INTRODUCTION TO MANIFOLDS

TSOGTGEREL GANTUMUR

Abstract. We discuss manifolds embedded in $\mathbb{R}^n$ and the implicit function theorem.

1. The implicit function theorem

In this section, we want to investigate if the equation $g(x,y)=0$ can be solved as $y=y(x)$. Our approach will be based on differentiability, meaning that we fix some point (x_*,y_*), and approximate g as

$$g(x,y) \approx g(x_*,y_*) + \partial_x g(x_*,y_*)(x-x_*) + \partial_y g(x_*,y_*)(y-y_*), \tag{1}$$

for $y \approx y_*$ and $x \approx x_*$. If $\partial_y g(x_*,y_*) \neq 0$, this approximate equation can be solved for y:

$$y - y_* \approx \frac{g(x,y) - g(x_*,y_*) - \partial_x g(x_*,y_*)(x-x_*)}{\partial_y g(x_*,y_*)}. \tag{2}$$

Solving this equation for $g(x,y) \neq g(x_*,y_*)$ would not yield a good approximation, because then $x \approx x_*$ would *not* imply $y \approx y_*$. Thus we put $g(x,y) = g(x_*,y_*) = 0$, and get

$$y - y_* \approx -\frac{\partial_x g(x_*,y_*)}{\partial_y g(x_*,y_*)}(x-x_*). \tag{3}$$

This suggests that the conditions $g(x_*,y_*)=0$ and $\partial_y g(x_*,y_*) \neq 0$ might be sufficient to solve $g(x,y)=0$ for a function $y=y(x)$, at least when x is in a small interval containing x_*. In the following remark, we will justify this expectation in full detail.

Remark 1.1. Let $Q_a = (-a,a)^2 \subset \mathbb{R}^2$ be an open square, with $a>0$, and let $g : Q_a \to \mathbb{R}$ be a continuously differentiable function, satisfying $g(0,0)=0$ and $\partial_y g(0,0) \neq 0$. We want to find a function $y = h(x)$, defined for $x \in (-\delta,\delta)$ with some $\delta > 0$, such that $g(x,h(x))=0$ for all $x \in (-\delta,\delta)$. Note that the point (x_*,y_*) from the previous discussion is now the origin. This is no loss of generality, since we may think of $g(x,y)$ as $\tilde{g}(x_*+x, y_*+y)$ for some function $\tilde{g}$. To proceed further, we introduce the auxiliary map $f : Q_a \to \mathbb{R}^2$, given by $f(x,y) = (x, g(x,y))$ for $(x,y) \in Q_a$. The motivation for considering such a map is that if we can solve $f(x,y) = (\alpha, 0)$ for (x,y) depending on α, then we would have $x=\alpha$ and $g(\alpha, y(\alpha)) = 0$. In order to invert f near the origin, we shall invoke the inverse function theorem. The Jacobian of f is

$$J(x,y) = \begin{pmatrix} 1 & 0 \\ \partial_x g(x,y) & \partial_y g(x,y) \end{pmatrix}, \tag{4}$$

and since g is continuously differentiable, J is continuous in Q_a, and hence we conclude that f is continuously differentiable in Q_a with $Df = J$. At the origin, Df is invertible, and

$$(Df)^{-1} = \begin{pmatrix} 1 & 0 \\ -\partial_x g/\partial_y g & 1/\partial_y g \end{pmatrix}, \tag{5}$$

where all functions are evaluated at the origin $0 \in \mathbb{R}^2$. Now the inverse function theorem guarantees that there exist of $r>0$ and $f^{-1} : f(Q_r) \to \mathbb{R}^2$, satisfying $f^{-1}(f(x,y)) = (x,y)$ for all $(x,y) \in Q_r$. Note that $f^{-1}(0,0) = (0,0)$. Moreover, $Df(x,y)$ is nonsingular for each

Date: 2016/10/29 22:55 (EDT). Incomplete draft.

