

On the Stationary Distribution of GAs with Fixed Crossover Probability

U. Chandimal de Silva
Department of Mathematics
Osaka University, Toyonaka
Osaka 560-0043, Japan
chandi@jwh.jp

Joe Suzuki
Department of Mathematics
Osaka University, Toyonaka
Osaka 560-0043, Japan
suzuki@math.sci.osaka-u.ac.jp

ABSTRACT

We analyse the convergence of a GA when the mutation probability is low and the selection pressure is high, for arbitrary crossover types and probabilities. We succeed in mathematically proving that the stationary distribution associated with the Markov chain concentrates on uniform populations of the best individuals, as would be expected.

Categories and Subject Descriptors

G.1.6 [Numerical Analysis]: Optimization—*simulated annealing, stochastic programming*; G.3 [Probability and Statistics]: *markov processes*; I.2.8 [Artificial Intelligence]: Problem Solving, Control Methods, and Search

General Terms

Algorithms, Performance, Reliability, Theory

Keywords

genetic algorithms, stationary distribution, crossover

1. INTRODUCTION

Genetic Algorithms have been used in various disciplines as an optimization tool since they were introduced by Holland in 1975. Notwithstanding their broad landscape of usage and practically impressive results, little has been done to prove their efficiency or boost their credibility in a rigorous mathematical framework until recently. It was mainly in the past decade that such mathematically interesting results were discovered: notably by De Jong (1985), Nix and Vose (1992), Rudolph (1994), Banzhaf and Reeves (1999) etc.

A GA can be modelled using a Markov chain, which would be ergodic as long as the mutation probability is positive, which in turn guarantees the convergence of the chain with

a unique stationary distribution. This is the model used in most mathematical analyses of GAs.

Davis and Principe (1991), one of the first in this endeavour, tried to extrapolate results from the already mathematically rich area of Simulated Annealing into GAs. Davis proved that the stationary distribution of his GA was asymptotically positive over the uniform population subset of the search space when mutation tends to zero.

Suzuki (1997) later extended this work to the limit distribution when the selective pressure is arbitrarily large. He proved that this limit concentrates around the uniform populations of the fittest individuals.

Cerf (1996) extensively studied the asymptotic dynamics of the stationary distribution of his mutation-selection algorithm in the context of generalized simulated annealing, using the tools developed by Freidlin and Wentzell (1984).

More recently, Schmitt (2001) was able to prove that his scaled GA, which included crossover, asymptotically converged to the optimal solution under certain conditions. He was able to prove that if the mutation rate converges to zero slow enough to ensure weak ergodicity of the Markov chain concerned, and if the selection pressure increases fast enough (satisfying a convergence condition), the GA asymptotically behaves like a certain take-the-best algorithm, which in turn converges to the optimal solution, if the population size is larger than the length of the genome, and if the selection pressure satisfies a certain convergence condition. The fitness function is assumed to be injective, and mutation and crossover commute.

Out of the above-mentioned papers, Davis's, Suzuki's and Schmitt's were the only papers to at least consider crossover, which seems to be quite an 'unpopular' parameter when it comes to mathematical analysis of GAs. Suzuki proved his result in the advent of crossover at an asymptotically vanishing level. We will prove that this same result (with certain limitations) is valid under an arbitrary crossover probability. We believe that our result, which sheds light on the convergence of the GA to an optimal solution with crossover probability positive, will help bring the mathematical analysis of GAs one little step closer to the Simple Genetic Algorithm, where crossover plays an integral role. Schmitt's result, although chronologically positioned before ours, is an excellent example illustrating the practical achievability of this stationary distribution, although under restricting conditions.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GECCO '05, June 25–29, 2005, Washington, DC, USA.
Copyright 2005 ACM 1-59593-010-8/05/0006 ...\$5.00.

