Elementary Landscape Analysis

SEP592 AI-Based Software Engineering

Shin Yoo

Contents

Given a finite set of solutions, X:

- **Neighbourhood**: let $N : X \rightarrow 2^X$ be the neighbourhood function.
- **Fitness**: let $f : X \rightarrow \mathbb{R}$ be the fitness function.

A *landscape* is a triple (X, N, f): the candidate solution set, the neighbourhood function that imposes connectivity, and the fitness function that assigns value to each point.

- In the lecture, we focus on *regular* and *symmetric* landscapes. For discussions of non-symmetric and non-regular neighbourhoods (which are more complicated), see Barnes et al. [**?**].
- The following proof is a summary of Chicano et al. [**?**].

Definition 1

Laplacian Δ of graph G is $A - D$, where A is the adjacency matrix of G and D is the diagonal degree matrix of G:

$$A_{xy} = \begin{cases} 1 & \text{if } y \in N(x) \\ 0 & \text{otherwise} \end{cases}$$

$$D_{xy} = \begin{cases} |N(x)| & \text{if } x == y \\ 0 & \text{otherwise} \end{cases}$$

(for regular and symmetric landscape, $D = dI$, where d is the degree of the neighbourhood graph)

Laplacian matrix on function f:

$$\Delta f = \begin{pmatrix} \sum_{y \in N(x_1)} (f(y) - f(x_1)) \\ \sum_{y \in N(x_2)} (f(y) - f(x_2)) \\ \sum_{y \in N(x_3)} (f(y) - f(x_3)) \dots \\ \sum_{y \in N(x_{|X|})} (f(y) - f(x_{|X|})) \end{pmatrix}$$

The component for solution x is:

$$(\Delta f)(x) = \sum_{y \in N(x)} (f(y) - f(x)) \quad (1)$$

Definition 2

Elementary Landscape: Let (X, N, f) be a landscape, and Δ be the Laplacian of N. Function f is *elementary* if there exists a constant b and eigenvalue λ of $-\Delta$ s.t. $(-\Delta)(f - b) = \lambda(f - b)$.

- The f is originally the fitness function, $f(x)$. However, we can capture the fitness function as a vector of length $|X|$: the ith member of vector f is $f(x_i), x_i \in X$.

What is the implication??

Proposition 1 (Elementary Properties)

1. *If f is a constant function, i.e. $f(x) = b, \forall x \in X$, then $(-\Delta)f = 0$ and f is eigenfunction of $-\Delta$ with eigenvalue $\lambda = 0$.*
2. *If f is elementary for N with eigenvalue λ, then there exists a constant b such that:*

$$\mathbb{E}_{y \in N(x)}(f(y)) = f(x) + \frac{\lambda}{d}(b - f(x)) \tag{2}$$

Elementary Properties.

For the first property, we use Eq **??** and write:

$$(-\Delta f)(x) = \sum_{y \in N(x)} (f(x) - f(y)) = \sum_{y \in N(x)} (b - b) = 0$$

$\therefore (-\Delta f)(x) = 0$ and f is eigenfunction of $-\Delta$ with eigenvalue $\lambda = 0$.

Elementary Properties (Cont.)

For the second property, start from Eq **??**:

$$(\Delta f)(x) = \sum_{y \in N(x)} (f(y) - f(x)) = \sum_{y \in N(x)} f(y) - df(x)$$

Divide with d:

$$\frac{1}{d}(\Delta f)(x) = \frac{1}{d} \sum_{y \in N(x)} f(y) - f(x) = \mathbb{E}_{y \in N(x)}(f(y)) - f(x)$$

Elementary Properties (Cont.)

Since f is elementary, there exists a constant b such that $-\Delta(f-b) = \lambda(f-b)$. Dividing this with d, we get:

$$\frac{1}{d}(\Delta(f-b))(x) = -\frac{\lambda}{d}(f(x)-b)$$

Using the first elementary property, we can lose b from the left hand side (since it is a constant function and therefore $(-\Delta)b = \lambda b$ with $\lambda = 0$):

$$\frac{1}{d}(\Delta(f-b))(x) = \frac{1}{d}(\Delta f)(x)$$

The right hand side is $\mathbb{E}_{y \in N(x)}(f(y)) - f(x)$ in the previous slide.

$$\mathbb{E}_{y \in N(x)}(f(y)) - f(x) = -\frac{\lambda}{d}(f(x)-b)$$

$$\therefore \mathbb{E}_{y \in N(x)}(f(y)) = f(x) + \frac{\lambda}{d}(b - f(x))$$

□

This is not so helpful **yet**, as we do not know b.

