- **Limits of functions**
  - Def: Let X and Y be metric space; suppose $D \subset X$, $f : D \to Y$, and x_0 is a limit point of D. We write $f(x) \to q$ as $x \to x_0$, or $\lim_{x \to x_0} f(x) = q$
    - if there is a point $q \in Y$ with the following property:
      $\forall e > 0\ \exists d > 0$ such that $\forall x \in D$, $\underset{x \neq x_0}{0} < d^X(x, x_0) < d \Rightarrow d^Y(f(x), q) < e$
      - x_0 need not be in D. Even if $x_0 \in D$, we may have $f(x_0) \neq \lim_{x \to x_0} f(x)$.

    $\equiv$ $\lim_{n \to \infty} f(x_n) = q$ for every sequence $\{x_n\}$ in D such that $x_n \neq x$, and $\lim_{n \to \infty} x_n = x$.

    Proof. "$\Rightarrow$": Given e, can find d such that $\underset{x \neq x_0}{0} < d^X(x, x_0) < d \Rightarrow$ $d^Y(f(x), q) < e$. Also, given $e' = d$, can find N such that $\forall n \geq N$, $d^X(x_n, x_0)$ $< e' = d$, which implies $d^Y(f(x_n), q) < e$. "$\Leftarrow$": Suppose $\lim_{n \to \infty} f(x_n) \neq q$, then $\exists e\ \ \forall d = \frac{1}{n}$, $\underset{x \neq x_0}{0} < d^X(x_n, x_0) < d$ and $d^Y(f(x_n), q) > e$. The sequence $x_n \to x_0$ but $f(x_n) \not\to q$.
  - If f has a limit at p, this limit is unique.

    Proof. If two limits, then any sequence has to converge to both limits, which implies the limits are equal.
  - Let x_0 be a limit point of D, $Y \subset \mathbb{R}^n$, $\lim_{x \to x_0} f(x) = y_1$, and $\lim_{x \to x_0} g(x) = y_2$, then (1) $\lim_{x \to x_0} (f + g)(x) = y_1 + y_2$, and (2) $\lim_{x \to x_0} (f \cdot g)(x) = y_1 \cdot y_2$.

    Proof. Consider any sequence $x_n \to x_0$. We have sequences in $\mathbb{R}^k$ $f(x_n) \to y_1$ and $g(x_n) \to y_2$; thus, $f(x_n) + g(x_n) \to y_1 + y_2$, $f(x_n) \cdot g(x_n) \to y_1 \cdot y_2$.
  - Let x_0 be a limit point of D, $Y \subset \mathbb{R}^n$. $\lim_{x \to x_0} f(x) = y \Leftrightarrow \lim_{x \to x_0} f_k(x) = (y)_k$.

    Proof. "$\Rightarrow$" Consider any sequence $x_n \to x_0$; we have $\lim_{n \to \infty} f(x_n) = y$. In $\mathbb{R}^n$, convergence means convergence for each component. So, $\lim_{n \to \infty} f_k(x_n) = (y)_k$. This is true for any sequence $x_n \to x_0$. "$\Leftarrow$"Consider any sequence $x_n \to x_0$. For all k, $\lim_{x \to x_0} f_k(x) = (y)_k$; so, $\lim_{n \to \infty} f_k(x_n) = (y)_k$ for all k. Thus, $\lim_{n \to \infty} f(x_n) = y$. This is true for any sequence $x_n \to x_0$.

Alternative proof. "$\Rightarrow$" $\lim_{x\to x_0} f(x) = y$ means $\forall e\ \exists d$ such that $\underset{x\neq x_0}{0} < |x-x_0| < d \Rightarrow$ $|f(x)-y| < e$. Hence we have $|f_k(x)-y_k| \le \sqrt{\sum_{k=1}^{n} |f_k(x)-y_k|^2} = |f(x)-y| < e$.

"$\Leftarrow$" $\lim_{x\to x_0} f_k(x) = y_k$ means $\forall e\ \exists d$ such that $\underset{x\neq x_0}{0} < |x-x_0| < d \Rightarrow |f_k(x)-y_k| < \frac{e}{\sqrt{n}}$.

Thus, $|f(x)-y| = \sqrt{\sum_{k=1}^{n} |f_k(x)-y_k|^2} < e$

- Euclidean: Let x_0 be a limit point of D, $Y \subset \mathbb{R}^n$. $\lim_{x\to x_0} |f(x)| = 0 \Leftrightarrow \lim_{x\to x_0} f(x) = 0$.

  Proof. By def. The right hand side means $\forall e > 0\ \exists d > 0$ such that $\forall x \in D$, $\underset{x\neq x_0}{0} < |x-x_0| < d \Rightarrow |f(x)-0| < e$. The left hand side has $\big||f(x)|\big| < e$; same.

Euclidean Space

- $\forall x\ \ ax = 0 \Leftrightarrow a = 0$ matrix.

  Proof. Choose $x = e^{(j)}$. Then, ax is the j^{th} column of A. $ax = 0$ implies the j^{th} column of A is zero.

- Let A be any $m \times n$ matrix, $(a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$, then $\forall x \in \mathbb{R}^n\ \exists c$ such that $|Ax| \le c|x|$.
  - Take $c = \sqrt{\sum_{k=1}^{m}\sum_{j=1}^{n} a_{kj}^2}$.
- $\lim_{x\to y} A(x-y) = 0$

  Proof. $\exists c$ such that $|A(x-y)| \le c|x-y|$.

