Chapter 6

Curves and Surfaces

In Chapter 2 a plane is defined as the zero set of a linear function in $\mathbb{R}^3$. It is expected a surface is the zero set of a differentiable function in $\mathbb{R}^n$. To motivate, graphs and their tangent planes are discussed in Section 6.1. The definition of a surface is defined and shown to be the zero set of a differentiable function whenever it has non-vanishing gradient in Section 6.2. A space curve is defined as the common zero set of two differentiable functions with linearly independent gradients in Section 6.3. The general case concerning common zero sets of finitely many differentiable functions in $\mathbb{R}^n$ is described by the Implicit Function Theorem in Section 6.4. In Section 6.5 the Inverse Function Theorem which gives a criterion for the existence of a local inverse to a differentiable map is present.

6.1 Graphs as Surfaces

Everybody is likely to know what a surface is and recognize it whenever they see one. However, to make it precise in mathematics we need a definition. One must be careful when it comes to a definition. Here we start with a special class of geometric objects which should be called surfaces no matter what the definition would be.

Let φ be a continuous function defined in an open set G in $\mathbb{R}^2$. Everybody agrees that its graph

$$\Sigma = \{(x, y, \varphi(x, y)) : \ (x, y) \in G\} \subset \mathbb{R}^3 \ ,$$

is a surface. This surface may be called a differentiable surface when its defining function is differentiable. We can talk about the tangent plane of a differentiable surface. Let $p = (x, y, \varphi(x, y))$ be a point on such a surface Σ. We define its tangent vector at p first. The idea is to look at an arbitrary regular parametric curve $\gamma : (-a, a) \to \mathbb{R}^2$ passing through (x, y) at $t = 0$ in $\mathbb{R}^2$. It lifts to a parametric curve $c(t) \equiv (\gamma(t), \varphi(\gamma(t)))$ on Σ passing through $(\gamma(0), \varphi(\gamma(0))) = p$. Its velocity vector $c'(0)$ is called a **tangent vector** of Σ at p. Geometrically it is evident that it is tangential to Σ at p. The **tangent plane** of Σ at p is the collection of all these tangent vectors. In other words, we set

$$T_p\Sigma \equiv \{c'(0): \ c(t) = (\gamma(t), \varphi(\gamma(t)))\},$$

where γ's range over all regular parametric curves described above. The definition is messy at first sight. The following theorem shows that it could be described in a very simple manner.

Theorem 6.1. *Setting as above, $T_p\Sigma$ is the vector space spanned by the vectors*

$$(1, 0, \varphi_x(x, y)) \text{ and } (0, 1, \varphi_y(x, y)) .$$

Furthermore, $(-\varphi_x, -\varphi_y, 1)$ points in the normal direction.

A non-zero vector is said to point to the normal direction (or simply a normal vector) at a point on a surface if it is perpendicular to the tangent plane of the surface at this point.

Proof. We first show that $(1, 0, \varphi_x(x, y))$ is a tangent vector. To this end we choose the parametric curve to be the straight line $l(t) = (x, y) + t(1, 0)$. Then $l'(0) = (1, 0)$ and, by the Chain Rule,

$$c'(0) = (1, 0, \nabla\varphi(x, y) \cdot (1, 0)) = (1, 0, \varphi_x(x, y)) \in T_p\Sigma .$$

Similarly we can show that $(0, 1, \varphi_y(x, y))$ is also a tangent vector. Now, given any curve $\gamma(t) = (x(t), y(t)),\ (x(0), y(0)) = (x, y)$, and $c(t) = (x(t), y(t), \varphi(x(t), y(t)))$,

$$\begin{aligned} c'(0) &= (x'(0), y'(0), \varphi_x(x, y)x'(0) + \varphi_y(x, y)y'(0)) \\ &= x'(0)(1, 0, \varphi_x(x, y)) + y'(0)(0, 1, \varphi_y(x, y)) , \end{aligned}$$

which shows that every tangent vector can be expressed as the linear combination of $(1, 0, \varphi_x(x, y))$ and $(0, 1, \varphi_y(x, y))$. It follows that all tangent vectors form a subspace in $\mathbb{R}^3$ spanned by these two obviously linearly independent tangent vectors. Now, observing that

$$(-\varphi_x, -\varphi_y, 1) \cdot (1, 0, \varphi_x) = 0 , \quad (-\varphi_x, -\varphi_y, 1) \cdot (0, 1, \varphi_y) = 0 ,$$

we see that $(-\nabla\varphi, 1)$ points to the normal direction at p.

