Tropical Geometry of Deep Neural Networks

May 13, 2020

Contents

0 Introduction

1 Tropical Algebra

Our basic object of study is the following object $(\mathbb{R} \cup \{-\infty\}, \oplus, \odot)$. As a set this is the real numbers $\mathbb{R}$, together with an extra element $-\infty$ whitch represents minus infinity. In this (semiring) the tropical sum of real numbers is ther maximum and the tropical product of real numbers is their usual sum

$$x \oplus y := \max(x, \ y) \quad \& \quad x \odot y := x + y$$

In Tropical Geometry often infinity instead of minus infinity and min in stead of max are used.

This does not change any of the underlying theories of tropical algebras, as the the two tropical semirings are exchangeable.
Möglicherweise noch eine transoformation zwischen den semiringen angeben.

Definition 1.1. *A semiring is a set R equipped with two binary operations + and ·, called addition and multiplication, such that (Wikipedia.en Semiring)*

- *(R, +) is a cummutative monoid with identity element 0:*

  - *$(a + b) + c = a + (b + c)$*
  - *$0 + a = a + 0 = 0$*
  - *$a + b = b + a$*

- *(R, ·) is a monoid with identity element 1:*

  - *$(a \cdot b) \cdot c = a \cdot (b \cdot c)$*
  - *$1 \cdot a = a \cdot 1 = a$*

- *Multiplication left and right distributes over addition*

  - *$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$*
  - *$(a + b) \cdot c = (a \cdot c) + (b \cdot c)$*

- *Multiplication by 0 annihilates*

  - *$0 \cdot a = a \cdot 0 = 0$*

Proposition 1.1. $\mathbb{T} := (\mathbb{R} \cup \{-\infty\}, \oplus, \odot)$ *is a semiring called the tropical semiring. (Maclagan & Sturmfels, 2015, p. 10)*

Proof.

(1): The neutral element for the tropical sum is $-\infty$ since for $x \in \mathbb{R} \cup \{-\infty\}$ the following stands $x \oplus \infty = \max(x, \ -\infty) = x$ and with $x \odot 0 = x + 0 = x$ for $x \in \mathbb{R}$, 0 is the neutral elemet of tropical multiplication.

(2): Both addition and multiplication are commutative. To prove this we take $x, y \in \mathbb{R} \cup \{-\infty\}$ and do a case distinct. Because $\mathbb{R}$ is a field w.l.o.g. we set $x = -\infty, y \in \mathbb{R}$

$$-\infty \oplus y = \max(-\infty,\ y) = y = \max(y,\ -\infty) = y \oplus \infty$$
$$\infty \odot y = \infty + y = \infty = y + \infty = y \odot \infty$$

(3): Tropical multiplication distributes over addition. Take $x, y, z \in \mathbb{R}$ then

$$x \odot (y \oplus z) = x + \max(y,\ z) = \max(x + y,\ x + z) = (x \odot y) \oplus (x \odot z)$$
$$(y \oplus z) \odot x = \max(y,\ z) + x = \max(y + x,\ z + x) = (y \odot x) \oplus (z \odot x)$$

(4): Multiplication by $-\infty$ annihilates $-\infty \odot x = -\infty\ \forall x \in \mathbb{R} \cup \{-\infty\}$.

$\square$

A essential feature of tropical arithmetics is that there is no subtraction. Take $a, b \in \mathbb{R} \cup \{-\infty\}$ with $a < b$ then the equation $a \oplus x = b$ has no solution x at all. (Maclagan & Sturmfels, 2015, p. 11)

Definition 1.2. *(Zhang, Naitzat, & Lim, 2018, p. 2) A tropical monomial in d variables $x_1, \ldots, x_d$ is an expression of the from*

$$x \odot x_1^{a_1} \odot x_2^{a_2} \cdots \odot x_d^{a_d}$$

where $c \in \mathbb{R} \cup \{-\infty\}$ and $a_1, \ldots, a_d \in \mathbb{N}$. As a convenient shorthand, we will also write a tropical monomial in multiindex notation as cx_α where $\alpha = (a_1, \ldots, a_d) \in \mathbb{N}_d$ and $x = (x_1, \ldots, x_d)$. Note that $x^\alpha = 0 \odot x^\alpha$.

