技術解說
17—1—1

二項教習過程의 수렴에 對해서

On the Convergence of Binomial Learning Process

姜 麟 求*
(In Ku Kang,)

〔Abstract〕

The convergence rate of Bayes' learning process is investigated for a binomial random variable. A measure. of the rate of convergence is proposed and it is found that such a measure can be approximated by an exponential function of the number of observations.

서 론

근래에 Bayes의 교습과정은 많은 注目을 받아왔다 [1~7]. 비록 이러한 교습과정이 極限에서 收斂함은 이미 증명된 바 있으나 收斂度에 대한 연구는 별로 없다. 이 논문에서 二項 亂變數에 대한 Bayes 교습과정의 收斂度를 조사하였다.

二項 教習課程은 二元 對稱 차넬을 통해 작동되고 誤差傳達 確率에. 따라 그 傳達率이나 解讀裝置의 機構를 最適化시키는 適應性 通信系統에 利用될 수 있다[2, 5]. 이경우 二項亂變數는 시간에 따라 변하므로 分布가 一定하다는 假定밑에 이루어지는 관찰의 회수를 제한하거나 혹은 다른 방법으로 過度教習의 폐단을 피해야한다. 어떤 방법을 쓰든 收斂度는 그 방법의 演算法 決定에 중요한 要素이다.

二項 亂變數에 對한 Bayes의 教習過程

X_k는 다음과 같다. 즉

$$X_k=\begin{cases}1 & K\text{째 관찰에서 어느 현상이 나타난 경우}\\ 0 & \text{그렇지 않은 경우}\end{cases}$$

각각의 관찰은 통계상 독립적이고 X의 평균은 관찰 전부를 통해 일정하다고 가정한다. 즉

$$E[X_k]=q_o$$

윗값은 물론 미리 모른다.

m가

$$m=\sum_{k=1}^{n}X_k \text{ 이면}$$

Bayes의 법칙에서 q의 事後 신빙도 함수**는

$$f_n(q|m)=\frac{f_o(q)q^m(1-q)^{n-m}}{\int\lambda^m(1-\lambda)^{n-m}f_o(\lambda)d\lambda}\quad\cdots\cdots\cdots\cdots(1)$$

여기서 Φ는 q의 標本 存在空間이다. 만약 事前 신빙도 함수 $f_o(q)$도 事後 신빙도 함수와 그 형식이 같다면 즉

$$f_o(q)=\frac{q^j(1-q)^i}{\int_\Phi\lambda^j(1-\lambda)^i d\lambda}\quad\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(2)$$

여기서 i와 j는 임의의 整數로서 事前知識에 따라 선택된 것이다. 그리고 $\Phi=[0,\ 1)$이면 $f_n\ (q|m)$은 다음과 같은 재생(Ωeproducing) 형[3]이 된다. 즉

$$f_n(q|m)=A(m+j,\ n-m+i)\times q^{m+j}(1-q)^{n-m+i}\quad\cdots\cdots\cdots\cdots\cdots\cdots\cdots(3)$$

여기서

$$A(a,b)=\frac{(a+b+1)!}{a!\ b!}$$

收 斂 度

관찰회수 n가 무한대로 접근함에 따라 q의 事後 신빙도 함수도 그 極限형태에 근접한다. 만약 q의 事後 신빙도 함수가 등식(3) 같은 형태라면 그 極限형태는 $q=q_o$에 있어서의 "델터"함수로 나타난다[2]. 위의 사실은 또한 다른 분들이 [1,3,4] 증명한 일반적인 收斂 정리에 의해서도 증명된다.

실제 계통에 응용하려면 極限에서의 收斂만으로는 不充分하다. 그 까닭은 q의 값이 그렇게 오랫동안 일정하지 않다는 전제때문에 적응성 계통의 존재 가치가 있기 때문이다. q가 시간에 따라 변한다면 그 시간적 변동을 명확히 파악하려면 작은 관찰회수가 요망되는 반면 너무 작으면 매우 낮은 신빙도밖에 못얻는다. 그래서 收斂度는 한 教習過程에서의 관찰회수를 最適하게 얻는데 중요한 요소가 되며 또한 다른 곳에서 설명된 시

*正會員 : 해군사관학교

**신빙도 함수란 "確信度" 또는 未知의 g의 가능한 값마다에 주워진 우리의 지식 정도의 數的 測度이다. 渡邦 〔7〕이 主唱한 用語이다.

간에 따라 변하는 q에 관련된 다른 방법에서도 중요한 역할을 한다[5, 6].

