School of Mathematics and Statistics

MATH2111 - Higher Several Variable Calculus

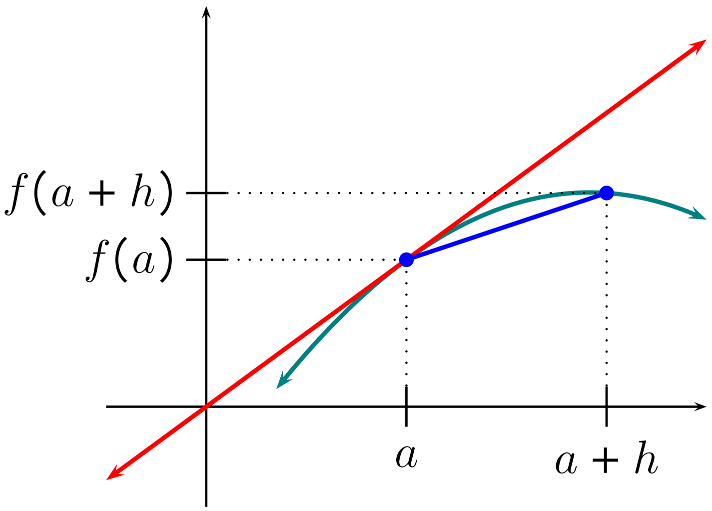
6. Differentiability

Dr Alan Stoneham

a.stoneham@unsw.edu.au

2025 Term 1

Differentiation on $\mathbb{R}$



The slope of the chord is

$$\frac{f(a+h)-f(a)}{h}$$

and the slope of the tangent line is

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$$

As $h \to 0$, the gradient of the chord approaches the gradient of the tangent. We can also write

$$\lim_{x\to a} \frac{f(x)-f(a)-f'(a)(x-a)}{x-a} = 0,$$

but how do we generalise this to functions from $\mathbb{R}^n \to \mathbb{R}^m$?

Linear and Affine Maps

Definition

A function $L : \mathbb{R}^n \to \mathbb{R}^m$ is **linear** if for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ we have that

$$L(\lambda\mathbf{x} + \mathbf{y}) = \lambda L(\mathbf{x}) + L(\mathbf{y}).$$

Recall. Any linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ can be represented by an $m \times n$ matrix A_L satisfying $L(\mathbf{x}) = A_L\mathbf{x}$.

We are slightly abusing notation since

$$L(\mathbf{x}) = \left(A_L\mathbf{x}^T\right)^T = \left(\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right)^T = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{1m}x_1 + \cdots + a_{mn}x_n \end{pmatrix}^T$$

In first-year maths, we often think of $\mathbf{x}$ as a column vector, but here we are thinking of $\mathbf{x}$ as a point in $\mathbb{R}^n$, or row vector. I'm not going to fuss over the differences in these lectures.

Conventions and Affine Maps

- For $\mathbf{x} \in \mathbb{R}^2$ or $\mathbf{x} \in \mathbb{R}^3$, we write $\mathbf{x} = (x, y)$ or $\mathbf{x} = (x, y, z)$.
- For $\mathbf{x} \in \mathbb{R}^n$ and $n \geqslant 1$, we write $\mathbf{x} = (x_1, x_2, \ldots, x_n)$.
- The standard basis of $\mathbb{R}^n$ is $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ where $\mathbf{e}_i$ is 1 in the ith co-ordinate and 0 elsewhere.
- It is also common to write in $\mathbb{R}^2$ or $\mathbb{R}^3$ that $\mathbf{i} = \mathbf{e}_1$, $\mathbf{j} = \mathbf{e}_2$ and $\mathbf{k} = \mathbf{e}_3$.
- We'll often write $A\mathbf{x}$ when we really mean $A\mathbf{x}^T$ or $(A\mathbf{x}^T)^T$.

Definition

A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is called **affine** if there is a $\mathbf{y}_0 \in \mathbb{R}^m$ and a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$T(\mathbf{x}) = L(\mathbf{x}) + \mathbf{y}_0 \qquad \forall \mathbf{x} \in \mathbb{R}^n.$$

Equivalently, there is $\mathbf{y}_0 \in \mathbb{R}^m$ and an $m \times n$ matrix A such that $T(\mathbf{x}) = A\mathbf{x} + \mathbf{y}_0$ for all $\mathbf{x} \in \mathbb{R}^n$.

Affine Approximation

Definition

A function $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ **has an affine approximation at** $\mathbf{x}_0 \in \Omega$ if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

bars $\Leftarrow\Rightarrow$ norm

$|\underset{\sim}{v}| = d(\underset{\sim}{v}, 0)$

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{|f(\mathbf{x}) - f(\mathbf{x}_0) - L(\mathbf{x} - \mathbf{x}_0)|}{|\mathbf{x} - \mathbf{x}_0|} = 0.$$

If f has an affine approximation at $\mathbf{x}_0 \in \Omega$, then f is said to be **differentiable at** $\mathbf{x}_0$.