$(x, y) \in Q_r$, and f^{-1} is continuously differentiable with $Df^{-1} \circ f = (Df)^{-1}$ in Q_r. If we let $f^{-1}(\alpha, \beta) = (x(\alpha, \beta), y(\alpha, \beta))$, then from $f(f^{-1}(\alpha, \beta)) = (\alpha, \beta)$, we infer that $x(\alpha, \beta) = \alpha$ and $g(\alpha, y(\alpha, \beta)) = \beta$ for $(\alpha, \beta) \in f(Q_r)$. In addition to what we have already mentioned, the inverse function theorem tells us that there is $\delta > 0$ such that $Q_\delta \in f(Q_r)$, implying that we have $g(\alpha, y(\alpha, \beta)) = \beta$ for all $(\alpha, \beta) \in Q_\delta$. In particular, setting $h(\alpha) = y(\alpha, 0)$, we get $g(\alpha, h(\alpha)) = 0$ for all $\alpha \in (-\delta, \delta)$. From $f^{-1}(0, 0) = (0, 0)$, we get $h(0) = 0$.

The function h we found in the preceding paragraph in fact solves our problem, but our assumptions are strong enough to yield additional results. As a component of f^{-1}, the function $y = y(\alpha, \beta)$ is *continuously differentiable* in Q_δ, and we have

$$Df^{-1} = \begin{pmatrix} 1 & 0 \\ \partial_\alpha y & \partial_\beta y \end{pmatrix}. \tag{6}$$

Comparing this with (5), we get $\partial_\alpha y \circ f = -\partial_x g / \partial_y g$ and $\partial_\beta y \circ f = 1/\partial_y g$. In particular, taking into account that $h'(\alpha) = \partial_\alpha y(\alpha, 0)$, we conclude that

$$h'(x) = -\frac{\partial_x g(x, h(x))}{\partial_y g(x, h(x))}, \qquad \text{for} \quad x \in (-\delta, \delta). \tag{7}$$

Before closing this remark, we make one crucial observation. Fix $x \in (-\delta, \delta)$, and consider $I = \{(x, y) : y \in (-r, r)\}$. The map f sends I to $f(I) = \{(x, g(x, y)) : y \in (-r, r)\} \subset f(Q_r)$. Since f is invertible in Q_r, the only point $(x, y) \in I$ with $g(x, y) = 0$ is $(x, h(x))$. In other words, apart from the curve $\{(x, h(x)) : x \in (-\delta, \delta)\}$, there are *no other* points (x, y) exist in the rectangle $(-\delta, \delta) \times (-r, r)$ satisfying $g(x, y) = 0$.

The preceding remark is the *implicit function theorem* in two dimensions.

Example 1.2. (a) Let us apply the implicit function theorem to the equation $x^2 + y^2 = 1$. Thus we set $g(x, y) = x^2 + y^2 - 1$, and compute $\partial_y g(x, y) = 2y$. This means that as long as (x_*, y_*) satisfies $g(x_*, y_*) = 0$ and $y_* \neq 0$, we can apply the result at the point (x_*, y_*), and infer the existence of $\delta > 0$ and $h : (x_* - \delta, x_* + \delta) \to \mathbb{R}$ such that $g(x, h(x)) = 0$ for all $x \in (x_* - \delta, x_* + \delta)$. We can also compute the derivative of h as

$$h'(x) = -\frac{\partial_x g(x, y)}{\partial_y g(x, y)} = -\frac{2x}{2y} = -\frac{x}{h(x)}, \qquad \text{for} \quad x \in (x_* - \delta, x_* + \delta). \tag{8}$$

The intuitive reason why the case $y_* = 0$ must be excluded is the fact that then the derivative $h'(x_*)$ would have to become infinity.

(b) Let $g(x, y) = y^3 - x$, and let us try to solve $g(x, y) = 0$ for $y = y(x)$ near $(x, y) = (0, 0)$. We have $g(0, 0) = 0$, but $\partial_y g(0, 0) = (3y^2)|_{y=0} = 0$. Therefore the implicit function theorem *cannot* be applied, even though we can explicitly solve the equation as $y(x) = \sqrt[3]{x}$. This has of course to do with the fact that $\sqrt[3]{x}$ is *not* differentiable at $x = 0$.

