Theorem 90. *(Taylor's Theorem) Let $f : (a,b) \subset \mathbb{R} \to \mathbb{R}$ be n times differentiable on the open interval (a,b). For all $x, x+h \in (a,b)$*

$$f(x+h) = f(x) + \sum_{k=1}^{n-1} \frac{1}{k!} f^{(k)}(x) h^k + e_n(x) \tag{20}$$

where $e_n(x) = \frac{1}{n!} f^{(n)}(x + \lambda h) h^n$ for some $\lambda \in (0,1)$.

Proof. Let $y = x + h$ and define $F(z) = f(y) - f(z) - \sum_{k=1}^{n-1} \frac{1}{k!} f^{(k)}(z)(y-z)^k$. The theorem says that for some $x + \lambda(y-x)$ (strictly) between x and y

$$F(x) = f(x+h) - f(x) - \sum_{k=1}^{n-1} \frac{1}{k!} f^{(k)}(x) h^k = \frac{1}{n!} f^{(n)}(x + \lambda(y-x))(y-x)^n.$$

Observe that $F(y) = 0$ and

$$\begin{aligned} F^{(1)}(z) &= -f^{(1)}(z) - f^{(2)}(z)(y-z) + f^{(1)}(z) - \frac{1}{2} f^{(3)}(z)(y-z)^2 + f^{(2)}(z)(y-z) + \cdots \\ &\quad - \frac{1}{(n-1)!} f^{(n)}(z)(y-z)^{n-1} + f^{(n-1)}(z)(y-z)^{n-2} = -\frac{1}{(n-1)!} f^{(n)}(z)(y-z)^{n-1}. \end{aligned}$$

Next, define $G(z) = F(z) - \left(\frac{y-z}{y-x}\right)^n F(x)$. Note that $G(x) = F(x) - F(x) = 0$, $G(y) = F(y) = 0$ and

$$G^{(1)}(z) = F^{(1)}(z) - n(-1)\left(\frac{y-z}{y-x}\right)^{n-1} \frac{F(x)}{y-x}.$$

By Theorem 88, there exists $\lambda \in (0,1)$ such that $G^{(1)}(x + \lambda(y-x)) = 0$. But

$$G^{(1)}(x + \lambda(y-x)) = F^{(1)}(x + \lambda(y-x)) + n\left(\frac{y - x - \lambda(y-x)}{y-x}\right)^{n-1} \frac{F(x)}{y-x} = 0$$

implies

$$\frac{f^{(n)}(x + \lambda(y-x))}{(n-1)!}(y - x - \lambda(y-x))^{n-1} = n\left(\frac{y - x - \lambda(y-x)}{y-x}\right)^{n-1} \frac{F(x)}{y-x}$$

or

$$\frac{f^{(n)}(x + \lambda(y-x))}{n(n-1)!} = \frac{F(x)}{(y-x)^n} \iff F(x) = \frac{1}{n!} f^{(n)}(x + \lambda(y-x))(y-x)^n.$$

$\square$

Remark 25. *The term denoted by $e_n(x)$, i.e., the difference $f(x+h)-f(x)-\sum_{k=1}^{n-1}\frac{1}{k!}f^{(k)}(x)h^k$ is usually called the "remainder" term in Taylor's Theorem. The particular structure assumed by the remainder in this case is known as the LaGrange form of the remainder. Other expressions exist for this difference.*

The next two theorems provide alternative forms of the remainder in Taylor's Theorem. The proof of the next theorem relies on Riemann integration and the fundamental theorem of calculus, which we have not yet studied. Thus, you can postpone its reading. The proof of the second theorem, which is omitted in these notes, depends additionally on Fubini's Theorem, which will be studied when we deal with abstract integration in Econ 7818.

Theorem 91. *If f satisfies the conditions in theorem 90, we have*

$$f(x+h) = f(x) + \sum_{k=1}^{n-1}\frac{1}{k!}f^{(k)}(x)h^k + \int_x^{x+h}\int_x^{s_1}\cdots\int_x^{s_n-1} f^{(n)}(s_n)ds_n\cdots ds_1.$$

Proof. The proof is by induction. First, by the fundamental theorem of calculus, $f(x+h) - f(x) = \int_x^{x+h} f^{(1)}(s_1)ds_1$. Similarly, $f^{(1)}(s_1) = f^{(1)}(x) + \int_x^{s_1} f^{(2)}(s_2)ds_2$, and this gives

$$f(x+h) = f(x) + f^{(1)}(x)h + \int_x^{x+h}\int_x^{s_1} f^{(2)}(s_2)ds_2ds_1.$$

Continuing in a similar fashion, we have

$$\begin{aligned} f(x+h) = f(x) + f^{(1)}(x)h + f^{(2)}(x)\int_x^{x+h}\int_x^{s_1}\int_x^{s_2} ds_3ds_2ds_1 + \cdots \\ + f^{(n-1)}(x)\int_x^{x+h}\int_x^{s_1}\cdots\int_x^{s_{n-1}} ds_n\cdots ds_2ds_1 \\ + \int_x^{x+h}\int_x^{s_1}\cdots\int_x^{s_{n-1}} f^{(n)}(s_n)ds_n\cdots ds_2ds_1. \end{aligned}$$

It is easy to verify by induction that $\int_x^{x+h}\int_x^{s_1}\cdots\int_x^{s_{k-1}} ds_k\cdots ds_2ds_1 = \frac{h^k}{k!}$ (verify it trivially for $k=1$, assume it holds for k and show that it holds for $k+1$), which proves the theorem.

$\square$

Theorem 92. *(Taylor's Theorem with error in Cauchy form) If f satisfies the conditions in Theorem 90, we have*

$$f(x+h) = f(x) + \sum_{k=1}^{n-1} \frac{1}{k!} f^{(k)}(x) h^k + \frac{1}{(n-1)!} \int_x^{x+h} f^{(n)}(y)(x+h-y)^{n-1} dy.$$

16.3 Gâteaux and Fréchet differentiability

We first consider extending the notion of a differentiability for functions defined on $\mathbb{R}^n$ and taking values in $\mathbb{R}$.

Definition 60. *Let S be an open subset of $\mathbb{R}^n$ and $f : S \to \mathbb{R}$. The directional derivative of f in the direction $u \in \mathbb{R}^n$ at x is defined by the following limit, whenever it exists,*

$$Df(x; u) = \lim_{h \to 0} \frac{1}{h}(f(x+hu) - f(x))$$

where $h \in \mathbb{R} - \{0\}$, $\|u\|_E = 1$.

The limit $Df(x; u)$, whenever it exists, clearly depends on x and the direction u. There is no need to define a norm or a metric on $\mathbb{R}^n$ to obtain $Df(x; u)$. When u is a unit coordinate vector, i.e.,

$$u^T := e_i^T = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix},$$

with the number 1 on the i^{th} position, we have the *partial derivative* of f with respect to its i^{th} argument. In this case, we denote this particular directional derivative by

$$\frac{\partial}{\partial_i} f(x) = \lim_{h \to 0} \frac{1}{h}(f(x+hu) - f(x)).$$

Remark 26.

1. *The vector $\nabla f(x) = \begin{pmatrix} \frac{\partial}{\partial_1} f(x) & \cdots & \frac{\partial}{\partial_n} f(x) \end{pmatrix}$ is called the gradient of f at x.*

2. *If f is linear, the limit always exists and does not depend on x, only on u. In this case $Df(x; u) = f(u)$.*

3. *It is clear that if $Df(x;u)$ exists in all directions u, then all partial derivatives exist. The converse, however, is not true. For example, let $f : \mathbb{R}^2 \to \mathbb{R}$*

$$f(x_1, x_2) = \begin{cases} x_1 + x_2, & \textit{if } x_1 = 0 \textit{ or } x_2 = 0 \\ 1, & \textit{otherwise.} \end{cases}$$

By definition

$$\frac{\partial f}{\partial_1}(x) = Df(x;(1,0)) = \lim_{h \to 0} \frac{1}{h}(f(x_1 + h, x_2) - f(x_1, x_2)).$$

Thus, $\frac{\partial f}{\partial_1}((0,0)) = \lim_{h \to 0} \frac{1}{h}(h - 0) = 1$. Similarly, $\frac{\partial}{\partial_2} f(0,0) = 1$. Thus, the partial derivatives exists at $(0,0)$.

Now, consider a direction $u = (u_1, u_2)$ where $u_1, u_2 \neq 0$. Now, letting $\theta = (0,0)$ we have

$$\frac{f(\theta + hu) - f(\theta)}{h} = \frac{f(hu) - f(\theta)}{h} = \frac{1 - 0}{h} = \frac{1}{h} \to \infty$$

as $h \to 0$. Hence, no directional derivative for $u \neq \theta$ exists.

4. *It is interesting that existence of all directional derivatives at a certain point x does not imply continuity of the function at that point. This is, of course, in contrast to the case of functions defined on the real line, where the existence of a derivative at a point implies continuity. For example, let $f : \mathbb{R}^2 \to \mathbb{R}$ with*

$$f(x) := f(x_1, x_2) = \begin{cases} \frac{x_1 x_2^2}{x_1^2 + x_2^4}, & \textit{if } x_1 \neq 0 \\ 0, & \textit{otherwise.} \end{cases}$$

Then,

$$\frac{f(\theta + hu) - f(\theta)}{h} = \frac{f(hu)}{h} = \frac{u_1 u_2^2}{u_1^2 + h^2 u_2^4}$$

and,

$$\lim_{h \to 0} \frac{f(\theta + hu) - f(\theta)}{h} = \begin{cases} \frac{u_2^2}{u_1}, & \textit{if } u_1 \neq 0 \\ 0, & \textit{otherwise.} \end{cases}$$

Therefore, $Df(\theta; u)$ exists at θ for all directions. But note that $f(x_2^2, x_2) = \frac{1}{2}$ for all x_2, except when $x_2 = 0$, where $f(x_2^2, x_2) = 0$. Hence, f is discontinuous at θ in spite of the fact that all directional derivatives exist at θ.

We now give two more general definitions of differentiability.

Definition 61. *(Gâteaux differential) Let $x \in D \subseteq \mathbb{X}$, $f : D \subset \mathbb{X} \to (\mathbb{Y}, ||\cdot||_{\mathbb{Y}})$, $u \in \mathbb{X}$ and $h \in \mathbb{R} - \{0\}$. If the limit*

$$\delta_g(x; u) := \lim_{h \to 0} \frac{f(x + hu) - f(x)}{h}.$$

exists, $\delta_g(x; u)$ is called the Gâteaux differential of f at x with increment u. If it exists for all $u \in \mathbb{X}$, we say that f is Gâteaux differentiable at x.

The set $\left\{\frac{f(x+hu)-f(x)}{h}\right\}_{h \in \mathbb{R}-\{0\}} \subset \mathbb{Y}$ and as mentioned above, the limit in the definition means that for all $\epsilon > 0$ there exist $\delta(\epsilon) > 0$ such that whenever $|h| < \delta(\epsilon)$

$$\frac{1}{|h|} \| f(x + hu) - f(x) - h\delta_g(x; u) \|_{\mathbb{Y}} < \epsilon.$$

Also, for fixed $x \in D$, $\delta_g(x; u)$ can be thought of as a function from $\mathbb{X}$ ($u \in \mathbb{X}$) to $\mathbb{Y}$. Here, there is no need to have a norm on $\mathbb{X}$.

Definition 62. *(Fréchet differential) A function $f : D \subset (\mathbb{X}, \| \cdot \|_{\mathbb{X}}) \to (\mathbb{Y}, \| \cdot \|_{\mathbb{Y}})$, where D is an open set, is Fréchet differentiable at $x \in D$ if for each $h \in \mathbb{X}$ there exists $\delta f(x, h)$ which is linear and continuous with respect to h, taking values in $\mathbb{Y}$ such that*

$$\lim_{\|h\|_{\mathbb{X}} \to 0} \frac{\| f(x + h) - f(x) - \delta f(x, h) \|_{\mathbb{Y}}}{\|h\|_{\mathbb{X}}} = 0.$$

If f is Fréchet differentiable everywhere in D, we say that f is Fréchet differentiable on D. $\delta f(x, h)$ is called the Fréchet differential of f at x with increment h.

Remark 27.

1. *Since the Fréchet differential is linear and continuous on $\mathbb{X}$ with respect to h (for fixed x), by Theorem 48 it is bounded. That is, $\delta f(x, h) \in B(\mathbb{X}, \mathbb{Y}) \subset L(\mathbb{X}, \mathbb{Y})$ as a function of h ($h \neq \theta$).*

2. *In general, we can write that there exists a bounded linear function A_x (indexed by $x \in D$) that acts on $h \in \mathbb{X}$ so that $\delta f(x, h) = A_x(h)$. A_x (or $A(x)$) can be viewed as a*

function from $D \subset \mathbb{X}$ to the space of bounded linear functions from $\mathbb{X}$ to $\mathbb{Y}$. That is, for every x we have an $A(x) \in B(\mathbb{X}, \mathbb{Y})$ that operates on $h \in \mathbb{X}$ to produce an element of $\mathbb{Y}$. We call $A(x)$ the Fréchet derivative of f at x and denote it by $f'(x)$. That is, $f' : D \to B(\mathbb{X}, \mathbb{Y})$ with $\delta f(x,h) = f'(x)(h)$.

3. *By Theorem 46, if dim$(\mathbb{X}) = n$ and dim$(\mathbb{Y}) = m$, with $n, m \in \mathbb{N}$, the space $L(\mathbb{X}, \mathbb{Y})$ is isomorphic with $\mathcal{M}^{m \times n}$. Hence, there exists a unique matrix $A(x) := f'(x)$ such that $\delta f(x,h) = A(x)h$.*

4. *Differentiability, as given on Definition 57, is a special case of Fréchet differentiability. In that case, $D = (a,b) \subset \mathbb{X} = \mathbb{R}$ and differentiability at $x \in (a,b)$ meant that the following limit,*
$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$
denoted by $f^{(1)}(x)$, exists. But this implies that
$$\lim_{h \to 0} \left(\frac{f(x+h) - f(x) - hf^{(1)}(x)}{h} \right) = \lim_{h \to 0} \left(\frac{E(x,h)}{h} \right) = 0.$$
Therefore, the definition of differentiability requires that $\frac{E(x,h)}{h} \to 0$ as $h \to 0$.

5. *We say that the Fréchet derivative $f' : D \to B(\mathbb{X}, \mathbb{Y})$ is continuous at x_0 if for all $\epsilon > 0$ there exists $\delta(x_0, \epsilon) > 0$ such that $\|f'(x) - f'(x_0)\|_B < \epsilon$ whenever $\|x - x_0\|_{\mathbb{X}} < \delta(x_0, \epsilon)$. If f' is continuous on an open ball $B \subset \mathbb{X}$ we say that f is continuously differentiable on B. Note that the norm $\| \cdot \|_B$ is a norm on $B(\mathbb{X}, \mathbb{Y})$.*

   For example, suppose $f : D \subset (\mathbb{X}, \| \cdot \|_{\mathbb{X}}) \to (\mathbb{R}, | \cdot |)$. The space of all bounded linear functions from $\mathbb{X}$ to $\mathbb{R}$ ($B(\mathbb{X}, \mathbb{R})$) is called the normed dual of $\mathbb{X}$ and denoted by $\mathbb{X}^*$. Therefore, $f' \in \mathbb{X}^*$. If $\mathbb{X} = \mathbb{R}^n$, $f'(x) = A(x)$ a $1 \times n$ vector.

6. *Note that here, continuity follows from differentiability since if*
$$f(x+h) = f(x) + A(x)h + E(x,h) \text{ and } \lim_{h \to 0} f(x+h) = f(x)$$

since $A(x)h \to 0$ as $h \to 0$ and $E(x,h) \to 0$, by differentiability.