$\square$

Example 6.1. Find the tangent plane and the normal direction of the surface defined by the equation $xy^2 - xz + y = 0$ at $(1, 2, 6)$. We express this equation as a graph

$$z = \frac{xy^2 + y}{x} ,$$

which passes the point $(1, 2, 6)$. From

$$\frac{\partial z}{\partial x} = -\frac{y}{x^2}\ , \quad \frac{\partial z}{\partial y} = \frac{2xy+1}{x}\ ,$$

we get

$$\frac{\partial z}{\partial x}(1,2) = -2\ , \quad \frac{\partial z}{\partial y}(1,2) = 5\ .$$

Therefore, the tangent plane of the surface at $(1, 2, 6)$ is spanned by the vectors $(1, 0, z_x)$ and $(0, 1, z_y)$ at $(1, 2, 6)$, that is,

$$(1, 0, -2)\ , \quad (0, 1, 5)\ ,$$

and the vector $(-\nabla z, 1) = (2, -5, 1)$ points in the normal direction at $(1, 1, 2)$. The tangent plane is given by

$$(2, -5, 1) \cdot ((x, y, z) - (1, 2, 6)) = 0, \qquad \text{i.e.,} \quad 2x - 5y + z + 2 = 0\ .$$

6.2 Level Sets as Surfaces

Consider a function f defined in some open set of $\mathbb{R}^3$. We would like to examine the set $\{(x, y, z) : f(x, y, z) = c\}$, that is, the level set of this function at c. By replacing the function f by $f - c$, it is without loss of generality to assume c is equal to 0. In view of this, in the following discussion we usually assume $c = 0$. Just like the level sets of linear equations define planes and those of quadratic equations define quadric surfaces, we would like to study what geometric object that is associated to the level set of a general differentiable function. Let us start with an example.

Example 6.2. Study the zero sets of the following functions.

(a) $z - x^2 + y^3 - 1 = 0$,

(b) $xyz - 4 = 0$,

(c) $x^2 + y^2 + z^2 - 1 = 0$,

(d) $x + 2y + ze^{x+y+z} = 0$.

The equation in (a) can be written as $z = x^2 - y^3 + 1$, so it is the graph of the function $\varphi(x, y) = x^2 - y^3 + 1$ over $\mathbb{R}^2$. It should be a surface according to our discussion in the previous section. For (b), we see that no x, y or z can be 0, so it is the graph of the function $f(x,y) = 4/xy$ over the open set $\{(x,y):\ x,y \neq 0\}$. Of course, we may use the other two variables as the independent ones. For instance, the set can also be described as the graph of $g(y,z) = 1/yz$ over the open set $\{(y,z):\ y,z \neq 0\}$. Now (c) describes the unit sphere. By writing the equation as

$$z = \sqrt{1-x^2-y^2}\ , \text{ or } z = -\sqrt{1-x^2-y^2}\ ,$$

the unit sphere can be described as the union of the graphs of two functions over the closed unit disk $\{(x,y) : x^2+y^2 \leq 1\}$. Finally, the equation in (d) is so messy that it is not clear how to express one variable in terms of the other two, hence we are not sure whether it is surface or not.

From these examples, we see that the level sets defined by an equation can be a graph over the entire plane as in (a), over some open set as in (b), expressed as the union of more than one graphs as in (c), or too messy to describe as in (d). In any case, they hint at how a surface should be defined, namely, a surface should be an subset of $\mathbb{R}^3$ that looks like a graph locally. To be precise let us fix some notations. For $p = (x,y,z)$, denote $C_a(p)$ or simply C a cube centered at p with side $2a$, that is, $(x-a,x+a)\times(y-a,y+a)\times(z-a,z+a)$ and let $C_{xy} = (x-a,x+a)\times(y-a,y+a)$ be its projection on the xy-plane. Similarly we define C_{yz} and C_{zx}. A subset Σ of $\mathbb{R}^3$ is a **surface** if, for every point $p \in \Sigma$, at least one of the following cases holds:

(a) there exists a cube C centered at p and a function φ from C_{xy} to $(z-a,z+a)$ such that

$$\Sigma \cap C = \{(x,y,\varphi(x,y)):\ (x,y) \in C_{xy}\}\,,$$

or

(b) there exists a cube C centered at p and a function ψ from C_{zx} to $(y-a,y+a)$ such that

$$\Sigma \cap C = \{(x,\psi(x,z),z):\ (x,z) \in C_{zx}\}\,,$$

or

(c) there exists a cube C centered at p and a function η from C_{yz} to $(x-a,x+a)$ such that

$$\Sigma \cap C = \{(\eta(y,z),y,z):\ (y,z) \in C_{yz}\}\,.$$

It is called a **differentiable surface** if the functions φ,ψ,η can be chosen to be differentiable. It is called a **smooth surface** if they are smooth, that is, infinitely many times differentiable.

