Bulletin of Aichi Univ. of Education, 59 (Natural Sciences), pp. 7 -11, March, 2010

On a Jacobi Matrix Associated with a Simple Genetic Algorithm

Yukihiro HASHIMOTO

Department of Mathematics Education, Aichi University of Education, Kariya, 448-8542, Japan

1. Introduction

We consider a map associated with the simple genetic algorithm (SGA), in which the successive generation is reproduced through only the crossover (recombination) operations, without changing the length of genomes, and no mutation occurs. In this situation, a deterministic map describes the evolution of probability distributions of genomes, which induces a dynamics on a simplex.

In this note, we show that the Jacobian of the map associated with SGA is upper triangular matrix at each vertices of the simplex, in a suitable order of genomes. As a result, we show that if the fitness function has the maximum value at a unique genome, the vertex of the simplex corresponding to the genome is local attractive point (Theorem 6.3), as expected.

2. Notations and definitions

Let $\Sigma = \{0, 1, \ldots, l-1\}$ be a set of alphabets for $l \geq 2$. A *genome* $\sigma = (\sigma_1 \sigma_2 \cdots \sigma_N)$ with length $N \geq 2$ is an element of Σ^N. Particularly we put $\mathbf{0} = (00\cdots0)$ We fix the length N of genomes throughout the paper. The *fitness function* is a positive function $f : \Sigma^N \to (0, \infty)$. Each genome $\sigma \in \Sigma^N$ is taken as an l-adic representation of a natural number $v(\sigma) = \sum_{i=1}^{N} \sigma_i l^{i-1}$, which induces an *order* of genomes as $\sigma < \tau$ if and only if $v(\sigma) < v(\tau)$.

A *mask* M is a proper subset $\emptyset \neq M \subsetneqq \{1, \ldots, N\}$ of $\{1, \ldots, N\}$. $\mathcal{M}$ stands for the set of all masks: $\mathcal{M} = \{M \mid \emptyset \neq M \subsetneqq \{1, \ldots, N\}\}$. The *complement* $\overline{M}$ of a mask $M \in \mathcal{M}$ is given by $\overline{M} = \{1, \ldots, N\} \setminus M$. For genomes $\sigma, \tau \in \Sigma^N$, we define a set $M(\sigma, \tau) = \{i \in \{1, \ldots, N\} \mid \sigma_i = \tau_i\}$. Note that $M(\sigma, \tau)$ may be $\{1, \ldots, N\}$ or $\emptyset$. For a genome $\sigma \in \Sigma^N$ and a mask $M \in \mathcal{M}$, we define a set of genomes

$$[\sigma]_M = \{\tau \in \Sigma^N \mid \tau_i = \sigma_i \text{ for all } i \in M\}.$$

By definition, we see the followings.

Lemma 2.1. *The statements i)* $\tau \in [\sigma]_M$ *ii)* $\sigma \in [\tau]_M$ *iii)* $[\sigma]_M = [\tau]_M$ *iv)* $M \subseteq M(\sigma, \tau)$ *are equivalent to each other.*

When a function ϕ on Σ^N is given, we often use the abbreviation $\phi([\sigma]_M) = \sum_{\tau \in [\sigma]_M} \phi(\tau)$, for a genome σ and a mask M.

3. Selection and crossover

Our SGA evolves the population ratios of each genome deterministic way. Let $p^t(\sigma)$ be the population ratio of a genome σ at time t. Then $\mathbf{p}^t = (p^t(\sigma))_{\sigma \in \Sigma^N}$ is a probability vector of genomes. The t-th generation of genomes reproduces the next generation, where genomes with more larger fitness are likely to be selected, as 'parents' of the next. The probability $q^t(\sigma)$ to select a genome σ at time t is given by

$$q_\sigma^t = \frac{f(\sigma) p^t(\sigma)}{S^t}, \tag{3.1}$$

where $S^t = \sum_{\tau \in \Sigma^N} f(\tau) p^t(\tau)$ is the expectation value of f under the distribution $\mathbf{p}^t$. Note that $\mathbf{q}^t = (q^t(\sigma))_{\sigma \in \Sigma^N}$ is a probability vector, and we define the *selection map* by $Q(\mathbf{p}^t) = \mathbf{q}^t$.

