CHAPTER 3: Differential Calculus

1 Differentiable Functions

First, we revise the concept of differentiability of a real valued function.

Definition 1 *Let $f : S \longrightarrow \mathbb{R}$ be a function where $S \subseteq \mathbb{R}$. The function f is differentiable at $x \in S$ if*

$$f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x} \in \mathbb{R}$$

exists. The function f is differentiable on S if it is differentiable at each $x \in S$.

Lemma 1 *Let $f : S \longrightarrow \mathbb{R}$ be a function where $S \subseteq \mathbb{R}$. If f is differentiable at a point x, then it is continuous at x.*

Proof. Take two points x and $x+h$ in S. Hence,

$$\lim_{h \to 0}[f(x+h) - f(x)] = \left[\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}\right]\left[\lim_{h \to 0} h\right] = f'(x).0 = 0.$$

The above implies that $\lim_{h \to 0} f(x+h) = f(x)$, and hence f is continuous at x. $\square$

The converse of the above lemma is not necessarily true. The function $f(x) = |x|$ is continuous on $[-1, 1]$, but is not differentiable at $x = 0$.

Definition 2 (Local maximizer) *A point x^0 is a local maximizer of a function $f : S \longrightarrow \mathbb{R}$, where $S \subseteq \mathbb{R}$, if there exists some $\delta > 0$ such that $f(x^0) \geq f(x)$ for all $x \in B_\delta(x^0)$.*

Theorem 1 *Let $f : (a, b) \longrightarrow \mathbb{R}$ be differentiable on (a, b) and x^0 be a local maximizer (minimizer) of f. Then $f'(x^0) = 0$.*

Proof. Suppose that f has a local maximum at x^0. Then we have $f(x^0 + h) - f(x^0) \leq 0$ for all h with $|h| < \delta$, and therefore,

$$\begin{aligned} \frac{f(x^0+h) - f(x^0)}{h} &\leq 0, \text{ for } h \in (0, \delta), \\ &\geq 0, \text{ for } h \in (-\delta, 0). \end{aligned}$$

Thus, we have

$$\lim_{h\to 0^+} \frac{f(x^0+h)-f(x^0)}{h} \le 0, \text{ and } \lim_{h\to 0^-} \frac{f(x^0+h)-f(x^0)}{h} \ge 0.$$

Differentiability of f implies that

$$0 \le f'(x^0) = \lim_{h\to 0} \frac{f(x^0+h)-f(x^0)}{h} \le 0,$$

and hence $f'(x^0) = 0$. $\square$

Theorem 2 (Rolle's theorem) *Let $f : [a, b] \longrightarrow \mathbb{R}$ be differentiable on (a, b) such that $f(a) = f(b) = \alpha$. Then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. Because f is continuous on a compact set $[a, b]$, by Weirstrass Theorem, there exist two points x_{min} and x_{max} in $[a, b]$ such that $f(x_{min}) = \min f(x)$ and $f(x_{max}) = \max f(x)$. If $f(x_{min}) = f(x_{max}) = \alpha$, then f is constant, and hence $f'(x) = 0$ for all $x \in [a, b]$. Otherwise, $f(x_{min}) < \alpha$ for $x_{min} \in (a, b)$ and $f'(x_{min}) = 0$ (because x_{min} is a local minimizer) or $f(x_{max}) > \alpha$ for $x_{max} \in (a, b)$ and $f'(x_{max}) = 0$ (because x_{max} is a local maximizer), or both. $\square$

Theorem 3 (Mean value theorem) *Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a differentiable function. If a and b are two points in $\mathbb{R}$ with $a < b$, then there exists some $c \in (a, b)$ such that*

$$f'(c) = \frac{f(b)-f(a)}{b-a}.$$

Proof. Define the following function

$$g(x) = f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x-a).$$

Because g satisfies the assumptions of Rolle's theorem, there exists some point c in (a, b) such that $g'(c) = 0$, i.e.,

$$g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a} = 0.$$

The above completes the proof. $\square$

If a function $f : \mathbb{R} \longrightarrow \mathbb{R}$ is differentiable at a point $x \in \mathbb{R}$, its derivative at x, $f'(x)$ is interpreted as the slope of the tangent to the function at the point x. Let $g(y) = my + c$ be the tangent to $f(y)$ at x. Intuitively, the derivative of f at x is the best linear approximation of f around x by the function g. This motivates the following generalized notion of differentiability.

