1. Local consequences of differentiation

(1) Optimization

(a) Let U be a subset of $\mathbb{R}^n$.

(b) We say that $f : U \to \mathbb{R}$ *attains a maximum* at $p \in U$ iff

$$f(p) = \sup\{f(x) \mid x \in U\}.$$

(c) We say that $f : U \to \mathbb{R}$ *attains a minimum* at $p \in U$ iff

$$f(p) = \inf\{f(x) \mid x \in U\}.$$

(d) Critical points

(i) Let U be an open subset of $\mathbb{R}^n$.

(ii) p is a *critical point of* f if and only if $df_p(v) = 0$ for all $v \in \mathbb{R}^n$.

(iii) *Nota Bene:* Since df_p is linear, need only check that $df_p(e_i) = 0$ where $\{e_i\}$ is a basis. In other words, only need to check that Jacobian is the 0 matrix.

(iv) *Theorem:* Let $f : U \to \mathbb{R}$ be differentiable. If f attains a maximum or minimum at p, then p is a critical point of f.

(A) Consider $g_v(t) = g(p + tv)$, $g_v : U' \subset \mathbb{R} \to \mathbb{R}$.

(B) By assumption g attains maximum at 0.

(C) Thus if $t > 0$, then

$$\frac{g_v(t) - g_v(0)}{t - 0} \leq 0$$

and hence $g_v'(0) \leq 0$,

(D) and if $t < 0$, then

$$\frac{g_v(t) - g_v(0)}{t - 0} \geq 0$$

and hence $g_v'(0) \geq 0$.

(E) Thus $df_p(v) = g'(0) = 0$.

(e) *Problem:* How do we determine whether a critical point is a maximum, a minumum, or neither?

(i) In one dimension: second derivative test or first derivative test.

(ii) In higher dimensions: use second derivative test in every direction.

(f) Second derivatives and Hessian

(i) Fix a direction $v = \sum_i a_i \cdot e_i$

(ii) We have

$$\begin{aligned} g_v'(t) &= df_{p+tv}(v) \\ &= \sum_i a_i df_{p+tv}(e_i) \\ &= \sum_i a_i \frac{\partial f}{\partial x_i}(p+tv) \end{aligned}$$

(iii) and then

$$\begin{aligned} g_v''(t) &= d\left(\sum_i a_i \frac{\partial f}{\partial x_i}\right)_{p+tv}(v) \\ &= \sum_i a_i \cdot d\left(\left(\frac{\partial f}{\partial x_i}\right)_{p+tv}(v)\right) \\ &= \sum_i a_i \sum_j a_j \cdot d\left(\frac{\partial f}{\partial x_i}\right)(e_j) \\ &= \sum_i a_i \sum_j a_j \cdot \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(p+tv) \\ &= \sum_{i,j} a_i a_j \frac{\partial^2 f}{\partial x_j \partial x_i}(p+tv) \end{aligned}$$

(iv) The Hessian of f at p is the $n \times n$ matrix of second partials:

$$\mathrm{Hess}(f)_p = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(p)\right].$$

(A) From above computation we have

$$g_v''(0) = v^T \cdot \mathrm{Hess}(f)_p \cdot v.$$

(B) If the second derivative of f is continuous at p (i.e. second partials continuous) then the Hessian is a symmetric matrix.

(C) Thus, it has n real eigenvalues: $\lambda_1 \leq \ldots \leq \lambda_n$. (Spectral Thm)

(D) By usual proof of spectral theorem λ_1 (resp. λ_n) is the smallest (largest) value of $v \mapsto v^T \cdot \mathrm{Hess}(f)_p \cdot v$ on unit sphere.

(E) Among unit vectors v, the function $v \mapsto g_v''(0)$ is maximized at the eigenvector(s) associated to λ_n.

(F) Among unit vectors v, the function $v \mapsto g_v''(0)$ is minimized at the eigenvector(s) associated to λ_1.

(g) *Second derivative test*: Suppose f has second partials defined and continuous on U and $df_p \equiv 0$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of the Hessian of f at p.

(i) If $\forall i$, we have $\lambda_i > 0$, then f attains a minimum at p.

(ii) If $\forall i$, we have $\lambda_i < 0$, then f attains a maximum at p.

(iii) If $\lambda_i < 0 < \lambda_j$ for some i and j, then p is neither a minimum or maximum.

(iv) If $\lambda_i = 0$ for some i then must find additional informaton e.g. third derivatives.

(h) *Exercise:* Does $f(x, y) = x^2 + 4xy + 3y^2$ attain a (local) maximum or minimum at $(x, y) = (0, 0)$.

(i) Functions without maxima or minima

(i) Let $\lambda_\pm(p)$ denote the least/greatest eigenvalue of the Hessian of f at p.

