DIRICHLET LANDSCAPES

JONATHAN ATWELL

1. Defining the Dirichlet Dot Product Landscape

Let the fitness landscape be defined as follows. Let $\lambda \in \Lambda \subseteq \mathbb{R}^N$ be a location on the landscape and $X \sim Dir(\alpha)$ be an N-dimensional random vector on the sample space Ω. Consider a state $\omega \in \Omega$, we denote by $x_i = X_i(\omega)$ a realization of the i-th random variable in X and, accordingly, $x = (x_1, \dots, x_N)$ is a realization of X. In what follows we restrict our analysis to the case $\Lambda = \{0,1\}^N$, that is, where the set of possible locations on the landscape is discrete and corresponds to the set of the 2^N edges of an N-th dimensional hypercube.

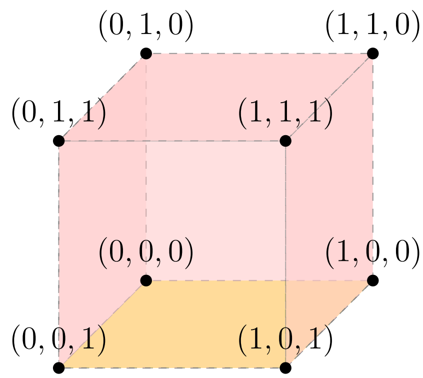

Figure 1.1. The set Λ of locations when $N = 3$

We define the fitness f of a given location as

$$\begin{aligned} f :& \Lambda \to \mathbb{R}^+ \\ f :& \lambda \mapsto \lambda \cdot x \end{aligned}$$

This definition of fitness, on the one hand, allows to conceive the performance of a given configuration as a convex combination of the values of its dimensions: this conveys the idea that not all dimensions contribute equally to fitness. On the other hand, an *indirect*[1] idea of tradeoff can also be inferred by this formulation: at any fixed number of non-zero dimensions, the fact that different dimensions contribute differently may direct search with the aim of remaining at least at the same value of fitness.

Date: September 30, 2020.

[1]A direct notion of tradeoff would be embedded by constraining the total number of resources to be allocated.

Remark 1. Note that this setup prescribes the use of a unique Dirichlet random vector to determine the value of the random weights at every location of the landscape. Another possibility is that of considering a different random vector be assigned to each different locations. This operationalization would translate the consideration that importance of the various dimensions can change as one moves from a location to another one.[2] *@@This setup was used in the original note – I think it's an interesting extension but I suggest to first clarify what happens when weights are "location-independent".@@* □

Marginal contributions (or the *added value*) of each dimension to fitness can be naturally defined in this setup by considering the contribution of the focal dimension to total fitness. Formally, let $f_i(\lambda_i)$ denote the section of the fitness function at the i-th dimension. The contribution of dimension i to fitness is defined as

$$\Delta^i = f_i(1) - f_i(0)$$

which isolates the variation in overall fitness as dimension i varies from 0 to 1.[3]

About interdependence and complementarities – Our definition of fitness is in accordance with the NK modeling literature in Management in that it maintains the assumption of fitness be **increasing in all of the components** of the N-dimensional vector of locations coordinates. However, this hypothesis makes it hard to explore negative complementarities without imposing a polynomial functional form for fitness (Rahmandad, 2019).

In fact, the "microeconomics-way" of defining complementarities (also exploited in Rahmandad, 2019) is through cross-derivatives of the fitness function (the sign of the cross derivatives dictates the direction of the complementarity between dimensions). This direction is not pursuable in this setup due to the definition of the fitness function (all cross-derivatives are null here).

In this setting, interdependence between dimensions is also indirect, as opposed to the NK-landscape framework, and derives from the nature of the Dirichlet weights.

A possible solution is to let not only the weights, but also the locations coordinates, be random. In particular we can model the coordinates locations as Bernoulli random variables. Interdependence among them can then me modeled through the use of a copula. This construction not only allows to tune interdependence between variables more directly than in the NK-landscape framework, but also allows to obtain an indirect study of complementarities between dimensions. In fact, the structure of correlation between dimensions embeds considerations of mutual change in the location coordinates. Due to the assumption of monotonicity of the fitness function, this

[2]Note that this reasoning can be akin to the idea of inconsistency of preferences in decision theory.

[3]Note that the definition extends also to the case of multi-valued coordinates and to the continuous case.

can be interpreted as two components being synergetic or antagonistic, in probabilistic terms.