数学入門公開講座

平成4年8月4日(火)から8月13日(木)まで

京都大学数理解析研究所

講師及び内容

1. 確率論の話題から（7時間）

京都大学数理解析研究所・助教授　楠　岡　成　雄

確率論に関連した4つの話題（シャノンの情報理論、確率制御の話題、シミュレーテッド・アニーリング、乱数の理論）について話を行う予定である。どのように問題をとらえるかといった考え方について話の力点を置いて話していく。

2. 「保証書」付き数値計算法（7時間）

京都大学数理解析研究所・助教授　室　田　一　雄

計算機のできる数値計算は、所栓は誤差を含んだ近似計算である。それでも、その計算結果に従って橋は作られ列車は走る。数値計算誤差につきまとう不安感を一掃し、ぼんやりとした計算結果に基づいてはっきりとした結論を導き出す手法について解説する。

3. 「時間とミクロ世界・マクロ世界」（7時間）

京都大学数理解析研究所・助教授　小　嶋　　泉

「時間とは何か？人が問わなければ、私はそれを知っている。だが人に説明しようとすると、私は最早それを知っていない」――時間を巡る謎のとらえどころのなさは、聖アウグスティヌスのこの有名な言葉によく表わされている。自然法則を書き下す上で時間概念は不可欠であり、今世紀の物理学革命も時空間概念の根本的変革と軌を一にして達成された。それによってこの「謎」はどう解き明されたのか？時間概念を軸に、基本的な物理法則とその論理構造を概観し、そこから、自然の運動・構造・歴史を理論的・数学的に記述することの意味を改めて考え直してみたい。

4. グラフと組合せ論（7時間）

京都大学数理解析研究所・助手　松　本　　眞

グラフ理論や組合せ論の世界は、まだいまいち体系的ではありませんが、たとえばかつての四色問題などのように、誰でも問題の内容がすぐ理解できるのに、証明はいまだ満足にされていないような楽しい予想が小石のように転がっています。こうした楽しそうな部分をつまみぐいして、平易にグラフと組合せ論の紹介をします。

時　間　割

日 / 時間	8月4日(火)	5日(水)	6日(木)	7日(金)	8日(土) 9日(日)	10日(月)	11日(火)	12日(水)	13日(木)
13:15～15:00	楠岡	楠岡	楠岡	楠岡	休講	小嶋	小嶋	小嶋	小嶋
15:00～15:15	休憩					休憩			
15:15～17:00	室田	室田	室田	室田		松本	松本	松本	松本

1．確率論の話題から　（7時間）

京都大学数理解析研究所・助教授 楠 岡 成 雄

1992, AUGUST 4,5,6,7　　13:15-15:00

確率論の話題から

楠岡成雄

1 考える問題

この講座では次のような問題について論じていく。

1.1 伝えることのできる情報の大きさ

今、甲氏から乙氏に対して0または1の二つの数字からなる信号を n 個送るとする。甲氏と乙氏との間であらかじめ信号について相談してあるとすればどれだけの情報を送ることができるであろうか。その答は簡単で、2^n 通りの場合を区別できるわけであるから 2^n の情報が送れるというのが答である。ではもし通信の途中でノイズが入り、信号が完全には送れないという状況ではどれだけの情報を送ることができるであろうか。

設定をはっきりさせよう。ここで通信にノイズがはいるとは次のような状況であるとする。

『$p \in (0,1)$ が与えられており、信号0（または1）が1（または0）と誤って伝えられる確率が p であり、各信号が誤って伝えられるという事象は独立。』

この時、2^n の情報を伝えられると思うのは余りにも楽天的である。2^n はすべての場合の数だから、一つでも信号が誤って伝えられれば全体の情報が誤りということになる。すべての信号が誤りなく伝えられる確率は $(1-p)^n$ であるので、n が増大するにつれて急速に0に近づく。よってほとんど確実に情報は誤りであることになる。