Lemma 1

Let N be a symmetric neighbourhood over X, and Δ its Laplacian. If f is an eigenvector of $-\Delta$ with $\lambda \neq 0$, then $\bar{f} = 0$.

Proof.

Two eigenvectors of a symmetric matrix with different eignevalues are orthogonal. Any constant function is an eigenvector of $-\Delta$ with $\lambda = 0$ (Prop. **??**). So, if $\lambda \neq 0$, f is orthogonal to any constant function, including $(1, 1, \ldots, 1)$. Consequently:

$$\bar{f} = \frac{1}{|X|} \sum_{x \in X} f(x) = \frac{1}{|X|}(1, 1, 1, \ldots, 1) \cdot f = 0$$

□

Theorem 1 (Grover's Wave Equation)

Let (X, N, f) be a landscape where N is regular and symmetric. Then, f is elementary if and only if there exists a constant λ such that the following expression holds:

$$\mathbb{E}_{y \in N(x)}(f(y)) = f(x) + \frac{\lambda}{d}(\bar{f} - f(x)), \forall x \in X \quad (3)$$

AND λ is the eigenvalue of f for its Laplacian.

Proof.

- $\rightarrow$: If N is regular and symmetric, and f is elementary, Equation **??** from Proposition **??** holds: we only need to show that the constant b is $\bar{f}$. Also, from Proposition **??**, $(-\Delta)(f - b) = \lambda(f - b)$. If we let $g = f - b$, g is an eigenvector of $-\Delta$ with eigenvalue λ. If $\lambda = 0$, Equation **??** trivially holds. If $\lambda \neq 0$, Lemma **??** states that $\bar{g} = 0$, therefore, $\bar{f} = \bar{g} + b = b$, so Equation **??** holds.
- $\leftarrow$: If Eq. **??** holds, multiplying it by d yields $\sum_{y \in N(x)} f(y) = df(x) + \lambda(\bar{f} - f(x))$. In vector form, $-\Delta f = \lambda(f - \bar{f})$. Since $-\Delta \bar{f} = 0$, we can write $-\Delta(f - \bar{f}) = \lambda(f - \bar{f})$. $\therefore$ f is elementary.

□

Elementary landscape: what is the implication?

- We can get the average fitness of neighbouring solutions of $x \in X$ (i.e. $\mathbb{E}_{y \in N(x)}(f(y))$) just by knowing the average fitness of all solutions (i.e. $\bar{f}$) and the fitness of the current solution (i.e. $f(x)$).
- Wait, usually there are too many solutions in X: how are we supposed to know $\bar{f}$ if we haven't evaluated all solutions in X?

Compositional View: an alternative approach to the elementary landscape. This is best given by Whitley et al. [**?**].

Intra/intercomponents: in most practical elementary landscapes, the fitness function for a candidate solution is a linear combination of a subset of a collection of *components*, C, such as edge weights in TSP. Then there exists $C_x \subset C$ s.t.:

$$f(x) = \sum_{c \in C_x} c$$

We refer to C_x as the *intracomponents* of solution x; $C - C_x$ as the *intercomponents* of x.

Building Neighbours with Components: when a local search algorithm *moves* from x to $y \in N(x)$, a subset of intracomponents $c_{out} \subset C_x$ is removed, and a subset of intercomponents, $c_{in} \subset C - C_x$ is added. That is:

$$f(y) = f(x) - \sum_{c \in c_{out}} c + \sum_{c \in c_{in}} c$$

If we fix x and let y be a uniformly random neighbour-move, we can compute the exepected value of $f(y)$:

$$\begin{aligned}\mathbb{E}[f(y)] &= \mathbb{E}[f(x) - \sum_{c \in c_{out}} c + \sum_{c \in c_{in}} c] \\ &= f(x) - \mathbb{E}[\sum_{c \in c_{out}} c] + \mathbb{E}[\sum_{c \in c_{in}} c]\end{aligned}$$

TSP under 2-OPT: Now, let's see if TSP under 2-OPT forms an elementary landscape. First, what is the average fitness of a TSP problem with n cities?

