

12 Individual fitness

個体群の適応度とは、その存続可能性である。結果ではなくて可能性の総体が個体群の適応度である。小惑星の衝突のような極端な事象が原因で、完全に fit した個体群が絶滅してしまうことは容易に想像できる。しかし安定した状況においても、個体群の存続は *stochasticity* の影響を受けやすい。適応度は（存続・絶滅の）結果以前に存在している個体群の特徴であり、適応度について論じるには個体群変化の数学的モデルが必要である。適応度は潜在的な特徴 (*latent feature*) であり、モデルパラメータの関数として定義される。

地理的に閉じた個体群の存続は個体の生存率および新しい個体の加入率に依存する。生存率や加入率は個体数の比である（ある期間で、サイズ N の個体群で X 個体が生存した場合、生存率は X/N であり、加入個体が Y 個体であれば加入率は Y/N である。これらの率そのものは、前もって決まった (*a priori*) 潜在的量ではなく、実際に起きた事象と結びついているので、適応度ではない。数学的モデルのパラメータを推定するのにこれらを使うことはできる。適応度はこれらのパラメータの関数である。

実際に起きたことと潜在的な可能性の区別、推定値とパラメータの区別は重要である。たとえば X を生起回数 N 、生起確率 P の二項確率変数、 Y を平均 $F \times N$ のポアソン確率変数とする事が出来る。しかしこれは単純化のしすぎで、個体群では齢別にバリエーションがある可能性がある。そこで、モデルの記述を齢段階別のパラメータ F_i 、 P_i と拡張する。

個体群適応度は F_i 、 P_i で記述される。生存率や加入率は齢段階以外の要因の影響を受けるかもしれない。時間変化する事象が影響するなら、時間的バリエーションを考慮したモデルを作り、それに応じて適応度の定義を変える必要がある。しかも生存や加入に与える要因には個体差 (*individual heterogeneity*) がある。

この章では、個体群適応度のアナロジーをもちいて、個体適応度について考える。個体特異的な生存や繁殖パラメータを考える。このような文脈では実際に起きた事 (*realization*) と潜在的な可能性 (*capability*) の区別は重要である。これは各個体の生涯は一度しかないからである。個体適応度は *derived parameter* であり、階層モデルで最もよく扱える。

12.1 Population fitness

地理的に閉鎖していて、最大寿命 T の齢構造化個体群において、時間 t における齢段階 i の個体数を $n_i(t)$ とすると、

$$\begin{bmatrix} F_1 & F_2 & F_3 & \cdots & F_{T-1} & F_T \\ P_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & P_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & P_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \cdots & P_{T-1} & 0 \end{bmatrix} \begin{bmatrix} n_1(t) \\ n_2(t) \\ n_3(t) \\ \vdots \\ n_T(t) \end{bmatrix} = \begin{bmatrix} n_1(t+1) \\ n_2(t+1) \\ n_3(t+1) \\ \vdots \\ n_T(t+1) \end{bmatrix}, \quad (12.1)$$

または

$$\mathbf{Mn}(t) = \mathbf{n}(t+1)$$

累積生存率を $S_1 = 1$, $S_i = P_{i-1}S_{i-1}$, $i = 2, 3, \dots, T$ と定義する。

$S_i F_i$ の少なくとも一つが正の値を取るなら、 \mathbf{M} は正の固有値をもつ。これは次式の解 x と等しい。

$$\sum_{i=1}^T (S_i F_i) x^{T-i} = x^T \quad (12.2)$$

成分同士の積を $\{\mathbf{SF}\}$ とする。これにより Eq. (12.2) は $\lambda(\{\mathbf{SF}\})$ と書ける。

行列 \mathbf{A} の固有値 λ と固有ベクトル \mathbf{x} は $\mathbf{Ax} = \lambda\mathbf{x}$ という関係がある。したがって

$\mathbf{n}(t)$ が固有ベクトルであるとき、次の時間における個体数は $\mathbf{n}(t+1)$ である。二つのベクトルは比例関係にあり、 $\mathbf{n}(t+1)$ は $\lambda(\{\mathbf{SF}\})$ でスケール化されている。この個体群は安定年齢分布にある、という。 $\lambda(\{\mathbf{SF}\})$ は個体群増加率である。