Theorem 93. *If a function f has a Fréchet differential, it is unique.*

Proof. Suppose that $\delta f(x,h)$ and $\delta' f(x,h)$ satisfy the requirements of Definition 62. Then,

$$\begin{aligned}\|\delta f(x,h) - \delta' f(x,h)\| &= \|f(x+h) - f(x) - \delta' f(x,h) - (f(x+h) - f(x) - \delta f(x,h))\| \\ &\leq \|f(x+h) - f(x) - \delta' f(x,h)\| + \|f(x+h) - f(x) - \delta f(x,h)\| \\ &= o(\|h\|).\end{aligned}$$

Since $\delta f(x,h) - \delta' f(x,h)$ is bounded and linear in h, its norm must be zero. □

Theorem 94. *If the Fréchet differential of f exists at x, then the Gâteaux differential exists at x and they are equal.*

Proof. Let $\delta f(x,h)$ be the Fréchet differential. By definition, for any h

$$\lim_{a \to 0} \frac{\|f(x+ah) - f(x) - \delta f(x,ah)\|}{\|ah\|} = 0.$$

Since $\delta f(x,ah)$ is linear we have $\delta f(x,ah) = a\delta f(x,h)$. Hence,

$$\begin{aligned}\lim_{a \to 0} \frac{\|f(x+ah) - f(x) - \delta f(x,ah)\|}{\|ah\|} &= \lim_{a \to 0} \frac{1}{|a|} \frac{\|a\frac{f(x+ah)-f(x)}{a} - a\delta f(x,h)\|}{\|h\|} \\ &= \lim_{a \to 0} \frac{\|\frac{f(x+ah)-f(x)}{a} - \delta f(x,h)\|}{\|h\|}\end{aligned}$$

and the last equality implies that

$$\lim_{a \to 0} \|\frac{f(x+ah) - f(x)}{a} - \delta f(x,h)\| = 0 \text{ or } \lim_{a \to 0} \frac{f(x+ah) - f(x)}{a} = \delta f(x,h) = \delta_g f(x,h).$$

□

Theorem 95. *If $f : D \subseteq \mathbb{X} \to \mathbb{Y}$, where D is an open set has a Fréchet differential at x, then f is continuous at x.*

Proof. Given any $\epsilon > 0$, by definition of a limit, there exists a ball centered at x such that for $x + h$ in the ball we have

$$\|f(x+h) - f(x) - \delta f(x,h))\| < \|h\|\epsilon. \tag{21}$$

Consequently,

$$\begin{aligned}\|f(x+h) - f(x)\| &= \|f(x+h) - f(x) - \delta f(x,h) + \delta f(x,h)\| \\ &\leq \|f(x+h) - f(x) - \delta f(x,h)\| + \|\delta f(x,h)\| \\ &\leq \|h\|\epsilon + C\|h\| = M\|h\|,\end{aligned}$$

where the last inequality follows from equation (21) and the fact that $\delta f(x,h)$ is a bounded linear function of h. □

Theorem 96. *Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, where D is an open set and*

$$f(x)^T = \begin{pmatrix} f_1(x) & f_2(x) & \cdots & f_m(x) \end{pmatrix}$$

with $f_i : D \subseteq \mathbb{R}^n \to \mathbb{R}$. Then,

1. *f is Fréchet differentiable at x if, and only if, f_i is Fréchet differentiable at x for all $i = 1, \cdots, m$.*

2. *f Fréchet differentiable at x implies that*

$$A(x) = \begin{pmatrix} \nabla f_1(x) \\ \nabla f_2(x) \\ \vdots \\ \nabla f_m(x) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial_1}(x) & \frac{\partial f_1}{\partial_2}(x) & \cdots & \frac{\partial f_1}{\partial_n}(x) \\ \frac{\partial f_2}{\partial_1}(x) & \frac{\partial f_2}{\partial_2}(x) & \cdots & \frac{\partial f_2}{\partial_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial_1}(x) & \frac{\partial f_m}{\partial_2}(x) & \cdots & \frac{\partial f_m}{\partial_n}(x) \end{pmatrix}.$$

Proof. 1. ($\Rightarrow$) f Fréchet differentiable at x implies there exists $A(x)$ such that

$$\lim_{\|h\|_E \to 0} \frac{\|f(x+h) - f(x) - A(x)h\|_E}{\|h\|_E} = 0.$$

If $A(x) = [A_{ik}(x)]_{i,k=1}^{m,n}$ and $h^T = \begin{pmatrix} h_1 & h_2 & \cdots & h_n \end{pmatrix}$, then

$$A(x)h = \begin{pmatrix} \sum_{k=1}^{n} A_{1k}(x)h_k \\ \sum_{k=1}^{n} A_{2k}(x)h_k \\ \vdots \\ \sum_{k=1}^{n} A_{mk}(x)h_k \end{pmatrix}$$

and the i^{th} component of $f(x+h)-f(x)-A(x)h$ is given by $f_i(x+h)-f_i(x)-\sum_{k=1}^{n} A_{ik}(x)h_k$. Since for any vector $a \in \mathbb{R}^n$ we have $|a_i| \leq \|a\|_E$, we have that

$$|f_i(x+h) - f_i(x) - \sum_{k=1}^{n} A_{ik}(x)h_k| \leq \|f(x+h) - f(x) - A(x)h\|.$$

Hence,

$$\lim_{\|h\|_E \to 0} \frac{|f_i(x+h) - f_i(x) - \sum_{k=1}^{n} A_{ik}(x)h_k|}{\|h\|_E} = 0$$

which establishes the Fréchet differentiability of f_i at x.

($\Leftarrow$) Exercise.

2. Now, let $h \to \theta$ along the s^{th} coordinate, i.e., put $h_k = 0$ for all $k \neq s$. Then, the last limit implies

$$\lim_{h_s \to 0} \frac{f_i\left(x_1, \cdots, x_s + h_s, \cdots, x_n\right) - f_i(x) - A_{is}(x)h_s}{h_s} = 0,$$

which is equivalent to

$$\lim_{h_s \to 0} \frac{f_i\left(x_1, \cdots, x_s + h_s, \cdots, x_n\right) - f_i(x)}{h_s} = A_{is}(x) = \frac{\partial}{\partial_s} f_i(x).$$

$\square$

Theorem 97. *Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, where D is an open set and*

$$f(x)^T = \begin{pmatrix} f_1(x) & f_2(x) & \cdots & f_m(x) \end{pmatrix}$$

with $f_j : D \subseteq \mathbb{R}^n \to \mathbb{R}$. If the partial derivatives $\frac{\partial}{\partial_s} f_i(x)$ exists for $i = 1, \cdots, m$ and $s = 1, \cdots, n$ and are continuous on D, f is Fréchet differentiable on D.

Proof. In part 1 of Theorem 96 we showed that f is Fréchet differentiable at x if, and only if, all component functions f_i are Fréchet differentiable. So, it suffices to prove that f_i is Fréchet differentiable on D.

Since D is open, for fixed x there exists an $r > 0$ such that $B(x, r) \subset D$ and by continuity of the partial derivatives

$$|\frac{\partial}{\partial_j} f_i(x+h) - \frac{\partial}{\partial_s} f_i(x)| < \epsilon$$

whenever $x + h \in B(x, r)$ or $\|x + h - x\| = \|h\| < r$. If f_i is Fréchet differentiable at x its derivative will be the gradient vector. So, we would like to show that

$$\lim_{\|h\| \to 0} \frac{|f_i\,(x+h) - f_i(x) - \nabla f_i(x)h|}{\|h\|} = 0.$$

Since $h \in \mathbb{R}^n$ we can write $h = \sum_{i=1}^{n} e_i h_i$ where e_i is the unit coordinate vector. Now, let $v_0 = (0, \cdots, 0)'$ and $v_j = (h_1, \cdots, h_j, 0 \cdots 0) = \sum_{k=1}^{j} e_k h_k$ for $j = 1, \cdots n$. Then,

$$f_i(x+h) - f_i(x) = \sum_{j=1}^{n} (f_i\,(x+v_j) - f_i\,(x+v_{j-1})).$$

Because $\|v_j\| \leq \|h\| < r$ the points $x + v_j$ and $x + v_{j-1}$ are inside $B(x, r)$. Also, since $B(x, r)$ is a convex set, any convex combination of these two points is in $B(x, r)$. But since the partial derivatives of f_i exist and are continuous in D, they exist and are continuous in $B(x, r)$ and on the convex combination of $x + v_j$ and $x + v_{j-1}$. Hence, for $\lambda_j \in (0, 1)$ by the Mean Value Theorem

$$f_i\,(x+v_j) - f_i\,(x+v_{j-1}) = \frac{\partial}{\partial_j} f_i(x + v_{j-1} + \lambda_j h_j v_j) h_j.$$

Hence,

$$\begin{aligned}
|f_i\,(x+h) - f_i(x) - \nabla f_i(x)h| &= \left|\sum_{j=1}^{n}(f_i\,(x+v_j) - f_i\,(x+v_{j-1})) - \nabla f_i(x)h\right| \\
&= \left|\sum_{j=1}^{n}(f_i\,(x+v_j) - f_i\,(x+v_{j-1})) - \sum_{j=1}^{n}\frac{\partial f_i}{\partial_j}(x)h_j\right| \\
&= \left|\sum_{j=1}^{n}\frac{\partial}{\partial_j}f_i(x+v_{j-1}+\lambda_j h_j v_j)h_j - \sum_{j=1}^{n}\frac{\partial f_i}{\partial_j}(x)h_j\right| \\
&\leq \sum_{j=1}^{n}|h_j|\left|\frac{\partial}{\partial_s}f_i(x+v_{j-1}+\lambda_j h_j v_j) - \frac{\partial}{\partial_j}f_i(x)\right| \\
&< \sum_{j=1}^{n}|h_j|\epsilon \leq n\|h\|\epsilon \text{ by continuity.}
\end{aligned}$$

Dividing both sides by $\|h\|$ gives the result. □

Theorem 98. *Let $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{Y} \subseteq \mathbb{R}^m \to \mathbb{R}^p$, where $\mathbb{X}$ and $\mathbb{Y}$ are open sets with $f(\mathbb{X}) \subseteq \mathbb{Y}$. Let $x \in \mathbb{X}$ and put $y = f(x)$. Define $F(x) = (g \circ f)(x) = g(f(x))$ for all $x \in \mathbb{X}$. If f is Fréchet differentiable at x and g is Fréchet differentiable at y, then F is Fréchet differentiable at x. Furthermore,*

$$A_F(x) := A_g(f(x))A_f(x).$$

Proof. First, write for arbitrary $h \in \mathbb{R}^n$ such that $x + h \in \mathbb{X}$

$$\begin{aligned}
F(x+h) - F(x) &= g(f(x+h)) - g(f(x)) \\
&= g(f(x+h) - f(x) + f(x)) - g(f(x)) = g(f(x+h) - f(x) + y) - g(y) \\
&= g(m+y) - g(y) \text{ where } m = f(x+h) - f(x).
\end{aligned}$$

By Fréchet differentiability of g at y we have $\frac{1}{\|m\|}\|g(m+y) - g(y) - A(y)m\| \to 0$ as $\|m\| \to 0$. Now, $m = f(x+h) - f(x)$ and by differentiability of f at x we have $\|f(x+h) - f(x) -$

$B(x)h\| \to 0$ as $\|h\| \to 0$. Note that,

$$\begin{aligned} F(x+h) - F(x) - A(y)B(x)h &= g(m+y) - g(y) - A(y)B(x)h - A(y)m + A(y)m \\ &= g(m+y) - g(y) - A(y)m + A(y)(m - B(x)h), \text{ hence} \\ \|F(x+h) - F(x) - A(y)B(x)h\| &\leq \|g(m+y) - g(y) - A(y)m\| + \|A(y)\|\|m - B(x)h\| \\ &= o(\|m\|) + o(\|h\|) \end{aligned}$$

where the orders after the equality follow from the differentiability of g and f and from the fact that $\|A(y)\|$ is bounded. Dividing both sides of the last inequality by $\|h\|$, we have

$$\frac{1}{\|h\|}\|F(x+h) - F(x) - A(y)B(x)h\| \leq \frac{\|m\|}{\|h\|}o(1) + o(1).$$

Lastly, we need to show that $\frac{\|m\|}{\|h\|} = \frac{\|f(x+h)-f(x)\|}{\|h\|}$ is bounded. But this follows from the fact that f is differentiable at x and therefore continuous at x by Theorem 95. □

Remark 28. *Theorem 98 is valid when the domain and co-domain are arbitrary normed vector spaces. In this case we write, where applicable, $A_y(m)$ and $B_x(h)$.*

Theorem 99. *Let $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be Fréchet differentiable on the open set $\mathbb{X}$. Let $x, y \in \mathbb{X}$ be such that $L(x,y) = \{z \in \mathbb{R}^n : z = \lambda y + (1-\lambda)x,\ \lambda \in (0,1)\} \subseteq \mathbb{X}$. Then, for all $a \in \mathbb{R}^m$ there exists $z \in L(x,y)$ such that $a^T(f(y) - f(x)) = a^T(Df(z)(y-x))$, where $Df(z) = A(z)$ as in Theorem 96. (Recall that $\mathbb{X}$ is a convex set if $L(x,y) \subseteq \mathbb{X}$).*

Proof. Consider $z = \lambda y + (1-\lambda)x = x + \lambda(y-x) = x + \lambda h$ where $h = y - x$. By assumption $z \in \mathbb{X}$ and since $\mathbb{X}$ is open, there exists $\delta > 0$ such that $z \in \mathbb{X}$ for $\lambda \in (-\delta, 1+\delta)$. Fix $a \in \mathbb{R}^m$ and define $\phi_a : (-\delta, 1+\delta) \to \mathbb{R}$ as

$$\phi_a(\lambda) = a^T f(x + \lambda h) = \sum_{i=1}^{m} a_i f_i(x + \lambda h).$$

By construction $\phi_a(\lambda)$ is differentiable on $(-\delta, 1+\delta)$ and by theorem 96

$$\phi_a^{(1)}(\lambda) = \sum_{i=1}^{m} a_i \left(\begin{pmatrix} \frac{\partial}{\partial_1} f_i(x+\lambda h) & \frac{\partial}{\partial_2} f_i(x+\lambda h) & \cdots & \frac{\partial}{\partial_n} f_i(x+\lambda h) \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} \right)$$

$$= \sum_{i=1}^{m} a_i \sum_{k=1}^{n} \frac{\partial}{\partial_k} f_i(x+\lambda h) h_k = \sum_{i=1}^{m} a_i \nabla f_i(x+\lambda h) h = a^T Df(x+\lambda h)h.$$

By the Mean Value Theorem (applied to ϕ_a), for $\lambda \in (0,1)$

$$a^T(f(y)-f(x)) = \phi_a(1) - \phi_a(0) = \phi_a^{(1)}(\lambda)(1-0) = \phi_a^{(1)}(\lambda) = a^T Df(x+\lambda h)h.$$

Putting $z = x + \lambda h$, then $a^T(f(y)-f(x)) = a^T Df(z)(y-x)$. □

Theorem 100. *Let $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be differentiable on the open set $\mathbb{X}$. Let $x, y \in \mathbb{X}$ be such that $L(x,y) = \{z \in \mathbb{R}^n : z = \lambda x + (1-\lambda)y,\ \lambda \in (0,1)\} \subseteq \mathbb{X}$. Then, there exists $z \in L(x,y)$ such that $\|f(y)-f(x)\| \leq \|Df(z)(y-x)\| \leq \|Df(z)\|\|y-x\|$.*

Proof. By the Mean Value Theorem, for all $a \in \mathbb{R}^m$ there exists $z \in L(x,y)$ such that

$$a^T(f(y)-f(x)) = a^T Df(z)(y-x).$$

Now,

$|a^T(f(y)-f(x))| = |a^T(Df(z)(y-x))| \leq \|a\|\|Df(z)(y-x)\|$ by the Cauchy-Schwarz inequality.

Take $a = \frac{1}{\|f(y)-f(x)\|}(f(y)-f(x))$ and note that $\|a\| = 1$ and $a^T(f(y)-f(x)) = \frac{\|f(y)-f(x)\|^2}{\|f(y)-f(x)\|} = \|f(y)-f(x)\|$. Then,

$$\|f(y)-f(x)\| \leq \|Df(z)(y-x)\| \leq \|Df(z)\|\|y-x\|.$$

□

Definition 63. *The space of all bounded linear functions $f : (\mathbb{X}, \|\cdot\|_{\mathbb{X}}) \to (\mathbb{R}, |\cdot|)$ is called the normed dual of $\mathbb{X}$ and denoted by $\mathbb{X}^*$. The norm of $\mathbb{X}^*$ is taken to be $\|f\| = \sup\limits_{\|x\|_{\mathbb{X}}=1} |f|$.*

Definition 64. *$f \in \mathbb{X}^*$ is said to be aligned with $x \in \mathbb{X}$ if $f(x) = \|f\|\|x\|_{\mathbb{X}}$.*

We know that by definition it is always the case that $|f(x)| \leq \|f\|\|x\|_{\mathbb{X}}$.

Theorem 101. *Let $f : (\mathbb{X}, \|\cdot\|_{\mathbb{X}}) \to (\mathbb{Y}, \|\cdot\|_{\mathbb{Y}})$ be Fréchet differentiable at on an open $D \subset \mathbb{X}$. Let $x \in D$ and suppose that $x + ah \in D$ for all $a \in [0,1]$. Then,*

$$\|f(x+h) - f(x)\| \leq \|h\| \sup_{0<a<1} \|A_f(x+ah)\|.$$

Proof. Let $g : (\mathbb{Y}, \|\cdot\|_{\mathbb{Y}}) \to (\mathbb{R}, |\cdot|)$ be linear and bounded with $g \neq 0$ and g aligned with $f(x+h) - f(x) \in \mathbb{Y}$, so that

$$g(f(x+h) - f(x)) = \|g\|\|f(x+h) - f(x)\|.$$

Let $\phi(a) = g(f(x+ah))$ be a function defined on $[0,1]$ and note that by the chain rule

$$\phi^{(1)}(a) = g(A_f(x+ah)h)$$

since because g is linear $\delta g(y,h) = g(h)$ for all y. By the Mean Value Theorem for function in $\mathbb{R}$

$$\phi(1) - \phi(0) = \phi^{(1)}(a_0), \text{ for } a_0 \in (0,1).$$

Hence, by linearity of g

$$|g(f(x+h) - f(x))| = |\phi(1) - \phi(0)|.$$

Thus,

$$|g(f(x+h) - f(x))| = |g(A_f(x+ah)h)|(1-0) \leq \|g\| \sup_{0<a<1} \|A_f(x+ah)\|\|h\|.$$

But since g is aligned with $f(x+h) - f(x)$,

$$\|f(x+h) - f(x)\| \leq \|h\| \sup_{0<a<1} \|A_f(x+ah)\|.$$

□

Definition 65. *$f : D \subseteq (\mathbb{X}, \|\cdot\|_{\mathbb{X}}) \to (\mathbb{Y}, \|\cdot\|_{\mathbb{Y}})$ is continuously Fréchet differentiable on the open set D if, it is Fréchet differentiable on D and, for all $x \in D$ and for all $\epsilon > 0$ there exists some $\delta > 0$ such that its derivative A satisfies*

$$\|A_y - A_x\| < \epsilon \textit{ whenever } \|y - x\|_{\mathbb{X}} < \delta_x.$$

Note that the norm on the left-hand side of the previous inequality is a norm on a space of linear functions.

