Généralisation des réseau de neurones par l'algèbre différentielle

Paul LANDRIER

1 Introduction

On cherche à abstraire le fonctionnement d'un réseau de neurones. On se fonde sur l'approche des graphes de calcul.

2 Solution simpliste

La solution la plus simple, n'utilisant pas l'algèbre différentielle, consiste à séparer les annotations en deux types, un pour les fonctions et un pour les dérivées partielles.

Formellement, on définit un monoïde de fonction $(F, \circ, id)$ et semi-anneau $(S, +, \cdot, 0, 1)$ représentant les dérivées partielles. On annote alors les arrêtes (u, v) du graphe avec $F \times S$, où le premier élément du couple représente la fonction permettant de passer de u à v et le second représente la valeur de la dérivée de cette fonction (on considère exclusivement des fonctions $\mathbb{R} \to \mathbb{R}$, lorsque l'on a besoin de fonctions de plusieurs variable, comme $+$, on précise l'annotation sur le nœud correspondant).

Prenons l'exemple du calcul de la fonction $f(x) = e^{x^2}(x^2 + 2(x + 1))$, représenté par le graphe :

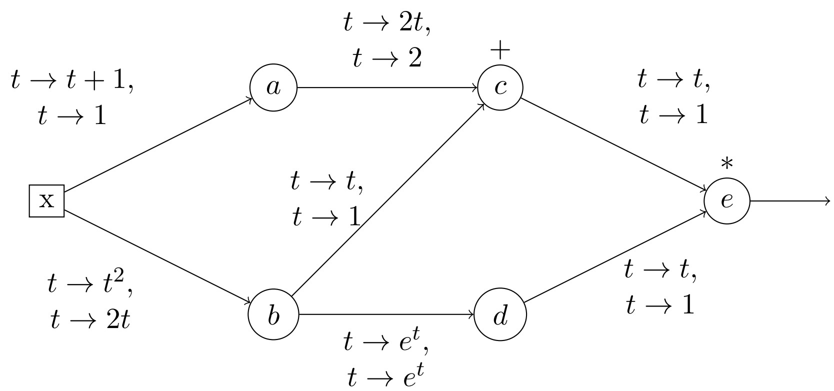

Figure 1: Graphe de calcul de la fonction f.

La figure 2 représente le calcul annoté de la fonction f.

3 Semi-anneau différentiel

On introduit ici une notion de semi-anneau différentiel :

Definition 1 *Un semi-anneau différentiel* $(F, +, \cdot, 0, 1, d)$ *est un semi-anneau commutatif* $(F, +, \cdot, 0, 1)$ *muni d'un application* $d : F \to F$ *vérifiant :*

- $\forall f, g \in F,\ d(f + g) = d(f) + d(g)$
- $\forall f, g \in F,\ d(fg) = d(f)g + fd(g)$

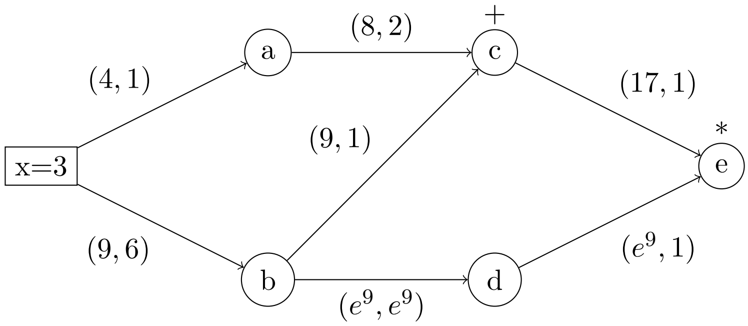

Figure 2: Graphe de calcul de la fonction f annoté sur l'exécution $x = 3$.

L'égalité $d(0 \cdot 0) = d(0) \cdot 0 + 0 \cdot d(0)$ implique $d(0) = 0$. En revanche, aucun phénomène analogue ne se produit pour $d(1)$ (c'est le cas dans un anneau différentiel en revanche, car on a l'égalité $d(1) = d(1) + d(1)$).

4 Executing a computation

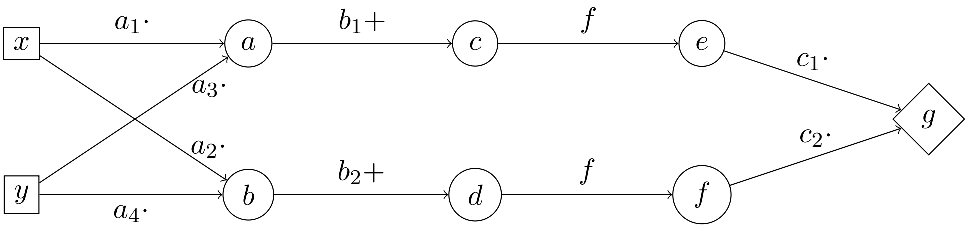

Figure 3: Executing a computation

4
We have a composition and a sum for aggregation. No order for aggregation : + assoc and commut. Composition assoc to treat some modules as black-boxes. Identity and 0 with *1 and *0 (zero-weight is important (no edge)) not sure for *1. $(f + g) \circ h = f \circ h + g \circ h$?

5 Training a model

Want to add a derivative d. Forces to add a $\cdot$. What identities ? -¿ $d(f \circ g)$?

6 Cas général

6.1 Définition théorique

Definition 2 *On définit une structure $(F, +, \cdot, \circ, 0, 1, id, d)$ telle que :*

- *$(F, +, \cdot, 0, 1, d)$ est un semi-anneau différentiel.*
- *$(F, \circ, id)$ est un monoïde.*
- *Pour tout $(f, g) \in F^2$, $d(f \circ g) = d(g) \cdot (d(f) \circ g)$.*
- *Equalities on $f + g \circ h$?*

On modélise un graphe de calcul par un graphe $G = (V, E)$ muni de deux fonctions $fun : E \to F$ et $op : V \to \{+, \cdot\}$.

La fonction fun sert à instancier le calcul.

La fonction *op* permet de détermier comment aggréger plusieurs fonctions en entrée d'un noeud. Par convention, pour les nœuds ayant un degré entrant de 1, le choix d'opérateur n'a pas d'importance et le nœud est l'identité.

Cette définition de la fonction *op* force l'associativité et la commutativité de $+$ et $\cdot$ (pas d'ordre sur les arrêtes).

6.2 Exemples

$\mathcal{C}^\infty(\mathbb{R})$,$\mathcal{C}^\infty(\mathbb{R}^+)$,$\mathbb{R}[X]$,$\mathbb{N}[X]$

6.3 Objet universel

6.4 Remarques

1. La factorisation à travers l'objet universel ne peut pas faire apparaitre plus d'informations que l'architecture.
2. Si *op* est constant égal à $+$, on a la provenance classique.
3. 12