Background for Lesson 4

1 Products and Exponents

In Lesson 1, we introduced the summation notation $\sum_{i=1}^{n} x_i = x_1 + x_2 + \ldots + x_n$. Similarly, we can define product notation as

$$\prod_{i=1}^{n} x_i = x_1 \cdot x_2 \cdot \ldots \cdot x_n \,.$$

Example: We can re-write the factorial function as $n! = \prod_{i=1}^{n} i$ for $n \geq 1$.

Example: Suppose f is a function which returns $f(x) = 3x + 1$. Suppose x can take on only the discrete values $x \in \{-1, 2, 4\}$. Then

$$\begin{aligned}\prod_{x} f(x) &= (3 \cdot (-1) + 1) \cdot (3 \cdot (2) + 1) \cdot (3 \cdot (4) + 1) \\ &= (-2) \cdot (7) \cdot (13) = -182\,.\end{aligned}$$

Exponents are of the form a^x where a (called the base) and x (called the exponent) are any real numbers. Recall that $a^0 = 1$. Exponents have the following useful properties

1. $a^x \cdot a^y = a^{x+y}$
2. $(a^x)^y = a^{x \cdot y}$

Note that the first property requires that both terms have the same base a. Thus we cannot simplify $a^x \cdot b^y$ if $a \neq b$.

One common base is the number e which is approximately equal to 2.7183. The function e^x is so common in mathematics that it has its own symbol $e^x = \exp(x)$. Because $e > 0$, we have $e^x > 0$ for all real numbers x, although $\lim_{x \to \infty} e^{-x} = 0$.

Example: Using Property 1 above, we have $\prod_{x=1}^{5} e^x = \exp(\sum_{x=1}^{5} x) = e^{15}$.

2 Natural Logarithm

Logarithms can be defined as the inverse of exponential functions. That is, if $y = a^x$ then $\log_a(y) = x$. The natural logarithm function has base e and is written without the subscript $\log_e(y) = \log(y)$. Because $e^x > 0$ for all x, $\log(y)$ is only defined for $y > 0$. We always have $\exp(\log(y)) = \log(\exp(y)) = y$.

We can use the properties of exponents from the previous section to obtain some important properties of logarithms:

1. $\log(x \cdot y) = \log(x) + \log(y)$
2. $\log(\frac{x}{y}) = \log(x) - \log(y)$
3. $\log(x^b) = b\log(x)$
4. $\log(1) = 0$

Because the natural logarithm is a monotonically increasing one-to-one function, finding the x which maximizes any (positive-valued function) $f(x)$ is equivalent to maximizing $\log(f(x))$. This is useful because we often take derivatives to maximize functions. If $f(x)$ has product terms, then $\log(f(x))$ will have summation terms, which are usually simpler when taking derivatives.

Example: $\log\left(\frac{5^2}{10}\right) = 2\log(5) - \log(10) \approx 0.916$.

3 Argmax

When we want to maximize a function $f(x)$, there are two things we may be interested in:

1. The value $f(x)$ achieves when it is maximized, which we denote $\max_x f(x)$.
2. The x-value that results in maximizing $f(x)$, which we denote $\hat{x} = \arg\max_x f(x)$.

Thus $\max_x f(x) = f(\hat{x})$.

Example: Suppose $f(x) = \exp(-x^2)$. Then $\log(f(x)) = -x^2$ which is maximized at $x = 0$. Hence, $\arg\max_x f(x) = \hat{x} = 0$ and $\max_x f(x) = f(\hat{x}) = f(0) = 1$.