Definition 66. *$f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ defined on an open set D belongs to the class C^k in D for $k \in \mathbb{N}$ and we write $f \in C^k(D)$, if the first k partial derivatives of the component functions of f exist and are continuous on $\mathbb{X}$.*

Theorem 102. *Let $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be defined on the open set $\mathbb{X}$. f is continuously differentiable on $\mathbb{X}$ if, and only if, $f \in C^1(\mathbb{X})$.*

Proof. ($\Rightarrow$) f continuously differentiable on $\mathbb{X}$ implies f is differentiable on $\mathbb{X}$. Let $x, x+h \in \mathbb{X}$, then there exists $A(x)$ and $A(x+h)$ associated with linear functions in $L(\mathbb{X}, \mathbb{R}^m)$ (space of linear functions from $\mathbb{X}$ to $\mathbb{R}^m$). Also, their difference $A(x+h) - A(x)$ is a linear function in $L(\mathbb{X}, \mathbb{R}^m)$ and has matrix A with typical element $A_{ik}(h)$ where $A_{ik} = \frac{\partial}{\partial_k} f_i(x+h) - \frac{\partial}{\partial_k} f_i(x)$ by Theorem 96 (item 2). For fixed x, define $\kappa(h) = \max_{i,k} |A_{ik}(h)|$ and note that from Theorem 50

$$0 \le \kappa(h) \le \|A(x+h) - A(x)\| \le \kappa(h)\sqrt{mn}$$

Hence, $\kappa(h) \to 0$ as $h \to \theta$ if, and only if, $\|A(x+h) - A(x)\| \to 0$ as $\|h\| \to \theta$ which shows that all partial derivatives are continuous.

($\Leftarrow$) $f \in C^1$ implies that all partial derivatives exist and are continuous, which implies that f is differentiable (Theorem 97). Now, the continuity of the partial derivatives implies that for all i, k, $A_{ik}(h) \to 0$ as $h \to \theta$. But $\kappa(h) \to 0$ as $h \to \theta$ implying that $\|A(x+h) - A(x)\| \to 0$ and hence we have continuity of the Fréchet derivative. □

Theorem 103. *Let $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}$ be defined on the open and convex set $\mathbb{X}$ with $f \in C^2(\mathbb{X})$. If $x, x+h \in \mathbb{X}$ then,*

$$f(x+h) = f(x) + Df(x)h + \frac{1}{2}h^T D^2 f(x+\lambda h)h$$

for some $\lambda \in (0,1)$.

Proof. Left as an exercise. $\square$

17 The Inverse Function Theorem

When we studied linear transformations we learned that if $Ax = y$ and if A is invertible we can solve the equation for x. Put differently, if the linear function T associated with A is one-to-one and onto then we can find a unique solution for the equation. In this case, $x = A^{-1}y$. If we are dealing with a nonlinear function $f : \mathbb{R}^n \to \mathbb{R}^n$, when does $f(x) = y$ have a solution in terms of y?

Definition 67. *Let $(\mathbb{X}, d_{\mathbb{X}})$ be a metric space and $T : (\mathbb{X}, d_{\mathbb{X}}) \to (\mathbb{X}, d_{\mathbb{X}})$ be an operator. We say that T is a contraction of modulus β if for some $\beta \in (0,1)$ we have that for all $x, y \in \mathbb{X}$, $d(T(x), T(y)) \leq \beta d(x,y)$.*

Remark 29. *A function that satisfies this definition is, of course, Lipschitz with a constant restricted to be in $(0,1)$.*

Theorem 104. *Let $(\mathbb{X}, d_{\mathbb{X}})$ be a Banach space and T be a contraction of modulus β. Then,*

1. *(Fixed Point) There exists only one $x^* \in \mathbb{X}$ such that $T(x^*) = x^*$,*

2. *The sequence $\{x_1 = T(x_0), x_2 = T(x_1), \cdots\}$ is such that $\{x_i\}_{i=1,2,\cdots} \to x^*$ for any starting point $x_0 \in \mathbb{X}$.*

Proof. First we prove 2.

$$d(x_{n+1}, x_n) = d(T(x_n), T(x_{n-1})) \leq \beta d(x_n, x_{n-1}) = \beta d(T(x_{n-1}), T(x_{n-2}))$$
$$\leq \beta^2 d(x_{n-1}, x_{n-2}) \leq \cdots \leq \beta^n d(x_1, x_0) \tag{22}$$

Also, for $n > m$

$$\begin{aligned} d(x_n, x_m) &\leq \sum_{i=m}^{n-1} d(x_{i+1}, x_i) \text{ by the triangle inequality} \\ &\leq \sum_{i=m}^{n-1} \beta^i d(x_1, x_0) = \beta^m d(x_1, x_0) \sum_{i=0}^{n-m-1} \beta^i \text{ by equation (22).} \\ &\leq \beta^m d(x_1, x_0) \sum_{i=0}^{\infty} \beta^i = \frac{\beta^m}{1-\beta} d(x_1, x_0). \end{aligned}$$

Where the last inequality follows from $\beta < 1$. Also, $\frac{\beta^m}{1-\beta} \to 0$ as $m \to \infty$. Hence, for all $\epsilon > 0$ we can choose m, n sufficiently large such that $d(x_m, x_n) < \epsilon$. Hence, $T(x_0), T(x_1), \cdots$ is a Cauchy sequence for any x_0. Since $\mathbb{X}$ is a Banach space, for any x_0 we have $T(x_0), T(x_1), \cdots$ converges in $\mathbb{X}$. Take one such limit and call it x^*. Since T is a contraction, T is continuous (why?). Hence,

$$T(x^*) = T(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} T(x_n) = \lim_{n\to\infty} x_{n+1} = x^*$$

1. Now, suppose x', x'' are such that $T(x') = x'$ and $T(x'') = x''$. Since T is a contraction, we have for some $\beta \in (0, 1)$

$$d(T(x'), T(x'')) = d(x', x'') \leq d(x', x'')\beta.$$

But since $\beta \in (0, 1)$ this can hold only if $d(x', x'') = 0$, otherwise $d(x', x'') < d(x', x'')$, clearly a contradiction. So, there must be only one x such that $x = T(x)$. □

Theorem 105. *(Inverse Function Theorem) Let $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ where $\mathbb{X}$ is open and f be continuously differentiable. Let $x_0 \in \mathbb{X}$ and suppose that the Fréchet derivative $A(x_0)$ is invertible. Then, there exists an open neighborhood $\mathcal{N}$ of x_0 such that,*

1. *f is one-to-one on $\mathcal{N}$. Hence, the inverse f^{-1} exists as a function from $\mathcal{V} = f(\mathcal{N})$ to $\mathcal{N}$,*

2. *$\mathcal{V} = f(\mathcal{N})$ is an open set and $y_0 = f(x_0) \in \mathcal{V}$,*

3. *$f^{-1} \in C^1$ and $Df^{-1}(y_0) = (Df(x_0))^{-1}$,*

4. *if $f \in C^k$, $k > 1$ then $f^{-1} \in C^k$.*

Proof. 1. Since $A(x_0)$ is invertible, there exist a matrix $A(x_0)^{-1}$. Put

$$\lambda = \frac{1}{2\|A(x_0)^{-1}\|} \tag{23}$$

where $\|A(x_0)^{-1}\|$ is the norm of $A(x_0)^{-1}$. Or, if you prefer, the norm of the linear operator associated with $A(x_0)^{-1}$.

$\mathbb{X}$ open implies that there exists $\delta > 0$ such that $B(x_0, \delta) \equiv \mathcal{N} \subset \mathbb{X}$. f continuously differentiable implies that for all $\lambda > 0$, there exists $\delta > 0$ such that

$$\|A(x) - A(x_0)\| < \lambda \text{ whenever } x \in B(x_0, \delta) \equiv \mathcal{N}. \tag{24}$$

Now, for all $y \in \mathbb{R}^n$ let

$$\phi_y(x) = x + A(x_0)^{-1}(y - f(x)) \tag{25}$$

and note that $y = f(x)$ if, and only if, $\phi_y(x) = x$.

To show that f is one-to-one on $\mathcal{N}$ we must show that there exits at most one $x \in \mathcal{N}$ for which $y = f(x)$, or equivalently, there exists at most one x for which $\phi_y(x) = x$ for each y. The Fréchet derivative of ϕ_y at x, denoted by $\phi'_y(x)$ is given by (see part 2 of Theorem 96)

$$\phi'_y(x) = I_n + A(x_0)^{-1}(-A(x)) = A(x_0)^{-1}(A(x_0) - A(x)).$$

Then, by (23)

$$\|\phi'_y(x)\| = \|A(x_0)^{-1}(A(x_0) - A(x))\| \leq \|A(x_0)^{-1}\|\|A(x_0) - A(x)\| < \frac{1}{2\lambda}\lambda = \frac{1}{2}$$

for $x \in \mathcal{N}$. From Theorem 100 we have for all $x, x' \in \mathcal{N}$

$$\|\phi_y(x') - \phi_y(x)\| \leq \|\phi_y'(z)\|\|x' - x\| = \frac{1}{2}\|x' - x\|. \tag{26}$$

This establishes that ϕ_y is a contraction modulus $1/2 < 1$. Now, suppose $\phi_y(x) = x$ **and** $\phi_y(x') = x'$, then equation (26) cannot hold, so it must be that that there exists at most one x for which $\phi_y(x) = x$. Thus, f is one-to-one and onto on $\mathcal{N}$.

2. Let $\mathcal{V} = f(\mathcal{N})$ and take an arbitrary $y' \in \mathcal{V}$, then $f(x') = y'$ for some $x' \in \mathcal{N}$. To show that $\mathcal{V}$ is open we show that $\|y - y'\| < \lambda r$ implies $y \in \mathcal{V}$. That is, there exists a ball centered at y' that is entirely in $\mathcal{V}$. From equation (23) we have that $2\|A(x_0)^{-1}\|\lambda = 1$ and if $\bar{B}(x', r)$ is a closed ball such that $\bar{B}(x', r) \subset \mathcal{N}$,

$$\|\phi_y(x') - x'\| = \|A(x_0)^{-1}(y - y')\| \leq \|A(x_0)^{-1}\|\|y - y'\| < \|A(x_0)^{-1}\|\lambda r = r/2.$$

For $x \in \bar{B}(x', r)$ we have

$$\begin{aligned}
\|\phi_y(x) - x'\| &= \|\phi_y(x) - \phi_y(x') + \phi_y(x') - x'\| \\
&\leq \|\phi_y(x) - \phi_y(x')\| + \|\phi_y(x') - x'\| \\
&\leq \frac{1}{2}\|x - x'\| + \frac{r}{2} \leq r \text{ since } x \in B(x', r).
\end{aligned}$$

Hence, $\phi_y(x) \in \bar{B}(x', r)$ whenever x is close to x' and y close to y'. From equation (26) we see that for suitably chosen y, $\phi_y(x)$ is a contraction modulus $1/2$ on $\bar{B}(x', r)$. Since $\bar{B}$ is closed in $\mathbb{R}^n$ it is complete and by the contraction mapping theorem ϕ_y is such that $\phi_y(x^*) = x^*$ for $x^* \in \bar{B}$. For this $x^* \in \bar{B} \subset \mathcal{N}$, we have $f(x^*) = y$ implying that $y \in f(\bar{B}) \subseteq f(\mathcal{N}) = \mathcal{V}$, whenever $\|y - y'\| < \lambda r$. That is, for all $y' \in \mathcal{V}$ there exists $B(y, \lambda r) \subset \mathcal{V}$, so $\mathcal{V}$ is open.

3. Let $y, y + k \in \mathcal{V} = f(\mathcal{N})$. Then, there exists $x, x + h \in \mathcal{N}$ such that $y = f(x)$ and $y + k = f(x + h)$. Again, consider $\phi_z(x) = x + A(x_0)^{-1}(z - f(x))$. Then,

$$\begin{aligned}
\phi_z(x + h) - \phi_z(x) &= x + h + A(x_0)^{-1}(z - f(x + h)) - x - A(x_0)^{-1}(z - f(x)) \\
&= h + A(x_0)^{-1}(f(x) - f(x + h)) = h - A(x_0)^{-1}k, \text{ and consequently,}
\end{aligned}$$

$$\|\phi_z(x+h) - \phi_z(x)\| = \|h - A(x_0)^{-1}k\| \le \frac{1}{2}\|h\|$$

by equation (26) and the fact that $x, x+h \in \mathcal{N}$. Now, $\|h - A^{-1}(x_0)k\| \ge \|h\| - \|A^{-1}(x_0)k\|$, therefore $\|h\| - \|A^{-1}(x_0)k\| \le \frac{1}{2}\|h\|$ which implies that

$$\|A^{-1}(x_0)k\| \ge \|h\| - \frac{1}{2}\|h\| = \frac{1}{2}\|h\|.$$

Using equation (23)

$$\|h\| \le 2\|A^{-1}(x_0)k\| \le 2\|A^{-1}(x_0)\|\|k\| = \frac{1}{\lambda}\|k\|.$$

So, whenever $\|k\| = \|f(x) - f(x+k)\| < \delta$ we have that $\|h\| = \|x+h-x\| < \delta/\lambda$, establishing continuity of f^{-1}. Now, observe that equations (23) and (24) imply that for all $x \in \mathcal{N}$

$$\|A(x) - A(x_0)\| < \frac{1}{2\|A(x_0)^{-1}\|}.$$

By Theorem 53, $A(x)$ is invertible for all $x \in \mathcal{N}$. Now, suppose that the Fréchet differential of f^{-1} is $\delta f^{-1}(y,k) = A(x)^{-1}k$, then

$$\begin{aligned} f^{-1}(y+k) - f^{-1}(y) - A(x)^{-1}k &= x + h - x - A(x)^{-1}k = Ih - A(x)^{-1}k \\ &= A(x)^{-1}A(x)h - A(x)^{-1}(y+k-y) \\ &= -A(x)^{-1}(f(x+h) - f(x) - A(x)h) \end{aligned}$$

and

$$0 \le \frac{\|f^{-1}(y+k) - f^{-1}(y) - A(x)^{-1}k\|}{\|k\|} \le \frac{\|A(x)^{-1}\|\|f(x+h) - f(x) - A(x)h\|}{\lambda\|h\|}.$$

Now, as $\|k\| \to 0$, $\|h\| \to 0$ and since f is differentiable at x

$$\frac{\|f(x+h) - f(x) - A(x)h\|}{\|h\|} \to 0 \text{ as } \|h\| \to 0.$$

Hence, f^{-1} is differentiable at y, with $Df^{-1}(y) = A(x)^{-1} = (Df(f^{-1}(y)))^{-1}$.

To finish the proof of 3, it suffices to show that $Df^{-1}(y)$ is continuous. From the last equality, we see that $Df^{-1}(y)$ is the composition of three functions: f^{-1}, which is continuous, since it is differentiable; Df which is continuous by assumption; and the inverse $(\cdot)^{-1}$ which is continuous by Theorem 54. Hence, $Df(y)^{-1}$ is continuous.

Finally, item 4, follows directly from 3. $\square$

Definition 68. *(Homogeneity) $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}$ is homogeneous of degree k in $\mathbb{X}$ if $f(\lambda x) = \lambda^k f(x)$ for all $\lambda > 0$.*

Note that λx must be in $\mathbb{X}$, that is, $\mathbb{X}$ is what we call a cone.

Theorem 106. *(Euler's Theorem) Let $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}$ with continuous partial derivatives on an open cone $\mathbb{X}$. f is homogeneous of degree k in $\mathbb{X}$ if, and only if, $\sum_{i=1}^n \frac{\partial}{\partial_i} f(x) x_i = kf(x)$ for all $x \in \mathbb{X}$.*

Proof. Exercise. $\square$

18 Static models and comparative statics

Many economic models can be reduced to a system of equations

$$F(x, a) = \theta \tag{27}$$

where $F : \mathbb{X} \times A \subseteq \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^m$ and θ is the null vector. a is normally interpreted as a set of "parameters" describing the "environment" or the state of the world in which the economic model is embedded and x are "endogenous" variables. We normally seek values for x that satisfy equation (27), that is,

$$S(a) = \{x \in \mathbb{X} : F(x, a) = \theta\}.$$

In this context, here are some important questions: a) For given a, do solutions for (27) exist? That is, is $S(a) \neq \emptyset$?; b) If so, how many elements are in $S(a)$?; c) How does $S(a)$ change with a? (comparative statics).