Definition 3 *Let $f : S \longrightarrow \mathbb{R}^m$ be a function where S is an open set in $\mathbb{R}^n$. The function f is differentiable at $x \in S$ if there exists an $m \times n$ matrix M such that for all $\varepsilon > 0$, there exists a $\delta > 0$ such that $y \in S$ and $||x-y|| < \delta$ implies*

$$||f(x) - f(y) - M(x-y)|| < \varepsilon ||x-y||.$$

Equivalently, f is differentiable at $x \in S$ if

$$\lim_{y\to x} \frac{||f(y)-f(x)-M(y-x)||}{||y-x||} = 0.$$

The function f is differentiable on S if it is differentiable at each $x \in S$.

The matrix M is called the derivative of f at x and is denoted $Df(x)$. In case of $n = m = 1$, we denote $Df(x)$ by $f'(x)$. Let $g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ an affine function of the form $g(y) = My + c$, where M is an $m \times n$ matrix and $c \in \mathbb{R}^m$. The derivative of f at x is the best affine approximation of f around the point x by the function g. Here, the best means the ratio

$$\frac{||f(y) - g(y)||}{||y - x||}$$

goes to zero as $y \to x$. Since the values of f and g must coincide at x, we must have $g(x) = Mx + c = f(x)$ or $c = f(x) - Mx$. Thus, we may write the approximation function g as

$$g(y) = My - Mx + f(x) = M(y - x) + f(x).$$

Given this value for $g(y)$, the task of identifying the best affine approximation to f at x now amounts to identifying a matrix M such that

$$\frac{||f(y) - g(y)||}{||y - x||} = \frac{||f(y) - f(x) - M(y - x)||}{||y - x||} \to 0 \text{ as } y \to x.$$

This is precisely the definition of derivative given above.

When f is differentiable on S, the derivative Df itself forms a function from S to $\mathbb{R}^{m\times n}$. If $Df : S \longrightarrow \mathbb{R}^{m\times n}$ is a continuous function, then f is said to *continuously differentiable* on S, and we write f is $\mathscr{C}^1$. Consider now the following example.

Example 1 Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be given by

$$f(x) = \begin{cases} 0, & \text{if } x = 0, \\ x^2 \sin\left(1/x^2\right), & \text{if } x \neq 0. \end{cases}$$

For $x \neq 0$, we have

$$f'(x) = 2x \sin\left(\frac{1}{x^2}\right) - \left(\frac{2}{x}\right) \cos\left(\frac{1}{x^2}\right).$$

Since $|\sin(\cdot)| \leq 1$ and $|\cos(\cdot)| \leq 1$, but $2/x \to \infty$ as $x \to 0$, it is clear that $\lim_{x\to 0} f'(x)$ is not well defined. However, $f'(0)$ does exist! Indeed

$$f'(0) = \lim_{x\to 0} \frac{f(x) - f(0)}{x} = \lim_{x\to 0} x \sin\left(\frac{1}{x^2}\right) = 0.$$

The above implies that f is differentiable at $x = 0$, but Df is not continuous at this point. Thus, f is not $\mathscr{C}^1$ on $\mathbb{R}_+$. ■

Next, given functions $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and $h : \mathbb{R}^k \longrightarrow \mathbb{R}^n$, their composition is given by the function $f \circ h : \mathbb{R}^k \longrightarrow \mathbb{R}^m$ whose value at any $x \in \mathbb{R}^k$ is given by $f(h(x))$. Then