$\lambda(\{\mathbf{SF}\})$ は 1 より大きければ個体群は増加し、1 より小さければ個体群は減少することが示唆される。このように $\lambda(\{\mathbf{SF}\})$ は生存と繁殖の個体群への影響をまとめてあらわしたもので、個体群の適応度の良い指標であると考えられる。Panel 12.1 は $\lambda(\{\mathbf{SF}\})$ を計算する R コードである。

行列モデルの詳細は Caswell (2001) を参照

脚注 1 個体群が安定年齢分布（あるいはその近傍）に達していなければ $\lambda(\{\mathbf{SF}\}) > 1$ でも個体群が増加している保証はない。さらに、パラメータは実際の変化ではなく、変化率の期待値を表しているにすぎない。

脚注 2 もし常に $S_i F_i = 0$ なら、全ての個体は繁殖前に死んでしまう。このときは $\lambda(\{\mathbf{SF}\}) = 0$ である。

12.2 Individual fitness

適応度は遺伝子型が次世代にどれくらい広がるかを表している。もし個体群が遺伝的に均一な個体で構成されていれば、遺伝子型の適応度は $\lambda(\{\mathbf{SF}\})$ となる。

$\lambda(\{\mathbf{SF}\}) > 1$ ならその遺伝子型は次世代で頻度を増やすし、 $\lambda(\{\mathbf{SF}\}) < 1$ ならその遺伝子型は消えていくだろう。

しかしながら、遺伝的に均一な個体群というのはあまりない。個体固有の産子数と生存率を、個体ごとの遺伝的構成に対応した潜在的な特性 (latent feature) と考えるほうが適切である。

個体 j の生存率と繁殖成功は生まれたときに決まるパラメータによって支配されているとする。年齢階 i から $i+1$ への移行は成功率 $P_i^{(j)}$ のベルヌーイ試行である。年齢階 i において産出される子の数は期待値 $F_i^{(j)}$ の確率変数 $f_i^{(j)}$ である。累積生存率を $S_i^{(j)} = S_{i-1}^{(j)} P_{i-1}^{(j)}$ と定義する ($S_1^{(j)} = 1$)。

個体の適応度は個体群の適応度と同じように定義できる。

$$\lambda^{(j)} = \lambda(\{S^{(j)} F^{(j)}\})$$

個体の適応度は潜在的な量で、遺伝子型毎に決まる潜在的特質をまとめたものである。

* 以下添字 j を省略するが、個体ごとの値であることを注意。

個体の適応度は観測可能な量なのか？適応度は一度きりの生活史を支配するパラメータで決まるので反復して観測することができない。しかしそれに関わらず個体の適応度は単一の生活史から得られた情報だけで推定することができる。これは $S_i F_i$ が推定可能であるからであ

る。齢段階 i で生存しているかどうかを I_i であらわすと、 $I_i f_i$ は $S_i F_i$ の偏りのない推定値になる。 $S_i F_i$ の推定値を Eq. (12.2)に代入し、固有値を求め個体の適応度とする。

脚注 3 個体がクローン繁殖する場合には反復が取れる。しかしこうした状況はまれである。

脚注 4 個体が死んだあとは確率変数 f_i は観測されないが、 $I_i = 0$ の時は $I_i f_i = 0$ と計算できる。 $I_i f_i$ は $S_i F_i$ の推定値であり F_i の推定値ではないことに注意が必要である。

MacGraw&Caswell (1996)は基本的にこの立場をとり、Eq. (12.1)の M のような個体ベースのレズリー行列の固有値として個体の適応度を定義している。生存年の S_i を1、 F_i を f_i とし、死亡年の S_i を0としている。個体が齢段階 X まで生存してから死ぬとする。 X 個の1と $T-X$ 個の0からなるベクトルを $\mathbf{1}_X, F_1, F_2, \dots, F_X$ と $T-X$ 個の0からなるベクトルを \mathbf{f} と書くと、 $\lambda(\{\mathbf{1}_X \mathbf{f}\})$ のようになる。

