Idiosyncrasies

Vladimir Feinberg

July 1, 2017

1 Notation

Landau notation is used in the standard sense for single variables: $f = O(g)$ if there exists a c such that $f(x) \leq cg(x)$ for all x sufficiently large, with x perhaps restricted to the support of g as required by the context. We use the analogous sense for $\Omega, \Theta, o, \omega$.

In multiple variables, big-O and friends require $f(\mathbf{x}) \leq cg(\mathbf{x})$ and similar inequalities, respectively, for all $\mathbf{x}$ in the support for the current context with $\|\mathbf{x}\|_\infty$ sufficiently large. To capture the essence of asymptotics, a tilde will crudely absorb behavior up to logarithmic factors: $f = \tilde{O}(g)$ if $f = O(g \log^k g)$ for some $k \in \mathbb{N}$ and possibly multivariate function g.

When some iterative process is clear, such as $\mathbf{x}_t$ arising from iterative optimization of an objective f, then the gradient will be abbreviated $\nabla_t = \nabla f(\mathbf{x}_t)$. Similarly we use $H_t = \nabla^2 f(\mathbf{x}_t)$.

I'll use $x \simeq y$ to claim that equality holds up to some fixed multiplicative constants.

2 Abbreviations

- iid: independent and identically distributed
- rv: random variable
- wp: with probability
- wrt: with respect to
- wlog: without loss of generality