Lemma 2 (Chain rule of derivative) *Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and $h : \mathbb{R}^k \longrightarrow \mathbb{R}^n$, and let $x \in \mathbb{R}^k$. If h is differentiable at x, and f is differentiable at $h(x)$, then $f \circ h$ is itself differentiable at x, and its derivative is obtained through the "chain rule" as:*

$$D(f \circ h)(x) = Df(h(x))Dh(x).$$

2 Partial Derivatives

Definition 4 (Partial derivative) *Let $f : S \longrightarrow \mathbb{R}$, where $S \subset \mathbb{R}^n$ is open. Let e_j denote the vector in $\mathbb{R}^n$ that has a 1 in the j-th place and zeros elsewhere ($j = 1, \ldots, n$). Then the j-th partial derivative of f is said to exist at a point x if there is a number $\partial f(x)/\partial x_j$ such that*

$$\lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t} = \frac{\partial f}{\partial x_j}(x) \ \text{or} \ f_j(x).$$

For the partial derivatives, the following theorem is true.

Theorem 4 *Let $f : S \longrightarrow \mathbb{R}$, where $S \subset \mathbb{R}^n$ is open. Define the gradient vector of f at x by the vector of partial derivatives of f at x as $\triangledown f(x) := [f_1(x), \ldots, f_n(x)]$.*

(a) *If f is differentiable at x, then all partials $f_j(x)$ exist at x, and $Df(x) = \triangledown f(x)$.*

(b) *If all the partials exist and are continuous at x, then the derivative of f at x exists and is given by $Df(x) = \triangledown f(x)$.*

(c) *f is $\mathscr{C}^1$ on S if and only if all partials $f_j(x)$ exist and are continuous on S.*

Thus to check if f is $\mathscr{C}^1$, we only need to figure out if (a) the partial derivatives all exist on S, and (b) if they are all continuous on S. On the other hand, the requirement that the partial derivatives not only exist but be continuous at x is very important for the coincidence of the vector of partials with $Df(x)$. In the absence of this condition, all partials could exist at some point without the function itself being differentiable at that point. Consider the following example.

Example 2 Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be given by

$$f(x, y) = \begin{cases} 0, & \text{if } (x, y) = (0, 0), \\ \frac{xy}{\sqrt{x^2 + y^2}}, & \text{if } (x, y) \neq (0, 0). \end{cases}$$

We will show that f has all partial derivatives everywhere, but that these partials are not continuous at $(0, 0)$. Then we will show that f is not differentiable at $(0, 0)$. Since $f(x, 0) = 0$ for any $x \neq 0$, it is immediate that for all $x \neq 0$,

$$\frac{\partial f}{\partial y}(x, 0) = \lim_{\hat{y} \to 0} \frac{f(x, \hat{y}) - f(x, 0)}{\hat{y}} = \lim_{\hat{y} \to 0} \frac{x}{\sqrt{x^2 + \hat{y}^2}} = 1.$$

Similarly, at all points $(0, y)$ for $y \neq 0$, we have $\partial f(0, y)/\partial x = 1$. However, note that

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{x \to 0} \frac{f(x, 0) - f(0, 0)}{x} = 0 = \lim_{y \to 0} \frac{f(0, y) - f(0, 0)}{y} = \frac{\partial f}{\partial y}(0, 0).$$

So, $\partial f(0, 0)/\partial x$ and $\partial f(0, 0)/\partial y$ exist at $(0, 0)$. But

$$\lim_{x \to 0} \frac{\partial f}{\partial y}(x, 0) = \lim_{y \to 0} \frac{\partial f}{\partial x}(0, y) = 1 \neq 0.$$

Thus, the partials are not continuous at $(0,0)$. Now suppose that f were differentiable at $(0,0)$. Then we must have $Df(0,0)=(0,0)$. Take the points (x,y) of the form (a,a) for some $a>0$, and note that every open neighborhood of $(0,0)$ must contain at least one such point. Since $f(0,0)=0$, $Df(0,0)=(0,0)$ and $||(x,y)||=\sqrt{x^2+y^2}$, we have

$$\lim_{a\to 0}\frac{||f(a,a)-f(0,0)-Df(0,0).(a,a)||}{||(a,a)-(0,0)||}=\lim_{a\to 0}\frac{a^2}{2a^2}=\frac{1}{2}\neq 0.$$

Thus, f is not differentiable at $(0,0)$. ■

The failure of the existence of derivative in the above example induces a generalized notion of derivative which is studied in the following section. In what follows we extend the concept of derivative of a vector-valued function.

