

自主課題レポート

相関と偏りのある情報源を用いた乱数生成

63：守 優輔

平成 19 年 2 月 7 日

1 目的

相関と偏りのある情報を入れて全く独立した乱数を生成する。そのために有名な J.von Neumann のアルゴリズムと M.Blum のアルゴリズムを用い、それらの性質などの理解を深める事を目的としている。

2 内容

アルゴリズム vN：

入力：0 と 1 の列、それを $\sigma_1, \sigma_2, \dots$ とする。

出力： H と T の列。

vN は同じ確率で H と T を出力する。

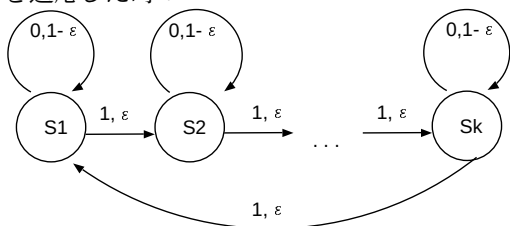
マルコフ連鎖：離散時間、離散状態の確率過程 $X_n, n = 0, 1, 2, \dots$ において、「次の時点の状態確率は現在の状態のみに依存する」と仮定した時、次の時点の状態確率は、確率変数を用いて表すと、

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots) = P(X_{n+1} = j \mid X_n = i)$$

となります。この式が成り立つ時の確率過程はマルコフ性を持つといい、マルコフ性を持つ離散状態の確率過程 $X_n; n = 0, 1, 2, \dots$ をマルコフ連鎖という。

定理 A (アルゴリズム A は悪い)： MC を $S_1, \dots, S_k$ の k 状態から成るマルコフ連鎖としよう。各状態から出る 0 の辺は確率 $q_1 = 1 - \epsilon$ が関連づけられ、自分自身の状態に戻る；各状態から出る 1 の辺は確率 $q_i = \epsilon$ が関連づけられ、 $i < k$ に対して S_i を S_{i+1} に写し、 S_k を S_1 へ写像する。

この MC によって生成される系列にアルゴリズム A を適応した時：



1, (アルゴリズム A の出力列は、初め、偏りがある) : $\epsilon \rightarrow 0$ (または $\epsilon = 1/2$) に対して、アルゴリズム A は、最初の H を出力する前に、約 $k/2$ 個の T (約 $k/3$ 個の T) を出力すると期待される。

2, (アルゴリズム A は独立でない列を発生させる) : A の t 番目の出力を H_t または T_t で表す。固定した $k > 1$ に対して $\epsilon \rightarrow 0$ の時、 $\Pr[T_{t+1}T_{t+2} \mid H_{t-k} \dots H_t] \rightarrow 0$ となる。特に、 $k = 2$ に対して、 MC が出力する最後の二つのシンボルが、共に H であるという条件の下、 MC が二つの連なる T を出力する確率は漸近的に 0 となる (漸近的に $1/4$ になるべきであるが)。

定理 B (アルゴリズム B は良い)。 MC をマルコフ連鎖とする。 MC により生成された列をアルゴリズム B への入力に用いられるとする。このとき：

1. B の出力は独立した偏りのない (おそらく有限の) H と T の列である。

2. MC が、全ての状態から、確率 = 1 を随伴する単一の辺が存在するような、決定論的ループに入らないならば、 B の出力は確率 1 で無限に長い。

3. 辺に付けられたすべての確率が $1/2$ ならば、 B は、4 入力ビットごとに対して、一つの (独立な偏りのない) 出力ビットを生成すると期待される。

3 まとめ

$vN \rightarrow A \rightarrow B$ とアルゴリズムは複雑になるが、相関と偏りのある情報源から乱数を生成するのにいいアルゴリズムとなっていく事が分かった。