**Question 1** (20%) Here we investigate the demography of a mustard plant, *Boechera stricta*, following Cotto *et al.* (2019).

B. *stricta* has two key stages: immature (non-reproductive) and mature (reproductive). A fraction $s_I$ of immature individuals survive each year, of which a proportion $m$ mature. Each mature individual produces $f$ seeds that survive to become immature individuals in the following year, and a proportion $s_M$ of mature individuals survive and remain mature.

(a) (4%) Write down recursion equations for the number of immature, $I$, and matrure, $M$, individuals. Write out the corresponding transition matrix, $\mathbf{M}$.

(b) (4%) Write down the characteristic polynomial for $\mathbf{M}$.

Cotto *et al.* (2019) used published data to estimate parameters values for B. *stricta* and found that the two eigenvalues are approximately $\lambda = 3$ and $\lambda = 1/2$. With their estimated parameter values the right eigenvector associated with $\lambda = 3$ is approximately $\begin{pmatrix} 1/10 \\ 9/10 \end{pmatrix}$ and the right eigenvector associated with $\lambda = 1/2$ is approximately $\begin{pmatrix} 11/10 \\ -1/10 \end{pmatrix}$. The left eigenvector associated with $\lambda = 3$ is approximately $\begin{pmatrix} 1 & 15 \end{pmatrix}$ and the left eigenvector associated with $\lambda = 1/2$ is approximately $\begin{pmatrix} 1 & -1/10 \end{pmatrix}$. Use these estimates in your answers below.

(c) (4%) What is the leading eigenvalue? Is the population expected to grow or decline?

(d) (2%) What fraction of the population do we expect to be reproductively mature in the long-run?

(e) (2%) What is the reproductive value of a mature individual relative to an immature individual?

(f) (4%) The general solution for this system can be written $\vec{n}(t) = \mathbf{A}\mathbf{D}^t\mathbf{A}^{-1}\vec{n}(0)$. Write out the entries of $\mathbf{D}$ and $\mathbf{A}$.

**Question 2** (30%) Here we analyze a model to understand how species compete for essential nutrients (e.g., plants need certain inorganic nutrients to grow). Our approach follows Loreau (2011).

Let $R$ represent the available stock of an essential nutrient, which flows into the system at rate $I$ and is lost from the system at rate $qR$. Let $N_1$ be the biomass of a plant species that requires this nutrient. We assume the plants uptake the nutrient at a rate of $a_1 R N_1$, of which $e_1$ is converted into new plant biomass. Plant biomass is lost at a rate of $m_1 N_1$. The rate of change in the nutrient stock and plant biomass is then

$$\frac{\mathrm{d}R}{\mathrm{d}t} = I - qR - a_1 R N_1$$

$$\frac{\mathrm{d}N_1}{\mathrm{d}t} = e_1 a_1 R N_1 - m_1 N_1.$$

(a) (10%) Find all equilibria of this model.

(b) (10%) Use the Jacobian to determine when the equilibrium with the plant present, $\hat{N}_1 > 0$, is stable (assuming all parameters are positive). [Hint: look at the sign of the trace and determinant.] We'll call the value of $R$ at this equilibrium the "R star" of species 1, $R_1^*$.

(c) (2%) Now consider a second plant species whose dynamics follow

$$\frac{\mathrm{d}N_2}{\mathrm{d}t} = e_2 a_2 R N_2 - m_2 N_2.$$

If it were the only plant species present, with $N_2 > 0$, use this equation to get the equilibrium value of the resource. We call this the "R star" of the second species, $R_2^*$.

(d) (4%) Now to understand how competition for essential nutrients works, replace $R$ with $R_1^*$ in the equation for $\frac{\mathrm{d}N_2}{\mathrm{d}t}$. Dividing by $N_2$ gives the growth rate of species 2 when it attempts to establish in a habitat where only species 1 was present beforehand. Show that species 2 can only establish when $R_2^* < R_1^*$.

(e) (4%) What does it mean, biologically, that a species can only replace another when it has a lower $R^*$?

**Question 3** (40%) Here we analyze a "Levene-type" model to understand how environmental heterogeneity can produce genetic polymorphism. This treatment follows Doebeli (2011).