Definition 5 (Jacobian matrix) *Let $f:S\longrightarrow\mathbb{R}^m$, where $S\subseteq\mathbb{R}^n$ is open, assigns to each $x=(x_1,\ldots,x_n)\in\mathbb{R}^n$ a vector $f(x)=(f^1(x),\ldots,f^m(x))$ in $\mathbb{R}^m$. The Jacobian matrix of f at $x\in S$ is the $m\times n$ matrix of partial derivatives which is given by*

$$J_f(x):=\begin{bmatrix}\frac{\partial f^1}{\partial x_1}(x) & \ldots & \frac{\partial f^1}{\partial x_n}(x)\\ \vdots & \ddots & \vdots\\ \frac{\partial f^m}{\partial x_1}(x) & \ldots & \frac{\partial f^m}{\partial x_n}(x)\end{bmatrix}$$

Following in an extension of Theorem 4 in case of a vector-valued function.

Theorem 5 *Let $f:S\longrightarrow\mathbb{R}^m$, where $S\subset\mathbb{R}^n$ is open.*

(a) *If f is differentiable at x, then all partials $\partial f^i/\partial x_j$ for $i=1,\ldots,m$ and $j=1,\ldots,n$ exist at x, and $Df(x)=J_f(x)$.*

(b) *If all the partials $\partial f^i/\partial x_j$ for $i=1,\ldots,m$ and $j=1,\ldots,n$ exist and are continuous at x, then the derivative of f at x exists and is given by $Df(x)=J_f(x)$.*

(c) *f is $\mathscr{C}^1$ on S if and only if all partials $\partial f^i/\partial x_j$ for $i=1,\ldots,m$ and $j=1,\ldots,n$ exist and are continuous on S.*

Example 3 Let $f:\mathbb{R}^2\longrightarrow\mathbb{R}^2$ such that $f(x,y)=(x^2+x^2y+10y,\ x+y^3)$. The Jacobian of f at $(x,y)\in\mathbb{R}^2$ is given by

$$Df(x,y)=\begin{bmatrix}2x(1+y) & x^2+10\\ 1 & 3y^2\end{bmatrix}.\quad ■$$

3 Directional Derivatives

Definition 6 (Directional derivative) *Let $f:S\longrightarrow\mathbb{R}$, where $S\subset\mathbb{R}^n$ is open. Let x be a point in S and let $h\in\mathbb{R}^n$. The directional derivative of f at x in the direction h is defined as*

$$Df(x;h)=\lim_{t\to 0}\frac{f(x+th)-f(x)}{t},\quad \text{where } t\in\mathbb{R} \text{ and } ||h||=1,$$

whenever this limit exists.

Theorem 6 *Suppose f is differentiable at $x \in S$. Then, for any $h \in \mathbb{R}^n$, the directional derivative $Df(x;h)$ of f at x in the direction h exists, and we have $Df(x;h) = \nabla f(x) \cdot h$.*

Example 4 Let $f(x_1,x_2) = x_1x_2$, $h = (3/5, 4/5)$ and $x^0 = (1,2)$. First we compute $Df(x^0;h)$, and then verify the above result. The directional derivative of f at x in the direction h is given by

$$Df(x_1,x_2;h) = \lim_{t\to 0}\frac{f(x+th)-f(x)}{t} = \lim_{t\to 0}\frac{\left(x_1+\frac{3t}{5}\right)\left(x_2+\frac{4t}{5}\right)-x_1x_2}{t} = \frac{4x_1}{5}+\frac{3x_2}{5}.$$

