OPTIMIZATION OF MEAN FITNESS OF A POPULATION VIA ARTIFICIAL SELECTION

I. López[1], M. Gámez[2], R. Carreño[3] and Z. Varga[4]

[1, 2, 3] Dept. of Statistics and Applied Mathematics
University of Almería, Spain
[4]Institute of Mathematics and Informatics
Szent István University, Gödöllő, Hungary
[1]e-mail: milopez@ual.es
[2]e-mail: mgamez@ual.es
[3]e-mail: ramon@ual.es
[4]e-mail: zvarga@mszi.gau.hu

Abstract: The proposed control-theoretical model of artificial selection is based on the classical Fisher model of natural selection. Artificial selection is realized by changing the fitness of certain genotypes in a Mendelian population. Time-dependent fitness parameters are considered as control functions (artificial selection strategies). Under certain conditions on the model parameters the mean fitness of the population attains a maximum at the equilibrium of the selection dynamics. Thus the problem of optimization of the mean fitness via artificial selection is an optimal control problem. Using a sufficient condition for local controllability of nonlinear systems with invariant manifold, the existence of optimal artificial selection strategy is obtained.

Keywords: Bio control, optimality, controllability

1. INTRODUCTION

A synthesis of Darwinian theory of natural selection and Mendelian genetics is due to Fisher (1930, 1958), one of the "founding fathers" of population genetics, who formulated the Fundamental Theorem of Natural Selection: "The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time". This statement has been the subject of controversy involving many population geneticists, making clear its correct conditions, see, e. g., (Ewens, 1992), (Lessard, 1997). It has also been a starting point of very fruitful investigations concerning selection processes, e. g. (Akin, 1979), (Hofbauer and Sigmund, 1988 and 1998)

In mathematical terms, an important implication of Fisher's Fundamental Theorem is that during the infinite selection process the mean fitness of the population increases. Later in the paper we shall recall conditions, under which mean fitness attains a maximum at the equilibrium of the selection dynamics. In this case, the increase of the mean fitness of the population results in approaching its maximum by asymptotic stability of the equilibrium.

In our model of artificial selection we wish to control the population into the state of maximum mean fitness, controlling the population by changing the fitness parameters of certain genotypes. If the mean fitness of the population attains its maximum at equilibrium, it is enough to control the population state into this equilibrium.

In (Varga, 1989) a sufficient condition was given under which a system with invariant manifold can be controlled to equilibrium in a given time. In (Scarelli and Varga, 2002) this controllability result was applied to the analysis of selection-mutation processes. In the present paper this condition for controllability will provide the existence of optimal control for the artificial selection model.

2. A CONTROL-THEORETICAL MODEL OF ARTIFICIAL SELECTION

We shall consider a diploid Mendelian population with random mating and with alleles $A_1, \ldots, A_n$ at an autosomal locus. Assume that the diploid zygote individuals undergo a selection described in terms of a symmetric fitness matrix $W \in \mathbf{R}^{n\times n}$ with non-negative components where, for each $i, j \in \overline{1,n}$, w_{ij} is the Malthusian fitness value of an A_iA_j zygote (defined as the difference of the birth rate and the death rate of A_iA_j individuals). Then according to Fisher's classical model of *natural selection*, for the time-dependent frequency x_i of allele A_i we have

$$\dot{x}_i = x_i\left[(Wx)_i - \langle x, Wx\rangle\right] \quad (i \in \overline{1,n}) \qquad (1)$$

where

$$w_i(x) := (Wx)_i = \sum_{j=1}^{n} w_{ij}x_j$$

is the *potential (marginal) fitness* of allele A_i and

$$\overline{w}(x) := \langle x, Wx\rangle = \sum_{k,l=1}^{n} w_{kl}x_k x_l$$

is the *mean fitness* of the population in state x. The biological interpretation of Fisher's model is the following: if the potential fitness of A_i is greater than the average fitness of the whole population, then the frequency of A_i will increase, in the contrary case it will decrease.