We will soon see that if such a linear map L exists, then it is unique, and we call it the **derivative of f at** $\mathbf{x}_0$, denoted by $Df(\mathbf{x}_0)$ or $D_{\mathbf{x}_0}f$.

Remark. With the definition above, it should be reasonably clear that f being differentiable at $\mathbf{x}_0$ implies that f is continuous at $\mathbf{x}_0$ as well.

Partial Derivatives

Recall that in $\mathbb{R}^n$, we can notate the co-ordinates by $x_1, x_2, \ldots, x_n$ and the standard basis vectors by $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$.

not jth order partial derivative

Definition

Let $\mathbf{x}_0$ be an interior point of $\Omega \subseteq \mathbb{R}^n$ and $f : \Omega \to \mathbb{R}$. The j**th partial derivative of** f **at** $\mathbf{x}_0$ is defined as

$$\frac{\partial f}{\partial x_j}(\mathbf{x}_0) = \lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\mathbf{e}_j) - f(\mathbf{x}_0)}{h},$$

w.r.t x_j

and $\dfrac{\partial f}{\partial x_j}$ is the function that takes $\mathbf{x}_0$ to $\dfrac{\partial f}{\partial x_j}(\mathbf{x}_0)$.

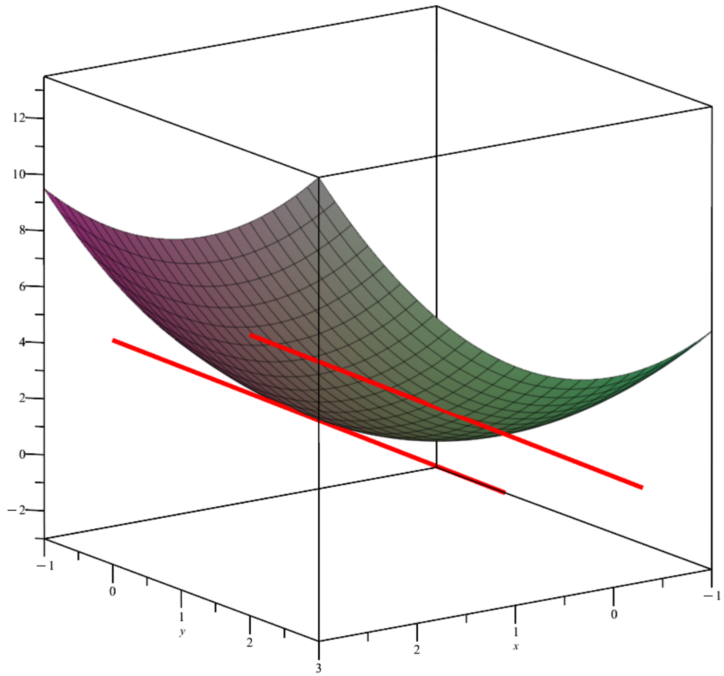
Some other common notations are $\dfrac{\partial f}{\partial x_j} = \partial_{x_j} f = f_{x_j} = f_j = \partial_j f$.

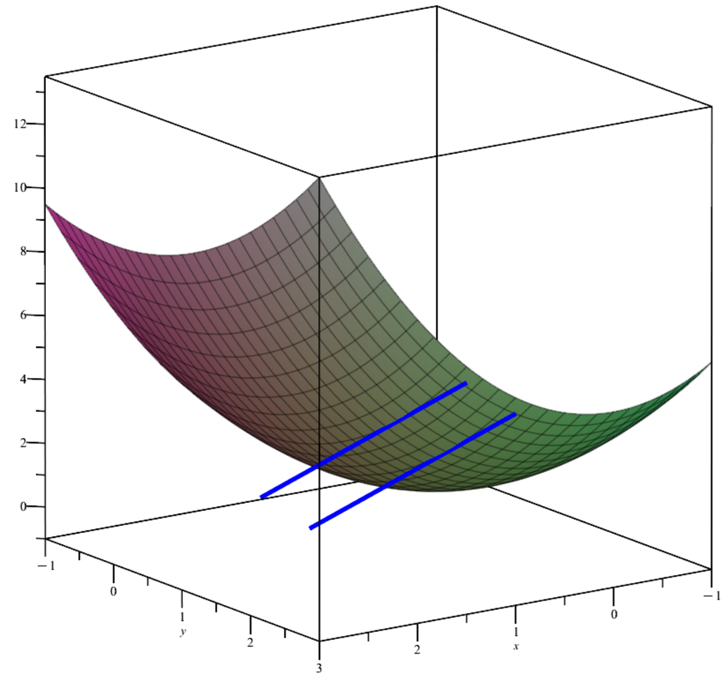
I don't like the "f_j" notation since it looks like the jth component of f.