Therefore,

$$Df(x^0;h) = D(1,2;(3/5,4/5)) = 2.$$

On the other hand,

$$\nabla f(x_1,x_2) = (x_2,x_1)$$

Therefore,

$$\nabla f(x_1^0,x_2^0)\cdot h = (2,\,1)\begin{pmatrix}3/5\\4/5\end{pmatrix} = 2 = Df(1,2;h). \quad \blacksquare$$

4 Higher Order Derivatives

We have discussed earlier that, for a function $f : S \longrightarrow \mathbb{R}$ where $S \subset \mathbb{R}^n$ is open, which is differentiable on S, the derivative Df is itself a function from S to $\mathbb{R}^n$. Suppose now that there is an $x \in S$ such that Df is differentiable at x, i.e., such that for each $i = 1,\ldots,n$, the function $f_j : S \longrightarrow \mathbb{R}$ is differentiable at x. Denote the partial derivative of f_i in the direction e_j at x by $f_{ij}(x)$ or $\partial^2 f(x)/\partial x_j \partial x_i$ if $i \neq j$, and by $f_{ii}(x)$ or $\partial^2 f(x)/\partial^2 x_j$ if $i = j$. The *Hessian matrix* of f at x is given by

$$H[f(x)] := \begin{bmatrix} f_{11}(x) & \ldots & f_{1n}(x)\\ \vdots & \ddots & \vdots\\ f_{n1}(x) & \ldots & f_{nn}(x)\end{bmatrix}$$

Definition 7 *A function $f : S \longrightarrow \mathbb{R}$, where $S \subset \mathbb{R}^n$ is open, is twice-differentiable at x if the second derivative $D^2f(x)$ equals the Hessian matrix of f at x, i.e., $D^2f(x) = H[f(x)]$. For $n = 1$, we denote $D^2f(x)$ by $f''(x)$. If f is twice-differentiable at each $x \in S$, then f is twice-differentiable on S. If for each i, the cross partial $f_{ij} : S \longrightarrow \mathbb{R}$ is continuous, then f is twice continuously differentiable on S, and we write f is $\mathscr{C}^2$.*

Theorem 7 (Young's theorem) *If $f : S \longrightarrow \mathbb{R}$ is a $\mathscr{C}^2$ function, then D^2f is a symmetric matrix, i.e., we have*

$$f_{ij}(x) = f_{ji}(x) \text{ for all } i,j = 1,\ldots,n, \text{ and } x \in S.$$

The above asserts a one-way implication. The matrix D^2f may fail to be symmetric if a function is not $\mathscr{C}^2$.

5 Taylor's Theorem

In this section we discuss a generalization of the Mean value theorem, known as Taylor's theorem. The idea is that a many times differentiable function can be approximated by a polynomial. The notation $f^{(k)}(z)$ denotes the k-th derivative of f at a point z, and $k = 0$ implies that $f^{(k)}(z) = f(z)$.

Theorem 8 (Taylor's theorem in $\mathbb{R}$) *Let $f : (a, b) \longrightarrow \mathbb{R}$ be an m-times continuously differentiable function. Suppose also that $f^{(m+1)}(z)$ exists for every $z \in (a, b)$. Then for any $x, y \in (a, b)$, there is a $z \in (x, y)$ such that*

$$f(y) = \sum_{k=0}^{m} \frac{f^{(k)}(x)(y-x)^k}{k!} + \frac{f^{(m+1)}(z)(y-x)^{(m+1)}}{(m+1)!}.$$

Example 5 We would like to approximate $f(y) = e^y$ around $x = 0$ by a polynomial $P_m(y)$. Notice that $f^{(k)}(x) = e^x$ for all $k = 0, \ldots, m$. Then $f^{(k)}(0) = 1$ for all $k = 0, \ldots, m$. Thus, applying Taylor's theorem for $\mathbb{R}$ and ignoring the remainder term, we have

$$e^y \approx 1 + y + \frac{y^2}{2!} + \ldots + \frac{y^m}{m!} \equiv P_m(y). \quad \blacksquare$$

