1 Iterated Derivatives

In this section X and Y are Banach spaces over $\Bbbk$, $U = U(u) \subset X$ is an open neighbourhood of u. Suppose $f : U \to Y$ is differentiable on U. Then $f'(\cdot) : U \to \mathcal{L}(X, Y)$ and hence, $f'(u) \in \mathcal{L}(X, Y)$, i.e., $f'(u) : X \to Y$ and $f'(u)h \in Y$ for every element $h \in X$. If $f'(\cdot)$ is also differentiable on U, then $(f'(\cdot))' : U \to \mathcal{L}(X, \mathcal{L}(X, Y))$. In particuluar, $(f'(u))' \in \mathcal{L}(X, \mathcal{L}(X, Y))$, i.e., $(f'(u))' : X \to \mathcal{L}(X, Y)$ and $(f'(u))'h \in \mathcal{L}(X, Y)$ for each $h \in X$, i.e., $(f'(u))'h : X \to Y$ and $(f'(u))'hk \in Y$ for every element $k \in X$. This consideration shows the hierarchical order of structure of higher derivatives. This is in contrast to the previoiusly defined derivatives as multilinear operators. Clearly it is much easier to think of higher derivatives as multilinear operators than operators on operator spaces. In this section we show the equivalence of the two definitions. This equivalence will be shown here for second derivatives but the ideas can easily be carried over to higher derivatives.

Proposition 1 *Suppose $f : U \to Y$ is differentiable on U.*

1. *$f''(u)$ exists if and only if $(f'(u))'$ exists. In this case*

$$f''(u)\,hk = (f'(u))'\,hk \ \forall h, k \in X$$

2. *$f''(\cdot)$ is continuous at u if and only if $(f'(\cdot))'$ is continuous at u.*

Proof.

1. Suppose $(f'(u))'$ exists. Define the operator

$$B : X \times X \to Y$$

by

$$B(h, k) = (f'(u))'\,hk$$

Clearly B is bilinear and

$$\begin{aligned} \|B(h,k)\| &= \left\|(f'(u))'\,hk\right\| \le \left\|(f'(u))'\,h\right\| \|k\| \\ &\le \left\|(f'(u))'\right\| \|h\| \|k\|. \end{aligned}$$

Therefor, B is also bounded. Now

$$f'(u+h) - f'(u) = (f'(u))'\,h + r(u; h)$$

with $\|r(u;h)\| = o(h)$ as $h \to 0$. The above equation means that $r(u;h)$ is also a bounded linear operator from X to Y because the other terms of the equation are. Evaluating at $k \in X$ gives

$$\begin{aligned} f'(u+h)\,k - f'(u)\,k &= (f'(u))'\,hk + r(u;h)\,k \\ &= B(h,k) + r(u;h)\,k \end{aligned}$$

An since

$$\sup_{\|k\|\leq 1} \|r(u;h)k\| = \|r(u;h)\| = o(h)$$

as $h \to 0$, it follows that $f''(u)$ and

$$f''(u)hk = B(h,k) = (f'(u))' hk.$$

The other direction can be proved similarly.

2. This statement follows form the equality

$$\|f''(u) - f''(v)\| = \left\|(f'(u))' - (f'(v))'\right\|$$

which follows from the equality

$$\|f''(u)hk - f''(v)hk\| = \left\|(f'(u))' hk - (f'(v))' hk\right\|.$$

Indeed this last equality can be used to show the inequality of the operator norms both ways. For example

$$\begin{aligned} \left\|(f'(u))' hk - (f'(v))' hk\right\| &= \|f''(u)hk - f''(v)hk\| \\ &= \|(f''(u) - f''(v))hk\| \\ &\leq \|f''(u) - f''(v)\| \|h\| \|k\| \end{aligned}$$

gives

$$\left\|(f'(u))' h - (f'(v))' h\right\| \leq \|f''(u) - f''(v)\| \|h\|$$

which, in turn, gives

$$\left\|(f'(u))' - (f'(v))'\right\| \leq \|f''(u) - f''(v)\|.$$

■

Remark 2 *Notice the meaning of the notation in the above proposition. With full notation we may restate part 1 of the proposition as*

$$f''(u)(h,k) = (f'(u))'(h)(k).$$

Both sides represent elements of Y.

2 Chain Rules

We have proved before the chain rule theorem:

> If $f : U \subset X \to Y$ and $g : V \subset Y \to Z$ such that $f(U) \subset V$ and if f is differentiable at $u \in U$ and g is differentiable at $v = f(u)$, then $g \circ f$ is differentiable at u and $(g \circ f)'(u) = g' \circ f(u) f'(u)$

In this section we wish to extend this result and discuss the computation of higher derivatives for composite functions.