Partial Derivatives

Consider $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2 + \frac{1}{2}y^2$. Then

$$\frac{\partial f}{\partial x} = 2x \qquad \frac{\partial f}{\partial y} = y$$





Jacobian Matrix

Definition

If all partial derivatives of $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ exist at $\mathbf{x}_0 \in \Omega$, then the **Jacobian of** f **at** $\mathbf{x}_0$ is the $m \times n$ matrix defined by $\big(Jf(\mathbf{x}_0)\big)_{ij} = \dfrac{\partial f_i}{\partial x_j}(\mathbf{x}_0)$. That is,

J_f

$$J_f(\mathbf{x}_0) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \dfrac{\partial f_1}{\partial x_2}(\mathbf{x}_0) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\[2ex] \dfrac{\partial f_2}{\partial x_1}(\mathbf{x}_0) & \dfrac{\partial f_2}{\partial x_2}(\mathbf{x}_0) & \cdots & \dfrac{\partial f_2}{\partial x_n}(\mathbf{x}_0) \\[2ex] \vdots & \vdots & \ddots & \vdots \\[2ex] \dfrac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \dfrac{\partial f_m}{\partial x_2}(\mathbf{x}_0) & \cdots & \dfrac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{pmatrix}$$

which may also be denoted by $J_{\mathbf{x}_0} f$, $Jf(\mathbf{x}_0)$, or $\dfrac{\partial(f_1, \ldots, f_m)}{\partial(x_1, \ldots, x_n)}$.

Jacobian Matrix

Exercise. Find the Jacobian at the points $\mathbf{x}_1 = (1, -1)$ and $\mathbf{x}_2 = (0, \pi)$ for the function $f : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$f(x, y) = (\overset{f_1}{\sin(x^2 + y)},\ \overset{f_2}{x^4 - y^3},\ \overset{f_3}{xe^y + ye^x}).$$

$$\frac{\partial f_1}{\partial x} = 2x\cos(x^2+y^2) \qquad \frac{\partial f_1}{\partial y} = \cos(x^2+y^2)$$

$$\frac{\partial f_2}{\partial x} = 4x^3 \qquad \frac{\partial f_2}{\partial y} = -3y^2$$

$$\frac{\partial f_3}{\partial x} = e^y + ye^x \qquad \frac{\partial f_3}{\partial y} = xe^y + ye^x$$

$$J_f\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix} 2x\cos(x^2+y) & \cos(x^2+y) \\ 4x^3 & -3y^2 \\ e^y+ye^x & xe^y+e^x \end{pmatrix}$$

$$J_f(1,-1) = \begin{pmatrix} 2 & 1 \\ 4 & -3 \\ e^{-1}-e & e^{-1}+e \end{pmatrix}$$

$$J_f(0,\pi) = \begin{pmatrix} 0 & -1 \\ 0 & -3\pi^2 \\ e^{\pi}+\pi & 1 \end{pmatrix}$$

Best Affine Approximation

Theorem

Let $\Omega \subseteq \mathbb{R}^n$, $\mathbf{x}_0$ be an interior point of Ω and $f : \Omega \to \mathbb{R}^m$ be a function. If f is differentiable at $\mathbf{x}_0$, then its derivative at $\mathbf{x}_0$ is unique, all partial derivatives at $\mathbf{x}_0$ exist, and $Df(\mathbf{x}_0)$ is represented by the $m \times n$ matrix $J_f(\mathbf{x}_0)$.

Proof later on slide 12. When a function is differentiable, we might use 'derivative' and 'Jacobian' interchangeably.

Definition

Under the above hypotheses, the function $T_{\mathbf{x}_0} f : \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$T_{\mathbf{x}_0} f(\mathbf{x}) = J_f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + f(\mathbf{x}_0)$$

Basically degree 1 Taylor polynomial

is **the best affine approximation of f at $\mathbf{x}_0$**.

The graph of $T_{\mathbf{x}_0} f$ is the **tangent plane of f at $\mathbf{x}_0$.**

UNSW SYDNEY

Best Affine Approximation

Example. Find the best affine approximation of $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2 + \frac{1}{2}y^2$, at $(1,1)$. Then

