Chapter 9

Functions of Several Variables

9.1 Notes

2/14:

- Plan:
  1. Warm-up with matrices.
  2. The total derivatives of $f : \mathbb{R}^n \to \mathbb{R}^m$ ($n = m = 2$, i.e., $f : \mathbb{C} \to \mathbb{C}$).
  3. Basic properties: Chain rule, relation with partial derivatives, implicit function theorem.
- Let V, W be finite-dimensional vector spaces over $\mathbb{R}$. We let $L(V, W)$ be the vector space of all linear transformations $\phi : V \to W$.
- If we pick bases $\mathbf{v}_1, \dots, \mathbf{v}_n$ of V and $\mathbf{w}_1, \dots, \mathbf{w}_m$ of W, then $V \cong \mathbb{R}^n$ and $W \cong \mathbb{R}^m$. It follows that $L(V, W) \cong \mathbb{R}^{mn}$.
- $L(V, W) \times L(W, U) \xrightarrow{\text{compose}} L(V, U)$, i.e., $\mathbb{R}^{mn} \times \mathbb{R}^{nl} \xrightarrow[\text{mult.}]{\text{matrix}} \mathbb{R}^{ml}$.
- Sup norm: If A is an $m \times n$ real matrix, then $\|A\| = \sup_{\substack{\mathbf{x} \in \mathbb{R}^n \\ |\mathbf{x}| = 1}} |A\mathbf{x}|$.
  - Basic properties:
    1. $|A\mathbf{x}| \leq \|A\||x|$.
    2. $\|A\| < \infty$ and all $A : \mathbb{R}^n \to \mathbb{R}^m$ are uniformly continuous.
    3. $\|A\| = 0 \iff A = 0$.
    4. $\|cA\| = |c|\|A\|$.
    5. $\|A + B\| \leq \|A\| + \|B\|$.
    6. $\|AB\| \leq \|A\|\|B\|$.
  - Note that we get a metric space structure on $L(V, W)$ by defining $d(A, B) = \|A - B\|$.
- Proves that 1 and 2 imply the uniform continuity of all A (via Lipschitz continuity).
- **Differentiable** (function $\mathbf{f}$ at $\mathbf{x}_0$): A function $\mathbf{f} : U \to \mathbb{R}^m$ ($U \subset \mathbb{R}^n$) such that to $\mathbf{x}_0 \in U$ there corresponds some linear transformation $A : \mathbb{R}^n \to \mathbb{R}^m$ such that

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

- **Total derivative** (of $\mathbf{f}$ at $\mathbf{x}_0$): The linear transformation A in the above definition. *Denoted by* $\mathbf{f'(x_0)}$, $\boldsymbol{D}\mathbf{f(x_0)}$, $\boldsymbol{d}\mathbf{f(x_0)}$.
- "An proof and progress in mathematics" - Thurston.

  - Relating to the old one dimensional derivative.
  - A paper we'd find rather impressionistic right now.
- Propositions ahead of us.
  - Proposition: Suppose that $\mathbf{f}$ is differentiable at $\mathbf{x}_0 \in U$ and A, B are both derivatives of $\mathbf{f}$ at $\mathbf{x}_0$. Then $A = B$.
  - Proposition: Differentiable implies continuous.
  - Proposition: Sum rule, product rule, quotient rule.

2/16:

- Plan: Derivatives of functions $\mathbf{f} : U \to \mathbb{R}^m$ where $U \subset \mathbb{R}^n$.
  - Basic properties: Differentiability implies continuity, $(\mathbf{f}+\mathbf{g})' = \mathbf{f}' + \mathbf{g}'$, $(c\mathbf{f})' = c\mathbf{f}'$, chain rule, $\mathbf{f}' = 0$ iff $\mathbf{f}$ is constant.
  - Relationship with partial derivatives (how we compute everything and anything).
  - When is $\mathbf{f}$ differentiable?
  - Inverse function theorem.
  - Implicit function theorem.
- **Continuously differentiable** (function $\mathbf{f}$): A function $\mathbf{f} : U \to \mathbb{R}^m$ that is differentiable for all $\mathbf{x}_0 \in U$ and such that $\mathbf{f}' : U \to L(\mathbb{R}^n, \mathbb{R}^m)$ is continuous. *Also known as* $\boldsymbol{\mathscr{C}^1}$.
- Proposition: Let $\mathbf{f} : U \to \mathbb{R}^m$ be differentiable at $\mathbf{x}_0 \in U$. Then $\mathbf{f}$ is continuous at $\mathbf{x}_0$.
  - The proof makes use of the fact that $\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0) = \mathbf{f}'(\mathbf{x}_0)\mathbf{h} + \mathbf{r}(\mathbf{h})$.
- Proposition: Given $\mathbf{f}, \mathbf{g} : U \to \mathbb{R}^m$ both differentiable at $\mathbf{x}_0 \in U$, then $\mathbf{f}+\mathbf{g}$ is also differentiable at $\mathbf{x}_0$ with
$$(\mathbf{f}+\mathbf{g})'(\mathbf{x}_0) = \mathbf{f}'(\mathbf{x}_0) + \mathbf{g}'(\mathbf{x}_0)$$
  - The proof is immediate via the triangle inequality.
- Theorem (Chain Rule): Given $\mathbf{f} : U \to \mathbb{R}^m$ and $\mathbf{g} : V \to \mathbb{R}^k$, where $U \subset \mathbb{R}^n$ and $\mathbf{f}(U) \subset V \subset \mathbb{R}^m$, with $\mathbf{f}$ differentiable at $\mathbf{x}_0 \in U$ and $\mathbf{g}$ differentiable at $\mathbf{f}(\mathbf{x}_0)$, the composition $\mathbf{g} \circ \mathbf{f}$ is differentiable at $\mathbf{x}_0$ with
$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}_0) = \mathbf{g}'(\mathbf{f}(\mathbf{x}_0)) \cdot \mathbf{f}'(\mathbf{x}_0)$$
  - The proof is rather subtle.
- **Partial derivative** (of f_i wrt. x_j at $\mathbf{x}_0$): The following limit, if it exists, where $f_i : \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq m$, and $1 \leq j \leq n$. *Denoted by* $\boldsymbol{(\partial f_i / \partial x_j)(\mathbf{x}_0)}$, $\boldsymbol{(D_j f_i)(\mathbf{x}_0)}$. *Given by*
$$\frac{\partial f_i}{\partial x_j}(\mathbf{x_0}) = \lim_{t \to 0} \frac{f_i(\mathbf{x}_0 + t\mathbf{e}_j) - f_i(\mathbf{x}_0)}{t}$$
- **Directional derivative** (of f_i toward $\mathbf{u} \in \mathbb{R}^n$): The following limit, if it exists, where $f_i : \mathbb{R}^n \to \mathbb{R}$ and $1 \leq i \leq m$. *Denoted by* $\boldsymbol{D_{\mathbf{u}} f_i}$. *Given by*
$$D_{\mathbf{u}} f_i = \lim_{t \to 0} \frac{f_i(\mathbf{x}_0 + t\mathbf{u}) - f_i(\mathbf{x}_0)}{t}$$
- **Jacobian**: The following matrix. *Given by*
$$\left[\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) \right]$$

- Theorem: Let $\mathbf{f} = (f_1, \dots, f_m) : U \to \mathbb{R}^m$, where $U \subset \mathbb{R}^n$, be differentiable at some $\mathbf{x}_0 \in U$. Then the partial derivatives $\partial f_i / \partial x_j$ $(1 \leq i \leq m;\ 1 \leq j \leq n)$ exist at $\mathbf{x}_0$ and, with respect to the usual choice of bases,
$$\mathbf{f}'(\mathbf{x}_0) = \left[\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)\right]$$

2/18: – We have that
$$\mathbf{f}(\mathbf{x}_0 + t\mathbf{e}_j) - \mathbf{f}(\mathbf{x}_0) = \mathbf{f}'(\mathbf{x}_0)(t\mathbf{e}_j) + \mathbf{r}(t\mathbf{e}_j)$$
  - Since $\mathbf{f}$ is differentiable at $\mathbf{x}_0$, $\mathbf{f}(t\mathbf{e}_j)/t \to 0$ as $t \to 0$.
  - Additionally, $\mathbf{f}'(\mathbf{x}_0)(t\mathbf{e}_j)/t = \mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j)$.
  - Therefore,
$$\lim_{t \to 0} \frac{\mathbf{f}(\mathbf{x}_0 + t\mathbf{e}_j) - \mathbf{f}(\mathbf{x}_0)}{t} = \lim_{t \to 0} \frac{\mathbf{f}'(\mathbf{x}_0)(t\mathbf{e}_j) - \mathbf{r}(t\mathbf{e}_j)}{t} = \mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j) - \lim_{t \to 0} \frac{\mathbf{r}(t\mathbf{e}_j)}{t} = \mathbf{f}'(\mathbf{x}_0)(\mathbf{e}_j)$$
    as desired.
  - Unpacking the definition of the linear transformation as a matrix gives the rest of the proof.

