Continuous Mathematics Memorandum

1 Derivatives and Taylor's theorem

- A function f is **continuous** at x if
$$\lim_{h\to 0} f(x+h) = f(x).$$
f is continuous if it is continuous at every point of its domain.

- A function $f : D \to \mathbb{R}$, where $D \subseteq \mathbb{R}$, is called **differentiable** at x if
$$\lim_{h\to 0} \frac{f(x+h) - f(x) - dh}{h} = 0$$
holds for some $d \in \mathbb{R}$. The value d is called the derivative at x. f is called differentiable if it is differentiable at every point of D.

- If f and g are continuous (differentiable*), then so are:

  $f + g$, *
  cf, for a constant $c \in \mathbb{R}$, *
  f^n, for a constant $n \in \mathbb{N}$, *
  fg, *
  $f \circ g$, *
  $\max(f, g)$,
  $|f|$,
  $\exp(f)$, *
  f^α, for a constant $\alpha \in \mathbb{R}$, **where f is strictly positive**, *
  $\frac{f}{g}$, **where g is nonzero**, *
  $\log(f)$, **where f is strictly positive**. *

- Rules for differentiation:

  $\frac{d}{dx}(f+g) = \frac{df}{dx} + \frac{dg}{dx}$,
  $\frac{d}{dx}(cf) = c\frac{df}{dx}$,
  $\frac{d}{dx}(fg) = f\frac{dg}{dx} + g\frac{df}{dx}$,
  $\frac{d}{dx}(\frac{f}{g}) = \frac{g\frac{df}{dx} - f\frac{dg}{dx}}{g^2}$,
  $\frac{d}{dx}(g \circ f) = (\frac{dg}{dx} \circ f)\frac{df}{dx}$.

- **Taylor's Theorem:** Let $f : D \to \mathbb{R}$ be a function, $k \geq 0$ an integer and $x_0, x_0 + h \in D$. Then, if f and its first $k+1$ derivatives exist and are continuous on an interval containing x_0 and $x_0 + h$, the following holds:
$$f(x_0+h) = f(x_0) + h\frac{df}{dx}(x_0) + \frac{h^2}{2!}\frac{d^2 f}{dx^2}(x_0) + \cdots + \frac{h^k}{k!}\frac{d^k f}{dx^k}(x_0) + \frac{h^{k+1}}{(k+1)!}\frac{d^{k+1} f}{dx^{k+1}}(\xi),$$

for some $\xi \in (x0, x0 + h)$.
Equivalently, if we let $h = x - x_0$, we can write

$$f(x) = \hat{f}_k(x) + e_{k+1}(x, x_0),$$

where $\hat{f}_k(x)$ is the **Taylor polynomial of order** k,

$$\hat{f}_k(x) = \sum_{i=0}^{k} \frac{(x - x_0)^k}{k!} \frac{d^k f}{dx^k}(x_0),$$

and $e_{k+1}(x, x_0)$ is the **error term**,

$$e_{k+1}(x, x_0) = \frac{(x - x_0)^{k+1}}{(k+1)!} \frac{d^{k+1} f}{dx^{k+1}}(\xi).$$

We can Bound the approximation: if $\underline{C} \leq \frac{d^{k+1} f}{dx^{k+1}} \leq \overline{C}$, then

$$\hat{f}_k(x) + \underline{C} \leq f(x) \leq \hat{f}_k(x) + \overline{C}.$$

- A function $f : D \to \mathbb{R}$, for some $D \subseteq \mathbb{R}^n$, is called a **multivariate function** or a **scalar field**. Differentiating f with respect to a variable x, while holding all others constant, is called a **partial derivative** and written $\frac{\partial f}{\partial x}$.
- **Clairaut's theorem:** if f is continuous in x and y, then $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$.
- We can collect all the partial derivatives in a vector and define the derivative of the function f:

$$\frac{df}{d\boldsymbol{x}} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

- The **Hessian** of f, which is the equivalent of the second derivative, is a matrix that collects all the second partial derivatives:

$$\boldsymbol{H}(f) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

- A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is sometimes called a **vector field**. We can break down f in the form of a vector, $\begin{pmatrix} f_1(\boldsymbol{x}) \\ f_2(\boldsymbol{x}) \\ \vdots \\ f_n(\boldsymbol{x}) \end{pmatrix}$. What acts as the derivative of f is called the **Jacobian**:

$$\boldsymbol{J}(\boldsymbol{f}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

.