Example 27. *Suppose we have two markets, one for good* 1 *and another, for good* 2*. The demand and supply equations are given by*

$$\begin{aligned} Q_1^D &= a_0 + a_1 p_1 + a_2 p_2 \\ Q_1^S &= b_0 + b_1 p_1 + b_2 p_2 \\ Q_2^D &= \alpha_0 + \alpha_1 p_1 + \alpha_2 p_2 \\ Q_2^S &= \beta_0 + \beta_1 p_1 + \beta_2 p_2 \end{aligned}$$

The equilibrium conditions are $Q_1^D - Q_1^S = 0$ *and* $Q_2^D - Q_2^S = 0$*. Or, alternatively,*

$$\begin{aligned} (a_0 - b_0) + (a_1 - b_1)p_1 + (a_2 - b_2)p_2 &= 0 \\ (\alpha_0 - \beta_0) + (\alpha_1 - \beta_1)p_1 + (\alpha_2 - \beta_2)p_2 &= 0. \end{aligned}$$

Setting $a_i - b_i = c_i$ *and* $\alpha_i - \beta_i = \gamma_i$ *for* $i = 0, 1, 2$*, we have*

$$\begin{pmatrix} c_1 & c_2 \\ \gamma_1 & \gamma_2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + \begin{pmatrix} c_0 \\ \gamma_0 \end{pmatrix} = \theta$$

which gives $F(p_1, p_2; c_0, \gamma_0, c_1, c_2, \gamma_1, \gamma_2) = \theta$*. In the notation of (27),* $m = n = 2$ *and* $p = 12$*.*

18.1 Linear models

If F is linear, then we can write

$$Ax + Ba = \theta \tag{28}$$

or $Ax = -Ba = y$. We interpret $Ax = y$ as a system of m equations with n unknowns. We start with the case where $y = \theta$, that is, a homogeneous system. Let T be the the linear function associated with A. From Theorem 41

$$dim(\mathbb{X}) = dim(null(T)) + dim(im(T)) = dim(null(T)) + rank(T).$$

Recall that $null(T) = \{x \in \mathbb{X} : T(x) = \theta\}$ is the set of solutions and $im(T) = \{y \in \mathbb{R}^m : y = T(x) \text{ for } x \in \mathbb{R}^n\}$. Consequently,

$$dim(null(T)) = dim(\mathbb{X}) - rank(T) \geq n - m \text{ since } rank(T) \leq m.$$

The last inequality follows from the fact that the image of T is a subspace of $\mathbb{R}^m$ and therefore cannot have dimension greater than m. Hence, if $n > m$ (more unknowns than equations) we have that $dim(null(T)) \geq 1$ and $T(x) = \theta$ has nontrivial solutions. Now, for non-homogeneous systems.

Theorem 107. *Let $T : \mathbb{X} \to \mathbb{Y}$ be a linear function, x^p be such that $T(x^p) = y$ and S^H be a set containing the solutions for $T(x) = \theta$. Then, the set of solutions for $T(x) = y$ is,*

$$S^N = x^p + S^H = \{x^N \in \mathbb{X} : x^N = x^p + x^H,\ x^H \in S^H\}.$$

Proof. First, suppose $x^N \in S^N$. Then by definition $x^N = x^p + x^H$ with $T(x^p) = y$ and $x^H \in S^H$. Then, by linearity of T

$$T(x^N) = T(x^p + x^H) = T(x^p) + T(x^H) = y + \theta = y.$$

Hence, x^N is a solution for $T(x) = y$.

Second, suppose x^s is a solution for $T(x) = y$, that is, $T(x^s) = y$.

$$T(x^s) = y \Rightarrow T(x^s) = y + \theta = T(x^p) + T(x^H)$$

Now, by linearity

$$T(x^s - x^p - x^H) = \theta,$$

which implies $x^s = x^p + x^H$. Hence, $x^s \in S^N$. □

Remark 30. *1. This theorem shows that if $T(x) = y$ has a solution, then the dimension of S^N is the same as S^H. Also, once we know one solution for $T(x) = y$ we can find all solutions, provided we have the solutions for $T(x) = \theta$.*

2. *Since $im(T)$ is the set of vectors $y \in \mathbb{R}^m$ such that $T(x) = y$ for $x \in \mathbb{X}$, we can write that $im(T)$ is a vector sub-space of $\mathbb{R}^m$ generated by the columns of the $m \times n$ matrix A associated with T, that is*

$$\begin{pmatrix} A_{.1} & A_{.2} & \cdots & A_{.n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = y.$$

Hence, a solution exists if, and only if, $y \in \mathbb{R}^m$ can be written as a linear combination of the columns of A, that is, y is in the column space of A. Now, consider $(A\ y) = \begin{pmatrix} A_{.1} & A_{.2} & \cdots & A_{.n} & y \end{pmatrix}$. Note that if $rank(A) < rank((A\ y))$, y is not a linear combination of the columns of A. If $rank(A) = rank((A\ y))$, y is a linear combination of the columns of A and $Ax = y$ has a solution.

Hence, we have the following theorem.

Theorem 108. *The linear equation $Ax = y$ has at least one solution if, and only if, $rank(A) = rank((A\ y))$.*

If there exists a solution, and given that $dim(null(A)) = n - rank(A)$, the solution is unique if $rank(A) = n$, that is, $dim(null(A)) = 0$. This gives,

Theorem 109. *$Ax = y$ has a unique solution if, and only if, $rank(A) = n = rank((A\ y))$.*

Remark 31.

1. *If $rank(A) = m = n$ then we have $A^{-1}Ax = A^{-1}y$ which implies $x^N = A^{-1}y$.*

2. *Since $Ax + Ba = \theta$, we can write $x^*(a) = -A^{-1}Ba$. Hence, the solution is a linear function of a and*

$$Dx^*(a) = -A^{-1}B$$

which provides comparative statics for the equilibrium. In example 27 we have

$$\begin{pmatrix} p_1^* \\ p_2^* \end{pmatrix} = \frac{1}{c_1\gamma_2 - c_2\gamma_1} \begin{pmatrix} \gamma_2 & -c_2 \\ -\gamma_1 & c_1 \end{pmatrix} \begin{pmatrix} -c_0 \\ \gamma_0 \end{pmatrix}$$

18.2 Nonlinear models

The solution(s) for $F(x, \alpha) := f_\alpha(x) = \theta \in \mathbb{R}^m$ for a given value of α can be described as a relation/correspondence $s : A \subset \mathbb{R}^p \to \mathbb{X} \subseteq \mathbb{R}^n$ that assigns to each $\alpha \in A$ a set

$$s(\alpha) = f_\alpha^{-1}(\theta) = \{x \in \mathbb{X} : f_\alpha(x) = \theta\}.$$

We will concentrate on the case where $m = n$, that is f_α maps $\mathbb{R}^n$ to $\mathbb{R}^n$. As in the case of linear models, we are interested in giving a more precise description of $s(\alpha)$ and, in many instances, understanding how $s(\alpha)$ changes when α takes on different values in A. A key insight is to assume differentiability of F and use the inverse function theorem. Informally, if (x_0, α_0) is such that $F(x_0, \alpha_0) = f_{\alpha_0}(x_0) = \theta$, we note that by the inverse function theorem, if $f_{\alpha_0}(x)$ is continuously differentiable at $x_0 \in \mathbb{X} \subseteq \mathbb{R}^n$, $\mathbb{X}$ is open and $A_{f_{\alpha_0}}(x_0)$ is invertible, there exists an open neighborhood $\mathcal{N}$ of x_0 in which $f_{\alpha_0}(x)$ is one-to-one and onto. Consequently, $x_0 \in f_{\alpha_0}^{-1}(\theta)$, i.e., x_0 is a locally unique solution for $f_{\alpha_0}(x) = \theta$.

The Implicit Function Theorem, which we will study next, helps us deal with comparative statistics for these models. That is, it provides a framework to understand how the solutions change when α changes in a vicinity of α_0. Before we prove the theorem we make the following remarks.

Remark 32. *1. Suppose that the solution $x^*(\alpha)$ for some α is a differentiable function, and by virtue of it being a solution for the system of equations under consideration, we have*

$$F(x^*(\alpha), \alpha) := F\begin{pmatrix} x^*(\alpha) \\ \alpha \end{pmatrix} := \theta.$$

This is an identity in so far as it holds for all α. We can think of a composition $H : A \subseteq \mathbb{R}^p \to \mathbb{R}^n$ such that $H(\alpha) := \theta$ and $H(\alpha) = F(f(\alpha))$ where $f(\alpha) : A \subseteq \mathbb{R}^p \to \mathbb{R}^{n+p}$ and $f(\alpha) = \begin{pmatrix} \underset{(n\times 1)}{x^*(\alpha)} \\ \underset{(p\times 1)}{\alpha} \end{pmatrix}$. Note that since $H(\alpha) := \theta$ for all $\alpha \in A$ we can differentiate *both sides of the identity with respect to α and write $DH(\alpha) = 0$. By the chain rule $DH(\alpha) = \underbrace{DF(f(\alpha))}_{n\times(n+p)}\underbrace{Df(\alpha)}_{(n+p)\times p} = \theta$. Since there are two sets of arguments for f, i.e., x^* and α, we may write*

$$DH(\alpha) = \left(\underbrace{D_xF(f(\alpha))}_{n\times n} \quad \underbrace{D_\alpha F(f(\alpha))}_{n\times p} \right) \begin{pmatrix} \frac{\partial x_1^*(\alpha)}{\partial_1} & \frac{\partial x_1^*(\alpha)}{\partial_2} & \cdots & \frac{\partial x_1^*(\alpha)}{\partial_p} \\ \vdots & & & \\ \frac{\partial x_n^*(\alpha)}{\partial_1} & \frac{\partial x_n^*(\alpha)}{\partial_2} & \cdots & \frac{\partial x_n^*(\alpha)}{\partial_P} \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \theta$$

$$= (D_xF(x^*(\alpha),\alpha)\ D_\alpha F(x^*(\alpha),\alpha)) \begin{pmatrix} Dx^*(\alpha) \\ \mathrm{I} \end{pmatrix} = \theta.$$

Equivalently, we write

$$D_xF(x^*(\alpha),\alpha)Dx^*(\alpha) + D_\alpha F(x^*(\alpha),\alpha) = \theta$$

which implies

$$Dx^*(\alpha) = -\left[D_xF(x^*(\alpha),\alpha)\right]^{-1} D_\alpha F(x^*(\alpha),\alpha),$$

provided $D_xF(x^*(\alpha),\alpha)$ is invertible.

Theorem 110. *Let $F : \mathbb{X} \times A \subseteq \mathbb{R}^{n+p} \to \mathbb{R}^n$ be a continuously differentiable function on an open set $\mathbb{X} \times A$. Consider the system of equations $F(x,\alpha) = \theta$, and assume that it has a solution $x^0 \in \mathbb{X}$ for given $\alpha^0 \in A$. If $D_xF(x^0,\alpha^0)$ is invertible,*

1. *There exist U open in $\mathbb{R}^{n+p}$ and U_α open in $\mathbb{R}^p$ with $(x^0,\alpha^0) \in U$ and $\alpha^0 \in U_\alpha$ such that for all $\alpha \in U_\alpha$ there exists a unique $x(\alpha)$ such that $(x(\alpha),\alpha) \in U$ and $F(x(\alpha),\alpha) = \theta$.*

2. *The solution function $x(\alpha) : U_\alpha \to \mathbb{R}^n$ is continuously differentiable and its derivative is given by , $Dx(\alpha) = -[D_xF(x(\alpha),\alpha)]^{-1}D_\alpha F(x(\alpha),\alpha)$.*

3. *If $F(x,\alpha) \in C^k$, so is $x(\alpha)$.*

Proof. Let $G(x,\alpha):\mathbb{R}^{n+p}\to\mathbb{R}^{n+p}$ defined by $G(x,\alpha)=\begin{bmatrix} F(x,\alpha) \\ \alpha \end{bmatrix}$ with

$$G^i(x,\alpha)=\begin{cases} F^i(x,\alpha) & i=1,2,\cdots,n, \\ \alpha_i & i=1,2,\cdots,p. \end{cases}$$

Now, $G(x^0,\alpha^0)=\begin{pmatrix} F(x^0,\alpha^0) \\ \alpha^0 \end{pmatrix}=\begin{pmatrix} \theta \\ \alpha^0 \end{pmatrix}$, where the second equality follows from $F(x^0,\alpha^0)=\theta$. Then, since $D_xF(x^0,\alpha^0)$ exists, we have

$$DG(x^0,\alpha^0)=\left(\begin{array}{cccc:ccc} \frac{\partial G^1}{\partial x_1}(x^0,\alpha^0) & \frac{\partial G^1}{\partial x_2}(x^0,\alpha^0) & \cdots & \frac{\partial G^1}{\partial x_n}(x^0,\alpha^0) & \frac{\partial G^1}{\partial \alpha_1}(x^0,\alpha^0) & \cdots & \frac{\partial G^1}{\partial \alpha_p}(x^0,\alpha^0) \\ \vdots & & & & & & \\ \frac{\partial G^n}{\partial x_1}(x^0,\alpha^0) & \frac{\partial G^n}{\partial x_2}(x^0,\alpha^0) & \cdots & \frac{\partial G^n}{\partial x_n}(x^0,\alpha^0) & \frac{\partial G^n}{\partial \alpha_1}(x^0,\alpha^0) & \cdots & \frac{\partial G^n}{\partial \alpha_P}(x^0,\alpha^0) \\ -- & -- & -- & -- & -- & -- & -- \\ & 0 & & & & I_p & \end{array}\right)$$

$$=\begin{pmatrix} D_xG(x^0,\alpha^0) & D_\alpha G(x^0,\alpha^0) \\ 0 & I_p \end{pmatrix}=\begin{pmatrix} D_xF(x^0,\alpha^0) & D_\alpha F(x^0,\alpha^0) \\ 0 & I_p \end{pmatrix}$$

$DG(x^0,\alpha^0)$ is invertible, since $D_xF(x^0,\alpha^0)$ is invertible by assumption. Because F is continuously differentiable, so is G. Hence, by the inverse function theorem, there exists an open neighborhood U of (x^0,α^0) such that $G(x,\alpha)$ is one to one on U. Hence, G^{-1} exists as a function from $V=G(U)$ to U, where V is an open set in $\mathbb{R}^{n+p}$ and $G(x^0,\alpha^0)=\begin{bmatrix} F(x^0,\alpha^0) \\ \alpha^0 \end{bmatrix}=\begin{bmatrix} \theta \\ \alpha^0 \end{bmatrix}\in V$. Because V is open and $\begin{bmatrix} \theta \\ \alpha \end{bmatrix}\in V$, for all α sufficiently close to α^0, i.e., for all $\alpha\in U_\alpha$. Since G is invertible in U, for all $\alpha\in U_\alpha$ there exists a unique $(x(\alpha),\alpha)\in U$, such that $G(x,\alpha)=\begin{bmatrix} \theta \\ \alpha \end{bmatrix}$, which implies $F(x(\alpha),\alpha)=\theta$. Also, by the inverse function theorem $G^{-1}\in C^1$ and since,

$$G(x(\alpha),\alpha)=\begin{bmatrix} \theta \\ \alpha \end{bmatrix}\Longleftrightarrow\begin{bmatrix} x(\alpha) \\ \alpha \end{bmatrix}=G^{-1}(\theta,\alpha)$$

We see that $x(\alpha)$ is a component of a C^1 function, and therefore, it belongs to C^1 itself. □

Example 28. *The seller of a product pays a proportional tax at a flat rate $\theta\in(0,1)$. Hence, the price received by a seller is $(1-\theta)p$, where p is the market price.*

$$Q^d=D(p),\ D^{(1)}(p)<0,$$

$$Q^s = S((1-\theta)p),\ S^{(1)}(p) > 0.$$

Market clearing requires $D(p) = S((1-\theta)p)$*, which implicitly defines* $p^* = p(\theta)$*. Differentiating* $D(p(\theta)) = S((1-\theta)p(\theta))$ *we obtain*

$$D^{(1)}(p(\theta))p^{(1)}(\theta) = S^{(1)}((1-\theta)p(\theta))((1-\theta)p^{(1)}(\theta) - p(\theta)).$$

This implies that

$$p^{(1)}(\theta) = \frac{-\overbrace{p(\theta)}^{+}\overbrace{S^{(1)}((1-\theta)p(\theta))}^{+}}{\underbrace{D^{(1)}(p(\theta))}_{-} - \underbrace{(1-\theta)S^{(1)}((1-\theta)p(\theta))}_{+}} > 0,\ Q^* = D(p(\theta))$$

and $\frac{dQ^*}{d\theta} = \underbrace{D^{(1)}(p(\theta))}_{-}\underbrace{p^{(1)}(\theta)}_{+} < 0.$

It is important to note that the implicit function assumes the existence of a solution.