「ノイズがある限り誤りを限りなく小さくして情報を伝えることなどできるはずがない」と思われる方もいるかも知れない。しかし次のようにすれば少なくとも二つの情報を伝えることができる。今、甲氏がある事柄についてAであるかBであるかについて伝えたいとする。甲氏はもしAであれば0を n 回発信し、Bであれば1を n 回発信することにする。乙氏は信号の0と1の数を比較し、0が多ければA、1が多ければB、等しければ硬貨を投げて裏か表かで決めることにする。この時、誤りを犯す確率は

$$n \text{ が奇数}, \quad n = 2m+1 \text{ の時} \quad \sum_{k=0}^{m+1} \binom{n}{k} p^k (1-p)^{n-k}$$

$$n \text{ が偶数}, \quad n = 2m \text{ の時} \quad \sum_{k=0}^{m} \binom{n}{k} p^k (1-p)^{n-k}$$

となる。もし$0 < p < 1/2$ならば、これは指数的に0に減少することがわかる。

上にみたようにノイズのある場合の伝達可能な情報の数は2と2^nの間のどこかにあるはずである。では実際にどれだけの情報を送ることができるのか。これが問題である。

1.2 確率制御：最適停止時刻

1から10までの数字のかいたカードが10枚あるとする。そして次の作業を行うことにする。

- まず、このカードをよくきって1枚ひく。
- 気にいればそのカードを手元においてそれを自分のカードとして作業を終了する。
- そのカードが気にいらなければカードを返して、再びカード10枚をよくきって、また1枚カードをひく。
- この操作を気に入ったカードをひくまで繰り返すことができるが、8回目にカードをひいた時にはひいたカードを自分のカードとして作業を終了しなければならない。

さて作業を終えたときの自分のカードの数字をできるだけ大きくしたい。ここでこちらが選択できるのは、作業をいつ終了するかだけである。どうすれば、自分のカードの数字の期待値を最も大きくすることができるであろうか。

1.3 乱数

乱数とは「規則性のない数字の列」である。しかし、どうであれば数列に「規則性がない」といえるのであろうか。例えば次のような問題を考えてみよう。

0と1の二種類の数字n個よりなる数列を考える。このような数列は全部で2^n通りある。この中で最も規則性のないものはどの数列か。

講演では他にSimulated Annealingと呼ばれるモンテカルロ法による最適化問題についても解説する予定である。

2 Shannonの情報理論入門

2.1 Shannonの情報理論

最初の問題1.1は1948年にShannonにより初めて解析された。Shannonは情報伝達を最も効率よくするにはどうすればよいかを考え、ノイズのある通信路における情報伝送の問題、情報縮約の問題等、情報理論の基礎を与えた。Shannonの理論はそ

の後多くの人たちにより整理、精密化された。問題 1.1 は Shannon の理論の出発点であると言っていい。以下、Shannon の理論の基本的な考えを述べていく。

2.2 エントロピー

Ω は有限集合、$p : \Omega \to [0,1]$ は Ω の上の関数で

$$\sum_{\omega \in \Omega} p(\omega) = 1$$

を満たすものとする。$p(\omega)$ は ω の起こる確率と考える。Ω の部分集合 A に対して、

$$P(A) = \sum_{\omega \in A} p(\omega)$$

とおく。$P(A)$ は事象 A の確率である。ただし、$P(\emptyset) = 0$ とする。

A, B を事象とする。この時、事象 B の与えられたときの事象 A の条件つき確率 $P(A|B)$ を

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

で表す。

Ω の部分集合の族 $\mathcal{F}$ が

(1) $\quad A, B \in \mathcal{F}, A \neq B \quad$ ならば $\quad A \cap B = \emptyset$

(2) $\quad \bigcup_{A \in \mathcal{F}} A = \Omega$

を満たすとき、Ω の分割という。

$\mathcal{F}$ が Ω の分割であるとき、そのエントロピー $H(\mathcal{F})$ を

$$H(\mathcal{F}) = -\sum_{A \in \mathcal{F}} P(A) \log P(A)$$

により定義する。ただし、$0 \times \infty = 0$ とする。この時、次のことがわかる。

$$0 \leq H(\mathcal{F}) \leq \log(\sharp(\mathcal{F}))$$

いま、$\mathcal{F}, \mathcal{G}$ を Ω の分割とする。二つの分割の合成 $\mathcal{F} \vee \mathcal{G}$ を

$$\mathcal{F} \vee \mathcal{G} = \{A \cap B; A \in \mathcal{F}, B \in \mathcal{G}\}$$

