

Moran 過程の数値シミュレーション

2011SE195 夏川将実

指導教員：小藤俊幸

1 はじめに

有限集団での進化ダイナミクスの解析を始めよう。個体数は連続変数ではなく、整数で与えられる。その結果、進化ダイナミクスはもはや決定論的な微分方程式では記述されず、確率論的な定式化が必要である。

生物学的な問題を研究するための最良のアプローチは、まず決定論的な記述を試み、それが適切な側面を損なっているときのみ、確率論的な解析へ移行することである。この研究では有限サイズの集団における中立的浮動と一定の淘汰について調べ、ダーウィンの進化論に相対する中立説を指示していきたい。

2 中立的浮動

サイズ N で固定されたある集団を考えよう。個体の種類は A と B の 2 種類ある。個体は同じ速度で繁殖する。したがって、A と B は選択に関して中立な変異体である。すべての単位時間ステップにおいて、1 個体がランダムに選択されて繁殖し、かつ 1 個体がランダムに選択されて除去される。ここでは復元抽出を用いよう。すなわち、同一の個体が繁殖と死亡に選ばれることも起こり得る。繁殖は突然変異することなく起こる。すなわち A は A を、B は B を生産する。

この確率過程はオーストラリアの集団遺伝子学者 P.A.P.Moran にちなんで名付けられており、1958 年に考案された。各時間ステップで常に 1 個体の出生と 1 個体の死亡が起こるため、集団サイズは必ず一定にあることが保証されている。確率変数は個体 A の個体数のみであり、ここでは i と表そう。個体 B の個体数は $N - i$ である。1 変数の確率過程を調べることは、2 進数あるいはそれ以上の変数の過程よりも非常に簡単である。

Moran 過程は状態空間 $i = 0, \dots, N$ で定義される。個体 A を (出生または死亡で) 選択する確率は i/N で与えられる。個体 B を選択する確率は $(N - i)/N$ で与えられる。すべての単位時間ステップにおいて、次の 4 つの事象が起こる可能性がある。

(i) 繁殖と死亡に A の 1 個体を選択され、この事象が起こった後の個体 A の数は、起こる前の数と同じであり、 i は変化しない。この事象の確率は $(i/N)^2$ である。

(ii) 繁殖と脂肪に B の 1 個体を選択され、この事象が起こった後の個体 B の数は、起こる前の数と同じであり、 $N - i$ は変化しない。この事象の確率は $[(N - i)/N]^2$ である。

(iii) 繁殖に A の 1 個体が、死亡に B の 1 個体を選択され、この事象が起こった後、A は 1 個体増え、変数 i は $i + 1$ となる。この事象の確率は $i(N - i)/N^2$ である。

(iv) 繁殖に B の 1 個体が、死亡に A の 1 個体を選択され、この事象が起こった後、A は 1 個体減り、変数 i は $i - 1$ となる。この事象の確率は $i(N - i)/N^2$ である。

3 遷移確率

x_i^n を時刻 n で i 個である確率とし、

$$P_{i,i-1} = \frac{i(N-i)}{N^2} \quad (1)$$

$$P_{i,i} = 1 - \frac{i(N-i)}{N^2} - \frac{i(N-i)}{N^2} \quad (2)$$

$$P_{i,i+1} = \frac{i(N-i)}{N^2} \quad (3)$$

とおく。時刻 n で $i - 1$ 個のとき、時刻 $n + 1$ で i 個となる確率は $P_{i,i-1}x_{i-1}^n$ 、時刻 n で i 個のとき、時刻 $n + 1$ で i 個となる確率は $P_{i,i}x_i^n$ 、時刻 n で $i + 1$ 個のとき、時刻 $n + 1$ で i 個となる確率は $P_{i+1,i}x_{i+1}^n$ であるから、

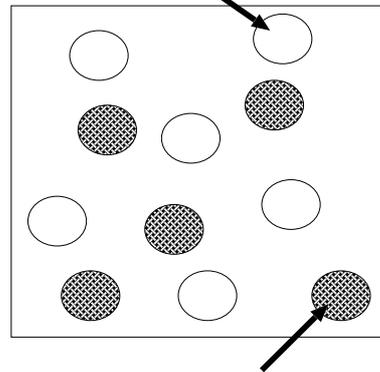
$$x_i^{n+1} = P_{i-1,i}x_{i-1}^n + P_{i,i}x_i^n + P_{i+1,i}x_{i+1}^n \quad (4)$$

