Chapter 4. Submanifolds in $\mathbb{R}^n$ and Constrained Extremum Problems

We are now going to introduce the notion of submanifolds of $\mathbb{R}^n$ which are generalizations to general dimensions of smooth surfaces in $\mathbb{R}^3$.

4.1. Submanifolds in $\mathbb{R}^n$

Let us begin by looking at *hypersurfaces* in $\mathbb{R}^n$. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a C^1-function and let $M = \{x \in \mathbb{R}^n : f(x) = 0\} = f^{-1}(\{0\})$, that is, the level set of f corresponding to 0. If $Df(a) \neq 0$ for some $a \in M$, then, by the Implicit Function Theorem, we can represent M in a neighborhood of a as a graph of a function of $n-1$ variables. We may write $x_n = h(x_1, \ldots, x_{n-1})$ after a suitable reordering of coordinates. We are now going to generalize this situation to general k-dimensional submanifolds of $\mathbb{R}^n$:

Definition 4.1 A set $M \subseteq \mathbb{R}^n$ is called a *k-dimensional submanifold* of $\mathbb{R}^n$ if for every $x_0 \in M$, there exists an open neighborhood Ω of $x_0 \in \mathbb{R}^n$ and a C^1-function $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^{n-k}$ such that

$$M \cap \Omega = f^{-1}(\{0\}) \quad \text{and} \quad \operatorname{rank} Df(x) = n - k \quad \text{for any } x \in M \cap \Omega.$$

The next theorem, which is a consequence of the Implicit Function Theorem, will tell us that a submanifold can be locally represented as graphs of functions.

Theorem 4.2 *Let $M \subseteq \mathbb{R}^n$. Then the following statements are equivalent.*

(a) M is a k-dimensional submanifold of $\mathbb{R}^n$.

(b) For each $x_0 \in M$, we can (after suitably relabeling the coordinates) write $x_0 = (y_0, z_0)$ with $y_0 \in \mathbb{R}^k$, $z_0 \in \mathbb{R}^{n-k}$ and find an open neighborhood U of y_0 in $\mathbb{R}^k$, an open neighborhood V of $z_0 \in \mathbb{R}^{n-k}$, and a C^1-function $g : U \longrightarrow V$ with $g(y_0) = z_0$ such that

$$M \cap (U \times V) = \{(y, g(y)) : y \in U\}.$$

Remarks:

(a) By Theorem 4.2, we see that for each point in a k-dimensional submanifold, there is a neighborhood which can be identified with an open set in $\mathbb{R}^k$. Thus we can think of submanifolds as subsets of $\mathbb{R}^n$ that locally look like open sets in Euclidean space $\mathbb{R}^k$.

(b) In fact, it is possible to use the idea above to define abstract k-dimensional manifolds without reference to an embedding in $\mathbb{R}^n$. Roughly speaking, such a manifold is a space covered by open sets ("charts"), each looks like open sets in $\mathbb{R}^k$, such that the charts fit together smoothly in a suitable sense. It is in fact possible to transfer the tools of differential and integral calculus to such abstract manifolds.

Example 4.3 (Curves in $\mathbb{R}^2$)

(a) The unit circle in $\mathbb{R}^2$ is given by

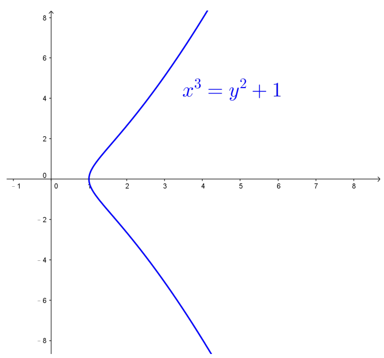
$$S^1 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

Let $f(x,y) = x^2 + y^2 - 1$. Then $S^1 = f^{-1}(\{0\})$ and $f'(x,y) = (2x \;\; 2y)$ has rank 1 at all points of the unit circle (it will have rank 0 at origin which does not lie on the circle). So it is a 1-dimensional submanifold of $\mathbb{R}^2$.

(b) Let

$$M = \{(x,y) \in \mathbb{R}^2 : x^3 = y^2 + 1\}.$$

Then $M = f^{-1}(\{0\})$, where $f(x,y) = x^3 - y^2 - 1$. Again, $f'(x,y) = (3x^2 \;\; -2y)$ has rank 1 at all points of M. So M defines a 1-dimensional submanifold of $\mathbb{R}^2$.