Proposition 3 *Suppose* $f : U \subset X \to Y$ *and* $g : V \subset Y \to Z$ *such that* $f(U) \subset V$*.If* f *is* C^m *on* U *and* g *is* C^m *on* V *then* $g \circ f$ *is* C^m *on* U*.*

Proof. The proof follows in the same way as that of the theorem qouted above, together with induction. ■

We now wish to justify the formula

$$(g \circ f)''(u)\,hk = g'' \circ f(u)\,f'(u)\,hf'(uh)\,k + g' \circ f(u)\,f''(u)\,hk$$

For this we use the usual technique

$$\begin{aligned}
(g \circ f)''(u)\,hk &= \frac{d}{dt}\,(g \circ f)'(u+th)\,k\big|_{t=0} \\
&= \frac{d}{dt}\,g' \circ f(u+th)\,f'(u+th)\,k\big|_{t=0} \\
&= g'' \circ f(u+th)\,f'(u+th)\,hf'(u+th)\,k \\
&\quad + g' \circ f(u+th)\,f''(u+th)\,hk\big|_{t=0} \\
&= g'' \circ f(u)\,f'(u)\,hf'(uh)\,k + g' \circ f(u)\,f''(u)\,hk
\end{aligned}$$

where in this derivation we used the formula for differentiating bilinear operators and the chain rule.

Example 4 *(The product rule) Suppose* $B : Y \times Y \to Z$ *is a bilinear operator and* $f, g : U \subset X \to Y$ *. Set*

$$H(u) = B(f(u), g(u)).$$

Then

$$\begin{aligned}
H'(u)\,h &= \frac{d}{dt}\,B(f(u+th), g(u+th))\big|_{t=0} \\
&= B(f'(u+th)\,h, g(u+th)) + B(f(u+th), g'(u+th)\,h)\big|_{t=0} \\
&= B(f'(u)\,h, g(u)) + B(f(u), g'(u)\,h).
\end{aligned}$$

We can use this pattern to compute the second derivative:

$$\begin{aligned}
H''(u)\,hk &= B(f''(u)\,hk, g(u)) + B(f'(u)\,h, g'(u)\,k) \\
&\quad + B(f'(u)\,k, g'(u)\,h) + B(f(u), g''(u)\,hk).
\end{aligned}$$

3 The Implicit Function Theorem

In this section we assume that $F : U \subset X \times Y \to Z$ and $(u_0, v_0) \in U$ such that

$$F(u_0, v_0) = 0.$$

We are concerned with solving the equation

$$F(u,v) = 0$$

for u a neighbourhood $B_X(u_0,\rho) \subset X$ and v in a neighbourhood $B_Y(v_0,r) \subset Y$. The implicit function theorem states that under certain assumptions on F, this equation has a unique solution. The proof of the theorem is based on a variant of Newton's method which we will motivate now in a formal manner.

Solving the equation

$$G(u) = 0$$

is equivalent to finding the fixed points of the operator

$$H(u) = u - G'^{-1}(u_0)\, G(u).$$

For equations in two variables

$$F(u,v) = 0$$

we fix a value u and find the solution $v = v(u)$ corresponding to that u. This leads us to define the function

$$G_u(\cdot) = F(u,\cdot).$$

Then the problem reduces to solving

$$G_u(v) = 0$$

The Newton's method corresponding to this equation is

$$H_u(v) = v - G_u'^{-1}(v_0)\, G(v).$$

Noting that $G_u'(v_0) = F_v(u,v)$ (the subscript u is used as a parameter, while the subscript v is used to denote the partial derivative of F with respect to its second variable) we may rewrite Newton's method as

$$H_u(v) = v - F_v^{-1}(u_0,v_0)\, F(u,v). \tag{1}$$

Theorem 5 *(The Implicit Function Theorem) Suppose $F : U \subset X \times Y \to Z$ is C^m on U, $1 \le m \le \infty$.Suppose further that $(u_0,v_0) \in U$ satisfies*

