Math 346 Lecture #1
6.1 The Directional Derivative

6.1.1 Tangent Vectors

Definition 6.1.1. For an open interval (a, b), a function $f : (a, b) \to \mathbb{R}$ is differentiable at $x \in (a, b)$ if the limit (of the rise over the run)

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$$

exists. When the limit exists we write $f'(x)$ for this limit. If f is differentiable at every point $x \in (a, b)$, we say f is differentiable on (a, b).

Remark 6.1.2. To see how that derivative $m = f'(x_0)$ defines a linear transformation $L(h) = mh$ that best approximates curve $y(h) = f(x_0 + h) - f(x_0)$ for h close to zero, we recast the limit for $f'(x_0)$ in terms of the standard norm-induced metric $d(x, y) = |x - y|$ on $\mathbb{R}$.

For every $\epsilon > 0$ there exists $\delta > 0$ such that for all $0 < |h| < \delta$ there holds

$$\left| \frac{f(x_0 + h) - f(x_0)}{h} - m \right| < \epsilon,$$

or equivalently

$$\left| \frac{f(x_0 + h) - f(x_0) - mh}{h} \right| < \epsilon.$$

This is precisely

$$\left| \frac{y(h) - L(h)}{h} \right| < \epsilon \text{ or } (m - \epsilon)h = L(h) - \epsilon h < y(h) < L(h) + \epsilon h = (m + \epsilon)h.$$

This says that the graph of $y(h)$ lies between the graphs of $L(h) - \epsilon h$ and $L(h) + \epsilon h$ over the interval $|h| < \delta$.

Definition 6.1.3. A curve $\gamma : (a, b) \to \mathbb{R}^n$ is differentiable at $t_0 \in (a, b)$ if

$$\lim_{h\to 0} \frac{\gamma(t_0 + h) - \gamma(t_0)}{h}$$

exists with respect to the standard norm-induced metrics on $\mathbb{R}$ and $\mathbb{R}^n$, i.e., there is $a \in \mathbb{R}^n$ such that for every $\epsilon > 0$ there exists $\delta > 0$ for which for all $0 < |h| < \delta$ there holds

$$\left\| \frac{\gamma(t_0 + h) - \gamma(t_0) - ah}{h} \right\|_2 < \epsilon.$$

If the limit exists, it is called the derivative of γ at t_0 and denoted by $\gamma'(t_0)$. If γ is differentiable at every $t \in (a, b)$, then we say that γ is differentiable on (a, b).

Remark 6.1.4. The derivative $a = \gamma'(t_0)$ defines a linear transformation $L : \mathbb{R} \to \mathbb{R}^n$ given by $L(h) = ah$ that best approximates $\gamma(t_0 + h) - \gamma(t_0)$ for $|h|$ small.

Proposition 6.1.5. A curve $\gamma : (a, b) \to \mathbb{R}^n$ represented in standard coordinates as $[\gamma_1(t), \ldots, \gamma_n(t)]^{\mathrm{T}}$ is differentiable at $t_0 \in (a, b)$ if and only if $\gamma_i : (a, b) \to \mathbb{R}$ is differentiable at t_0 for every $i = 1, \ldots, n$.

Proof. Since all norm-induced metrics on $\mathbb{R}^n$ are topologically equivalent, we can use any norm-induced metric on $\mathbb{R}^n$ to compute the limit. We will use the metric induced by the ∞-norm.

Suppose the derivative $\gamma_i'(t_0)$ exists for all $i = 1, \ldots, n$.

Then for $\epsilon > 0$ there exists $\delta_i > 0$ such that for all $0 < |h| < \delta_i$ there holds

$$\left| \frac{\gamma_i(t_0 + h) - \gamma_i(t_0)}{h} - \gamma_i'(t_0) \right| < \epsilon.$$

Set $\delta = \min\{\delta_1, \ldots, \delta_n\}$.

Then for all $0 < |h| < \delta$ there holds

$$\left\| \frac{\gamma(t_0 + h) - \gamma(t_0)}{h} - [\gamma_1'(t_0), \ldots, \gamma_n'(t_0)]^{\mathrm{T}} \right\|_\infty = \max_{i=1,\ldots,n} \left| \frac{\gamma_i(t_0 + h) - \gamma_i(t_0)}{h} - \gamma_i'(t_0) \right| < \epsilon.$$

This implies that γ is differentiable at t_0.

Now suppose that γ is differentiable at t_0, with derivative $\gamma'(t_0) = [y_1, \ldots, y_n]^{\mathrm{T}}$.

Then for $\epsilon > 0$ there exists $\delta > 0$ such that for all $0 < |h| < \delta$ there holds

$$\left| \frac{\gamma_i(t_0 + h) - \gamma_i(t_0)}{h} - y_i \right| \leq \left\| \frac{\gamma(t_0 + h) - \gamma(t_0)}{h} - [y_1, \ldots, y_n]^{\mathrm{T}} \right\|_\infty < \epsilon.$$

This implies that γ_i is differentiable at t_0. □

Application 6.1.6. A twice-differentiable curve $\gamma : (a, b) \to \mathbb{R}^n$ can represent the position of a particle as a function of time.

The derivative $\gamma'(t)$ is the instantaneous velocity (or simply the velocity), and its norm $\|\gamma'(t)\|_2$ is the speed.

The second derivative $\gamma''(t)$ is the acceleration.

Often the motion of the particle is governed by a second-order differential equation,

$$\gamma''(t) = F(t, \gamma(t), \gamma'(t))$$

for a function $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$.