19 Mathematical Programming

In this section we study maximization problems given by

$$\max_x \{f(x,\alpha) : x \in c(\alpha)\} \tag{29}$$

where $x \in \mathbb{X} \subseteq \mathbb{R}^n$, $\alpha \in A \subseteq \mathbb{R}^p$, $c(\alpha) \subseteq \mathbb{X}$, $f(x,\alpha) : \mathbb{X} \times A \subseteq \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$. In this context, we call $S : A \twoheadrightarrow X$ with $S(\alpha) = \underset{x}{\operatorname{argmax}}\{f(x,\alpha) : x \in c(\alpha)\}$ the best-response relation/correspondence. For a given $\alpha \in A$, it gives the set of $x \in c(\alpha)$ that provide that maximum vale of f. When there is a unique x, say x^*, that maximizes $f(x,\alpha)$ for $\alpha \in A$, we call $S(\alpha)$ the best response *function*. We call $v(\alpha) = \max_x\{f(x;\alpha) : x \in c(\alpha)\} = f(x^*,\alpha)$ for $x^* \in S(\alpha)$ the value function. We will consider three different maximization scenarios:

1. $c(\alpha)$ is a convex subset of $\mathbb{R}^n$,

2. $c(\alpha) = \{x \in \mathbb{X} :\ g(x,\alpha) = \theta\}$ (Lagrange Problem),

3. $c(\alpha) = \{x \in \mathbb{X} : \; g(x, \alpha) \geq \theta\}$ (Kuhn-Tucker Problem),

where $g : \mathbb{X} \times A \to \mathbb{R}^c$ for $c \in \mathbb{N}$, θ is the null element of $\mathbb{R}^c$ and in item 3, $\geq$ is taken element-wise.

19.1 Scenario 1

We start with the definition of a convex set.

Definition 69. *A set $\mathbb{X} \subseteq \mathbb{R}^n$ is convex if for all $x, x' \in \mathbb{X}$, $x^\lambda = (1-\lambda)x + \lambda x'$ is also in $\mathbb{X}$ for all $\lambda \in (0, 1)$.*

Fix $\alpha = \bar{\alpha}$ and consider

$$\max_x \{f(x, \bar{\alpha}) : x \in c(\bar{\alpha})\}. \tag{30}$$

We first give the following definition.

Definition 70. *Consider the maximization in (30). Let $x \in c(\bar{\alpha})$ and choose a direction (increment) $h \in \mathbb{R}^n$. We say that h is a feasible direction from x if there exists $\delta_h > 0$ such that for all $a \in (0, \delta_h)$, $x + ah \in c(\bar{\alpha})$.*

Theorem 111. *Assume that $f \in C^1$ and let x^* be a solution for (30). Then, $Df(x^*)h \leq 0$ for all $h \in \mathbb{R}^n$ feasible from x^*.*

Proof. Because x^* is a solution for (30) and h is a feasible direction, for a as in definition 70, $f(x^* + ah) \leq f(x^*)$. Since $a > 0$, we have $\frac{f(x^*+ah)-f(x^*)}{a} \leq 0$. Since $f \in C^1$, from Theorem 102, f is continuously differentiable at x^*, which implies that f is differentiable. But since f is differentiable at x^*, for all h feasible, $\lim_{a\to 0} \frac{f(x^*+ah)-f(x^*)}{a} = Df(x^*; h) \leq 0$. Since $f \in C^1$ we can write $Df(x^*; h) = Df(x^*)h$ where $Df(x^*)$ is the Fréchet derivative. $\square$

Remark 33. *1. In the last step of the proof we use the fact that when both Gâteaux and Fréchet differentials exist, they coincide; 2. The theorem gives first order necessary conditions for a maximum; 3. If $c(\bar{\alpha})$ is open, then for all $x \in c(\bar{\alpha})$, all directions are feasible.*

Then, note that

$$Df(x^*)h = \begin{pmatrix} \frac{\partial}{\partial_1} f(x^*) & \cdots & \frac{\partial}{\partial_n} f(x^*) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \sum_{i=1}^{n} \frac{\partial}{\partial_i} f(x^*) h_i \leq 0.$$

Since $\sum_{i=1}^{n} \frac{\partial}{\partial_i} f(x^*) h_i \leq 0$*, for all* h_i*, it must be that* $\frac{\partial}{\partial_i} f(x^*) = 0$ *for all* i*.*

Corollary 6. *Let* $f \in C^1$*,* $c(\bar{\alpha})$ *open and* x^* *be a solution for (30). Then,* $Df(x^*)h = \theta$ *for all* h*. That is,* $\frac{\partial f}{\partial i}(x^*) = 0$ *for all* i*.*

It should be noted that $\frac{\partial f}{\partial i}(x^*) = 0$ is not sufficient for x^* to be a maximum. We now look at the property of concavity and some related concepts.

Definition 71. *Let* $\mathbb{X}$ *be convex and* $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}$*.*

1. f *is concave if for all* $x, x' \in \mathbb{X}$*, we have*

   $$(1-\lambda) f(x) + \lambda f(x') \leq f((1-\lambda)x + \lambda x') = f(x^\lambda)$$

   for all $\lambda \in (0, 1)$*.*

2. f *is strictly concave if*

   $$(1-\lambda) f(x) + \lambda f(x') < f((1-\lambda)x + \lambda x') = f(x^\lambda)$$

   for all $\lambda \in (0, 1)$*.*

Remark 34. *Reverse the inequalities to obtain the definition of convex and strictly convex functions.*

Theorem 112. *Let* $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}$ *belong to* C^1 *defined on an open and convex set* $\mathbb{X}$*. Then,*

1. f *concave* $\Leftrightarrow$ *for all* $x_0, x \in \mathbb{X}$*,* $f(x) \leq f(x_0) + Df(x_0)(x - x_0)$*;*

2. *f is strictly concave $\Leftrightarrow$ for all $x_0, x \in \mathbb{X}$, $f(x) < f(x_0) + Df(x_0)(x - x_0)$.*

Proof. First, sufficiency for both 1 and 2. For 1, let $x, x_0 \in \mathbb{X}$ and write $x^\lambda = (1-\lambda)x_0 + \lambda x = x_0 + \lambda(x - x_0)$ with $\lambda \in (0,1)$. By concavity $(1-\lambda)f(x_0) + \lambda f(x) \leq f(x_0 + \lambda(x - x_0))$, or $f(x) - f(x_0) \leq \frac{f(x_0+\lambda(x-x_0))-f(x_0)}{\lambda}$, since $\lambda \neq 0$. Now, take the limit on both sides as $\lambda \to 0$. Then,

$$f(x) - f(x_0) \leq Df(x_0)(x - x_0). \tag{31}$$

For 2, we have $f(x) - f(x_0) < \frac{f(x^\lambda)-f(x_0)}{\lambda}$ and concavity implies $f(x^\lambda) - f(x_0) \leq Df(x_0)(x^\lambda - x_0)$ (letting $x = x^\lambda$ in (31)). Then, $f(x) - f(x_0) < \frac{1}{\lambda} Df(x_0)(x^\lambda - x_0) = Df(x_0)\frac{1}{\lambda}(x_0 + \lambda(x - x_0) - x_0) = Df(x_0)(x - x_0)$. Necessity is left as an exercise for both cases 1 and 2. □

Definition 72. *Let $\mathbb{X}$ be convex and $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}$. f is quasi-concave if for all $x, x' \in \mathbb{X}$ and for all $\lambda \in (0,1)$*

$$f((1-\lambda)x + \lambda x') \geq \min\{f(x), f(x')\}.$$

f is strictly quasi-concave if for all $x, x' \in \mathbb{X}$ and for all $\lambda \in (0,1)$

$$f((1-\lambda)x + \lambda x) > \min\{f(x), f(x')\}.$$

Remark 35. *It is clear that concavity implies quasi-concavity and strict concavity implies strict quasi-concavity. To see this, note that by concavity*

$$f(\lambda x + (1-\lambda)x') \geq \lambda f(x) + (1-\lambda)f(x') \text{ for all } \lambda \in (0,1). \tag{32}$$

Without loss of generality let $\min\{f(x), f(x')\} = f(x')$. Then,

$$\lambda f(x) + (1-\lambda)f(x') = \lambda(f(x) - f(x')) + f(x').$$

But since $f(x) \geq f(x')$ we have that $f(x) - f(x') \geq 0$. Thus, $\lambda(f(x) - f(x')) \geq 0$ and

$$\lambda f(x) + (1-\lambda)f(x') \geq f(x'). \tag{33}$$

Using equation (32) we have

$$f(\lambda x + (1-\lambda)x') \geq f(x') = \min\{f(x), f(x')\}.$$

establishing quasi-concavity.

Strict concavity means that

$$f(\lambda x + (1-\lambda)x') > \lambda f(x) + (1-\lambda)f(x') \text{ for all } \lambda \in (0,1).$$

Following the same steps from above, we have

$$f(\lambda x + (1-\lambda)x') > \lambda f(x) + (1-\lambda)f(x') \geq f(x') = \min\{f(x), f(x')\}.$$

establishing that strict concavity implies strict quasi-concavity.

Theorem 113. *Let $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}$ where $\mathbb{X}$ open and convex and $f \in C^1$. Then, f is quasi-concave in $\mathbb{X}$ if, and only if, for all $x', x'' \in \mathbb{X}$, we have*

$$f(x'') - f(x') \geq 0 \Rightarrow Df(x')(x'' - x') \geq 0.$$

If $x' \neq x''$ with $f(x'') - f(x') \geq 0 \Rightarrow Df(x')(x'' - x') > 0$, then f is strictly quasi-concave (the converse is not always true).

Proof. (the first sufficiency) For all $x', x'' \in \mathbb{X}$ and $\lambda \in [0,1]$ define $g(\lambda) = f(x^\lambda = x' + \lambda(x'' - x'))$. Because $f \in C^1$, g is differentiable and $g^{(1)}(\lambda) = Df(x' + \lambda(x'' - x'))(x'' - x')$. Assume f is quasi-concave and $f(x') \leq f(x'')$. Then, by definition $f(x^\lambda) = g(\lambda) \geq f(x') = g(\lambda = 0)$ for all $\lambda \in [0,1]$ and at $\lambda = 0$, $g^{(1)}(\lambda)|_{\lambda=0} = Df(x')(x'' - x') \geq 0$ (g increasing).

If f is strictly quasi-concave, $g(\lambda) > f(x') = g(\lambda = 0)$ for all $\lambda \in [0,1]$ and g is strictly increasing at $\lambda = 0$. But this does not imply that $g^{(1)}(0) > 0$ ($f(x) = x^3$ is strictly increasing at 0, but $f^{(1)}(x = 0) = 0$).

(Necessity) Assuming $f(x') \leq f(x'') \Rightarrow Df(x')(x'' - x') \geq 0$ for all $x', x'' \in \mathbb{X}$, we want to show that this implies,

$$g(\lambda) = f(x' + (x'' - x')\lambda) \geq f(x') = g(0)$$

for all $\lambda \in (0,1)$, i.e., quasi-concavity. Suppose not, i.e., suppose there exists some $\lambda^0 \in (0,1)$ such that $g(\lambda^0) < g(0)$. Because $g(1) = f(x'') \geq f(x') = g(0)$, we can choose λ^0 such that $g^{(1)}(\lambda^0) > 0$. Now, if that is the case, put $x^0 = x' + \lambda^0(x'' - x')$. Since $f(x') = g(0) > g(\lambda^0) = f(x^0)$ and $x^0 \in \mathbb{X}$, we have by assumption

$$Df(x^0)(x' - x^0) = Df(x^0)(-\lambda^0)(x'' - x') \geq 0$$

and hence $Df(x^0)(x'' - x') \leq 0$. But $g^{(1)}(\lambda^0) = Df(x_0)(x'' - x') > 0$, which contradicts the previous inequality (same works for strict quasi-concavity). □

Definition 73. *Let $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}$, $\mathbb{X}$ open and convex. $f \in C^1$ is pseudo-concave in $\mathbb{X}$, if for all $x', x'' \in \mathbb{X}$*

$$f(x'') - f(x') > 0 \Rightarrow Df(x')(x'' - x') > 0.$$

Remark 36. *From Theorem 112 we have that f concave $\Leftrightarrow$ for all $x_0, x \in \mathbb{X}$, $f(x) - f(x_0) \leq Df(x_0)(x - x_0)$. Thus, if $f(x) - f(x_0) > 0$, then $0 < Df(x_0)(x - x_0)$. So, f is pseudo-concave. Also, strict concavity implies pseudo-concavity and strict quasi-concavity implies pseudo-concavity.*

Theorem 114. *Let $f \in C^1$ be pseudo-concave. If $x^* \in c(\bar{\alpha})$ a convex set and for every direction $h \in \mathbb{R}^n$ feasible from x^* we have $Df(x^*)h \leq 0$ then x^* is a solution for (30).*

Proof. We will show that if x^* is not a solution for (30) then $Df(x^*)h > 0$, for all feasible h. x^* not a solution for (30) implies there exists $x \in c(\bar{\alpha})$ such that $f(x) > f(x^*)$. By pseudo-concavity of f, $f(x) > f(x^*) \Rightarrow Df(x^*)(x - x^*) > 0$. But by convexity of $c(\bar{\alpha})$, $x^\lambda = (1-\lambda)x^* + \lambda x = x^* - \lambda x^* + \lambda x = x^* + \lambda(x - x^*) \in c(\bar{\alpha})$ and $(x - x^*)$ is a feasible direction from x^*. □

Remark 37. *1. Since concavity and strict quasi-concavity imply pseudo-concavity, the theorem holds if f is assumed to have either of these properties; 2. The theorem gives sufficient* *conditions for a global maximum; 3. If $f \in C^1$ is concave and x^* is such that $\nabla f(x^*) = \theta$, then it is a global maximum of f .*

Theorem 115. *Let $f \in C^2$ with $c(\bar{\alpha})$ open and convex. Let $x^* \in c(\bar{\alpha})$ and $Df(x^*) = \theta$. If $D^2 f(x^*)$ is negative definite then f achieves a strict local maximum at x^*.*

Proof. Let $h \in \mathbb{R}$ be an arbitrary direction. By the convexity and openness of $c(\bar{\alpha})$ there exists $\delta > 0$ such that $x^* + ah \in c(\bar{\alpha})$ for all $a \in (0, \delta)$. Then,

$$f(x^* + ah) - f(x^*) = Df(x^*)ah + \frac{1}{2}(ah)^{'} D^2 f(x^* + \lambda_a ah)ah$$

for some $\lambda_a \in (0, 1)$. But since $Df(x^*) = \theta$

$$f(x^* + ah) - f(x^*) = \frac{a^2}{2} h^{'} D^2 f(x^* + \lambda_a ah)h.$$

Now, $h^{'} D^2 f(x^*)h \equiv h^{'} D^2 f(x^* + \lambda_a ah)h|_{a=0} \equiv Q(a = 0) < 0$. Since $Q(a)$ is continuous in a, then

$$|h^{'} D^2 f(x^* + \lambda_a ah)h - h^{'} D^2 f(x^*)h| \leq \epsilon$$

whenever $a < \gamma$. Hence, $h^{'} D^2 f(x^* + \lambda_a ah)h < 0$. Therefore, $f(x^* + ah) - f(x^*) < 0$ for all $a < \gamma$. $\square$

Theorem 116. *Let x^*be a solution for (30) with $c(\bar{\alpha})$ convex. If f is strictly quasi-concave (strict concavity implies strict quasi-concavity) then x^* is unique.*

Proof. Suppose there exists two solutions in $c(\bar{\alpha})$, say x', x''. Then, $f(x') = f(x") = M$. By strict quasi-concavity of f, for any $\lambda \in (0, 1)$

$$f((1 - \lambda)x' + \lambda x'') > \min\{f(x'),\ f(x'')\} = M$$

where $(1 - \lambda)x' + \lambda x''$ is feasible by convexity of $c(\alpha)$. Because $f(x^\lambda) > M$, then neither x' nor x'' solve (30). $\square$

Theorem 117. *Let $f : \mathbb{X} \subseteq \mathbb{R}^n \to \mathbb{R}$, be such that $f \in C^2$, and $\mathbb{X}$ is open and convex. Then, f is concave if, and only if, for all $x \in \mathbb{X}$ and for all $h \in \mathbb{R}^n$, $h^{'}D^2f(x)h \leq 0$. If $h^{'}D^2f(x)h < 0$ for all $x \in \mathbb{X}$ and for all $h \in \mathbb{R}^n$, f is strictly concave.*

