Multivariate Calculus Rough Notes

Continuity

- **A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous at $\vec{\mathbf{x}}$ iff $\forall \epsilon > 0 : \exists \delta > 0 :$ if $||\vec{\mathbf{y}} - \vec{\mathbf{x}}|| < \delta$ then $||f(\vec{\mathbf{y}}) - f(\vec{\mathbf{x}})|| < \epsilon$.** In other words, a function f is continuous at some point $\vec{\mathbf{x}}$ iff for all open balls centred around $f(\vec{\mathbf{x}})$, there is some open ball centred around $\vec{\mathbf{x}}$ such that f maps all points in the latter circle to some point in the former circle. Informally, points near $\vec{\mathbf{x}}$ map to points near $f(\vec{\mathbf{x}})$. (See Moreno de Barreda, 2023 - Lecture on Multivariate Calculus, p. 8 for intuition.)
  - This definition is generally useful for proving continuity.
- Equivalently, a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous at $\vec{\mathbf{x}}$ if for any sequence $\{\vec{\mathbf{x}}_k\}_{k=1}^{\infty}$ converging to $\vec{\mathbf{x}}$, $\{f(\vec{\mathbf{x}}_k)\}_{k=1}^{\infty}$ converges to $f(\vec{\mathbf{x}})$.
  - This definition is generally useful for proving non-continuity.

Limit Laws

- $\lim_{x\to a} x = a, \lim_{x\to a} c = c$.
- $\lim_{x\to a}[f(x) + g(x)] = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$ and likewise for subtraction, multiplication, and division.
- $\lim_{x\to a} cf(x) = c \lim_{x\to a} f(x), \lim_{x\to a}[f(x)]^n = [\lim_{x\to a} f(x)]^n$.
- Common techniques for evaluating $\lim_{x\to a} \frac{f(x)}{g(x)}$ where $g(a) = 0$ include factoring the numerator and denominator, multiplying by a conjugate, and applying L'Hopital's rule: $\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}$.

Common Results

- Consider $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^n \to \mathbb{R}^m$. If each of f and g is continuous at $\vec{\mathbf{x}}$, then each of $f + g$, $f - g$, $f \times g$ is continuous at $\vec{\mathbf{x}}$.
- Consider $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$. If $g(\vec{\mathbf{x}}) \neq 0$, and each of f and g is continuous at $\vec{\mathbf{x}}$, then f/g is continuous at $\vec{\mathbf{x}}$.
- Consider $f = (f_1, f_2, \dots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$. f is continuous at $\vec{\mathbf{x}}$ iff each of $f_1, f_2, \dots, f_m$ is continuous at $\vec{\mathbf{x}}$.
- Consider $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^k$. If f is continuous at $\vec{\mathbf{x}}$ and g is continuous at $f(\vec{\mathbf{x}})$, then $g \circ f(\vec{\mathbf{x}}) \equiv g(f(\vec{\mathbf{x}}))$ is continuous at $\vec{\mathbf{x}}$.

Weierstrass Extreme Value Theorem

- The Weierstrass extreme value theorem states that a real-valued function $f : S \to \mathbb{R}$ attains a maximum and a minimum iff f is continuous for all $\vec{\mathbf{x}} \in S$ and S is a compact (i.e. closed and bounded) set.
  - A set is closed iff any sequence of points, each in the set, converges to some point in the set. Informally, a set is closed if the boundaries of that set are contained in that set. For example, $[0, 1]$ is a closed set and $(0, 1)$ is an open set.
  - A set is bounded iff there exists some ball that contains that set. Equivalently, a set is bounded iff there exists some upper bound and some lower bound of the set in each dimension. For example, $[0, \infty)$ is not bounded.