- For $f : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{H}(f) = \boldsymbol{J}(\frac{df}{dx})^T$

- Rules for differentiation of multivariate functions:

  $\frac{d}{d\boldsymbol{x}}(f+g) = \frac{df}{d\boldsymbol{x}} + \frac{dg}{d\boldsymbol{x}}$,
  $\frac{d}{d\boldsymbol{x}}(cf) = c\frac{df}{d\boldsymbol{x}}$,
  $\frac{d}{d\boldsymbol{x}}(fg) = f\frac{dg}{d\boldsymbol{x}} + g\frac{df}{d\boldsymbol{x}}$,
  $\frac{d}{d\boldsymbol{x}}(\frac{f}{g}) = \frac{g\frac{df}{d\boldsymbol{x}} - f\frac{dg}{d\boldsymbol{x}}}{g^2}$,
  If $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ then,
  $\frac{d}{d\boldsymbol{x}}(g \circ f) = (\frac{dg}{d\boldsymbol{x}} \circ f)\frac{df}{d\boldsymbol{x}}$,
  If $\boldsymbol{f} : \mathbb{R}^m \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}$ then,
  $\frac{d}{d\boldsymbol{x}}(g \circ \boldsymbol{f}) = \boldsymbol{J}(\boldsymbol{f})^T(\frac{dg}{d\boldsymbol{x}} \circ \boldsymbol{f})$,
  $\boldsymbol{J}(\boldsymbol{f}+\boldsymbol{g}) = \boldsymbol{J}(\boldsymbol{f}) + \boldsymbol{J}(\boldsymbol{g})$,
  $\boldsymbol{J}(c\boldsymbol{f}) = c\boldsymbol{J}(\boldsymbol{f})$,
  $\boldsymbol{J}(\boldsymbol{A}\boldsymbol{f}) = \boldsymbol{A}\boldsymbol{J}(\boldsymbol{f})$,
  $\boldsymbol{J}(\boldsymbol{f}^T\boldsymbol{g}) = \boldsymbol{g}^T\boldsymbol{J}(\boldsymbol{f}) + \boldsymbol{f}^T\boldsymbol{J}(\boldsymbol{g})$,
  $\boldsymbol{J}(f\boldsymbol{g}) = \boldsymbol{g}\frac{df}{d\boldsymbol{x}}^T + f\boldsymbol{J}(\boldsymbol{g})$,
  $\boldsymbol{J}(\boldsymbol{g} \circ \boldsymbol{f}) = (\boldsymbol{J}(\boldsymbol{g}) \circ \boldsymbol{f})\boldsymbol{J}(\boldsymbol{f})$.

- Standard derivatives:

  $\frac{d}{d\boldsymbol{x}}\boldsymbol{a}^T\boldsymbol{x} = \frac{d}{d\boldsymbol{x}}\boldsymbol{x}^T\boldsymbol{a} = \boldsymbol{a}$,
  $\frac{d}{d\boldsymbol{x}}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{A} + \boldsymbol{A}^T)\boldsymbol{x}$,
  $\boldsymbol{J}(\boldsymbol{x}) = \boldsymbol{I}$,
  $\boldsymbol{J}(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{A}$.

- **Taylor's theorem (multivariate functions):** Let $D \subseteq \mathbb{R}^n$. Fix a vector $\boldsymbol{x_0}$ and $\boldsymbol{h} = (h_1, h_2, \ldots, h_n)$ be another vector. Let $k \geq 0$ be an integer. If f and its first $k+1$ derivatives exist and are continuous on a region including $\boldsymbol{x_0}$ and $\boldsymbol{x_0} + \boldsymbol{h}$, then

$$\begin{aligned}
f(\boldsymbol{x}+\boldsymbol{h}) = f(\boldsymbol{x}) & \\
&+ \left[\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right) f\right](\boldsymbol{x}) \\
&+ \frac{1}{2!}\left[\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)^2 f\right](\boldsymbol{x}) \\
&+ \ldots \\
&+ \frac{1}{k!}\left[\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)^k f\right](\boldsymbol{x}) \\
&+ \frac{1}{(k+1)!}\left[\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)^{k+1} f\right](\boldsymbol{x}+\xi\boldsymbol{h})
\end{aligned}$$