- Little o
  - $f(x) = o(h(x))$ and $g(x) = o(h(x))$ as $x \to x_0$, then $f(x)+g(x) = o(h(x))$ as $x \to x_0$.
  - $f(x) = o(h(x))$ as $x \to x_0$

    $\equiv\ \ \lim_{x\to x_0} \frac{f(x)}{h(x)} = 0$ (don't care about $f(x_0), h(x_0)$)

    $\equiv\ \ \forall e > 0\ \exists d > 0$ such that $\forall x \in D\ \ |x-x_0| < d \Rightarrow |f(x)| \le |h(x)|e$.

    $\Rightarrow f(x_0) = 0$.

    $\equiv\ \ f(x) = o(h(x))$ as $x \to x_0 \Leftrightarrow \forall k\ \ f_k(x) = o(h(x))$ also as $x \to x_0$.

    Proof1. "$\Rightarrow$": $\forall e > 0\ \exists d > 0$ such that $\forall x \in D\ \ |x-x_0| < d \Rightarrow |f(x)| \le |h(x)|e$.

- $f(x) = o(|x-y|)$ as $x \to y$
  - iff $\forall e > 0\ \exists d > 0$ such that $\forall x \in D$, $|x-y| < d \Rightarrow |f(x)| \le |x-y| e$.

$\equiv$ $\lim_{x \to y} \frac{f(x)}{|x-y|} = 0$.

$\equiv$ for all k, $f_k(x) = o(|x-y|)$ as $x \to y$.

Proof. Let $g(x) = \frac{f(x)}{|x-y|}$. Then, $g_k(x) = \left(\frac{f(x)}{|x-y|}\right)_k = \frac{f_k(x)}{|x-y|}$.

$\lim_{x \to y} g(x) = 0 \Leftrightarrow \lim_{x \to y} g_k(x) = 0$.

$\Rightarrow$ $\lim_{x \to y} f(x) = 0$.

Proof. By def of $f(x) = o(|x-y|)$, given e, use $d' = \min\{d, 1\}$. Then $|x-y| < d'$ still $< d$; thus, $|f(x)| \le |x-y| e < e$.

$\Rightarrow$ $f(y) = 0$ if $f(x)$ is continuous.

$\Rightarrow$ If $f(x) = a(x-y) + b$, then $f(x) = 0\ \forall x$.

Proof. $f(y) = 0 \Rightarrow b = 0$. We then have $\lim_{x \to y} \frac{a(x-y)}{|x-y|} = 0$. This is true $\forall x \to y$.

So, consider x only of the form $x = y + tw$ where $w \neq 0$. As $t \to 0^+$, $x \to y$.

$$\lim_{t \to 0^+} \frac{a(y+tw-y)}{|y+tw-y|} = \lim_{t \to 0^+} \frac{a(tw)}{|tw|} = \lim_{x \to y} \frac{a(tw)}{|t||w|} = \lim_{t \to 0^+} \frac{a(w)}{|w|} = a(w) = 0 \Rightarrow a = 0.$$

- For scalars a, $a o(|x-y|)$ is still $o(|x-y|)$.
- For scalars a, b, $a o(|x-y|) + b o(|x-y|)$ is still $o(|x-y|)$.
- $f : D \to \mathbb{R}^m$, $g : D \to \mathbb{R}$. $f(x) = o(|x-y|)$ and $g(x) = o(|x-y|)$ implies $g \cdot f(x)$ is $o(|x-y|)$.

Proof. This implies $\lim_{x \to y} \frac{f_k(x)}{|x-y|} \frac{g(x)}{|x-y|} = 0$. Thus, $\frac{(g \cdot f)_k(x)}{|x-y|^2} = o\left(|x-y|^2\right)$ and $\frac{g \cdot f_k(x)}{|x-y|} = o(|x-y|)$. The later part implies $\lim_{x \to y} \frac{g \cdot f_k(x)}{|x-y|} = 0$. So, $\frac{g \cdot f_k(x)}{|x-y|} = o(|x-y|)$.

Differential Calculus in Euclidean Space

- $f:\mathbb{R}^n \to \mathbb{R}^m$
- **Limit of a function**: $\lim_{x\to x_0} f(x)$ exists if $\exists y \left(=\lim_{x\to x_0} f(x)\right)$, such that

  $\forall e>0\ \exists d>0$ such that $\forall x\in D$, $|x-x_0|<d$ and $x\neq x_0 \Rightarrow |f(x)-y|<e$.

The Differential

- Convention: all vectors are considered as column vectors in any equation involving matrix multiplication.
- a: $m\times n$ matrix (m rows and n columns)
- Affine function: $g(x)=ax+b=\begin{pmatrix} g_1(x)\\ \vdots \\ g_m(x)\end{pmatrix}$, $g_k(x)=\sum_{j=1}^{n} a_{kj}x_j+b_k$.
- $f:D\to\mathbb{R}^m$ with $D\subseteq\mathbb{R}^n$. D is an open set.
- $f(x)=g(x)+o(|x-y|)$ as $x\to y$.