Definition 6.1.7. For a differentiable curve $\gamma : (a, b) \to \mathbb{R}^n$, the tangent vector of γ at $t \in (a, b)$ is the derivative $\gamma'(t)$.

Example 6.1.8. A particle moving according to $\gamma(t) = [\cos t \ \ \sin t]^{\mathrm{T}}$ traces out the circle of radius 1 centered at the origin.

The velocity $\gamma'(t) = [-\sin t, \cos t]^{\mathrm{T}}$ is orthogonal to $\gamma(t)$, and the acceleration $\gamma''(t) = [-\cos t, -\sin t]^{\mathrm{T}}$ satisfies the differential equation $\gamma''(t) = -\gamma(t)$.

Proposition 6.1.9. If $f, g : \mathbb{R} \to \mathbb{R}^n$ and $\varphi : \mathbb{R} \to \mathbb{R}$ are differentiable, and $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^n$, then

(i) $(f+g)' = f' + g'$ (sum rule),
(ii) $(\varphi f)' = \varphi' f + \varphi f'$ (product rule),
(iii) $\langle f, g\rangle' = \langle f', g\rangle + \langle f, g'\rangle$, and
(iv) $(f \circ \varphi)'(t) = \varphi'(t) f'(\varphi(t))$ (chain rule).

The proof of this is HW (Exercise 6.2).

6.1.2 Directional Derivatives

The directional, or Gâteaux, derivative is a generalization of the scalar-variable derivative to multivariable functions.

It is obtained a function $f : \mathbb{R}^n \to \mathbb{R}^m$ by composing f, for a point $x \in \mathbb{R}^n$ and a vector $\mathrm{v} \in \mathbb{R}^n$, with a curve $\gamma(t) = x + t\mathrm{v}$ in $\mathbb{R}^n$, i.e., $(f \circ \gamma)(t) = f(x + t\mathrm{v})$, which gives a curve in $\mathbb{R}^m$, and then taking the derivative with respect to t and evaluating it at $t = 0$.

Definition 6.1.10. Let $f : \mathbb{R}^n \to \mathbb{R}^m$. The directional derivative of f at $x \in \mathbb{R}^n$ in the direction $\mathrm{v} \in \mathbb{R}^n$ is the limit

$$\lim_{t \to 0} \frac{f(x + t\mathrm{v}) - f(x)}{t}$$

if it exists. The limit, denoted by $D_{\mathrm{v}} f(x)$, assesses the change in the value of f in the direction v from x.

Note. In multivariable calculus the vector v is always taken to be a unit vector when computing the directional derivative. We will not assume this here.

Remark 6.1.11. We show in the next section for fixed x that $\mathrm{v} \to D_{\mathrm{v}} f(x)$ is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. For now we illustrate this by way of example.

Example (in lieu of 6.1.12). Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x, y) = x^2 y^3$. Then for $\mathrm{v} = [\mathrm{v}_1, \mathrm{v}_2]^{\mathrm{T}}$, we have

$$\begin{aligned}
D_{\mathrm{v}} f(x, y) &= \frac{d}{dt} f(x + t\mathrm{v}_1, y + t\mathrm{v}_2)\Big|_{t=0} \\
&= \frac{d}{dt}\big((x + t\mathrm{v}_1)^2 (y + t\mathrm{v}_2)^3\big)\Big|_{t=0} \\
&= \Big(2(x + t\mathrm{v}_1)(\mathrm{v}_1)(y + t\mathrm{v}_2)^3 + (x + t\mathrm{v}_1)^2 3(y + t\mathrm{v}_2)^2(\mathrm{v}_2)\Big)\Big|_{t=0} \\
&= 2xy^3\mathrm{v}_1 + 3x^2y^2\mathrm{v}_2.
\end{aligned}$$

We recognize this as the inner product of the vectors $[2xy^3, 3x^2y^2]^{\mathrm{T}}$ and $\mathrm{v} = [\mathrm{v}_1, \mathrm{v}_2]^{\mathrm{T}}$, and so $D_{\mathrm{v}} f(x, y)$ is indeed linear in v.

6.1.3 Partial Derivatives

Partial derivatives of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ are the directional derivatives of f along the standard basis vectors $e_1, \ldots, e_n$ of $\mathbb{R}^n$.

Definition 6.1.13. Let $f : \mathbb{R}^n \to \mathbb{R}^m$. The i^{th} partial derivative of f at a point $x \in \mathbb{R}^n$ is the limit

$$\lim_{h \to 0} \frac{f(x + he_i) - f(x)}{h}$$

if it exists, and is denoted by $D_i f(x)$.

Example 6.1.14. Unfortunately, the existence of all of the partial derivatives of a function at a point does not imply the continuity of the function at that point. For

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0), \end{cases}$$

we have

$$D_1 f(0,0) = \lim_{h\to 0} \frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0} \frac{0}{h} = 0,$$
$$D_2 f(0,0) = \lim_{h\to 0} \frac{f(0,h)-f(0,0)}{h} = \lim_{h\to 0} \frac{0}{h} = 0.$$

But along the sequence $(1/n, 2/n)$ which approaches the origin as $n \to \infty$ we have

$$f(1/n, 2/n) = \frac{2/n^2}{1/n^2 + 4/n^2} = \frac{2}{5}$$

while $f(0,0) = 0$, so that f is not continuous at $(0,0)$.

Remark 6.1.15. In the i^{th} partial derivative $D_i f(x)$, it is only the i^{th} coordinate that is changing while the other coordinates remain fixed. We may thus use all of the differentiation rules for single-variable functions when computing $D_i f(x)$ as long as the rules apply.