Definition 1. Let $D$ be a compact set of $\mathbb{R}^n$. We say that the function $f : D \to \mathbb{R}$ has a local maximum (respectively, local minimum) on $D$ at the point $p \in D$ if and only if there exists an open ball $B \subset D$ centered at $p$ such that $f(x) \leq f(p)$ [respectively, $f(x) \geq f(p)$] for all points $x \in B$.

Recall the well-known result from single-variable calculus that if the differentiable function $f : \mathbb{R} \to \mathbb{R}$ has a local maximum or local minimum at $p \in \mathbb{R}$, then $f'(p) = 0$.

Lemma 1. Let $S \subset \mathbb{R}^n$, and $\varphi : \mathbb{R} \to S$ be a differentiable curve with $\varphi(0) = a$. If $f$ is a differentiable real-valued function defined on some open set containing $S$, and $f$ has a local maximum (or local minimum) on $S$ at $a$, then the gradient vector $\nabla f(a)$ is orthogonal to the velocity vector $\varphi'(0)$.

Proof. On the blackboard. \qed

Corollary 1. If $U$ is an open set of $\mathbb{R}^n$ and $a \in U$ is a point at which the differentiable function $f : U \to \mathbb{R}$ has a local maximum or local minimum, then $\nabla f(a) = 0$.

Definition 2. A set $M \subset \mathbb{R}^n$ is said to have a $k$-dimensional tangent plane at the point $a \in M$ if the union of all tangent lines to differentiable curves on $M$ passing through $a$ is a $k$-dimensional plane.

Definition 3. The projection mapping $\pi_i : \mathbb{R}^n \to \mathbb{R}^{n-1}$ is defined by removing the $i$th coordinate:

$$\pi_i(x_1, \ldots, x_n) = (x_1, \ldots, \hat{x}_i, \ldots, x_n) = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \mathbb{R}^{n-1}.$$
1 Single-constraint Optimization

Definition 4. The set $P \subset \mathbb{R}^n$ is called an $(n-1)$-dimensional patch if and only if for some integer $i$, $1 \leq i \leq n$, there exists a differentiable function $h : U \rightarrow \mathbb{R}$ for $U \subset \mathbb{R}^{n-1}$ open, such that

$$P = \{ x \in \mathbb{R}^n : \pi_i(x) \in U \text{ and } x_i = h(\pi_i(x)) \}.$$ 

NOTE: The concept of an $(n-1)$-dimensional patch is equivalent to having a permutation $x_i^1, \ldots, x_i^n$ of the coordinates $x_1, \ldots, x_n$ and a differentiable function $h : U \rightarrow \mathbb{R}$ on an open set $U \subset \mathbb{R}^{n-1}$ such that:

$$P = \{ x \in \mathbb{R}^n : (x_i^1, \ldots, x_i^{n-1}) \in U \text{ and } x_i^n = h(x_i^1, \ldots, x_i^{n-1}) \}.$$ 

Definition 5. The set $M \subset \mathbb{R}^n$ is called an $(n-1)$-dimensional manifold if and only if each point $a \in M$ lies in an open subset $U \subset \mathbb{R}^n$ such that $U \cap M$ is an $(n-1)$-dimensional patch.

Theorem 1. If $M$ is an $(n-1)$-dimensional manifold in $\mathbb{R}^n$, then at each of its points $M$ has an $(n-1)$-dimensional tangent plane.

Proof. On the blackboard. 

Theorem (Implicit Function Theorem). Let $g : \mathbb{R}^n \rightarrow \mathbb{R}$ be continuously differentiable and suppose that $g(a) = 0$ and $D_n g(a) \neq 0$. Then there exists a neighborhood $U$ of $a$ and a differentiable function $f : V \rightarrow \mathbb{R}$, with $V \subset \mathbb{R}^{n-1}$ a neighborhood of $(a_1, \ldots, a_{n-1})$, such that

$$U \cap g^{-1}(0) = \{ x \in \mathbb{R}^n : (x_1, \ldots, x_{n-1}) \in V \text{ and } x_n = f(x_1, \ldots, x_{n_1}) \}.$$ 

Theorem 2. Suppose that $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable. If $M$ is the set of all points $x \in S = g^{-1}(0)$ at which $\nabla g(x) \neq 0$, then $M$ is an $(n-1)$-manifold. Given $a \in M$, the gradient vector $\nabla g(a)$ is orthogonal to the tangent plane to $M$ at $a$.

