Matrix Interpretation of Generalized Embedded Landscape

Shude Zhou
State Key Laboratory of
Intelligent Technology and Systems
Department of Computer Science and
Technology, Tsinghua University
Beijing, 100084, China
zsd03@mails.tsinghua.edu.cn

Zengqi Sun
State Key Laboratory of
Intelligent Technology and Systems
Department of Computer Science and
Technology, Tsinghua University
Beijing, 100084, China
szq-dcs@tsinghua.edu.cn

ABSTRACT

This paper gives a matrix interpretation of generalized embedded landscape which is a class of additive decomposable functions mapping from *high-cardinality* alphabets domain to real numbers. Discrete Fourier transform is used to analyze the epistatic structure. We theoretically show the close relationship between Fourier coefficients and the underlying epistatic structure.

Categories and Subject Descriptors: I.2.8 [Artificial Intelligence]: Problem Solving, Control Methods, Search; F.2 [Analysis of Algorithms and Problem Complexity]: General

General Terms: Algorithms, Theory

Keywords: embedded landscape, epistisis, Fourier coefficients

1. INTRODUCTION

Epistasis is the nonlinear interaction between variable positions in the domain of a function with respect to computing the value of the function. Epistasis, in some form, is a necessary but not sufficient cause of problem difficulty for genetic algorithm. Walsh analysis provides a powerful way to quantify the epistasis of a function defined over a *binary* domain. A function in binary domain has k-bounded epistasis iff all of its Walsh coefficients of order greater than k are zero [1]. The paper extends some theoretical results of embedded landscapes to a domain of *high-cardinality* alphabets $\widetilde{M}^L = \{0, 1, \cdots, M-1\}^L$. By analyzing Fourier matrix, we show the relation between the epistases of GEL and its Fourier coefficients.

2. MATRIX INTERPRETATION OF GEL

The notation used is the same as that in [2]. A *generalized embedded landscape* (*GEL*) is a function $f : \widetilde{M}^L \mapsto \mathbb{R}$ which can be written as $f(x) = \sum_{i=1}^{P} g_i(pack(x, m_i))$ where g_i has a support mask m_i. A function $unpack$ is defined as $\widetilde{M}^A \times \widetilde{M}^B \times B^L \mapsto \widetilde{M}^L$, where $A + B = L$, $unpack(a, b, m)$ will generate a L-length string x that $pack(x, m) = b$ and $pack(x, zero(m)) = a$. Later, we will use $u(a, b)$ short for $unpack(a, b, m)$. Discrete Fourier basis function in domain $\widetilde{M}^L$ is $\psi_i(x)^{(M)} = e^{\frac{2\pi\sqrt{-1}}{M}(i\cdot x)}$ where $i, x \in \widetilde{M}^L$. Its complex conjugate is denoted by $\psi_i^{(M)*}(x)$. We have $\boldsymbol{f} = \boldsymbol{\Psi}\boldsymbol{\omega}$ and $\boldsymbol{\omega} = \frac{1}{M^L}\boldsymbol{\Psi}^*\boldsymbol{f}$, where $\boldsymbol{\Psi}$ is the Fourier matrix. When necessary we will use $\boldsymbol{\Psi}^{(L)}$ to indicate $\boldsymbol{\Psi}$ in L-dimension space. For $m \in \widetilde{M}^L$, $\|m\| = H$, $a, b \in \widetilde{M}^H$, the submatrix $\boldsymbol{V}_{(a,b)}$ of $\boldsymbol{\Psi}$ is a $M^{L-H} \times M^{L-H}$ matrix: $\boldsymbol{V}_{(a,b)} =$

$$\begin{pmatrix} \psi_{u(0,a)}(u(0,b)) & \cdots & \psi_{u(0,a)}(u(M^{L-H}-1,b)) \\ \vdots & \ddots & \vdots \\ \psi_{u(M^{L-H}-1,a)}(u(0,b)) & \cdots & \psi_{u(M^{L-H}-1,a)}(u(M^{L-H}-1,b)) \end{pmatrix} \quad (1)$$

GECCO'07, July 7–11, 2007, London, England, United Kingdom.
ACM 978-1-59593-697-4/07/0007.

