The Derivative

Eric Kostelich

Arizona State University
Dept. of Mathematics and Statistics

Jan. 30, 2009

Reading for this week

- Sections 2.2–2.4

Differentiability in $\mathbb{R}^n$

Definition

The function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\mathbf{x}_0 \in \mathbb{R}^n$ if $\nabla f(\mathbf{x}_0)$ exists and

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{[f(\mathbf{x}) - f(\mathbf{x}_0)] - \nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0)}{\|\mathbf{x}-\mathbf{x}_0\|} = 0.$$

- Informally speaking, the error in the approximation $f(\mathbf{x}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0)$ must go to zero faster than $\|\mathbf{x}-\mathbf{x}_0\|$ as $\mathbf{x} \to \mathbf{x}_0$
- Compact notation parallels familiar 1-d formulas

Differentiability implies continuity

- The one-variable case: $f : \mathbb{R} \to \mathbb{R}$
- If $f'(a)$ exists, then given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\left| \frac{f(a+h) - f(a)}{h} - f'(a) \right| < \varepsilon$$

whenever $|h| < \delta$. Therefore,

$$\left| \frac{f(a+h) - f(a)}{h} \right| < |f'(a)| + \varepsilon$$

so

$$|f(a+h) - f(a)| < |h| \left(|f'(a)| + \varepsilon \right)$$

whenever $|h| < \delta$.

Differentiability implies continuity, 2

- Given $\varepsilon' > 0$, choose

$$\delta' < \frac{\varepsilon'}{|f'(a)| + \varepsilon}.$$

- Then $|f(a+h) - f(a)| < \varepsilon'$ whenever $0 < |h| < \delta'$.
- Hence f is continuous at a.

Differentiability implies continuity, 3

- A similar statement is true in $\mathbb{R}^n$: if

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{[f(\mathbf{x}) - \mathbf{x}(\mathbf{x}_0)] - \nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0)}{\|\mathbf{x}-\mathbf{x}_0\|} = 0$$

  as discussed above, then f is continuous at $\mathbf{x}_0$.
- Difficulty: This condition can be hard to check in practice.
- Alternative test: If all the partial derivatives of f exist and are continuous in a neighborhood of $\mathbf{x}_0$, then f is continuous at $\mathbf{x}_0$.

Parametric representations in $\mathbb{R}^n$

- A parametric representation of a line L through the point $\mathbf{p}_0$ in the direction $\mathbf{v}$ is $L(t) = \mathbf{p}_0 + t\mathbf{v}$.
- The representation is not unique because any nonzero multiple of $\mathbf{v}$ can be used

Parametric representations of 2-planes

- In $\mathbb{R}^2$, two vectors determine a plane through the origin
- A 2-plane through the origin is the graph of the function $P(s,t) = s\mathbf{u} + t\mathbf{v}$
- This idea can be extended to $\mathbb{R}^n$ for $n > 3$
- $P(s,t) = \mathbf{p}_0 + s\mathbf{u} + t\mathbf{v}$ passes through the "base point" $\mathbf{p}_0$ instead of the origin

Intersections of lines and planes in $\mathbb{R}^3$

- The plane $P(s,t) = \mathbf{p}_0 + s\mathbf{u} + t\mathbf{v}$ can be represented as the matrix-vector product

$$P(s,t) = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} + \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \\ u_3 & v_3 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix}$$

- Given the line $L(r) = \mathbf{q}_0 + r\mathbf{w}$
- Where do L and P intersect?

- Must find values of r, s, t (if they exist) such that $L(r) = P(s,t)$
- This is equivalent to

$$\mathbf{0} = P(s,t) - L(r) = (\mathbf{p}_0 - \mathbf{q}_0) + s\mathbf{u} + t\mathbf{v} - r\mathbf{w}$$

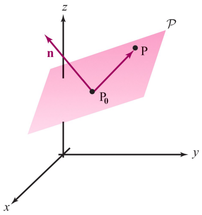
- The solution is the same as

$$\begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \\ q_3 - p_3 \end{pmatrix} = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} \begin{pmatrix} s \\ t \\ -r \end{pmatrix}$$

- If the matrix
$$\begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}$$
is nonsingular, then there is a unique solution
- So the intersection typically is a single point
- This requires that $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ not be collinear or coplanar

Intersections of lines and planes in $\mathbb{R}^3$, 4

- Important special case: In $\mathbb{R}^3$, a 2-plane is completely determined by its normal vector
- Given $\mathbf{u}$ and $\mathbf{v}$ in a parametric representation, one normal vector is $\mathbf{n} = \mathbf{u} \times \mathbf{v}$
- The plane P through the base point $\mathbf{p}_0$ is defined by $\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}_0) = \mathbf{0}$



