

深度学习（三十）——Graph NN（2），自动求导

2018-01-15

Graph NN（续）

https://mp.weixin.qq.com/s/5DmpgPN4t3p3H53Xu7_-3A

北大、微软亚洲研究院：高效的大规模图神经网络计算

https://mp.weixin.qq.com/s/BFJD8i_yg1Y6fxZS5or-rw

Bengio最新论文提出GibbsNet：深度图模型中的迭代性对抗推断

<https://zhuanlan.zhihu.com/p/48834333>

GCN in 2018：2018年顶会论文中的图卷积神经网络的理论与应用

<https://mp.weixin.qq.com/s/zg3yW7e4UKls9-m6WmcbvA>

GraphWave：一种全新的无监督网络嵌入方法

https://mp.weixin.qq.com/s/mamet6l_IA7fhoYkysZ7PQ

华为联合LSE提出KONG：有序近邻图的核函数

<https://mp.weixin.qq.com/s/OnRB44tliuTFcjlmuRG3Xw>

图神经网络“理论在哪里”？

<https://mp.weixin.qq.com/s/Uy2ekBiwkl2slo637b-16g>

北大、微软提出NGra：高效大规模图神经网络计算

<https://mp.weixin.qq.com/s/dilzbc0tpCW4xhblQu8mCw>

阿里凑单算法首次公开！基于Graph Embedding的打包购商品挖掘系统解析

<https://mp.weixin.qq.com/s/chiHw5gKnJyTJTQeF6gViw>

在向量空间中启用网络分析和推理，清华大学崔鹏博士最新分享

<https://mp.weixin.qq.com/s/kQlxLDHLI6xxFzwJVjFj7w>

GraRep: 基于全局结构信息的图结点表示学习

<https://mp.weixin.qq.com/s/c6ZhSk4r3pvnjHsvpwkkSw>

用图卷积网络(GCN)来做语义角色标注

<https://mp.weixin.qq.com/s/6vhFfSh2mveBiZXB1oZb1Q>

图分类：结合胶囊网络Capsule和图卷积GCN

<https://mp.weixin.qq.com/s/aGP8pcsCmEdjdCWVjA82Jg>

近期必读的5篇CVPR 2019图卷积网络相关论文和代码

<https://mp.weixin.qq.com/s/XApSbi-Pg-AeYGkPN3fldg>

旷视研究院提出ML-GCN：基于图卷积网络的多标签图像识别模型

https://mp.weixin.qq.com/s/49vnVOO0G_JvKrWcsN2_Ww

关系图注意力网络-Relational Graph Attention Networks

https://mp.weixin.qq.com/s/rvcj9-6KIBsVmF_CAsip2A

超越标准GNN！DeepMind、谷歌提出图匹配网络

<https://mp.weixin.qq.com/s/UotqgRjCTpjPrsIEWBRPxA>

基于随机游走的图匹配算法

<https://mp.weixin.qq.com/s/dDaFhssFEYxS7EIMy4ekJw>

基于图嵌入的深度图匹配

https://mp.weixin.qq.com/s/LZvxvDpxQEtlKuXoxT_gTQ

可变形曲面跟踪，亮风台新出基于图匹配的方法

<https://mp.weixin.qq.com/s/7DyPJ9LnqZ9XyAop33SxSw>

ST-GCN动作识别算法详解

<https://mp.weixin.qq.com/s/fxVsN2dDmayxJfxBRIXHhQ>

解读PingSage：图卷积神经网络在数十亿数据网络级别推荐系统的应用

<https://mp.weixin.qq.com/s/lyy3AhqLDBT88B2LSSlbZQ>

图表示解决长文本关系匹配问题：腾讯提出概念交互图算法

https://mp.weixin.qq.com/s/bvp3NIrrarJc_MesKy1x_A

崔泽宇：套装搭配推荐在图神经网络上的应用