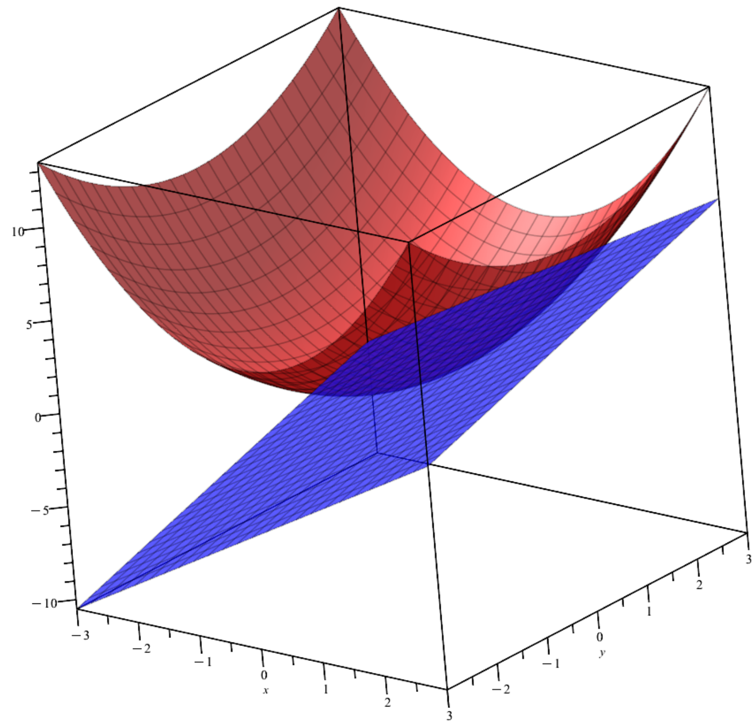
$$\frac{\partial f}{\partial x} = 2x \qquad \text{and} \qquad \frac{\partial f}{\partial y} = y.$$

The Jacobian of f at $(1,1)$ is

$$\begin{pmatrix} 2 & 1 \end{pmatrix},$$

so the best affine approximation is given by

$$\begin{aligned} T_{(1,1)}f(x,y) &= \begin{pmatrix} 2 & 1 \end{pmatrix}\begin{pmatrix} x-1 \\ y-1 \end{pmatrix} + f(1,1) \\ &= 2(x-1) + (y-1) + \frac{3}{2}. \end{aligned}$$



Best Affine Approximation

Exercise. Find the best affine approximations to $f : \mathbb{R}^2 \to \mathbb{R}^3$, defined by

$$f(x, y) = \left(\sin(x^2 + y), x^4 - y^3, xe^y + ye^x\right)$$

at the points $(1, -1)$ and $(0, \pi)$.

$$T_{(1,-1)} f(\underset{\sim}{x}) = J_f(1,-1)(\underset{\sim}{x} - (1,-1)) + f(1,-1)$$

$$= \begin{pmatrix} 2 & -1 \\ 4 & -3 \\ e^{-1} - e & e^{-1} + e \end{pmatrix} \begin{pmatrix} x-1 \\ x+1 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ e^{-1} - e \end{pmatrix}$$

$$T_{(0,\pi)} f(\underset{\sim}{x}) = J_f(0,\pi)(\underset{\sim}{x} - (0,\pi)) + f(0,\pi)$$

$$= \begin{pmatrix} 0 & -1 \\ 0 & -3\pi^2 \\ e^{\pi} + \pi & 1 \end{pmatrix} \begin{pmatrix} x \\ y-\pi \end{pmatrix} + \begin{pmatrix} 0 \\ -\pi^3 \\ \pi \end{pmatrix}$$

UNSW SYDNEY

Proof of Jacobian Uniqueness

Suppose that $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at an interior point $\mathbf{x}_0 \in \Omega$ with derivative $Df(\mathbf{x}_0)$ represented by $A = (a_{ij})_{i,j=1}^{m,n}$. As f is differentiable,

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{|f(\mathbf{x}) - f(\mathbf{x}_0) - A(\mathbf{x}-\mathbf{x}_0)|}{|\mathbf{x}-\mathbf{x}_0|} = 0.$$

In particular, if $\mathbf{x} = \mathbf{x}_0 + h\mathbf{e}_j$, then as $h \to 0$ we have

$$\lim_{h\to 0} \frac{|f(\mathbf{x}_0 + h\mathbf{e}_j) - f(\mathbf{x}_0) - A(h\mathbf{e}_j)|}{|h\mathbf{e}_j|} = 0.$$

simplifies to

$$A\underset{\sim}{e}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

Since $A\mathbf{e}_j$ is the jth column of A, this is equivalent to

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} = \begin{pmatrix} \lim_{h\to 0} \frac{f_1(\mathbf{x}_0+h\mathbf{e}_j)-f_1(\mathbf{x}_0)}{h} \\ \lim_{h\to 0} \frac{f_2(\mathbf{x}_0+h\mathbf{e}_j)-f_2(\mathbf{x}_0)}{h} \\ \vdots \\ \lim_{h\to 0} \frac{f_m(\mathbf{x}_0+h\mathbf{e}_j)-f_m(\mathbf{x}_0)}{h} \end{pmatrix} \iff a_{ij} = \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) \text{ for all } 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n.$$

$\square$

UNSW SYDNEY

Partial Derivatives $\Longrightarrow$ Differentiable?

$$\left| \frac{\sin(x)}{x} - 1 \right| \leq x^2$$

Exercise. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x, y) = \begin{cases} \dfrac{\sin(x^2 + y^2)}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 1 & \text{if } (x, y) = (0, 0). \end{cases}$$

Is f differentiable at $(0, 0)$?