(ii) *Nota Bene:* If for each $p \in U$, we have $\lambda_-(p) < 0 < \lambda_+(p)$, then f attains neither a maximum or a minimum in U.

(iii) A C^2 function $f : U \to \mathbb{R}$ is called *harmonic* if and only if for each $p \in U$ the sum of the eigenvalues of the Hessian of f at p equals zero.

(iv) If f harmonic, then $\lambda_-(p) \le 0 \le \lambda_+(p)$.

(v) *Maximum Principle:* Harmonic functions do not attain maxima or minima. (Does not follow immediately from above.)

(2) Local inversion

(a) We will say that $f : U \to \mathbb{R}^n$ is C^1 *locally invertible* at $p \in U$ iff

(i) $\exists$ open nbhd V of $f(p)$,

(ii) $\exists\ g : V \to U$ such that $\forall y \in V$

$$f(g(y)) = y,$$

and

(iii) g has a continuous first derivative with $dg_{f(p)} = df_p^{-1}$.

(b) Suppose that the derivative of $f : U \to \mathbb{R}^n$ exists and is continuous on $U \subset \mathbb{R}^n$. If df_p is invertible, then f is locally invertible at p.

(i) Proof: Given y, want to show that we have a unique solution x to $f(x) = y$. Then set $g(y)$ equal to this solution.

(ii) Solving $f(x) = y$ is equivalent to finding a fixed point of the mapping F defined by

$$F(x) = x + y - f(x).$$

(iii) Want to show F is contraction mapping.

(iv) Reduction to origin: By precomposing with $x \mapsto x+p$ and postcomposing with $x \mapsto df_p^{-1}(f(x) - f(p))$ can assume $p = f(p) = 0$ and df_p is the identity linear transformation. (Alternatively, could use $F(x) = x + y - df_p^{-1} \circ f(x)$).

(v) We have

$$|F(a) - F(b)| \leq M \cdot |a - b|$$

where M is $\sup\{||dF_x||\}$ over x in the line segment joining a and b.

(vi) We have $dF_x = I - df_x$, thus, since $x \mapsto df_x$ is continuous with $df_0 = I$, there exists $r > 0$ so that $|x| \leq r \Rightarrow ||dF_x|| \leq \frac{1}{2}$.

(vii) Thus,

$$|F(a) - F(b)| \leq \frac{1}{2} \cdot |a - b|$$

and F is a contraction mapping.

(viii) Exercise: Show that F maps $\{x \mid |x| \leq r\}$ into $\{x \mid |x| \leq r\}$.

(ix) Exercise: Show that g has a continuous first derivative near $f(p) = 0$.

(3) Local section

(a) Implicit functions

(i) Suppose $m \leq n$ and $U \subset \mathbb{R}^n$.

(ii) Given $f : U \to \mathbb{R}^m$, the q-level set of f is the set

$$f^{-1}(\{q\}) = \{x \in \mathbb{R}^n \mid f(x) = q\}.$$

(iii) Wish to realize $f^{-1}(\{q\})$ as the graph of a function g. A suitable function g is called an 'implicit function' because it is only implicitly defined.

(iv) *Definition:* Let $p \in f^{-1}(\{q\})$ and let V be an open set in $\mathbb{R}^{n-m}$. We say that $g : V \to \mathbb{R}^n$ is a C^1 *implicit function for f near p* iff

(A) $p \in g(V)$,

(B) g is injective,

(C) g has a continuous derivative, and

(D) $f(g(z), z) = q$ for all $z \in V$.

(v) Example:

(A) $f(x_1, x_2) = x_1^2 + x_2^2$.

(B) Let $q = 1$. The point $(1, 0)$ belongs to the 1-level set of f.

(C) Let $V = (-1,1)$. Then $g : V \to \mathbb{R}^2$ defined by

$$g(y) = (\sqrt{1+y^2}, y)$$

is an implicit function for f.

(b) Constructing implicit function using inverse function theorem:

(i) $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}$: Each vector $x \in \mathbb{R}^n$ can be written as (x_1, x_2) with $x_1 \in \mathbb{R}^n$ and $x_2 \in \mathbb{R}^{n-m}$.

(ii) Define $F : \mathbb{R}^n \to \mathbb{R}^n$ by

$$F\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ x_2 \end{pmatrix}.$$

(iii) If dF_p is invertible, then F has a local inverse G:

$$F \circ G(y) = y.$$

(iv) Write

$$G\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} g_1\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ g_2\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \end{pmatrix}.$$

(v) Then $F \circ G(y) = y$ becomes

$$\begin{pmatrix} f\begin{pmatrix} g_1\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ g_2\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \end{pmatrix} \\ g_2\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

(vi) In particular,

$$g_2\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = y_2,$$

(vii) and hence

$$f\begin{pmatrix} g_1\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ y_2 \end{pmatrix} = y_1.$$

(viii) Therefore, if we set

$$g(z) = g_1\begin{pmatrix} q \\ z \end{pmatrix},$$

then

$$f\begin{pmatrix} g(z) \\ z \end{pmatrix} = q.$$

(ix) That is, g is the desired implict function.