  $\equiv\quad \lim_{x\to y}\dfrac{f(x)-g(x)}{|x-y|}=0\in\mathbb{R}^m$

  $\equiv\quad \forall e>0\ \exists d>0$ such that $\forall x\in D$, $|x-y|<d \Rightarrow |f(x)-g(x)|\leq |x-y|e$.
- If there exists an affine function $g(x)$ such that $f(x)=g(x)+o(|x-y|)$ as $x\to y$, then it is unique.
- Zero is the only affine function such that $g(x)=o(|x-y|)$ as $x\to y$.
- Def: $f:D\to\mathbb{R}^m$, $y\in D$. $g(x)=ax+b$ is a best affine approximation of f at y provide that $f(x)=g(x)+o(|x-y|)$ as $x\to y$.
  - $\Rightarrow b=f(y)$.
- Def: $f:D\to\mathbb{R}^m$ is **differentiable** at $y\in D$
  - if there exists an $m\times n$ matrix $df(y)$, called the **differential** of f at y, such that
    - $f(x)=f(y)+df(y)(x-y)+o(|x-y|)$ as $x\to y$.

    $\equiv\quad \lim_{x\to y}\dfrac{f(x)-f(y)-df(y)(x-y)}{|x-y|}=0$

    $\equiv\quad \lim_{x\to y}\dfrac{|f(x)-f(y)-df(y)(x-y)|}{|x-y|}=0$

$\equiv \quad \forall e > 0\ \exists d > 0$ such that $|x - y| < d \Rightarrow |f(x) - f(y) - df(y)(x - y)| \le |x - y| e$.

$\equiv$ each of the coordinate function $f_k : D \to \mathbb{R}$ is differentiable at *y*.

- Differentiability $\Rightarrow$ continuity

  Proof. $\lim_{x\to y} f(x) = \lim_{x\to y}\left(f(y) + df(y)(x - y) + o(|x - y|)\right) = f(y)$.

- If *f* is real valued, then $df(y)$ is $1\times n$ (row vector). It is sometimes called the gradient of *f* and written $\nabla f(y)$.
- Def: If *f* is differentiable at every point of *D*, we say *f* is differentiable on *D*.

  We can regard the differential $df(y)$ as a function of *y*, taking values in the space of $m\times n$ matrices $\mathbb{R}^{m\times n}$.

  If $df : D \to \mathbb{R}^{m\times n}$ is continuous, we say *f* is continuously differentiable or *f* is C^1.
- Differentiability and differential are linear. If both $f : D \to \mathbb{R}^m$ and $g : D \to \mathbb{R}^m$ are differentiable at *y*, then so is $af + bg$ for scalars *a*, *b* and $d(af + bg) = adf + bdg$.
- If $f : D \to \mathbb{R}^m$ and $g : D \to \mathbb{R}$ are differentiable at *y*, then so is $g \cdot f$, and $d(g \cdot f)(y) = g(y)df(y) + f(y)dg(y)$.
- The **partial derivative** $\frac{\partial f_k}{\partial x_j} \in \mathbb{R}$ is said to exist at a point *y*
  - Def: if $f_k(y + te_j) = f_k(y) + \frac{\partial f_k}{\partial x_j}(y)t + o(t)$ as $t \to 0$.

  $\equiv \quad f_k(y + te_j) = f_k(y) + \frac{\partial f_k}{\partial x_j}(y)t + o(|t|)$ as $t \to 0$.

  $\equiv \quad f_k(y + te_j)$ as a function of *t* is differentiable at *t* = 0 with derivative $[df(y)]_{kj}$.

  $= \quad \lim_{t\to 0} \frac{f_k(y + te_j) - f_k(y)}{t}$.

  - Obtained by keeping all the variables $x_1, \ldots, x_n$ except x_j fixed and differentiating f_k as a function of x_j.
- Def: the **partial derivative of *f* with respect to x_j**:

  $$\frac{\partial f}{\partial x_j}(y) = d_{e_j} f(y) = \lim_{t\to 0}\frac{f(y + te_j) - f(y)}{t} : D \to \mathbb{R}^m$$
- If nonzero $u \in \mathbb{R}^n$, the **directional derivative** $d_u f \in \mathbb{R}^m$ is said to exist at a point *y*
  - Def: if $f(y + tu) = f(y) + d_u f(y)t + o(t)$ as $t \to 0$.

$= \lim_{t\to 0}\frac{f(y+tu)-f(y)}{t}$.

- If $u = 0$, then $d_u f = 0$.

- Def: $d_u f_k(y) = [d_u f(y)]_k$ = the k^{th} component of $d_u f(y)$.
- If f is differentiable at y, then all partial and directional derivative exists at y

$$df(y)=\begin{pmatrix}\nabla f_1(y)\\ \vdots\\ \nabla f_m(y)\end{pmatrix}=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(y) & \cdots & \frac{\partial f_1}{\partial x_n}(y)\\ \vdots & \ddots & \vdots\\ \frac{\partial f_m}{\partial x_1}(y) & \cdots & \frac{\partial f_m}{\partial x_n}(y)\end{pmatrix}=\left(\frac{\partial f}{\partial x_1}(y),\dots,\frac{\partial f}{\partial x_n}(y)\right)$$

$$\frac{\partial f}{\partial x_j}(y)=d_{e_j} f(y)=df(y)e_j = \text{ the } j^{\text{th}} \text{ column of } df(y)$$

$$\frac{\partial f_k}{\partial x_j}=[df(y)]_{kj}=[d_{e_j} f(y)]_k$$

$$d_u f(y)=df(y)u=df(y)\left(\sum_{j=1}^{n}u_j e_j\right)=\sum_{j=1}^{n}u_j\, df(y)e_j=\sum_{j=1}^{n}u_j\frac{\partial f}{\partial x_j}(y).$$

- $f(x,y)=\frac{xy}{\sqrt{x^2+y^2}}$ has directional derivatives in all directions at all points in the plane but is not differentiable at the origin.
- Let $f: D\to\mathbb{R}^m$ with $D\subset\mathbb{R}^n$ have partial derivatives $\frac{\partial f}{\partial x_j}: D\to\mathbb{R}^m$ for $j = 1, \dots, n$ that are continuous in a neighborhood of y. <u>Then</u>, f is differentiable at y.