Example 6.3. Example 6.2 (a) is graph over the xy-plane, so case (a) is satisfied for an arbitrary cube on the graph. Next the unit sphere is also a surface according to this definition. Near the point $(0,0,1)$ case (a) holds by taking $\varphi(x,y) = (1-x^2-y^2)^{1/2}$.

However, at $(1,0,0)$ we need to consider case (b) by taking $\psi(y,z) = (1-y^2-z^2)^{1/2}$. Near $(1,1,1)/\sqrt{3}$, all cases hold and the sphere could be described locally by

$$z = \sqrt{1-x^2-y^2}\ , \quad y = \sqrt{1-x^2-z^2}\ , \quad \text{or } x = \sqrt{1-y^2-z^2}\ .$$

The following fundamental result gives a simple criterion to test when the zero set of a differentiable function forms a differentiable surface.

Theorem 6.2. *Let f be differentiable in the open set G in $\mathbb{R}^3$ and Σ its zero set. Suppose $p = (x_0, y_0, z_0) \in G$ satisfies $f(p) = 0$ and $\nabla f(p) \neq (0,0,0)$. Then there is a cube C centered at p such that*

(a) If $f_z(p) \neq 0$, there exists a function φ from C_{xy} to C satisfying $\varphi(x_0, y_0) = z_0$ such that

$$\Sigma \cap C = \{(x, y, \varphi(x,y)) :\ (x,y) \in C_{xy}\}\,,$$

(b) If $f_x(p) \neq 0$, there exists a function η from C_{yz} to C satisfying $\eta(y_0, z_0) = x_0$ such that

$$\Sigma \cap C = \{(\eta(y,z), y, z) :\ (x,y) \in C_{yz}\}\,,$$

(c) If $f_y(p) \neq 0$, there exists a function ψ from C_{zx} to C satisfying $\psi(x_0, z_0) = y_0$ such that

$$\Sigma \cap C = \{(x, \psi(x,z), z) :\ (x,z) \in C_{zx}\}\,.$$

Corollary 6.3. *Let f be differentiable in the open G in $\mathbb{R}^3$ and $c \in \mathbb{R}$. Suppose that*

$$\Sigma = \{(x,y,z) :\ f(x,y,z) = c\ , \nabla f(x,y,z) \neq (0,0,0)\}$$

is non-empty. Then Σ is a differentiable surface.

Proof. Apply the Theorem 6.2 to the function $f_c(x,y,z) = f(x,y,z) - c$ and note that $\nabla f_c = \nabla f$.

□

Example 6.4. Looking at Example 6.2 (c) above, the gradient of the unit sphere is given by $2(x,y,z)$ which never vanishes on $x^2+y^2+z^2 = 1$. By this corollary it is a differentiable surface. At $(0,0,1)$ the gradient is $(0,0,2)$. Theorem 6.2 asserts that it can be described as a function $\varphi(x,y)$ near this point. In fact the function is given by

$$\varphi(x,y) = \sqrt{1-x^2-y^2}\ .$$

On the other hand, it cannot be represented as graphs of functions in (z,x) or (y,z).

Example 6.5. Now, let us examine Example 6.2 (d) by letting $f(x,y,z) \equiv x + 2y + ze^{x+y+z}$ so (d) is its zero set. We compute

$$\nabla f(x,y,z) = (1 + ze^{x+y+z}, 2 + ze^{x+y+z}, e^{x+y+z} + ze^{x+y+z}) \ .$$

By simply looking at the first two components of the gradient, ∇f never vanishes. By the corollary $x + 2y + ze^{x+y+z} = 0$ defines a differentiable surface despite we cannot solve this equation explicitly. It demonstrates the power of this criterion.

Example 6.6. Theorem 6.2 does not necessarily hold when the gradient vanishes. For instance, $\{(x,y,z) : \ x^2 + y^2 + z^2 = 0\}$ which consists of a single point $\{(0,0,0)\}$ cannot be a surface. Here we have $\nabla f(0,0,0) = (0,0,0)$. Similarly, $\{(x,y,z) : \ x^2 + y^2 = 0\}$ is the z-axis which is a straight line not a surface. We have $\nabla f(0,0,z) = (0,0,0)$ at every point on the z-axis.

Example 6.7. Let us look at $\{(x,y,z) : (x - y + z - 1)^2 = 0\}$. On one hand, this set is the same as $\{(x,y,z) : \ x - y + z - 1 = 0\}$, so it is a surface (in fact, a plane). On the other hand, $\nabla g = (2(x-y+z-1), 2(x-y+z-1)(-1), 2(x-y+1))$ identically vanishes on the set. This example shows that non-vanishing of the gradient is a sufficient but not a necessary condition for a local graph.