Latent individual fitness と Realized individual fitness の区別は重要である。 $\lambda(\{\mathbf{1}_X \mathbf{f}\})$ は realized individual fitness で、 $\lambda(\{\mathbf{S} \mathbf{F}\})$ は latent fitness である。Latent fitness は出生の時点で決まる。Realized fitness は latent なパラメータによって決められた分布を持つ確率変数であり、latent fitness は latent parameter で決定される関数である。

12.3 Realized individual fitness

Realized individual fitness $\lambda(\{\mathbf{1}_X \mathbf{f}\})$ はそれ自体興味の対象になりうるし、latent individual fitness と関係する量だが、latent individual fitness の推定値としてはあまり良いものではない。(省略)。同じ latent fitness を持つ2つの遺伝子型は、かなり異なる生活史を取りうる。そのときの realized fitness は、平均値でさえ、latent fitness とはかなり違う値になる。(省略)。

$T = 5$ (最大寿命) とし、年間生存率が80%とすると、

$$\mathbf{S} = (1.0000, 0.8000, 0.6400, 0.5120, 0.4096)'$$

簡単のために \mathbf{f} を固定する。

$$\mathbf{f} = \mathbf{F} = (0, 1, 2, 3, 2)'$$

Panel 12.1のRコードにより latent fitness は1.567である。Realized fitness はTable 12.1の5つの値のうちの一つの値をとる(確率 $\Pr(X)$)で。平均は1.282で標準偏差は0.702である。

この偏りは、特に固定された latent fitness の値に近ければ、そんなに悪くないかもしれない。しかし同じ latent fitness を導く別のベクトル \mathbf{S} と \mathbf{F} を作ることができる。

$$\mathbf{S} = \{1, 0.4000, 0.3200, 0.2560, 0.2048\}$$

$$\mathbf{F} = \{0, 2, 4, 6, 4\}$$

$\{\mathbf{S} \mathbf{F}\}$ は同じ値なので latent fitness は変化していない。しかしTable 12.1の最後の列の確率は0.6000, 0.0800, 0.0640, 0.0512, 0.2048となり、realized fitness は平均0.833 標準偏差1.032となる。後二回この操作を繰り返す(\mathbf{S} を1/2倍、 \mathbf{F} を2倍)と、平均値は0.54、0.35となる。Realized fitness は latent fitness の指標としては信頼性に欠ける。

12.4 Individual fitness in group context

Realized fitness の問題はサンプルサイズが1であることである。しかし、このようにする必要はなく、グループの中にいる個体を扱うべきである。この場合、各個体特有のパラメータは *stochastically* に関係していると考えられる。

グループの中の全ての個体のデータに基づき、 $P_i^{(j)}$ と $F_i^{(j)}$ の事後分布を得る階層モデルを作ることができる。アプローチは、ふたつの極端な方法の間をとることになる。一つは *realized fitness* を使うときのように、全ての個体は異なっていて別のものであるというものであり、もう一つは、個体群適応度 (*population fitness*) を計算するときのように、全ての個体を同一のものとしてとらえるものである。そして、個体適応度は *Derived parameter* として扱うことができ、その事後分布は $P_i^{(j)}$ と $F_i^{(j)}$ の事後サンプルから計算することができる。これらの量は個体がグループの中にいる場合にうまく推定され、単に *realized fitness* を使うよりも、グループ内の個体の適応度をうまく扱うことができる。

Cam *et al* (2002) は個体の効果を考慮し、kittiwakes (*Rissa trydactyla*) ミツユビカモメの生存率と繁殖に関する解析を行っている。それぞれの個体は0あるいは1個体の子を産むので、 $f_i^{(j)}$ は成功確率 $F_i^{(j)}$ のベルヌーイ試行としてモデル化することができる。Cam *et al* のモデルは $\text{logit}(P_i^{(j)}) = \alpha_p^{(j)} + \mathbf{x}'_{i,j} \boldsymbol{\beta}_p$ および $\text{logit}(F_i^{(j)}) = \alpha_f^{(j)} + \mathbf{y}'_{i,j} \boldsymbol{\beta}_f$ である。 $\mathbf{x}'_{i,j}$ と $\mathbf{y}'_{i,j}$ は個体固有の共変量 (*covariate*) であり、効果はグループレベルでモデル化される。個体効果 $\alpha_p^{(j)}$ および $\alpha_f^{(j)}$ は二変量正規分布からサンプリングされる。ベクトル $\boldsymbol{\beta}_p$ と $\boldsymbol{\beta}_f$ は年の効果と加齢による負の効果である。個体効果間の相関係数の事後分布は、平均 0.67、95% 確信区間 (*credible interval*) が (0.28 0.98) であり、このことから生存率が高い個体は、生存した時の繁殖確率 (条件付き確率) も高いことが分かる。個体間には生存率と繁殖確率にはかなりのばらつきがあった。Link *et al* (2002) はこのばらつきが個体の適応度に及ぼす影響を再解析して調べている。