- Today:
  - More on differentiation (recall the Jacobian).
  - Sufficient condition for differentiability.
  - $\mathbf{f}' = 0$ iff $\mathbf{f}$ is constant.
  - State the inverse function theorem.

- It is not true that having all partials exist implies that $\mathbf{f}$ is differentiable at $\mathbf{x}_0$.

- Theorem: $\mathbf{f}$ continuously differentiable at $\mathbf{x}_0$ iff all partials exist and are continuous at $\mathbf{x}_0$.

- Theorem (Inverse function theorem): If $E \subset \mathbb{R}^n$ open, $\mathbf{f} : E \to \mathbb{R}^n$ is differentiable at $\mathbf{x}_0 \in E$, and $\mathbf{f}'(\mathbf{x}_0)$ is invertible, then there exist $U \subset E$ open with $\mathbf{x}_0 \in U$ and $V \subset \mathbb{R}^n$ open with $\mathbf{f}(\mathbf{x}_0) \in V$ such that $\mathbf{f}|_U : U \to V$ is a bijection and $(\mathbf{f}|_U)^{-1}$ is continuously differentiable.

9.2 Chapter 9: Functions of Several Variables

From Rudin (1976).

2/15: - Defines a vector space by the closure of its elements under addition and scalar multiplication.

- Defines a linear combination, span, independence and dependence, dimension, basis, coordinates, and the standard basis.

- Theorem 9.2: If X is spanned by r vectors, $\dim X \leq r$.

- Corollary: $\dim \mathbb{R}^n = n$.

- Theorem 9.3: Let X a vector space with $\dim X = n$.
  (a) $E \subset X$ containing n vectors spans X iff E is independent.
  (b) X has a basis, and every basis contains n vectors.
  (c) If $1 \leq r \leq n$ and $\{\mathbf{y}_1, \dots, \mathbf{y}_r\}$ is independent in X, then X has a basis containing $\{\mathbf{y}_1, \dots, \mathbf{y}_r\}$.

- Defines linear transformation, linear operator.

- Notes that $A\mathbf{0} = \mathbf{0}$ if A is a linear transformation, and that A is completely determined by its action on any basis.

- **Invertible** (linear operator): A linear operator A that is one-to-one and onto.
- Theorem 9.5: A a linear operator on X finite-dimensional is one-to-one iff it is onto.
- Defines $L(X,Y)$, $L(X)$, the product BA of two linear transformations, and the supremum norm of a linear transformation.
- Theorem 9.7:
  (a) $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ implies $\|A\| < \infty$ and $A : \mathbb{R}^n \to \mathbb{R}^m$ uniformly continuous.
  (b) $A, B \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $c \in \mathbb{C}$ implies
  $$\|A + B\| \leq \|A\| + \|B\| \qquad\qquad \|cA\| = |c|\|A\|$$
  Defining $d(A,B) = \|A - B\|$ makes $L(\mathbb{R}^n, \mathbb{R}^m)$ a metric space.
  (c) $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $B \in L(\mathbb{R}^m, \mathbb{R}^k)$ implies
  $$\|BA\| \leq \|B\|\|A\|$$
- Theorem 9.8: Let Ω be the set of all invertible linear operators on $\mathbb{R}^n$.
  (a) $A \in \Omega$, $B \in L(\mathbb{R}^n)$, and $\|B - A\| \cdot \left\|A^{-1}\right\| < 1$ implies $B \in \Omega$.

