SAL

论文：《SAL: Sign Agnostic Learning of Shapes from Raw Data》
地址：https://arxiv.org/abs/1911.10414
年份：CVPR 2020

Introduction

任务：表面重建
技术贡献：
(1) 提出了重建物体表面时将 SDF 初始化为球的几何初始化方法；
(2) 证明了基于论文的方法训练的 SDF 具有 plane reproduction 的性质。

Method

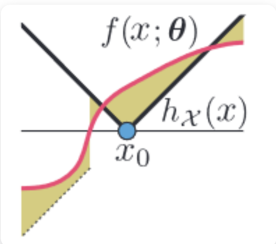
首先是问题设定，文章希望从 raw data (如不带法向的点云，triangle soup) 中学习一个 SDF，用于重建表面。
文章定义了训练的 loss：

$$\operatorname{loss}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{x}\sim D_{\mathcal{X}}} \tau(f(\boldsymbol{x};\boldsymbol{\theta}), h_{\mathcal{X}}(\boldsymbol{x}))$$

其中 $D_{\mathcal{X}}$ 是由输入数据 $\mathcal{X}$ 确定的分布，$h_{\mathcal{X}}(\boldsymbol{x})$ 是度量到 $\mathcal{X}$ 的无符号距离，τ 是一个度量相似性的函数。
要求 τ 满足以下两个性质：

1. Sign agnostic：$\tau(-a,b)=\tau(a,b)$
2. Monotonic：$\frac{\partial\tau}{\partial a}(a,b)=\rho(a-b)$

其中 ρ 是一个单调递增函数，且 $\rho(0)=0$。τ 的一个例子是 $\tau(a,b)=||a|-b|$



这是一个一维的例子，$\mathcal{X}=\{x_0\}, h_{\mathcal{X}}(x)=|x-x_0|$ 是无符号距离，$\tau(a,b)=||a|-b|$，上面定义的 loss 就会要求最小化图上黄色区域的面积，最终得到的结果 $f(\boldsymbol{x};\boldsymbol{\theta})$ 会是 $h_{\mathcal{X}}(x)$ 的有符号版本。
$h_{\mathcal{X}}(x)$ 由两种选择：

$$h_2(z)=\min_{x\in\mathcal{X}}||z-x||_2$$

$$h_0(z) = \begin{cases} 0 & z \in \mathcal{X} \\ 1 & z \notin \mathcal{X} \end{cases}$$

对于 $D_{\mathcal{X}}$，选择是在每一处点 $\boldsymbol{x} \in \mathcal{X}$ 投射一个各向同性的高斯 $\mathcal{N}(\boldsymbol{x}, \sigma^2 I)$ 以及狄拉克 δ 函数 (对于 $h_0(z)$)，这样的话 loss 就是：

$$\operatorname{loss}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{z} \sim \mathcal{N}_\sigma(\mathcal{X})} \left| |f(\boldsymbol{x}; \boldsymbol{\theta})| - 1 \right|^l + \mathbb{E}_{\boldsymbol{x} \sim \mathcal{X}} |f(\boldsymbol{x}; \boldsymbol{\theta})|^l$$

对于 $h_2(z)$，有

$$\operatorname{loss}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{z} \sim \mathcal{N}_\sigma(\mathcal{X})} \left| |f(\boldsymbol{x}; \boldsymbol{\theta})| - h_2(z) \right|^l$$

网络选用 MLP

Neural architecture. Although SAL can work with different parametric models, in this paper we consider a multilayer perceptron (MLP) defined by

$$f(\boldsymbol{x}; \boldsymbol{\theta}) = \varphi\big(\boldsymbol{w}^T f_\ell \circ f_{\ell-1} \circ \cdots \circ f_1(\boldsymbol{x}) + b\big), \qquad (8)$$

and

$$f_i(\boldsymbol{y}) = \nu(\boldsymbol{W}_i \boldsymbol{y} + \boldsymbol{b}_i), \boldsymbol{W} \in \mathbb{R}^{d_i^{out} \times d_i^{in}}, \boldsymbol{b}_i \in \mathbb{R}^{d_i^{out}}, \quad (9)$$

where $\nu(a) = (a)_+$ is the ReLU activation, and $\boldsymbol{\theta} = (\boldsymbol{w}, b, \boldsymbol{W}_\ell, \boldsymbol{b}_\ell, \dots, \boldsymbol{W}_1, \boldsymbol{b}_1)$; φ is a strong non-linearity, as defined next:

Geometric network initialization

接下来希望去初始化 MLP 使得一开始为半径为 r 的球体。
下面就是证明 2 个结论，选用其中给出的初始化参数即可将 MLP 几何初始化为球体。

(a) (b) (c)

上图是 MLP 中线性层宽度为 100，200 和 2000 的情况，可以看到线性层宽度越宽，初始化结果越接近球体。这是由于后面结论的推导用到了大数定律，宽度越宽则越接近期望。

Theorem 1. *Let f be an MLP (see equations 8-9). Set, for $1 \le i \le \ell$, $\boldsymbol{b}_i = 0$ and $\boldsymbol{W}_i$ i.i.d. from a normal distribution $\mathcal{N}(0, \frac{\sqrt{2}}{\sqrt{d_i^{out}}})$; further set $\boldsymbol{w} = \frac{\sqrt{\pi}}{\sqrt{d_\ell^{out}}}\mathbf{1}$, $c = -r$. Then, $f(\boldsymbol{x}) \approx \varphi(\|x\| - r)$.*

Proof of Theorem 1

定理 1 的大致意思是，先不看最后的 $\boldsymbol{w}$ 和 c，通过指定中间层的参数，使最后一层隐藏层的输出的模长大小为 $||\boldsymbol{x}||_2$。
具体是让 $\boldsymbol{W}_i$ 的值是 $\mathcal{N}(0, \frac{2}{d_i^{out}})$ 的独立同分布，$\boldsymbol{b}_i = 0$。我们要证明的是经过隐藏层 $g(\boldsymbol{x}) = f_{l-1} \circ \cdots \circ f_1$ 后，有 $||g(\boldsymbol{x})|| \approx ||\boldsymbol{x}||$。
考虑一层隐藏层的情况即可，因为经过一层后有上述性质，那么经过任意层后模长还会是一样的。所以考虑 $h(\boldsymbol{x}) = \nu(\boldsymbol{W}\boldsymbol{x} + \boldsymbol{b})$，$\boldsymbol{W} \in \mathbb{R}^{d^{out} \times d^{in}}, \boldsymbol{x} \in \mathbb{R}^{d^{in}}, \boldsymbol{b} \in \mathbb{R}^{d^{out}}$。记 $k = d^{out}$
先把矩阵 $\boldsymbol{W}$ 拆成行来看，那么就会有 k 个 d^{in} 维的向量 $\boldsymbol{W}_{i,:}$，然后与 $\boldsymbol{x}$ 点乘后就是标量。这里相当于提前将隐藏层结果变为标量方便证明模长。

$$||h(\boldsymbol{x})||^2 = \sum_{i=1}^{k} \nu(\boldsymbol{W}_{i,:} \cdot \boldsymbol{x})^2 = \frac{1}{k} \sum_{i=1}^{k} \nu(\sqrt{k}\boldsymbol{W}_{i,:} \cdot \boldsymbol{x})^2$$

然后是利用大数定律将离散求和转换为算期望连续的形式，有

$$||\boldsymbol{x}||^2 \int_{\mathbb{R}^k} \nu(\boldsymbol{y} \cdot \frac{\boldsymbol{x}}{||\boldsymbol{x}||})^2 \mu(\boldsymbol{y}) d\boldsymbol{y}$$

相当于 $\sqrt{k}\boldsymbol{W}_{i,:}$ 是随机变量序列，其满足独立同分布的 $\mathcal{N}(0, 2)$。
然后是做了一步变换 $\boldsymbol{y} = \boldsymbol{R}\boldsymbol{y}'$，其中 $\boldsymbol{R}$ 是 $k \times k$ 的正交阵，且使 $\boldsymbol{R}^T \frac{\boldsymbol{x}}{||\boldsymbol{x}||} = (1, 0, \cdots, 0)^T$，这样做的目的是消除 $\boldsymbol{x}$ 这个变量。这样就会有

$$\boldsymbol{y} \cdot \frac{\boldsymbol{x}}{||\boldsymbol{x}||} = \boldsymbol{R}\boldsymbol{y}' \cdot \frac{\boldsymbol{x}}{||\boldsymbol{x}||} = \boldsymbol{y}'^T \boldsymbol{R}^T \frac{\boldsymbol{x}}{||\boldsymbol{x}||} = \boldsymbol{y}'^T (1, 0, \cdots, 0)^T = y_1$$