この例を使うのは以下の点を強調するためである。未知の量全ての事後分布からなる $[\boldsymbol{\theta}|\text{Data}]$ が与えられたときに個体 j の適応度の推定はとても簡単になる。 $\boldsymbol{\theta}$ が与えられると、個体 j の $S^{(j)}$ と $F^{(j)}$ の計算に必要な全ての情報を得たことになる。

以下のように書ける。

$$\lambda^{(j)} \equiv \lambda^{(j)}(\boldsymbol{\theta}) = \lambda(\{S^{(j)} F^{(j)}\})$$

すなわち、個体 j の適応度を、未知のパラメータからなる $\boldsymbol{\theta}$ の既知の関数として考えることができる。

$[\boldsymbol{\theta}|\text{Data}]$ からのサンプリング、たとえば、 $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_B$ のマルコフ連鎖が与えられると、 $\lambda_1^{(j)} = \lambda^{(j)}(\boldsymbol{\theta}_1), \lambda_2^{(j)} = \lambda^{(j)}(\boldsymbol{\theta}_2), \dots, \lambda_B^{(j)} = \lambda^{(j)}(\boldsymbol{\theta}_B)$ を計算できる。これらの値は事後分布 $[\lambda^{(j)}|\text{Data}]$ からのサンプリングであり、これがベイズ推定の基礎である。 $\lambda^{(j)}$ は偏った分布になる傾向があるので中央値を点推定値に使うことが推奨される。(省略)。

12.5 Analysis of individual fitness: an example

Table 12.2 は 15 年間の生活史の 25 のシミュレーションである。これらのデータはミツユビカモメのモデルをもとに計算されている。ただし、年の効果と加齢による負の効果は考慮されていない。1 年目には繁殖せず、2 年目以降 15 年まで 0 または 1 の子を産む。(省略)。

個体特異的パラメータ $\text{logit}(P_i^{(j)}) \equiv \alpha_p^{(j)}$ および $\text{logit}(F_i^{(j)}) = \alpha_f^{(j)}$ としデータを生成する。これらの個体効果のペアは平均 $\mu_p = 3, \mu_f = -2$ 、標準偏差 $\sigma_p = \sigma_f = 0.50$ 、相関 $\rho = 0.50$ の二変量正規分布からサンプリングされている。

以下このデータの解析。 μ_p と μ_f には vague な正規分布 (mean = 0, standard deviation = 1000) を事前分布とする。 σ_p と σ_f には $U(0,5)$ 、 ρ には $U(-1,1)$ を事前分布とする。

実際のデータ (特にサンプルサイズの大きな) があれば、我々は典型的に時間の効果をモデルに含めて、それを天候のような観測値および時間毎の (time-stationary な) ランダム効果と関係させる、ということをする。そのような時間の効果は、個体の適応度 (latent fitness) とは関係ないとみなして、個体適応度の計算では 0 とすることもあるかもしれない。あるいは、個体の適応度を時間の効果が影響した結果であるとしたり、もしくは時間の効果のランダムな反復の平均と考えるかもしれない。ただし後者はかなり複雑な計算を必要とする。また、加齢が生存や繁殖に及ぼす効果をモデリングしたい場合があるかもしれないが、これらと関係のない時間効果から区別して解析するには、複数のコホートからのデータが必要である。

ここでの例は derived parameter を用いた推論について例示するためのもので、簡略化されている。そのような解析を頻度主義的な方法でどのように行うかという検討は読者に任せる。その時には、必要とされる近似を行う回数とそれに伴う不確実性について意識する必要がある。