It is easy to see that the standard simplex $\Delta_n \subset \mathbf{R}^n$ and its interior $\overset{\circ}{\Delta}_n$ are positively invariant for dynamics (1). A state $x^* \in \overset{\circ}{\Delta}_n$ is called a *polymorphic equilibrium* of the population if in this state all alleles have the same potential fitness, or equivalently, $w_i(x^*) = \overline{w}(x^*)$ $(i \in \overline{1,n})$. Assume that the following *regularity condition* holds: W is invertible, with $\mathbf{1} := (1,\ldots,1) \in \mathbf{R}^n$ $\langle W^{-1}\mathbf{1}, \mathbf{1}\rangle \neq 0$ holds and for

$$x^* := \frac{W^{-1}\mathbf{1}}{\langle W^{-1}\mathbf{1}, \mathbf{1}\rangle} \qquad (2)$$

we have $x^* > 0$. Then x^* is the unique polymorphic equilibrium, see Varga (1989).

Now, in order to model the *artificial selection* process, let us suppose that certain fitness parameters may change in function of time, as a result of human intervention: $w_{ij} + u_{ij}(t)$. For a technically convenient, structured description of the effect of artificial selection on the fitness parameters, for each $i, j \in \overline{1,n}$ define an $n \times n$ matrix R_{ij} with all items equal to zero except that with indices *ij,* which is 1, if genotype A_iA_j is artificially selected and zero otherwise. In terms of the mapping

$$\Psi : \mathbf{R}^{n\times n} \to \mathbf{R}^{n\times n}, \qquad \Psi(u) := \sum_{i,j=1}^{n} u_{ij}R_{ij},$$

the effect of artificial selection on the fitness matrix can be described in the form $W + \Psi(u)$. (The *ij* item of matrix R_{ij} will be denoted by r_{ij}).

Now, with the identification $\mathbf{R}^{n\times n} = \mathbf{R}^{n^2}$, considering u as control variable, our control system takes the form

$$\begin{aligned}\dot{x}_i &= x_i\left[(Wx)_i - \langle x, Wx\rangle\right] + \\ &\quad x_i\left[(\Psi(u)x)_i - \langle x, \Psi(u)x\rangle\right] \quad (i \in \overline{1,n})\end{aligned} \qquad (3)$$

3. OPTIMAL CONTROL VIA ARTIFICIAL SELECTION

According to Fisher's Fundamental Theorem quoted above, the mean fitness of a population is an important indicator of the evolution of a population. A natural population has the tendency to be in a state of maximum mean fitness. If by any undesired intervention this state suffers a change, a perturbation, an important task of *conservation biology* may be to control the population into this state of maximum fitness in a given time, applying a "soft" artificial selection. The corresponding optimal control problem can be set up as follows:

Fix a time interval $[0,T]$, and for given $\varepsilon > 0$ consider the set $U_\varepsilon[0,T]$ of $\mathbf{R}^{n\times n}$-valued C^∞ control functions with point-wise matrix norm less than ε. Given an initial state $x(0) = z \in \overset{\circ}{\Delta}_n$, define the terminal functional $J(u) := \overline{w}(x_u(T))$ $(u \in U_\varepsilon)$, where x_u is the solution of system (3) corresponding to this initial value. Formally,

$$J(u) \to \max \quad (u \in U_\varepsilon) \qquad (4)$$

$$x(0) = z,$$

$$\dot{x}_i = x_i\left[(Wx)_i - \langle x, Wx\rangle\right] + x_i\left[(\Psi(u)x)_i - \langle x, \Psi(u)x\rangle\right] \quad (i \in \overline{1,n}) \qquad (5)$$

Consider the following matrices defined in terms of the model parameters:

$$A := \left[x_i^*(w_{ij} - 2\overline{w}(x^*))\right]_{n\times n},$$

$$B := \left[x_i^*(\delta_{i\lambda(j)} - x^*_{\lambda(j)})r_{\lambda(j)\mu(j)}x^*_{\mu(j)}\right]_{n\times n}$$

with

$$\lambda(j) := ent\,\frac{j-1}{n} + 1, \; \mu(j) := (j-1)_{\operatorname{mod} n} + 1$$

and

$$D := \left[B \,|\, AB \,|\ldots|\, A^{n-1}B\right].$$

Theorem. Suppose that the above regularity condition holds, the matrix

$$P := \left[w_{kl} - w_{nl} - w_{kn} + w_{nn}\right]_{(n-1)\times(n-1)}$$

is negative definite and

$$\text{rank } D = n-1. \qquad (6)$$

Then for ε small enough and $\left|z - x^*\right| < \varepsilon$, the optimal control problem (4) – (5) has a solution.