其中 y_1 是 $\boldsymbol{y}'$ 的第 1 个元素。
继续前面的式子，代换之后有

$$||\boldsymbol{x}||^2 \int_{\mathbb{R}^k} \nu(y_1)^2 \mu(\boldsymbol{R}\boldsymbol{y}') d(\boldsymbol{R}\boldsymbol{y}')$$

先看 μ ，参照 https://zhuanlan.zhihu.com/p/650679177 ，多元正态分布有密度旋转不变性，因此 $\mu(\boldsymbol{R}\boldsymbol{y}') = \mu(\boldsymbol{y})$，再看 d 后面的项，$\boldsymbol{R}$ 相当于就是 Jacobian，即有 $d\boldsymbol{y} = d(\boldsymbol{R}\boldsymbol{y}') = |\det(\boldsymbol{R})| d\boldsymbol{y}'$ ，正交阵的行列式为 1，所以 $d\boldsymbol{y} = d\boldsymbol{y}'$。(论文中这里的推导并没有这么具体，这是按照我的个人理解来的)
经过变换后，上式会变为

$$||\boldsymbol{x}||^2 \int_{\mathbb{R}^k} \nu(y_1)^2 \mu(\boldsymbol{y}') d\boldsymbol{y}'$$

后面一步是将多元高斯的积分转换为一元高斯的积分，这一步具体原理没在网上找到。

$$||\boldsymbol{x}||^2 \int_{\mathbb{R}} \nu(y_1)^2 \mu(y_1) dy$$

ν 用的是 ReLU，只在正半轴有值，所以做个一个对称

$$\frac{||\boldsymbol{x}||^2}{2}\int_{\mathbb{R}} y_1^2\mu(y_1)dy$$

最后积分里的式子相当于算期望 $E[x^2]$，因此值为 $\mu^2+\sigma^2$ 即 2。所以最终结果为 $||\boldsymbol{x}||^2$
论文最后还考虑了有 skip connection 的情况，比较简单。

Theorem 2. *Let $f:\mathbb{R}^d\rightarrow\mathbb{R}$ be an MLP with ReLU activation, ν, and a single hidden layer. That is, $f(\boldsymbol{x})=\boldsymbol{w}^T\nu(\boldsymbol{W}\boldsymbol{x}+\boldsymbol{b})+c$, where $\boldsymbol{W}\in\mathbb{R}^{d^{out}\times d}$, $\boldsymbol{b}\in\mathbb{R}^{d^{out}}$, $\boldsymbol{w}\in\mathbb{R}^{d^{out}}$, $c\in\mathbb{R}$ are the learnable parameters. If $\boldsymbol{b}=0$, $\boldsymbol{w}=\frac{\sqrt{2\pi}}{\sigma d^{out}}\mathbf{1}$, $c=-r$, $r>0$, and all entries of $\boldsymbol{W}$ are i.i.d. normal $\mathcal{N}(0,\sigma^2)$ then $f(\boldsymbol{x})\approx\|\boldsymbol{x}\|-r$. That is, f is approximately the signed distance function to a $d-1$ sphere of radius r in $\mathbb{R}^d$, centered at the origin.*

定理 2 的大致意思是，将隐藏层参数设为定理 1 中给出的值，令最后一层的 $\boldsymbol{w}=\frac{\sqrt{2\pi}}{\sigma d^{out}}\mathbf{1}, c=-r$，就会使输出的 $f(\boldsymbol{x})=||\boldsymbol{x}||-r$。
这里也记 $k=d^{out}$，$\boldsymbol{W}$ 的值属于 $\mathcal{N}(0,\sigma^2)$。前面的推导过程基本与定理 1 思想一致，主要是看最后一个等式

$$\frac{\sqrt{2\pi}||\boldsymbol{x}||}{2\sigma}\int_{\mathbb{R}}|v_1|\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{v_1^2}{2\sigma^2}}dv_1$$

积分内是偶函数，利用对称性就有

$$\frac{\sqrt{2\pi}||\boldsymbol{x}||}{\sigma}\int_0^{+\infty}v_1\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{v_1^2}{2\sigma^2}}dv_1$$

积分里的原函数很好求，最终结果仍是 $||\boldsymbol{x}||$。

Properties

TBD.

Experiments

Failure case 是不能重建细的物体，可能是因为细的部分采样的点少，模型不能很好的建模这一部分的距离关系。