Selection in class-structured populations

Up to now: homogeneous populations

In a homogeneous population, a rare mutant can only ever be found in a single demographic state. Its success can therefore be reduced to the individual fitness of a representative mutant carrier:

$$\rho(y, x) = \text{individual fitness} = w(y, x). \tag{11}$$

This has made fitness bookkeeping easy: invasion fitness measures success at the level of the mutant lineage, while individual fitness measures the success of a mutant carrier—and the two are equal.

It also makes adaptation easy to think about: natural selection favours traits that improve individual fitness (although this does not necessarily mean maximising the growth rate of the population!)

Class structure changes the fate of a mutant lineage

But many populations are heterogeneous. Mutant carriers may differ in sex, age, habitat or physiological state: different contexts in which they can survive and reproduce.

The success of the mutant lineage must then combine what happens in all these contexts. A descendant entering one class may contribute much more to the future than a descendant entering another, and mutant copies may occupy some classes more often than others.

So how should these different contributions be combined to obtain $\rho(y, x)$?

The mean matrix follows mutant copies across classes

For a population divided into C classes, let $n_{j,t}$ be the number of mutant copies in class j at time t, and collect these numbers in the column vector

$$\mathbf{n}_t = \begin{pmatrix} n_{1,t} & \cdots & n_{C,t} \end{pmatrix}^{\mathsf{T}}. \tag{12}$$

The expected numbers of mutant copies at the next time step are

$$\mathrm{E}[\mathbf{n}_{t+1} \mid \mathbf{n}_t] = \mathbf{W}(y, x)\mathbf{n}_t, \qquad \mathbf{W}(y, x) = \big[w_{ij}(y, x)\big]. \tag{13}$$

where $w_{ij}(y, x)$ is the expected number of class-i mutant descendants produced by one class-j mutant (including itself if it survives).

The long-term fate of a mutant lineage

Let $\rho(y, x)$ be the leading eigenvalue of $\mathbf{W}(y, x)$, with right and left eigenvectors $\mathbf{q}(y, x)$ and $\mathbf{v}(y, x)$. We normalise them so that

$$\sum_j q_j(y, x) = 1, \qquad \mathbf{v}(y, x) \cdot \mathbf{q}(y, x) = 1. \tag{14}$$

Under the usual regularity conditions, for large t,

$$\mathrm{E}[\mathbf{n}_t \mid \mathbf{n}_0] = \mathbf{W}(y, x)^t \mathbf{n}_0 \sim \rho(y, x)^t \mathbf{q}(y, x) \big[\mathbf{v}(y, x)\mathbf{n}_0\big]. \tag{15}$$

- $\rho(y, x)$ determines whether the mutant lineage grows or declines asymptotically.
- The entry $q_i(y, x)$ is the probability that a randomly sampled mutant copy belongs to class i.
- The entry $v_i(y, x)$ is the **reproductive value** of a mutant copy in class i: its relative contribution to the distant future of the lineage. Thus, $v(y, x)\mathbf{n}_0$ is the total reproductive value of the mutant copies present initially.

The leading eigenvalue gives mutant invasion fitness

The previous result showed that, once the transient class dynamics have died away, the mutant lineage grows approximately as $\rho(y, x)^t$, where $\rho(y, x)$ is the leading eigenvalue of $\mathbf{W}(y, x)$.

In fact, according to standard multi-type branching process theory:

- if $\rho(y, x) > 1$, the mutant has a positive probability of establishing;
- if $\rho(y, x) \leq 1$, it eventually goes extinct.

Then we're back on familiar ground: differentiate $\rho(y, x)$, solve $S(x) = 0$, assess convergence stability and then examine $H(x^*)$.

Selection depends on class frequency and reproductive value

Let $\mathbf{q}^\circ$ and $\mathbf{v}^\circ$ be the right and left eigenvectors at neutrality ($y = x$), normalised so that $\sum_j \mathbf{q}_j^\circ = 1$ and $\mathbf{v}^\circ \mathbf{q}^\circ = 1$.

Differentiating the eigenvalue equation gives

$$S(x) = \left.\frac{\partial \rho(y,x)}{\partial y}\right|_{y=x} = \sum_{i,j} \mathbf{v}_i^\circ \left.\frac{\partial w_{ij}(y,x)}{\partial y}\right|_{y=x} \mathbf{q}_j^\circ . \tag{16}$$

Read each term from right to left:

1. $\mathbf{q}_j^\circ$: frequency of mutant copies in parent class j;
2. $\partial w_{ij}/\partial y$: effect of the trait on descendants entering class i;
3. $\mathbf{v}_i^\circ$: reproductive value of those descendants.