Imagine a species occupying two environments, which we'll call patch 1 and patch 2. Each generation, an individual with trait value $x$ survives to reproductive age with probability $w_1(x) = e^{-(x+d)^2/(2\sigma^2)}$ if it is in patch 1 or with probability $w_2(x) = e^{-(x-d)^2/(2\sigma^2)}$ if it is in patch 2. That is, trait value $x = -d$ maximizes survival in patch 1 and trait value $x = d$ maximizes survival in patch 2. The parameter $d$ determines how different the two patches are while $\sigma$ (assumed to be positive) describes how quickly survival drops off as trait values deviate from those maximizing survival (larger $\sigma$ cause slower drop offs). Survivors produce offspring, asexually, of which $cN$ from patch 1 and $(1-c)N$ from patch 2 are chosen to start the next generation (e.g., $c$, which is between 0 and 1, could be the quality of patch 1 relative to patch 2). We assume the total population size across the two patches, $N$, is constant. The offspring then randomly disperse into the two habitat patches to start the next generation.

Now imagine a population where individuals have either trait $x$ or trait $y$. The frequency of $y$ in the next generation is then determined by the sum of the classic haploid selection recursion in each patch, weighted by the contribution of each patch to the next generation,

$$q(t+1) = c\frac{q(t)w_1(y)}{(1-q(t))w_1(x) + q(t)w_1(y)} + (1-c)\frac{q(t)w_2(y)}{(1-q(t))w_2(x) + q(t)w_2(y)}.$$

Here we'll do an evolutionary invasion analysis, assuming $x$ is the resident trait value and $y$ is the trait value of a rare mutant. To find the invasion fitness of a rare mutant we take the derivative of $q(t+1)$ with respect to $q(t)$ and evaluate at $q(t) = 0$, giving

$$\lambda(y,x) = c\frac{w_1(y)}{w_1(x)} + (1-c)\frac{w_2(y)}{w_2(x)}.$$

(a) (10%) Calculate the selection gradient, $D(x) = \frac{\partial \lambda(y,x)}{\partial y}\big|_{y=x}$, which describes the direction of evolution from $x$. [Hint: using the chain rule, the derivative of $f(x) = e^{-(x-a)^2/(2\sigma^2)}$ is $(-(x-a)/\sigma^2)f(x)$.]

(b) (4%) At a singular strategy, $\hat{x}$, there is no directional selection. Show that $\hat{x} = d(1-2c)$ is a singular strategy.

(c) (4%) Calculate $\frac{dD(x)}{dx}\big|_{x=\hat{x}}$, describing how the selection gradient changes with the resident strategy near the singular point.

(d) (2%) When is $\hat{x}$ convergence stable? (Hint: use your answer in (c).)

(e) (10%) Calculate $\frac{\partial^2 \lambda(y,x)}{\partial y^2}\big|_{y=\hat{x},x=\hat{x}}$, describing the curvature of the fitness function when the resident trait value is at the singular strategy.

(f) (4%) Assuming the two patches contribute equally, $c = 0.5$, show that $\hat{x}$ is evolutionarily *unstable* when $\sigma < d$. (Hint: use your answer in (e).)

(g) (4%) Give a biological interpretation for why we see evolutionary instability at $\hat{x}$ when $d$ is large relative to $\sigma$ (assuming $c = 0.5$). Note that $\hat{x} = d(1-2c)$ represents a strategy that is a comprimise between the two patches.

(h) (2%) Given the parameter values are such that $\hat{x}$ is convergent stable but not evolutionarily stable, how does environmental heterogeneity affect genetic polymorphism?

**Question 4** (10%) Here we use probability theory derive a classic result in the coalescent model, a model which describes how alleles sampled from present-day individuals have lineages that "coalesce" into common ancestors as we look into the past.

(a) (2.5%) In a diploid population of constant size $N$, the probability the lineages of two sampled alleles coalesce in the previous generation is the probability they have the same parent allele. When we choose parent alleles at random this is $1/(2N)$. If the two sample lineages do not coalesce in that generation the probability of them coalescing in the generation before that is again $1/(2N)$, and so on. What is the name (or equation) of the probability distribution that most accurately describes the number of generations until the two lineages coalesce, $\Pr(T = t)$?

(b) (2.5%) Given it takes $T = t$ generations until the two lineages coalesce, there will be a branch of length $t$ leading from each sampled allele to their most recent common ancestor. That means the two sampled alleles are separated by $2t$ generations. Given mutations – all of which are unique – occur *continuously* at a rate of $\mu$ per generation, what is the name (or equation) of the probability distribution that most accurately describes the number of mutations that occur in $2t$ generations, $P(M = m|2t)$?

(c) (2.5%) Use the law of total expectation to write the expected number of mutations, $\mathbb{E}(M)$, as a sum of conditional expectations, where the conditioning is the time until coalescence, $\mathbb{E}(M|T = t)$.

(d) (2.5%) We find that the answer to (c) can be written $\mathbb{E}(M) = 2\mu\mathbb{E}(T)$. Based on your answer in (a), what is $\mathbb{E}(T)$? You now have the expected number of mutational differences between two randomly sampled alleles!