Proof. (Sufficiency) Let $x \in \mathbb{X}$. Since $\mathbb{X}$ is open, there exists $\alpha \in (-\delta, \delta)$ for $\delta > 0$ such that for any $h \in \mathbb{R}^n$, $x + h\alpha \in \mathbb{X}$. Let $g : (-\delta, \delta) \to \mathbb{R}$, with

$$g(\alpha) = f(x + h\alpha) - f(x) - Df(x)h\alpha. \tag{34}$$

Since $f \in C^2$, so is g. Also, $g(0) = 0$. By Theorem 112, since f is concave, $f(x+h\alpha)-f(x) \leq Df(x)h\alpha$ for all $\alpha \in (-\delta, \delta)$. Using (34) we conclude $g(\alpha) \leq 0$ for all $\alpha \in (-\delta, \delta)$. Since $g(0) = 0$, g is maximized at 0, a point in the interior of $(-\delta, \delta)$. Then, by the corollary to Theorem 111

$$g^{(1)}(\alpha) = Df(x + h\alpha)h - Df(x)h|_{\alpha=0} = 0,$$

and for $\alpha = 0$ to be a maximum,

$$g^{(2)}(\alpha) = h^{'}D^2f(x + h\alpha)h|_{\alpha=0} \leq 0, \text{ or } h^{'}D^2f(x)h \leq 0.$$

(Necessity) Since $f \in C^2$, by Theorem 103, for some $\alpha \in (0, 1)$

$$f(x + h) - f(x) - Df(x)h = \frac{1}{2}h^{'}D^2f(x + h\alpha)h.$$

Since $h^{'}D^2f(x+h\alpha)h \leq 0$, for all $x+h\alpha \in \mathbb{X}$, we have $f(x+h) - f(x) - Df(x)h \leq 0$ which implies $f(x + h) - f(x) \leq Df(x)h$. The last inequality implies concavity by theorem 112. The proof of the last statement in the theorem follows easily from above. □

Remark 38. *1. The theorem characterizes concavity and strict concavity of a function $f \in C^2$ by $h^{'}D^2f(x)h$; 2. In Theorem 13.5 the condition that $h^{'}D^2f(x^*)h < 0$ can be replaced (is equivalent) with strict concavity.*

Example 29. *Let $y = f(x_1, x_2) = x_1^\alpha x_2^\beta$ be a production function where y is output, x_1, x_2 are inputs. $\alpha + \beta < 1, \beta > 0$ and $\alpha > 0$. Firm maximize profits.*

$$\pi(x_1, x_2) = px_1^\alpha x_2^\beta - w_1 x_1 - w_2 x_2.$$

p is price of output, w_1, w_2 are input prices. Obtain, and let partial derivatives equal to zero.

$$\frac{\partial \pi}{\partial x_1} = p\alpha x_1^{\alpha-1} x_2^\beta - w_1 = 0$$

$$\frac{\partial \pi}{\partial x_2} = px_1^\alpha \beta x_2^{\beta-1} - w_2 = 0$$

For these to characterize a maximum, we need $D^2 f$ to be negative definite, that is,

$$D^2 f = \begin{bmatrix} p\alpha(\alpha-1)x_1^{\alpha-2}x_2^\beta & p\alpha x_1^{\alpha-1}x_2^{\beta-1} \\ p\alpha\beta x_1^{\alpha-1}x_2^{\beta-1} & p\beta(\beta-1)x_1^\alpha x_2^{\beta-2} \end{bmatrix} = py \begin{bmatrix} \frac{\alpha(\alpha-1)}{x_1^2} & \frac{\alpha\beta}{x_1 x_2} \\ \frac{\alpha\beta}{x_1 x_2} & \frac{\beta(\beta-1)}{x_2^2} \end{bmatrix}$$

where $y = x_1^\alpha x_2^\beta$, or, $D^2 f = py \begin{bmatrix} \frac{1}{x_1} & 0 \\ 0 & \frac{1}{x_2} \end{bmatrix} \begin{bmatrix} \alpha(\alpha-1) & \alpha\beta \\ \alpha\beta & \beta(\beta-1) \end{bmatrix} \begin{bmatrix} \frac{1}{x_1} & 0 \\ 0 & \frac{1}{x_2} \end{bmatrix}$. Now,

$$\begin{bmatrix} \alpha(\alpha-1) & \alpha\beta \\ \alpha\beta & \beta(\beta-1) \end{bmatrix} = \alpha \begin{bmatrix} \alpha-1 & \beta \\ \beta & \frac{\beta(\beta-1)}{\alpha} \end{bmatrix} = \alpha\beta \begin{bmatrix} \frac{\alpha-1}{\beta} & 1 \\ 1 & \frac{\beta-1}{\alpha} \end{bmatrix}.$$

Since $\frac{\alpha-1}{\beta} < 0$, and $\frac{(\alpha-1)}{\beta}\frac{(\beta-1)}{\alpha} - 1 = \frac{\alpha\beta-\alpha-\beta+1}{\alpha\beta} - 1 = \frac{1-(\alpha+\beta)}{\alpha\beta} > 0$ since $\alpha + \beta < 1$, $\alpha, \beta > 0$. Hence, $h' D^2 f h < 0$. Now, $w_1 = \frac{p\alpha y}{x_1}$, $w_2 = \frac{p\alpha y}{x_2}$, and $\frac{w_1}{w_2} = \frac{\alpha x_2}{\beta x_1}$ implies $x_2 = \frac{\beta x_1 w_1}{\alpha w_2}$. Substituting back into the first order conditions

$$w_1 = p\alpha x_1^{\alpha-1} x_2^\beta = p\alpha x_1^{\alpha-1} \left(\frac{\beta w_1 x_1}{\alpha w_2}\right)^\beta \Rightarrow x_1^* = \left(\frac{p\alpha^{1-\beta}\beta^\beta w_1^\beta}{w_2^\beta}\right)^{\frac{1}{1-\alpha-\beta}}.$$

Note that we use the fact that, for $A_{n\times n}$ (symmetric)

1. *$h'Ah > 0 \Leftrightarrow d_k > 0$ for $k = 1, ..., n$, where $d_k = det(A_k = [a_{ij}]_{i=1}^k)$, $k \le n$.*

2. *$h'Ah < 0 \Leftrightarrow d_k < 0$ for k odd, $d_k > 0$ for k even.*

3. *$h'Ah \ge 0 \Leftrightarrow$ if $d_k \ge 0$ for all k*

Alternatively,

1. $h'Ah > 0 \Leftrightarrow$ *all eigenvalues* $\lambda_1, ..., \lambda_n > 0$,

2. $h'Ah \geq 0 \Leftrightarrow$ *all eigenvalues* $\lambda_1, ..., \lambda_n \geq 0$,

3. $h'Ah < 0 \Leftrightarrow$ *all eigenvalues* $\lambda_1, ..., \lambda_n < 0$,

4. $h'Ah \leq 0 \Leftrightarrow$ *all eigenvalues* $\lambda_1, ..., \lambda_n \leq 0$.

Example 30. *If* $f(x) = -x^2 + 2x$*, then* $\frac{df(x)}{dx} = -2x + 2$, $\frac{d^2 f(x)}{dx^2} = -2$. $\frac{df(x)}{dx} = -2x + 2 = 0 \Rightarrow x = 1$. $\frac{d^2 f(x)}{dx^2} = -2$.

Example 31. *Let* $f(\beta; Y, X) = -\|Y - X\beta\|$ *where* $Y \in \mathbb{R}^n, X \in \mathbb{R}^n, \beta \in \mathbb{R}$. *Then,*

$$-\|Y - X\beta\| = -\left(\sum_{i=1}^{n} (Y_i - X_i\beta)^2\right)^{1/2}.$$

and we would like to maximize f *for fixed* Y *and* X *by choosing* β.

$$\begin{aligned}\frac{d}{d\beta} f(\beta; Y, X) = -\frac{d}{d\beta}\|Y - X\beta\| &= -\frac{1}{2}\left(\sum_{i=1}^{n} (Y_i - X_i\beta)^2\right)^{-\frac{1}{2}} \sum_{i=1}^{n} 2(Y_i - X_i\beta)(-X_i) \\ &= \left(\sum_{i=1}^{n} (Y_i - X_i\beta)^2\right)^{-\frac{1}{2}} \sum_{i=1}^{n} X_i(Y_i - X_i\beta).\end{aligned}$$

Choosing $\hat{\beta}$ *that sets the derivative equal to* 0*, i.e.,* $\frac{d}{d\beta} f(\hat{\beta}; Y, X) = 0$ *implies* $\sum_{i=1}^{n} X_i(Y_i - X_i\beta) = 0 \Leftrightarrow \hat{\beta} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}$. *Now,*

$$\frac{d^2}{d\beta^2} f(\hat{\beta}; Y, X) = \frac{-\sum_{i=1}^{n} X_i^2 \|Y - X\hat{\beta}\| + 0 \cdot \frac{d}{d\beta}(\sum_{i=1}^{n} (Y_i - X_i\hat{\beta})^2)^{-\frac{1}{2}}}{\|Y - X\hat{\beta}\|^2} = \frac{-\sum_{i=1}^{n} X_i^2}{\|Y - X\hat{\beta}\|} < 0$$

If $\beta \in B$*, that is convex and open,* $\hat{\beta}$ *is a strict local maximum. The maximization problem in this example is known as the least squares problem. It is normally posed as the minimization of a norm, we chose to state it as the maximization of the negative of a norm.*

We will now study some concepts that allow us to speak of continuity of correspondences. Continuity of correspondences is more involved than that of functions. We start with the notions of upper semi-continuity and lower semi-continuity for functions.

Definition 74. *Let $f : \mathbb{X} \subset \mathbb{R}^n \to \mathbb{R}$ and consider $x \in \mathbb{X}$.*

1. *f is upper semi-continuous at x if, for all $\{x_n\}_{n=1,2,\cdots} \in \mathbb{X}$ and for all $\epsilon > 0$ there exists a number $N \in \mathbb{N}$ such that*

$$x_n \to x \Rightarrow f(x_n) < f(x) + \epsilon \textit{ for all } n \geq N$$

2. *f is lower semi-continuous at x if, for all $\{x_n\}_{n=1,2,\cdots} \in \mathbb{X}$ and for all $\epsilon > 0$ there exists a number $N \in \mathbb{N}$ such that*

$$x_n \to x \Rightarrow f(x_n) > f(x) - \epsilon \textit{ for all } n \geq N$$

3. *f is continuous at x if f is upper semi-continuous and lower semi-continuous at x.*

Note that if f is upper semi-continuous and lower semi-continuous at x, then for all $\epsilon > 0$ there exists a number $N \in \mathbb{N}$ such that

$$f(x) - \epsilon < f(x_n) < f(x) + \epsilon \text{ or } |f(x_n) - f(x)| < \epsilon$$

whenever $x_n \to x$ and for all $n \geq N$, conforming with our usual definition of continuity. We now extend the upper and lower semi-continuity from functions to correspondences.

Definition 75. *Let $\mathbb{X} \subset \mathbb{R}^m$, $\mathbb{Y} \subset \mathbb{R}^p$, $f : \mathbb{X} \twoheadrightarrow 2^{\mathbb{Y}}$ and $x \in \mathbb{X}$.*

1. *f is upper semi-continuous at x if, for all $\{x_n\}_{n=1,2,\cdots} \in \mathbb{X}$ and $\{y_n\}_{n=1,2,\cdots} \in \mathbb{Y}$ where $y_n \in f(x_n)$*

$$x_n \to x,\, y_n \to y \Rightarrow y \in f(x)$$

2. f is lower semi-continuous at x if, for all $\{x_n\}_{n=1,2,\cdots} \in \mathbb{X}$ and $y \in f(x)$ such that

$$x_n \to x \Rightarrow \text{ there exists } \{y_n\}_{n=1,2,\cdots} \text{ where } y_n \in f(x_n) \text{ such that } y_n \to y.$$

3. f is continuous at x if f is upper semi-continuous and lower semi-continuous at x.

Definition 76. *Let $\mathbb{X}, \mathbb{Y}$ be finite dimensional Euclidean spaces and let $\psi : \mathbb{X} \twoheadrightarrow 2^{\mathbb{Y}}$ be a correspondence.*

1. ψ is usc (upper-semi-continuous) at $x \in \mathbb{X}$ if for all open sets V in $\mathbb{Y}$ such that $\psi(x) \subseteq V$ there exists a neighborhood U of x such that $\psi(x') \subseteq V$ for all $x' \in U$.

2. ψ is lsc (lower-semi-continuous) at $x \in \mathbb{X}$ if for all open sets V in $\mathbb{Y}$ with $\psi(x) \bigcap V \neq \emptyset$ there exists a neighborhood U of x such that $\psi(x') \bigcap V \neq \emptyset$ for all $x' \in U$.

3. ψ is continuous at x if it is usc and lsc at x.

Remark 39. *Let $\mathbb{X}$ and $\mathbb{Y}$ be normed spaces and $f : \mathbb{X} \to \mathbb{Y}$. f is continuous if for every open set $O_{\mathbb{Y}}$, $f^{-1}(O_{\mathbb{Y}}) = O_{\mathbb{X}}$ is open in $\mathbb{X}$. If we focus on a point $x \in \mathbb{X}$ such that $x \in f^{-1}(O_{\mathbb{Y}})$, there is a neighborhood $\mathcal{N}_x$ of x such that every $x' \in \mathcal{N}_x$ is also in $f^{-1}(O_{\mathbb{Y}})$, i.e., $x' \in f^{-1}(O_{\mathbb{Y}})$.*

Now, in the case of correspondences, we may require (usc) that for

$$x' \in \mathcal{N}_x,\ \psi(x') \subseteq V$$

whenever $\psi(x) \subseteq V$. That is, in this case we require that points in a neighborhood of x have images that are entirely contained in an open set that contains $\psi(x)$.

Me may also require (lsc) that

$$x' \in \mathcal{N}_x,\ \psi(x') \bigcap V \neq \emptyset$$

that is, overlaps with V, whenever $\psi(x) \bigcap V \neq \emptyset$.

We now give two characterization theorems for lower and upper semi-continuity.

Definition 77. *A correspondence $f : \mathbb{X} \twoheadrightarrow 2^{\mathbb{Y}}$ where $\mathbb{X}$ and $\mathbb{Y}$ are metric spaces is said to be closed value at $x \in \mathbb{X}$ if $f(x)$ is a closed set. It is said to be compact valued at x if $f(x)$ is compact valued in $\mathbb{Y}$.*

Theorem 118. *A compact-valued correspondence $\psi : \mathbb{X} \twoheadrightarrow 2^{\mathbb{Y}}$ is upper semi-continuous at x if, and only if, for all $\{x_n\}_{n=1,2,\cdots}$ such that $x_n \to x$, every $\{y_n\}_{n=1,2,\cdots}$ such that $y_n \in \psi(x_n)$, for all n, has a convergent subsequence $\{y_{n_k}\}$ with limit in $\psi(x)$.*

Proof. ($\Rightarrow$) $\psi(x)$ is bounded by compactness. Hence, there exists an open set B such that $\psi(x) \subset B$. By upper semi-continuity there exists an open neighborhood $\mathcal{U}$ of x such that for all $z \in \mathcal{U}$, $\psi(z) \subseteq B$. Since $x_n \to x$, there exists N such that $x_n \in \mathcal{U}$ for all $n > N$, and consequently $\psi(x_n) \subseteq B$. Thus, $\{y_n : y_n \in \psi(x_n)\}$ is bounded and contains a convergent subsequence by the Bolzano-Weierstrass Theorem. Let $\{y_{n_k}\}$ be such sequence and $y_{n_k} \to y$ as $n \to \infty$. We need to show that $y \in \psi(x)$.

Suppose $y \notin \psi(x)$. Since $\psi(x)$ is compact, y cannot be a boundary point of $\psi(x)$ and $\|y - z\| > 0$ for any point $z \in \psi(x)$. Thus, there exists closed ϵ-balls such that

$$y \notin \bigcup_{a \in \psi(x)} S(a, \epsilon) := \mathcal{B}$$

with $0 < \epsilon < \inf_{a \in \psi(x)} d(a, y)$ with interior that is open.

Since ψ is usc at x and $x_n \to x$, $\psi(x_n) \subset int\,(\mathcal{B})$ for n sufficiently large. Thus, $\{y_n : y_n \in \psi(x_n)\} \subset \mathcal{B}$. Since $\mathcal{B}$ is closed, y (the limit of y_n) is such that $y \in \mathcal{B}$, which contradicts $y \notin \mathcal{B}$.

($\Leftarrow$) We establish the contrapositive by construction. That is, if ψ not usc at x we will construct a sequence $x_n \to x$ and $y_n \in \psi(x_n)$ with no subsequence converging to a point in $\psi(x)$.

ψ not usc at x implies there exists V open such that $\psi(x) \subseteq V$ where for $\mathcal{U}$ a neighborhood of x and a $x' \in \mathcal{U}$, $\psi(x')$ is not a subset of V. Choose $U_{n+1} \subseteq U_n \subseteq \cdots$ to be balls with radius $1/n$ centered at x and for each there exists z_n such that $\psi(z_n) \nsubseteq V$. By construction $z_n \to x$, but there exists $y_n \in \psi(z_n)$ with $y_n \in V^c$. Suppose there exists y_{n_k} as subsequence of y_n such that $y_{n_k} \to y$. Since $\{y_{n_k}\} \in V^c$ (V^c is closed as V is open), $y \in V^c$ and, of course, $y \notin V$. Since $\psi(x) \subset V$ we have that $y \in \psi(x)$ □

The following theorem provides a characterization of lower semi-continuity.

