

پروژه سوم – بررسی شبیه

شما بخشی از یک تیم هستید که برای در دسترس قرار دادن پرداخت‌های تلفن همراه در سطح جهانی کار می‌کنید و از شما خواسته می‌شود یک مدل یادگیری عمیق برای کشف تقلب پیاده سازی کنید - هر زمان که شخصی پرداختی را انجام می‌دهد، می‌خواهید ببینید که آیا پرداخت جعلی است یا نه.

اما پیاده سازی پس انتشار (Backpropagation) شبکه کاملاً چالش برانگیز است و گاهی اشکالاتی نیز دارد. از آنجا که ماموریت مهم و حساسی در دست دارید، مدیر عامل شرکت شما می‌خواهد در مورد صحیح بودن اجرای پس انتشار شما اطمینان حاصل کند. مدیر عامل شرکت شما می‌گوید: "به من اثبات کنید که پس انتشار شما به خوبی کار می‌کنند!" برای دادن این اطمینان، شما می‌خواهید از "بررسی شبیه" استفاده کنید.

مرحله صفر: اضافه کردن کتابخانه‌های لازم

محتویات پوشه Requirements را در مسیر پروژه اضافه کنید.

مرحله اول: بررسی شبیه چگونه کار می‌کند؟

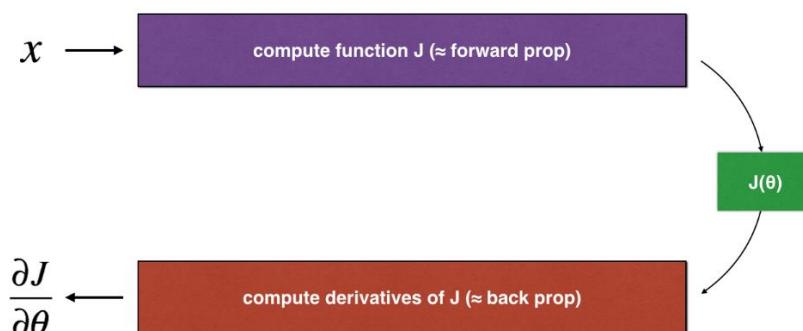
پس انتشار، شبیه‌ها را محاسبه می‌کند، جایی که θ پارامترهای مدل را نشان می‌دهد و J تابع خطای شما را نشان میدهد که در مرحله انتشار رو به جلو محاسبه می‌شود.

از آنجا که پیاده سازی انتشار رو به جلو به آسانی انجام می‌شود، اطمینان دارید که این کار را درست انجام داده اید و بنابراین تقریباً ۱۰۰٪ مطمئن هستید که هزینه J را به درستی محاسبه می‌کنید. در نتیجه، میتوانید از محاسبه خود برای J کمک بگیرید و

$$\text{محاسبه شبیه } \frac{dJ}{d\theta} \text{ را تایید کنید.}$$

$$\frac{\partial J}{\partial \theta} = \lim_{\epsilon \rightarrow 0} \frac{J(\theta + \epsilon) - J(\theta - \epsilon)}{2\epsilon}$$

تابع خطی یک بعدی $J(\theta) = \theta x$ را در نظر بگیرید، این مدل فقط یک مقدار واقعی (θ) دارد و مقدار x را به عنوان ورودی می‌گیرد.



نمودار قبل مراحل محاسبه را نشان میدهد. در مرحله انتشار رو به جلو ورودی x دریافت شده و مقدار J محاسبه می‌شود. در مرحله پس انتشار مشتق $\frac{dJ}{d\theta}$ محاسبه می‌شود.

تمرین اول: توابع انتشار روبه جلو و پس انتشار را برای تابع یک بعدی گفته شده پیاده سازی کنید.

مرحله دوم: بررسی شبیه یک بعدی

برای اینکه بررسی کنیم محاسبه پس انتشار به درستی کار می کند از بررسی شبیه استفاده می کنیم. پارامترهای زیر را در نظر بگیرید:

1. $\theta^+ = \theta + \epsilon$
2. $\theta^- = \theta - \epsilon$
3. $J^+ = J(\theta^+)$
4. $J^- = J(\theta^-)$
5. $gradapprox = \frac{J^+ - J^-}{2\epsilon}$

تخمین مشتق از روش زیر به دست می آید:

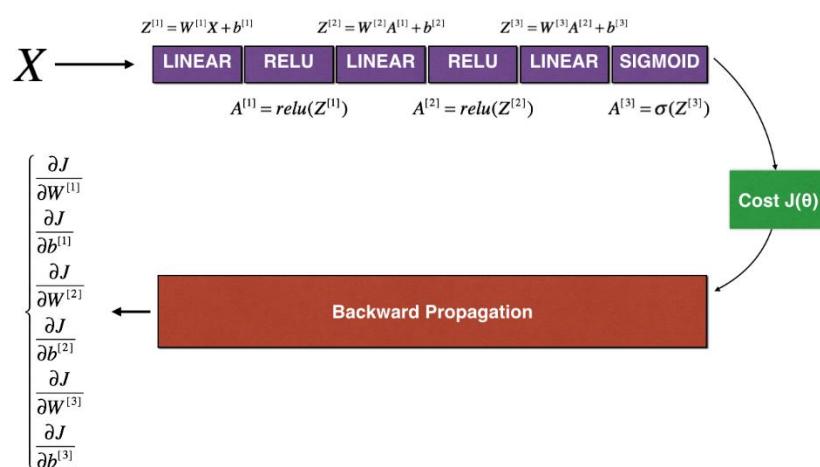
$$difference = \frac{\| grad - gradapprox \|_2}{\| grad \|_2 + \| gradapprox \|_2}$$

کاری که در روش بررسی شبیه انجام میدهیم مقایسه خروجی پس انتشار با تخمین مشتق است. حد آستانه اختلاف آنها را برابر با 10^{-7} در نظر میگیریم.

تمرین دوم: تابع بررسی شبیه را پیاده سازی کنید. (برای محاسبه $\|x\|_2$ از تابع $np.linalg.norm(x)$ کمک بگیرید.)

مرحله سوم: بررسی شبیه چند بعدی

مراحل قبل را به حالت چند بعدی تعمیم میدهیم:



تمرین سوم: تابع بررسی شبیه را برای حالت چند بعدی مشابه قبل پیاده سازی کنید.