数学特論 B4 のレジュメ

1 月 23 日 (木曜 2 限)

本日の目標 (最終回):
前回はトランプのリフル・シャッフル (riffle shuffle) を 52 次の置換群（対称群）を状態空間とするマルコフ連鎖としてモデル化しました (GSR 模型). 引き続きこれを $(X_n)_{n\geq 0}$ と表す. ちなみに文献 [IOTM] によると, $52! \approx 8.066 \times 10^{67}$ だそうです.

さてこのマルコフ連鎖 $(X_n)_{n\geq 0}$ は既約かつ非周期的だったので, 一意的に定常分布が存在して, マルコフ連鎖の時刻 n における確率分布は $n \to \infty$ のときにその定常分布に収束するのでした. しかも群の対称性があるために, その定常分布が置換群上の一様分布になります.

今回はこの X_n が誘導する分布の一様分布への収束の速さ を問題にします.

[注意] 確率論に非常によく出てくる考え方ですが, 考えている系の「状態」というのを状態空間の上の確率分布だと思っています. この考え方を採用すれば「トランプの束がよく切れている状態」を「置換群上の一様分布」とみなすことはごく自然です（そもそもそれ以外の選択肢はなさそうだ).

したがってランダムな操作をトランプの束に n 回繰り返した「状態」は X_n が誘導する分布のことなのである（つまり X_n を写像または関数だと思ったときの値そのものを問題にしているわけではない). この X_n が誘導する分布と一様分布の「距離」が近い時に「n 回のランダムな操作によりトランプの束がよく切れた」とみなしてシャッフル問題を定式化するわけだ.

確率分布の収束 (または確率分布間の距離) を問題にしてるので, 例えば次のような疑問 (指摘) はナンセンスである.
「任意の n に対して X_n が整列状態（すなわち単位元）である確率は 0 でないので, この方法ではいつまでたってもよく切れない」

確率分布間の距離: S を有限集合として, S 上の確率分布全体の集合 $\mathrm{Prob}(S)$ 上に距離を定義する. これは**全変動距離 (total variation distance)** と呼ばれる. μ, ν を S 上の確率分布とするとき,

$$\|\mu - \nu\|_{\mathrm{TV}} := \max_{A : A \subset S} |\mu(A) - \nu(A)|, \qquad \text{ただし } \mu(A) := \textstyle\sum_{x\in A} \mu(x).$$

と定める. ただし max は S の全ての部分集合 A を渡って取る. この全変動距離が確率分布の間の距離として適切である. すると確率分布の列の収束 "$\lim_{n\to\infty} \mu_n = \nu$" は当然ながら $\lim_{n\to\infty} \|\mu_n - \nu\|_{\mathrm{TV}} = 0$ として定義される.

自明ではないが, 少々計算すると

$$\|\mu - \nu\|_{\mathrm{TV}} = \frac{1}{2} \sum_{x\in S} |\mu(x) - \nu(x)| = \sum_{x\in S : \mu(x) > \nu(x)} \{\mu(x) - \nu(x)\}$$

となるので, これは三角不等式をみたしていることがわかり, 本当に距離になっていることが見える.

• 実は一般的に既約かつ非周期的なマルコフ連鎖の定常分布への収束は非常に速い（指数的に速い）ことが比較的簡単に証明できる. すなわち, ある $0 < \alpha < 1$ と $n_0 \in \mathbf{N}$ が存在して, 任意の $n \in \mathbf{N}$ と $x \in S$ に対して,

$$\|P^n(x, \bullet) - \pi\|_{\mathrm{TV}} \leq \min\{\alpha^{n-n_0}, 1\}$$

が成立する. ここで π は一意的な定常分布で, $P^n(x, \bullet)$ は初期値 x(あるいは初期分布 δ_x といってもよい) の場合に X_n が誘導する分布である.

♠ さて 52 枚のカードのリフル・シャッフルの話に戻ろう．初期値は恒等置換 Id（すなわち「整列状態」）である．定常分布は置換群上の一様分布であった．これを π と書く．

実はこの場合の収束の速さは前ページで述べた「指数的な収束」よりもさらにずっと速い，きわめて急激な収束であることが知られている．マルコフ連鎖のこの種の振る舞いは「急混合現象」と呼ばれて最近では活発に研究されている．（ここでの説明はかなりの手抜きである．もし詳細に興味があれば [LPW, IOTM] などを参照せよ．）

ダイアコニスに代表される多くの優秀な数学者が時刻 n におけるマルコフ連鎖の誘導する分布と π との距離

$$\|P^n(\mathrm{Id}, \bullet\,) - \pi\|_{\mathrm{TV}}$$

を具体的に計算した結果，以下のようなことがわかった．

参考文献

[LPW] Levin, David A.; Peres, Yuval; Wilmer, Elizabeth L.; Markov chains and mixing times. American Mathematical Society, Providence, RI, 2009.

[IOTM] 池田信行，小倉幸雄，高橋陽一郎，眞鍋昭治郎: 確率論入門 2, 培風館, 2015.