A pair of 'parents' reproduces 'children' through the crossover operation. Let σ and τ be selected genomes, and M be a mask. The *crossover* $\sigma \oplus_M \tau \in \Sigma^N$ *of* σ *and* τ *through* M is defined by

$$\sigma \oplus_M \tau = \begin{cases} \sigma_i, & \text{if } i \in M, \\ \tau_i, & \text{otherwise.} \end{cases}$$

We see the following properties by definition.

Lemma 3.1. *For genomes* $\sigma, \tau \in \Sigma^N$ *and a mask* $M \in \mathcal{M}$*, we have*

(1) $\sigma \oplus_M \tau = \tau \oplus_{\overline{M}} \sigma$.
(2) $\forall \sigma' \in [\sigma]_M,\ \forall \tau' \in [\tau]_{\overline{M}},\quad \sigma' \oplus_M \tau' = \sigma \oplus_M \tau$.
(3) *Any genome* σ *has a decomposition* $\tau \oplus_M \tau' = \sigma$*, where* $\tau \in [\sigma]_M$ *and* $\tau' \in [\sigma]_{\overline{M}}$.

Then the population ratio $p^{t+1}(\sigma)$ of a genome σ at time $t+1$ is given by

$$p^{t+1}(\sigma) = \sum_{M \in \mathcal{M}} \sum_{\substack{\tau, \tau' \in \Sigma^N \\ \tau \oplus_M \tau' = \sigma}} q^t(\tau) q^t(\tau') Prob(M) = \sum_{M \in \mathcal{M}} q^t([\sigma]_M) q^t([\sigma]_{\overline{M}}) Prob(M), \tag{3.2}$$

where $Prob(\cdot)$ is a suitable probability distribution on $\mathcal{M}$. For simplicity, we adopt $Prob(M) = 1/\#\mathcal{M}$ throughout this paper. We define the *crossover map* by $P(\mathbf{q}^t) = \mathbf{p}^{t+1}$.

Lemma 3.2. *Let* $\mathbf{q} = (q(\sigma))_{\Sigma^N}$ *be a probability vector. Then* $\mathbf{p} = (p(\sigma))_{\Sigma^N} = P(\mathbf{q})$ *is also probability vector.*

Proof. It follows from lemma 2.1 that a mask $M \in \mathcal{M}$ induces an equivalence relation on the set Σ^N of genomes by $\sigma \sim_M \tau$ if and only if $[\sigma]_M = [\tau]_M$. Let $\mathcal{R}_M$ be a complete system of representatives under the relation $\sim_M$, giving a decomposition $\Sigma^N = \amalg_{\sigma \in \mathcal{R}_M} [\sigma]_M$. By definition, we also see $\Sigma^N = \amalg_{\tau \in [\sigma]_M} [\tau]_{\overline{M}}$ and hence

$$\sum_{\tau \in [\sigma]_M} q([\tau]_{\overline{M}}) = \sum_{\tau \in \Sigma^N} q(\tau) = 1.$$

Then

$$\begin{aligned} \sum_{\sigma \in \Sigma^N} p(\sigma) &= \sum_{\sigma \in \Sigma^N} \frac{1}{\#\mathcal{M}} \sum_{M \in \mathcal{M}} q([\sigma]_M) q([\sigma]_{\overline{M}}) = \frac{1}{\#\mathcal{M}} \sum_{M \in \mathcal{M}} \sum_{\sigma \in \mathcal{R}_M} \sum_{\tau \in [\sigma]_M} q([\tau]_M) q([\tau]_{\overline{M}}) \\ &= \frac{1}{\#\mathcal{M}} \sum_{M \in \mathcal{M}} \sum_{\sigma \in \mathcal{R}_M} q([\sigma]_M) \sum_{\tau \in [\sigma]_M} q([\tau]_{\overline{M}}) = \frac{1}{\#\mathcal{M}} \sum_{M \in \mathcal{M}} \sum_{\sigma \in \Sigma^N} q(\sigma) = 1. \end{aligned}$$

□

Summing up, the evolution of population ratios of genomes under our SGA is described by a map $P \circ Q : \Delta \to \Delta$ on a simplex $\Delta = \{\mathbf{p} \in \mathbf{R}^{l^N} \mid \sum_{i=1}^{l^N} p_i = 1,\ p_i \geq 0\}$.