THEOREM 1. *(Properties of Fourier Matrix) a)* $\boldsymbol{\Psi\Psi^*} = M^L\boldsymbol{I}$; *b)* $\boldsymbol{\Psi} = \boldsymbol{\Psi}'$; *c)* $\boldsymbol{\Psi}^{-1} = \frac{1}{M^L}\boldsymbol{\Psi}^*$; *d)* $\sum_{j=0}^{M^L-1}\boldsymbol{\Psi}_j = (M^L, \underbrace{0, 0, \ldots, 0}_{M^L-1})'$

THEOREM 2. $\boldsymbol{V}_{(a,b)}$ *is the submatrix of* $\boldsymbol{\Psi}^{(L)}$, $\boldsymbol{V}_{(a,b)} = \psi_a^{(H)}(b)\boldsymbol{\Psi}^{(L-H)}$

THEOREM 3. *(Sum of Submatrix) For simplicity,* $\boldsymbol{V}$ *stands for* $\boldsymbol{V}_{(a,b)}$*, and* $\boldsymbol{V}_i$ *is the* i^{th} *column vector of* $\boldsymbol{V}_{(a,b)}$*. We have:*

$$\sum_{i=0}^{M^{L-H}-1}\boldsymbol{V}_i = (M^{L-H}\psi_a^{(H)}(b), \underbrace{0, 0, \ldots, 0}_{M^{L-H}-1})' \quad (2)$$

Thm.1-3 are the properties of Fourier matrix $\boldsymbol{\Psi}$, based on which we can analyze the epistases of GEL. Usually, $\boldsymbol{f}$ is written in such order: $\boldsymbol{f} = (f(0), f(1), \ldots, f(M^L-1))'$. Here, we rewrite $\boldsymbol{f}$ as a column vector: $\boldsymbol{f} = (\boldsymbol{f}_0, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_{M^{L-H}-1})'$, where $\boldsymbol{f}_i = (f(u(i, 0)), f(u(i, 1)), \ldots, f(u(i, M^H-1)))'$, $i = 0, \ldots, M^{L-H}-1$. Accordingly, the elements in Fourier transform matrix are rearranged. In this way, Thm.4-5 can be obtained by analyzing linear projections and injections between function space and Fourier space. Proof omitted.

THEOREM 4. *Let* $f : \widetilde{M}^L \mapsto \mathbb{R}$, $g : \widetilde{M}^H \mapsto \mathbb{R}$ *and* $f(x) = g(pack(x, m))$*, where* $H < L$*, Fourier coefficients of* f *can be calculated:*

$$\omega_i^f = \begin{cases} \omega^g_{pack(i,m)} & \text{if } pack(i, zero(m)) = \boldsymbol{0} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where ω^f *and* ω^g *denote Fourier coefficients of* f *and* g *respectively.*

THEOREM 5. $f(x) = \sum_{i=1}^{P} g_i(pack(x, m_i))$ *is a GEL, then its Fourier coefficients can be calculated:*

$$\omega_j^f = \begin{cases} \alpha & \text{if } \exists i \in \{1, \cdots, P\}, pack(j, zero(m_i)) = \boldsymbol{0} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where $x, j \in \tilde{M}^L$, $m_i \in B^L$ *and* α *is* $\sum_{i\in\Lambda}\omega^{g_i}_{pack(j,m_i)}$, $\Lambda = \{i : 1 \le i \le P \,\&\, pack(j, zero(m_i)) = \boldsymbol{0}\}$.

Thm.4 and 5 show the close relation between the values of Fourier coefficients and the underlying epistatic structure of GEL by analyzing Fourier matrix. Embedding a lower dimensional function in a higher dimensional space neither increases the number of nonzero Fourier coefficients nor the maximum level of epistasis. The Embedding Theorem about binary embedded landscape [1] is extended to a more general domain over high-cardinality alphabets. A high-cardinality function is of order-k bounded epistasis iff its Fourier coefficients of order greater than k are zero.

Acknowledgment: The work is funded by the National Key Project for Basic Research of China (G2002cb312205).

3. REFERENCES

[1] R. B. Heckendorn. Embedded landscapes. *Evolutionary Computation*, 10(4):345–369, 2002.

[2] S. Zhou, Z. Sun and R. B. Heckendorn. Extended probe method for linkage discovery over high-cardinality alphabets. In *GECCO 2007 Proceedings*. ACM Press, 2007.