Definition 1.3. *(Zhang et al., 2018, p. 2) Following notations above, a tropical polynomial $f(x) = f(x_1, \ldots, x_d)$ is a finite tropical sum of tropical monomials*

$$f(x) = c_1 x^{\alpha_1} \oplus \cdots \oplus c_r x^{\alpha_r}$$

where $\alpha_i = (\alpha_{i1}, \ldots, \alpha_{id}) \in \mathbb{N}^d$ and $c_i \in \mathbb{R} \cup \{-\infty\}$, $i = 1, \ldots, r$. We will assume that a monomial of a given multiindex appears at most once in the sum, i.e. $\alpha_i \neq \alpha_j$ for any $i \neq j$.

Lemma 1.0.1. *Let f be a tropical polynomial*

$$f(x) = c_1^{\alpha_1} \oplus \cdots \oplus c_r x^{\alpha_r}$$

as in 1.3. Then f has three important properties:

(1) f is continuous.

(2) f is piecewise-linear, where the number of pieces is finite.

(3) f is convex, i.e. $p(\frac{x+y}{2}) \leq \frac{1}{2}(p(x) + p(y)) \forall x, y \in \mathbb{R}$

Proof.

(1): The minimum of continuous functions is still continuous.

(2): Every monomial $c_i x^{\alpha_i} = c_i + xy\alpha_{i1} + \cdots + x_r \alpha_{ir}$ is per definition linear. Because of (1), linearity can only be broken where $c_i x^{\alpha_i} = c_j x^{\alpha_j}$ for $i \leq j$ and $i, j = 1, \ldots, r$.

A piece is one whole piece of $c_i x^{\alpha_i}$ where linearity is not broken. If we introduce $c_i x^{\alpha_i}$ one after another, then in the i-th step not more than i^2 new pieces can be created, so there can only be $\sum_{i=1}^{r} i^2$ or less pieces.

(3): On one piece f is concave. Moving from one piece to another the slope can only increase, with means f is still concave.

$\square$

Definition 1.4. *(Zhang et al., 2018, p. 3) Following notations above, a tropical rational function is a standard difference, or, equivalently, a tropical quotient of two tropical polynomials $f(x)$ and $g(x)$:*

$$f(x) - g(x) = f(x) \oslash g(x)$$

Proposition 1.2. $\mathbb{T}[X_1, \ldots, X_d] := \{f : \mathbb{T}^d \to \mathbb{T}; f \text{ is tropical polynomial}\}$ *and* $\mathbb{T}(X_1, \ldots, X_d) := \{f : \mathbb{T}^d \to \mathbb{T}; f \text{is tropical rational funktion}\}$ *are both semifields. (Zhang et al., 2018, p. 3)*

Proof. Let $g, f, h \in \mathbb{T}(X_1, \ldots, X_d)$ with

$$f(x) = f_1(x) \oslash f_2(x) = (\oplus_{i=0}^{r} c_{1i} x^{\alpha_{1i}}) \oslash (\oplus_{i=0}^{r} c_{2i} x^{\alpha_{2i}})$$
$$g(x) = g_1(x) \oslash g_2(x) = (\oplus_{i=0}^{r} d_{1i} x^{\beta_{1i}}) \oslash (\oplus_{i=0}^{r} d_{2i} x^{\beta_{2i}})$$
$$h(x) = h_1(x) \oslash h_2(x) = (\oplus_{i=0}^{r} e_{1i} x^{\gamma_{1i}}) \oslash (\oplus_{i=0}^{r} e_{2i} x^{\gamma_{2i}})$$

We begin the proof by showing, that for two tropical polynomials $a(x) = \oplus_{i=0}^{r} z_i x^{\zeta_i}), b(x) = \oplus_{i=0}^{r} o_i x^{\omega_i}) \in \mathbb{T}[X_1, \ldots, X_r]$ the normal sum is a tropical polynomial $(a + b)(x) \in \mathbb{T}[X_1, \ldots, X_r]$.

$$(a + b)(x) = a(x) + b(x)$$
$$= (\oplus_{i=0}^{r} z_i x^{\zeta_i}) + (\oplus_{i=0}^{r} o_i x^{\omega_i})$$
$$= \oplus_{i,j=0,\ldots,r} (z_i + \zeta_i * x + o_j + \omega_i * x)$$
$$= \oplus_{i,j=0,\ldots,r} ((z_i + o_j) + x^{\zeta_i + \omega_i}) \in \mathbb{T}[X_1, \ldots X_r]$$

Other than with the proof of the Tropical semiring we will show that topical tropical addition and tropical multiplication of tropical polynomials as tropical rational functions stay tropical polynomials respectively tropical rational functions. All other axioms stay pointwise the same.