- Let $f: D\to\mathbb{R}^m$ with $D\subset\mathbb{R}^n$ open, $y\in D$. If $\forall k\forall j$ partial derivatives $\frac{\partial f_k}{\partial x_j}$ exists in a neighborhood of y and continuous at y, then f is differentiable at y. And df is continuous at y.

- A function $f: D\to\mathbb{R}^m$ with $D\subset\mathbb{R}^n$ open is C^1 <u>iff</u> the partial derivatives $\left(\frac{\partial f}{\partial x_j}: D\to\mathbb{R}^m \text{ for } j=1,\dots,n\right)$ exist and are continuous on D.
- Note
  - $f: D\to\mathbb{R}^m$ is differentiable at a point $(df(y) \text{ exists})$ if and only if each of the coordinate functions $f_k: D\to\mathbb{R}$ is differentiable at that point $\left(df_k(y)=\underset{1\times n}{\nabla f_k(y)} \text{ exists}\right)$.
  - $f: D\to\mathbb{R}^m$ is continuous if and only if all the $f_k: D\to\mathbb{R}$ are continuous.

- $f: D \to \mathbb{R}^m$, open $D \subset \mathbb{R}^n$ is C^1 iff all $f_k: D \to \mathbb{R}$ are C^1.

  Proof. "$\Rightarrow$" df exists and continuous $\Rightarrow$ $\forall x \forall k \forall j$ $\frac{\partial f_k}{\partial x_j}(x)$ exists and continuous $\Rightarrow$ for each k, $\forall x \forall j$ $\frac{\partial f_k}{\partial x_j}(x)$ exists and continuous. Hence, df_k exists and continuous.

- $\frac{\partial f}{\partial x_j}: D \to \mathbb{R}^m$ exists $\Leftrightarrow$ $\frac{\partial f_k}{\partial x_j}$ exists $\forall k = 1, \ldots, m$.

  $$\frac{\partial f}{\partial x_j}(y) = \begin{pmatrix} \frac{\partial f_1}{\partial x_j} \\ \vdots \\ \frac{\partial f_m}{\partial x_j} \end{pmatrix}$$ (always, need not assume continuity.)

  Proof. Let $g(t) = \frac{f(y + te_j) - f(y)}{t}$. Then,

  $$\frac{\partial f}{\partial x_j}(y) = \lim_{t \to 0} \frac{f(y + te_j) - f(y)}{t} = \lim_{t \to 0} g(t), \text{ and}$$

  $$\frac{\partial f_k}{\partial x_j}(y) = \lim_{t \to 0} \frac{f_k(y + te_j) - f_k(y)}{t} = \lim_{t \to 0} g_k(t).$$

  Then, use $\lim_{t \to 0} g(t) = w \Leftrightarrow \lim_{t \to 0} g_k(t) = w_k$.

- $\frac{\partial f}{\partial x_j}: D \to \mathbb{R}^m$ exists and is continuous if and only if $\frac{\partial f_k}{\partial x_j}$ is exists and is continuous $\forall k = 1, \ldots, m$.

  Proof. Existence relationship follows from above. Continuity relationship follows because $\frac{\partial f_k}{\partial x_j}$ is the k^{th} component of $\frac{\partial f}{\partial x_j}$ by the formula given above.

- Thus, theorem(s) using existence or continuity of $\frac{\partial f}{\partial x_j}$ is equivalent to theorem using existence or continuity of df $\forall k = 1, \ldots, m$.

- **Pointwise Lipschitz condition**: If f is differentiable at y, then $\exists d > 0$ and constant M_y (depending on y) such that $|x - y| < d \Rightarrow |f(x) - f(y)| \le M_y |x - y|$.

Proof. $\left|f(x)-f(y)\right|=\left|df(y)(x-y)+o\left(\left|x-y\right|\right)\right|\leq\left|df(y)(x-y)\right|+\left|o\left(\left|x-y\right|\right)\right|$ $\leq c\left|x-y\right|+\left|x-y\right|$. The inequality of the second part of the sum requires that *x* and *y* are close enough.

- Let $f:D\to\mathbb{R}^m$ with $D\subseteq\mathbb{R}^n$ and $g:A\to\mathbb{R}^p$ with $f(D)\subset A\subset\mathbb{R}^m$.

  $g\circ f:D\to\mathbb{R}^p$ is defined by $g\circ f(x)=g(f(x))$ for $x\in D$.

- Recall, $g(t):\mathbb{R}\to\mathbb{R}$. $g'(t_0)$ exists. $g(t)$ attains its max or min at $t=t_0$, then $g'(t_0)=0$.

  Proof. Assume max. $g'(t_0)=\lim_{t\to t_0}\frac{g(t)-g(t_0)}{t-t_0}$ exists. $\lim_{t\to t_0^+}\frac{\overbrace{g(t)-g(t_0)}^{\leq 0}}{\underbrace{t-t_0}_{>0}}\leq 0$, and

  $\lim_{t\to t_0^-}\frac{\overbrace{g(t)-g(t_0)}^{\leq 0}}{\underbrace{t-t_0}_{<0}}\geq 0$. Hence, $\lim_{t\to t_0}\frac{g(t)-g(t_0)}{t-t_0}=0$.