(c) Let

$$M = \{(x,y) \in \mathbb{R}^2 : x^2 - y^2 = 0\}.$$

Then $M = f^{-1}(\{0\})$, where $f(x,y) = x^2 - y^2$. This time $f'(x,y) = (2x \;\; -2y)$ has rank 1 except at the origin, which is on the curve. So the curve M is a 1-dimensional submanifold away from the origin. Geometrically, we can see that the curve is the union of the two lines $y = x$ and $y = -x$ which meet at $(0,0)$. $\square$

Example 4.4 (Ellipsoids and Unit Sphere in $\mathbb{R}^3$) An *ellipsoid* in $\mathbb{R}^3$ is given by

$$M = \left\{(x,y,z) \in \mathbb{R}^3 : f(x,y,z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0\right\}$$

for some $a, b, c > 0$. Note that this defines a 2-dimensional submanifold of $\mathbb{R}^3$. Indeed,

$$f'(x,y,z) = \begin{pmatrix} \frac{2x}{a^2} & \frac{2y}{b^2} & \frac{2z}{c^2} \end{pmatrix}$$

has rank 1 for any $(x,y,z) \in M$. The *unit sphere* in $\mathbb{R}^3$, denoted by S^2, is an ellipsoid with $a = b = c = 1$. $\square$

Example 4.5 (Torus in $\mathbb{R}^3$) Let $r, R \in \mathbb{R}$ with $0 < r < R$. A *torus* in $\mathbb{R}^3$ is given by

$$T = \Big\{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = \big(\sqrt{x^2 + y^2} - R\big)^2 + z^2 - r^2 = 0\Big\}.$$

Thus the torus consists of the points in $\mathbb{R}^3$ which have distance r to a circle on the xy-plane centered at the origin with radius R. Since $r < R$, the torus T does not contain any point on the z-axis, that is, points of the form $(0, 0, z)$ do not lie on T. Then

$$f'(x, y, z) = \begin{pmatrix} \dfrac{2x\big(\sqrt{x^2 + y^2} - R\big)}{\sqrt{x^2 + y^2}} & \dfrac{2y\big(\sqrt{x^2 + y^2} - R\big)}{\sqrt{x^2 + y^2}} & 2z \end{pmatrix}$$

is well-defined on T and we see that $f'(x, y, z)$ has rank 1 for any $(x, y, z) \in T$. Hence T is a 2-dimensional submanifold of $\mathbb{R}^3$. $\square$

We now define the tangent space at a point of the submanifold.

Definition 4.6 Let $M \subseteq \mathbb{R}^n$ be a k-dimensional submanifold of $\mathbb{R}^n$ and let $x_0 \in M$.

(a) A vector v is called a *tangent vector* to M at x_0 if there exist $\epsilon > 0$ and a C^1-function $\gamma : (-\epsilon, \epsilon) \longrightarrow \mathbb{R}^n$ such that $\gamma(t) \in M$ for any $t \in (-\epsilon, \epsilon)$, $\gamma(0) = x_0$ and $\gamma'(0) = v$.

(b) The set of all tangent vectors to M at x_0 is called the *tangent space* to M at x_0 and is denoted by $T_{x_0}M$.

(c) A vector w is called a *normal vector* to M at x_0 if $\langle w, v\rangle = 0$ for all $v \in T_{x_0}M$. Thus the set of all normal vectors to M at x_0 is precisely the orthogonal complement $(T_{x_0}M)^\perp$ of $T_{x_0}M$ in $\mathbb{R}^n$.

Next we prove the generalization of the property that "the gradient is perpendicular to the level sets of a function".

Theorem 4.7 *Let $M \subseteq \mathbb{R}^n$ be a k-dimensional submanifold of $\mathbb{R}^n$ and let $x_0 \in M$. Let Ω and f be the same as in Definition 4.1. Then we have*

$$T_{x_0}M = \ker Df(x_0),$$

that is, the tangent space equals the kernel of the linear map $Df(x_0) : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-k}$.

Remarks:

(a) This result in particular shows that $T_{x_0}M$ is indeed a k-dimensional vector space, and as a consequence the space of normal vectors is an $(n-k)$-dimensional vector space.

(b) Define the *tangent k-plane* to M at x_0 by

$$\{x_0 + v : v \in T_{x_0}M\}.$$

This is a k-dimensional affine space in $\mathbb{R}^n$. By Theorem 4.7, the tangent k-plane to M at x_0 is given by

$$\{x \in \mathbb{R}^n : Df(x_0)(x - x_0) = 0\}.$$

Example 4.8 From Example 4.4, the ellipsoid M is a 2-dimensional submanifold of $\mathbb{R}^3$. Furthermore, we have

$$f'(x,y,z) = \begin{pmatrix} \frac{2x}{a^2} & \frac{2y}{b^2} & \frac{2z}{c^2} \end{pmatrix}.$$

According to Theorem 4.7 and its remark, the tangent 2-plane at any $(x_0, y_0, z_0) \in M$ is given by

$$\left\{ (x,y,z) \in \mathbb{R}^3 : \begin{pmatrix} \frac{2x_0}{a^2} & \frac{2y_0}{b^2} & \frac{2z_0}{c^2} \end{pmatrix} \cdot \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} = 0 \right\},$$

which can be simplified to

$$\left\{ (x,y,z) \in \mathbb{R}^3 : \frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} + \frac{z_0 z}{c^2} = 1 \right\}. \qquad \square$$

4.2. Constrained Extremum Problems

In this section we want to consider the problem of extremizing a function on a submanifold M of $\mathbb{R}^n$ given by a system of equations. The following theorem will provide a necessary condition for an extremal point.