で与える。Ω の分割の合成は再び Ω の分割となる。また、分割 $\mathcal{G}$ が与えられたときの分割 $\mathcal{F}$ の条件つきエントロピー $H(\mathcal{F}|\mathcal{G})$ を

$$H(\mathcal{F}|\mathcal{G}) = -\sum_{B \in \mathcal{G}} P(B) \{ \sum_{A \in \mathcal{F}} P(A|B) \log P(A|B) \}$$

この時次のことが成立する。

$$H(\mathcal{F}\vee\mathcal{G}) = H(\mathcal{G}) + H(\mathcal{F}|\mathcal{G})$$

$$H(\mathcal{F}) \geq H(\mathcal{F}|\mathcal{G})$$

さらに、二つの分割の相互情報量 $I(\mathcal{F};\mathcal{G})$ を

$$I(\mathcal{F};\mathcal{G}) = H(\mathcal{F}) - H(\mathcal{F}|\mathcal{G})$$

で定める。

2.3 通信路の容量

今、$X = \{x_1,\ldots,x_M\}, Y = \{y_1,\ldots,y_L\}$ とおき、X は送信される信号の集合、Y は受信される信号の集合とする。通信にノイズが混じるので、同じ信号が発信されても同じ信号が受信されるとは限らない。そこで、信号 x_i が発信されたという条件の下で信号 y_j が受信される確率を p_{ij} とする。もちろん、

$$p_{ij} \geq 0, \qquad \sum_{j=1}^{L} p_{ij} = 1, \quad i = 1,\ldots,M$$

である。

いま、$\xi = (q_1,\ldots,q_M)$ はベクトルで $q_i \geq 0, \sum_{i=1}^{M} q_i = 1$ を満たすとする。この時、$\Omega = X\times Y$, $p:\Omega\to[0,1]$ を

$$p((x_i,y_j)) = q_i p_{ij}$$

で定めると、$\sum_{\omega\in\Omega} p(\omega) = 1$ を満たす。Ω の分割 $\mathcal{X},\mathcal{Y}$ を

$$\mathcal{X} = \{\{(x,y); y\in Y\}; x\in X\}, \qquad \mathcal{Y} = \{\{(x,y); x\in X\}; y\in Y\}$$

で定義する。この時、分割 $\mathcal{X}$ と $\mathcal{Y}$ の相互情報量 $I(\mathcal{X};\mathcal{Y})$ は ξ に依存するので ξ の関数となる。それを $f(\xi)$ と書くことにすると、通信路の容量 C は

$$C = \sup\{f(\xi);\ \xi = (q_1,\ldots,q_M)\in[0,1]^M,\ \sum_{i=1}^{M} q_i = 1\}$$

により定義される。この時、Shannon らによる情報理論の結論は粗っぽくいうと次のようになる。

定理 1 *(1)* 情報の大きさが $\exp(C)$ より十分小さければ、誤り確率をいくらでも小さくして、この通信路で情報を送ることができる。
(2) 情報の大きさが $\exp(C)$ より十分大きい時は、どのように工夫してこの通信路で情報を送っても、誤り確率をある正数以下にはできない。

この定理の正確な意味については講演の中で述べる。以下の節では、この定理を示すにあたって必要な概念を述べておく。

2.4 誤り確率

送信信号 $x_i, i = 1, \ldots, M$ の現れる確率が q_i であることがわかっているとしよう。この時、もし受信信号 $y_j, j = 1, \ldots, L$ を受けたならば、Bayes の定理を用いて x_i が送信信号であったという確率は

$$p(x_i|y_j) = \frac{q_i p_{ij}}{\sum_{k=1}^{M} q_k p_{kj}}$$

である。よって、$p(x_i|y_j)$ を最大にする x_i を選ぶのが誤り確率を最小にする最適の方法である。$p(x_i|y_j)$ を最大にする x_i は $q_i p_{ij}$ を最大にする x_i に等しい。したがって、誤りをおかす確率を P_e とすると

$$1 - P_e = \sum_{j=1}^{L} \max\{q_i p_{ij}; \quad i = 1, \ldots, M\}$$

となる。この時、Fano の不等式と呼ばれる次のような式が成立する。

$$H(\mathcal{X}|\mathcal{Y}) \leq -\{P_e \log P_e + (1 - P_e)\log(1 - P_e)\} + P_e \log(M - 1)$$