- Let $f:D\to\mathbb{R}$ for $D\subseteq\mathbb{R}^n$, and let *y* be a point in the interior of *D*.
  If *f* assumes its maximum or minimum value at *y* and *f* is differentiable at *y*, then
  $df(y)=\nabla f(y)=0$ $\left(\frac{\partial f}{\partial x_k}(y)=0\ \forall\ k=1,\ldots,n\right)$.

  Proof. $g(t)=f\left(y+te_j\right)$ is differentiable at 0 because $g'(0)=\lim_{t\to 0}\frac{g(t)-g(0)}{t}=$ $\lim_{t\to 0}\frac{f\left(y+te_j\right)-f(y)}{t}=d_{e_j}f(y)$, and *f* is differentiable at *y*. Also, $d_{e_j}f(y)=\frac{\partial f}{\partial x_j}(y)$.

  Because $g(t)$ attains its max or min at $t=0$, $g'(0)=0$.

- Let $f(x)=Ax+b$, then $df(x)=A$.

  Proof. $\lim_{x\to x_0}\frac{f(x)-f(x_0)-A(x-x_0)}{\left|x-x_0\right|}=\lim_{x\to x_0}\frac{Ax+b-Ax_0-b-A(x-x_0)}{\left|x-x_0\right|}=\lim_{x\to x_0}\frac{0}{\left|x-x_0\right|}$ $=0$.

- If $f:D\to\mathbb{R}^m$ with $D\subset\mathbb{R}^n$ is differentiable, then we regard ***df* as a function** $df:D\to\mathbb{R}^{m\times n}$ taking values in the space of $m\times n$ matrices.

- *f* is differentiable at *y*. $h(w,z)=o\left(\left|w-z\right|\right)$ as $w\to z$.

  Then, $h\left(f(x),f(y)\right)=o\left(\left|x-y\right|\right)$ as $x\to y$.

Proof. By the differentiability at y, $|x-y|<d \Rightarrow |f(x)-f(y)| \le M_y|x-y|$. From def, $h(f(x),f(y)) = o(|f(x)-f(y)|)$ as $f(x)\to f(y)$. By continuity of f, $x\to y$ implies $f(x)\to f(y)$. So, $\lim\limits_{x\to y}\dfrac{h(f(x),f(y))}{|f(x)-f(y)|} = \lim\limits_{f(x)\to f(y)}\dfrac{h(f(x),f(y))}{|f(x)-f(y)|} = 0$. Thus,

$$\lim_{x\to y}\frac{h(f(x),f(y))}{|x-y|} = \lim_{x\to y}\frac{h(f(x),f(y))}{|x-y|} = \lim_{x\to y}\underbrace{\frac{h(f(x),f(y))}{|f(x)-f(y)|}}_{\to 0}\underbrace{\frac{|f(x)-f(y)|}{|x-y|}}_{\text{bounded for } x,y \text{ close enough}}. = 0.$$

- **Chain rule**: If f is differentiable at y and g is differentiable at $z=f(y)$, then $g\circ f$ is differentiable at y and $\underset{p\times n}{d(g\circ f)(y)} = \underset{p\times m}{dg(z)}\,\underset{m\times n}{df(y)}$ (matrix multiplication.)

  - If g is real-valued, then

    $$\underset{1\times n}{d(g\circ f)(y)} = \underset{1\times m}{dg(z)}\,\underset{m\times n}{df(y)}$$

    $$\frac{\partial}{\partial x_j}(g\circ f)(y) = \sum_{k=1}^{m}\frac{\partial g}{\partial z_k}(z)\frac{\partial f_k}{\partial x_j}(y).$$

- Ex. $f:\mathbb{R}^n\to\mathbb{R}$, $df(y)$ exists, $g(t) = f\left(y+te^{(j)}\right):\mathbb{R}\to\mathbb{R}$, then,

  $dg(t) = df\left(y+te^{(j)}\right)e^{(j)} = \dfrac{\partial f}{\partial x_j}\left(y+te^{(j)}\right)$. Hence, $dg(0) = df(y)e^{(j)} = \dfrac{\partial f}{\partial x_j}(y)$.

  Alternatively, can have $g'(0) = \lim\limits_{s\to 0}\dfrac{f\left(y+se^{(j)}\right)-f\left(y+0e^{(j)}\right)}{s} = \lim\limits_{s\to 0}\dfrac{f\left(y+se^{(j)}\right)-f(y)}{s} = \dfrac{\partial f}{\partial x_j}(y)$ (by definition of partial derivative); exists because $df(y)$ exists.

- If f and g are differentiable (respectively C^1) on their domains, then so is $g\circ f$.
- If f and g are differentiable on their domains, $g\circ f$ is differentiable on its domain.
- If dg and df are continuous, $d(g\circ f)$ is continuous.

- **MVT0: Mean Value theorem**:

  Let open $\Omega\subset\mathbb{R}^n$, $f:\Omega\to\mathbb{R}$ differentiable.

  $[a,b]\subset\Omega \Rightarrow \exists c\in(a,b)$ such that $f(b)-f(a) = df(c)(b-a) = \nabla f(c)\cdot(b-a)$.