Taylor's theorem gives us a formula for constructing a polynomial approximation to a differentiable function. For $m = 0$, we obtain the Mean Value Theorem. With $m = 2$, and omitting the remainder, we get

$$f(x+h) \approx f(x) + f'(x)h.$$

With f differentiable, the remainder term will be very small. Thus, the linear function on the right-hand-side of the above equation seems to be a good approximation to $f(\cdot)$ around x. Following is a generalization of the above theorem.

Theorem 9 (Taylor's theorem in $\mathbb{R}^n$) *Let $f : S \longrightarrow \mathbb{R}$ be a $\mathscr{C}^1$ function, where $S \subset \mathbb{R}^n$ is open. Then for any $x, y \in S$, we have*

$$f(y) = f(x) + Df(x) \cdot (y-x) + R_1(x, y), \text{ where } \lim_{y \to x} \frac{R_1(x, y)}{||y-x||} = 0.$$

Proof. See Sundaram (1996, pp. 64-65). □

Example 6 Let i, r and π denote the nominal rate of interest, the real rate of interest and the rate of inflation. The nominal rate of interest is given by the formula $1 + i = (1+r)(1+\pi)$. Define by $f(r, \pi) = (1+r)(1+\pi)$. Let $(r^0, \pi^0) = (0, 0)$. Notice that $Df(0, 0) = (1, 1)$. Then by Taylor's expansion of $f(r, \pi)$ around (r^0, π^0) we have

$$1 + i = f(r, \pi) \approx f(r^0, \pi^0) + Df(r^0, \pi^0) \cdot (r - r^0, \pi - \pi^0) = 1 + (1, 1)^T \cdot (r, \pi) = 1 + r + \pi$$

The above implies $i \approx r + \pi$. $\blacksquare$

Example 7 (Rule of 70) With compound interest rate, the time it takes for an initial investment to double is $70/100r$ years. Let T is the time taken, i.e.,

$$\begin{aligned} & (1+r)^T I = 2I \\ \implies & T\ln(1+r) = \ln 2 \\ \implies & T = \frac{\ln 2}{\ln(1+r)} \approx \frac{0.693147}{r} = \frac{70}{100r} \end{aligned}$$

since $\ln(1+r) \approx r$. ■

Definition 8 (Total derivative) *$f : S \longrightarrow \mathbb{R}$ be a $\mathscr{C}^1$ function, where $S \subset \mathbb{R}^n$. The total derivative f at $x \in S$ is defined as*

$$df(x) = \nabla f(x) \cdot dx = \sum_{i=1}^{n} f_i(x)\, dx_i.$$

Example 8 (Indifference curves) Let $u : \mathbb{R}^2_+ \longrightarrow \mathbb{R}$ be the continuously differentiable utility function of a consumer derived from the consumption of two goods 1 and 2 in quantities $x = (x_1, x_2)$. The indifference curve at level α is the set $\{x \in \mathbb{R}^2_+ \mid u(x) = \alpha\}$. The total derivative of u at x is given by

$$du(x) = u_1(x)dx_1 + u_2(x)dx_2 = 0.$$

The above equation implies that

$$\frac{dx_2}{dx_1} = -\frac{u_1(x)}{u_2(x)} = MRS_{12}(x)\,.$$

Given that the marginal utilities are positive, the indifference curve between goods 1 and 2 is negatively sloped. ■

6 Inverse and Implicit Function Theorems

Given two sets A and B, if a function $f : A \longrightarrow B$ is one-to-one and onto, then there is a unique function $f^{-1} : B \longrightarrow A$ such that $f(f^{-1}(b)) = b$ for all $b \in B$. The function g is called the *inverse function* of f.