Theorem 119. *A correspondence $\psi : \mathbb{X} \twoheadrightarrow 2^{\mathbb{Y}}$ is lsc at x if, and only if, for every sequence $\{x_n\}_{n=1,2,\cdots}$ converging to x and every point $y \in \psi(x)$ there exists a sequence $y_n \in \psi(x_n)$ for all n such that $y_n \to y$.*

Proof. See Aliprantis and Border, 2006, p. 565. □

19.2 Comparative Statics and Value Functions

Consider the problem

$$\max_{x \in c(\alpha)} f(x, \alpha) \tag{35}$$

and solutions $x^*(\alpha)$ depending on α. Write

$$S(\alpha) = \underset{x \in c(\alpha)}{\operatorname{argmax}} f(x, \alpha).$$

Since, for all α there might be multiple solutions for (35), $S(\alpha)$ defines a correspondence $S : A \twoheadrightarrow 2^{\mathbb{R}^n}$. As defined earlier, the value function is $V(\alpha) = f(x^*, \alpha)$, for $x^* \in S(\alpha)$. Note that V is always a well defined function $V : A \to \mathbb{R}$.

Theorem 120. *(Berge). Let $\mathbb{X} \subseteq \mathbb{R}^n$, $A \subseteq \mathbb{R}^p$, $f : \mathbb{X} \times A \to \mathbb{R}$ be a continuous function and $c : A \twoheadrightarrow 2^{\mathbb{X}}$ be a nonempty, compact valued and continuous correspondence. Consider the maximization problem*

$$\max_{x \in c(\alpha)} f(x, \alpha).$$

Then, $V(\alpha) = \max_{x \in c(\alpha)} f(x, \alpha)$ *is continuous, and the solution correspondence* $S(\alpha) = \underset{x \in c(\alpha)}{argmax}\{x \in c(\alpha) : f(x, \alpha) = V(\alpha)\}$ *is nonempty, compact valued and upper-semi-continuous.*

Proof. Fix $\alpha \in A$. By assumption $c(\alpha)$ is compact, nonempty and f is continuous. Hence, by Weierstrass Theorem f achieves a maximum on $c(\alpha)$ and $S(\alpha)$ is nonempty. Since, $S(\alpha) \subseteq c(\alpha)$ it is bounded, since $c(\alpha)$ is compact. Now, consider $\{x_n\}$ a sequence of maximizers for given α ($x_n \in S(\alpha)$) and suppose $x_n \to x$. We want to show that $x \in S(\alpha)$. If that is the case, $S(\alpha)$ is closed, and together with boundedness we have that $S(\alpha)$ is compact. Because $c(\alpha)$ is closed and $\{x_n\} \in c(\alpha)$ then $x \in c(\alpha)$. Also, $V(\alpha) = f(x_n, \alpha)$ for all x_n, and since f is continuous,

$$lim_{n \to \infty} f(x_{n,} \alpha) = f(x, \alpha) = V(\alpha).$$

That is, x is a maximizer, and therefore $x \in S(\alpha)$. Now, let's show upper-semi-continuity.

Since S is compact-valued, we use Theorem 118 to establish upper-semi-continuity of S. Fix α and consider α_n such that $\alpha_n \to \alpha$, and choose $x_n \in S(\alpha_n) \subseteq c(\alpha_n)$ for each n. To establish that S is usc, we must show that there exists a subsequence x_{n_k} of x_n with limit in $S(\alpha)$. Because c is continuous (usc), there exists $\{x_{n_k} \in S(\alpha_{n_k}) \subseteq C(\alpha_{n_k})\}$ converging to some $x \in c(\alpha)$. The question that remains is whether $x \in S(\alpha)$.

Let Z be an arbitrary point in $c(\alpha)$. Because c is lsc, there exists a sequence $Z_{n_k} \in c(\alpha_{n_k})$ for all n such that $Z_{n_k} \to Z \in c(\alpha)$. Because x_{n_k} is optimal ($x_{n_k} \in S(\alpha_{n_k})$) we have that

$$f(x_{n_k}, \alpha_{n_k}) \geq f(Z_{n_k}, \alpha_{n_k}), \text{ for all } k.$$

Since f is continuous, taking limits on both sides of the inequality gives

$$f(x, \alpha) \geq f(Z, \alpha),$$

where Z is an arbitrary point. So, x is a maximizer of f in $c(\alpha)$. Hence, $x \in S(\alpha)$ and S is usc.

(Continuity of V) Note that

$$V(\alpha) = \max_{x \in c(\alpha)} f(x, \alpha) = f(S(\alpha), \alpha).$$

Hence, V is the composition of a continuous function, viz., f and a usc correspondence, viz., S. Since f can be viewed as a single valued usc correspondence, V is the composition of two usc correspondences, and therefore, it is itself usc. But since it is single valued, this implies continuity. $\square$

Remark 40. *1. In essence, the theorem says that the value function and the solutions for the maximization vary continuously with α provided that f is continuous and the constraint correspondence, $c(\alpha)$, is compact valued and continuous; 2. The most difficult condition to check is that $c(\alpha)$ is continuous in A.*

19.3 Envelope Theorem

Theorem 121. *Let $f(x, \alpha) \in C^2$ and let x^* satisfy the conditions for a strict local maximum for f at α_0. That is, $Df(x^*, \alpha_0) = \theta$ and $D^2 f(x^*, \alpha_0)$ is negative definite.*

$$V(\alpha) = \max_{x} f(x, \alpha) = f(x(\alpha), \alpha)$$

is differentiable at α_0 and $DV(\alpha_0) = D_\alpha f(x^*, \alpha_0)$.

Proof. We note that $f(x, \alpha) : \mathbb{X} \times A \to \mathbb{R}$. The assumption that $f \in C^2$ implies f is continuously differentiable. $x^*(\alpha_0)$ satisfying the conditions for a strict local maximum at α_0 means

$$D_x f(x^*, \alpha_0) = \theta$$

and $h' D^2 f(x^*, \alpha_0) h < 0$ for all $h \in \mathbb{R}^n$ for $(x^*, \alpha_0) \in int(\mathbb{X} \times A)$. But $h' D^2 f(x^*, \alpha_0) h < 0$ implies that $D^2 f(x^*, \alpha_0)$ is invertible. Hence, by the implicit function theorem, there exists a unique $x(\alpha) : U_{\alpha_0} \to \mathbb{R}^n$ (U_{α_0} is an open neighborhood of α_0) that is continuously differentiable. Now, $V(\alpha) = f(x(\alpha), \alpha)$ and note that $x(\alpha_0) \equiv x^*(\alpha_0)$. Since f is differentiable, by the chain rule

$$D_\alpha V(\alpha) = D_x f(x(\alpha), \alpha) D_\alpha x(\alpha) + D_\alpha f(x(\alpha), \alpha).$$

and since $D_x f(x(\alpha_0), \alpha_0) = \theta$

$$D_\alpha V(\alpha_0) = D_x f(x(\alpha_0), \alpha_0) D_\alpha x(\alpha_0) + D_\alpha f(x(\alpha_0), \alpha_0) = D_\alpha f(x(\alpha_0), \alpha_0).$$

□

Remark 41. *1. This is an "Envelope Theorem" for unconstrained optimization.*

19.4 Scenario 2

We now consider a maximization problem described by

$$\max_x f(x, \bar{\alpha}) \text{ such that } g(x, \bar{\alpha}) = 0 \tag{36}$$

where $g : \mathbb{R}^n \times A \to \mathbb{R}^c$ with $c \leq n$. For fixed $\bar{\alpha}$ we could view $g : \mathbb{R}^n \to \mathbb{R}^c$.

We start by providing a motivational example that helps understand the method we will consider. The main idea is to maximize the function of interest without the constraint but introduce a penalization for violating of the constraint. Take, $n = 2$, $c = 1$ and consider the following example

$$\max_{x_1, x_2} f(x_1, x_2, \bar{\alpha}) \text{ subject to } g(x_1, x_2) = k. \tag{37}$$

Let the penalization be governed by λ and consider

$$\mathcal{L}(x_1, x_2, \lambda) = f(x_1, x_2, \bar{\alpha}) - \lambda(k - g(x_1, x_2)).$$

Here, λ is the penalty and $k - g(x_1, x_2)$ is the deviation of $g(x_1, x_2)$ from k. Suppose $f, g \in C^1$, then for $i = 1, 2$

$$\frac{\partial}{\partial_i} \mathcal{L}(x_1^*, x_2^*, \lambda^*) = \frac{\partial f}{\partial_i}(x_1^*, x_2^*, \bar{\alpha}) + \lambda^* \frac{\partial g}{\partial_i}(x_1^*, x_2^*) = 0.$$

Of course, there is no guarantee that the pair (x_1^*, x_2^*) that solves this two-equation system, does so satisfying $g(x_1^*, x_2^*) = k$. This can be attained by adding

$$\frac{\partial \mathcal{L}}{\partial \lambda}(x_1^*, x_2^*, \lambda^*) = -k + g(x_1^*, x_2^*) = 0.$$

In this case, the triple $(x_1^*, x_2^*, \lambda^*)$ is a stationary point for $\mathcal{L}$.

Now, an example where $n = 3$ and $c = 1$. Let $f(x, y, z) = 2x - 2y + z$ and $g(x, y, z) = x^2 + y^2 + z^2 - 9$. Then,

$$\mathcal{L}(x, y, z; \lambda) = 2x - 2y + z - \lambda(x^2 + y^2 + z^2 - 9)$$

and we have

$$\frac{\partial \mathcal{L}}{\partial x}(x^*, y^*, z^*, \lambda^*) = 2 - 2x^*\lambda^* = 0 \Rightarrow x^*\lambda^* = 1, \tag{38}$$

$$\frac{\partial \mathcal{L}}{\partial y}(x^*, y^*, z^*, \lambda^*) = -2 - 2y^*\lambda^* = 0 \Rightarrow y^*\lambda^* = -1, \tag{39}$$

$$\frac{\partial \mathcal{L}}{\partial z}(x^*, y^*, z^*, \lambda^*) = 1 - 2z^*\lambda^* = 0 \Rightarrow z^*\lambda^* = \frac{1}{2}. \tag{40}$$

Also,

$$\frac{\partial \mathcal{L}}{\partial \lambda}(x^*, y^*, z^*, \lambda^*) = x^{*2} + y^{*2} + z^{*2} - 9 = 0$$

implies $\frac{1}{\lambda^{*2}} + \frac{1}{\lambda^{*2}} + \frac{1}{4\lambda^{*2}} = 9 \Rightarrow \lambda^* = \pm\frac{1}{2}$. If $\lambda^* = \frac{1}{2}$, then $x^* = 2$, $y^* = -2$, $z^* = 1$, and if $\lambda^* = -\frac{1}{2}$, $x^* = -2$, $y^* = 2$, $z^* = -1$. Substituting back into the objective function we have $f(-2, 2, -1) = -9$ and $f(2, -2, 1) = 9$.

Theorem 122. *Let x^* be a solution for $\max\limits_x f(x, \bar{\alpha})$ subject to $g(x) = \theta$ where $f, g \in C^1$ and $rank(Dg(x^*)) = c \leq n$. Then, there exists a unique $\lambda^* \in \mathbb{R}^c$ such that*

$$Df(x^*) + \lambda^{*'} Dg(x^*) = \theta. \tag{41}$$

Proof. Let $x^* = \begin{bmatrix} \underset{c\times 1}{x_\alpha^*} \\ \underset{n-c\times 1}{x_\beta^*} \end{bmatrix}$ be a solution for (36). Because x^* is a solution it satisfies $g(x^*) = \theta$ or $g(x_\alpha^*, x_\beta^*) = \theta$. By the rank assumption, $rank(\underset{c\times n}{Dg}(x^*)) = rank\begin{pmatrix} \underset{c\times c}{D_\alpha g(x^*)} & \underset{c\times n-c}{D_\beta g(x^*)} \end{pmatrix} =$

c. Hence, $\det(D_\alpha g(x^*_\alpha, x^*_\beta)) \neq 0$ or $D_\alpha g(x^*_\alpha, x^*_\beta)$ is invertible. Hence, by the Implicit Function Theorem, for any given x_β in a neighborhood of x^*_β there exists a unique C^1 function (differentiable) $h : U_\beta \to U_\alpha$ with $h(x_\beta) = x_\alpha$, such that $g(x_\alpha, x_\beta) = \theta$ where U_α, U_β are open balls centered at x^*_α, and x^*_β. Now,

$$g(h(x_\beta), x_\beta) = 0 \Rightarrow D_\alpha g(x) Dh(x_\beta) + D_\beta g(x) = 0 \Rightarrow Dh(x_\beta) = -\left[D_\alpha g(x)\right]^{-1} D_\beta g(x).$$

Now, let $F : U_\beta \to \mathbb{R}$ such that $F(x_\beta) = f(h(x_\beta), x_\beta)$. If $x^* = (x^*_\alpha, x^*_\beta)$ is a solution for (36) then x^*_β will be a solution for

$$\max_{x_\beta} \{F(x_\beta); x_\beta \in U_\beta\}$$

and therefore, $DF(x^*_\beta) = \theta$ (U_β is open, so if x^*_β is a maximum, it must be this value). Now,

$$\begin{aligned}
\theta &= DF(x^*_\beta) \\
&= Df(h(x^*_\beta), x^*_\beta) \\
&= D_\alpha f(x^*) Dh(x^*_\beta) + D_\beta f(x^*) \\
&= D_\alpha f(x^*)(-)[D_\alpha g(x^*)]^{-1} D_\beta g(x^*) + D_\beta f(x^*) \\
&= \lambda^{*'} D_\beta g(x^*) + D_\beta f(x^*).
\end{aligned}$$

Hence, $D_\beta f(x^*) + \lambda^{*'} D_\beta g(x^*) = \theta$. $\square$

Remark 42. *If $c = n$, then $-Df(x^*)Dg(x^*)^{-1} = \lambda^{*'}$ and the proof is obvious.*

We now give sufficient conditions for a global maximum.

Theorem 123. *(Sufficient conditions for a global maximum) Let f be pseudo-concave and all $g^j(x)$ be quasi-concave. If (x^*, λ^*) satisfy the Lagrange condition,*

$$Df(x^*) + \lambda^{*'} Dg(x^*) = \theta,$$

with x^* feasible in the constraint set and $\lambda^* \geq \theta$, then x^* is an optimal solution to the Lagrange problem (36).

Proof. We show that if x^* is not a solution for (36) then it does not satisfy (41). If x^* does not solve (36) there exists $x \neq x^*$ such that $f(x) > f(x^*)$. Pseudo-concavity implies if $f(x) > f(x^*)$ then $Df(x^*)(x - x^*) > 0$. Note that $g(x) = g(x^*) = \theta$. By quasi-concavity of g^j, we have

$$Dg^j(x^*)(x - x^*) \geq 0 \text{ for all } j, \text{ or } Dg(x^*)(x - x^*) \geq \theta.$$

So, combining the last inequality with $Df(x^*)(x - x^*) > 0$ and $\lambda \geq \theta$

$$0 < Df(x^*)(x - x^*) + \lambda^{*'} Dg(x^*)(x - x^*) = \left(Df(x^*) + \lambda^{*'} Dg(x^*)\right)(x - x^*)$$

and we conclude $Df(x^*) + \lambda^{*'} Dg(x^*) \neq \theta$. □

Theorem 124. *(Sufficient conditions for a strict local maximum) Let x^* be feasible, satisfying the Lagrange condition for some λ^*. Assume that*

$$D_x^2 \mathcal{L}(x^*, \lambda^*) = D_x^2 f(x^*, \lambda^*) + \sum_{j=1}^{c} \lambda_j^* D^2 g^j(x^*)$$

is negative definite subject to the constraints $Dg(x^*)h = \theta$, viz., $h' D_x^2 \mathcal{L}(x^*, \lambda^*)h < 0$ for all $h \in \mathbb{R}^n$ such that $Dg(x^*)h = \theta$. Then x^* is a strict local maximum of f subject to $g(x) = \theta$.