Proof. Let $u=b-a$. Define a real-valued function $g(t) = f(a+tu)$. Then, by chain rule, $g'(t) = df(a+tu)u$, exists. This is true $\forall t\in[0,1]$.

By the mean value theorem, $\exists t_0 \in (0,1)$ such that $g'(t_0) = \dfrac{g(1)-g(0)}{1-0} = f(b) - f(a)$. Let $c = a + t_0(b-a) = a + t_0 u$. Note that $c \in (a,b)$ because $t_0 \in (0,1)$. Hence, $\exists c \in (a,b)$ such that $g'(t_0) = df(c)(b-a) = f(b) - f(a)$.

- Let $f : \mathbb{R}^n \to \mathbb{R}^m$, then
  1) $df \equiv 0$ iff f is constant
  2) df is constant if and only if f is an affine function $(Ax+b)$.

Proof 1): "$\Leftarrow$" f_k is constant. $\forall x \forall k \forall j$ $\dfrac{\partial f_k}{\partial x_j}(x) = 0$, continuous. "$\Rightarrow$" Consider $g(x) = f_k(x)$. f is differentiable $\Rightarrow$ f_k is differentiable $\Rightarrow$ g is differentiable, $dg(x) = df_k(x) = 0$ $\forall x$. Consider any $x, y \in \Omega$, $x \neq y$. Then, by MVT0, $\exists z_0 \in (x,y)$ $f(y) - f(x) = \underbrace{df(z_0)}_{0 \text{ for any } z_0}(y-x) = 0$. Hence, f_k is constant. This is true $\forall k$.

Proof 2): "$\Leftarrow$" Let $f(x) = Ax + b$, then $df(x) = A$ constant. "$\Rightarrow$" Let $f(0) = b$. Consider $g(x) = f_k(x)$. So, $g(0) = b_k$. $dg(x) = df_k(x) = a^{(k)}$, the k^{th} row of A. By MVT0, $\exists z_0 \in (0,x)$ $g(x) - g(0) = \underbrace{dg(z_0)}_{a^{(k)} \text{ for any } z_0}(x-0) = a^{(k)}x$. Therefore, $f_k(x) = g(x) = a^{(k)}x + b_k$.

- Let $f : [a,b] \to \mathbb{R}^m$ be continuous. Also, $\forall t \in (a,b)$ $\underset{m\times 1}{df(t)} = 0$. Then, $f(a) = f(b)$.

Proof. Let $z = f(b) - f(a)$, and $g(t) = z \cdot f(t) = \sum_{i=1}^{m} z_i f_i(t) : [0,1] \to \mathbb{R}$. Because $f(t)$ is continuous on in $[a,b]$, $g(t)$ is also continuous on $[a,b]$. Also, because $\forall t \in (a,b)$ $df(t) = 0$, we also have $\dfrac{\partial f_i}{\partial t}(t) = f_i'(t) = 0$. Thus, $g'(t) = \sum_{i=1}^{m} z_i f_i'(t) = 0$ $\forall t \in (a,b)$. By the mean value theorem, $\exists t_0 \in (a,b)$ such that $\dfrac{g(b)-g(a)}{a-b} = g'(t_0) = 0$. Hence, $g(b) - g(a) = 0$. Note that

$$g(b) - g(a) = \big(f(b) - f(a)\big)\cdot f(b) - \big(f(b) - f(a)\big)\cdot f(a)$$
$$= \big(f(b) - f(a)\big)\cdot\big(f(b) - f(a)\big) = \left|f(b) - f(a)\right|^2$$

Thus, $f(b) = f(a)$.

- Def: Let (a,b) denote the line segment joining a and b. The points on this line segment can be expressed as $v = (1-t)a + tb = a + t(b-a)$, $t \in [0,1]$.

- Def: We say a set $B \subset \mathbb{R}^n$ is **convex** if $\forall x, y \in B \; \forall l \in [0,1]$ we have $l\,x + (1-l)\,y \in B$.
    - Open balls are convex. Consider $B = B_r(x_0)$. Then $\forall x, y \in B \;\; \forall l \in [0,1]$

$$\begin{aligned} \left| l\,x + (1-l)\,y - x_0 \right| &= \left| l\,(x - x_0) + (1-l)(y - x_0) \right| \\ &\le l\left|x - x_0\right| + (1-l)\left|y - x_0\right| < l\,r + (1-l)\,r = r \end{aligned}$$

- Let $f : D \to \mathbb{R}^m$ where $D \subset \mathbb{R}^n$ is open and convex. Then, $df(x) = 0 \; \forall x \in D \Rightarrow f$ is constant.

Proof. Let x be any point in D. Let $f(x) = a$. Consider any y in D.
Define $g(t) = (1-t)x + ty : [0,1] \to \mathbb{R}^n$. Because D is convex, $\forall t \in [0,1]$, $g(t) \in D$. Let $h(t) = f(g(t)) : [0,1] \to \mathbb{R}^m$. Then, $\forall t \in [0,1]$ $\underset{m\times 1}{dh(t)} = df(g(t))\,dg(t) = 0$ because $df(g(t)) = 0$. This implies $h(0) = h(1)$, or equivalently, $f(x) = f(y)$.