https://mp.weixin.qq.com/s/_8K0s9WceJ-xIRViHhz2Zw

Google图挖掘团队最新博客《图表示学习中的创新》

<https://mp.weixin.qq.com/s/c3SBGlxzJOYhQBrJ2h3j0g>

呼奋宇：深度层次化图卷积神经网络

<https://mp.weixin.qq.com/s/YPV2BR6eayKUIPazUeZVnQ>

何时能懂你的心——图卷积神经网络（GCN）

<https://mp.weixin.qq.com/s/sRKW8DLXZXWLUUVTb12F4Q>

“AI新贵”图神经网络算法及平台在阿里的大规模实践

<https://mp.weixin.qq.com/s/tAfTmGWqG6lR8SOP0uKW6g>

什么限制了GNN的能力？ 首篇探究GNN普适性与局限性的论文出炉！

<https://mp.weixin.qq.com/s/Orv47r4EchVIR7VcleoJ5Q>

谷歌图表征学习创新：学习单个节点多个嵌入&自动学习最优超参数

<https://mp.weixin.qq.com/s/DNePTCpyjrIzEixw5L7w5A>

GraphSAGE：我寻思GCN也没我牛逼

<https://mp.weixin.qq.com/s/1DHvLLysMU24dBeLzbSpUA>

GraphSAGE

<https://mp.weixin.qq.com/s/lcLk-fMjKO19BaHbuUCeXg>

GraphSAGE 算法原理，实现和应用

<https://mp.weixin.qq.com/s/C-Pa1jznQntyhocdxS-4Hg>

节点嵌入训练加快300倍！ 解读开源高性能图嵌入系统GraphVite

<https://mp.weixin.qq.com/s/9MWoCmtKPPVs3Rmko-7adQ>

10亿节点异构网络中，GCN如何应用？

<https://mp.weixin.qq.com/s/gE0jGuduKOQYqkTAuhC2Nw>

基于空间对齐的图卷积网络

https://mp.weixin.qq.com/s/xJ05ctIYhyVy__6CaOm4WQ

利用图Transformer实现基于知识图谱的文本生成

<https://zhuanlan.zhihu.com/p/81350196>

GAT（图注意力模型）

<https://mp.weixin.qq.com/s/nGPUcDHTrG6KwAqDCkfA1w>

基于超图网络模型的图网络进化算法

<https://mp.weixin.qq.com/s/XoVUJln3oyhk2jMOtWwfog>

基于图神经网络的协同过滤算法

[https://mp.weixin.qq.com/s?
__biz=MzU2OTA0NzE2NA==&mid=2247515242&idx=2&sn=7497044ec2f80215061fb0de8abb2f72](https://mp.weixin.qq.com/s?__biz=MzU2OTA0NzE2NA==&mid=2247515242&idx=2&sn=7497044ec2f80215061fb0de8abb2f72)

用于Web级推荐系统的图卷积神经网络

[https://mp.weixin.qq.com/s?
__biz=MzU2OTA0NzE2NA==&mid=2247515402&idx=2&sn=726b417a71789e65430cbd931c78959e](https://mp.weixin.qq.com/s?__biz=MzU2OTA0NzE2NA==&mid=2247515402&idx=2&sn=726b417a71789e65430cbd931c78959e)

基于门控图网络实现图到序列学习

<https://mp.weixin.qq.com/s/aIU8mP5nlnqR4Qt-4RfMgQ>

孙付伟：Graph Embedding在知乎的应用实践

<https://blog.csdn.net/yyl424525/article/details/100057907>

Cluster-GCN: An Efficient Algorithm for Training Deep and Large Graph Convolutional Networks

