demandé conseil à Leo Moser, qui (avec J. R. Pounder) a fourni une analyse correcte pour publication dans le magazine. Peu de temps après, plusieurs mathématiciens ont écrit à Gardner pour lui parler du problème équivalent qu'ils avaient entendu via la vigne, et dont tous peuvent très probablement être attribués aux travaux originaux de Flood. (wikipedia).

0.2 Contexte du problème

Imaginez un administrateur qui souhait recruter le meilleur secrétaire de N candidats (classable)pour un post. Les candidats sont interviewés un par un successivement dans un ordre aléatoire. Une décision concernant chaque candidat doit être pris immédiatement après l'entretien. Une fois rejeté, un candidat ne peut pas être rappelé. Au cours de l'entretien, l'administrateur obtient des informations nécessaires pour classer les candidats, mais n'est pas conscient de la qualité des candidats encore invisibles. La question est de trouver le stratégie optimale (règle d'arrêt) pour maximiser la probabilité de sélectionner le meilleur candidat. Si on peut décider à la fin, cela peut être résolu pour un simple d'algorithme de sélection du maximum consistant à suivre le maximum en cours et à sélectionner le maximum global à la fin. La difficulté est que la décision doit être prise immédiatement. (Thomson, Jonny (21 April 2022). "Mathematicians suggest the "37

0.3 Formulation du problème

Bien qu'il existe de nombreuses variantes, le problème fondamental peut être énoncé comme suit:

* Il y a un seul poste à pouvoir.
* Il y a N candidats et la valeur de N est inconnu.
* Les candidats, vus tous ensemble, peuvent être classés sans ambiguïté du meilleur au pire.
* Les candidats sont interrogés séquentiellement dans un ordre aléatoire, chaque ordre étant également probable.
* Immédiatement après un entretien, le candidat interviewé est soit accepté, soit rejeté, et la décision est irrévocable.
* La décision d'accepter ou de rejeter un candidat ne peut être fondée que sur le rang relatif des candidats interviewés jusqu'à présent.
* L'objectif de la solution générale est d'avoir la plus grande probabilité de sélectionner le meilleur candidat de l'ensemble du groupe. Cela revient à maximiser le gain attendu, le gain étant défini comme étant un pour le meilleur candidat et zéro sinon.

2 Règle d'arrêt optimal

La stratégie optimale pour le problème est une règle d'arrêt. Temps d'arrêt ou règle d'arrêt est une variable aléatoire dont la valeur est interprétée comme le moment auquel un processus stochastique donné présente un certain comportement intéressant. Un temps d'arrêt est souvent défini par une règle d'arrêt, un mécanisme permettant de décider de poursuivre ou d'arrêter un processus en fonction de la position actuelle et des événements passés, et qui conduira presque toujours à une décision d'arrêt à un moment donné.

0.1 Récurrence

Dans cette session, nous détaillons la formulation de la stratégie optimale. Soient $S = 0, 1$ l'ensemble des états et N le nombre de candidats disponibles: 1 signifie que *le candidat actuel est le meilleur (car de rang plus proche de* 1*)* et 0 signifie que *le candidat précédent était le meilleur* . Quelque soit les états, on définit l'action Q par *choisir le candidat actuel* (i.e attribuer l'offre à l'actuel candidat) et l'action C par *ne pas choisir le candidat actuel et poursui*
La récompense est perçu seulement à l'arrêt; c'est-à-dire choisir l'action Q. Ainsi, les actions Q et C correspondent à la probabilité de choisir le meilleur candidat. Sous les hypothèses ci-dessus, on définit le coût continu par $f_t(s) = 0, \forall s \in \{0, 1\}$, la récompense à l'arrêt par $g_t(s) = \begin{cases} 0 & \text{si } s = 0 \\ \frac{1}{N} & \text{si } s = 1 \end{cases}$ et la récompense terminale par $h(s) = \begin{cases} 0 & \text{si } s = 0 \\ 1 & \text{si } s = 1 \end{cases}$