After knowing that $f(x,y,z) = 0$ defines z as a function of x and y, that is, $z = \varphi(x,y)$, we may use the technique of **implicit differentiation** to find the derivatives of φ. Usually people do not introduce new notation. Imagining z now a function of (x,y), we differentiate the equation $f(x,y,z) = 0$ by the Chain Rule:

$$f_x + f_z z_x = 0, \quad f_y + f_z z_y = 0 \ ,$$

so

$$z_x = \frac{-f_x}{f_z} \ , \quad z_y = \frac{-f_y}{f_z} \ .$$

By further differentiating,

$$f_{xx} + 2f_{zx}z_x + f_{zz}z_x^2 + f_z z_{xx} = 0, \quad f_{xy} + f_{xz}z_y + f_{zy}z_x + f_{zz}z_x z_y + f_z z_{xy} = 0 \ ,$$

$$f_{yx} + f_{yz}z_x + f_{zx}z_y + f_{zz}z_x z_y + f_z z_{yx} = 0, \quad f_{yy} + 2f_{yz}z_y + f_{zz}z_y^2 + f_z z_{yy} = 0 \ .$$

Thus z_{xx}, z_{xy}, and z_{yy} can be expressed in terms of z_x, z_y and the first and second derivatives of f. All higher order partial derivatives of z can be obtained by differentiating the resulting equations in a recursive way. In practise, no one would memorize these formulas. One just goes ahead to differentiate the equation many times.

Example 6.8. Consider the equation

$$x^2 + y^2 + z^2 - xz - x = 0 \ .$$

Find the derivative of z as a function of x, y at the point $(0, 2, -1)$. We have

$$\nabla f = (2x - z - 1, 2y, 2z - x) \ ,$$

so $\nabla f(0, 2, -1) = (0, 4, -2)$ is non-zero. According to Theorem 6.2, the z-component can be represented as a function $\varphi(x, y)$. We differentiate the equation in x and y respectively to get

$$2x + 2z\frac{\partial z}{\partial x} - z - x\frac{\partial z}{\partial x} - 1 = 0, \quad 2y + 2z\frac{\partial z}{\partial y} - x\frac{\partial z}{\partial y} = 0 \ .$$

We can further express it as

$$\frac{\partial z}{\partial x} = \frac{1 - 2x + z}{x + 2z} \ , \quad \frac{\partial z}{\partial y} = \frac{2y}{x - 2z} \ .$$

Since there is no explicit expression of z in terms of x and y, this is the farthest we could reach. Evaluating at $(0, 2, -1)$, we get

$$\frac{\partial z}{\partial x}(0, 2) = 0, \quad \frac{\partial z}{\partial y}(0, 2) = 2 \ .$$

Higher order derivatives can also be found by implicit differentiation. For instance, to determine $\partial^2 z/\partial x^2(0, 2)$, we differentiate

$$2x + (2z + x)\frac{\partial z}{\partial x} - z - 1 = 0 \ ,$$

once more to get

$$2 + 2\frac{\partial z}{\partial x}\frac{\partial z}{\partial x} + \frac{\partial z}{\partial x} + (2z + x)\frac{\partial^2 z}{\partial x^2} - \frac{\partial z}{\partial x} = 0 \ .$$

Plugging in $(0, 2, -1)$ and $z_x = 0$, we get $z_{xx} = 1$ at $(0, 2)$.

One may also obtain the second derivatives by differentiating the expression of z_x and z_y listed above, but implicit differentiation is usually more efficient.

We will define the tangent plane of Σ at p to be the tangent plane of the local graph representation. Since there could be more than one representations, it is not clear whether the tangent planes found from different representation are the same. The following theorem not only gives a positive answer to this question but also yield a simple description of the tangent plane.

Theorem 6.4. *Let f be differentiable in some open set G in $\mathbb{R}^3$. Suppose that $\nabla f(p) \neq (0,0,0)$ at some $p \in G$. Then it is normal to the tangent plane of the level set passing p. As a result, the tangent plane of $\{(x,y,z) : \ f(x,y,z) = f(p)\}$ is given by*

$$\nabla f(p) \cdot ((x,y,z) - p) = 0 \ .$$

For clarity the **tangent plane** of $\Sigma = \{(x,y,z) : \ f(x,y,z) = 0\}$ at p is defined to be the hyperplane

$$T_p\Sigma = \{(x,y,z) : \ \nabla f(p) \cdot ((x,y,z) - p) = 0\} \, .$$

Proof. Without loss of generality assume $f_z(p) \neq 0$. By Theorem 6.2 for some small cube C, the set $\{(x,y,x) : f(x,y,z) = c\} \cap C, \ c = f(p)$, coincides with the graph $\{(x,y,\varphi(x,y)) : \ (x,y) \in C_{xy}\}$ for some φ. We have $f(x,y,\varphi(x,y)) = c$. Taking partial derivatives of this relation,

$$(1,0,\varphi_x) \cdot \nabla f(p) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}(p)\frac{\partial \varphi}{\partial x}(x,y) = 0 \ ,$$

and

$$(0,1,\varphi_y) \cdot \nabla f(p) = \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}(p)\frac{\partial \varphi}{\partial y}(x,y) = 0 \ ,$$

so ∇f is perpendicular to the tangent plane of the level surface at p.