(1): The Tropical sum of a Tropical rational funktions is a tropical tropical rations funktion

$$(f \oplus g)(x) = f(x) \oplus g(x)$$
$$= (f_1(x) \oslash f_2(x)) \oplus (g_1(x) \oslash g_2(x))$$
$$= \min\{f_1(x) - f_2(x), g_1(x) - g_2(x)\}$$
$$= \min\{f_1(x) + g_2(x), g_1(x) + f_2(x)\} - f_2(x) - g_2(x)$$
$$= (f_1(x) + g_2(x) \oplus g_1(x) + f_2(x)) \oslash (f_2(x) + g_2(x)) \in \mathbb{T}(X_1, \ldots, X_r).$$

Since Addition as tropical addition of tropical polynomials is a tropical polynomial.

(2): The Tropical product of a Tropical rational functions is a tropical tropical rational function

$$(f \odot g)(x) = f(x) \odot g(x)$$
$$= (f_1(x) \oslash f_2(x)) \odot (g_1(x) \oslash g_2(x))$$
$$= f_1(x) - f_2(x) + g_1(x) - g_2(x)$$
$$= (f_1(x) + g_1(x)) - (f_2(x) + g_2(x)) \in \mathbb{T}(X_1, \ldots, X_r)$$

(3): The neutral element for the tropical sum is $-\infty = -\infty \oslash x = x \oslash -\infty \ x \in \mathbb{T}$ since for $f(x) \in \mathbb{T}(X_1, \ldots, X_r)$ as above, the following stands $f(x) \oplus \infty = \max(f(x), -\infty) = f(x)$ and with $f(x) \odot 0 = f(x) + 0 = f(x) \forall f(x) \in \mathbb{T}$ 0 is the neutral element of tropical multiplication.

$\square$

Bemerkung 1.1. *We regard a tropical polynomial $f = f \oslash 0$ as a special case of a tropical rational function and thus $\mathbb{T}[X_1, \ldots, X_r] \subseteq \mathbb{T}(X_1, \ldots, X_r)$ (Zhang et al., 2018, p. 3).*

Bemerkung 1.2.

- *A d-variate tropical polynomial $f(x)$ defines a function $f : \mathbb{T}^d \to \mathbb{T}$ that is a convex function in the usual sense as taking $\max$ and $\sum$ of convex functions preserve convexity (Boyd, Boyd, & Vandenberghe, 2004).*

- *As such, a tropical rational function $f \oslash g : \mathbb{T}^d \to \mathbb{T}$ is a DC function or differenceconvex function (Hartman et al., 1959).*

Definition 1.5. *$R : \mathbb{R}^d \to \mathbb{R}^p, x = (x_1, \ldots, x_d) \mapsto (f_1(x), \ldots, f_p(x))$, is called a tropical polynomial map if each $f_i : \mathbb{R}^d \to \mathbb{R}$ is a tropical polynomial, $i = 1, \ldots, p$, and a tropical rational map if $f_1, \ldots, f_p$ are tropical rational functions. We will denote the set of tropical polynomial maps by $Pol(d, p)$ and the set of tropical rational maps by $Rat(d, p)$. So $Pol(d, 1) = \mathbb{T}[X_1, \ldots, x_d]$ and $Rat(d, 1) = \mathbb{T}(x_1, \ldots, x_d)$ (Zhang et al., 2018, p. 3).*

2 Tropical hypersurfaces

Definition 2.1. *The tropical hypersurface of a tropical polynomial $f(x) = c_1 x^{\alpha_1} \oplus \cdots \oplus c_r x^{\alpha_r}$ is*

$$\Gamma(f) := \{x \in \mathbb{R}^d : c_i x^{\alpha_i} = c_j x^{\alpha_j} = f(x) \, for some \, \alpha_i \neq \alpha_j\}$$

i.e., the set of points x at which the value of f at x is attained by two or more monomials in f (Zhang et al., 2018, p. 3).