Theorem 10 (Inverse function theorem) *Let $f : S \longrightarrow \mathbb{R}^n$ be a $\mathscr{C}^1$ function, where $S \subset \mathbb{R}^n$ is open. Suppose there is a point $x \in S$ such that the $n \times n$ matrix $Df(x)$ is invertible. Let $y = f(x)$. Then*

(a) *There are open sets U and V in $\mathbb{R}^n$ such that $x \in U$, $y \in V$, f is one-to-one on V, and $f(U) = V$.*

(b) *The inverse function $f^{-1} : V \longrightarrow U$ of f is a $\mathscr{C}^1$ function, whose derivative at any point $y^0 \in V$ satisfies*

$$Df^{-1}(y^0) = (Df(x^0))^{-1}, \text{ where } f(x^0) = y^0.$$

Example 9 Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be defined by $f(x, y) = (x^2 + x^2 y + 10y, x + y^3)$. We will show that f has an inverse in the neighborhood of $(1, 1)$. We have

$$Df(x, y) = J_f(x, y) = \begin{bmatrix} 2x(1+y) & x^2 + 10 \\ 1 & 3y^2 \end{bmatrix}$$

Thus, $f(1, 1) = (12, 2)$ and

$$Df(1, 1) = \begin{bmatrix} 4 & 11 \\ 1 & 3 \end{bmatrix},$$

and $det(Df(1, 1)) = 1 \neq 0$. Therefore, $Df(1, 1)$ is invertible. By the Inverse function theorem, we deduce that there is an open set $U \subset \mathbb{R}^2$ containing $(1, 1)$ such that f when restricted to U has a continuously differentiable inverse f^{-1}, and

$$Df^{-1}(1, 1) = \begin{bmatrix} 3 & -11 \\ -1 & 4 \end{bmatrix} = (Df(1, 1))^{-1}. \quad \blacksquare$$

Now, consider the function $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by $F(x, y) = (x-2)^3 y + xe^{y-1}$, and suppose we are interested in solving the equation $F(x, y) = 0$. We will ask the question whether it is possible to define y as a function of x in some neighborhood of (x^*, y^*). This question motivates the following theorem. We introduce some additional notations. Given integers $m \geq 1$ and $n \geq 1$, let a typical point in $\mathbb{R}^{m+n}$ be denoted by (x, y), where $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$. For a $\mathscr{C}^1$ function F mapping some subset of $\mathbb{R}^{m+n}$ into $\mathbb{R}^n$, let $DF_y(x, y)$ denote the portion of the matrix $DF(x, y)$, which is an $n \times (m+n)$ matrix, corresponding to the last n variables. Notice that $DF_y(x, y)$ is a $n \times n$ matrix. Define $DF_x(x, y)$ similarly, which is an $n \times m$ matrix.

Theorem 11 (Implicit function theorem) *Let $F : S \longrightarrow \mathbb{R}^n$ be a $\mathscr{C}^1$ function, where $S \subset \mathbb{R}^{m+n}$ is open. Let (x^*, y^*) be a point in S such that $DF_y(x^*, y^*)$ is invertible, and let $F(x^*, y^*) = c$. Then, there is a neighborhood $U \subset \mathbb{R}^m$ of x^* and a $\mathscr{C}^1$ function $g : U \longrightarrow \mathbb{R}^n$ such that*

(a) *$(x, g(x))$ is in S for all $x \in U$,*

(b) *$g(x^*) = y^*$,*

(c) *$F(x, g(x)) = c$ for all $x \in U$.*

The derivative of g at any $x \in U$ is obtained from the chain rule:

$$Dg(x) = -(DF_y(x, y))^{-1} DF_x(x, y).$$

Example 10 Consider the equation $F(x, y) = (x-2)^3 y + xe^{y-1} = 0$. We will show that y can be defined implicitly as a function of x in the neighborhood of $(0, 0)$ but not around $(1, 1)$. First notice that $F(0, 0) = F(1, 1) = 0$. We have $DF_y(x, y) = \partial F(x, y)/\partial y = (x-2)^3 + xe^{y-1}$. Now, $DF_y(0, 0) = -8 \neq 0$, and hence $DF_y(0, 0)$ is invertible. But $DF_y(1, 1) = 0$, and hence $DF_y(1, 1)$ is not invertible. $\blacksquare$

7 Homogeneous Functions

This is a special class of functions used frequently in economics. In what follows, we will study some important properties associated with homogeneous functions.