If f is differentiable, then $Df(0,0)$ must be represented by $J_f(0,0)$.

$$0 \leq \left| \frac{\partial f}{\partial x}(0,0) \right| = \lim_{h \to 0} \left| \frac{f(0+h, 0) - f(0,0)}{h} \right|$$

$$= \lim_{h \to 0} \left| \frac{\frac{\sin(h^2)}{h^2} - 1}{h} \right| \leq \lim_{h \to 0} \frac{(h^2)^2}{h}$$

$$= 0$$

f

Partial Derivatives $\Longrightarrow$ Differentiable?

similarly $\frac{\partial f}{\partial y}(0,0) = 0$, so $J_f(0,0) = (0 \quad 0)$

check differentiability:

$$\frac{|f(x,y) - f(0,0) - J_f(0,0)(\underset{\sim}{x} - (0,0))|}{|\underset{\sim}{x} - (0,0)|}$$

$$= \frac{|\sin(x^2+y^2) - 1|}{\sqrt{x^2+y^2}} \leq \frac{(x^2+y^2)^2}{\sqrt{x^2+y^2}} = |\underset{\sim}{x}|^{3/2}$$

$\longrightarrow 0$ as $|\underset{\sim}{x}| \longrightarrow 0$

So f is differentiable at $(0,0)$.

Partial Derivatives $\not\Longrightarrow$ Differentiable

It is pretty easy to provide an example of $f : \mathbb{R}^2 \to \mathbb{R}$ where the partial derivatives exist but f is not differentiable. In fact, f can be constructed to not even be continuous!

Consider $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x,y) = \begin{cases} 1 & \text{if } xy = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\dfrac{\partial f}{\partial x}(0,0) = \dfrac{\partial f}{\partial y}(0,0) = 0$ but f is not continuous at $(0,0)$, and hence not differentiable.

Ex. Show that $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x,y) = \sqrt[3]{xy}$ is continuous with partial derivatives at $(0,0)$, but $D_{(0,0)}f$ does not exist.

Continuous Partial Derivatives $\implies$ Differentiable

Theorem

Suppose $\Omega \subseteq \mathbb{R}^n$ is open and $f : \Omega \to \mathbb{R}^m$. If the partial derivatives $\dfrac{\partial f_i}{\partial x_j}$ exist and are continuous on Ω for all $i = 1, \dots, m$ and $j = 1, \dots, n$, then f is differentiable on Ω.

If you have taken MATH2621, this theorem is similar to the theorem stating that the Cauchy-Riemann equations holding along with continuous partial derivatives implies complex differentiability.

Proof. (Included for completion sake, but very difficult!)

Suppose $\varepsilon > 0$ and $\mathbf{a} = (a_1, \dots, a_n) \in \Omega$. Since Ω is open, there is an $r > 0$ such that $d_1(\mathbf{x}, \mathbf{a}) < r$ implies[1] $\mathbf{x} \in \Omega$. For each $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$, since $\frac{\partial f_i}{\partial x_j}$ is continuous, we can find $0 < \delta_{i,j} < r$ such that

$$\left| \frac{\partial f_i}{\partial x_j}(\mathbf{x}) - \frac{\partial f_i}{\partial x_j}(\mathbf{a}) \right| < \frac{\varepsilon}{m} \qquad (*)$$

whenever $d_1(\mathbf{x}, \mathbf{a}) = \displaystyle\sum_{j=1}^{n} |x_j - a_j| < \delta_{i,j}$. Let $\delta = \displaystyle\min_{i,j} \delta_{i,j}$.

[1] Recall that the d_1 metric is equivalent to the Euclidean metric.

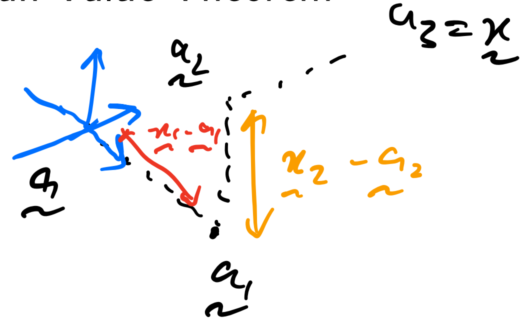
d_1 : diamond in $\mathbb{R}^2$

Continuous Partial Derivatives $\implies$ Differentiable

Telescoping sum

Suppose that $\mathbf{x} = (x_1, \ldots, x_n) \in \Omega$ satisfies $d_1(\mathbf{x}, \mathbf{a}) < \delta$. Let $\mathbf{a}_0 = \mathbf{a}$ and $\mathbf{a}_j = \mathbf{a} + \sum_{k=1}^{j}(x_k - a_k)\mathbf{e}_k$ for $1 \leqslant j \leqslant n$ so that $f_i(\mathbf{x}) - f_i(\mathbf{a}) = \sum_{j=1}^{n}\big(f_i(\mathbf{a}_j) - f_i(\mathbf{a}_{j-1})\big)$ for all $1 \leqslant i \leqslant m$. Since the partial derivatives exist everywhere, the Mean Value Theorem yields $c_{i,j}$ between 0 and $x_j - a_j$ such that

