Characterization of Computation-Activation Networks by sufficient statistics

Research Notes in the ENEXA and QROM Projects

November 14, 2025

Sufficient statistics are treated in mathematical statistics and in information theory. We here choose a definition of information theory and apply a factorization theorem of mathematical statistics to relate with Computation-Activation Networks.

Definition 1 (Sufficient statistics). *Let $\{\mathbb{P}^\theta\left[X_{[d]}\right] : \theta \in \Theta\}$ be a family of probability distributions and*

$$\mathcal{S} : \bigtimes_{k\in[d]} [m_k] \to \bigtimes_{l\in[p]} [p_l]$$

be a function. We say that $\mathcal{S}$ is sufficient for θ, if for any distribution $\tilde{\mathbb{P}}[\theta]$ of θ, when drawing $X_{[d]}$ from $\mathbb{P}^\theta\left[X_{[d]}\right]$ with probability $\tilde{\mathbb{P}}[\theta]$, we have that

$$\left(\theta \perp X_{[d]}\right)\big| \, \mathcal{S}\left(X_{[d]}\right) \, .$$

Theorem 1 (Characterization of Computation-Activation Networks). *Let $\{\mathbb{P}^\theta\left[X_{[d]}\right] : \theta \in \Theta\}$ be a family of probability distributions with a sufficient statistic $\mathcal{S}$. Then there is a non-negative (possibly non-Boolean) base measure $\nu\left[X_{[d]}\right]$ and a map*

$$h : \Theta \to \bigotimes_{l\in[p]} \mathbb{R}^{p_l}$$

such that for all $\theta \in \Theta$

$$\mathbb{P}^\theta\left[X_{[d]}\right] = \left\langle h(\theta)[Y_{[p]}], \beta^{\mathcal{S}}\left[Y_{[p]}, X_{[d]}\right], \nu\left[X_{[d]}\right]\right\rangle \left[X_{[d]}|\varnothing\right] \, .$$

We further have that for a set $\{\mathbb{P}^\theta\left[X_{[d]}\right] : \theta \in \Theta\}$ $\mathcal{S}$ is a sufficient statistic, if and only if there is a non-negative (possibly non-Boolean) base measure $\nu\left[X_{[d]}\right]$ with

$$\{\mathbb{P}^\theta\left[X_{[d]}\right] : \theta \in \Theta\} \subset \Lambda^{\mathcal{S},\mathrm{MAX},\nu} \, .$$

Proof. By the factorization theorem of mathematical statistics (see [Hogg - The. 2.7.1]) we have that $\mathcal{S}$ is a sufficient statistic if and only if there are real-valued functions k_1 on $\left(\bigtimes_{l\in[p]}[p_l]\right) \times \Theta$ and k_2 on $\bigtimes_{k\in[d]}[m_k]$ such that

$$\mathbb{P}^\theta\left[X_{[d]} = x_{[d]}\right] = k_1(\mathcal{S}^{x_{[d]}}, \theta) \cdot k_2(x_{[d]}) \, . \tag{1}$$

We define a base measure by the coordinate encoding of k_2 by

$$\nu\left[X_{[d]}\right] = \sum_{x_{[d]}\in\bigtimes_{k\in[d]}[m_k]} k_2(x_{[d]})\epsilon_{x_{[d]}}\left[X_{[d]}\right]$$

and for each $\theta \in \Theta$ an activation tensor

$$\xi^\theta\left[Y_{[p]}\right] = \sum_{y_{[p]}} k_1(y_{[p]}, \theta)\epsilon_{y_{[p]}}\left[Y_{[p]}\right] \, .$$

With this we have for any $\theta \in \Theta$

$$\left\langle h(\theta)[Y_{[p]}], \beta^{\mathcal{S}}\left[Y_{[p]}, X_{[d]}\right], \nu\left[X_{[d]}\right]\right\rangle [\varnothing] = 1$$

and thus for any $x_{[d]} \in \bigtimes_{k\in[d]}[m_k]$ applying basis calculus

$$\begin{aligned}
\left\langle h(\theta)[Y_{[p]}], \beta^{\mathcal{S}}\left[Y_{[p]}, X_{[d]}\right], \nu\left[X_{[d]}\right]\right\rangle \left[X_{[d]} = x_{[d]} | \varnothing\right] &= h(\theta)[Y_{[p]} = \mathcal{S}^{x_{[d]}}] \cdot \nu\left[X_{[d]} = x_{[d]}\right] \\
&= k_1(\mathcal{S}^{x_{[d]}}, \theta) \cdot k_2(x_{[d]}) \\
&= \mathbb{P}^{\theta}\left[X_{[d]} = x_{[d]}\right] .
\end{aligned}$$

We therefore find for any $\mathbb{P}^{\theta}\left[X_{[d]}\right]$ a representation as a Computation-Activation Network in $\Lambda^{\mathcal{S},\mathrm{MAX},\nu}$ with the activation tensor $h(\theta)[Y_{[p]}]$.

To show the second claim, we are left to show that any set of Computation-Activation Networks in $\Lambda^{\mathcal{S},\mathrm{MAX},\nu}$ has $\mathcal{S}$ as a sufficient statistic. Let us thus consider a parametric family

$$\{\mathbb{P}^{\theta}\left[X_{[d]}\right] \,:\, \theta \in \Theta\} \subset \Lambda^{\mathcal{S},\mathrm{MAX},\nu} .$$

By this inclusion we find for any $\theta \in \Theta$ an activation core $\alpha^{\theta}[Y_{[p]}]$. We then construct functions k_1 and k_2 by

$$k_1(y_{[p]}, \theta) = \alpha^{\theta}[Y_{[p]} = y_{[p]}] \quad \text{and} \quad k_2(x_{[d]}) = \nu\left[X_{[d]} = x_{[d]}\right]$$

and notice that the equivalent condition (1) to $\mathcal{S}$ being a sufficient statistic is satisfied. $\square$