2. THE MODEL AND PREVIOUS WORK

2.1 Notation

L : Length of the bit-string representing each individual
 $N = 2^L$: Number of possible individuals
 M : Number of individuals in each population
 S : State space of M -individual populations
 $J = \{0, 1, \dots, N-1\}$: Set of individuals
 i_j : Individual with the j 'th largest fitness value
 $z[i, s]$: Number of occurrences of $i \in J$ in $s \in S$
 $f(i)$: Fitness value of $i \in J$
 l : $f(i_1) = \dots = f(i_l) > f(i_{l+1}) \geq \dots \geq f(i_N)$
 μ : Mutation probability
 χ : Crossover probability
 $H(i, j)$: Hamming distance between $i, j \in J$
 $|i|$: Number of one's in $i \in J$
 $U \subset S$: set of uniform populations $u(i)$, consisting of $i \in J$
 $V = \{s \in S | 1 \leq z[i_j, s] \leq M-1, \exists j = 1, 2, \dots, l\}$
 $V_j = \{s \in S | 1 \leq z[i_j, s] \leq M-1, z[i_k, s] = 0, k \neq j \in \{1, \dots, l\}\}, j = 1, \dots, l$
 $S[u]$: Set of states $t \in S$ such that for some $j \in J$,
 1) $z[i, t] = M-1$;
 2) $z[j, t] = 1$;
 3) $H(i, j) = 1$;
 for $u = u(i) \in U$

2.2 The Markov Chain Model

The transition probability $Q(t|s)$ from $s \in S$ to $t \in S$ takes the form of a multinomial distribution,

$$Q(t|s) = \frac{M!}{\prod_{j \in J} z[j, t]!} \prod_{i \in J} P(i|s)^{z[i, t]} \quad (1)$$

where $P(i|s)$ is the probability of generating an individual $i \in J$ from a state $s \in S$. When proportional selection is assumed (with no mutation and crossover), we have,

$$P_1(i|s) = \frac{z[i, s]f(i)}{\sum_{j \in J} z[j, s]f(j)} \quad (2)$$

When crossover and mutation are positive, we have,

$$P_2(k|s) = \sum_{i, j \in J} P_1(i|s)P_1(j|s)r_{i, j}(k), k \in J, s \in S \quad (3)$$

where $r_{i, j}(k)$ is the probability that $k \in J$ is generated from $i, j \in J$ by crossover and mutation, and $r_{i, j}(k) = r_{i \oplus k, j \oplus k}(0)$ holds (Vose (1999)). Suzuki (1998) introduced a formula for $r_{i, j}(0)$ when uniform crossover was used.

2.3 Previous Results

Davis determined an expression for the stationary distribution $q^{(\mu)}$ over S for fixed μ , and its limit q as $\mu \rightarrow 0$. for $s \in S$, we have

$$q^{(\mu)}[s] = \frac{\det(Q[s] - I)}{\sum_{t \in S} \det(Q[t] - I)} \quad (4)$$

by solving the linear equation $q^{(\mu)}Q = q^{(\mu)}$ using Cramer's rule, where the (t, s) -element of Q is the transition probability $Q(s|t)$, and $Q[s]$ is obtained by replacing in Q the s 'th row by the zero vector $(0, \dots, 0)$.

The limit is given as

$$q[s] = \lim_{\mu \rightarrow 0} q^{(\mu)}[s] = \begin{cases} \frac{\lim_{\mu \rightarrow 0} \det(Q^*[s] - I)}{\sum_{t \in U} \lim_{\mu \rightarrow 0} \det(Q^*[t] - I)} & [s \in U] \\ 0 & [s \in S - U] \end{cases} \quad (5)$$

where the elements of $Q^*[u]$, $u \in U$, are defined by

$$Q^*[u](t|s) = \begin{cases} 0 & [s = u] \\ Q(t|s) + \sum_{v \in U - u} \delta(t \in S[v]) \frac{Q(v|s)}{L} & [s \neq u] \end{cases}$$

for $s, t \in S - U \cup \{u\}$. Note that $|S[u]| = L$ for $u \in U$, and that the matrix $Q^*[u]$ has the following property:

$$\begin{aligned} & \sum_{t \in S - U \cup \{u\}} Q^*[u](t|s) \\ &= \sum_{t \in S - U \cup \{u\}} Q(t|s) + \sum_{v \in U - \{u\}} \sum_{t \in S[v]} \frac{Q(v|s)}{L} \\ &= \sum_{t \in S - U \cup \{u\}} Q(t|s) + \sum_{v \in U - \{u\}} Q(v|s) \\ &= \sum_{t \in S} Q(t|s) \\ &= 1 \end{aligned} \quad (6)$$

for $s \in S - U$.