Proof. Let x be any point in D. Let $f(x) = a$. Consider any y in D. Because D is convex, D contains the line segment joining x and y.
For each k in $\{1,\dots,m\}$, consider $f_k : D \to \mathbb{R}$. $df(x) = 0 \; \forall x \in D \Rightarrow$
$\nabla f_k = \begin{pmatrix} \frac{\partial f_k}{\partial x_1} & \cdots & \frac{\partial f_k}{\partial x_n} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 \end{pmatrix}$. So, f_k is C^1, and $\exists z$ on the line joining x and y such that $f_k(x) = f_k(y) + \nabla f(z) \cdot (y - x)$. For any value of z on the line segment, $\nabla f(z) = 0$. So, $f_k(x) = f_k(y)$. This is true for all k, so $f(x) = f(y)$.

- Let $f : D \to \mathbb{R}^m$ where $D \subset \mathbb{R}^n$ is open and connected (so arcwise-connected). Then, $df(x) = 0 \; \forall x \in D \Rightarrow f$ is constant.

Let x be any point in D. Let $f(x) = a$. Consider any y in D.

Let $A = \{ t \in [0,1];\; f(g(t)) = a \}$, and $t_0 = \sup A$.

$0 \le t_0 \le 1$ because $0 \in A$ $\left(f(g(0)) = f(x) = a \right)$, and 1 is an upper bound of A. Claim: $f(g(t_0)) = a$.

Because t_0 is the sup of $A \subset [0,1]$, $\exists$ sequence $\{t_n\}$ in $A \subset [0,1]$ converging to t_0. Because f and g are continuous, $f \circ g$ is continuous. Thus, $\lim_{n\to\infty} f \circ g(t_n) = f \circ g\left(\lim_{n\to\infty} t_n \right) = f \circ g(t_0) = f(g(t_0))$. Because $\lim_{n\to\infty} f \circ g(t_n) = \lim_{n\to\infty} f(g(0)) = f(g(0)) = a$, we conclude that $f(g(t_0)) = a$.

Claim: $t_0 = 1$.

Assume $0 \le t_0 < 1$, then because D is an open set in $\mathbb{R}^n$, $\exists r$ such that $B_r\left(g\left(t_0\right)\right) \subset D$. (If $r > \left|y - g\left(t_0\right)\right|$, set $r = \left|y - g\left(t_0\right)\right|$, and $B_r\left(g\left(t_0\right)\right) \subset D$, still.) Because $B_r\left(g\left(t_0\right)\right)$ is convex, and $\forall x \in B_r\left(g\left(t_0\right)\right)$, $df\left(x\right) = 0$, we conclude that $\forall x \in B_r\left(g\left(t_0\right)\right)$, $f\left(x\right) = f\left(g\left(t_0\right)\right) = a$. By continuity of g, $\exists d$ such that $\forall t \in [0,1]\ \left|t - t_0\right| < d \Rightarrow \left|g(t) - g\left(t_0\right)\right| < r$. Hence, $\exists t' \in [0,1]$, $t' > t_0$, $g\left(t'\right) \in B_r\left(g\left(t_0\right)\right)$, which implies $f\left(g\left(t'\right)\right) = a$; so, $t' > t_0 \in A$. This contradict the assumption that $t_0 = \sup A$.

We have shown that $f\left(g\left(t_0\right)\right) = a$. Because $t_0 = 1$, $f\left(g(1)\right) = f\left(y\right) = a$.

- If D is not connected, then f may not be constant. Ex. Let $D = (0,1) \cup (2,3)$, not connected because it is not an interval. Let $f\left(x\right) = \begin{cases} 0 & x \in (0,1) \\ 1 & x \in (2,3) \end{cases} : D \to \mathbb{R}$. Then, $f'\left(x\right) = 0\ \ \forall x \in D$, but $f(x)$ is not constant.

Differentiating a general function defined by an integral

- If $g : \mathbb{R}^2 \to \mathbb{R}$ is continuous, then $G\left(x\right) = \int_a^b g\left(x,y\right) dy$ is continuous.

  Proof. Consider at x_0. g is continuous; thus, uniformly continuous on compact $\left[x_0 - d', x_0 + d'\right] \times [a,b]$. Thus, given $e > 0$, $\exists d > 0$ such that $\forall x \in \left[x_0 - d', x_0 + d'\right]$ $\forall y \in [a,b]$ $\left|x - x_0\right| < d \Rightarrow \left|(x,y) - \left(x_0, y\right)\right| < d \Rightarrow \left|g\left(x,y\right) - g\left(x_0, y\right)\right| < e$. So,

  $$\left|\sum_{j=1}^{n} g\left(x, y_j\right)\Delta y_j - \sum_{j=1}^{n} g\left(x_0, y_j\right)\Delta y_j\right| \le \sum_{j=1}^{n}\left|g\left(x, y_j\right) - g\left(x_0, y_j\right)\right|\Delta y_j \le \sum_{j=1}^{n} e\Delta y_j = e\left(b-a\right).$$

  Taking the limit as the max interval length of the partition goes to zero, the sum become integrals, and $\left|G\left(x\right) - G\left(x_0\right)\right| \le e\left(b-a\right)$.

- $g : \mathbb{R}^2 \to \mathbb{R}$ is C^1. Let $h_n \to 0$. $\forall x_0$ $G_n\left(x_0, y\right) = \dfrac{g\left(x_0 + h_n, y\right) - g\left(x_0, y\right)}{h_n}$. Then,

  $G_n\left(x_0, y\right) \xrightarrow{\text{uniformly}} \dfrac{\partial g}{\partial x}\left(x_0, y\right)$ over $y \in [a,b]$. Hence, $\lim\limits_{n\to\infty} \displaystyle\int_a^b \dfrac{g\left(x + h_n, y\right) - g\left(x,y\right)}{h_n} dy = \int_a^b \dfrac{\partial g}{\partial x}\left(x,y\right) dy$.