Life cycle in the two-patch example

Consider a population living in two patches, each containing a fixed number of adult breeding positions. The two patches present different environments, so the same phenotype may perform differently in each patch.

1. Adults in patch i reproduce with fecundity $f_i(z)$.
2. Each offspring moves to the other patch with probability m. Otherwise, it remains in its natal patch.
3. Adults die.
4. Offspring in each patch compete for the vacant breeding positions.

For a rare mutant, its class is the patch it currently occupies. On the board, we will translate this life cycle into the entries $w_{ij}(y, x)$ of the mean matrix $\mathbf{W}(y, x)$.

Two sources of disruptive selection in heterogeneous populations

At a singular strategy, $S(x^*) = 0$, and the Hessian can be decomposed as

$$H(x^*) = H_{\mathrm{w}}(x^*) + 2H_{\mathrm{q}}(x^*). \tag{17}$$

The two components are

$$\begin{aligned} H_{\mathrm{w}}(x^*) &= \sum_{i,j} \mathbf{v}_i^\circ \left.\frac{\partial^2 w_{ij}(y, x^*)}{\partial y^2}\right|_{y=x^*} \mathbf{q}_j^\circ, \\ H_{\mathrm{q}}(x^*) &= \sum_{i,j} \mathbf{v}_i^\circ \left.\frac{\partial w_{ij}(y, x^*)}{\partial y}\right|_{y=x^*} \left.\frac{\partial \mathbf{q}_j(y, x^*)}{\partial y}\right|_{y=x^*}. \end{aligned} \tag{18}$$

- H_{w}: disruptive selection generated by non-linear effects of the trait on fitness.
- H_{q}: disruptive selection generated by feedback between the trait and the classes occupied by the mutant lineage.

Summary: selection in class-structured populations

- Invasion fitness is the leading eigenvalue of the mutant mean matrix.
- The right eigenvector locates neutral mutant copies; the left eigenvector gives class reproductive values.
- The selection gradient weights class-specific effects by both quantities.
- The Hessian also includes selection-induced changes in the lineage class distribution.

Extra : Between-generation structure – when environmental fluctuations affect the whole population

Temporal fluctuations are common

In nature, conditions often vary from one generation to the next.

- Resource availability may follow seasonal cycles
- Climate variables such as temperature or rainfall fluctuate year to year
- Epidemics can wax and wane
- Host or habitat availability may change unpredictably

These fluctuations affect the entire population. A rare mutant will also experience these changes and, to determine whether it can invade, we need to ask whether it is on average doing better than the resident. But what kind of average should we take?

Invasion fitness in temporally fluctuating populations

Suppose the environment switches between states $c \in \{1, \ldots, M\}$ (e.g. hot and cold, wet and dry). The probability of experiencing state c is $q_c(x)$, which may depend on the resident trait x (e.g. if the population modifies its habitat). Let $w_c(y, x)$ be the expected number of offspring of a mutant with trait y in environment c, when the resident is fixed for x.

Invasion fitness is given by the geometric mean of mutant fitness across environments

$$\rho(y, x) = \prod_{c=1}^{M} w_c(y, x)^{q_c(x)}. \tag{19}$$

Directional selection: favouring traits exposed most often

The selection gradient in a fluctuating environment is

$$S(x) = \sum_{c=1}^{M} \frac{\partial}{\partial y} \log w_c(y, x)\Big|_{y=x} q_c(x). \tag{20}$$

This expression shows that:

- the contribution of each environment comes from how the trait affects fitness there,
- and each contribution is weighted by how often that environment occurs.

As a result, selection tends to favour trait values that are best suited to the most common environments.

Disruptive selection: not so easy due to bet-hedging

The Hessian is

$$H(x^*) = \sum_{c=1}^{M} \frac{\partial^2}{\partial y^2} \log w_c(y, x^*)\Big|_{y=x^*} q_c(x^*). \tag{21}$$

As before, the contribution of each environment is weighted by how often that environment occurs. Note that it involves the derivative of *log* fitness making it more difficult for $H(x^*) > 0$ and thus for polymorphism to emerge where different morphs are better adapted to different environments.

One way to interpret this is in terms of **bet-hedging**: because of the geometric mean form of invasion fitness, strategies that avoid very low fitness (or worse, zero!) in bad years are favoured. This tends to favour generalists who are never too bad in any particular environment.

An example of adaptation to temporal heterogeneity

Consider a simple case where the environment alternates between two states each year (e.g. wet vs dry, hot vs cold).