- There are $\frac{n(n-1)}{2}$ edges, out of which we choose n to form a tour. Consequently,

$$\bar{f} = \sum_{e_{i,j} \in E} w_{i,j} \frac{n}{n(n-1)/2} = \frac{2}{n-1} \sum_{e_{i,j} \in E} w_{i,j}$$

Intracomponents: under 2-OPT, two edges change in each neighbour. That is, two intracomponents go out for each neighbour. There are n edges in the current solution (i.e. a tour). Consequently, the probability of one intracomponent being removed is $\frac{2}{n}$, while the sum of intracomponents is simply $f(x)$ (again, the current tour). As a result, the expected value of intracomponents that go out is:

$$\mathbb{E}[\sum_{c \in c_{out}} c] = \frac{2}{n} f(x)$$

Intercomponents: The sum of intercomponents is simpy the sum of all edges minus the sum of all the intracomponent: $\sum_{e_{i,j} \in E} w_{i,j} - f(x)$. There are $\frac{n(n-1)}{2} - n = n(n-3)/2$ intercomponents (i.e. edges that are not part of the current solution), out of which we choose 2 to add. As a result, the expected value of the intercomponents that come in is:

$$\mathbb{E}[\sum_{c \in c_{in}} c] = \frac{2}{n(n-3)/2}(\sum_{e_{i,j} \in E} w_{i,j} - f(x))$$

- Note that $\bar{f} = \frac{2}{n-1} \sum_{e_{i,j} \in E} w_{i,j}$. That is, $\sum_{e_{i,j} \in E} w_{i,j} = \frac{n-1}{2} \bar{f}$

$$
\begin{aligned}
\mathbb{E}[f(y)] &= f(x) - \frac{2}{n}f(x) + \frac{2}{n(n-3)/2}\left(\sum_{e_{i,j}\in E} w_{i,j} - f(x)\right) \\
&= f(x) - \frac{2}{n}f(x) + \frac{2}{n(n-3)/2}\left(\frac{n-1}{2}\bar{f} = f(x)\right) \\
&= \dots \\
&= f(x) + \frac{n-1}{n(n-3)/2}(\bar{f} - f(x)) \\
&= f(x) + \frac{k}{d}(\bar{f} - f(x))
\end{aligned}
$$

It is elementary, Watson.

Observations on Elementary Landscape: Codenotti and Margara [**?**] state that:

- if $f(x) < \bar{f}$ then $f(x) < \mathbb{E}[f(y)] < \bar{f}$.
- if $f(x) = \mathbb{E}[f(y)]$ then $f(x) = \bar{f} = f(y)$.
- if $f(x) > \bar{f}$ then $f(x) > \mathbb{E}[f(y)] > \bar{f}$.

That is, all the local minima lie below $\bar{f}$ and all the local optima lie above $\bar{f}$. There is also an interesting observation about *plateau* in Whitley et al. [**?**] that we will not go into during the lecture.

Evaluation of Partial Neighbourhood: Suppose $M \subseteq N(x)$ is a subset of the neighbour of x. Expected fitness of r uniformly randomly drawn from the *remaining* neighbours in $N(X)$ is:

$$\mathbb{E}[f(r)] = \frac{1}{d - |M|}\left(\sum_{z \in N(x)} f(z) - \sum_{z \in M} f(z)\right)$$
$$= \frac{1}{d - |M|}\left(d\left(f(x) + \frac{k}{d}(\bar{f} - f(x))\right) - \sum_{z \in M} f(z)\right)$$

- We can even compute the expected improvement from M!

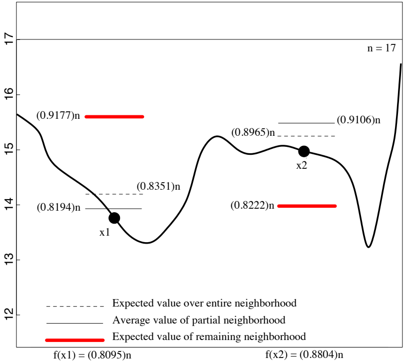

Figure 5: A stylized sketch of partial and whole neighborhood relationships for two solutions x_1 and x_2. The average evaluation of the partial neighborhood of x_1 is worse than x_2; the expected evaluation of remaining neighbors of x_1 are *better* than those of x_2. Numerical data were generated using two random solutions from (normalized) 17-city TSP `gr17.tsp` instance.