□

In $\mathbb{R}^n$ a **hypersurface** is defined as a surface in $\mathbb{R}^3$ where now the cases (a), (b) and (c) have to be replaced by n-many cases. Note that it is called a curve when $n = 2$ and a surface when $n = 3$. With these definitions, analogs of Theorem 6.2, Corollary 6.3 and Theorem 6.4 hold, where now the **tangent hyperplane** is defined as $\{x \in \mathbb{R}^n : \ \nabla f(p) \cdot (x - p) = 0 \ \}$. In particular, the level set $\{x \in G : \ f(x) = c\}, \ G \subset \mathbb{R}^n$, defines a differentiable hypersurface near a point p if f is differentiable with non-zero gradient at p.

6.3 Space Curves

In Section 2.3 we defined the straight line as the intersection of two planes or the zero set of a system of two linear equations. Now we would like to describe a curve as the common zero set of a system of two differentiable functions. Parallel to what has been done for surfaces, we first define what a curve is.

A subset Γ of $\mathbb{R}^3$ is a **curve** if, for every point $p = (x,y,z) \in \Gamma$, at least one of the following cases holds:

(a) there exists a cube C centered at p and a function γ from $(x-a, x+a)$ to $(y-a, y+a) \times (z-a, z+a)$ such that

$$\Gamma \cap C = \{(t, \gamma(t)) : \ t \in (x-a, x+a) \ \},$$

or

(b) there exists a cube C centered at p and a function β from $(y-a, y+a)$ to $(x-a, x+a) \times (z-a, z+a)$ such that

$$\Gamma \cap C = \{(\beta_1(t), t, \beta_2(t)) : \ t \in (y-a, y+a) \ \},$$

or

(c) there exists a cube C centered at p and a function τ from $(z-a, z+a)$ to $(x-a, x+a) \times (y-a, y+a)$ such that

$$\Gamma \cap C = \{(\tau(t), t)) : \ t \in (z-a, z+a) \ \}.$$

It is called a **differentiable curve** (resp. **smooth curve**) if the functions γ, β, τ can be chosen to be differentiable (resp. smooth). From this definition we see that a differentiable curve is the image of a regular parametric curve. For instance, in (a) the parametric curve is given by

$$\tilde{\gamma}(t) = (t, \gamma(t)), \quad t \in (x-a, x+a) \ ,$$

so $|\tilde{\gamma}(t)| = \sqrt{1 + \gamma_1'^2(t) + \gamma_2'^2(t)} > 0$.

To formulate our analog of Theorem 6.2 for curves, for any two vectors (a_1, a_2, a_3) and (b_1, b_2, b_3), we form the matrix

$$M = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$$

and use M_j to denote the 2×2-submatrix obtained by removing the j-column of M. When these two vectors are linearly independent, M has rank 2, so at least one of M_1, M_2 and M_3 is non-singular. In the following theorem M stands for

$$\begin{bmatrix} f_x & f_y & f_z \\ g_x & g_y & g_z \end{bmatrix} .$$

Theorem 6.5. *Let f and g be differentiable in the open set G in $\mathbb{R}^3$. Suppose $p = (x_0, y_0, z_0)$ satisfies $f(p) = 0$, $g(p) = 0$ and $\nabla f(p)$, $\nabla g(p)$ are linearly independent. There is a cube C centered at p such that*

(a) If M_1 is non-singular at p, there exists a function γ from $(x_0 - a, x_0 + a)$ to $(y_0 - a, y_0 + a) \times (z_0 - a, z_0 + a)$ satisfying $\gamma(x_0) = (y_0, z_0)$ such that

$$\{(x, y, z) : \ f(x, y, z) = 0, \ g(x, y, z) = 0\} \cap C = \{(x, \gamma(x)) : \ x \in (x_0 - a, x_0 + a) \ \},$$

(b) If M_2 is non-singular at p, there exists a function β satisfying $\beta(y_0) = (x_0, z_0)$ from $(y_0 - a, y_0 + a)$ to $(x_0 - a, x_0 + a) \times (z_0 - a, z_0 + a)$ such that

$$\{(x,y,z):\ f(x,y,z)=0,\ g(x,y,z)=0\} \cap C = \{(\beta_1(y), y, \beta_2(y)):\ y \in (y_0 - a, y_0 + a)\ \},$$