2.5 ランダム符号化

通信したい情報源を $S = \{s_1, \ldots, s_N\}$ とする。この情報源をどのように送信信号に割り当てるのが効率的であろうか。実はこの問題はかなり複雑な問題である。本来は $q_i p_{ij}$ に依存して決定しなくてはならないが、容易には決められない。Shannon はこれについて次のようなランダム符号化のアイデアを提唱した。

ランダム符号化とは情報源と送信信号の対応を確率的に決定するということである。最適のやり方を選ばねばならないのに、ランダムに選ぶのはきわめて非能率と思われるかも知れないが、漸近的にはほぼ最良の結果を得ることができる。詳しくは講演の中で説明していく。

なお、Shannon の情報理論については

1. 有本卓　情報理論　(共立数学講座　22) 共立出版　1976

2. 国沢清典・梅垣寿春編　情報理論の進歩　岩波書店　1965

を参考にして解説していく。

3 Simulated Annealing について

E を有限集合、$U : E \to \mathrm{R}$ は関数として、「$U(x)$ を最小にする x を見つけよ」という問題は現実によく出会う問題である。もちろん、集合 E や関数 U に特別な構造がある場合はそれに応じた解法があるだろう。しかし、あまりはっきりした構造がない場合は解法の決定版というものはないかもしれない。

例を一つあげる。今、N は自然数とし、$a_{ij} > 0, i, j = 1, \ldots, N$ が与えられているとする。$x = (i_1, \ldots, i_N)$ は $(1, \ldots, N)$ の順番を入れ換えた順列でもので $i_1 = 1$ を満たすものとする。この時、$U(x)$ を

$$U(x) = \sum_{k=1}^{N} a_{i_k i_{k+1}} \text{ただし } i_{N+1} = 1 \text{ とする}$$

で定める。問題は「$U(x)$ を最小にする順列を求めよ」というものである（これは "Traveling Salesman" の問題と呼ばれる）。

この問題の解法の決定版は知られていない。これを解く方法はいろいろなものが提案されているが、この中で最近注目されている "Simulated Annealing" と呼ばれる確率のアイデアを用いた方法を講演の中で紹介する。

以下にそのアイデアを簡単に述べておく。Γ は $E \times E$ の部分集合で

$$(x, x) \in \Gamma, x \in E$$

$$(x, y) \in \Gamma \quad \text{ならば} \quad (y, x) \in \Gamma$$

をみたすものとする。$(x, y) \in \Gamma$ の時、x と y はグラフで結ばれているものと考える。そしてこのグラフは連結であるとする。いま、$n : E \to \mathrm{N}$ および $p : E \times E \to [0, 1]$ を

$$n(x) = \sharp\{y \in E;\ (x, y) \in \Gamma\}, \quad x \in E$$

$$p(x, y) = \begin{cases} n(x)^{-1}, & (x, y) \in \Gamma \\ 0, & (x, y) \notin \Gamma \end{cases}$$

により定める。さらに、$\beta \geq 0$ に対して $p_\beta : E \times E \to [0, 1]$ を

$$p_\beta(x, y) = \begin{cases} \exp(-\beta(U(x) - U(y)))p(x, y), & x \neq y \\ 1 - \sum_{z \in E \setminus \{x\}} \exp(-\beta(U(x) - U(z)))p(x, z), & x = y \end{cases}$$

により定める。さて、$\beta > 0$ を一つ固定し、$\beta_n = \beta \cdot \log n$, $n = 1, 2, \ldots$ とおく。さらに次のように確率過程 $\{X_n; n = 0, 1, 2, \ldots\}$ を構成する。

$$P(X_1 = x) = (\sharp E)^{-1}$$

$$P(X_n = y | X_0 = x_0, \ldots X_{n-1} = x_{n-1}) = p_{\beta_n}(x_{n-1}, y), \quad n \geq 1$$

このような確率過程はもし理想的な乱数が存在すれば簡単に構成できる。

Simulated Annealing の基本定理は以下のようなものである。

定理 2 次のような β_0 が存在する。

(1) もし、$\beta > \beta_0$ ならば X_n は確率 1 で U の最小点に近づく。

(2) もし、$\beta < \beta_0$ ならば X_n が最小点以外のものに近づく確率が正である。