Supposons à présent qu'après avoir observé t candidats le décideur choisit le candidat actuel comme le meilleur parmi ceux déjà interrogé. Alors, $g_t(1)$, la probabilité que le candidat choisi est le meilleur parmi tous les autres est donné par:

$$\begin{aligned} P[\text{meilleur candidat}] &= \frac{\text{Nombre de sous ensemble de } \{1, 2, ..., N\} \text{ de taille t contenant 1}}{\text{Nombre de sous ensemble de } \{1, 2, ..., N\} \text{ de taille t}} && (2.1) \\ &= \frac{\binom{N-1}{t-1}}{\binom{N}{t}} && (2.2) \\ &= \frac{t}{N} && (2.3) \end{aligned}$$

, où $\binom{n}{r}$ désigne la combinaison de r dans n. Si tous les $N - 1$ condidats ont été interrogés, le candidat N doit être choisi. Si celui-ci est le meilleur, c'est-à-dire $s = 1$, alors, la probabilité de choisir le meilleur candidat est 1 de sorte que $h(1) = 1$, sinon $h(0) = 0$. Par conséquent, $p_t(j|s) = p_t(j), \forall s = \{0, 1\}$. La probabilité que le prochain candidat soit le meilleur parmi les premiers les $t + 1$ premiers est

$$p_t(1|s) = \frac{1}{t+1}$$

et la probabilité qu'il ne le soit pas parmi les $t + 1$ premiers est

$$p_t(0|s) = \frac{t}{t+1}$$

, $\forall s = \{0, 1\}$. Nous rappélons que les variantes de ce modèle incluent un nombre aléatoire de candidats, une maximisation de la probabilité de choisir l'un des k meilleurs candidats, une minimisation du rang attendu du candidat qui reçoit l'offre, ou une règle de sélection des candidats dans laquelle, après avoir interviewé le candidat actuel, l'employeur peut faire une offre au candidat actuel et à un sous-ensemble désigné de candidats précédents dont la disponibilité dépend d'une distribution de probabilité précise. Considérons à présent $u_t^*(0)$ la probabilité maximale de choisir le meilleur candidat quand le candidat actuel a le rang relatif le plus élevé parmi les premiers t intérogés, et $u_t^*(1)$ la probabilité maximale de choisir le meilleur candidat quand le candidat actuel n'a pas le rang relatif le plus élevé parmi les premiers t intérogés. On pose Δ, l'état d'arrêt.

Ainsi, u_t^* satisfait la formule de récurrence suivante: $u_N^*(1) = h(1) = 1$, $u_N^*(0) = h(0) = 0$, $u_N^*(\Delta) = 0$, et $\forall t < N$,

$$\begin{aligned} u_t^*(1) &= max\{g_t(1) + u_{t+1}^*(\Delta), -f_t(0) + p_t(1|1)u_{t+1}^*(1) + p_t(0|1)u_{t+1}^*(0)\} && (2.4)\\ &= max\{\frac{1}{N}, \frac{1}{t+1}u_{t+1}^*(1) + \frac{t}{t+1}u_{t+1}^*(0)\} && (2.5)\end{aligned}$$

$$\begin{aligned} u_t^*(0) &= max\{g_t(0) + u_{t+1}^*(\Delta), -f_t(0) + p_t(1|0)u_{t+1}^*(0) + p_t(0|0)u_{t+1}^*(0)\} && (2.6)\\ &= max\{0, \frac{1}{t+1}u_{t+1}^*(1) + \frac{t}{t+1}u_{t+1}^*(0)\}, && (2.7)\end{aligned}$$

$u_t^*(\Delta) = u_{t+1}^*(\Delta) = 0.$

Remarquons que $\forall t \in \{1, 2, ..., N\}$, $u_t^* \geq 0$. Cela nous permet de simplifier les équations (2.5) et (2.7). On obtient donc les résultats ci-dessous:

$$\begin{aligned} u_t^*(0) &= \frac{1}{t+1}u_{t+1}^*(1) + \frac{t}{t+1}u_{t+1}^*(0) && (2.8)\\ u_t^*(1) &= max\{\frac{1}{N}, u_t^*(0)\}. && (2.9)\end{aligned}$$

La solution des équations (2.8) et (2.9) nous mènnent à une stratégie optimale comme décrite ci-dessous:

. Dans l'état 1, quand $\frac{t}{N} > u_t^*(0)$, la stratégie optimale est d'arrêter, tandis que quand $\frac{t}{N} < u_t^*(0)$ la stratégie optimale est de continuer. De plus, quand $\frac{t}{N} = u_t^*(0)$, n'importe quelle stratégie est optimale.