Suzuki proved the following convergence theorem.

Proposition1: For the 3 operator algorithms, the stationary distribution focuses on the best population as the mutation probability $\mu \rightarrow 0$, crossover probability $\chi \rightarrow 0$, and the fitness ratio

$$F = \max_{1 \leq j \leq N-1, f(i_j) \neq f(i_{j+1})} f(i_{j+1})/f(i_j) \rightarrow 0 \quad (7)$$

ie.

$$\sum_{s \in \{u(i_1), \dots, u(i_l)\}} \lim_{F \rightarrow 0} \lim_{\chi \rightarrow 0} \lim_{\mu \rightarrow 0} q^{(\mu)}[s] = 1 \quad (8)$$

Notice that Proposition 1 essentially assumes $\chi = 0$. Suzuki proves this theorem by showing that,

$\lim_{F \rightarrow 0} \lim_{\chi \rightarrow 0} \lim_{\mu \rightarrow 0} \det(Q^*[u(i_j)] - I) \neq 0, j = 1, 2, \dots, l$, and

$\lim_{F \rightarrow 0} \lim_{\chi \rightarrow 0} \lim_{\mu \rightarrow 0} \det(Q^*[u(i_j)] - I) = 0, j = l+1, l+2, \dots, N$.

Albuquerque and Mazza (2000) derived a similar result when $\chi = 0$. However, they consider a GA with a noise-function added to the fitness function so as to break any existing ties and get a unique fittest individual.

3. EXTENSION FOR POSITIVE CROSSOVER CASE

3.1 Unique Best Individual Case

Suzuki[13] proved the following lemma in the course of the proof of Proposition 1.

Lemma1:

$\forall s \in V + \cup_{j=1}^l \{u(i_j)\}$

$$\lim_{F \rightarrow 0} \lim_{\chi \rightarrow 0} \lim_{\mu \rightarrow 0} \sum_{t \in V + \cup_{j=1}^l \{u(i_j)\}} Q(t|s) = 1 \quad (9)$$

and for any pair of $t \in S - V - \cup_{j=1}^l \{u(i_j)\}$ and $s \in V + \cup_{j=1}^l \{u(i_j)\}$,

$$\lim_{F \rightarrow 0} \lim_{\chi \rightarrow 0} \lim_{\mu \rightarrow 0} Q(t|s) = 0 \quad (10)$$

proof

$$\begin{aligned}
& \lim_{F \rightarrow 0} \lim_{\chi \rightarrow 0} \lim_{\mu \rightarrow 0} \sum_{t \in V \cup \cup_{j=1}^l \{u(i_j)\}} Q(t|s) \\
&= \lim_{F \rightarrow 0} [1 - \{1 - \sum_{k=1}^l P_1(i_k|s)\}^M] \\
&= \lim_{F \rightarrow 0} [1 - \{ \frac{\sum_{k=l+1}^N z[i_k, s] f(i_k)}{\sum_{i \in J} z[i, s] f(i)} \}^M] = 1
\end{aligned}$$

Q.E.D.

With a fixed positive crossover probability, and $l = 1$, $\sum_{k=1}^l P_1(i_k|s)$ will be replaced by

$$\begin{aligned}
P_2(i_1|s) &= \sum_{i, j \in J} P_1(i|s) P_1(j|s) r_{i, j}(i_1) \\
&\rightarrow P_1(i_1|s) P_1(i_1|s) r_{i_1, i_1}(i_1) = 1
\end{aligned}$$

as $F \rightarrow 0$. The lemma is still true. Thus, Proposition 1 would still hold under positive crossover probability, provided there is a unique best individual.

3.2 Two or More Best Individuals

Lemma2:

$\forall s \in V_j \cup \{u(i_j)\}$

$$\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} Q(u(i_j)|s) = 1 \quad (11)$$

and for any pair of $t \in S - \{u(i_j)\}$ and $s \in V_j \cup \{u(i_j)\}$,

$$\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} Q(t|s) = 0 \quad (12)$$

The proof is trivial if it is noted that $\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} P_1(i_j|s) = 1$ for $s \in V_j \cup \{u(i_j)\}$.

Q.E.D.