  Proof. Let $\left\|A^{-1}\right\| = 1/\alpha$, and let $\|B - A\| = \beta$. Then
  $$\begin{aligned} \|B - A\| \cdot \left\|A^{-1}\right\| &< 1 \\ \beta \cdot \frac{1}{\alpha} &< 1 \\ \beta &< \alpha \end{aligned}$$
  To prove that $B \in \Omega$, the definition of invertibility and Theorem 9.5 tell us that it will suffice to show that B is 1-1. To do so, it will suffice to show that $B\mathbf{x} = \mathbf{0}$ iff $\mathbf{x} = \mathbf{0}$. Let's begin. Let $\mathbf{x} \in \mathbb{R}^n$ be arbitrary. Then
  $$\begin{aligned} \alpha|\mathbf{x}| = \alpha|A^{-1}A\mathbf{x}| \leq \alpha\left\|A^{-1}\right\| \cdot |Ax| = |A\mathbf{x}| \leq |(A - B)\mathbf{x}| + |B\mathbf{x}| &\leq \beta|\mathbf{x}| + |B\mathbf{x}| \\ (\alpha - \beta)|\mathbf{x}| &\leq |B\mathbf{x}| \end{aligned}$$
  It follows that if $\mathbf{x} \neq \mathbf{0}$, then $|B\mathbf{x}| > 0$. This combined with the fact that $B\mathbf{0} = \mathbf{0}$ implies the desired result. ☐

  (b) Ω is open in $L(\mathbb{R}^n)$ and $A \mapsto A^{-1}$ is continuous on Ω.

  Proof. To prove that Ω is open in $L(\mathbb{R}^n)$, it will suffice to show that for all $A \in \Omega$, there exists $N_r(A)$ such that if $\|B - A\| < r$, then $B \in \Omega$. Let's begin. Let $A \in \Omega$ be arbitrary. Choose $N_\alpha(A)$ to be our neighborhood, where α is defined as in part (a). Let $B \in L(\mathbb{R}^n)$ satisfy $\|B - A\| < \alpha$. Then $\|B - A\| \cdot \left\|A^{-1}\right\| < 1$, so $B \in \Omega$ by part (a), as desired.
  To prove that $A \mapsto A^{-1}$ is continuous, it will suffice to show that $\left\|B^{-1} - A^{-1}\right\| \to 0$ as $B \to A$. First off, we have by part (a) and the substitution $\mathbf{x} = B^{-1}\mathbf{y}$ ($\mathbf{y} \in \mathbb{R}^n$) that
  $$\begin{aligned} (\alpha - \beta)|B^{-1}\mathbf{y}| &\leq |BB^{-1}\mathbf{y}| = |\mathbf{y}| \\ \left|B^{-1}\left(\frac{\mathbf{y}}{|\mathbf{y}|}\right)\right| &\leq (\alpha - \beta)^{-1} \end{aligned}$$
  Thus, since $|B^{-1}\mathbf{u}|$ is bounded by $(\alpha - \beta)^{-1}$ for every unit vector $\mathbf{u} \in \mathbb{R}^n$, $\left\|B^{-1}\right\|$ is bounded by $(\alpha - \beta)^{-1}$. This combined with the fact that
  $$\begin{aligned} B^{-1} - A^{-1} &= B^{-1}I - IA^{-1} \\ &= B^{-1}AA^{-1} - B^{-1}BA^{-1} \\ &= B^{-1}(A - B)A^{-1} \end{aligned}$$

implies by Theorem 9.7c that

$$\left\|B^{-1} - A^{-1}\right\| \leq \left\|B^{-1}\right\|\|A - B\|\left\|A^{-1}\right\| \leq (\alpha - \beta)^{-1} \cdot \beta \cdot \frac{1}{\alpha} = \frac{\beta}{\alpha(\alpha - \beta)}$$

Therefore, since $\beta \to 0$ as $B \to A$, the above inequality establishes the desired result. □

- Note that the mapping $A \mapsto A^{-1}$ defined in Theorem 9.8b is a 1-1 mapping of Ω onto Ω and its own inverse.
- Defines matrices, column vectors, and matrix multiplication.
- From the Schwarz inequality, we can show that

