Argmax and Max Calculus

Mark Schmidt

February 19, 2016

1 Argmax, Max, and Supremum

We define the *argmax* of a function f defined on a set D as

$$\operatorname*{argmax}_{x \in D} f(x) = \{x | f(x) \geq f(y), \forall y \in D\}.$$

In other words, it is the set of inputs x from the domain D that achieve the highest function value. For example, $\operatorname{argmax}_{x \in \mathbb{R}} -x^2 = \{0\}$, since $-x^2$ is maximized when $x = 0$. Note that the output of the argmax function is a set, since more than one value might achieve the maximum. The result of the argmax function might even be an infinite set. For example, the argmax of $cos(x)$ is the set containing all integer multiples of 2π.

We can define the *max* of a function f defined on a set D as

$$\max_{x \in D} f(x) = f(x^*), \quad \text{for any } x^* \in \operatorname*{argmax}_{x \in D} f(x).$$

The max function gives the largest possible value of $f(x)$ for any x in the domain, and it is thus the function value achieved by any element of the argmax. Unlike the argmax, the max function is unique since all elements of the argmax achieve the same value. However, the max may not exist because the argmax may not exist. For example, if $f(x) = x$ then f just keeps growing as we increase x (there is no largest value). A more subtle issue happens if we consider $f(x) = -e^x$. In this case, the function is never larger than 0 but the maximum doesn't exist since we can always move it closer to 0 by decreasing x.

To handle functions like $f(x) = -e^x$, we define the *sup* function ('supremum') as the smallest value of the set $\{y | y \geq f(x), \forall x \in D\}$. That is, it's the smallest value that is greater than or equal to $f(x)$ for any x in D. Often the sup is equal to the max, but the sup is sometimes defined even when the max is not defined. For example, $\sup_{x \in \mathbb{R}} -x^2 = \max_{x \in \mathbb{R}} -x^2 = 0$ but $\sup_{x \in \mathbb{R}} -e^x = 0$ even though the max is not defined.

If we are talking about smallest elements instead of largest elements, we replace argmax/max/sup by argmin/min/inf. We can relate minimization to maximization (or even define it), by maximizing the negation of f,

$$\begin{aligned}
\operatorname*{argmin}_{x \in D} f(x) &= \operatorname*{argmax}_{x \in D} -f(x), \\
\min_{x \in D} f(x) &= -\max_{x \in D} -f(x), \\
\inf_{x \in D} f(x) &= -\sup_{x \in D} -f(x).
\end{aligned}$$

When the variable/domain are obvious from context, we sometimes use a simpler notation for the max and argmax,

$$\begin{aligned}
\operatorname{argmax} f(x) &= \{x | f(x) \geq f(y), \forall y \in D\} \\
\max f(x) &= f(x^*), \quad \text{for any } x^* \in \operatorname*{argmax}_{x \in D} f(x) \\
\sup f(x) &= \min_{y | y \geq f(x), \forall x \in D} y.
\end{aligned}$$

2 Operations that Preserve Argmax

There are a variety of operations that do not change the argmax of a function. Since the argmax is defined by an inequality, these are basically operations that preserve inequalities. Below are some examples.

1. If θ is a constant then we have
$$\operatorname{argmax} f(x) = \operatorname{argmax} f(x) + \theta.$$

2. If $\theta > 0$ we have
$$\operatorname{argmax} f(x) = \operatorname{argmax} \theta f(x).$$

3. If $\theta < 0$ we have
$$\operatorname{argmax} f(x) = \operatorname{argmin} \theta f(x).$$

4. If $\operatorname{argmax} f(x) > 0$, then
$$\operatorname{argmax} f(x) = \operatorname{argmin} \frac{1}{f(x)}.$$

5. If g is strictly monotonic, meaning that $\alpha > \beta$ implies $g(\alpha) > g(\beta)$, then
$$\operatorname{argmax} g(f(x)) = \operatorname{argmax} f(x).$$

The last property is used frequently when doing calculations involving probabilities, since the logarithm (a strictly monotonic function) transforms multiplication of probabilities into addition of log-probabilities,

$$\operatorname{argmax} \prod_{i=1}^{n} p_i(x) = \operatorname{argmax} \sum_{i=1}^{n} \log p_i(x).$$

3 Max Identities

We also have a set of related identities for the max function.

1. If θ is a constant we have
$$\max\{f(x) + \theta\} = \theta + \max f(x).$$

2. If $\theta > 0$ we have
$$\max \theta f(x) = \theta \max f(x).$$

3. If $\theta < 0$ we have
$$\max \theta f(x) = \theta \min f(x).$$

4. If $\operatorname{argmax} f(x) > 0$, then
$$\max \frac{1}{f(x)} = \frac{1}{\min f(x)}.$$

5. If g is strictly monotonic then
$$\max g(f(x)) = g(\max f(x)).$$

6. If have multiple variables x and y, then we have
$$\max_{x,y} f(x, y) = \max_{x}\{\max_{y} f(x, y)\}.$$

As before, we assume that the max functions exist. Similar identities hold for the sup funciton.

4 Inequalities

A variety of inequalities follow from the definition of the max function.

1. If $y \in D$, then
$$f(y) \leq \max f(x),$$
and similarly for any $x^* \in \operatorname{argmax}_f(x)$ we have
$$f(y) \leq f(x^*).$$
A generalization of this is that for a subset $D' \subseteq D$ we have
$$\max_{x \in D'} f(x) \leq \max_{x \in D} f(x).$$

2. We also have a triangle-like inequality
$$\max\{f(x) + g(x)\} \leq \max\{f(x)\} + \max\{g(x)\}.$$

3. A looser version of the triangle inequality is
$$\max\{f(x) + g(x)\} \leq 2 \max\{\max f(x), \max g(x)\}.$$
If you have n functions instead of 2, you replace the 2 by n and take the outer max over the maxima of all functions. A notable variant of this is that if we have n numbers x_i (for $i = 1, 2, \ldots, n$), then we have
$$\sum_{i=1}^{n} x_i \leq n \max_i x_i,$$
or written another way, that the maximum is larger than the average,
$$\max_i x_i \geq \frac{1}{n} \sum_{i=1}^{n} x_i.$$

4. We also have a type of reverse triangle inequality,
$$|\max\{f(x)\} - \max\{g(x)\}| \leq \max\{|f(x) - g(x)|\}.$$

5. Inequalities are preserved under maximization, so if $f(x) \leq g(x)$ then
$$\max\{f(x)\} \leq \max\{g(x)\}.$$

6. We have a special case of the Cauchy-Schwartz inequality,
$$\max\{f(x)g(x)\} \leq \max\{|f(x)|\} \max\{|g(x)|\},$$

7. If $f(x) > 0$ for all x and $g(x) > 0$ for all x then we can also do a reciprocal version,
$$\max\left\{\frac{f(x)}{g(x)}\right\} \leq \frac{\max f(x)}{\min g(x)}.$$

8. A more subtle inequality is that if we have n positive numbers x_i and n numbers y_i then
$$\max \frac{x_i}{y_i} \geq \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i}.$$