Proof. Consider first the finite dimensional maximization problem

$$\overline{w}(x) \to \max \qquad (x \in \overset{\circ}{\Delta}_n)\,.$$

Under the conditions that W is invertible and $\left\langle W^{-1}\mathbf{1}, \mathbf{1}\right\rangle \neq 0$, the Lagrange multiplier rule provides that the unique possible solution to this maximization problem is x^*. An easy calculation shows that the negative definiteness of matrix P implies that x^* is a strict maximum point of the mean fitness in $\overset{\circ}{\Delta}_n$. Therefore, it is enough to show that there exist an $\varepsilon > 0$ such that from any initial state $z \in \overset{\circ}{\Delta}_n$ with $\left|z - x^*\right| < \varepsilon$, system (5) can be controlled to this equilibrium within $\overset{\circ}{\Delta}_n$ by a control $u \in U_\varepsilon[0, T]$. In other words, we have to show that the control system (5) is locally controllable to x^* within $\overset{\circ}{\Delta}_n$. For ε small enough, $\overset{\circ}{\Delta}_n$ is an $(n-1)$-dimensional regular submanifold positively invariant with respect to controls $u \in U_\varepsilon$, see (Varga 1989). In this paper a sufficient condition was obtained for local controllability (into equilibrium) of such systems, as a generalization of a classical result in (Lee and Markus, 1967), to control systems with invariant manifold. This sufficient condition is that the Kalman controllability matrix D corresponding to the system linearized around the equilibrium has rank equal to the dimension of the invariant manifold. In our case for the linearization, identifying $\mathbf{R}^{n\times n} = \mathbf{R}^{n^2}$, let

$$v := (u_{11}, u_{12}, \ldots, , u_{1n}, \ldots, u_{n1}, u_{12}, \ldots, u_{nn})'$$

and define $F_i(x, v)$ as the right-hand side of equation (5). Denoting by F the corresponding vector-valued function of two vector variables, we easily calculate

$$\frac{\partial F(x^*, 0)}{\partial x} = A \quad \text{and} \quad \frac{\partial F(x^*, 0)}{\partial v} = B\,.$$

Thus, if the rank condition (6) holds, our optimal control problem has a solution.

Remark 1. In (Varga, 1989) the quoted sufficient condition is announced for controls of L^∞ type. It is easy to see that a similar proof works for C^∞ type, too.

Remark 2. Under the conditions of the above theorem, $\overline{w}(x^*) - \overline{w}(x) \quad (x \in \overset{\circ}{\Delta}_n)$ is a positive definite Lyapunov function with a negative definite derivative with respect to system (1), providing global asymptotic stability of the polymorphic equilibrium. Thus, if the population starts from any polymorphic state, then by natural selection (by the zero-control dynamics) the population arrives into a neighborhood of the maximum fitness state x^* in finite time, from which it can be controlled into x^* by artificial selection in given time.

4. EXAMPLES

In the following illustrative examples we consider a population with three alleles A_1, A_2, A_3 at an autosomal locus. For an operational model of artificial selection it is reasonable to suppose that we can not distinguish individuals of the same phenotype. For a concrete illustration suppose that homozygotes have different phenotypes, A_1 is dominant over A_2 and A_3, while A_2 is dominant over A_3. We shall suppose that the above regularity condition holds and the matrix

$$P := \left[w_{kl} - w_{3l} - w_{k3} + w_{33}\right]_{2\times 2}$$

is negative definite. For a numerical example we can take

$$W := \begin{bmatrix} 1 & 3 & 4 \\ 3 & 1 & 3 \\ 4 & 3 & 1 \end{bmatrix}, \qquad (7)$$

providing $x^* = \begin{bmatrix} 2/5 & 1/5 & 2/5 \end{bmatrix}'$ and a negative definite $P = \begin{bmatrix} -6 & -3 \\ -3 & -4 \end{bmatrix}$.