Codimension

- Consider two lines $L_1(s) = \mathbf{a} + s\mathbf{u}$ and $L_2(t) = \mathbf{b} + t\mathbf{v}$ in $\mathbb{R}^2$
- They intersect if $L_1(s) = L_2(t)$ for some values of s and t
- Equivalently, $\mathbf{a} - \mathbf{b} = t\mathbf{v} - s\mathbf{u}$ or

$$\begin{pmatrix} a_1 - b_1 \\ a_2 - b_2 \end{pmatrix} = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} \begin{pmatrix} -s \\ t \end{pmatrix}$$

- There is a solution provided $\mathbf{u}$ and $\mathbf{v}$ aren't multiples

Now consider $\mathbb{R}^3$

- L_1 and L_2 intersect in $\mathbb{R}^3$ only if we can solve

$$\begin{pmatrix} a_1 - b_1 \\ a_2 - b_2 \\ a_3 - b_3 \end{pmatrix} = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \\ u_3 & v_3 \end{pmatrix} \begin{pmatrix} -s \\ t \end{pmatrix}$$

- This cannot be done except in special cases

How do 2-planes intersect in $\mathbb{R}^4$?

- Given $P_1(s,t) = \mathbf{a} + s\mathbf{u} + t\mathbf{v}$ and $P_2(q,r) = \mathbf{b} + q\mathbf{w} + r\mathbf{x}$

- Given $P_1(s,t) = \mathbf{a} + s\mathbf{u} + t\mathbf{v}$ and $P_2(q,r) = \mathbf{b} + q\mathbf{w} + r\mathbf{x}$
- We must solve $P_1(s,t) = P_2(q,r)$, or

$$\begin{pmatrix} a_1 - b_1 \\ a_2 - b_2 \\ a_3 - b_3 \\ a_4 - b_4 \end{pmatrix} = \begin{pmatrix} u_1 & v_1 & w_1 & x_1 \\ u_2 & v_2 & w_2 & x_2 \\ u_3 & v_3 & w_3 & x_3 \\ u_4 & v_4 & w_4 & x_4 \end{pmatrix} \begin{pmatrix} -s \\ -t \\ q \\ r \end{pmatrix}$$

- Typically there is a unique solution (so P_1 and P_2 intersect in a single point)

Codimension predicts the typical intersections

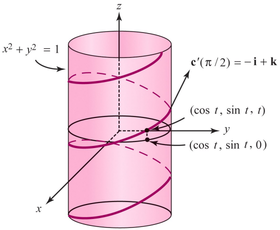
- In $\mathbb{R}^2$, two lines (1-planes) intersect in a point (0 dimensional set). Note: $1+1-2=0$
- In $\mathbb{R}^3$, a 2-plane and a line intersect in a point ($2+1-3=0$)
- In $\mathbb{R}^3$, two lines don't intersect ($1+1-3<0$)
- In $\mathbb{R}^3$, two 2-planes intersect in a line ($2+2-3=1$)
- In $\mathbb{R}^4$, two 2-planes intersect in a point ($2+2-4=0$)

Parametrized paths

- A path is a continuous function $f : [a,b] \to \mathbb{R}^n$
- Interval notation: $[a,b] = \{t : a \leq t \leq b\}$, $(a,b) = \{t : a < t < b\}$, $[a,b) = \{t : a \leq t < b\}$, etc.
- Example: $\mathbf{c}(t) = (\cos t, \sin t)$ in $\mathbb{R}^2$ for $0 \leq t \leq 2\pi$ (equivalently: $t \in [0, 2\pi]$)
- In general, $\mathbf{c}(t) = \big(c_1(t), c_2(t), \ldots, c_n(t)\big)$
- We define $\mathbf{c}'(t) = \big(c_1'(t), c_2'(t), \ldots, c_n'(t)\big)$ whenever all the derivatives exist

Important examples

- The unit circle in $\mathbb{R}^2$: $\mathbf{c}(t) = (\cos t, \sin t)$
- A line in $\mathbb{R}^n$: $\mathbf{c}(t) = \mathbf{p} + t\mathbf{v} = (p_1 + tv_1, \ldots, p_n + tv_n)$
- A helix in $\mathbb{R}^3$: $\mathbf{c}(t) = (\cos t, \sin t, t)$



The chain rule

- Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable and $\mathbf{c}(t)$ is a differentiable path in $\mathbb{R}^n$
- Example: $f(\mathbf{x})$ is the temperature at $\mathbf{x} \in \mathbb{R}^3$ and $\mathbf{c}(t)$ is a path in $\mathbb{R}^3$
- Then $f(\mathbf{c}(t))$ is a function of t
- Important special case of the chain rule:

$$\frac{d}{dt} f(\mathbf{c}(t)) = \nabla f(\mathbf{c}(t)) \cdot \mathbf{c}'(t)$$

- Generalizes the 1-d formula: $[f(x(t))]' = f'(x(t)) \cdot x'(t)$