$$\frac{f_i(\mathbf{a}_j) - f_i(\mathbf{a}_{j-1})}{x_j - a_j} = \frac{\partial f_i}{\partial x_j}(\mathbf{a}_{j-1} + c_{i,j}\mathbf{e}_j).$$



Then, noting that $\mathbf{a} = \mathbf{a}_0$ and $\mathbf{x} = \mathbf{a}_n$,

$$\begin{aligned}
\frac{d_1\big(f(\mathbf{x}) - f(\mathbf{a}), J_f(\mathbf{a})(\mathbf{x} - \mathbf{a})\big)}{d_1(\mathbf{x}, \mathbf{a})} &= \frac{\sum_{i=1}^{m}\left|f_i(\mathbf{x}) - f_i(\mathbf{a}) - \sum_{j=1}^{n}\frac{\partial f_i}{\partial x_j}(\mathbf{a})(x_j - a_j)\right|}{\sum_{j=1}^{n}|x_j - a_j|} \\
&= \frac{\sum_{i=1}^{m}\left|\sum_{j=1}^{n}\left(f_i(\mathbf{a}_j) - f_i(\mathbf{a}_{j-1}) - \frac{\partial f_i}{\partial x_j}(\mathbf{a})(x_j - a_j)\right)\right|}{\sum_{j=1}^{n}|x_j - a_j|} \\
&= \frac{\sum_{i=1}^{m}\left|\sum_{j=1}^{n}\left(\frac{\partial f_i}{\partial x_j}(\mathbf{a}_{j-1} + c_{i,j}\mathbf{e}_j) - \frac{\partial f_i}{\partial x_j}(\mathbf{a})\right)(x_j - a_j)\right|}{\sum_{j=1}^{n}|x_j - a_j|}.
\end{aligned}$$

Continuous Partial Derivatives $\implies$ Differentiable

Since

$$d_1(\mathbf{a}_{j-1} + c_{i,j}\mathbf{e}_j, \mathbf{a}) = \sum_{k=1}^{j-1} |x_k - a_k| + |c_{i,j}| \leqslant \sum_{k=1}^{n} |x_k - a_k| = d_1(\mathbf{x}, \mathbf{a}) < \delta,$$

then applying $({\color{blue}*})$ yields

$$\left| \frac{\partial f_i}{\partial x_j}(\mathbf{a}_{j-1} + c_{i,j}\mathbf{e}_j) - \frac{\partial f_i}{\partial x_j}(\mathbf{a}) \right| < \frac{\varepsilon}{m}.$$

Hence whenever $d_1(\mathbf{x}, \mathbf{a}) < \delta$,

$$\frac{d_1\big(f(\mathbf{x}) - f(\mathbf{a}), J_f(\mathbf{a})(\mathbf{x} - \mathbf{a})\big)}{d_1(\mathbf{x}, \mathbf{a})} < \frac{m\left(\frac{\varepsilon}{m}\right) \sum_{j=1}^{n} |x_j - a_j|}{\sum_{j=1}^{n} |x_j - a_j|} = \varepsilon.$$

Thus

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{d_1\big(f(\mathbf{x}) - f(\mathbf{a}), J_f(\mathbf{a})(\mathbf{x} - \mathbf{a})\big)}{d_1(\mathbf{x}, \mathbf{a})} = 0,$$

and so f is differentiable at $\mathbf{a}$. $\square$

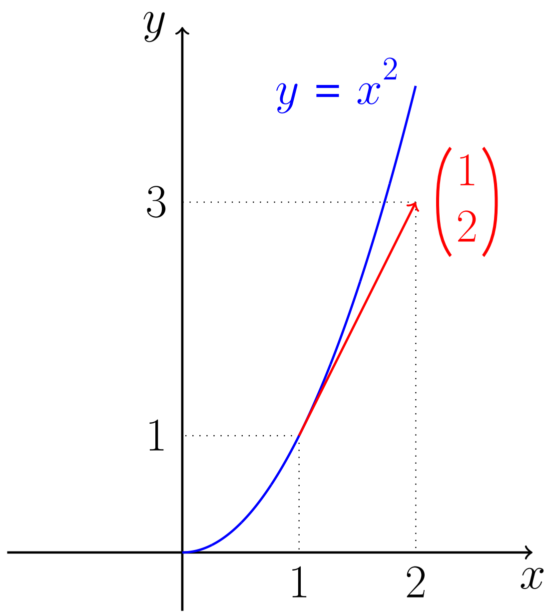
Differentiation of Curves

Let's return to Earth with an example. Consider the curve $\mathbf{c} : [0, 2] \to \mathbb{R}^2$ defined by $\mathbf{c}(t) = (t, t^2)$.