Let us prove that $\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} \det(Q^*[u(i_1)] - I) \neq 0$.

Let W_p^1 and D_p^1 be submatrices of $Q^*[u(i_1)]$ which consist of rows and columns indexed by V_1 and $S - U - V_1$ respectively. Then,

$$\begin{aligned}
\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} \det(Q^*[u(i_1)] - I) &= -\det(W^1 - I_{|V_1|}) \\
&\quad \times \det(D^1 - I_{|S-U-V_1|}) \quad (13)
\end{aligned}$$

holds. Here, $W^1 := \lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} W_p^1$ and $D^1 := \lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} D_p^1$. For any $s \in V_1$, since $\{u(i_1)\} \cap S[u] = \emptyset$ for all $u \in U$,

$$\begin{aligned}
& \sum_{t \in V_1 \cup \{u(i_1)\}} Q^*[u(i_1)](t|s) \\
&= \sum_{t \in V_1 \cup \{u(i_1)\}} Q(t|s) \\
&+ \sum_{v \in U - \{u(i_1)\}} \sum_{t \in (V_1 \cup \{u(i_1)\}) \cap S[v]} \frac{Q(v|s)}{L} \\
&= \sum_{t \in V_1 \cup \{u(i_1)\}} Q(t|s) + \sum_{j=l+1}^N |V_1 \cap S[u(i_j)]| \frac{Q(u(i_j)|s)}{L} \\
&\rightarrow 1 \quad (14)
\end{aligned}$$

from Lemma2 and $u(i_j) \in S - V_1 - u(i_1), j = l+1, \dots, N$. Combining this with (6), it turns out that all (s, t) elements of $Q^*[u(i_1)], s \in V_1, t \in S - U - V_1$ diminishes as in Fig.1 below.

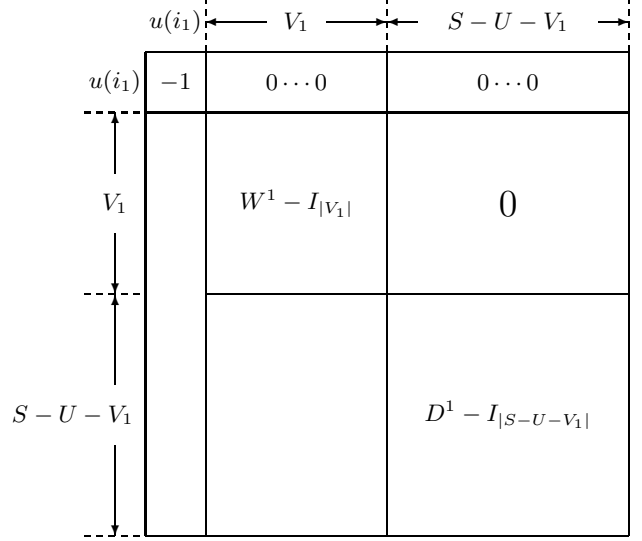


Figure 1: limit of the matrix $Q^*[u(i_1)] - I$

Now, we need to prove that the sum of some row in W^1 and in D^1 is strictly less than one, in order to show that there exists primitive stochastic matrices W^* and D^* such that $0 \leq W^1 \leq W^*, W^1 \neq W^*$ and $0 \leq D^1 \leq D^*, D^1 \neq D^*$ (see Appendix). This in turn would prove that all spectral radii of W^1 and D^1 are strictly less than one, meaning that the two determinants on the right hand side of eq.(13) are nonzero.

(1) For W^1 : for $s \in V_1$,

$$\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} Q^*[u(i_1)](u(i_1)|s) = \lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} Q(u(i_1)|s) = 1 \quad (15)$$

from Lemma1.

(2) For D^1 :

$$\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} Q^*[u(i_1)](u(i_1)|s) = \lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} Q(u(i_1)|s) > 0 \quad (16)$$

where we choose s from $V - \cup V_j$ such that i_1 occurs at least once. Note that here we use the fact that $l > 1$.

These complete the proof of $\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} \det(Q^*[u(i_1)] - I) \neq 0$. Similarly, $\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} \det(Q^*[u(i_j)] - I) \neq 0, j = 2, \dots, l$ can be proved.