. Quant à l'état 0, la stratégie optimale est de continuer

0.2 Formulation avec τ

Les résultats de la partie précédente nous mennent à penser que la stratégie optimale est de la forme "Observer les τ premiers candidats, puis selectionner ensuite le premier candidat meilleur que tous les précédents". Formellement, notons $d_t^*(s)$ la décision optimale à l'instant t en étant dans l'état s, i.e. sélectionner le candidat actuel (Q) ou passer au suivant (C). On veut montrer que la suite des actions optimales s'écrit $(d_1^*, d_2^*, ..., d_{N-1}^*)$, avec $\forall t \in \{1, 2, ..., N-1\}, d_t^*(0) = C$ et $d_t^*(1) = \begin{cases} C & \text{si } t \leq \tau \\ Q & \text{si } t > \tau \end{cases}$, pour un certain τ à déterminer.

- Rappelons que $\forall t < N, u_t^*(0) = \max\{0, \frac{1}{t+1}u_{t+1}^*(1) + \frac{t}{t+1}u_{t+1}^*(0)\} = \frac{1}{t+1}u_{t+1}^*(1) + \frac{t}{t+1}u_{t+1}^*(0)$, donc l'action correspondant à la probabilité maximale de choisir le meilleur candidat en étant dans l'état 0 est toujours de passer au candidat suivant. On a donc $d_t^*(0) = C, \forall t < N$.
  Si on se place maintenant dans l'état 1, $u_t^*(1) = \max\{\frac{t}{N}, u_t^*(0)\}$, le premier terme correspondant à l'action Q et le deuxième à l'action C. Intuitivement, pour des valeurs de t petites (premiers candidats), on aura $u_t^*(1) > \frac{t}{N}$ (cf graphique .....) et donc les premières actions de notre stratégie optimale seront C.

- Supposons que pour un certain τ, $u_\tau^*(1) > \frac{\tau}{N}$ ou que $u_\tau^*(1) = \frac{\tau}{N} = u_\tau^*(0)$. Dans les deux cas, on a $u_\tau^*(1) = \max\{\frac{\tau}{N}, u_\tau^*(0)\} = u_\tau^*(0) \geq \frac{\tau}{N}$. Autrement dit, on a supposé qu'à l'instant τ, quel que soit l'état dans lequel on se trouve, l'action optimale est C. On peut alors exprimer $u_{\tau-1}^*$ par la formule ... :
  $\begin{cases} u_{\tau-1}^*(0) = \frac{1}{\tau}u_\tau^*(1) + \frac{\tau-1}{\tau}u_\tau^*(0) = (\frac{1}{\tau} + \frac{\tau-1}{\tau})u_\tau^*(0) = u_\tau^*(0) \\ u_{\tau-1}^*(1) = \max\{\frac{\tau-1}{N}, u_{\tau-1}^*(0)\} = \max\{\frac{\tau-1}{N}, u_\tau^*(0)\} = u_\tau^*(0) \quad \text{car } u_\tau^*(0) \geq \frac{\tau}{N} > \frac{\tau-1}{N} \end{cases}$
  On vient donc de montrer que s'il est optimal de passer au candidat suivant à l'instant τ, c'est également le cas à l'instant $\tau - 1$. En itérant ce raisonnement, on obtient que pour tout $t < \tau$, $u_{t-1}^*(1) = u_t^*(0)$. Autrement dit, $\forall t < \tau, u_t^*(1) > \frac{t}{N}$ et il est optimal de passer au candidat suivant. En définitive, s'il est optimal de passer au candidat suivant à un instant donné, c'est également le cas aux instants antérieurs.
  Cela exclut les stratégies du type $d_t^*(1) = \begin{cases} C & \text{si } t \leq t' \\ Q & \text{si } t' < t \leq t'' \\ C & \text{si } t > t'' \end{cases}$.
  La stratégie optimale a donc bien la forme $d_t^*(1) = \begin{cases} C & \text{si } t \leq \tau \\ Q & \text{si } t > \tau \end{cases}$.

