MATHÉMATIQUES ET SCIENCES HUMAINES

M. VATE

Un modèle d'apprentissage asymétrique des probabilités

Mathématiques et sciences humaines, tome 55 (1976), p. 37-44

<http://www.numdam.org/item?id=MSH_1976__55__37_0>

Math. Sci. hum. (14e année, n°55, 1976, p.37-44)

UN MODELE D'APPRENTISSAGE ASYMETRIQUE DES PROBABILITES

M. VATE★

L'élaboration d'une stratégie adaptative en avenir aléatoire pose le problème de la révision des jugements de vraisemblance sous l'effet de l'information progressivement acquise, ou apprentissage des probabilités. Deux solutions classiques de ce problème (solution psychologique linéaire, solution bayésienne) proposent des modèles d'apprentissage défini, c'est-à-dire dans lesquels la sensibilité à l'apprentissage (propension à réviser les estimations) est soit constante, soit fixée par le nombre d'expériences passées. Dans le modèle RNC proposé ici, la sensibilité dépend de l'histoire des résultats des expériences passées.

I. SOLUTION PSYCHOLOGIQUE LINEAIRE (voir [2] [5] [6] [7])

Soit une expérience à deux issues : R_1 (succès) et R_2 (échec). R_1 peut être la réalisation d'un événement E, et R_2 celle d'un élément du complément de {E} dans un ensemble d'événements. $\hat{P}_t$ désigne l'estimation, après t expériences, de la probabilité P de R_1 ; $\tilde{P}_t$ est l'estimation a priori correspondante. On a :

$$\hat{P}_t = \tilde{P}_{t+1} \qquad (1)$$

La révision de $\hat{P}$ par le modèle linéaire est définie par :

$$\left|\begin{array}{l} \hat{P}_{t+1} = \alpha\, \tilde{P}_{t+1} + (1-\alpha)\, \lambda_j \\ \text{avec } \lambda_j = 1 \text{ si } R_1 \\ \qquad\quad = 0 \text{ si } R_2 \\ \text{et } 0 < \alpha < 1 \end{array}\right. \qquad (2)$$

★ Institut des Etudes Economiques (section "Prévisions-Choix-Planification"), Université Lyon-II.

L'apprentissage n'est pas commutatif : il dépend de l'ordre des événements (succès ou échec). La sensibilité à l'apprentissage, mesurée par $1-\alpha$, est constante.

II. SOLUTION BAYESIENNE (voir [1] [3] [4])

P est une variable aléatoire de densité f(p). En prenant comme loi a priori de P une loi bêta de paramètres a et b, on a :

$$\left| \begin{array}{l} f_o(p) = \beta(a, b\ ;\ p) \\ \hat{P}_o = \tilde{P}_1 = \dfrac{a}{a+b} \end{array} \right. \qquad (3)$$

Après t expériences, dont m succès et n = t−m échecs, le théorème de Bayes donne la nouvelle distribution a priori pour la (t+1)ème expérience :

$$\left| \begin{array}{l} f_t(p) = \beta(a+m, b+n\ ;\ p) \\ \hat{P}_t = \tilde{P}_{t+1} = \dfrac{a+m}{a+b+t} \end{array} \right. \qquad (4)$$

Les relations (4) indiquent que la révision des estimations est d'autant plus rapide que les paramètres a et b sont plus petits ; cela signifie que la conviction du décideur est faible : a et b seront appelés paramètres de conviction.

En posant $h_t = \dfrac{a+b+t}{a+b+t+1}$, il vient :

$$\left| \begin{array}{l} P_{t+1} = h_t\, P_{t+1} + (1-h_t)\, \lambda_j \\ \text{avec } \lambda_j = 1 \text{ si } R_1 \\ \qquad\quad = 0 \text{ si } R_2 \end{array} \right. \qquad (5)$$

où, naturellement, $0<h_t<1$. On remarque l'analogie des relations (2) et (5).

On obtient une autre description intéressante de ce modèle en posant $a_t = a+m$ et $b_t = b+n$:

$$\left| \begin{array}{l} f_t(p) = \beta(a_t, b_t\ ;\ p) \\ \hat{P}_t = \tilde{P}_{t+1} = \dfrac{a_t}{a_t + b_t} \end{array} \right. \qquad (6)$$

à comparer avec (3).

L'apprentissage est commutatif. La sensibilité à l'apprentissage, mesurée par $1-h_t$, dépend seulement de t et non du partage de t entre m et n ni, a fortiori, de l'ordre des événements ; elle est uniformément décroissante.

