# Notes on log sum exp tricks

Robert DiPietro
Department of Computer Science
Johns Hopkins University
rdipietro@jhu.edu

October 8, 2014

## Obtaining the logarithm of a sum of exponentials

Computing

$$f(\mathbf{x}) = \log\Big(\sum_i \exp(x_i)\Big)$$

will fail numerically if the $x_i$ are very small, because all of the terms in the sum will be very close to 0 (or equal to 0). We can fix this by shifting all of the $x_i$ by some constant $s$:

$$\begin{aligned} f(\mathbf{x}) &= \log\Big(\sum_i \exp(x_i + s - s)\Big) \\ &= \log\Big(\exp(s)\sum_i \exp(x_i - s)\Big) \\ &= s + \log\Big(\sum_i \exp(x_i - s)\Big) \end{aligned}$$

A common choice is to let $s = \max_i x_i$; then the largest $x_i$ maps to 0 in log space, or 1 after we exponentiate. Some small $x_i$ may still disappear, but this is okay because their contributions to the sum are negligible compared to the largest $x_i$'s contribution.

Example: consider an array of length 1, $x_1 = -1000$. If we exponentiate this directly using 64-bit precision, we get 0, and therefore $\log \exp x_1 = -\infty$. Instead, if we use the above result with $s = -1000$, we obtain $\log \exp x_1 = -1000 + \log \exp 0 = -1000$.

## Obtaining the logarithm of a product of two matrices

We can also extend this method to matrix multiplication; the goal is now to obtain

$$\mathbf{F}(\mathbf{X}, \mathbf{Y}) = \log\Big(\exp(\mathbf{X})\exp(\mathbf{Y})\Big)$$

where the exponentiation is carried out element wise. We can rewrite this as

$$\begin{aligned}\mathbf{F}(\mathbf{X}, \mathbf{Y}) &= \log\Big( \exp(\mathbf{X} + s_X - s_X) \exp(\mathbf{Y} + s_Y - s_Y)\Big) \\ &= \log\Big( \exp(s_X) \exp(\mathbf{X} - s_X) \exp(s_Y) \exp(\mathbf{Y} - s_Y)\Big) \\ &= s_X + s_Y + \log\Big( \exp(\mathbf{X} - s_X) \exp(\mathbf{Y} - s_Y)\Big)\end{aligned}$$

(Here, adding a scalar to a matrix is carried out by broadcasting the scalar to every element of the matrix.)