Next, we shall prove that $\det(Q^*[u(i_{l+1})]) \rightarrow 0$ as $F \rightarrow 0$ after $\mu \rightarrow 0$.

Let W_p^{l+1} and D_p^{l+1} be submatrices of $Q^*[u(i_{l+1})]$ which consist of rows and columns indexed by V_1 and $S - U - V_1$ respectively. Let $W^{l+1} := \lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} W_p^{l+1}$ and $D^{l+1} := \lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} D_p^{l+1}$. Now, if we can prove

$$\begin{aligned}
\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} \det(Q^*[u(i_{l+1})] - I) &= -\det(W^{l+1} - I_{|V_1|}) \\
&\quad \det(D^{l+1} - I_{|S-U-V_1|}) \quad (17)
\end{aligned}$$

then it is sufficient to show $\det(W^{l+1} - I_{|V_1|}) = 0$. Thus, all we have to prove is

$$\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} \sum_{t \in V_1} Q^*[u(i_{l+1})](t|s) = 1 \quad (18)$$

for any $s \in V_1$. In fact, then, W^{l+1} is a stochastic matrix, and thus includes one as an eigenvalue (see Appendix). For any $s \in V_1$

$$\begin{aligned} \sum_{t \in V_1} Q^*[u(i_{l+1})](t|s) &= \sum_{t \in V_1} Q(t|s) \\ &+ \sum_{v \in U - \{u(i_{l+1})\}} \sum_{t \in V_1 \cap S[v]} \frac{Q(v|s)}{L} \\ &= \sum_{t \in V_1} Q(t|s) + Q(u(i_1)|s) + \sum_{j=l+2}^N [H(i_1, i_j) = 1] \frac{Q(u(i_j)|s)}{L} \\ &= \sum_{t \in V_1 \cup \{u(i_1)\}} Q(t|s) \rightarrow 1 \end{aligned} \quad (19)$$

from Lemma1.

This completes the proof of $\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} \det(Q^*[u(i_{l+1})] - I) = 0$.

Similarly, $\lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} \det(Q^*[u(i_j)] - I) = 0$, $j = l + 2, \dots, N$.

Thus, we have proven that Proposition1 holds for arbitrary values of crossover. Let us reformulate:

Theorem1

For the 3 operator algorithms, the stationary distribution focuses on the best uniform populations as the mutation probability $\mu \rightarrow 0$ and the fitness ratio

$$F = \max_{1 \leq j \leq N-1, f(i_j) \neq f(i_{j+1})} f(i_{j+1})/f(i_j) \rightarrow 0 \quad (20)$$

ie.

$$\sum_{s \in \{u(i_1), \dots, u(i_1)\}} \lim_{F \rightarrow 0} \lim_{\mu \rightarrow 0} q^{(\mu)}[s] = 1 \quad (21)$$

REMARK 1. *This result is not necessarily true when the limit is taken in the reverse order.*

Example ¹: $L = 2, M = 2$.

The fitness function f is defined as,

$$f(00) = 4 \times t^3, f(01) = 1, f(10) = 2 \times t, f(11) = 3 \times t^2 \text{ for } t = 1.$$

Thus, letting t go to infinity is equivalent to letting F go to zero. Direct computation shows that,

$$\lim_{\mu \rightarrow 0} \lim_{F \rightarrow 0} q^{(\mu)} = (2/3, 0, 0, 0, 0, 0, 0, 0, 1/3).$$

where the two nonzero positions correspond to $u(00)$ and $u(11)$ respectively.

This is also true regarding the main theorem in [13].

REMARK 2. *We cannot suggest how fast this convergence takes place, as our analysis does not touch on convergence rates.*

¹This example is due to Schmitt[12], section8.3

4. DISCUSSION

This paper tries to extend Suzuki's previous results on GAs to the nonzero crossover case. We have proven that the population asymptotically goes to the uniform population of an optimal solution (best individual) when the two parameters of selection and mutation are very small, for positive crossover probability. It should be noted that this result is valid whatever the crossover type is, and whether the crossover probability stays fixed or varies (decreases or even increases!).