(x) Note: In practice, one chooses the variables $y_2 \in \mathbb{R}^{n-m}$ to be the ones that are to be the domain variables of the impicit function.

(c) *Implicit Function Theorem:*

(i) If dF_p invertible, then there exists a neighborhood U of p such that the restriction $f|_U$ has an implicit function.

(ii) Alternate version of hypothesis:

(A) Define $L_f : \mathbb{R}^m \to \mathbb{R}^m$ by

$$L_f(x) = \begin{pmatrix} df_p(x) \\ 0 \end{pmatrix}.$$

(B) Exercise: dF_p invertible $\Leftrightarrow L_f$ invertible.

(d) Example:

(i) Show that there exist functions $u = u(x, y)$ and $v = v(x, y)$ defined in a neighborhood of $(1, 1)$ so that $u(1, 1) = v(1, 1) = 1$,

$$e^{x^2-y^2} \cdot u^5 - v^3 = 0,$$

and

$$e^{u^2-v^2} \cdot x^2 - y^2 = 0$$

(ii) Define $f : \mathbb{R}^4 \to \mathbb{R}^2$ by

$$f \begin{pmatrix} u \\ v \\ x \\ y \end{pmatrix} = \begin{pmatrix} e^{x^2-y^2} \cdot u^5 - v^3 \\ e^{u^2-v^2} \cdot x^2 - y^2 \end{pmatrix}.$$

(iii) Note that $f(1, 1, 1, 1) = (0, 0)$.

(iv) Reinterpretation of problem: Construct implicit function for the $(0, 0)$-level set of f near $(u, v, x, y) = (1, 1, 1, 1)$.

(v) Define $F : \mathbb{R}^4 \to \mathbb{R}^4$ by

$$F \begin{pmatrix} u \\ v \\ x \\ y \end{pmatrix} = \begin{pmatrix} e^{x^2-y^2} \cdot u^5 - v^3 \\ e^{u^2-v^2} \cdot x^2 - y^2 \\ x \\ y \end{pmatrix}.$$

(vi) The Jacobian matrix of F at $p = (1, 1, 1, 1)$ is

$$\begin{pmatrix} 5 & -3 & 2 & -2 \\ 2 & -2 & 2 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(vii) The determinant of this matrix is $-4 \neq 0$ and hence dF_p invertible.

(e) Example continued: Compute $\frac{\partial u}{\partial y}$ at $(x, y) = (1, 1)$.

(i) $\frac{\partial u}{\partial y}(p)$ is the a_{12} entry of the Jacobian of g:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

(ii) Differentiating $f(g(x, y), x, y) = 1$ gives

$$\begin{pmatrix} 5 & -3 & 2 & -2 \\ 2 & -2 & 2 & -2 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

(iii) Thus, in particular, we have

$$\begin{aligned} -2 + 5a_{12} - 3a_{22} &= 0 \\ -2 + 2a_{12} - 2a_{22} &= 0. \end{aligned}$$

(iv) Solving gives $a_{12} = 1/2$.

1.1. Additional Exercises.

(1) The function $\frac{1}{32}x^4 + x^2y^2 - x^3 - y^3 - xy^3$ has critical points at $(24, 0)$ and $(0, 0)$. Determine whether these criticial points are local maximums, local minimums, or neither.

(2) Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a differentiable map. Show that if $||df_p - I||_2 < \frac{1}{2}$ for all $p \in \mathbb{R}^n$, then f is one-to-one, onto, and f^{-1} is also differentiable. Here $|| \cdot ||_2$ denotes the Hilbert-Schmidt norm of a linear transformation: Namely, for each linear transformation $L : \mathbb{R}^n \to \mathbb{R}^n$, define $||L|| = \sum_{i,j} L(e_i)^T \cdot L(e_j)$ where $\{e_i\}$ is the standard basis.

(3) Consider the following system of equations

$$\begin{aligned} x \cdot e^y &= u, \\ y \cdot e^x &= v. \end{aligned}$$

(a) Show that there exists an $\epsilon > 0$ such that given any u and v with $|u| < \epsilon$ and $|v| < \epsilon$, the above system has a unique solution $(x, y) \in \mathbb{R}^2$.

(b) Exhibit a pair $(u, v) \in \mathbb{R}^2$ such that there exist two distinct solutions to this system. Justify your answer.