Example 1. Let us consider an artificial selection situation in which the fitness of homozygotes A_3A_3 is controlled. Now the only nonzero matrix R_{ij} is R_{33} and we have

$$R_{33} := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \Psi(u_{33}) := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & u_{33} \end{bmatrix}.$$

In order to check rank condition (6), it is enough to see that this rank is at least 2. Indeed, if the rank is 3 then by the above quoted controllability theorem of (Lee and Markus, 1967) our system is locally controllable, and also reachable in the sense that the end points of the solutions starting from the equilibrium x^* fill a whole neighborhood of this equilibrium, which contradicts to the positive invariance of the interior of the simplex.

Now for a general fitness matrix W a routine calculation shows that

$$B = \begin{bmatrix} -\left(x_3^*\right)^2 x_1^* & -\left(x_3^*\right)^2 x_2^* & \left(x_3^*\right)^2\left(1 - x_3^*\right) \end{bmatrix}'.$$

To satisfy condition rank D=2 it is enough to check, e. g., the determinant of the upper-left minor of D, which is

$$x_1^* x_2^* (x_3^*)^4 \left[w_{13} - w_{23} - \sum_{k=1}^{3} (w_{1k} - w_{2k}) x_k^* \right]. \quad (8)$$

Since x^* is a polymorphic equilibrium, we have

$$\sum_{k=1}^{3}(w_{1k}-w_{2k})x_k^* = w_1(x^*)-w_2(x^*)=0 \,.$$

Therefore, a sufficient condition for the existence of an optimal artificial selection strategy in biological terms is that genotypes A_{13} and A_{23} have different fitness, which is the case for matrix given in (7).

Example 2. Consider now a case when we carry out an artificial selection among individuals of different phenotypes. Namely, in accordance with our above assumption that we can not distinguish individuals of the same phenotype, we insert the same control u_1 into the fitness of genotypes A_{22}, A_{23}, A_{32} and another control u_2 into the fitness of genotype A_{33}. Now by a reasoning similar to the considerations of Example 1, we conclude that the determinant of the upper-left minor of *D* is

$$x_1^* x_2^* (x_3^*)^2 (x_2^* + x_3^*) > 0 \,.$$

Thus, apart from the general conditions formulated at the beginning of the present section, we don't need any particular condition on the fitness parameters to ensure the existence of an optimal control.

Remark 3. For an effective realization of the existing optimal selection strategy, the application of efficient optimality conditions of control theory will be needed. A difficulty in applying Pontryagin's Maximum Principle is that, as calculations show, the optimal control is necessarily singular.

5. CONCLUSION

The presented control-affine dynamic model of artificial selection provides a method to find conditions which guarantee that by artificial selection we can control the population to the maximum state of mean fitness from nearby states, in given time. Under the same conditions the mean fitness can also be maximized on a finite time interval, starting out from any state in which all alleles are present. The illustrative examples show that, if in the course of the selection we intervene in several phenotypes, we may have more chance to maximize the mean fitness of the population by artificial selection.

ACKNOWLEDGEMENTS

This work has been supported by the Hungarian National Research Fund (OTKA No. T037271) and project UAL-CAJAMAR 2001 CR-UAL-011.

REFERENCES

Akin, E. (1979). *The Geometry of Population Genetics.* Lect. Notes Biomath. **31**.Springer, Berlin.

Ewens, W. J. (1992). An optimizing principle of natural selection in evolutionary population genetics. *Theor. Pop. Biol*. **36**, 333-346

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection. Clarendon*, Oxford

Fisher, R. A. (1958). *The Genetical Theory of Natural Selection*. Dover Publications, New York

Hofbauer, J. and Sigmund, K. (1988). *The Theory of Evolution and Dynamical Systems*. Cambridge University Press, Cambridge.

Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics.* Cambridge University Press, Cambridge

Lee, E. B., and Markus, L. (1967). *Foundations of Optimal Control Theory*. Wiley, New York

Lessard, S. (1997) Fisher's fundamental theorem of natural selection revisited. *Theor. Pop. Biol*. **52**, 119-136.

Scarelli, A. and Varga, Z. (2002). Controllability of selection-mutation systems. *BioSystems*. **65,** No 2-3, 113-121.

Varga, Z. (1989). On controllability of Fisher's model of selection. In "Differential Equations" (Eds. C. M. Dafermos, G. Ladas, G. Papanicolau) Marcel Dekker, New York, 717-723.