Proof of the Implicit Function Theorem: $\boldsymbol{f}\colon \mathbb{R}^2 \mapsto \mathbb{R}$.

In this lecture we explore how to prove the following result:

> Let $f\colon \mathbb{R}^2 \mapsto \mathbb{R}$ be a C^k, $k \geqslant 1$, function such that $f(x_0, y_0) = 0$. Further assume that $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. Then there is $\delta > 0$ and a function $Y\colon (x_0 - \delta, x_0 + \delta)$ such that
>
> i. $f(x, Y(x)) = 0$ for all $x \in (x_0 - \delta, x_0 + \delta)$;
>
> ii. $Y(x)$ is C^k.

We first note that

- No generality is lost by assuming $x_0 = y_0 = 0$;
- No generality is lost by further assuming $\frac{\partial f}{\partial y}(0, 0) = 1$.

Thus in the following we will proceed under these assumptions.

Proof of the result when $\boldsymbol{k} \geqslant 2$.

We first present a simple proof under the stronger assumption that f is C^k for some $k \geqslant 2$.

- Define
$$G(x, y) = -\left(\frac{\partial f}{\partial y}\right)^{-1} \left(\frac{\partial f}{\partial x}\right). \tag{1}$$
As $\frac{\partial f}{\partial y}(0, 0) = 1$ and $f \in C^2$, there is $\delta_1 > 0$ such that
$$\frac{\partial f}{\partial y}(x, y) \in \left(\frac{1}{2}, \frac{3}{2}\right), \qquad (x, y) \in I_1 := (-\delta_1, \delta_1) \times (-\delta_1, \delta_1). \tag{2}$$
- As a consequence, we have $G(x, y) \in C^1$ on the same interval I_1. In particular, $G(x, y)$ is Lipschitz with respect to the variable y.
- Now consider the first order differential equation
$$y' = G(x, y), \qquad y(0) = 0. \tag{3}$$
By the existence/uniqueness theorem of ODEs, there is a unique solution $Y(x)$ defined on $(-\delta_1, \delta_1)$.
- We have
$$\frac{\mathrm{d}}{\mathrm{d}x} f(x, Y(x)) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} Y' = 0 \tag{4}$$
for all $x \in (-\delta_1, \delta_1)$. Together with $f(0, Y(0)) = 0$ this shows $f(x, Y(x)) = 0$ for all $x \in (-\delta_1, \delta_1)$.
- By the theory of ODEs, we know that $Y \in C^1$. Together with $G(x, y) \in C^1$ we see that $Y' = G(x, Y(x)) \in C^1$, consequently $Y \in C^2$.
- Now it is easy to show that $Y \in C^k$, if $k > 2$. For example, when $k = 3$ we have $G(x, y) \in C^2$ and
$$Y'' = \frac{\partial G}{\partial x} + \frac{\partial G}{\partial y} Y' \in C^1 \Longrightarrow Y \in C^3. \tag{5}$$

Exercise 1. Prove for general k by induction.

Question 1. *Would this approach work if $k = 1$? In this case we have $G(x, y)$ merely continuous. Theory of ODEs only gives existence of Y but not uniqueness. Is this*

Proof of the result when $\boldsymbol{k} = 1$.

Now we assume that f is only C^1.

- The plan is as follows. We try to find $\delta_1, \delta_2 > 0$ such that
$$f(x, \delta_2) > 0, \qquad f(x, \delta_2) < 0 \qquad \text{for all } x \in (-\delta_1, \delta_1). \tag{6}$$

By the intermediate value theorem for the function $f(x,\cdot)$ we have the existence of $y \in (-\delta_2, \delta_2)$ such that $f(x,y)=0$. The proof ends after we further show the uniqueness of such y and that if we define $Y=y$, the function Y is C^1.