Definition 9 *A set S in $\mathbb{R}^n$ is a cone if given any $x \in S$, the point λx belongs to S for any $\lambda > 0$.*

Definition 10 (Homogeneous function) *A function $f : S \longrightarrow \mathbb{R}$, where S is a cone in $\mathbb{R}^n$, is homogeneous of degree r in S if, for all $\lambda > 0$,*

$$f(\lambda x) = \lambda^r f(x).$$

Consider the response of a consumer to an equiproportional increase in income and prices of all commodities in the market. In this case, the consumer's choice is, in general, not altered. Formally, the demand function $x(p, m)$ will be homogeneous of degree zero. Following theorem provides an useful characterization of homogeneous functions.

Theorem 12 (Euler's theorem) *Let $f : S \longrightarrow \mathbb{R}$ be a function with continuous partial derivatives defined on an open cone S in $\mathbb{R}^n$. Then f is homogeneous of degree r in S if and only if*

$$\sum_{i=1}^{n} f_i(x)x_i = rf(x), \text{ for all } x \in S. \tag{1}$$

Proof. Assume that f is homogeneous of degree r, and fix an arbitrary x in S. Then we have, for all $\lambda > 0$,

$$f(\lambda x) = \lambda^r f(x).$$

The continuity of the partials guarantees the differentiability of f. Differentiating the above with respect to λ and using the chain rule we get

$$\sum_{i=1}^{n} f_i(\lambda x)x_i = r\lambda^{r-1} f(x).$$

Putting $\lambda = 1$, we get Condition (1). Conversely, suppose that (1) holds for all $x \in S$. Fix an arbitrary x, and define the function ϕ for all $\lambda > 0$ by

$$\phi(\lambda) = f(\lambda x).$$

Then

$$\phi'(\lambda) = \sum_{i=1}^{n} f_i(\lambda x)x_i,$$

and multiplying both sides of the above expression by λ gives

$$\lambda\phi'(\lambda) = \sum_{i=1}^{n} f_i(\lambda x)\lambda x_i = rf(\lambda x) = r\phi(\lambda), \tag{2}$$

where the second equality is obtained by applying (1) at the point λx. Next, define the function F for $\lambda > 0$ by

$$F(\lambda) = \frac{\phi(\lambda)}{\lambda^r}, \tag{3}$$

and observe that, using (2),

$$F'(\lambda) = \frac{\lambda^{r-1}}{(\lambda^r)^2}[\lambda\phi'(\lambda) - r\phi(\lambda)] = 0.$$

Hence, F is a constant function. Putting $\lambda = 1$ in (3), we have $F(1) = \phi(1)$, and therefore

$$F(\lambda) = \frac{\phi(\lambda)}{\lambda^r} = \phi(1) \implies \phi(\lambda) = \lambda^r \phi(1).$$

Finally, since $\phi(\lambda) = f(\lambda x)$, we have $f(\lambda x) = \lambda^r f(x)$. $\square$

Example 11 Following are the examples of homogeneous functions.

(a) The Cobb-Douglas function: $f(x) = Ax_1^{\alpha_1} \dots x_n^{\alpha_n}$.

(b) The CES function: $f(x) = A\left(\sum_{i=1}^{n} \alpha_i x_i^{-\rho}\right)^{-\frac{1}{\rho}}$, where $A > 0$, $\rho > -1$, $\rho \neq 0$, $\alpha_i > 0$ for all i, and $\sum_i \alpha_i = 1$. ■

Homogeneous functions have nice geometric properties. The indifference curves of a homogeneous function are parallel to each other.