We see that $\mathbf{c}$ is differentiable at $t_0 = 1$ with derivative

$$J_{\mathbf{c}}(t_0) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

since the partial derivatives $\dfrac{\partial \mathbf{c}_1}{\partial t} = 1$ and $\dfrac{\partial \mathbf{c}_2}{\partial t} = 2t$ are continuous.



We see that $J_{\mathbf{c}}(1) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is tangent to the curve at the point $\mathbf{c}(1) = (1, 1)$. If we interpret $\mathbf{c}(1)$ as the position of a particle at time $t = 1$, then $\mathbf{c}'(1)$ can be interpreted as the particle's velocity at time $t = 1$, and its speed is $\sqrt{5}$, the length of $\mathbf{c}'(1)$.

Differentiation of Curves

Combining the definition of differentiability and the fact that the limit of a vector-valued function exists if and only if the limit of all the components exist gives us the following.

Proposition

A function $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\mathbf{x}_0$ if and only if all of its components $f_1, f_2, \ldots, f_m$ are differentiable at $\mathbf{x}_0$.

So for a path $\mathbf{c} : I \to \mathbb{R}^m$ with components $c_i : I \to \mathbb{R}$, $1 \leqslant i \leqslant m$, $\mathbf{c}$ is differentiable at $t_0 \in I$ if and only if c_i is differentiable at t_0 for all $1 \leqslant i \leqslant m$. This means that $c_i'(t_0)$ exists for all $1 \leqslant i \leqslant m$ and **the derivative of $\mathbf{c}$ at t_0** is

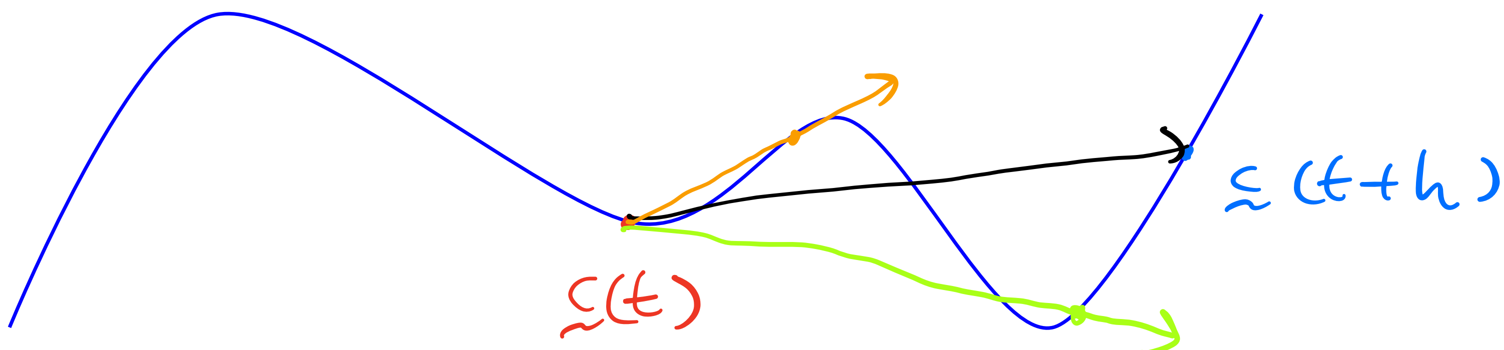
$$J_{\mathbf{c}}(t_0) = \mathbf{c}'(t_0) = \frac{d\mathbf{c}}{dt}(t_0) = \begin{pmatrix} c_1'(t_0) \\ c_2'(t_0) \\ \vdots \\ c_m'(t_0) \end{pmatrix}.$$

Geometric Interpretation of $\mathbf{c}'(t)$

For $\mathbf{c} : I \to \mathbb{R}^m$, consider the following:

$$
\begin{aligned}
\mathbf{c}'(t) &= \left(\mathbf{c}_1'(t), \ldots, \mathbf{c}_m'(t)\right)^T \\
&= \left(\lim_{h \to 0} \frac{c_1(t+h) - c(t)}{h}, \ldots, \lim_{h \to 0} \frac{c_m(t+h) - c(t)}{h}\right)^T \\
&= \lim_{h \to 0} \left(\frac{c_1(t+h) - c(t)}{h}, \ldots, \frac{c_m(t+h) - c(t)}{h}\right)^T \\
&= \lim_{h \to 0} \frac{1}{h}\left(\mathbf{c}(t+h) - \mathbf{c}(t)\right).
\end{aligned}
$$

$h < 1$



Geometric Interpretation of $\mathbf{c}'(t)$

As $h \to 0$, $\frac{1}{h}\big(\mathbf{c}(t+h) - \mathbf{c}(t)\big)$ approaches a vector which is parallel to a tangent line to $\mathbf{c}$ at $\mathbf{c}(t)$.