$$\|A\| \leq \left(\sum_{i,j} a_{i,j}^2\right)^{1/2}$$

- "If S is a metric space, if $a_{11}, \dots, a_{mn}$ are real continuous functions on S, and if for each $p \in S$, A_p is the linear transformation of $\mathbb{R}^n$ into $\mathbb{R}^m$ whose matrix has entries $a_{ij}(p)$, then the mapping $p \to A_p$ is a continuous mapping of S into $L(\mathbb{R}^n, \mathbb{R}^m)$" (Rudin, 1976, p. 211).
- Rudin (1976) spends some time motivating the definition of the total derivative. He also discusses the natural 1-1 correspondence between $\mathbb{R}^1$ and $L(\mathbb{R}^1)$.
- Defines differentiability in $\mathbb{R}^n$.
- Theorem 9.12: A_1, A_2 the derivative of $\mathbf{f}$ at $\mathbf{x}$ implies $A_1 = A_2$.
- If $\mathbf{f} : E \to \mathbb{R}^m$ where $E \subset \mathbb{R}^n$, then $\mathbf{f}' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$.
- $\mathbf{f}$ differentiable implies $\mathbf{f}$ continuous.
- Example ($\mathbf{f}$ is linear):
  - If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, then $A'(\mathbf{x}) = A$ for all $\mathbf{x} \in \mathbb{R}^n$. Note that this means that $A' : \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^m)$, as expected.
- Theorem 9.15 (Chain Rule): E open in $\mathbb{R}^n$, $\mathbf{f} : E \to \mathbb{R}^m$ differentiable at $\mathbf{x}_0 \in E$, $I \supset \mathbf{f}(E)$ open in $\mathbb{R}^m$, and $\mathbf{g} : I \to \mathbb{R}^k$ differentiable at $\mathbf{f}(\mathbf{x}_0)$ implies $\mathbf{F} : E \to \mathbb{R}^k$ defined by

$$\mathbf{F}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$$

is differentiable at $\mathbf{x}_0$ with

$$\mathbf{F}'(\mathbf{x}_0) = \mathbf{g}'(\mathbf{f}(\mathbf{x}_0))\mathbf{f}'(\mathbf{x}_0)^{[1]}$$

Proof. Largely symmetric to that of the one-dimensional chain rule in Chapter 5. □

- **Components** (of $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$): The real functions $f_1, \dots, f_m$ defined by

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{m} f_i(\mathbf{x})\mathbf{u}_i$$

for all $\mathbf{x} \in E$ or, equivalently, by $f_i(\mathbf{x}) = f(\mathbf{x}) \cdot \mathbf{u}_i$ $(1 \leq i \leq m)$, where $\mathbf{u}_1, \dots, \mathbf{u}_m$ is the standard basis of $\mathbb{R}^m$.

[1] Note that the right-hand side of this equation contains the product of two linear transformations.

- Defines partial derivatives.
- Theorem 9.17: $E \subset \mathbb{R}^n$ open and $\mathbf{f} : E \to \mathbb{R}^m$ differentiable at $\mathbf{x} \in E$ imply the partial derivatives $(D_j f_i)(\mathbf{x})$ exist and
$$\mathbf{f}'(\mathbf{x})\mathbf{e}_j = \sum_{i=1}^{m} (D_j f_i)(\mathbf{x})\mathbf{u}_i$$
for $1 \leq j \leq n$.
- It follows that
$$[\mathbf{f}'(\mathbf{x})] = \begin{bmatrix} (D_1 f_1)(\mathbf{x}) & \cdots & (D_n f_1)(\mathbf{x}) \\ \vdots & & \vdots \\ (D_1 f_m)(\mathbf{x}) & \cdots & (D_n f_m)(\mathbf{x}) \end{bmatrix}$$
- Discusses the gradient and the directional derivative.
- Theorem 9.19: $E \subset \mathbb{R}^n$ convex and open, $\mathbf{f} : E \to \mathbb{R}^m$ differentiable in E, and there exists M such that
$$\|\mathbf{f}'(\mathbf{x})\| \leq M$$
for all $\mathbf{x} \in E$ implies
$$|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})| \leq M|\mathbf{b} - \mathbf{a}|$$
for all $\mathbf{a}, \mathbf{b} \in E$.
- Corollary: If, in addition, $\mathbf{f}'(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in E$, then $\mathbf{f}$ is constant.
- **Continuously differentiable** (mapping $\mathbf{f} : E \to \mathbb{R}^m$): A function $\mathbf{f} : E \to \mathbb{R}^m$ such that $\mathbf{f}' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$ is continuous. *Also known as* $\mathscr{C}^1$**-mapping**. *Denoted by* $\mathbf{f} \in \mathscr{C}^1(E)$.
- Theorem 9.21: Let $E \subset \mathbb{R}^n$ open and $\mathbf{f} : E \to \mathbb{R}^m$. Then $\mathbf{f} \in \mathscr{C}^1(E)$ iff the partial derivatives $D_j f_i$ $(1 \leq i \leq m; 1 \leq j \leq n)$ exist and are continuous on E.