

集団遺伝学における拡散過程の 数値解法とその応用について

遺伝研 高畑尚之 丸山毅夫

集団遺伝学の究極的な目的は集団中に存在する遺伝的変異保有の機構を明らかにし更に生物の進化機構を解明することにある。集団の遺伝的変異に変化を及ぼす要因は突然変異、自然淘汰、移住、遺伝子内の組換えなどいくつか考えられるが、本質的に確率論的取り扱いが必要になるのは実際の生物集団が有限であって、有限個の配偶子しか次世代に寄与しない効果を考慮する場合である。特に近年、遺伝子レベルでの変異が実験的に同定できるようになり、それを理解するにはこの確率的要因が無視できない事が強調されてきた。この要因は数学的にはS. WrightやR.A. Fisherによって定式化され、M. Kimuraによって広く分子レベルの遺伝的変異と進化を理解する上で適用されてきた。その方法はKolmogorov方程式を利用するものでこの方法以外では取り扱いが困難な問題も比較的容易に解を得ることが出来るようになった。しかし遺伝子座

あたりの対立遺伝子の数が多い場合とか2つ以上の遺伝子座に相互作用がある場合などを扱う時は一般的な解を得ることはむづかしい。

ここでは上の Kolmogorov 方程式に等価な Ito の確率微分方程式を計算機を用いて数値解析する方法を考えその具体的な応用を2, 3行なったのでその結果について述べる。

まずその方法の概略を述べる。今ある時点まで1つの集団中に $K+1$ 個の対立遺伝子が分離しているとし, i 番目の遺伝子頻度を x_i とすると, Kolmogorov 後向き方程式は

$$\frac{\partial \phi}{\partial t} = \sum_{i=1}^K M_{\delta x_i} \frac{\partial \phi}{\partial x_i} + \frac{1}{2} \sum_{i=1}^K \sum_{j=1}^K x_i (\delta_{ij} - x_j) \frac{\partial^2 \phi}{\partial x_i \partial x_j} \quad (1)$$

となる。ここで ϕ は遷移確率密度, $M_{\delta x_i}$ は決定論的な要因による遺伝子頻度の変化量, $\delta_{ij} = 0$ ($i \neq j$), $\delta_{ij} = 1$ ($i=j$)
そして $\sum_{i=1}^{K+1} x_i = 1$ である。これに対応する Ito の確率微分方程式は

$$a_{ij} = x_i (\delta_{ij} - x_j) \quad (2)$$

を要素にもつ $K \times K$ の正方行列 $A = \{a_{ij}\}$ と K 個の独立な Brown 運動 $dB_i(t)$ を用いて

$$dx_i = M_{\delta x_i} dt + \sum_{j=1}^K \sqrt{A_{ij}} dB_j(t) \quad (3)$$

$i=1, \dots, K$

と書ける。ただし $E\{dB_i(t)\} = 0$, $E\{dB_i(t)dB_j(t)\} = \delta_{ij}dt$

で $\sqrt{A_{ij}}$ は $K \times K$ の正方行列 $\sqrt{A}$ の i, j 要素を表わすものとする。

われわれの方法は(3)式に対応する確率差分方程式

$$x_i(t + \Delta t) = x_i(t) + M_{\delta x_i} \Delta t + \sum_{j=1}^K \sqrt{A_{ij}} \Delta B_j(t), \quad (4)$$

$$E\{\Delta B_i(t)\} = 0$$

$$E\{\Delta B_i(t) \Delta B_j(t)\} = \delta_{ij} \Delta t$$

を計算機を用いて解くことである。ここで注意すべきことは(4)式で十分 Δt を小さくすれば確率 1 で(4)は(3)に収束させうることで実際の計算を行なう場合にも Δt を十分小さくとることが必要である。また

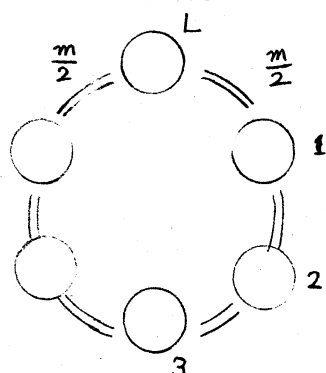
$$\sum_{i=1}^K x_i \leq 1, \quad x_i \geq 0 \quad (i = 1, \dots, K)$$

