Converting Sigmoids to CDFs

Wayne Hayes

April 2022

Assume we have an analytical sigmoid $s(x)$ whose range is (min, max)—that is,

$$\lim_{x\to-\infty} s(x) = min, \quad \text{and} \quad \lim_{x\to+\infty} s(x) = max.$$

We want to *rescale* $s(x)$ in the *vertical* direction to something like a CDF so that no output value lies outside the interval [0,1]. To do this re-scaling of the vertical direction, we will define

$$\sigma(x) \equiv \frac{s(x) - min}{max - min}.$$

So for example, if $s(x) = \arctan(x)$, then we know that $s(x) \in (-1, 1)$, so $min = -1$ and $max = 1$. So for any x, the numerator is $\arctan(x) - (-1) = \arctan(x) + 1$, which always lies in the interval [0,2], and the denominator is $1 - (-1) = 1 + 1 = 2$. Thus $\sigma(x)$ always lies in [0,1].

Furthermore, assume that the sigmoid is "halfway up" at $x = M$, $s(M)$ is the average of min and max: $s(M) = \frac{max+min}{2}$. In the case of arctan, the "halfway" point occurs at $x = M = 0$ where $\arctan(0) = 0$, which is the average of min and max, because $(1 + (-1))/2 = 0$. Technically this would be the *median* value of x, because half the distribution of x lies below $x = M$ and half above. This is where, ideally, we'd like our CDF to have a value of 0.5 (halfay between 0 and 1), because by definition the median of a distribution occurs at the x value where $\text{CDF}(x) = 0.5$.

Given the $\sigma(x)$ above, we now need to scale it *horizontally* to a new function $\hat{F}(x)$ that fits our "training data". Thus, assume we have data points (aka "training points") $x_1 \leq x_2 \leq x_3 \leq \ldots \leq x_{n-1} \leq x_n$, sorted smallest to largest. In our example, x_i is the i^{th} nearest-neighbor distance, where x_1 is the smallest observed value and x_n is the largest.

We want our CDF $\hat{F}(x)$ to be "fit" to the training data so that the following properties hold:

1. $\hat{F}(x) = 0$ if $x \leq 0$. (Note: this *does* mean that $\hat{F}(x) = 0$.)
2. $\hat{F}(x_1) \approx \frac{1}{n+1}$.
3. $\hat{F}(x_{n/2}) \approx 0.5$, where $x_{n/2}$ is the value in the middle of the list of x_i (ie., the *median*, **not** the average!)
4. $\hat{F}(x_n) \approx \frac{n}{n+1}$.
5. $\hat{F}(x) < 1$ for all x and $\hat{F}(x) \to 1$ as $x \to \infty$.

Although we have been curve-fitting to the entire list of $\{x_i\}_{i=1}^n$, you may also want to try performing a least-squares fit of the above $\sigma(x)$ *only* to the middle three criteria: $\hat{F}(x_1) = 1/(n+1)$, $\hat{F}(x_{n/2}) = 0.5$, $\hat{F}(x_n) = n/(n+1)$. It would be interesting to see how we do fitting only those 3 points.

PS: I *think* we want to fit $\hat{F}(x) = \sigma(\beta x - \alpha)$, finding the best α and β to make $\hat{F}(x)$ fit the above criteria.