1. $F(u_0,v_0) = 0,$
2. $F_v(u_0,v_0) : Y \to Z$ *is bijective.*

Then there exist $\rho > 0$, $r > 0$ such that

1. *for each $u \in B_X(u_0,\rho)$ the equation*

   $$F(u,v) = 0$$

   has a unique solution $v = v(u) \in \overline{B}_Y(v_0,r)$,

2. *the function* $u \mapsto v(u)$ *is* C^m *on* $B_X(u_0,\rho)$. *In particular*

$$v'(u) = -F_v^{-1}(u, v(u)) F(u,v) \;\; \forall\; u \in B_X(u_0,\rho).$$

Proof. For simplicity, assume $u_0 = 0$, $v_0 = 0$. The bijectivity of $F_v(0,0)$ and the closed graph theorem mean that $F_v(0,0)$ has a continuous inverse. The continuity of $F_v(u,v)$ on U and the vonNoemann theorem imply that $F_v(u,v)$ is continuously invertible for (u,v) in a neighbourhood $V \subset U$ of $(0,0)$. Let $\rho > 0$, $r > 0$ be small enough that the the mapping $H_u(\cdot)$ of equation (1) is well defined for u fixed but arbitry in $B_X(0,\rho)$ and all $v \in \overline{B}_Y(0,r)$. We will show that for sufficiently samll $\rho > 0$, $r > 0$ the mapping $H_u(\cdot)$ is a contraction. Since F is C^m, $m \geq 1$, we may write

$$\begin{aligned} F(u,v) &= F(0,0) + F'(0,0)\,uv + o((u,v)) \\ &= F_u(0,0)\,u + F_v(0,0)\,v + o((u,v)). \end{aligned}$$

Hence,

$$H_u(v) = -F_v^{-1}(0,0) F_u(0,0)\,u + o((u,v))$$

and

$$\begin{aligned} \|H_u(v)\| &\leq \left\|F_v^{-1}(0,0)\right\| \|F_u(0,0)\| \|u\| + o(\|(u,v)\|) \\ &\leq \left\|F_v^{-1}(0,0)\right\| \|F_u(0,0)\| \rho + o(\|(u,v)\|). \end{aligned}$$

Also,

$$\begin{aligned} \|H_u(v) - H_u(w)\| &\leq \left\|F_v^{-1}(0,0)\right\| \|F_v(0,0)(v-w) - (F(u,v) - F(u,w))\| \\ &\leq \left\|F_v^{-1}(0,0)\right\| \|(F_v(0,0) - F_v(u,v))(v-w)\| + o(\|v-w\|) \\ &\leq \left\|F_v^{-1}(0,0)\right\| \|F_v(0,0) - F_v(u,v)\| \|v-w\| + o(\|v-w\|). \end{aligned}$$

Therefore, for sufficiently small $\rho > 0$, $r > 0$ we get

1. $\|H_u(v)\| \leq r$,
2. $\|H_u(v) - H_u(w)\| \leq \frac{1}{2}\|v-w\|$.

It follows that H_u is a contraction on $\overline{B}_Y(0,r)$ for any $u \in B_X(0,\rho)$. Thus the equation $F(u,v) = 0$ has a unique solution $v = v(u) \in \overline{B}_Y(0,r)$ for every $u \in B_X(0,\rho)$. Next, we show continuity of the map $u \mapsto v(u)$. Let $u_1, u_2 \in B_X(0,\rho)$ and set $v_1 = v(u_1), v_2 = v(u_2)$. Then

$$\begin{aligned} \|v_1 - v_2\| &= \|H_{u_1}(v_1) - H_{u_2}(v_2)\| \\ &\leq \|H_{u_1}(v_1) - H_{u_1}(v_2)\| + \|H_{u_1}(v_2) - H_{u_2}(v_2)\| \\ &\leq \frac{1}{2}\|v_1 - v_2\| + \left\|F_v^{-1}(0,0)(F(u_1,v_2) - F(u_2,v_2))\right\|. \end{aligned}$$

Hence

$$\begin{aligned}\|v_1 - v_2\| &\leq 2\left\|F_v^{-1}(0,0)\left(F(u_1,v_2) - F(u_2,v_2)\right)\right\| \\ &\leq 2\left\|F_v^{-1}(0,0)\right\| \sup_{0<\theta<1} \|F_u(u_1 + \theta(u_2 - u_1), v_2)\| \|u_2 - u_1\| \\ &\leq 2M\left\|F_v^{-1}(0,0)\right\| \|u_2 - u_1\| \qquad (2)\end{aligned}$$

where

$$M = \sup_{\substack{v\in\overline{B}_Y(0,r),\\ u\in B_X(0,\rho)}} \|F_u(u,v)\|.$$

