دانشگاه صنعتی اصفهان

دانشکده مهندسی برق و کامپیوتر

عنوان: تکلیف اول درس یادگیری عمیق

نام و نام خانوادگی: علیرضا ابره فروش

شماره دانشجویی: ۹۸۱۶۶۰۳

نیم سال تحصیلی: پاییز ۱۴۰۲

مدرّس: دکتر سمانه حسینی سمنانی

دستیاران آموزشی: مریم محمدی-علی بزرگ زادارباب

۱

۲

۳

۱.۳ الف

تابع تانژانت هایپربولیک ($tanh$) اغلب به عنوان نسخه مقیاس‌شده‌ای از تابع سیگموید توصیف می‌شود، به خصوص تابع سیگموید لجستیک. این رابطه به دلیل شباهت‌های تابع تانژانت و تابع سیگموید وجود دارد، اما در بازه و مقیاس‌شان تفاوت دارند.
تابع سیگموید که اغلب با نماد $\sigma(x)$ نشان داده می‌شود، به شرح زیر تعریف می‌شود:

$\sigma(x) = \frac{1}{1+e^{-x}}$

این تابع هر عدد حقیقی را به یک مقدار بین ۰ و ۱ نگاشت می‌کند. وقتی x یک عدد مثبت بزرگ است، $\sigma(x)$ به ۱ نزدیک می‌شود و وقتی x یک عدد منفی بزرگ است، $\sigma(x)$ به ۰ نزدیک می‌شود. این به این معناست که تابع سیگموید ورودی خود را در بازه (۰، ۱) فشرده می‌کند که برای مسائل دسته‌بندی دودویی مفید است، چون می‌توان از آن تعبیر احتمالاتی کرد.
تابع تانژانت هایپربولیک، به شکل زیر تعریف می‌شود:

$tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x}-1}{e^{2x}+1}$

تابع تانژانت هایپربولیک هر عدد حقیقی را به یک مقدار بین ۱- و ۱ نگاشت می‌دهد. وقتی x یک عدد مثبت بزرگ است، $tanh(x)$ به ۱ نزدیک می‌شود و وقتی x یک عدد منفی بزرگ است، $tanh(x)$ به ۱- نزدیک می‌شود.

رابطه بین توابع تانژانت هایپربولیک و سیگموید به شرح زیر است:

۱. مقیاس‌دهی: تابع تانژانت هایپربولیک، انتقال داده شده و مقیاس شده‌ی تابع سیگموید به منظور داشتن رنج $(-1, 1)$ به جای $(0, 1)$ است. این مقیاس‌دهی با کم کردن 0.5 از تابع سیگموید و سپس ضرب آن در ۲ انجام می‌شود:

$2\sigma(2x) - 1 = 2 \times \frac{1}{1+e^{-2x}} - 1 = \frac{1-e^{-2x}}{1+e^{-2x}} = \frac{1-e^{-2x}}{1+e^{-2x}} \times \frac{e^{2x}}{e^{2x}} = \frac{e^{2x}-1}{e^{2x}+1} = tanh(x)$
$\Rightarrow tanh(x) = 2\sigma(2x) - 1$

۲. تقارن: یکی از تفاوت‌های کلیدی این است که تابع تانژانت هایپربولیک در اطراف مبدأ متقارن است (در واقع $tanh(-x) = -tanh(x)$)، درحالی که تابع سیگموید این تقارن را ندارد.

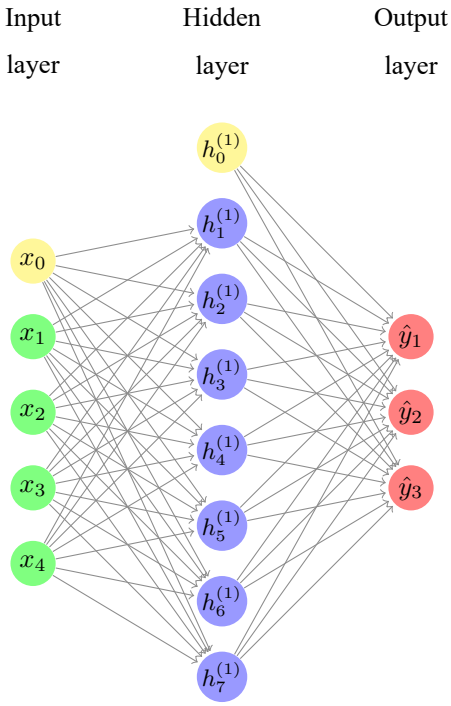
تابع تانژانت هایپربولیک اغلب به عنوان یک جایگزین برای تابع سیگموید در شبکه‌های عصبی استفاده می‌شود. زیرا به دلیل ماهیت مرکز شده حول صفری که دارد، یادگیری در شبکه‌های عمیق را سریع‌تر می‌کند. اینکه تابع تانژانت هایپربولیک مقادیر منفی را نیز خروجی دهد به معنای این است که می‌تواند تغییرات مثبت و منفی را در واحدهای مخفی شبکه در طول آموزش ایجاد کند که می‌تواند به همگرایی کمک کند. با این حال، هر دو تابع هنوز در متنوعی از زمینه‌ها استفاده می‌شوند و انتخاب بین آن‌ها بستگی به مسئله خاص و معماری شبکه دارد.

۲.۳ ب

$p(x) = x \log(1 + tanh(e^x))$

$$\frac{d}{dx}p\left(x\right) = \left(\frac{d}{dx}x\right) \times \log\left(1 + tanh\left(e^x\right)\right) + \left(\frac{d}{dx}\left(\log\left(1 + tanh\left(e^x\right)\right)\right)\right) \times x$$
$$= \log\left(1 + tanh\left(e^x\right)\right) + x\left(\frac{d}{dx}\left(1 + tanh\left(e^x\right)\right)\right) \times \frac{1}{1+tanh(e^x)}$$
$$= \log\left(1 + tanh\left(e^x\right)\right) + \frac{x}{1+tanh(e^x)} \times \left(\frac{d}{dx}\left(tanh\left(e^x\right)\right)\right)$$
$$= \log\left(1 + tanh\left(e^x\right)\right) + \frac{x}{1+tanh(e^x)} \times \left(\frac{d}{dx}\left(e^x\right)\right) \times \left(1 - tanh^2\left(e^x\right)\right)$$
$$= \log\left(1 + tanh\left(e^x\right)\right) + xe^x\left(1 - tanh\left(e^x\right)\right)$$

۴



با توجه به شبکه‌ی بالا، نورون‌های سبز، بنفش، و قرمز به ترتیب لایه‌ی ورودی، لایه‌ی مخفی، و لایه‌ی خروجی را تشکیل می‌دهند و همچنین نورون‌های زرد biasها هستند که همگی مقدار ۱ دارند. پارامترهای قابل یادگیری شبکه وزن‌های موجود بین نورون‌هاست که تعدادشان برابر است با: $4 \times 7 + 7 + 7 \times 3 + 3 = 59$

۵

۶

منابع

[1] Stošić, Lazar, and Milena Bogdanović. ”RC4 stream cipher and possible attacks on WEP.” Editorial Preface 3.3 (2012).