Differentiation

- The k^{th} partial derivative of $f : \mathbb{R}^n \to \mathbb{R}$ at $\vec{\mathbf{x}} \in \mathbb{R}^n$ is defined as $\frac{\partial f}{\partial x_k}(\vec{\mathbf{x}}) = \lim_{h\to 0} \frac{f(x_1,\dots,x_k+h,\dots,x_n) - f(\vec{\mathbf{x}})}{h}$. If this limit does not exist, then the k^{th} partial derivative at $\vec{\mathbf{x}}$, $\frac{\partial f}{\partial x_k}(\vec{\mathbf{x}})$ does not exist.
- The k^{th} partial derivative of f is denoted by each of $\frac{\partial f}{\partial x_k}, f_k, f_{x_k}, \partial_k f, \partial_{x_k} f, D_k f$.
- The second-order partial derivative $\frac{\partial^2 f}{\partial x_i \partial x_j}$ of the function f is equal to $\frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}$.
- The second-order partial derivative $\frac{\partial^2 f}{\partial x_i \partial x_j}$ is also denoted by each of $\partial_i \partial_j f, f_{ij}$.
- Young's theorem states that if each of $\partial_i \partial_j f$ and $\partial_j \partial_i f$ exists and is continuous, the two are equal.

Differentiability

- The function $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\overrightarrow{\mathbf{x}}$ iff it can be approximated by a linear function around $\overrightarrow{\mathbf{x}}$. Formally, this is iff there exists a $m \times n$ matrix A such that $\lim_{||\overrightarrow{\mathbf{h}}|| \to 0} \frac{||f(\overrightarrow{\mathbf{x}}+\overrightarrow{\mathbf{h}})-f(\overrightarrow{\mathbf{x}})-A\overrightarrow{\mathbf{h}}||}{||\overrightarrow{\mathbf{h}}||} = 0$ (where $\overrightarrow{\mathbf{h}}$ is n-dimensional).
- If f is differentiable at $\overrightarrow{\mathbf{x}}$, i.e. there exists such A, then f is continuous at $\overrightarrow{\mathbf{x}}$ and the unique such A is the Jacobian matrix Df of f, i.e. $f(\overrightarrow{\mathbf{x}}+\overrightarrow{\mathbf{h}}) \approx f(\overrightarrow{\mathbf{x}}) + Df(\overrightarrow{\mathbf{x}}) \cdot \overrightarrow{\mathbf{h}} \Rightarrow f(\overrightarrow{\mathbf{x'}}) \approx f(\overrightarrow{\mathbf{x}}) + Df(\overrightarrow{\mathbf{x}}) \cdot (\overrightarrow{\mathbf{x'}} - \overrightarrow{\mathbf{x}})$ near $\overrightarrow{\mathbf{x}}$.
  - From the above, $\lim_{||\overrightarrow{\mathbf{h}}|| \to 0} \frac{||f(\overrightarrow{\mathbf{x}}+\overrightarrow{\mathbf{h}})-f(\overrightarrow{\mathbf{x}})-Df(\overrightarrow{\mathbf{x}})\overrightarrow{\mathbf{h}}||}{||\overrightarrow{\mathbf{h}}||} = 0$ is necessary for the differentiability of f.
- **A function f is differentiable at $\overrightarrow{\mathbf{x}}$ if (but not only if) (1) in a neighbourhood of $\overrightarrow{\mathbf{x}}$ all partial derivatives of f exist, and (2) all partial derivatives of f are continuous at $\overrightarrow{\mathbf{x}}$.**
  - This is a sufficient (but not necessary) condition
- **A function f is continuously differentiable (i.e. C^1) iff at all points in the domain of f, the partial derivatives of f exist, and are continuous.**
- **A function f is twice continuously differentiable (i.e. C^2) iff at all points in the domain of f, the second-order partial derivatives of f exist, and are continuous.
- $C^2 \Rightarrow C^1 \Rightarrow$ **differentiable (at all points)** $\Rightarrow$ **continuous (at all points).**
  - Then, C^1 is a sufficient condition for differentiability and continuity, and continuity is a necessary condition for differentiability and C^1.
- **Differentiable (at some point)** $\Rightarrow$ **continuous (at that point).**

Gradient Vector

- The gradient vector ∇f of the function $f : \mathbb{R}^n \to \mathbb{R}$ is well-defined iff $\forall k \in \{1, \ldots, n\} : \partial_k f$ exists. If ∇f exists, $\nabla f = \begin{pmatrix} \partial_1 f \\ \partial_2 f \\ \vdots \\ \partial_n f \end{pmatrix}$.
- $\nabla f(\overrightarrow{\mathbf{x}})$ is a vector that points in the direction from $\overrightarrow{\mathbf{x}}$ in which the rate of change of f is greatest. The gradient vector at some point is perpendicular, at that point, to the level curve that intersects that point. (See Moreno de Barreda, 2023 - Lecture on Multivariate Calculus, p. 17 for illustration.)