4. Analysis of selection map Q

Let $\mathbf{e}_\sigma$ be a probability vector $(e_\sigma(\tau))_{\tau \in \Sigma^N} = (\delta_{\sigma\tau})_{\tau \in \Sigma^N}$ supported on a unique genome σ, corresponding to a vertex of the simplex Δ.

Lemma 4.1. *A probability vector* $\mathbf{p} \in \Delta$ *is a fixed point of the selection map* Q *if and only if the fitness function* f *is constant on the support of* $\mathbf{p}$*, that is,* $f(\sigma) = f(\tau)$ *for any genomes* σ *and* τ *with* $p(\sigma) \neq 0$ *and* $p(\tau) \neq 0$*. In particular,* $\mathbf{e}_\rho$ *is a fixed point of* Q *for any genome* ρ*.*

Proof. $Q(\mathbf{p}) = \mathbf{p}$ means $f(\sigma)p(\sigma) = Sp(\sigma)$ (where $S = \sum_\tau f(\tau)p(\tau)$), hence $f(\sigma) = S$ whenever $p(\sigma) \neq 0$. Conversely suppose f has constant value c on the support of $\mathbf{p} \in \Delta$. Then $S = \sum_\tau f(\tau)p(\tau) = c\sum_\tau p(\tau) = c$ and hence $Q(\mathbf{p}) = (f(\sigma)p(\sigma)/S)_{\Sigma^N} = (cp(\sigma)/c)_{\Sigma^N} = \mathbf{p}$. □

Proposition 4.2. *For probability vectors* $\mathbf{p} = (p_t)_{\Sigma^N}$ *and* $\mathbf{q} = (q_\tau)_{\Sigma^N} = Q(\mathbf{p})$,

$$\frac{\partial q_\tau}{\partial p_\sigma} = \frac{f(\sigma)}{S}(\delta_{\sigma\tau} - q_\tau),$$

where $S = \sum_\tau f(\tau)p_\tau$. *In particular, the evaluation at* $\mathbf{e}_\rho$ *equals*

$$\left.\frac{\partial q_\tau}{\partial p_\sigma}\right|_{\mathbf{e}_\rho} = \frac{f(\sigma)}{f(\rho)}(\delta_{\sigma\tau} - \delta_{\rho\tau}) \tag{4.1}$$

for any genome $\rho \in \Sigma^N$.

Proof. A direct calculus gives

$$\frac{\partial}{\partial p_\sigma}\left(\frac{f(\tau)p_\tau}{S}\right) = \frac{1}{S^2}\left\{\delta_{\sigma\tau}f(\sigma)S - f(\tau)p_\tau \cdot f(\sigma)\right\} = \frac{f(\sigma)}{S}\left(\delta_{\sigma\tau} - \frac{f(\tau)p_\tau}{S}\right) = \frac{\partial q_\tau}{\partial p_\sigma} = \frac{f(\sigma)}{S}(\delta_{\sigma\tau} - q_\tau).$$

Note that $S = f(\rho)$ and $q_\tau = \delta_{\rho\tau}$ at $\mathbf{p} = \mathbf{e}_\rho$. □

5. Analysis of crossover map P

Lemma 5.1. $\mathbf{e}_\rho$ *is a fixed point of the crossover map* P *for any genome* ρ.