ここで考えているモデルは、Table 12.2 の 25 個体それぞれについて、2 つの確率変数を指定している。生存率は年齢特異的ではないので、個体 j の生存期間 s_j はパラメータ $P^{(j)}$ を持つ幾何分布に従う確率変数である。個体 j の子の総個体数は二項分布に従う確率変数で試行回数は $n^{(j)} = \min\{s_j, 14\}$ である。

Panel 12.2 は BUGS コードである。ここでは、もし X_p と X_f が独立で、標準的で、正規分布に従う確率変数であるならば、 $\alpha_p = \mu_p + \sigma_p X_p$ 及び、 $\alpha_f = \mu_f + \sigma_p(\rho X_p + \sqrt{1 - \rho^2} X_f)$ は平均 μ_p, μ_f 、標準偏差 σ_p, σ_f 、相関 ρ の二変量正規分布を持つ、という性質を利用している。

幾何確率変数には二項尤度 (binomial likelihood) が使われていることにも注意。 $j = 1, 2, \dots, 16$ の個体は s_j 回生存したのちに死亡しているので、尤度として試行回数 (index) $s_j + 1$ 、観測値 s_j の二項確率変数を持つ。 $j = 17, 18, \dots, 25$ の個体は s_j 回生存し、かつまだ死亡していない。これらの個体の幾何尤度は、試行回数 (index) s_j 、観測値 s_j の二項確率変数である。個体 $j = 1, 2$ は繁殖齢に達する前に死亡しているので繁殖率に関する情報は含んでいない。BUGS により、burn-in 1000 の後に、マルコフ連鎖を 20,000 取った。次に個体ごとに、20,000 個の $\lambda(\{S^{(j)} F^{(j)}\})$ を計算した。3 個体の latent fitness の事後分布を Fig. 12.1 に作図した (1 (黒) 18 (赤) 25 (青))。

Table 12.2 にあるように、#1 は幼鳥のときに繁殖せずに死亡しており、#18 は最大の繁殖齢まで生存しているが、一度も繁殖していない。#25 は最大の繁殖齢まで生存しかつ、もっとも

多くの子を残している。

これらの周辺事後分布の間の重なりは少し誤解を生むかもしれない。 $\lambda^{(25)} > \lambda^{(18)}$ は強く支持され、 $\lambda^{(25)} > \lambda^{(1)}$ はそれなりに強く支持されている。この確からしさを評価するため、 $\lambda^{(25)} - \lambda^{(18)}$ と $\lambda^{(25)} - \lambda^{(1)}$ という **derived parameter** を新たに導入する。 $\Pr(\lambda^{(25)} - \lambda^{(18)} \geq 0 | \text{Data}) = \Pr(\lambda^{25} \geq \lambda^{18} | \text{Data}) = 0.993$ であり、 $\Pr(\lambda^{25} - \lambda^{1} \geq 0 | \text{Data}) = \Pr(\lambda^{25} \geq \lambda^1 | \text{Data}) = 0.954$ である。

別の **derived parameter** として、個体の適応度の他の個体の適応度に対する順位を考える。 $\lambda^{(25)}$ の他の $\lambda^{(j)}$ に対する大きさを調べるため、 $\lambda^{(j)}$ の中の $\lambda^{(25)}$ の順位を計算する（比較するのはMCMC中の2000個の各値）。

$$\Pr(\lambda^{(25)} = \max\{\lambda^{(j)} : j = 1, 2, \dots, 25\} | \text{Data}) = 0.616$$

さらに $\Pr(\text{Rank}(\lambda^{(25)}) \geq 22 | \text{Data}) = 0.948$ である。

Table 12.3 と Fig. 12.2 において **realized fitness** をベイズ推定による推定値を比較した。両者は真の値とよく相関していた (**realized fitness** と真の **latent fitness** の相関は 0.51 であり、真の **latent fitness** と事後分布の中央値は 0.48 であった)。しかし、平均二乗誤差はベイズ推定値では 0.098 であり、**realized fitness** では 0.701 であった。ベイズ推定値は 25 個体中 21 個体で真の値により近くなっている。