- Montrons que $N > 2 \implies \tau \geq 1$.
Supposons que $\tau = 0$, alors pour tout t>0, on s'arrête quand on tombe sur le meilleur candidat vu jusqu'à présent. Donc $\forall t > 0, u_t^*(1) = \frac{t}{N}$, et $u_t^*(0) = \frac{1}{t+1}u_{t+1}^*(1) + \frac{t}{t+1}u_{t+1}^*(0) = \frac{1}{N} + \frac{t}{t+1}u_{t+1}^*(0)$. Puis en itérant cette expression : $u_t^*(0) = \frac{t}{N}(\frac{1}{t} + \frac{1}{t+1}) + \frac{t}{t+2}u_{t+2}^*(0) = ... = \frac{t}{N}(\frac{1}{t} + \frac{1}{t+1} + ... + \frac{1}{N-1}) + \frac{t}{N}u_N^*(0)$. Enfin, en utilisant que $u_N^*(0) = 0$, on obtient $u_t^*(0) = \frac{t}{N}(\frac{1}{t} + \frac{1}{t+1} + ... + \frac{1}{N-1}), 1 \leq t < N$. Si $N > 2$, on a $\frac{1}{N-1} \leq \frac{1}{2}$ donc $(1 + \frac{1}{2} + ... + \frac{1}{N-1}) \geq (1 + \frac{1}{2}) > 1$. On en déduit qu'à t=1, $u_1^*(0) = \frac{1}{N}(1 + \frac{1}{2} + ... + \frac{1}{N-1}) > \frac{1}{N}$. Mais $\frac{1}{N} = u_1^*(1) = \max\{\frac{1}{N}, u_1^*(0)\} \geq u_1^*(0)$, on obtient donc une contradiction ($u_1^*(0) > u_1^*(0)$). Cela prouve que $N > 2 \implies 1 \leq \tau < N$ (par définition $\tau < N$).
- Si $N > 2$: on a montré au point 2 que $u_1^*(0) = u_1^*(1) = ... = u_\tau^*(0) = u_\tau^*(1)$, et $\forall t \leq \tau, d_t^*(1) = C$.
D'autre part, on a montré au point précédent que $\forall t > \tau, \begin{cases} u_t^*(1) = \frac{t}{N} \\ u_t^*(0) = \frac{t}{N}(\frac{1}{t} + \frac{1}{t+1} + ... + \frac{1}{N-1}) \end{cases}$.
L'action optimale en étant dans l'état 1 est C tant que $u_t^*(0) > \frac{t}{N}$. On en déduit que $\tau = \max\{t \geq 1 \mid \frac{t}{N}(\frac{1}{t} + \frac{1}{t+1} + ... + \frac{1}{N-1}) > \frac{t}{N}\} \iff \tau = \max\{t \geq 1 \mid \frac{1}{t} + \frac{1}{t+1} + ... + \frac{1}{N-1} > 1\}$.
- Si $N \leq 2$, toute stratégie est optimale (on peut prendre par exemple $\tau = 0$).

En résumé, pour un nombre de candidats N>2 donné, la stratégie optimale est de choisir à chaque instant t : $d_t^*(0) = C$ et $d_t^*(1) = \begin{cases} C & \text{si } t \leq \tau \\ Q & \text{si } t > \tau \end{cases}$, avec $\tau = \max\{t \geq 1 \mid \frac{1}{t} + \frac{1}{t+1} + ... + \frac{1}{N-1} > 1\}$.

3 Conclusion

Remerciements

Références