Proof. On the blackboard.
Theorem 3. Suppose \( g : \mathbb{R}^n \rightarrow \mathbb{R} \) is continuously differentiable and let \( M \) be the set of points \( x \in \mathbb{R}^n \) at which \( g(x) = 0 \) and \( \nabla g(x) \neq 0 \). If the differentiable function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) attains a local maximum or minimum on \( M \) at the point \( a \in M \), then
\[
\nabla f(a) = \lambda \nabla g(a)
\]
for some number \( \lambda \), denoted as the Lagrange multiplier.

Proof. On the blackboard. \( \square \)

2 Multiple-constraint Optimization

Definition 6. The set \( P \subset \mathbb{R}^n \) is called a \( k \)-dimensional patch if and only if there exists a permutation \( x_{i_1}, \ldots, x_{i_n} \) of \( x_1, \ldots, x_n \), and differentiable function \( h : U \rightarrow \mathbb{R}^{n-k} \) for \( U \subset \mathbb{R}^k \) open, such that
\[
P = \{ x \in \mathbb{R}^n : (x_{i_1}, \ldots, x_{i_k}) \in U \text{ and } (x_{i_{k+1}}, \ldots, x_{i_n}) = h(x_{i_1}, \ldots, x_{i_k}) \}
\]

Definition 7. The set \( M \subset \mathbb{R}^n \) is called a \( k \)-dimensional manifold if and only if each point \( a \in M \) lies in an open subset \( U \subset \mathbb{R}^n \) such that \( U \cap M \) is a \( k \)-dimensional patch.

Theorem 4. If \( M \) is an \( k \)-dimensional manifold in \( \mathbb{R}^n \) then, at each of its points, \( M \) has a \( k \)-dimensional tangent plane.

Proof. On the blackboard. \( \square \)

Theorem (Implicit Mapping Theorem). Let \( g : \mathbb{R}^n \rightarrow \mathbb{R}^m \) \( (m < n) \) be a continuously differentiable map. Suppose that \( g(a) = 0 \) and that the rank of the derivative matrix \( g'(a) \) is \( m \). Then there exists a permutation \( x_{i_1}, \ldots, x_{i_n} \) of the coordinates in \( \mathbb{R}^n \), an open set \( U \subset \mathbb{R}^n \) containing \( a \), an open subset \( V \subset \mathbb{R}^{n-m} \) containing \( b = \pi_{n-m}(a_{i_1}, \ldots, a_{i_n}) \), and a differentiable mapping \( h : V \rightarrow \mathbb{R}^m \) such that each point \( x \in U \) lies on \( S = g^{-1}(0) \) if and only if \( (x_{i_1}, \ldots, x_{i_{n-m}}) \in V \) and
\[
(x_{i_{n-m+1}}, \ldots, x_{i_n}) = h(x_{i_1}, \ldots, x_{i_{n-m}}).
\]
Theorem 5. Suppose that \( g : \mathbb{R}^n \rightarrow \mathbb{R}^m \) is continuously differentiable. If \( M \) is the set of all points \( x \in S = g^{-1}(0) \) for which the rank of \( g'(x) \) is \( m \), then \( M \) is an \((n-m)\)-manifold. Given \( a \in M \), the gradient vectors \( \nabla g_1(a), \ldots, \nabla g_m(a) \) are all orthogonal to the tangent plane to \( M \) at \( a \).

Theorem 6. Suppose \( g : \mathbb{R}^n \rightarrow \mathbb{R}^m \) \((m < n)\) is continuously differentiable and let \( M \) be the set of points \( x \in \mathbb{R}^n \) such that \( g(x) = 0 \) and the gradient vectors \( \nabla g_1(a), \ldots, \nabla g_m(a) \) are linearly independent. If the differentiable function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) attains a local maximum or minimum on \( M \) at the point \( a \in M \), then there exist real numbers \( \lambda_1, \ldots, \lambda_m \) (called Lagrange multipliers) such that:

\[
\nabla f(a) = \lambda_1 \nabla g_1(a) + \ldots + \lambda_m \nabla g_m(a)
\]