となる。

4 プログラム

Moran 過程をシミュレートするプログラムを作成した。

1 個体を繁殖に選択する



そして 1 個体を死亡に選択する

2 番目に選択された個体は 1 番目に選択された個体の子供に置き換えられる

プログラムでは集団 A, B を数字の 1, 0 に置き換え、乱数を発生させるため rand 関数を用いた。

4.1 プログラム

以下にソースリストを示す。

```
main()
{
```

```

int x[N];
int i, n, p, q;

srand48(10000);

for(i = 0; i < N/2; i++)
{
    x[i] = 0;
}
for(i = N/2; i < N; i++)
{
    x[i] = 1;
}

for(n = 0; n < 100; n++)
{
    p = floor(N*drand48());
    q = floor((N - 1)*drand48());
    if (q >= p) q = q + 1;
    printf("n = %d, p = %d, q = %d\n", n, p, q);
    for(i = 0; i < N; i++)
    {
        printf("%d ", x[i]);
    }
    printf("\n");
    x[q] = x[p];
}
}

```

このプログラムは初期値として、0, 1 を 5 個体ずつ用意し、乱数により上記の (i)(ii)(iii)(iv) の事象を起こさせるように変化していくかを表したものである。このプログラムを用いて発展させていく。まずはこのプログラムを動かすかのようにするのかを検証する。

4.2 乱数について

drand48 は区間 [0, 1) 上の一様乱数 (疑似乱数) を生成する関数である。m = 248 = 281474976710656 を法とする乗数 (multiplicand) a = 0xfdece66d = 25214903917 加数 (addend) c = 0xb = 11 の線形合同法 (linear congruential method) (1) $rn+1 = arn + c \pmod{m}$ により生成される乱数を m で割って、[0, 1) 上の乱数としている。注 0x は 16 進数を表す。最初の 0x を除いた部分が「数字」で a, b, c, d, e, f は、それぞれ、10, 11, 12, 13, 14, 15 に対応する。

4.3 結果

4.1 のプログラムの実行結果を示す。0 と 1 では表しにくいので、0 を ●, 1 を ○ とする。

```

n=0 ●●●●●○○○○○
n=1 ●●●●●●○○○○
n=2 ○●●●●●○○○○
n=3 ○●●●●●●○○○
n=4 ●●●●●●●○○○
n=5 ●●●●○●●○○○
n=6 ●●●●○●●○○○
n=7 ●●●●○●●○○○

```

```

n=8 ●●●●●○○●○○
n=9 ●●●●●○○●○○

```

この結果では n=70 ですべてが ● になった。

5 比較

最初の 1 の数が半分 (5 個) としたとき、100 回目に 1 が i 個現れる確率 (変数 $x[i]$) を遷移確率から計算するプログラムと、実際に乱数を使いシミュレーションした結果を比較する。遷移確率での結果はこのようになった。

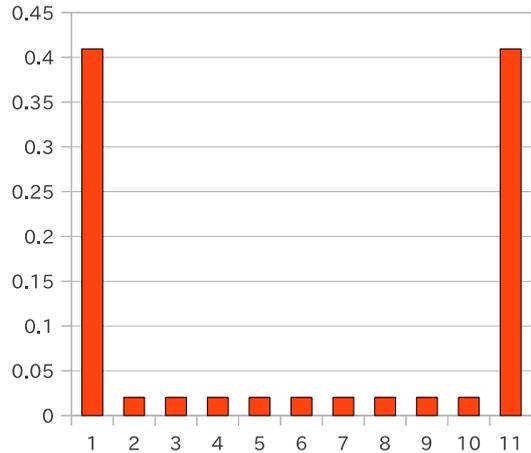


図1 シミュレーション結果

上記のプログラムを発展させ、n = 99 の結果のみを 100 個表示するプログラムを作った。更に結果の 1 の数を数えるように組み、遷移確率と比較できるようにした。0 と 1 の数は 1:9 とし、1 の個数はデータの平均をとった。

1 の数	0	1	2	3	4	
個数	7.67	0	0.33	0.33	0.66	
確率	0.07	0.01	0.01	0.01	0.01	
1 の数	5	6	7	8	9	10
個数	0	0	0	0	0.66	90.33
確率	0.01	0.01	0.01	0.01	0.01	0.87

6 終わりに

遷移確率と中立的浮動のシミュレーションの結果はほぼ一致するという結果になった。これより、遺伝子の変化は自然淘汰に対して有利でも不利でもなく中立であるという中立説を指示する。

7 参考文献

参考文献

[1] MARTIN A. NOWAK [著] 竹内康博・佐藤一憲・巖佐庸・中岡慎治 [監訳]: 『進化のダイナミクス』。共立出版, 東京, 2008.