III. JUSTIFICATION DU MODELE RNC

Une précédente recherche sur le temps de l'action économique et sur la séquentialité des décisions nous a conduit à l'idée que l'acquisition de l'information modifie les jugements de vraisemblance du décideur, mais aussi sa sensibilité à l'apprentissage. En conséquence, un schéma d'apprentissage des probabilités devrait respecter les propositions suivantes :

1/ La probabilité subjective se situe au plan de la connaissance. D'où il suit que tout événement (ou observation), porteur d'information, est facteur de probabilité ;

2/ L'expérience vécue est une fonction irréversible et asymétrique des événements. Elle dépend donc de la nature, de la fréquence et de l'ordre des événements ;

3/ Le degré de conviction du décideur est fonction de l'expérience vécue (au sens de 2/) ;

4/ L'apprentissage est fonction de la tension des convictions, c'est-à-dire de la plus ou moins grande aptitude du degré de conviction (au sens de 3/) à être modifié par l'expérience.

Le modèle psychologique linéaire et le modèle bayésien sont compatibles soit avec la deuxième, soit avec la première proposition, mais avec aucune des deux dernières. Le modèle proposé ici est construit en vue de satisfaire aux quatre conditions posées. Il sera désigné par sa propriété principale qui le distingue le plus nettement des modèles précédents : modèle d'apprentissage à révision non-commutative des convictions (en abrégé modèle RNC).

IV. STRUCTURE DU MODELE RNC

La loi a priori de P est une loi bêta. Les paramètres de la loi sont fixés par les convictions initiales du décideur, mais sont révisables ; les conditions de cette révision devront être précisées. Enfin, après chaque expérience, l'estimation de P devient l'espérance mathématique d'une nouvelle loi a priori de type bêta. On a donc, quel que soit t :

$$f_t(p) = \beta(a_t, b_t ; p) \qquad \text{et} \qquad \hat{P}_t = \frac{a_t}{a_t + b_t}$$

$$f_{t+1}(p) = \beta(a_{t+1}, b_{t+1} ; p) \qquad \text{et} \qquad \hat{P}_{t+1} = \frac{a_{t+1}}{a_{t+1} + b_{t+1}} \tag{7}$$

D'où il suit, d'après (1) :

$$\frac{a_{t+1}}{a_{t+1} + b_{t+1}} \equiv \frac{a_t + \lambda_j}{a_t + b_t + 1} \qquad \text{où } \lambda_j = 1 \text{ si } R_1 \quad = 0 \text{ si } R_2 \tag{8}$$

L'identité (8) est vérifiée par une infinité de solutions telles que :

$$\frac{a_{t+1}}{b_{t+1}} = \frac{a_t + \lambda_j}{b_t - \lambda_j + 1}$$

Le tableau suivant donne les solutions les plus simples, celles qui correspondent à la révision d'un seul paramètre de conviction après chaque expérience :

	Résultat R_1 ($\lambda_j = 1$)	Résultat R_2 ($\lambda_j = 0$)
Révision de a (b_t = cte)	$a_{t+1} = a_t + 1$	$a_{t+1} = a_t \frac{b_t}{b_t + 1}$
Révision de b (a_t = cte)	$b_{t+1} = b_t \frac{a_t}{a_t + 1}$	$b_{t+1} = b_t + 1$

La première diagonale désigne la solution du modèle bayésien (apprentissage commutatif ; fonction d'apprentissage h_t ne dépendant que de t). La

seconde diagonale rend h_t fonction des nombres m et n de succès et d'échecs, mais l'apprentissage demeure commutatif. La règle de révision du modèle RNC est fixée par l'une des lignes du tableau, au choix.

V. PROPRIETES DU MODELE RNC

A. Non-commutativité des révisions

A l'instant initial, les paramètres de conviction sont a_o et b_o. On choisit, par exemple, la révision de a (première ligne du tableau) :

$$a_{t+1} = a_t + 1 \qquad \text{si } R_1$$

$$a_{t+1} = a_t \frac{b_o}{b_o + 1} \qquad \text{si } R_2$$

Posons $k_b = b_o/(b_o + 1)$. A partir de a_o, le paramètre a va être l'objet de m majorations (+1) et de n multiplications par $k_b < 1$. Aussi, chaque majoration n'est-elle pas définitivement acquise (à la différence du modèle bayésien) si son inscription est suivie de un ou plusieurs échecs. Le "+1" introduit par le jème succès survenu en t_j sera multiplié par k_b autant de fois qu'il apparaîtra d'échecs entre t_{j+1} et t (bornes comprises). On notera $f_j(1)$ la valeur résiduelle en t de la majoration (+1) apparue en t_j. D'où :

$$a_t = a_o \cdot k_b^n + \sum_{j=1}^{m} f_j(1)$$

et l'on calcule : $f_j(1) = 1.k_b^{n+j-t_j}$. D'où finalement :

$a_t = a_o.k_b^n + \sum_{j=1}^{m} k_b^{n+j-t_j}$	$b_t = b_o$

(9)

Remarques :

1° - S'il y a t échecs consécutifs, le vecteur de conviction en t est $(a_t, b_t) = (a_o k_b^t, b_o)$; s'il y a t succès, on obtient $(a_t, b_t) = (a_o+t, b_o)$, identique à la solution bayésienne.