(c) If M_3 is non-singular at p, there exists a function τ satisfying $\tau(z_0) = (x_0, y_0)$ from $(z_0 - a, z_0 + a)$ to $(x_0 - a, x_0 + a) \times (z_0 - a, z_0 + a)$ such that

$$\{(x,y,z):\ f(x,y,z)=0,\ g(x,y,z)=0\} \cap C = \{(\tau(z), z):\ z \in (z_0 - a, z_0 + a)\ \}.$$

Corollary 6.6. *Let f and g be differentiable in the open set G in $\mathbb{R}^3$ and $c, d \in \mathbb{R}$. Suppose that*

$$\Gamma = \{p:\ f(p) = c,\ g(p) = d, \quad \nabla f \text{ and } \nabla g \text{ are linearly independent at } p\}$$

is non-empty. Then Γ is a differentiable curve.

Example 6.9. Consider

$$\begin{cases} x^2 + y^2 + z^2 & = 1 \\ x - y + 2z & = 0 \ . \end{cases}$$

Denote the first function by f and the second by g. Their gradients are given respectively by

$$\nabla f = 2(x,y,z) \ , \quad \nabla g = (1, -1, 2) \ .$$

One can check that they are always linearly independent so Theorem 6.5 can be applied to every point on the intersection. By Corollary 6.6 it is a differentiable curve. Let us consider the point $p = (1,1,0)/\sqrt{2}$ and determine its tangent line. We have

$$M = \begin{bmatrix} 2x & 2y & 2z \\ 1 & -1 & 2 \end{bmatrix} \ .$$

At the point $(1,1,0)/\sqrt{2}$ it is equal to

$$\begin{bmatrix} \sqrt{2} & \sqrt{2} & 0 \\ 1 & -1 & 2 \end{bmatrix} \ ,$$

whose 2×2-submatrices are non-singular. Therefore, near p the curve can be parametrized by either one of the x, y or z variables. Assuming that it is parametrized by x, we could write $\gamma(x)$ simply by $(y(x), z(x))$. We differentiate both $f = 0, g = 0$ to get

$$\begin{cases} 2x + 2yy_x + 2zz_x & = 0 \\ 1 - y_x + 2z_x & = 0 \ . \end{cases}$$

Plugging in $(x, y, z) = (1, 1, 0)/\sqrt{2}$, we solve this system to find that $y_x = -1, z_x = -1$ at $(1, 1, 0)/\sqrt{2}$, so the tangent vector of the curve at p is $(1, -1, -1)$. The tangent line of the curve at p is given by the parametric straight line

$$\frac{(1,1,0)}{\sqrt{2}} + (1,-1,-1)t\ , \quad t \in \mathbb{R}\ .$$

Alternatively, we could use y as the parameter. Then by differentiating the equations in y, we get

$$\begin{cases} 2xx_y + 2y + 2zz_y & = 0 \\ x_y - 1 + 2z_y & = 0\ , \end{cases}$$

which is solved to get $(-1, 1, 1)$. The tangent line is now described as

$$\frac{(1,1,0)}{\sqrt{2}} + (-1,1,1)t\ , \quad t \in \mathbb{R}\ .$$

Despite there is a change in the direction in the parametric form, it is the same geometric curve as the first one.

Example 6.10. Consider

$$\begin{cases} x^2 + y^2 + (z-1)^2 & = 1 \\ x^2 + y^2 + (z-2)^2 & = 4\ . \end{cases}$$

Geometrically it is clear that the solution set consists of a single point $\{(0, 0, 0)\}$. The gradients of these functions are $(2x, 2y, 2(z - 1))$ and $(2x, 2y, 2(z - 2))$ which are linearly dependent at $(0, 0, 0)$, and

$$M = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 0 & -4 \end{bmatrix}\ .$$

It shows that the rank 2 condition in Theorem 6.5 cannot be removed. On the other hand, consider

$$\begin{cases} x^2 - y & = 0 \\ 4x^2 - y & = 0\ , \end{cases}$$

as a system in $\mathbb{R}^3$. The zero set of the first equation is given by (x, x^2, z) and the second $(x, 4x^2, z)$. Hence the zero set of this system is given by $(0, 0, z), z \in \mathbb{R}$ which is the z-axis. However,

$$M = \begin{bmatrix} 0 & -1 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

is of rank 1 only. It shows that the rank 2 condition is sufficient but not necessary.