It is always easier to analyse GAs when the parameters are asymptotically small, the reasons being obvious. Thus, many results regarding the limit of the stationary distribution at this asymptotic level are already available. The only problem is that these algorithms are more of Simulated Annealing than GAs, with their variable parameters none of which represent anything close to crossover. Although even mutation may be sometimes regarded as analogous to the 'temperature' in SA theory, SA does not have any parameter that may be compared with crossover, which may be one reason why most asymptotic treatments of GAs (using techniques borrowed from the mathematically rich area of SA theory) conveniently forget crossover altogether.

It is reassuring to know that this result is still valid when crossover exists. We hope that this would be an encouragement to researchers who work on GAs with nonzero crossover.

Our result is very mathematical and, by itself, has little practical significance since the parameters of our GA are practically unattainable. It is preferably read in an abstract perspective as an asymptotic result ratifying and unifying any future convergence results towards this 'perfect' stationary distribution.

5. ACKNOWLEDGMENTS

The authors would like to thank the three reviewers for their valuable comments without which this paper would not have looked the same.

REFERENCES

1. Albuquerque, P. and Mazza, C.,(2000) Foundations of Genetic Algorithms-6, San Francisco, CA. Morgan Kaufmann.
2. Banzhaf, W. and Reeves, C., editors,(1999) Foundations of Genetic algorithms-5, San Francisco, CA. Morgan Kaufmann.
3. Belew, R., editor,(1997) Foundations of Genetic Algorithms-4, San Francisco, CA. Morgan Kaufmann.
4. Cerf, R. (1996) The dynamics of mutation-selection algorithms with large population sizes. *Annales de l'Institut Henri Poincare*, 32(4):455-508
5. Davis, T. (1991) Toward an extrapolation of the simulated annealing convergence theory for the simple genetic algorithm. Ph.D. dissertation, University of Florida, Gainesville.
6. Davis, T. and Principe, J.C. (1991) A simulated annealing like convergence theory for the simple genetic algorithm. In Belew, R. and Bookers, L., editors,

Proc. of the Fourth International Conference on genetic Algorithms, pages 174-181, San Mateo, CA. Morgan Kaufmann.

7. Francois, O.(2002) Global Optimization with Exploration/Selection Algorithms and Simulated Annealing. The Annals of Applied Probability, 12(1):248-271
8. De Jong, K; A. (1975) An analysis of the behaviour of a class of genetic adoptive systems. Ph.D. dissertation, University of Michigan.
9. Freidlin, M. and Wentzell, A.(1984) Random perturbations of dynamical systems. Springer-Verlag, New-York.
10. Nix, A. and Vose, M. (1999) Modeling Genetic Algorithms with Markov chains. Ann. Math. Art. Intell, 5(1):79-88.
11. Rudolph, G. (1994) Convergence analysis of canonical genetic algorithms. IEEE Trans. on Neural Networks, special issue on Evolutionary Computation, 5(1):96-101
12. Schmitt, Lothar M. (2001) Theory of genetic algorithms. Theoretical Computer Science, 259(1-2):1-61
13. Suzuki, J. (1998) A further result on the Markov chain model of genetic algorithms and its application to a simulated annealing-like strategy. IEEE Transactions on Systems, Man and Cybernatics, 28(1)
14. Vose, M. D. (1999) The Simple Genetic Algorithm, Foundations and Theory. The MIT Press.
15. Whitley, D., editor, (1993) Foundations of Genetic Algorithms-2, San Mateo, CA. Morgan Kaufmann.
16. Whitley, D., editor, (1995) Foundations of Genetic Algorithms-3, San Francisco, CA. Morgan Kaufmann.

APPENDIX

A. THE PERRON-FROBENIUS THEOREM

Let A be a stochastic primitive matrix. Then,

1. $\alpha = 1$ is an eigenvalue of A ;
2. $\alpha = 1$ has corresponding left and right eigenvectors with strictly positive components;
3. $\alpha = 1 > |\alpha'|$ for any eigenvalue $\alpha' \neq \alpha$;
4. the eigenvectors associated with $\alpha = 1$ are unique to constant multiples;
5. If $0 \leq B \leq A$ and β is an eigenvalue of B , then $|\beta| \leq \alpha = 1$. Moreover, $|\beta| = \alpha = 1$ implies $B = A$.
6. $\alpha = 1$ is a simple root of the characteristic polynomial of A .