(c) Let $g(x, y) = x^2 - y^2$, and let us try to solve $g(x, y) = 0$ for $y = y(x)$ near $(x, y) = (0, 0)$. We have $g(0, 0) = 0$, but $\partial_y g(0, 0) = (-2y)|_{y=0} = 0$, and hence the implicit function theorem *cannot* be applied. A close inspection reveals that the solution of $g(x, y) = 0$ is $y = \pm x$, which *cannot* be written as a function $y = y(x)$ near $(x, y) = (0, 0)$.

Let $\Omega \subset \mathbb{R}^n$ and $\Sigma \subset \mathbb{R}^m$ be open sets. Then their *product* $\Omega \times \Sigma \subset \mathbb{R}^{n+m}$ is given by

$$\Omega \times \Sigma = \{(x, y) : x \in \Omega,\ y \in \Sigma\}, \tag{9}$$

where $(x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{R}^{n+m}$. Let $g : \Omega \times \Sigma \to \mathbb{R}^m$ be a differentiable function. The value of g at $(x, y) \in \Omega \times \Sigma$ is denoted by $g(x, y) \in \mathbb{R}^m$. For any fixed $x \in \Omega$, the correspondence $y \mapsto g(x, y)$ is a function of $y \in \Sigma$, and its derivative will be denoted by $D_y g$ Similarly, we can introduce $D_x g$. In the following, sometimes it will be convenient to specify the dimension of a cube in the notation, as in $Q_r^n(a) = (a - r, a + r)^n \subset \mathbb{R}^n$.

Theorem 1.3. *Let $\Omega \subset \mathbb{R}^n$ and $\Sigma \subset \mathbb{R}^m$ be open sets, and let $g : \Omega \times \Sigma \to \mathbb{R}^m$ be continuously differentiable. Suppose that $(a, b) \in \Omega \times \Sigma$ satisfies $g(a, b) = 0$, and that $D_y g(a, b)$ is nonsingular. Then there exists $\delta > 0$ and $h : Q^n_\delta(a) \to \mathbb{R}^m$ with $h(a) = b$, such that $g(x, h(x)) = 0$ for all $x \in Q^n_\delta(a)$. Moreover, h is continuously differentiable in $Q^n_\delta(a)$, with*

$$Dh(x) = -(D_y g(x, h(x)))^{-1} D_x g(x, h(x)), \qquad x \in Q^n_\delta(a). \tag{10}$$

Proof. Let $f : \Omega\times\Sigma \to \mathbb{R}^{n+m}$ be defined by $f(x, y) = (x, g(x, y))$. This function is continuously differentiable, and

$$Df(x, y) = \begin{pmatrix} I & 0 \\ D_x g & D_y g \end{pmatrix}, \tag{11}$$

where $I \in \mathbb{R}^{n\times n}$ is the identity matrix. Since $\det Df(a, b) = \det D_y g(a, b) \neq 0$, the matrix $Df(a, b)$ is invertible. Consequently, the inverse function theorem guarantees that there exist of $r > 0$ and $f^{-1} : f(Q_r) \to \mathbb{R}^{n+m}$, satisfying $f^{-1}(f(x, y)) = (x, y)$ for all $(x, y) \in Q_r$, where $Q_r = Q^{n+m}_r(a, b)$. Note that $f^{-1}(a, 0) = (a, b)$. Moreover, $Df(x, y)$ is nonsingular for each $(x, y) \in Q_r$, and f^{-1} is continuously differentiable with $Df^{-1} \circ f = (Df)^{-1}$ in Q_r. If we let $f^{-1}(\alpha, \beta) = (x(\alpha, \beta), y(\alpha, \beta))$, then from $f(f^{-1}(\alpha, \beta)) = (\alpha, \beta)$, we infer that $x(\alpha, \beta) = \alpha$ and $g(\alpha, y(\alpha, \beta)) = \beta$ for $(\alpha, \beta) \in f(Q_r)$. In addition to what we have already mentioned, the inverse function theorem tells us that there is $\delta > 0$ such that $Q^{n+m}_\delta(a, 0) \in f(Q_r)$, implying that we have $g(\alpha, y(\alpha, \beta)) = \beta$ for all $(\alpha, \beta) \in Q^{n+m}_\delta(a, 0)$. In particular, setting $h(\alpha) = y(\alpha, 0)$, we get $g(\alpha, h(\alpha)) = 0$ for all $\alpha \in Q^n_\delta(a)$. From $f^{-1}(a, 0) = (a, b)$, we get $h(a) = b$.