Directional Derivative

- The directional derivative of a function $f : \mathbb{R}^n \to \mathbb{R}$ at $\overrightarrow{\mathbf{x}}$ in direction $\overrightarrow{\mathbf{v}}$ is defined as $Df(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{v}}) = \lim_{t \to 0} \frac{f(\overrightarrow{\mathbf{x}}+t\overrightarrow{\mathbf{v}})-f(\overrightarrow{\mathbf{x}})}{t}$.
  - For a C^1 function, this is equal to $\nabla f(\overrightarrow{\mathbf{x}}) \cdot \overrightarrow{\mathbf{v}}$.

Jacobian Matrix

The Jacobian matrix of the function $f : \mathbb{R}^n \to \mathbb{R}^m$ is the $m \times n$ matrix $Df(\overrightarrow{\mathbf{x}}) = \begin{pmatrix} \partial_1 f_1 & \partial_2 f_1 & \ldots & \partial_n f_1 \\ \partial_1 f_2 & \partial_2 f_2 & \ldots & \partial_n f_2 \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 f_m & \partial_2 f_m & \ldots & \partial_n f_m \end{pmatrix}$ (if each partial derivative exists). Each row of the Jacobian matrix consists of the partial derivatives of some f_i and each column of the Jacobian matrix consists of the partial derivatives of f with respect to some x_j.

Hessian Matrix

- The Hessian matrix of the function $f : \mathbb{R}^n \to \mathbb{R}$ is $D^2 f(\overrightarrow{\mathbf{x}}) = \begin{pmatrix} \partial_1\partial_1 f & \partial_1\partial_2 f & \ldots & \partial_1\partial_n f \\ \partial_2\partial_1 f & \partial_2\partial_2 f & \ldots & \partial_2\partial_n f \\ \vdots & \vdots & \ddots & \vdots \\ \partial_n\partial_1 f & \partial_n\partial_2 f & \ldots & \partial_n\partial_n f \end{pmatrix}$.
- If f is C^2, then $D^2 f(\overrightarrow{\mathbf{x}})$ is a symmetric matrix.

Chain Rule

- If each of $f : \mathbb{R}^n \to \mathbb{R}$ and $\overrightarrow{\mathbf{x}} : \mathbb{R}^m \to \mathbb{R}^n$ is a C^1 function, then $Z : \mathbb{R}^m \to \mathbb{R}$ such that $Z(\overrightarrow{\mathbf{t}}) = f(\overrightarrow{\mathbf{x}}(\overrightarrow{\mathbf{t}}))$ is C^1 and $\frac{\partial z}{\partial t_k} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_k} + \ldots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_k}$ for $1 \le k \le m$.

Taylor Approximation

- Taylor's (first order) theorem states that if function $f : U \to \mathbb{R}$ is C^1 and U is an open subset of $\mathbb{R}^n$, then $\forall \overrightarrow{\mathbf{a}}, \overrightarrow{\mathbf{x}} \in U : f(\overrightarrow{\mathbf{x}}) = f(\overrightarrow{\mathbf{a}}) + Df(\overrightarrow{\mathbf{a}})(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}}) + R_1(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{a}})$, where $\lim_{\overrightarrow{\mathbf{x}} \to \overrightarrow{\mathbf{a}}} \frac{R_1(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{a}})}{||(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}})||} \to 0$.
  - Informally, any $f(\overrightarrow{\mathbf{x}})$ is well approximated by $f(\overrightarrow{\mathbf{a}}) + Df(\overrightarrow{\mathbf{a}})(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}})$. **Intuitively, the first-order Taylor approximation $f(\overrightarrow{\mathbf{x}}) \approx f(\overrightarrow{\mathbf{a}}) + Df(\overrightarrow{\mathbf{a}})(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}})$ gives the equation of the plane that is tangent to the surface $z = f(\overrightarrow{\mathbf{x}})$ at $\overrightarrow{\mathbf{a}}$.**
- Taylor's (second order) theorem states that if function $f : U \to \mathbb{R}$ is C^2 and U is an open subset of $\mathbb{R}^n$, then $\forall \overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{a}} \in U : f(\overrightarrow{\mathbf{x}}) = f(\overrightarrow{\mathbf{a}}) + Df(\overrightarrow{\mathbf{a}})(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}}) + \frac{1}{2}(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}})^T D^2 f(\overrightarrow{\mathbf{a}})(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}}) + R_2(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{a}})$ where $\lim_{\overrightarrow{\mathbf{x}} \to \overrightarrow{\mathbf{a}}} \frac{R_2(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{a}})}{||\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}}||^2} \to 0$.
  - Informally, any $f(\overrightarrow{\mathbf{x}})$ is well approximated by $f(\overrightarrow{\mathbf{a}}) + Df(\overrightarrow{\mathbf{a}})(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}}) + \frac{1}{2}(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}})^T D^2 f(\overrightarrow{\mathbf{a}})(\overrightarrow{\mathbf{x}} - \overrightarrow{\mathbf{a}})$.

