Appendix B

Implicit function theorem (a quantitative version)

In this appendix we recall the implicit function Theorem. We provide an explicit proof because we use in the text a quantitative version of the theorem so it is important to keep track of the various constants.

B.1 The theorem

Let $n, m \in \mathbb{N}$ and $F \in \mathcal{C}^1(\mathbb{R}^{m+n}, \mathbb{R}^m)$ and let $(x_0, \lambda_0) \in \mathbb{R}^n \times \mathbb{R}^m$ such that $F(x_0, \lambda_0) = 0$. For each $\delta > 0$ let $V_\delta = \{(x, \lambda) \in \mathbb{R}^{n+m} \ : \ \|x - x_0\| \leq \delta, \|\lambda - \lambda_0\| \leq \delta\}$.

Theorem B.1.1 *Assume that $\partial_x F(x_0, \lambda_0)$ is invertible and choose $\delta > 0$ such that $\sup_{(x,\lambda)\in V_\delta} \|\mathbb{1} - [\partial_x F(x_0, \lambda_0)]^{-1} \partial_x F(x, \lambda)\| \leq \frac{1}{2}\}$. Let $B_\delta = \sup_{(x,\lambda)\in V_\delta} \|\partial_\lambda F(x, \lambda)\|$ and $M = \|\partial_x F(x_0, \lambda_0)^{-1}\|$. Set $\delta_1 = (2MB_\delta)^{-1}\delta$ and $\Lambda_{\delta_1} := \{\lambda \in \mathbb{R}^m \ : \ \|\lambda - \lambda\| < \delta_1\}$. Then there exists $g \in \mathcal{C}^1(\Lambda_{\delta_1}, \mathbb{R}^m)$ such that all the solutions of the equation $F(x, \lambda) = 0$ in the set $\{(x, \lambda) \in \mathcal{B}_1 \times \mathcal{B}_2 \ : \ \|\lambda - \lambda_0\| < \delta_1, \|x - x_0\| < \delta\}$ are given by $(g(\lambda), \lambda)$. In addition,*

$$\partial_\lambda g(\lambda) = -(\partial_x F(g(\lambda), \lambda))^{-1} \partial_\lambda F(g(\lambda), \lambda).$$

We will do the proof in several steps.

B.1.1 Existence of the solution

Let $A(x, \lambda) = \partial_x F(x, \lambda)$, $M = \|A(x_0, \lambda_0)^{-1}\|$.

We want to solve the equation $F(x, \lambda) = 0$, various approaches are possible. Here we will use a simplification of Newton method, made possible by the fact that we already know a good approximation of the zero we are looking for. Let λ be such that $\|\lambda - \lambda_0\| < \delta_1 \leq \delta$. Consider $U_\delta = \{x \in \mathbb{R}^n \ : \ \|x - x_0\| \leq \delta\}$

and the function $\Theta_\lambda : U_\delta \to \mathbb{R}^n$ defined by[1]

$$\Theta_\lambda(x) = x - A(x_0, \lambda_0)^{-1} F(x, \lambda). \tag{B.1.1}$$

Problem B.1 *Prove that, for $x \in U(\lambda)$, $F(x, \lambda) = 0$ is equivalent to $x = \Theta_\lambda(x)$.*

Next,

$$\|\Theta_\lambda(x_0) - \Theta_{\lambda_0}(x_0)\| \leq M\|F(x_0, \lambda)\| \leq MB_\delta \delta_1.$$

In addition, $\|\partial_x \Theta_\lambda\| = \|\mathbb{1} - A(x_0, \lambda_0)^{-1} A(x, \lambda)\| \leq \frac{1}{2}$. Thus,

$$\|\Theta_\lambda(x) - x_0\| \leq \frac{1}{2}\|x - x_0\| + \|\Theta_\lambda(x_0) - x_0\| \leq \frac{1}{2}\|x - x_0\| + MB_\delta \delta_1 \leq \delta.$$

The existence of $x \in U_\delta$ such that $\Theta_\lambda(x) = x$ follows then by the standard fixed point Theorem A.1.1. We have so obtained a function $g : \{\lambda \ : \ \|\lambda - \lambda_0\| \leq \delta_1\} = \Lambda_{\delta_1} \to \mathbb{R}^n$ such that $F(g(\lambda), \lambda) = 0$. it remains the question of the regularity.

B.1.2 Lipschitz continuity and Differentiability

Let $\lambda, \lambda' \in \Lambda_{\delta_1}$. By (B.1.1)

$$\|g(\lambda) - g(\lambda')\| \leq \frac{1}{2}\|g(\lambda) - g(\lambda')\| + MB_\delta|\lambda - \lambda'|$$

This yields the Lipschitz continuity of the function g. To obtain the differentiability we note that, by the differentiability of F and the above Lipschitz continuity of g, for $h \in \mathbb{R}^m$ small enough,

$$\|F(g(\lambda + h), \lambda + h) - F(g(\lambda), \lambda) + \partial_x F[g(\lambda + h) - g(\lambda)] + \partial_\lambda F h\| = o(\|h\|).$$

Since $F(g(\lambda + h), \lambda + h) = F(g(\lambda), \lambda) = 0$, we have that

$$\lim_{h \to 0} \|h\|^{-1} \|g(\lambda + h) - g(\lambda) + [\partial_x F]^{-1} \partial_\lambda F h\| = 0$$

which concludes the proof of the Theorem, the continuity of the derivative being obvious be the obtained explicit formula.

[1] The Newton method would consist in finding a fixed point for the function $x - A(x, \lambda)^{-1} F(x, \lambda)$. This gives a much faster convergence and hence is preferable in applications, yet here it would make the estimates a bit more complicated.

B.2 Generalization

First of all note that the above theorem implies the inverse function theorem. Indeed if $f : \mathbb{R}^n \to \mathbb{R}^n$ is a function such that $\partial_x f$ is invertible at some point x_0, then one can consider the function $F(x, y) = f(x) - y$. Applying the implicit function theorem to the equation $F(x, y) = 0$ it follows that $y = f(x)$ are the only solution, hence the function is locally invertible.

The above theorem can be generalized in several ways.

Problem B.2 *Show that if F in Theorem B.1.1 is $\mathcal{C}^r$, then also g is $\mathcal{C}^r$.*

Problem B.3 *Verify that if $\mathcal{B}_1, \mathcal{B}_2$ are two Banach spaces and in Theorem B.1.1 we have $\mathcal{B}_1$ instead of $\mathbb{R}^n$ and $\mathcal{B}_2$ instead of $\mathbb{R}^m$ the Theorem remains true and the proof remains exactly the same.*

As I mentioned the statement of Theorem B.1.1 is suitable for quantitative applications.

Problem B.4 *Suppose that in Theorem B.1.1 we have $F \in \mathcal{C}^2$, then show that we can chose*

$$\delta = [2\|D\partial_x F\|_\infty]^{-1} .$$