Proof. As one see $[\tau]_M \cap [\tau]_{\overline{M}} = \{\tau\}$ by definition for any genome τ and mask M, the τ-component of $P(\mathbf{e}_\rho)$ equals

$$\frac{1}{\#\mathcal{M}}\sum_{M\in\mathcal{M}} e_\rho([\tau]_M)\cdot e_\rho([\tau]_{\overline{M}}) = \frac{1}{\#\mathcal{M}}\sum_{M\in\mathcal{M}} \delta_{\rho\tau} = \delta_{\rho\tau} = e_\rho(\tau).$$

□

Proposition 5.2. *For probability vectors* $\mathbf{q} = (q_\tau)_{\Sigma^N}$ *and* $\mathbf{p} = (p_\tau)_{\Sigma^N} = P(\mathbf{q})$,

$$\frac{\partial p_\tau}{\partial q_\sigma} = \frac{2}{\#\mathcal{M}}\sum_{\substack{M\in\mathcal{M}\\ M\subseteq M(\sigma,\tau)}} q([\tau]_{\overline{M}}).$$

In particular, the evaluation at $\mathbf{e}_\rho$ *equals*

$$\left.\frac{\partial p_\tau}{\partial q_\sigma}\right|_{\mathbf{e}_\rho} = \frac{2}{\#\mathcal{M}}\#\{M\in\mathcal{M} \mid \overline{M(\rho,\tau)} \subseteq M \subseteq M(\sigma,\tau)\} \tag{5.1}$$

for any genome $\rho \in \Sigma^N$.

Proof. It follows from lemma 2.1 that

$$\frac{\partial}{\partial q_\sigma}\left(q([\tau]_M)q([\tau]_{\overline{M}})\right) = \begin{cases} q([\tau]_{\overline{M}}), & \text{if and only if } M \subseteq M(\sigma,\tau),\\ q([\tau]_M), & \text{if and only if } \overline{M} \subseteq M(\sigma,\tau).\end{cases}$$

Thus we see

$$\frac{\partial p_\tau}{\partial q_\sigma} = \frac{1}{\#\mathcal{M}}\left(\sum_{\substack{M\in\mathcal{M}\\ M\subseteq M(\sigma,\tau)}} q([\tau]_{\overline{M}}) + \sum_{\substack{M\in\mathcal{M}\\ \overline{M}\subseteq M(\sigma,\tau)}} q([\tau]_M)\right) = \frac{2}{\#\mathcal{M}}\sum_{\substack{M\in\mathcal{M}\\ M\subseteq M(\sigma,\tau)}} q([\tau]_{\overline{M}})$$

by replacing the variable $\overline{M}$ to M in the second term of the middle equality. In the case of $\mathbf{q} = \mathbf{e}_\rho$, we see

$$e_\rho([\tau]_{\overline{M}}) = \begin{cases} 1, & \text{if } \rho \in [\tau]_{\overline{M}},\\ 0, & \text{otherwise.}\end{cases}$$

By lemma 2.1, $\rho \in [\tau]_{\overline{M}}$ means $\overline{M} \subseteq M(\rho,\tau)$, that is, $\overline{M(\rho,\tau)} \subseteq M$. Hence

$$\left.\frac{\partial p_\tau}{\partial q_\sigma}\right|_{\mathbf{e}_\rho} = \frac{2}{\#\mathcal{M}}\sum_{\substack{M\in\mathcal{M}\\ \overline{M(\rho,\tau)}\subseteq M\subseteq M(\sigma,\tau)}} 1 = \frac{2}{\#\mathcal{M}}\#\{M\in\mathcal{M} \mid \overline{M(\rho,\tau)} \subseteq M \subseteq M(\sigma,\tau)\}.$$

□

6. Jacobian of map $P \circ Q$

Recall that our SGA map $P \circ Q$ fixes $\mathbf{e}_\rho$ for any genome ρ by lemma 4.1 and 5.1.

Proposition 6.1. *For a probability vector $\mathbf{p} = (p_\sigma)_{\Sigma^N}$, we put $\mathbf{p}' = (p'_\tau)_{\Sigma^N} = P \circ Q(\mathbf{p})$. Then*

$$\frac{\partial p'_\tau}{\partial p_\sigma} = \frac{2}{\#\mathcal{M}} \frac{f(\sigma)}{S} \sum_{\pi \in \Sigma^N} \sum_{\substack{M \in \mathcal{M} \\ M \subseteq M(\pi,\tau)}} \left(\sum_{\tau' \in [\tau]_{\overline{M}}} \frac{f(\tau') p_{\tau'}}{S} \right) \left(\delta_{\sigma\pi} - \frac{f(\pi) p_\pi}{S} \right).$$

In particular, the evaluation at $\mathbf{e}_\rho$ equals

$$\left. \frac{\partial p'_\tau}{\partial p_\sigma} \right|_{\mathbf{e}_\rho} = \frac{2}{\#\mathcal{M}} \#\{M \in \mathcal{M} \mid \overline{M(\rho,\tau)} \subseteq M \subseteq M(\sigma,\tau)\}(1 - \delta_{\rho\sigma}) \frac{f(\sigma)}{f(\rho)} \tag{6.1}$$

for any genome $\rho \in \Sigma^N$.