Theorem 4.9 (Theorem on Lagrange Multipliers) *Let Ω be an open subset of $\mathbb{R}^n$ and let $g : \Omega \longrightarrow \mathbb{R}$ and $f : \Omega \longrightarrow \mathbb{R}^{n-k}$ be C^1-functions. If $x_0 \in f^{-1}(\{0\})$ is a local extremum of g on $f^{-1}(\{0\})$, that is, if there exists an open neighborhood V of x_0 such that*

$$g(x) \geq g(x_0) \quad (\text{or } g(x) \leq g(x_0)) \quad \text{for any } x \in V \text{ with } f(x) = 0,$$

and if rank $Df(x_0) = n - k$, then there exist $\lambda_1, \ldots, \lambda_{n-k} \in \mathbb{R}$ called the Lagrange multipliers such that

$$\nabla g(x_0) = \sum_{i=1}^{n-k} \lambda_i \nabla f_i(x_0).$$

Remark: We can interpret this geometrically as saying that $\nabla g(x_0)$ is normal to the submanifold $f^{-1}(\{0\})$ and the normal space is spanned by $\nabla f_i(x_0)$ for $i = 1, \ldots, n-k$.

Example 4.10 Extremize $g(x,y,z) = 5x + y - 3z$ on the intersection of the plane $\{(x,y,z) \in \mathbb{R}^3 : x + y + z = 0\}$ with the unit sphere $S^2 = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$.

Solution. Define

$$f(x,y,z) = (x + y + z,\ x^2 + y^2 + z^2 - 1) \quad \text{and} \quad M = \{(x,y,z) \in \mathbb{R}^3 : f(x,y,z) = (0,0)\}.$$

Then

$$Df(x,y,z) = \begin{pmatrix} 1 & 1 & 1 \\ 2x & 2y & 2z \end{pmatrix}$$

and we have rank $Df(x,y,z) = 2$ for any $(x,y,z) \in M$. Hence M is a 1-dimensional submanifold of $\mathbb{R}^3$. Now since M is compact and g is continuous, it follows that g attains its maximum and minimum values on M. If (x_0, y_0, z_0) is an extremal point of g on M, then Theorem 4.9 implies that there exist $\lambda_1, \lambda_2 \in \mathbb{R}$ such that

$$\nabla g(x_0, y_0, z_0) = \lambda_1 \nabla f_1(x_0, y_0, z_0) + \lambda_2 \nabla f_2(x_0, y_0, z_0),$$

that is,

$$\begin{pmatrix} 5 \\ 1 \\ -3 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2x_0 \\ 2y_0 \\ 2z_0 \end{pmatrix}.$$

Solving this together with the constraints $f(x_0, y_0, z_0) = 0$ yields

$$(x_0, y_0, z_0) = \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}} \right) \quad \text{or} \quad \left(\frac{-1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right).$$

It follows that g attains its maximum at $(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}})$ with value $4\sqrt{2}$ and its minimum at $(\frac{-1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$ with value $-4\sqrt{2}$. □

Example 4.11 (Eigenvalues of symmetric matrices) Let A be an $n \times n$ symmetric matrix and let $S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$ be the unit sphere in $\mathbb{R}^n$. Define $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ by

$$g(x) = \langle x, Ax \rangle = x^T Ax.$$

Extremize $g(x)$ on the unit sphere S^{n-1}.

Solution. First of all, we observe that S^{n-1} is an $(n-1)$-dimensional submanifold of $\mathbb{R}^n$ since $S^{n-1} = f^{-1}(\{0\})$ where $f(x) = \|x\|^2 - 1$ with rank $Df(x) = 1$ for any $x \in S^{n-1}$. As S^{n-1} is compact and g is continuous, it follows that g attains its maximum and minimum values on S^{n-1}. If $x = x_0$ is an extremal point of g on S^{n-1}, then Theorem 4.9 implies that there exists $\lambda \in \mathbb{R}$ such that

$$\nabla g(x_0) = \lambda \nabla f(x_0), \quad \text{that is,} \quad Ax_0 = \lambda x_0.$$

This means that x_0 ($\neq 0$) is an eigenvector of A with eigenvalue λ. Furthermore, we have

$$\lambda = \lambda \langle x_0, x_0 \rangle = \langle x_0, \lambda x_0 \rangle = \langle x_0, Ax_0 \rangle = g(x_0).$$

Since $g(x_0)$ is an extreme value of g, thus we have shown that the maximum value of g on S^{n-1} equals the largest eigenvalue of A, and the minimum value of g on S^{n-1} equals the smallest eigenvalue of A. □