2° - Si on permute un succès avec un échec qui lui était postérieur, a_t est minoré ; a_t est majoré si le sens de la permutation est inversé : la révision est non-commutative.

3° - Si b_o est très grand, a_t tend vers a_o+m, mais le vecteur diffère encore de la solution bayésienne par la fixité de b_o.

4° - Le même raisonnement, appliqué à la révision de b, donne des résultats correspondants. Si, au cours de t expériences, on a observé n échecs aux dates t_i et m succès, on obtient :

$$\boxed{a_t = a_o \qquad b_t = b_o.k_a^m + \sum_{i=1}^{n} k_a^{m+i-t_i}} \tag{10}$$

où $k_a = a_o/(a_o+1)$. On note encore que t succès consécutifs donnent : $(a_t, b_t) = (a_o, b_o k_a^t)$ et t échecs : $(a_t, b_t) = (a_o, b_o+t)$.

B. Asymétrie de l'apprentissage

Cette propriété d'asymétrie exprime le remplacement du dénombrement des événements par leur chronologie. Elle découle de la propriété précédente (révision non commutative), comme on peut le montrer en écrivant la fonction d'apprentissage h_t qui correspond au modèle RNC. En effet, à partir des relations (7) et (8), on peut définir une fonction h_t telle que :

$$\hat{P}_{t+1} = h_t \tilde{P}_{t+1} + (1-h_t)\, \lambda_j$$

Elle s'écrit : $h_t = \dfrac{a_t + b_t}{a_t + b_t + 1}$

D'après (9) ou (10) :

ou
$$\boxed{h_t = h(a_o, b_o, n, t_j)} \qquad \boxed{h_t = h(a_o, b_o, m, t_i)} \tag{11}$$

alors que dans le modèle bayésien, on pouvait noter :

$$h_t = h(a_o, b_o, t).$$

La relation (11) exprime que l'apprentissage est une fonction asymétrique du nombre de succès obtenus, et en souligne le caractère "historique". Ainsi la fonction h_t a-t-elle une plus grande liberté de variation que dans le modèle bayésien. Cette liberté reste pourtant assez étroite. Dans le modèle bayésien, la valeur initiale $h_o = (a_o + b_o)/(a_o + b_o + 1)$ est aussi un minimum pour h_t. Mais h_t ne reviendra jamais à h_o car il tend nécessairement vers 1 quand t augmente, quels que soient les résultats observés et quel que soit le vecteur de conviction initial (a_o, b_o). Dans le modèle RNC, h_t ne tend vers 1 que si l'on observe une succession indéfinie de succès ; d'autre part, h_t peut devenir inférieur à h_o en cas d'échecs répétés. Sa limite inférieure, correspondant à une suite infinie d'échecs, est $b_o/(b_o + 1) < h_o$. L'intervalle de variation de h_t est donc d'autant plus étroit que les convictions initiales sont plus tendues (a_o et b_o grands).

Enfin, le modèle RNC met en évidence une caractéristique importante de l'apprentissage de la vraisemblance : quels que soient les résultats observés, il est plus difficile d'accéder à la quasi-certitude subjective que de s'en départir.

-:-:-:-:-:-:-:-:-

BIBLIOGRAPHIE

[1] BELLMAN, R., Adaptive Control Processes, a Guided Tour, Princeton, Princeton University Press, 1961, chap. 16 et 17.

[2] BUSH, R. R. and F. MOSTELLER, Stochastic Models for Learning, New York, Wiley, 1955.

[3] JACQUARD, A., Les probabilités, Paris, Presses Universitaires de France, 1974, chap. 5.

[4] MURPHY, R.E., Adaptive Processes in Economic Systems, New York, Academic Press, 1965, chap. 4.

[5] NORMAN, F., Mathematical Learning Theory, in Mathematics of the Decision Sciences, G.B. DANTZIG and A.F. VEINOTT ed., Amer. Math. Soc., Providence, 1968, 2ème Partie, pp. 283 à 313.

[6] ROUANET, H., Les modèles stochastiques d'apprentissage, Paris, Gauthier-Villars, 1967, chap. 3 et 4 spécialement.

[7] STERNBERG, S., "Stochastic Learning Theory", in Handbook of Mathematical Psychology, vol. II, LUCE, BUSH and GALANTER ed., New York, Wiley, 2ème éd., 1967, pp. 19 à 24.