  Proof. For a given x_0, let $H_n\left(y\right) = G_n\left(x_0, y\right)$. Then, for uniform convergence of $H_n\left(y\right)$ to $\dfrac{\partial g}{\partial x}\left(x_0, y\right)$ over $y \in [a,b]$, need $\forall e > 0\ \exists N \in \mathbb{N}\ \ \forall y \in [a,b]$ $\left|H_n\left(y\right) - \dfrac{\partial g}{\partial x}\left(x_0, y\right)\right| \le e$.

g is C^1. By the mean value theorem, $\forall y \forall n\ \exists z_{n,y}\ x_0 < z_{n,y} < x_0 + h$

$$G_n(x_0, y) = \frac{g(x_0 + h_n, y) - g(x_0, y)}{h_n} = \frac{\partial g}{\partial x}(z_{n,y}, y).$$

Given $e > 0$. Note that $\left| H_n(y) - \frac{\partial g}{\partial x}(x_0, y) \right| = \left| \frac{\partial g}{\partial x}(z_{n,y}, y) - \frac{\partial g}{\partial x}(x_0, y) \right|$. By the uniform continuity of $\frac{\partial g}{\partial x}(x,y)$ on compact $[x_0 - d', x_0 + d'] \times [a,b]$, $\exists d > 0\ \ \forall y \in [a,b]$ $\left| (z_{n,y}, y) - (x_0, y) \right| < d \Rightarrow \left| \frac{\partial g}{\partial x}(z_{n,y}, y) - \frac{\partial g}{\partial x}(x_0, y) \right| < e$. Note that $\left| z_{n,y} - x_0 \right| < h_n$, and $h_n \to 0$; thus, $\exists N \in \mathbb{N}$, $\forall n \geq N\ \ \left| z_{n,y} - x_0 \right| < d$.

- Recall:
  - $f_n(x) \xrightarrow{\text{uniformly}} f(x)$ iff $\forall e > 0\ \exists N \in \mathbb{N}\ \ \forall x \in D\ \ \forall k \geq N\ \ \left| f_k(x) - f(x) \right| \leq e$.
  - $f_n(x) \xrightarrow{\text{uniformly}} f(x)$ on $[a,b] \Rightarrow \lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx$.
- If $g : \mathbb{R}^2 \to \mathbb{R}$ is C^1, then $F(x) = \int_a^b g(x,y)\,dy$ is C^1 with $F'(x) = \int_a^b \frac{\partial g}{\partial x}(x,y)\,dy$.

  Proof. $F'(x) = \lim_{h\to0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to0} \frac{1}{h}\left(\int_a^b g(x+h,y)dy - \int_a^b g(x,y)dy \right) =$ $\lim_{h\to0} \int_a^b \frac{g(x+h,y) - g(x,y)}{h} dy$. By above, we have any sequence $h_n \to 0$ $\lim_{n\to\infty} \int_a^b \frac{g(x+h_n,y) - g(x,y)}{h_n} dy = \int_a^b \frac{\partial g}{\partial x}(x,y)dy$. Hence, $F'(x) = \int_a^b \frac{\partial g}{\partial x}(x,y)dy$.

  Because $\frac{\partial g}{\partial x}(x,y)$ is continuous, $\int_a^b \frac{\partial g}{\partial x}(x,y)dy$ is continuous.

- Let $g : \mathbb{R}^2 \to \mathbb{R}$, $a : \mathbb{R} \to \mathbb{R}$, and $b : \mathbb{R} \to \mathbb{R}$ be C^1. Then $f(x) = \int_{a(x)}^{b(x)} g(x,y)dy$ is C^1 and

$$f'(x) = b'(x)\,g(x, b(x)) - a'(x)\,g(x, a(x)) + \int_{a(x)}^{b(x)} \frac{\partial g}{\partial x}(x,y)dy.$$

Proof. Consider $F(x_1, x_2, x_3) = \int_{a(x_2)}^{b(x_1)} g(x_3, y)dy$. Then, by the 1-D chain rule and the fundamental theorem of the calculus (differentiation of the integral)

$$\frac{\partial}{\partial x_1} F(\bar{x}) = g(x_3, b(x_1))\,b'(x_1) \text{ and } \frac{\partial}{\partial x_2} F(\bar{x}) = -g(x_3, a(x_2))\,a'(x_2).$$

Note that both are continuous. Also,

$$\frac{\partial}{\partial x_3}F(\vec{x}) = \int_{a(x_2)}^{b(x_1)} \frac{\partial}{\partial x_3} g(x_3, y)dy = \int_{a(x_2)}^{b(x_1)} \frac{\partial}{\partial x} g(x, y)dy \text{, continuous.}$$

Let $h(x) = \begin{pmatrix} x \\ x \\ x \end{pmatrix} \Rightarrow dh(x) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

Thus, $f(x) = F(h(x)) = F(x, x, x) \Rightarrow$

$$df(x) = dF(h(x))dh(x) = \sum_{i=1}^{3} \frac{\partial}{\partial x_i} F(h(x)) = \sum_{i=1}^{3} \frac{\partial}{\partial x_i} F(x, x, x).$$
