Lecture 29, Mar 25, 2022

Redefining the Derivative

- A partial derivative only gives the rate of change along one of the axes; to define differentiability for a multivariable function, we need to consider all directions
- If we try to simply extend the definition of a derivative as $\lim_{\vec{h} \to \vec{0}} \frac{f(\vec{x}_0 + \vec{h}) - f(\vec{x})}{\vec{h}}$ we run into the problem of dividing by a vector; if we change it to $\left\|\vec{h}\right\|$ instead, we lose information about the direction; thus we need to reinvent the derivative
- Definition: Little o notation: $g(h) = o(h)$ if $\lim_{h \to 0} \frac{g(h)}{|h|} = 0$, i.e. $g(h)$ goes to 0 faster than h goes to 0
  - Big O is used to indicate that two things are the same order of magnitude; little o is used for different orders of magnitude
- In one dimension,

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ \implies \lim_{h \to 0} \frac{f(x+h) - f(x) - f'(x)h}{h} &= 0 \\ \implies (f(x+h) - f(x)) - f'(x)h &= o(h) \end{aligned}$$

  - Example: $f(x) = x^2$
    * $f(x+h) - f(x) = (x+h)^2 - x^2 = (2x)h + h^2$
    * By our definition $2x$ is f' if h^2 is $o(h)$
    * $\lim_{h \to 0} \frac{h^2}{h} = \lim_{h \to 0} h = 0 \implies h^2 = o(h)$

Differentiability in Multiple Dimensions

- Extending our definition of little o to multiple variables: $\lim_{\vec{h} \to \vec{0}} \frac{g(\vec{h})}{\left\|\vec{h}\right\|} = 0 \implies g(\vec{h}) = o(\vec{h})$
- Definition: f is differentiable at $\vec{x} \iff \exists \vec{y} \ni f(\vec{x} + \vec{h}) - f(\vec{x}) = \vec{y} \cdot \vec{h} + o(\vec{h})$
  - $\vec{y} = \nabla f(\vec{x})$ is the *gradient* of $\vec{f}$
- Example: $f(x, y) = x + y^2$
  - Let $\vec{h} = (h_1, h_2)$
  - $$\begin{aligned} f(\vec{x} + \vec{h}) - f(\vec{x}) &= f(x + h_1, y + h_2) - f(x, y) \\ &= x + h_1 + (y + h_2)^2 - x - y^2 \\ &= h_1 + 2yh_2 + h_2^2 \\ &= (1\hat{\imath} + 2y\hat{\jmath}) \cdot \vec{h} + h_2^2 \end{aligned}$$
  - Let $g(\vec{h}) = h_2^2 = h_2\hat{\jmath} \cdot \vec{h} \implies \frac{|g(\vec{h})|}{\left\|\vec{h}\right\|} = \frac{\left\|h_2\hat{\jmath}\right\|\left\|\vec{h}\right\| \cos\theta}{\left\|\vec{h}\right\|} \le |h_2| \implies \lim_{\vec{h} \to \vec{0}} \frac{|g(\vec{h})|}{\left\|\vec{h}\right\|} = 0$
  - Since g is $o(\vec{h})$ the gradient is $1\hat{\imath} + 2y\hat{\jmath}$
- Theorem: $\nabla f(x, y, z) = \frac{\partial f}{\partial x}\hat{\imath} + \frac{\mathrm{d}f}{\mathrm{d}y}\hat{\jmath} + \frac{\mathrm{d}f}{\mathrm{d}z}\hat{k}$
  - Note $f(\vec{x})$ is a vector but $\nabla f(\vec{x})$ is a vector
  - If the gradient exists, then the function is differentiable at that point
- The gradient points in the direction of steepest ascent

Directional Derivatives

- The idea of a partial derivative can be extended beyond just the axes
- Definition: The directional derivative of f at $\vec{x}_0$ in the direction of $\hat{u}$ is $f_{\hat{u}}(\vec{x}_0) = \lim_{h \to 0} \frac{f(\vec{x}_0 + h\hat{u}) - f(\vec{x}_0)}{h}$

- Note $\hat{u}$ is a unit vector by definition

- The direction derivative is related to the gradient: $f_{\hat{u}}(\vec{x}_0) = \nabla f(\vec{x}_0) \cdot \hat{u}$
  - Proof: $f(\vec{x} + \vec{h}) - f(\vec{x}) = \Delta f(\vec{x})\vec{h} + o(\vec{h})$
    * Choose $\vec{h} = h\hat{u}$ where $h = \left\|\vec{h}\right\| \implies f(\vec{x} + h\hat{u}) - f(\vec{x}) = \nabla f(\vec{x}) \cdot h\hat{u} + o(\vec{h}) \implies \lim_{h \to 0} \frac{f(\vec{x} + h\hat{u}) - f(\vec{x})}{h} = \lim_{h \to 0} \nabla f(\vec{x}) \cdot \hat{u} + \frac{o(\vec{h})}{h} \implies f_{\hat{u}} = \nabla f(\vec{x}_0) \cdot \hat{u}$
- Example: Parabolic hill $z(x, y) = 20 - x^2 - y^2$
- Note: $|f_{\hat{u}}(\vec{x})| = |\nabla f \cdot \hat{u}| = \|\nabla f\|\|\hat{u}\||\cos\theta| \leq \|\nabla f\|$
  - The rate of change along any direction is always less than or equal to the rate of change along the gradient
  - Max rate of change happens for $\theta = 0$, i.e. $\hat{u}$ pointing in the direction of ∇f
  - This shows that the gradient points in the direction with the greatest rate of change
- Example: Project the path of steepest descent to the xy plane: $z = f(x, y) = A + x + 2y - x^2 - 3y^2$ from $(0, 0, A)$
  - $\nabla f = (1 - 2x)\hat{i} + (2 - 6y)\hat{j} \implies -\nabla f = (2x - 1)\hat{i} + (6y - 2)\hat{j}$
  - Consider the curve $\vec{r}(t) = x(t)\hat{i} + y(t)\hat{j}$; we always want it to point in the direction of steepest descent so $x'(t), y'(t)$ should be in the opposite direction of the gradient
  - $\begin{cases} x'(t) = 2x(t) - 1 \\ y'(t) = 6y(t) - 2 \end{cases} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{6y - 2}{2x - 1}$
  - This is a separable differential equation; solving gives $6y - 2 = (2x - 1)^3 e^C$; plugging in the initial point we get $e^C = 2$
  - $y = \frac{(2x - 1)^3}{3} = \frac{1}{3}$