<https://zhuanlan.zhihu.com/p/86216369>

从3/4层拓展到56层，如何训练超级深层的图卷积神经网络

https://mp.weixin.qq.com/s/ceC1_39cmpqYOoHsu-4sCQ

HEC-Montreal唐建博士：图神经网络推理

https://mp.weixin.qq.com/s/emPDgw_c68JgbpYnem4JvA

一文读懂简化的图卷积网络GCN（SGC）

https://mp.weixin.qq.com/s/olt_i6JxfsIFSdbz0Dx4oA

用于深度时空图建模的Graph WaveNet

https://mp.weixin.qq.com/s/LzaP6BWaKXL4O_Jf9QN7Rw

基于层次化注意力图网络和多视角学习的商品推荐

<https://mp.weixin.qq.com/s/EKb9hN9cFrwNtSWWBpE4dw>

首次基于神经架构搜索自动生成图卷积结构，刷新人体动作识别准确率

自动求导

DL发展到现在，其基本运算单元早就不止CNN、RNN之类的简单模块了。针对新运算层出不穷的现状，各大DL框架基本都实现了自动求导的功能。

论文：

《Automatic Differentiation in Machine Learning: a Survey》

Numerical differentiation

数值微分最大的特点就是很直观，好计算，它直接利用了导数定义：

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h}$$

不过这里有一个很大的问题：h怎么选？选大了，误差会很大；选小了，不小心就陷进了浮点数的精度极限里，造成舍入误差。

第二个问题是由于参数比较多时，对深度学习模型来说，上面的计算是不够高效的，因为每计算一个参数的导数，你都需要重新计算 $f(x + h)$ 。

因此，这种方法并不常用，而主要用于做梯度检查（Gradient check），你可以用这种不高效但简单的方法去检查其他方法得到的梯度是否正确。

Symbolic differentiation

符号微分的主要步骤如下：

- 1.需要预置基本运算单元的求导公式。
- 2.遍历计算图，得到运算表达式。
- 3.根据导数的代入法则和四则运算法则，求出复杂运算的求导公式。

这种方法没有误差，是目前的主流，但遍历比较费时间。

Automatic differentiation

除此之外，常用的自动求导技术，还有Automatic differentiation。（请注意这里的AD是一个很狭义的概念。）

类比复数的概念：

$$x = a + bi \quad (i^2 = -1)$$

我们定义Dual number：

$$x \mapsto \tilde{x} = x + \dot{x}d \quad (d^2 = 0)$$

定义Dual number的运算法则：

$$\begin{aligned}(\tilde{x} + \dot{x}d) + (\tilde{y} + \dot{y}d) &= \tilde{x} + \tilde{y} + (\dot{x} + \dot{y})d \\(\tilde{x} + \dot{x}d)(\tilde{y} + \dot{y}d) &= \tilde{x}\tilde{y} + \dot{x}\dot{y}d + \tilde{x}\dot{y}d + \dot{x}\tilde{y}d^2 = \tilde{x}\tilde{y} + (\dot{x}\dot{y} + \tilde{x}\dot{y})d \\-(\tilde{x} + \dot{x}d) &= -\tilde{x} - \dot{x}d \\\frac{1}{\tilde{x} + \dot{x}d} &= \frac{1}{\tilde{x}} - \frac{\dot{x}}{\tilde{x}^2}d\end{aligned}$$

dual number有很多非常不错的性质。以下面的指数运算多项式为例：

$$f(x) = p_0 + p_1x + p_2x^2 + \dots + p_nx^n$$

用 $\tilde{x} + \dot{x}d$ 替换 x ，则有：

$$\begin{aligned}f(\tilde{x} + \dot{x}d) &= p_0 + p_1(\tilde{x} + \dot{x}d) + \dots + p_n(\tilde{x} + \dot{x}d)^n \\&= p_0 + p_1\tilde{x} + p_2\tilde{x}^2 + \dots + p_n\tilde{x}^n + \\&\quad p_1\dot{x}d + 2p_2\tilde{x}\dot{x}d + \dots + np_{n-1}\tilde{x}^{n-1}\dot{x}d \\&= f(\tilde{x}) + f'(\tilde{x})\dot{x}d\end{aligned}$$

可以看出d的系数就是 $f'(\tilde{x})$ 。

不可导函数的求导

不可导函数的求导，一般采用泰勒展开的方式。典型的算法有PGD（Proximal Gradient Descent）。

参考：

<https://blog.csdn.net/bingecuilab/article/details/50628634>

Proximal Gradient Descent for L1 Regularization

参考

<https://mp.weixin.qq.com/s/7Z2tDhSle-9MOslYEUUpq6g>

从概念到实践，我们该如何构建自动微分库

https://mp.weixin.qq.com/s/bigKoR3IX_Jvo-re9UjqUA

机器学习之——自动求导

<https://www.jianshu.com/p/4c2032c685dc>

自动求导框架综述

https://mp.weixin.qq.com/s/xXwbV46-kTobAMRwfKyk_w

自动求导—Deep Learning框架必备技术二三事

<https://mp.weixin.qq.com/s/f0xFfA1inOVOdJnSZR4k6Q>

自动微分技术

<https://mp.weixin.qq.com/s/0tTIPG4hd9hcHORkZF6w1A>

PyTorch的自动求导机制详细解析，PyTorch的核心魔法

<https://mp.weixin.qq.com/s/PELBuCvu-7KQ33XBtIYfYQ>

深度学习中的微分

<https://zhuanlan.zhihu.com/p/24709748>

矩阵求导术（上）

<https://zhuanlan.zhihu.com/p/24863977>

矩阵求导术（下）



您的打赏，是对我的鼓励

想在此留下评论，请访问 <https://github.com/antkillerfarm/antkillerfarm.github.com/issues/1> 提交评论

Loading

Recent Posts

- 22 Oct 2019 » [深度强化学习（七）——DRL参考资料（1）](#)
- 21 Oct 2019 » [世说新语专辑（二）](#)
- 08 Oct 2019 » [深度强化学习（六）——AlphaStar, 元学习](#)
- 06 Oct 2019 » [数学狂想曲（十四）——数学的深渊, Wavelet & Chirplet](#)
- 05 Oct 2019 » [高考](#)

Antkillerfarm
antkillerfarm@sohu.com

github.com/antkillerfarm