收斂의 測度는 여러가지 있겠지만 [2, 5] 여기서는 다음과 같은 신빙도를 쓰겠다. 즉

$$C_r[|q-q|\leq\epsilon]=\int_{\hat{q}-E}^{\hat{q}+E} f_n(q|m)dq$$

q는 $f_n(q|m)$의 모—드이다.

윗식을 고쳐 쓰며는

$$C_r[|q-\hat{q}|\leq\epsilon]=1-\int_o^{\hat{q}-E} f_n(q|m)dq - \int_{\hat{q}+E}^{1} f_n(q|m)dq \quad \cdots\cdots(4)$$

부록에서 $i=j=0$이면

$$C_r[|q-\hat{q}|\leq\epsilon]\sim\text{Sup}\left[0,\ \left\{1-2\exp\left(\frac{-(n+1)\epsilon^2}{2\hat{p}\hat{q}}\right)\right\}\right]$$

여기서

$$\hat{q}=\frac{m}{n}\text{이고 } \hat{p}=1-\hat{q}\text{ 이다.}$$

윗식 (5)에서 收斂은 관찰회수에 對해서 指數함수임을 알 수 있다. 뿐만아니라 그 指數는 未知인 q가 아니라 관찰한 값인 $\hat{q}$의 간단한 함수이다.

식 (5)의 오른 쪽변은 n가 n_t보다 적으면 영이다. 즉

$$n_t=\frac{2\hat{p}\hat{q}}{\epsilon^2}\ln 2$$

$n\angle n_t$이면 식(4)와의 비교에서 다음과 같은 근사식이 대부분의 $\hat{q}$에 대해서 성립됨이 알려지었다[6].

$$C_r[|q-\hat{q}|\leq\epsilon]\sim1-\exp\left[\frac{-(n+1)\epsilon^2}{2\hat{p}\hat{q}}\right] \quad \cdots\cdots(6)$$

i와 j가 영이 아니지만 $\hat{q}_t$를 다음과 같이 정의하면

$$\hat{q}_t=\frac{m+j}{n+i}$$

식 (5)와 근사한 식을 얻는다. 즉

$$C_r[|q-\hat{q}_t|\epsilon]\leq\sim\text{Sup}[0,\ B] \quad \cdots\cdots(7)$$

여기서

$$B=1-2\exp\left[\frac{-(n+j+1)\epsilon^2}{2\hat{p}\hat{q}}\right]$$

그러나 위의 경우에는 $\hat{q}$가 $\hat{q}_t$와는 판이하게 다를 수 있으며 그 차는 결국 없어지겠지만 상당한 관찰회수에 걸쳐 남아 있을 수가 있다.

結　　論

제의한 측도에 의거하면 二項 数習過程은 관찰회수의 指數함수의 형태로 收斂한다.

이 논문에서 얻은 근사식은 비교적 간단하고 $\hat{q}$라는 관측치의 함수이다. 이러한 특성때문에 이 식은 신빙도와 시간 변화의 가능성을 관찰회수의 함수로 나타내고 그들을 비교하여 q가 시간에 따라 변하는 상태하에서의 数習過程方式을 決定하는데 도움이 된다.

끝으로 본인의 지도교수이었던 A.H. Koschmann 박사의 지도에 감사드린다.