Note that $M < \infty$ because of the assumption that F is C^m on $U \supset B_X(0,\rho) \times \overline{B}_Y(0,r)$. Therefore, $\|v_1 - v_2\| \to 0$ as $\|u_2 - u_1\| \to 0$. We then proceed to show the differentiability of $u \mapsto v(u)$. For this we consider

$$\begin{aligned}F(u+h, v(u+h)) &= F(u, v(u)) + F_u(u, v(u))h + F_v(u, v(u))(v(u+h) - v(u)) + o(h,k) \\ &= F_u(u, v(u))h + F_v(u, v(u))(v(u+h) - v(u)) + o(h,k)\end{aligned}$$

where $k = v(u+h) - v(u)$. Since $F_v^{-1}(u,v)$ is uniformly bounded on $B_X(0,\rho) \times \overline{B}_Y(0,r)$, and $F(u+h, v(u+h)) = F(u, v(u)) = 0$, we get

$$v(u+h) - v(u) = -F_v^{-1}(u, v(u)) F_u(u, v(u)) h + o(h,k)$$

By equation (2) we see that $k = O(h)$ as $h \to 0$. It follows that $o(\|(h,k)\|) = o(\|h\|)$. Therefore, $v'(u)$ exists and

$$v'(u)h = -F_v^{-1}(u, v(u)) F_u(u, v(u)) h$$

for all $h \in Y$.

Higher derivatives of $u \mapsto v(u)$ can be shown similarly. ■

4 Application to Differential Equations

We want to investigate the solvability and continuous dependence on the initial data for the differential equaion

$$\left.\begin{array}{c} x'(t) = f(t, x(t)), \\ x(\tau) = y \end{array}\right\} \qquad (3)$$

for $t \in [\tau - a, \tau + a]$, $a > 0$ and (τ, y) vary around some nominal value (t_0, x_0). Here $f : \mathbb{R} \times X \to X$ is a C^1 function, $\tau,\ a \in \mathbb{R},\ y,\ x_0 \in X$.

Proposition 6 *Under the above assumptions, there exist $\rho > 0, r > 0$ such that for*

$$a, |\tau - t_0|, \|y - x_0\|_X < \rho$$

the system (3) has a unique solution $x = x(\cdot, a, \tau, y) \in \overline{B}(x_0, r)$ Moreover, the mapping $(a, \tau, y) \longmapsto x(a, \tau, y)$ is C^1 on $B((0, t_0, x_0), \rho)$.

Proof. The proof consists of two steps: 1- a rescaling, and 2- use of the implicit function theorem

Rescaling

Set $t = \tau + sa, \ s \in J = [-1, 1], \ v(s) = x(t) - y$. Then the system (3) is transformed into

$$\left.\begin{array}{c} v'(s) = af(\mu(s), v(s) + y) \\ v(0) = 0 \end{array}\right\} \tag{4}$$

where, $\mu(s) = \tau + sa$. Define the spaces

$$Y = C_0^1(J; X) = \left\{v \in C^1(J; X) : v(0) = 0\right\}$$

with the norm

$$\|v\|_1 = \max_{s \in J} \|v(s)\|_X + \max_{s \in J} \|v'(s)\|_X$$

and

$$Z = C(J; X)$$

with the norm

$$\|v\|_0 = \max_{s \in J} \|v(s)\|_X$$

Also, define the function

$$F : \mathbb{R} \times \mathbb{R} \times X \times Y \to Z$$

by

$$F(a, \tau, y, v) = v' - af(\mu, v - y).$$

Then equation (3) reduces to

$$F(a, \tau, y, v) = 0 \tag{5}$$

Now notice that

$$\begin{array}{rcl} F_a(a, \tau, y, v) & = & -f(\mu, v - y), \\ F_\tau(a, \tau, y, v) & = & -af_t(\mu, v - y), \\ F_y(a, \tau, y, v) & = & af_x(\mu, v - y), \\ F_v(a, \tau, y, v)h & = & h' - af_x(\mu, v - y)h \end{array}$$

(where the last equation is obtained by doing the formal evaluation $F_v(a, \tau, y, v)h = \frac{d}{dt} F(a, \tau, y, v + th)|_{t=0}$). These equations mean that F is C^1. Furthermore,

$$F(0, t_0, x_0, 0) = 0$$

and $F_v(0, t_0, x_0, 0) : Y \to Z$ is bijective since the equation

$$F_v(0, t_0, x_0, 0)h = w$$

or

$$\begin{array}{rcl} h' & = & w, \\ h(0) & = & 0 \end{array}$$

has the unique solution

$$h(s) = \int_0^s w(\sigma)\, d\sigma.$$

Therefore, by the implicit function theorem, there exist $\rho_1 > 0, r_1 > 0$ such that, for $(a, \tau, y) \in B((0, t_0, x_0), \rho_1)$, the equation (5) has a unique solution $v = v(\cdot, (a, \tau, y)) \in \overline{B}(0, r_1)$. The conclusions of the theorem follows from this. ■