Implicit Differentiation

- **The implicit function theorem states that if (1) $(\overrightarrow{\mathbf{x}}^*, y^*)$ solves $f(\overrightarrow{\mathbf{x}}, y) = 0$, (2) f is C^1 in an open ball around $(\overrightarrow{\mathbf{x}}^*, y^*)$, and (3) $f_y(\overrightarrow{\mathbf{x}}^*, y^*) \neq 0$, then there is a C^1 function $g(\overrightarrow{\mathbf{x}})$ defined on an open ball around $\overrightarrow{\mathbf{x}}^*$ such that $y^* = g(\overrightarrow{\mathbf{x}}^*)$, $f(\overrightarrow{\mathbf{x}}, g(\overrightarrow{\mathbf{x}})) = 0$, and $g_i(\overrightarrow{\mathbf{x}}) = -\frac{f_i(\overrightarrow{\mathbf{x}}, y)}{f_y(\overrightarrow{\mathbf{x}}, y)}$. In other words, there is some function $g(\overrightarrow{\mathbf{x}})$ such that each $(\overrightarrow{\mathbf{x}}, g(\overrightarrow{\mathbf{x}}))$ (in an open ball around $\overrightarrow{\mathbf{x}}^*$) solves $f(\overrightarrow{\mathbf{x}}, y) = 0$, and the change in y from y^* corresponding to a change in $\overrightarrow{\mathbf{x}}$ from $\overrightarrow{\mathbf{x}}^*$ can be found by implicit differentiation.**
  - **Note the negative sign on the left hand side of the equation for g_i.**
- The implicit function theorem generalises as follows. If (1) $(\overrightarrow{\mathbf{x}^*}, \overrightarrow{\mathbf{y}^*})$ solves $f(\overrightarrow{\mathbf{x}}, \overrightarrow{\mathbf{y}}) = \overrightarrow{\mathbf{0}}$, (2) f is C^1 in an open ball around $(\overrightarrow{\mathbf{x}^*}, \overrightarrow{\mathbf{y}^*})$, and (3) $D_{\overrightarrow{\mathbf{y}}} f$ is invertible at $(\overrightarrow{\mathbf{x}^*}, \overrightarrow{\mathbf{y}^*})$, then there is a C^1 function $g(\overrightarrow{\mathbf{x}})$ defined on an open ball around $\overrightarrow{\mathbf{x}^*}$ such that $\overrightarrow{\mathbf{y}^*} = g(\overrightarrow{\mathbf{x}^*})$, $f(\overrightarrow{\mathbf{x}}, g(\overrightarrow{\mathbf{x}})) = 0$, and $D_{\overrightarrow{\mathbf{x}}} g = -[D_{\overrightarrow{\mathbf{y}}} f(\overrightarrow{\mathbf{x}}, g(\overrightarrow{\mathbf{x}}))]^{-1} D_{\overrightarrow{\mathbf{x}}} f(\overrightarrow{\mathbf{x}}, g(\overrightarrow{\mathbf{x}}))$, i.e. $D_{\overrightarrow{\mathbf{x}}} g$ solves $AX = B$ where A is the Jacobian (in $\overrightarrow{\mathbf{y}}$) of f and B is the Jacobian (in $\overrightarrow{\mathbf{x}}$) of f, each evaluated at $(\overrightarrow{\mathbf{x}}, g(\overrightarrow{\mathbf{x}}))$.
  - **Note the negative sign on the left hand side of the equation for $D_{\overrightarrow{\mathbf{x}}} g$.**