부　　록

q의 참값이 선간$[\hat{q}-\epsilon,\ \hat{q}+\epsilon]$ 안에 존재하는 신빙도는

$$C_r[|q-\hat{q}|\leq\epsilon]=1-C_r[|q-\hat{q}|>\epsilon] \quad \cdots\cdots(\text{A1})$$

또

$$C_r[|q-\hat{q}|>\epsilon]=\int_o^{\hat{q}-E} f_n(q|m)dq+\int_{\hat{q}+E}^{1} f_n(q|m)dq \cdots(\text{A2})$$

위의 未完 베터積分은 二項分布의 합으로 표시할 수 있다. 즉 식(A 2)의 오른편 제 2항은 다음과 같이 쓸 수 있다.

$$\int_{\hat{q}+E}^{1} f_n(q|m)dq=P\left[\sum^{n+1} E_k<m|\hat{q}+\epsilon\right] \quad \cdots\cdots(\text{A 3})$$

여기서

$$E_k=\begin{cases}1\\0\end{cases}\text{이고 } E[E_k]=\hat{q}+\epsilon$$

식 (A 3)의 오른편은 다시 $\frac{m}{n+1}<\frac{m}{n}=\hat{q}$ 이니까

$$P\left[\frac{1}{n+1}\sum^{n+1}E_k<\frac{m}{n+1}|\hat{q}+\epsilon\right] < P\left[\frac{1}{n+1}\sum^{n+1}E_k\leq q_u-\epsilon|q_u\right] \cdots\cdots(\text{A 4})$$

여기서 $q_u=\hat{q}+\epsilon$

(A 4)의 식 오른편에 대해서 Chernoff의 上限을 이용하면 [8]

$$P\left[\frac{1}{n+1}\sum^{n+1}E_k\leq q_u-\epsilon\right] \leq\exp\left[(n+1)\ \ln\left(\frac{q_u}{\hat{q}}\right)^{\hat{q}}\left(\frac{1-q_u}{1-\hat{q}}\right)^{1-\hat{q}}\right] \cdots\cdots(\text{A 5})$$

윗식의 指數의 對數部分을 級數로 展開하면

$$\ln\left[\left(\frac{q_u}{\hat{q}}\right)^{\hat{q}}\left(\frac{1-q_u}{1-\hat{q}}\right)^{1-\hat{q}}\right]=\hat{q}\ln\left(\frac{\hat{q}+\epsilon}{\hat{q}}\right)+\hat{p}\ln\left(\frac{\hat{p}-\epsilon}{\hat{p}}\right) <-\frac{\epsilon^2}{2\hat{p}\hat{q}}+\frac{1}{3}\ \frac{(\hat{p}^2-\hat{q}^2)\epsilon^3}{\hat{p}^2\hat{q}^2}$$

따라서

$$\int_{\hat{q}+E}^{1} fn(q|m)dq<\exp\left[-\frac{(n+1)\epsilon^2}{2\hat{p}+\hat{q}} +\frac{(n+1)(\hat{p}^2-\hat{q}^2)\epsilon^3}{\hat{p}^2\hat{q}^2}\right] \cdots\cdots(\text{A7})$$

식 (A 2)의 오른편 첫항도 $P=1-q$란 값을 대입하여 정리하면

$$\int_o^{\hat{q}-E} f_n(q|m)dq=\int_{\hat{q}+E}^{1} f_n(p|n-m)dp =P\left[\frac{1}{n+1}\sum^{n+1}E_k<\frac{n-m}{n+1}|\hat{p}+\epsilon\right] <P\left[\frac{1}{n+1}\sum^{\hat{q}+1}E_k<P_b-\epsilon|P_b\right] \cdots\cdots(\text{A 8})$$

여기서 $P_b=\hat{p}+\epsilon$

식(A8)를 같은 방법으로 처리하면 식(A 7)와 비슷한 식을 얻는다. 즉

$$\int_0^{\hat{q}-E} f_n(q|m)dq<\exp\left[\frac{-(n+1)\epsilon^2}{2\hat{p}\hat{q}}\right.$$

$$-\frac{(n+1)(\hat{p}^2-\hat{q}^2)\epsilon^3}{\hat{p}^2-\hat{q}^2}\Big] \cdots\cdots\cdots\cdots\cdots (A\,9)$$