Definition 2.2. *The Newton polygon of a tropical polynomial $f(x) = c_1 x^{\alpha_1} \oplus \cdots \oplus c_r x^{\alpha_r}$ is the convex hull of $\alpha_1, \ldots, \alpha_r \in \mathbb{N}^d$, regarded as points in $\mathbb{R}^d$,*

$$\Delta(f) := Conv \alpha_i \in \mathbb{R}^d : c_i \neq -\infty, i = 1, \ldots, r$$

(Zhang et al., 2018, p. 3).

Definition 2.3. *A linear region of $F \in Rat(d, m)$ is a maximal connected subset of the domain on which F is linear. The number of linear regions of F is denoted $\mathcal{N}(f)$ (Zhang et al., 2018, p. 4).*

2.1 Transformations of tropical polynomial

Proposition 2.1. *Let f be a tropical polynomial and let $a \in \mathbb{N}$. Then*

$$\mathcal{P}(f^a) = a\mathcal{P}(f)$$

. $a\mathcal{P}(f) = \{ax : x \in \mathcal{P}(f)\} \subset \mathbb{R}^{d+1}$ is a scaled version of $\mathcal{P}(f)$ with the same shape but different volume (Zhang et al., 2018, p. 4).

Proof. □

Definition 2.4. *The Minkowski sum of two sets P_1 and P_2 in $\mathbb{R}^d$ is the set*

$$P_1 + P_2 := \{x_1 + x_2 \in \mathbb{R}^d : x_1 \in P_1, x_2 \in P_2\}$$

; and for $\lambda_1, \lambda_2 \geq 0$, their weighted Minkowski sum is

$$\lambda_1 P_1 + \lambda_2 P_2 := \{\lambda_1 x_1 + \lambda_2 x_2 \in \mathbb{R}^d : x_1 \in P_1, x_2 \in P_2\}$$

(Zhang et al., 2018, p. 4).

Proposition 2.2. *(Zhang et al., 2018, p. 4) Let $f, g \in Pol(d, 1) = \mathbb{T}[x_1, \ldots, x_d]$ be tropical polynomials. Then*

$$\mathcal{P}(f \odot g) = \mathcal{P}(f) + \mathcal{P}(g), \mathcal{P}(f \oplus g) \quad = Conv(\mathcal{V}(\mathcal{P}(f)) \cup \mathcal{V}(\mathcal{P}(g))).$$

Theorem 2.1. *(Gritzmann-Sturmfels). Let $P_1, \ldots, P_k$ be polytopes in $\mathbb{R}^d$ and let m denote the total number of nonparallel edges of $P_1, \ldots, P_k$. Then the number of vertices of $P_1 + \cdots + P_k$ does not exceed*

$$\sum_{j=0}^{d-1} \binom{m-1}{j}.$$

The upper bound is attained if all P_i's are zonotopes and all their generating line segments are in general positions. (Gritzmann & Sturmfels, 1993)

Corollary 2.1. *Let $\mathcal{P} \in \mathbb{R}^{d+1}$ be a zonotope generated by m line segments $P_1, \ldots, P_m$. Let $\pi : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ be the projection. Suppose P satisfies:*

(i) the generating line segments are in general positions;

(ii) the set of projected vertices $\{\pi(v) : v \in \mathcal{V}(\mathcal{P})\} \subseteq \mathbb{R}^d$ are in generalposition.

Then P has

$$\sum_{j=0}^{d} \binom{m}{j}$$

vertices on its upper faces. If either (i) or (ii) is violated, then this becomes an upper bound. (Zhang et al., 2018, p. 4)

References

Boyd, S., Boyd, S. P., & Vandenberghe, L. (2004). *Convex optimization.* Cambridge university press.

Gritzmann, P., & Sturmfels, B. (1993). Minkowski addition of polytopes: computational complexity and applications to gröbner bases. *SIAM Journal on Discrete Mathematics*, *6*(2), 246–269.

Hartman, P., et al. (1959). On functions representable as a difference of convex functions. *Pacific Journal of Mathematics*, *9*(3), 707–713.

Maclagan, D., & Sturmfels, B. (2015). *Introduction to tropical geometry* (Vol. 161). American Mathematical Soc.

Zhang, L., Naitzat, G., & Lim, L.-H. (2018). Tropical geometry of deep neural networks. *arXiv preprint arXiv:1805.07091*.