

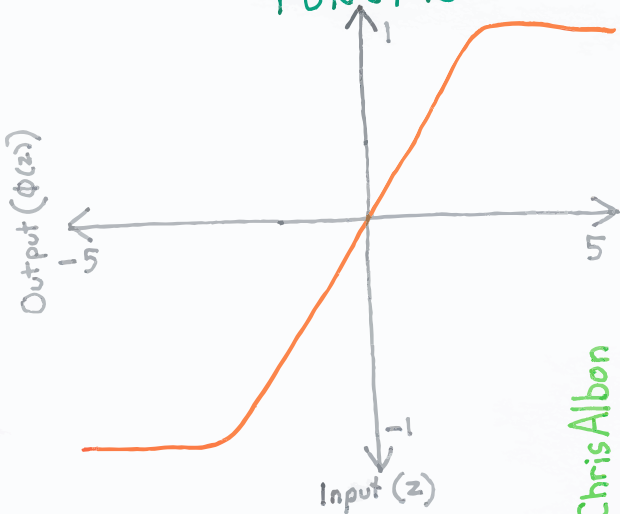
TANH

HYPERBOLIC
TANGENT
ACTIVATION
FUNCTION

Often used instead of
the sigmoid activation.

$$\phi(z) = \frac{\sinh(z)}{\cosh(z)}$$

Because tanh ranges
-1 and 1, it is less likely
to have gradients that are very
small compared to sigmoid.



ChrisAlbon