6.4 The Implicit Function Theorem

We are ready to generalize the results in the previous two sections. Let $f_1, f_2, \cdots, f_m, 1 \leq m \leq n-1$, be differentiable functions in some open set in $\mathbb{R}^n$. Suppose p is a point in the set

$$Z = \{x \in G : \ f_j(x) = 0, \ j = 1, 2, \cdots, m\} \subset \mathbb{R}^n \ .$$

We would like to investigate if the zero set is an "m-surface" passing p. When $m = 1$ and $n \geq 4$, it is a hypersurface. When $m = n-2$ and $n \geq 3$, it is a surface. When $m = n-1$, it is a curve. Let

$$M = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots\cdots & \dfrac{\partial f_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots\cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix} .$$

The implicit function theorem gives a criterion to this question.

Theorem 6.7 (Implicit Function Theorem). *Let $f_1, \cdots, f_m$ be differentiable functions in some open G of $\mathbb{R}^n$ where $1 \leq m \leq n-1$. Let Z be defined as above and $p = (z_1, \cdots, z_n) \in Z$. Assume that $\nabla f_j(p)$ are linearly independent, that is, M is of rank m. Without loss of generality let the last $m \times m$-submatrix of M be non-singular. There exist a cube C centered at p inside G and m many differentiable functions $\varphi_1, \cdots, \varphi_m$ from C' to C'' satisfying $\varphi_1(z') = z_{n-m+1}, \cdots, \varphi_m(z') = z_n$ such that the set*

$$\{(x', \varphi_1(x'), \cdots, \varphi_m(x')) : \ x' \in C'\}$$

coincides with $\{x \in C : \ f_j(x) = 0, \ j = 1, \cdots, m\}$, where $x = (x', x'')$ and

$$C = C' \times C'' \subset \mathbb{R}^{n-m} \times \mathbb{R}^m \ .$$

Example 6.11. Consider

$$\begin{cases} x^2 - \cos y + w^2 & = 0 \\ x + xw + y^2 - \log w + z & = 2 \ . \end{cases}$$

Does it define a surface in $\mathbb{R}^4$ near $(0, 0, 2, 1)$? The matrix M is given by

$$M = \begin{bmatrix} 2x & \sin y & 0 & 2w \\ 1+w & 2y & 1 & x - w^{-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 2 \\ 2 & 0 & 1 & -1 \end{bmatrix} ,$$

at $(0, 0, 2, 1)$. According to the Implicit Function Theorem, there are differentiable functions $z = \varphi(x, y), w = \psi(x, y)$ defined near $(0, 0, 1, 2)$ so that the surface locally coincides with the graph $\{(x, y, \varphi(x, y), \psi(x, y))\}$.

We conclude by reformulating the Implicit Function Theorem in a form that is usually stated.

Let F be a differentiable from some open $G \subset \mathbb{R}^n$ to $\mathbb{R}^m$. For $F = (f_1, f_2, \cdots, f_m)$, its **Jacobian matrix** ∇F is given by

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots\cdots & \dfrac{\partial f_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots\cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix} .$$

Now the Implicit Function Theorem becomes

Theorem 6.7' (Implicit Function Theorem) *Let F be a differentiable function in some open $G \subset \mathbb{R}^n$ to $\mathbb{R}^m, 1 \le m \le n-1$. Let $Z = F^{-1}(c)$ for some $c \in F(G)$ and $p \in Z$. Assume the Jacobian matrix of F at p has rank m. Without loss of generality let the last $m \times m$-submatrix of $\nabla F(p)$ be non-singular. There exist a cube C centered at p inside G and a differentiable function Φ from C' to C'' such that the set*

$$\{(x', \Phi(x')) : \ x' \in C'\}$$

coincides with $Z \cap C$.

6.5 Local Change of Coordinates*

In Section 5.3 changes of variables were applied to transform partial differentiable functions. Functions in the old variables are converted to functions in the new variables. One may wonder whether this procedure could be reversed so that functions in the new variables can be converted back to functions in the old variables. When this happens, we may call such change of variables a change of coordinates. To answer this question we need some preparation.

First of all, the most general Chain Rule.

Theorem 6.8 (**Chain Rule III**). *Let $G \subset \mathbb{R}^n$ and $D \subset \mathbb{R}^m$ be open, $F : G \to \mathbb{R}^m$ and $H : D \to \mathbb{R}^p$ be differentiable and $F(G) \subset D$. Then $H \circ F$ is differentiable in G and*

$$\nabla(H \circ F)(x) = \nabla H(F(x))\nabla F(x) , \quad \forall x \in G .$$

Proof. Let

$$F = (f_1, f_2, \cdots, f_m), \quad H = (h_1, h_2, \cdots, h_p) .$$

For each $i = 1, \cdots, p$, apply the Chain Rule II to $h_i(F(x)) = h_j(f_1(x), f_2(c), \cdots, f_m(x))$ to get

$$\frac{\partial h_i \circ F}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial h_i}{\partial y_k}(F(x)) \frac{\partial f_k}{\partial x_j}(x) \ , \quad j = 1, \cdots, n \ ,$$

which is the desired result after putting in matrix form. □

Next, we give a necessary condition for a differentiable map to have an inverse.