식(A 7)과 (A 8)를 식(A 2)에 대입 후 다시 식 (A 1)에 그 결과를 대입하면

$$C_r[|q-\hat{q}|\leq\epsilon]>1-2\exp\left[\frac{-(n+1)\epsilon^2}{2pq}\right]$$

$$\times\cosh\left[(n+1)\frac{(\hat{p}^2-\hat{q}^2)\epsilon^3}{\hat{p}^2\hat{q}^2}\cdots\cdots\cdots (A\,10)\right.$$

cosh 항의 크기는 指數부분보다 월등히 작으므로

$$C_r[|q-\hat{q}|\leq\epsilon]\sim1-2\exp\left[\frac{-(n+1)\epsilon^2}{2\hat{p}\hat{q}}\right]$$

이라 할 수 있고 신빙도가 결코 영이하는 될 수 없기 때문에 다음과 같이 다시 씀이 적절하다. 즉

$$C_r[1q-\hat{q}|\leq\epsilon]\sim\text{Sup}\left\{0,\left[1-2\exp\left\{\frac{-(n+1)\epsilon^2}{2\hat{p}\hat{q}}\right\}\right]\right\}$$

참 고 문 헌

1. Braverman, D. J., "Machine Learning and Automatic Pattern Recognition", Stanford Electronics Labs, Stanford, California, Tech. Rept. No. 2003—1, Feb. 1961.
2. Breipohl, A. M., and Koschmann, A.H., "A Communication System with Adaptive Decoding", Proc. of Third Symposium on Adaptive Process, pp. 72-85, Sept. 1964.
3. Spragains, J.D., Jr., "Reproducing Distributions for Machine Learning", Stanford Electronic Labs, Stanford, California, Tech. Rept. 6103-7, Nov. 1963; also a part of it as "A Note on the Interative Application of Bayes' Rule", IEEE Trans. on Information Theory Vol. IT-11, pp. 544-549, Oct. 1965.
4. Fralick, S.C., "Learning to Recognize Patterns Without a Teacher", Stanford Electronics Labs, Stanford, California, TR6103-10, March 1965.
5. Chao, C. F., and Koschmann, A. H., "On the Application of Bayesian Learning in Discrete Communication Systems", The Univ. of New Mexico, Albuqerque, N. Mexico, TREE-136, Aug. 1966.
6. Kang, I.K., and Koschmann, A. H., "A Study of Learning in Adaptive Systems", The Univ. of New Mexico, Albuquerque, N. Mexico, TREE-135, Aug. 1967.
7. Watanabe, S., "Information-Theoretical Aspects of Inductive and Deductive Inference", IBMJ of R&D, Vol. 4, pp. 208-231, April 1960.
8. Wozencraft, J. M and Jacobs, I. M., "Principles of Communication Engineering", John Wiley, New York, 1965. pp. 103

〈52p 에서 계속〉

理事, 成樂正技術部次長, 吳昌錫, 元俊喜 兩課長)을 招請하여 高見을 區分聽取해 보기로 하다.

2. 韓電側事業維持會費 增加를 要請키로하되 一但 事前에 折衝하기로 하다.
3. 事業維持會員을 年 2, 3回程度 招請하여 懇談會 또는 座談會形式으로 電氣界의 共同의 廣場으로 相互造成되기를 要望하다.
4. 電氣機器分野에 關한 講座進行經過報告
5. 電氣工學의 最近의 進步 定例講座 年間基本計劃案樹立
6. 物品購入 4段화일 Box(1個) 및 카비넷(1個)
7. 誘導障害問題의 圓滿한 意見交換을 하기 위하여 다음 機會에 遞信部, 韓電및 鐵道廳 三者連席會義의 橋梁的 役割의 討義準考를 進行하기로 하다.

電氣工學便覽編纂分科委員長會

西紀 1968年 4月 16日

出 席 者

禹亨疇, 韓萬春, 姜錫圭, 梁興錫, 徐錫仁, 李承院, 池哲根, 申龍澈, 尹日重

案 件

1. 原稿는 各分科別로 委員長의 最後承諾後에 印刷所에 보내기로하다(단 分委長은 所管分委의 原稿에 대하여 統計數字의 訂正內容의 調整 및 資料의 補充等에 注力하며, 이를 早速한 時日內에 마치기로 하다)
2. 臨時職員一人採用하여 電氣工學便覽出版에 協助키로 하다.