Definition

For a path $\mathbf{c} : I \to \mathbb{R}^m$ and $t_0 \in I$, if $\mathbf{c}'(t_0)$ exists and is not $\mathbf{0}$, then $\mathbf{c}'(t_0)$ is called the **tangent vector of $\mathbf{c}$ at t_0**, or the **velocity of $\mathbf{c}$ at t_0**, which can be denoted by $\mathbf{v}(t_0)$.

We can also define the **speed** of $\mathbf{c}$ at t_0 as $|\mathbf{v}(t_0)| = \sqrt{\mathbf{v}(t_0) \cdot \mathbf{v}(t_0)}$ and the **acceleration** of $\mathbf{c}$ at t_0 as $\mathbf{a}(t_0) = \mathbf{v}'(t_0) = \mathbf{c}''(t_0)$.

Higher order derivatives are known as '**jerk**', '**snap**', '**crackle**', and '**pop**'.

Differentiation Rules for Paths $\mathbf{c} : I \to \mathbb{R}^m$

Theorem

Suppose that $\lambda \in \mathbb{R}$ and $\mathbf{c}_1, \mathbf{c}_2 : I \to \mathbb{R}^m$ and $f : I \to \mathbb{R}$ are all differentiable. Then for all $t \in I$, we have that

- $\big(\mathbf{c}_1 + \mathbf{c}_2\big)'(t) = \mathbf{c}_1(t) + \mathbf{c}_2(t)$;
- $\big(\lambda \mathbf{c}_1\big)'(t) = \lambda \mathbf{c}_1'(t)$;
- $\dfrac{d}{dt}\big(f(t)\mathbf{c}_1(t)\big) = f'(t)\mathbf{c}_1(t) + f(t)\mathbf{c}_1'(t)$;
- $\dfrac{d}{dt}\big(\mathbf{c}_1 \cdot \mathbf{c}_2\big)(t) = \mathbf{c}_1'(t) \cdot \mathbf{c}_2(t) + \mathbf{c}_1(t) \cdot \mathbf{c}_2'(t)$;
- if $m = 3$, then $\dfrac{d}{dt}\big(\mathbf{c}_1(t) \times \mathbf{c}_2(t)\big) = \mathbf{c}_1'(t) \times \mathbf{c}_2(t) + \mathbf{c}_1(t) \times \mathbf{c}_2'(t)$.

If instead $f : I' \to I$ is differentiable, then $\mathbf{c}_1 \circ f : I' \to \mathbb{R}^m$ is differentiable with

$$\frac{d}{dt}\mathbf{c}\big(f(t)\big) = f'(t)\mathbf{c}_1'\big(f(t)\big).$$

The Chain Rule

Theorem (The Chain Rule)

Let $\Omega \subseteq \mathbb{R}^n$, $\Omega' \subseteq \mathbb{R}^m$ and $\mathbf{x}_0 \in \mathrm{Int}(\Omega)$. Suppose that $f : \Omega \to \mathbb{R}^m$ and $g : \Omega' \to \mathbb{R}^k$ with $f(\Omega) \subseteq \Omega'$. If f is differentiable at $\mathbf{x}_0$ and g is differentiable at $f(\mathbf{x}_0)$, then $g \circ f$ is differentiable at $\mathbf{x}_0$ and

$$D(f \circ g)(\mathbf{x}_0) = Dg\big(f(\mathbf{x}_0)\big)\, Df(\mathbf{x}_0).$$

Exercise. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by $f(x, y) = x^2 y$. How does $f(x, y)$ change as $|(x, y)| = \sqrt{x^2 + y^2}$ change?

Let $p : \mathbb{R}^2 \setminus \{0\} \longrightarrow \mathbb{R}^2$ be the polar map

$p(r, \theta) = (r\cos\theta, r\sin\theta)$

$$J_p(r, \theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

should be thinking of $f \circ g$

UNSW SYDNEY

The Chain Rule

Ex. cont. How does $f(x, y) = x^2 y$ change as $|(x, y)| = \sqrt{x^2 + y^2}$ change?

The Jacobian of f is

$$J_f(x, y) = \begin{pmatrix} 2xy & x^2 \end{pmatrix}$$

$x = r\cos\theta$

$y = r\sin\theta$

Applying the chain rule

$$J_{f \circ p}(r, \theta) = J_f(p(r, \theta))\, J_p(r, \theta)$$

$$= \big(2r^2 \cos\theta \sin\theta \quad r^2 \cos$$