Proof. A combination of proposition 4.2 and 5.2 shows the first assertion. It follows from (4.1) and (5.1) that

$$\left. \frac{\partial p'_\tau}{\partial p_\sigma} \right|_{\mathbf{e}_\rho} = \sum_{\pi \in \Sigma^N} \frac{2}{\#\mathcal{M}} \#\{M \in \mathcal{M} \mid \overline{M(\rho,\tau)} \subseteq M \subseteq M(\pi,\tau)\} \cdot \frac{f(\sigma)}{f(\rho)} (\delta_{\sigma\pi} - \delta_{\rho\pi}),$$

noticing that $\{M \in \mathcal{M} \mid \overline{M(\rho,\tau)} \subseteq M \subseteq M(\pi,\tau)\} = \emptyset$ whenever $\pi = \rho$,

$$= \frac{2}{\#\mathcal{M}} \frac{f(\sigma)}{f(\rho)} \sum_{\substack{\pi \in \Sigma^N \\ \pi \neq \rho}} \#\{M \in \mathcal{M} \mid \overline{M(\rho,\tau)} \subseteq M \subseteq M(\pi,\tau)\} \cdot \delta_{\sigma\pi}$$

$$= \frac{2}{\#\mathcal{M}} \frac{f(\sigma)}{f(\rho)} \#\{M \in \mathcal{M} \mid \overline{M(\rho,\tau)} \subseteq M \subseteq M(\sigma,\tau)\} \cdot (1 - \delta_{\rho\sigma}).$$

□

Thus we obtain the Jacobi matrix of the SGA map $P \circ Q$ explicitly. The following is a key lemma to observe the behavior of $P \circ Q$ at a vertex $\mathbf{e}_\rho$ of the simplex Δ.

Lemma 6.2. *For any genomes σ and τ with $\mathbf{0} < \sigma < \tau$, $\overline{M(\tau,\mathbf{0})} \not\subseteq M(\tau,\sigma)$ holds.*

Proof. Suppose that $\mathbf{0} < \sigma < \tau$ and $\overline{M(\tau,\mathbf{0})} \subseteq M(\tau,\sigma)$. Then for any $i \in \overline{M(\tau,\sigma)} \subseteq M(\tau,\mathbf{0})$, it holds $\tau_i = 0$ and hence $\sigma_i \neq \tau_i = 0$. Take $j = \max \overline{M(\tau,\sigma)}$, then $\sigma_j > \tau_j = 0$, which implies $v(\sigma) > v(\tau)$ while $\sigma < \tau$. Hence a contradiction. □

Theorem 6.3. *Take probability vectors $\mathbf{p} = (p_\sigma)_{\Sigma^N}$ and $\mathbf{p}' = (p'_\tau)_{\Sigma^N} = P \circ Q(\mathbf{p})$. Then*

1. $\left. \dfrac{\partial p'_\tau}{\partial p_\sigma} \right|_{\mathbf{e_0}} = 0$ *holds for any genomes σ and τ with $\mathbf{0} < \sigma < \tau$. Therefore the Jacobi matrix $D(P \circ Q)$ at $\mathbf{0}$ is upper triangular.*
2. *If a fitness function f takes a unique maximum value at $\mathbf{0}$, then $\mathbf{0}$ is a local attractive point.*

Proof. The assertion (1) is a direct result of (6.1) and lemma 6.2. In the case $\sigma = \tau$, as $M(\tau,\tau) = \mathcal{M}$, we have

$$\left. \frac{\partial p'_\tau}{\partial p_\tau} \right|_{\mathbf{e_0}} = \frac{2}{\#\mathcal{M}} \#\{M \in \mathcal{M} \mid \overline{M(\rho,\tau)} \subseteq M\}(1 - \delta_{\mathbf{0}\tau}) \frac{f(\tau)}{f(\mathbf{0})}.$$