- The species is *semelparous*: individuals reproduce once and then die.
- Fecundity in environment $i \in \{1, 2\}$:

$$\exp\left(-\frac{(z-\theta_i)^2}{2\sigma_i^2}\right) \tag{22}$$

- Thus fecundity is maximised by trait value θ_1 in environment 1 and θ_2 in environment 2, with the strength of selection set by the width parameters σ_1 and σ_2.
- The environment switches such that it is in state 1 with probability q_1 and in state 2 with probability $q_2 = 1 - q_1$.

Extra : Selection in age-structured populations

Age: a major axis of variation

Age is a fundamental sources of heterogeneity, especially in endotherms. It is associated with differences in physiology, behaviour, and morphology.

To model evolution in age-structured populations (in discrete time), we can use the same matrix formalism introduced in the previous section, with classes now corresponding to age classes.

Life history can be represented as a sequence of transitions between ages via survival, combined with reproduction at each stage.

Leslie matrix

In age-structured populations, all offspring are born into the same age class, and individuals move linearly through ages. As a result, the mean matrix $\mathbf{W}(y,x)$ takes the special form of a **Leslie matrix**:

$$\mathbf{W}(y,x) = \begin{pmatrix} b_1(y,x) & b_2(y,x) & \dots & b_{M-1}(y,x) & b_M(y,x) \\ s_1(y,x) & 0 & \dots & 0 & 0 \\ 0 & s_2(y,x) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & s_{M-1}(y,x) & 0 \end{pmatrix} \tag{23}$$

where

- $b_j(y,x)$: expected number of newborn offspring produced by a mutant of age j (all offspring enter age class 1).
- $s_j(y,x)$: probability that a mutant of age j survives to age $j+1$.

Decomposition of survival and reproduction

We can decompose the Leslie matrix as

$$\mathbf{W}(y,x) = \mathbf{S}(y,x) + \mathbf{R}(y,x) \tag{24}$$

where $\mathbf{S}(y,x)$ contains only survival terms and $\mathbf{R}(y,x)$ only reproduction:

$$\mathbf{S}(y,x) = \begin{pmatrix} 0 & 0 & \dots & 0 \\ s_1(y,x) & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & s_{M-1}(y,x) & 0 \end{pmatrix}, \quad \mathbf{R}(y,x) = \begin{pmatrix} b_1(y,x) & \dots & b_M(y,x) \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}.$$

Next generation theorem

For $\mathbf{W}(y,x) = \mathbf{S}(y,x) + \mathbf{R}(y,x)$ (with $\mathbf{S}(y,x) \geq 0$, $\mathbf{R}(y,x) \geq 0$, and $[\mathbf{I} - \mathbf{S}(y,x)]$ invertible), the leading eigenvalue of $\mathbf{W}(y,x)$ is less or equal to 1 if and only if the leading eigenvalue of

$$[\mathbf{I} - \mathbf{S}(y,x)]^{-1}\,\mathbf{R}(y,x)$$

is less or equal to 1, where $\mathbf{I}$ is the identity matrix.

Next generation theorem : Biological interpretation

The matrix

$$[\mathbf{I} - \mathbf{S}(y,x)]^{-1} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ s_1(y,x) & 1 & 0 & \dots & 0 \\ s_1(y,x)s_2(y,x) & s_2(y,x) & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \prod_{k=1}^{M-1} s_k(y,x) & \prod_{k=2}^{M-1} s_k(y,x) & \prod_{k=3}^{M-1} s_k(y,x) & \dots & 1 \end{pmatrix}$$

gives the expected time spent in each age (via survival): its (i,j) entry gives the expected number of time steps that an individual *starting life at age j* spends in age i. Meanwhile, $\mathbf{R}(y,x)$ gives the expected number of offspring produced in each age per time step.

R_0 : the lifetime reproductive number

Multiplying $(\mathbf{I} - \mathbf{S}(y,x))^{-1}$ and $\mathbf{R}(y,x)$ shows that the leading eigenvalue is

$$R_0(y,x) = \sum_{j=1}^{M} l_j(y,x)\, b_j(y,x) \tag{25}$$

where $l_j(y,x) = \prod_{k=1}^{j-1} s_k(y,x)$ is the probability a mutant survives at least to age j.

$R_0(y,x)$ is the **lifetime reproductive number**: the expected number of offspring produced over a lifetime. Therefore for having a chance to invade – $R_0(y,x) > 1$ – a mutant carrier must on average produce more that on offspring in its lifetime.