- Proof of (6). Due to the continuity of $\frac{\partial f}{\partial y}$, there is $\delta_2 > 0$ such that

$$\frac{\partial f}{\partial y}(x,y) \in \left(\frac{1}{2}, \frac{3}{2}\right), \qquad -\delta_2 < x, y < \delta_2. \tag{7}$$

Now denote

$$M := \sup_{-\delta_2 < x, y < \delta_2} \left|\frac{\partial f}{\partial x}(x,y)\right|. \tag{8}$$

We set

$$\delta_1 := \frac{\delta_2}{4M}. \tag{9}$$

Then for any $x_0 \in (-\delta_1, \delta_1)$, we have

$$\begin{aligned} f(x_0, \delta_2) &= f(x_0,\delta_2) - f(x_0,0) + f(x_0,0) - f(0,0) \\ &= \int_0^{\delta_2} \frac{\partial f}{\partial y}(x_0,y)\,\mathrm{d}y + \int_0^{x_0} \frac{\partial f}{\partial x}(x,0)\,\mathrm{d}x \\ &\geqslant \int_0^{\delta_2} \frac{\partial f}{\partial y}(x_0,y)\,\mathrm{d}y - \int_0^{x_0} \left|\frac{\partial f}{\partial x}(x,0)\right|\mathrm{d}x \\ &> \int_0^{\delta_2} \frac{\mathrm{d}y}{2} - \int_0^{\delta_1} M\,\mathrm{d}x \\ &= \frac{\delta_2}{2} - \frac{\delta_2}{4M} M = \frac{\delta_2}{4} > 0. \end{aligned} \tag{10}$$

Similarly we can prove

$$f(x_0, -\delta_1) < 0. \tag{11}$$

By the intermediate value theorem we see that for every $x \in (-\delta_1, \delta_1)$ there is at least one $y \in (-\delta_2, \delta_2)$ such that $f(x,y)=0$.

- Uniqueness of the intermediate value. Assume for some $x_0 \in (-\delta_1, \delta_1)$ there are $-\delta_2 < y_{01} < y_{02} < \delta_2$ such that $f(x_0, y_{01}) = f(x_0, y_{02}) = 0$, then by the mean value theorem there is $y_0 \in [y_{01}, y_{02}]$ such that

$$\frac{\partial f}{\partial y}(x_0, y_0) = 0. \tag{12}$$

Contradiction.

- $Y \in C^1$. We have $f(x, Y(x)) = 0$ for all $x \in (-\delta_1, \delta_1)$. Let $x_0 \in (-\delta_1, \delta_1)$ be arbitrary. Let $x \in (-\delta_1, \delta_1)$ be arbitrary. Define

$$g(t) := f((1-t)\,x_0 + t\,x,\ (1-t)\,Y(x_0) + t\,Y(x)). \tag{13}$$

By the mean value theorem we have

$$\begin{aligned} 0 &= f(x, Y(x)) - f(x_0, Y(x_0)) \\ &= g(1) - g(0) \\ &= g'(c) \\ &= \frac{\partial f}{\partial x}(x_c, y_c)\,(x - x_0) + \frac{\partial f}{\partial y}(x_c, y_c)\,(Y(x) - Y(x_0)) \end{aligned} \tag{14}$$

where $x_c := (1-c)\,x_0 + c\,x$, $y_c := (1-c)\,Y(x_0) + c\,Y(x)$. This gives

$$\frac{Y(x) - Y(x_0)}{x - x_0} = -\left[\frac{\partial f}{\partial y}(x_c, y_c)\right]^{-1} \frac{\partial f}{\partial x}(x_c, y_c). \tag{15}$$

As $f \in C^1$, and $c \in [0,1]$, we have

$$\lim_{x \longrightarrow x_0} \left[\frac{\partial f}{\partial y}(x_c, y_c)\right]^{-1} \frac{\partial f}{\partial x}(x_c, y_c) = \left[\frac{\partial f}{\partial y}(x_0, Y(x_0))\right]^{-1} \frac{\partial f}{\partial x}(x_0, Y(x_0)). \tag{16}$$

As a consequence Y is differentiable and satisfies

$$Y'(x) = \left[\frac{\partial f}{\partial y}(x, Y(x))\right]^{-1} \frac{\partial f}{\partial x}(x, Y(x)), \qquad x \in (-\delta_1, \delta_1). \tag{17}$$

By the differentiability of Y we have its continuity, together with $f \in C^1$ we conclude that the right hand side of (17) is continuous. Therefore Y' is continuous and $Y \in C^1$.

Exercise 2. Obtain a similar proof for the implicit function theorem for the equation $f(x_1, x_2, ..., x_M, y) = 0$.