The right hand side equals 0 when $\tau = \mathbf{0}$, while $\tau \neq \mathbf{0}$, we see $\overline{M(\tau,\mathbf{0})} \neq \emptyset$ and hence

$$\#\{M \in \mathcal{M} \mid \overline{M(\tau,\mathbf{0})} \subseteq M\} \leq \frac{\#\mathcal{M}}{2}.$$

As f takes a unique maximum, $f(\mathbf{0}) > f(\tau)$ holds for any genome τ. Then we have

$$0 \leq \left.\frac{\partial p'_\tau}{\partial p_\tau}\right|_{\mathbf{e_0}} \leq \frac{f(\tau)}{f(\mathbf{0})} < 1,$$

that is, all diagonal elements of upper triangular matrix $D(P \circ Q)$ at $\mathbf{e_0}$, which are to be eigen values of $D(P \circ Q)$, are non-negative and smaller than 1. Therefore $\mathbf{e_0}$ is locally attractive. □

Note that we lose no generality to set $\mathbf{0}$ a unique maximum of a fitness function f, whenever f has a maximum value on a unique genome.

EXAMPLE. In the case $\Sigma = \{0, 1\}$ and $N = 2$, the set of genomes is $\Sigma^2 = \{00, 01, 10, 11\}$ and the set of masks is $\mathcal{M} = \{\{0\}, \{1\}\}$. The order of genomes is defined by $00 < 01 < 10 < 11$. Given a fitness function f, the selection and crossover maps are described as

$$\mathbf{q} = Q(\mathbf{p}) = \frac{1}{S}\begin{pmatrix} f(00)p_{00} \\ f(01)p_{01} \\ f(10)p_{10} \\ f(11)p_{11} \end{pmatrix}, \qquad \mathbf{p}' = P(\mathbf{q}) = \begin{pmatrix} (q_{00}+q_{01})(q_{00}+q_{10}) \\ (q_{00}+q_{01})(q_{01}+q_{11}) \\ (q_{10}+q_{11})(q_{00}+q_{10}) \\ (q_{10}+q_{11})(q_{01}+q_{11}) \end{pmatrix},$$

where $S = f(00)p_{00} + f(01)p_{01} + f(10)p_{10} + f(11)p_{11}$. Then Jacobi matrices of P and Q at $\mathbf{e}_{00} = (1, 0, 0, 0)$ are given as

$$DQ|_{\mathbf{e}_{00}} = \frac{1}{f(00)}\begin{pmatrix} 0 & -f(01) & -f(10) & -f(11) \\ 0 & f(01) & 0 & 0 \\ 0 & 0 & f(10) & 0 \\ 0 & 0 & 0 & f(11) \end{pmatrix}, \qquad DP|_{\mathbf{e}_{00}} = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

hence

$$D(P \circ Q)|_{\mathbf{e}_{00}} = \frac{1}{f(00)}\begin{pmatrix} 0 & -f(01) & -f(10) & -2f(11) \\ 0 & f(01) & 0 & f(11) \\ 0 & 0 & f(10) & f(11) \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

One see that $\mathbf{e}_{00}$ is a local attractive point provided that f takes a unique maximum value at 00.

7. Concluding remarks

We see the SGA map fixes each vertex of the simplex Δ, however another fixed points might exist. Indeed one see that a probability vector $\mathbf{p} = (l^{-N})_{\Sigma^N}$ (l^N equals a cardinality of Σ^N) is a fixed point of P, thus a fixed point of $P \circ Q$ when the fitness function f is constant on Δ. One also see that $P \circ P = P$ when $\Sigma = \{0, 1\}$ and $N = 2$.

References

[1] B. Kitchenes, *Symbolic Dynamics,* Springer, 1998.
[2] M. Lothaire, *Algebraic combinatorics on words*, Encyclopedia of mathematics and its applications **90**, Cambridge University Press, 2002.

(Received September 14, 2009)