で定義される境界近傍のふるまいを正しくシミュレートするにはやはり Δt を小さくとる必要がある。しかし、このことは必ずしも一様に Δt を小さくとらなければならないということの意味しない。また問題によってはかなり大ざっぱに Δt を選んでもよい結果が得られる場合がある。以下の幾つかの応用例においては境界近傍を除いて Δt を経験的に

$$\max_{i, x_i} |M_{\delta x_i}| \Delta t \leq 10^{-2} \sim 10^{-3} \quad (5)$$

になるように決めた。

応用例 1, Geographically structured population



集団が図のような L 個の colonies から成り各 colony は N 個の個体から成るとする。colonies 間の migration rate を m , 対遺伝子の数は 2 個とすると自然淘汰の作用のもとで i 番目の colony の遺伝子頻度 x_i は

$$\begin{aligned} \dot{x}_i = & 2N \left\{ -mx_i + \frac{m}{2}(x_{i+1} + x_{i-1}) + sx_i(1 - x_i) \right\} dt \\ & + \sqrt{x_i(1 - x_i)} dB_i(t) \end{aligned} \quad (6)$$

$$0 \leq x_i \leq 1, \quad i = 1, \dots, L$$

なる Itô 方程式で書ける。集団構造は遺伝的変異の減少と地理的分化に関連した重要な問題で T. Maruyama によって詳しく解析されているので適当な初期条件のもとでシミュレーションの結果を比較することができる。特にこの場合は集団構造に不変な量として固定確率, ヘテロザイゴシティの和,

$$\sum_{t=0}^{\infty} \frac{2}{L} \sum_{i=1}^L x_i(t)(1 - x_i(t))$$

が知られているのでその比較検討

を試みた。

応用例 2, Two locus two allele model

1つの集団で2つの連鎖した遺伝子座を考えそれぞれの座に2つの対立遺伝子が分離している場合を考える。この問題は重複遺伝子の進化的運命を論じたり遺伝子間の相互作用の

問題を量的に把握する上で重要である。今、第1の座で A , a , 第2の座で B , b と2つの対立遺伝子があるとし, ab , aB , Ab , AB のタイプの染色体の頻度を x_1, x_2, x_3, x_4 とする。突然変異は $A \rightarrow a$, $B \rightarrow b$ に一方的に一定の rate v で起き, 20の座の組換え率を c とし自然淘汰は $aabb$ だけの個体が致死と仮定すると Itô 方程式は

$$\begin{aligned}
 dx_1(t) &= 2N[v(x_2+x_3) - \{x_1(1-x_1) + cD\}/(1-x_1^2)]dt \\
 &\quad + \sum_{j=1}^3 (\sqrt{A})_{1j} dB_j(t) \\
 dx_2(t) &= 2N[v(x_4-x_2) + \{x_1^2 x_2 + cD\}/(1-x_1^2)]dt \\
 &\quad + \sum_{j=1}^3 (\sqrt{A})_{2j} dB_j(t) \\
 dx_3(t) &= 2N[v(x_4-x_3) + \{x_1^2 x_3 + cD\}/(1-x_1^2)]dt \\
 &\quad + \sum_{j=1}^3 (\sqrt{A})_{3j} dB_j(t) \\
 dx_4(t) &= - \sum_{i=1}^3 dx_i(t)
 \end{aligned} \tag{7}$$

の形で与えられる。ここで

$$D = x_1 x_4 - x_2 x_3, \quad \sum_{i=1}^4 x_i = 1 \tag{8}$$

である。

調べた量は初期条件 $x_4(0) = 1$, $x_i(0) = 0$ ($i=1, 2, 3$) のもとでどちらか一方の遺伝子座が a または b に固定するまでの

持ち時間に関するものである。従来のシミュレーションは N が大きい場合はほとんど不可能に近かったが ($N < 10^3$)、この方法によっておおよそ $N = 10^6$ 位までは適当な時間内で実行可能になった。

応用例 3, K allele model

1つの集団で、 $K + 1$ 個の対立遺伝子が分離する 1 突然変異によってそれぞれの遺伝子は v の rate で自分以外の遺伝子に同等に変化する場合を考える。自然淘汰に関しては (i) 中立の場合、(ii) 超優性の場合、(iii) 淘汰強度がガンマ分布に従う場合を調べた。

中立の場合を例にとる。 i 番目の遺伝子頻度を前と同様に x_i とすると Itô 方程式は

$$dx_i = 2N[-vx_i + v(1-x_i)]dt + \sum_{j=1}^K (\sqrt{A})_{ij} dB_j(t) \quad (9)$$

となる。(自然淘汰がある場合には右辺第 1 項に対応する項が付加される。) この場合は (I) の定常解が存在するのでヘテロザイゴティ $1 - \sum_{i=1}^{K+1} x_i^2$ の平均、分散や遺伝子頻度の定常分布を求めた。中立の場合はその結果が知れているので方法のチェックに用い、(ii) と (iii) の場合は得られている近似解との比較を行なった。注意すべき点は $2Nv \ll 1$ のときは特に Δt を境界近傍のものに対しては小さくとることである。(4)

式で偶然大きく負の値になったものを0としたり, 1を超えたものを1とすることは結果的に中心方向に *bias* を生じ誤まった結果を得ることになる。

以上の例から境界近傍の取り扱いに注意すれば, この Ito 方程式を利用する方法は多くの今まで解きえなかった集団遺伝学的に興味ある問題を解析する 1つの有力な方法と思われる。