A Math Problem Defined

Morgan ∼ Bryant. Sept. 2020.

Here I define a math problem which underlies effective ρ-style neural networks. This might even become a neat avenue of open-ended research.

An essential citation: Urban, S., & van der Smagt, P. (2015). A neural transfer function for a smooth and differentiable transition between additive and multiplicative interactions. arXiv preprint arXiv:1503.05724.

Other creditors: P Smolensky, J McClelland, G Hinton, A Lampinen

Problem definition

This problem is an underdetermined challenge to find an 'optimal' function.

Let $\exp(y) = e^y$, real, elementwise on vectors. For a functional f, let $f^{\{\alpha\}}$ denote function iteration (or tetration when applied to $\exp(\cdot)$).[1,2] Let x be an N-tensor, indexed by i to one lower rank. Let $\oplus^\star_\beta(x) \doteq \exp^{\{\beta\}}\left(\sum_i \exp^{\{-\beta\}} x_i\right)$[3]. Find a $\oplus$ that 'algebraically' differs from $\oplus^*$ as little as possible[4], subject to $\oplus = \oplus^{\{1\}} \approx \oplus^*_0$ and $\oplus^{\{t\}} \to \oplus^*_1$ as $t \to N$ (and $N \to \infty$).

---

[1] $f : \mathbb{R}^m \mapsto \mathbb{R}^n$. In particular, $\mathbb{C}$ complicates our iteration of $\exp(\cdot)$ in undesired ways. Also, complex-valued models of synapses are not popular for being biological realistic.

[2] Definition details:

$$f^{\{\alpha\}} \doteq \underbrace{f \circ f \circ \cdots \circ f}_{\alpha \text{ times}}$$

Let $f^{\{0\}} = \text{Id}$. Let $f^{\{-1\}} = f^{-1}$, the (left) inverse of f; i.e., $\exp^{\{-1\}}(y) = \log(y)$.

As in (Urban and Smagt, 2016), permit $f^{\{1/z\}} : \underbrace{f^{\{1/z\}} \circ \cdots \circ f^{\{1/z\}}}_{z \text{ times}} = f, \; z \in \mathbb{N}$ and thus $z \in \mathbb{Z}$. Care: generally, $e^{\alpha y} \neq y e^\alpha \neq \exp^{\{\alpha\}}(y)$.

[3] Observe that $\exp^{\{0\}} = \Sigma$ and $\exp^{\{1\}} = \Pi$.

[4] such as $\oplus = \exp^{\{\epsilon_1\}}\left(\Sigma \exp^{\{-\epsilon_2\}}[\cdot]\right)$