Proposition 6.9. *Let $F : G \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable. Suppose it admits a differentiable inverse $H : F(G) \subset \mathbb{R}^m \to G$. Then $m = n$ and*

$$I = \nabla H(F(x)) \nabla F(x) \ , \quad x \in G \ . \tag{6.1}$$

Proof. Formula (6.1) follows from differentiating the relation $x = H(F(x))$ by the Chain Rule. It remains to show $n = m$. Assuming $n \geq m$, the matrix ∇F maps $\mathbb{R}^n$ to $\mathbb{R}^m$, so the subspace formed by its image has dimension $k \leq m$. As a result, the image of the product matrix $\nabla H \nabla F$ has dimension $j \leq k \leq m$. However, when F is invertible, (6.1) shows that $\nabla H \nabla F$ is the identity whose image has dimension n. From $n = j \leq k \leq m$ we conclude that $n = m$. When $n \leq m$, we just have to exchange the role of F and H to arrive at the same conclusion. □

(6.1) implies that the Jacobian matrix of an invertible map must be non-singular. This gives a necessary condition for invertibility. The inverse function theorem shows that this condition is also sufficient in a local sense.

Theorem 6.10 (Inverse Function Theorem). *Let F be a differentiable function in some open $G \subset \mathbb{R}^n$ to $\mathbb{R}^n$ satisfying $F(p) = q$. Suppose that its Jacobian matrix is nonsingular at p. There is some ball B containing p in which F is bijective. Furthermore, the image $F(B)$ is open and the inverse of F is a differentiable function from $F(B)$ to B.*

The inverse of F exists in a conceivably much smaller set B other than G, and this is why we call it admits a local inverse. Whether the inverse exists in the whole G must be analysed case by case. The Inverse Function Theorem can be regarded as the $m = n$ case of the Implicit Function Theorem.

In general, let G and D be two open sets in $\mathbb{R}^n$ and F is a differentiable map from G to D with a differentiable inverse. For any function f defined in D, the function $\tilde{f}(x) = f(F(x))$ becomes a function in G. Conversely, a function g defined in D there associates a function $\tilde{g}(y) = g(F^{-1}(y))$ in D. We call this a change of coordinates. In the following we examine the polar coordinates. The conversion between the rectangular coordinates and polar coordinates is one of the most common changes of coordinates.

Let $F : [0, \infty) \times \mathbb{R} \to \mathbb{R}^2$ be given by

$$F(r, \theta) = (r\cos\theta, r\sin\theta) \ .$$

The Jacobian matrix is

$$\nabla F = \begin{bmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{bmatrix} \ ,$$

and its Jacobian determinant is equal to r. By Inverse Function Theorem, F has a local inverse at every point except $(0, \theta), \theta \in \mathbb{R}$. Indeed, for any point (x, y), $F^{-1}(x, y)$ is not a single point but the infinite set

$$\left\{ (r, \theta + 2n\pi) : \ r = \sqrt{x^2 + y^2} \ , \ x = r\cos\theta, y = r\sin\theta, \theta \in [0, 2\pi), n \in \mathbb{Z} \ \right\},$$

where $\arctan y/x$ is the branch of the inverse tangent in $[0, 2\pi]$. Thus F is never a global inverse map. To obtain an inverse one way is to restrict the domain of F. For instance, F becomes a bijective map on $(0, \infty) \times [0, 2\pi)$ and now its image is $\mathbb{R}^2$ minus the origin. Or one can take $(0, \infty) \times [n\pi, (n + 2)\pi), n \in \mathbb{Z}$, instead of $(0, \infty) \times [0, 2\pi)$. One uneasy thing is that the domain is not an open set. To make it open one further restricts the map to $(0, \infty) \times (0, 2\pi)$ but then the image is $\mathbb{R}^2$ minus the semi-infinite x-axis $(x, 0), x \geq 0$. Life is not perfect. Luckily local changes of coordinates are sufficient for many purposes.

Comments on Chapter 6

The Implicit Function Theorem is used to verify whether a zero set is a curve, a surface, a hypersurface or an m-surface locally. It is a cornerstone of differential geometry and analysis. Its proof is harder and will be postponed to MATH3060. After going through this chapter, you are supposed to know:

- The concept of a surface and a curve in $\mathbb{R}^3$ with those examples illustrating the necessity of various assumptions in Theorems 6.2 and 6.5.
- Finding the tangent planes of surfaces and tangent lines of curves.
- The technique of implicit differentiation.

This theorem will be applied to constrained extremal problems in the next chapter.

Supplementary Readings

6.1 4.5 and 4.6 in [Au]. Chapters 16 and 17 in [Fitz].