Directional selection on age-specific effects

The selection gradient can then be written as

$$\tilde{S}(x) = \left.\frac{\partial R_0(y,x)}{\partial y}\right|_{y=x} = \sum_{a=1}^{M}\left(\left.\frac{\partial b_a(y,x)}{\partial y}\right|_{y=x} + \tilde{v}^{\circ}_{a+1}\left.\frac{\partial s_a(y,x)}{\partial y}\right|_{y=x}\right) l^{\circ}_a\,, \tag{26}$$

where

$$\tilde{v}^{\circ}_j = \sum_{k=j}^{M}\frac{l^{\circ}_k\, b^{\circ}_k}{l^{\circ}_j}\,, \tag{27}$$

i.e. the **expected number of offspring an individual will produce for the rest of its lifetime, given it has already survived to age** j. This is sometimes called the *current reproductive value*. This is like the reproductive value we saw earlier, but normalised so that $\tilde{v}^{\circ}_1 = 1$, so a newborn (age class 1) is the unit of measure, following Fisher's seminal work.

Disruptive selection in age-structured populations

Once the population expresses a singular trait value x^*, the two terms of disruptive selection

$$\tilde{H}(x^*) \;=\; \left.\frac{\partial^2 R_0(y,x)}{\partial y^2}\right|_{y=x^*} \;=\; \tilde{H}_{\mathrm{w}}(x^*) \;+\; 2\,\tilde{H}_{\mathrm{q}}(x^*)$$

are given by

$$\begin{aligned}
\tilde{H}_{\mathrm{w}}(x^*) &= \sum_{a=1}^{M} \left[\left.\frac{\partial^2 b_a(y,x^*)}{\partial y^2}\right|_{y=x^*} \;+\; \tilde{v}^{\circ}_{a+1} \left.\frac{\partial^2 s_a(y,x^*)}{\partial y^2}\right|_{y=x^*} \right] l^{\circ}_a \\
\tilde{H}_{\mathrm{q}}(x^*) &= \sum_{a=1}^{M} \left[\left.\frac{\partial b_a(y,x^*)}{\partial y}\right|_{y=x^*} \;+\; \tilde{v}^{\circ}_{a+1} \left.\frac{\partial s_a(y,x^*)}{\partial y}\right|_{y=x^*} \right] \left.\frac{\partial l_a(y,x^*)}{\partial y}\right|_{y=x^*},
\end{aligned} \tag{28}$$

where $\tilde{H}_{\mathrm{w}}(x^*)$ is selection from *non-linear effects* of traits on fecundity or survival and $\tilde{H}_{\mathrm{q}}(x^*)$ is selection from *context-by-fitness synergy*.

Selection strength is inversely proportional to generation time

Careful: $R_0(y, x)$ is sign-equivalent around 1 to invasion fitness $\rho(y, x)$ (the leading eigenvalue of $\mathbf{W}(y, x)$) but it's not equal to it! Therefore $R_0(y, x)$ is good at giving the *nature* of selection but not its overall *strength*. In fact,

$$S(x) = \frac{1}{T^\circ}\tilde{S}(x) \qquad \text{and} \qquad H(x^*) = \frac{1}{T^\circ}\tilde{H}(x^*)$$

where T° is the **generation time** in a resident population, i.e. the expected age of a parent. Selection is thus weaker in species with long generation times. This is because such species spread reproduction across many ages so the effect of a change at one age makes a smaller difference to reproduction over a whole generation. Remember: $R_0(y, x)$ is convenient for analysing the sign/structure of selection, but be careful if you wanna make statements about the absolute strength of selection (e.g. versus mutation or drift).

References

- Caswell, H. (1978). A general formula for the sensitivity of population growth rate to changes in life history parameters. *Theoretical Population Biology* 14, 215–230.
- Geritz, S. A. H., Kisdi, E., Meszéna, G. and Metz, J. A. J. (1998). Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology* 12, 35–57.
- Lehmann, L., Mullon, C., Akçay, E. and Van Cleve, J. (2016). Invasion fitness, inclusive fitness, and reproductive numbers in heterogeneous populations. *Evolution* 70, 1689–1702.

References

- Meszéna, G., Czibula, I. and Geritz, S. A. H. (1997). Adaptive dynamics in a 2-patch environment: a toy model for allopatric and parapatric speciation. *Journal of Biological Systems* 5, 265–284.
- Rueffler, C., Van Dooren, T. J. M., Leimar, O. and Abrams, P. A. (2006). Disruptive selection and then what? *Trends in Ecology and Evolution* 21, 238–245.
- Schmid, M., Rueffler, C., Lehmann, L. and Mullon, C. (2024). Resource variation within and between patches: where exploitation competition, local adaptation, and kin selection meet. *The American Naturalist* 203, E19–E34.