

# Implementation Tricks

---

## Numerical Stability

---

由于浮点数精度有限，在计算 softmax 时，由于采用  $\exp$ , 很可能由于数值过大导致 NaN

$$z_i = \frac{\exp(x_i)}{\sum_k \exp(x_k)} \quad (1)$$

而对于任意常数  $c$  有

$$z_i = \frac{\exp(x_i)}{\sum_k \exp(x_k)} = \frac{\exp(x_i - c)}{\sum_k \exp(x_k - c)} \quad (2)$$

所以可以取  $c = \max(x)$  以保证  $\exp$  的输入小于等于 0

这个技巧对 logsoftmax, logsumexp 等也有效