As a collection of components of f^{-1}, the function $y = y(\alpha, \beta)$ is continuously differentiable in $Q^{n+m}_\delta(a, 0)$, and we have

$$Df^{-1} = \begin{pmatrix} I & 0 \\ D_\alpha y & D_\beta y \end{pmatrix}. \tag{12}$$

Comparing this with

$$(Df)^{-1} = \begin{pmatrix} I & 0 \\ -(D_y g)^{-1} D_x g & (D_y g)^{-1} \end{pmatrix}, \tag{13}$$

we infer

$$D_\alpha y \circ f = -(D_y g)^{-1} D_x g, \qquad D_\beta y \circ f = (D_y g)^{-1}. \tag{14}$$

In particular, taking into account that $Dh(\alpha) = D_\alpha y(\alpha, 0)$, we conclude that

$$Dh(x) = -(D_y g(x, h(x)))^{-1} D_x g(x, h(x)), \tag{15}$$

for all $x \in Q^n_\delta(a)$. □

Example 1.4. (a) Consider the equation

$$g(x, y, z) \equiv \sin(xy + z) + \log(yz^2) = 0. \tag{16}$$

The triple $p = (x, y, z) = (1, 1, -1)$ is a solution: $g(1, 1, -1) = 0$, and g is continuously differentiable in the open set $\{(x, y, z) : x \in \mathbb{R}, y > 0, z < 0\}$. Can we express z as a function of x and y near p? This is exactly the kind of question that could be answered by the implicit function theorem. We have

$$\partial_z g(x, y, z) = \cos(xy + z) + \frac{2z}{yz^2} = \cos(xy + z) + \frac{2}{yz}, \tag{17}$$

and hence

$$\partial_z g(1, 1, -1) = \cos 0 - 2 = -1 \neq 0. \tag{18}$$

Thus there exist $\delta > 0$ and a continuously differentiable function $h : Q^2_\delta(1, 1) \to \mathbb{R}$ such that $g(x, y, h(x, y)) = 0$ for all $(x, y) \in Q^2_\delta(1, 1)$.

(b) Can we solve

$$\begin{aligned} xu^2 + yzv + x^2 z &= 3, \\ yv^5 + zu^2 - xv &= 1, \end{aligned} \tag{19}$$

for (u, v) near $(1, 1)$ as a function of (x, y, z) near $(1, 1, 1)$? We can formulate the problem as solving $g(\alpha, \beta) = 0$ for $\beta = \beta(\alpha)$, where $\alpha = (x, y, z)$, $\beta = (u, v)$, and

$$g(\alpha, \beta) = g(x, y, z, u, v) = \begin{pmatrix} xu^2 + yxv + x^2 z - 3 \\ yv^5 + 2zu - v^2 - 2 \end{pmatrix}. \tag{20}$$

Obviously, g is continuously differentiable in $\mathbb{R}^5$, and $g(1, 1, 1, 1, 1) = 0$. We can compute the relevant derivative as

$$D_\beta g(\alpha, \beta) = \begin{pmatrix} 2xu & yz \\ 2zu & 5yv^4 - x \end{pmatrix}. \tag{21}$$

so that the matrix

$$D_\beta g(1, 1, 1, 1, 1) = \begin{pmatrix} 2 & 1 \\ 2 & 4 \end{pmatrix}, \tag{22}$$

is invertible. Thus there exist $\delta > 0$ and $h : Q^3_\delta(1, 1, 1) \to \mathbb{R}^2$ continuously differentiable, such that $g(\alpha, h(\alpha)) = 0$ for all $\alpha \in Q^3_\delta(1, 1, 1)$.