Proof. Let h be a direction vector in $\mathbb{R}^n$ and consider a feasible point $x^* + \alpha h$ where $\alpha > 0$. By Taylor's Theorem

$$\mathcal{L}(x^* + \alpha h, \lambda^*) - \mathcal{L}(x^*, \lambda^*) = D_x \mathcal{L}(x^*, \lambda^*)(\alpha h) + \frac{\alpha^2}{2} h' D_x^2 \mathcal{L}(x^* + \theta_\alpha \alpha h, \lambda^*)h \tag{42}$$

for some $\theta_\alpha \in (0, 1)$. Writing (42) in detail, we have

$$\begin{aligned}\mathcal{L}(x^* + \alpha h, \lambda^*) - \mathcal{L}(x^*, \lambda^*) &\equiv f(x^* + \alpha h) + \lambda^{*'} g(x^* + \alpha h) - f(x^*) - \lambda^{*'} g(x*) \\ &= (Df(x^*) + \lambda^{*'} Dg(x^*))\alpha h + \frac{\alpha^2}{2} h' D_x^2 \mathcal{L}(x^* + \theta_\alpha \alpha h, \lambda^*)h.\end{aligned}$$

Since, by assumption, x^* and $x^* + \alpha h$ are feasible, we have that $g(x^*) = g(x^* + \alpha h) = \theta$ and x^* satisfies the Lagrange condition, i.e., $Df(x^*) + \lambda^{*'} Dg(x^*) = \theta$. We have,

$$f(x^* + \alpha h) - f(x^*) = \frac{\alpha^2}{2} h' D_x^2 \mathcal{L}(x^* + \theta_\alpha \alpha h, \lambda^*)h. \tag{43}$$

Now, suppose that x^* is not a strict local maximum of f, we want to show that if this is the case then $D_x^2\mathcal{L}(x^*, \lambda^*)$ cannot be negative definite subject to $Dg(x^*)h = \theta$ ($D_x^2\mathcal{L} \geq 0$).

Consider a decreasing sequence of real numbers $\delta_k > 0$ such that $\delta_k \to 0$ as $k \to \infty$, and construct balls $B(x^*, \delta_k)$ for $k = 1, 2, \cdots$ If x^* is not a strict local maximizer of $f(\cdot)$, then, for all such balls, we can choose $x^{\delta_k} \in B(x^*, \delta_k)$ such that $f(x^{\delta_k}) \geq f(x^*)$. We write the vectors x^{δ_k} in the form

$$x^{\delta_k} = x^* + \alpha_k h^k$$

where h^k is a unit norm direction vector. Observe that by construction $x^{\delta_k} \to x^*$ as $k \to \infty$, implying that $\alpha_k \to 0$. Now, $\{h^k\}$ is a bounded sequence in $\mathbb{R}^n$, so by the Bolzano -Weierstrass Theorem, the sequence $\{h^k\}$ has a convergent subsequence. For simplicity, assume that h^k itself converges to some h, then, for each $x^* + \alpha_k h^k$, we have

$$\frac{1}{\alpha_k}(g(x^* + \alpha_k h^k) - g(x^*)) = \theta \qquad (by\ feasibility).$$

Taking limits as $k \to \infty$, we have $Dg(x^*)h = \theta$. By (43) and the assumption that $f(x^{\delta_k}) \geq f(x^*)$ we have

$$\frac{\alpha_k^2}{2} h^{k'} D_x^2\mathcal{L}(x^* + \theta_k\alpha_k h^k, \lambda^*)h^k = f(x^* + \alpha_k h^k) - f(x^*) \geq 0,$$

or

$$\frac{1}{2} h^{k'} D_x^2\mathcal{L}(x^* + \theta_k\alpha_k h^k, \lambda^*)h^k = \frac{1}{\alpha_k^2}(f(x^* + \alpha_k h^k) - f(x^*)) \geq 0.$$

Taking limits, as $k \to \infty$

$$\frac{1}{2} h' D_x^2\mathcal{L}(x^*, \lambda^*)h \geq 0.$$

Hence, there exists a vector h such that $Dg(x^*)h = \theta$ and $D_x^2\mathcal{L}(x^*, \lambda^*)$ is positive semidefinite. This completes the proof. □

Theorem 125. *Let $V(\alpha) = \max_x\{f(x, \alpha);\ g(x, \alpha) = \theta\}$, where $f,\ g \in C^2$. If x^* satisfies the conditions for a strict local maximum, then V is differentiable at α_0 and*

$$DV(\alpha_0) = D_\alpha\mathcal{L}(x^*, \alpha_0) = D_\alpha f(x^*, \alpha_0) + \lambda_0' D_\alpha g(x^*, \alpha_0).$$

Proof.

$$D_x\mathcal{L}(x^*,\alpha_0) = D_x f(x^*,\alpha_0) + \lambda_0^{'} D_x g(x^*,\alpha_0) = \theta \tag{44}$$

$$D_\lambda\mathcal{L}(x^*,\alpha_0) = g(x^*,\alpha_0) = \theta \tag{45}$$

Under the assumptions of the theorem there exists $x(\alpha)$, a function which is differentiable and provides a solution for the optimization problem. Hence, we write $V(\alpha) = f(x(\alpha),\alpha)$ and

$$\begin{aligned} DV(\alpha_0) &= D_x f(x^*,\alpha_0)Dx(\alpha_0) + D_\alpha f(x^*,\alpha_0) \\ &= -\lambda_0^{'} D_x g(x^*,\alpha_0)Dx(\alpha_0) + D_\alpha f(x^*,\alpha_0). \end{aligned} \tag{46}$$

Also, substituting $x(\alpha)$ into (45) and differentiating we get

$$Dg(x(\alpha),\alpha_0) = D_x g(x^*,\alpha_0)Dx(\alpha_0) + D_\alpha g(x^*,\alpha_0) = \theta.$$

Substituting back into (46) gives

$$DV(\alpha_0) = \lambda_0^{'} D_\alpha g(x^*,\alpha_0) + D_\alpha f(x^*,\alpha_0).$$

□

19.5 Scenario 3

We now consider optimization problems where the optimal choice x^* is constrained by a weak inequality. This is what we referred to earlier as Kuhn-Tucker problems. For motivational purposes, suppose we want to maximize $f:\mathbb{R}\to\mathbb{R}$, where $f\in C^2$, $f^{(2)}(x)<0$ for all x, restricting the choice of x to the set $[0,\infty)$. Suppose, the maximum $x^*\in(0,\infty)$ and f is as depicted in figure A (see attached). In figure A, x^* is such that $f^{(1)}=0$. The fact that x^* must be in $[0,\infty)$ is of no consequence, that is the restriction is not binding.

If f is as depicted in figure B, again, x^* is such that $f^{(1)}(x^*)=0$. The restriction that $x\in[0,\infty)$ is of no consequence, not binding.

Now, suppose that f is as depicted in figure C. $f(0)$ is the maximum value of f over $[0, \infty)$ and at the maximum $x^* = 0$, $f^{(1)}(0) < 0$. So, the solution for the optimization problem can be characterized by,

$$f^{(1)}(x*) = 0, \text{ if } x^* > 0,$$

$$f^{(1)}(x*) = 0, \text{ if } x^* = 0,$$

$$f^{(1)}(x*) < 0, \text{ if } x^* = 0,$$

where in the last case the constraint is binding. This can be written as,

$$f^{(1)}(x^*) \leq 0,\ x^* \geq 0 \text{ and } f^{(1)}(x^*)x^* = 0 \text{ at least one of } x^* \text{ or } f^{(1)}(x^*) \text{ must be zero.}$$

These may be viewed as necessary conditions for a local (and global) maximum under the assumptions made on f and the permitted range of values for x.

Similarly, if $x \in \mathbb{R}^n$ we have,

$$\frac{\partial f}{\partial_i}(x^*) \leq 0,\ x_i^* \geq 0 \text{ and } \frac{\partial f}{\partial_i}(x^*)x_i^* = 0, \text{ for all } i.$$

Note that if we want to solve,

$$\max_x f(x_1, \cdots, x_n), \text{ subject to } g^1(x) \geq 0,\ g^2(x) \geq 0 \cdots g^c(x) \geq 0, \text{ and } x \geq \theta$$

we can write the maximization as

$$\max_x f(x_1, \cdots, x_n) \text{ subject to } g^1(x) - z_1 = 0,\ g^2(x) - z_2 = 0 \cdots g^c(x) - z_c = 0,$$

with $x \geq \theta$ and $z \geq \theta$. Furthermore, we could write

$$\mathcal{L}(x, z, \lambda) = f(x) + \lambda^{'}(g(x) - z).$$

Then, we set the first-order partial derivatives to zero and implicitly define x^*, z^*, λ^* such that $\frac{\partial \mathcal{L}}{\partial x}(x^*, z^*, \lambda^*) = 0$, $\frac{\partial \mathcal{L}}{\partial z}(x^*, z^*, \lambda^*) = 0$ and $\frac{\partial \mathcal{L}}{\partial \lambda}(x^*, z^*, \lambda^*) = 0$. But since it must be that $x \geq \theta$ and $z \geq \theta$, we have

$$\frac{\partial \mathcal{L}}{\partial x_i}(x^*, z^*, \lambda^*) \leq 0,\ x_i^* \geq 0,\ x_i^* \frac{\partial \mathcal{L}}{\partial x_i}(x^*, z^*, \lambda^*) = 0 \text{ for } i = 1, \cdots, n, \tag{47}$$

$$\frac{\partial \mathcal{L}}{\partial z_j}(x^*, z^*, \lambda^*) \leq 0,\ z_j^* \geq 0,\ z_j^* \frac{\partial \mathcal{L}}{\partial z_j}(x^*, z^*, \lambda^*) = 0 \text{ for } j = 1, \cdots, c, \tag{48}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_j}(x^*, z^*, \lambda^*) = g^j(x^*) - z_j^* = 0. \tag{49}$$

Note that $\frac{\partial \mathcal{L}}{\partial z_j}(x, z, \lambda) = -\lambda_j$, where $\frac{\partial \mathcal{L}}{\partial z_j}(x, z, \lambda) \leq 0$ by (48), hence $\lambda_j^* \geq 0$, $z_j^* \geq 0$ and $\lambda_j^* z_j^* = 0$. From (49) $z_j^* = g^j(x^*)$ so back to (48) we have

$$\lambda_j^* \geq 0,\ g^j(x^*) \geq 0,\ \lambda_j^* g^j(x^*) = 0.$$

We say that a constraint is biding or active at x if $g^j(x) = 0$ and a constraint is non-biding or inactive at x if $g^j(x) > 0$. Hence, the first order conditions for maximization are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_i}(x^*, z^*, \lambda^*) &= \frac{\partial f}{\partial x_i}(x^*) + \lambda' \frac{\partial g}{\partial x_i}(x^*) \\ &= \frac{\partial f}{\partial x_i}(x^*) + \sum_{j=1}^{c} \lambda_j \frac{\partial g^j}{\partial x_i}(x^*) \leq 0 \end{aligned}$$

with $x_i^* \geq 0$ and $x_i^* \frac{\partial \mathcal{L}}{\partial x_i}(x^*, z^*, \lambda^*) = 0$, $\lambda_j^* \geq 0$, $g^j(x^*) \geq 0$ and $\lambda_j^* g^j(x^*) = 0$. Note that we got rid of the the $z_j's$. These are the Kuhn-Tucker conditions. Alternatively, we can consider

$$\mathcal{L}^{KT}(x, \lambda) = f(x) + \lambda' g(x)$$

and set

$$\frac{\partial \mathcal{L}^{KT}}{\partial x_i}(x^*, \lambda^*) \leq 0,\ \frac{\partial \mathcal{L}^{KT}}{\partial \lambda_j}(x^*, \lambda^*) = g^j(x^*) \geq 0, \tag{50}$$

$$x_i^* \geq 0,\ \lambda_j^* \geq 0, \tag{51}$$

$$\frac{\partial \mathcal{L}^{KT}}{\partial x_i}(x^*, \lambda^*) x_i^* = 0,\ \frac{\partial \mathcal{L}^{KT}}{\partial \lambda_j}(x^*, \lambda^*) \lambda_j = 0. \tag{52}$$

Example 32. *Suppose we face the following maximization problem*

$$\max_{x_1, x_2} f(x_1, x_2) = -\left[(x_1 - 4)^2 + (x_2 - 4)^2\right] \textit{ subject to}$$

$$2x_1 + 3x_2 \geq 6 \Longleftrightarrow 2x_1 + 3x_2 - 6 \geq 0,\ \textit{and}\ -3x_1 - 2x_2 \geq -12 \Longleftrightarrow -3x_1 - 2x_2 + 12 \geq 0$$

with $x_1, x_2 \geq 0$. *Then, we write*

$$\mathcal{L}^{KT}(x,\lambda) = -(x_1-4)^2 - (x_2-4)^2 + \lambda_1(2x_1+3x_2-6) + \lambda_2(-3x_1-2x_2+12).$$

We must have: a) using (50)

$$\begin{aligned} \frac{\partial \mathcal{L}^{KT}}{\partial x_1} &= -2(x_1-4) + 2\lambda_1 - 3\lambda_2 \leq 0 \\ \frac{\partial \mathcal{L}^{KT}}{\partial x_2} &= -2(x_2-4) + 3\lambda_1 - 2x_2 \leq 0 \\ \frac{\partial \mathcal{L}^{KT}}{\partial \lambda_1} &= 2x_1 + 3x_2 - 6 \geq 0 \\ \frac{\partial \mathcal{L}^{KT}}{\partial \lambda_2} &= -3x_1 - 2x_2 + 12 \geq 0. \end{aligned} \tag{53}$$

b) using (51)

$$x_1, x_2 \geq 0, \lambda_1, \lambda_2 \geq 0 \tag{54}$$

c) using (52)

$$\begin{aligned} (-2(x_1-4) + 2\lambda_1 - 3\lambda_2)x_1 &= 0 \\ (-2(x_2-4) + 3\lambda_1 - 2\lambda_2)x_2 &= 0 \\ (2x_1 + 3x_2 - 6)\lambda_1 &= 0 \\ (-3x_1 + 2x_2 + 12)\lambda_2 &= 0. \end{aligned} \tag{55}$$

Note that in this case

$$f(x_{1,}x_2) = -\underbrace{[(x_1-4)^2 + (x_2-4)^2]}_{\text{always positive}},$$

and because of this it is easy to see that f *is maximized at* $(4,4)$. *Obviously, following the method to find a strict local maximum we have that*

$$\frac{\partial f}{\partial_1}(x_1,x_2) = -2(x_1-4),\ \frac{\partial f}{\partial_2}(x_1,x_2) = -2(x_2-4).$$

and setting both derivatives to zero we find $x_1^* = x_2^* = 4$. *Also,* $D^2 f(x^*) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$ *and* $\alpha' D^2 f(x^*)\alpha = (-2\alpha_1 - 2\alpha_2)\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = -2\alpha_1^2 - 2\alpha_2^2 = -2\sum_{i=1}^2 \alpha_i^2 < 0$. *Hence, we have an unconstrained maximum.*

Now, we incorporate the additional constraints, i.e., $2x_1 + 3x_2 \geq 6$, $-3x_1 - 2x_2 \geq -12$, $x_1 \geq 0$, $x_2 \geq 0$. *Then,* x_1 *and* x_2 *must be chosen from the shaded area* S *in figure KT. Note that by setting the Fréchet differential with increment* dx *to zero*

$$Df(x_1, x_2)dx = -2(x_1 - 4)dx_1 - 2(x_2 - 4)dx_2 = 0 \Rightarrow \frac{dx_2}{dx_1} = -\frac{x_1 - 4}{x_2 - 4}.$$

Setting the slope of the circles equal to $-\frac{3}{2}$, *i.e., the slope of the binding constraint we have*

$$-\left(\frac{x_1 - 4}{x_2 - 4}\right) = -\frac{3}{2} \Rightarrow 2x_1 - 3x_2 = -4,$$

which can be solved together with $3x_1 + 2x_2 = 12$ *to give* $(x_1^*, x_2^*) = (2\frac{2}{13}, 2\frac{10}{13})$.

This illustrates that fact that constraints are either binding or non-binding. If they are binding - met with equality - we can consider a Lagrangian problem. If they are non-binding, the associated optimal λ *must be zero. So, the problem can be re-stated as*

$$\mathcal{L}(x_1, x_2, \lambda_1) = -(x_1 - 4)^2 - (x_2 - 4)^2 + \lambda[-3x_1 - 2x_2 + 12]$$

with

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x_1}(x_1, x_2, \lambda_1) &= -2(x_1 - 4) - 3\lambda \\
\frac{\partial \mathcal{L}}{\partial x_2}(x_1, x_2, \lambda_1) &= -2(x_2 - 4) - 2\lambda \\
\frac{\partial \mathcal{L}}{\partial \lambda}(x_1, x_2, \lambda_1) &= -3x_1 - 2x_2 + 12.
\end{aligned}$$

Combining $\frac{\partial \mathcal{L}}{\partial x_1}, \frac{\partial \mathcal{L}}{\partial x_2} = 0$ *we have* $2(x_1^* - 4) = 3(x_2^* - 4) \Leftrightarrow 2x_1^* - 3x_2^* = -4$. *Together with* $\frac{\partial \mathcal{L}}{\partial \lambda}(x_1^*, x_2^*, \lambda^*) = 0 \Leftrightarrow 3x_1^* + 2x_2^* = 12$, *which gives exactly* $(x_1^*, x_2^*) = (2\frac{2}{13}, 2\frac{10}{13})$ *and* $\lambda^* = -(x_2^* - 4) = -(\frac{36}{13} - 4) = -(\frac{36-52}{13}) = \frac{16}{13} > 0$.

Obviously, in general if we do not know which constraints are binding, and we can maximize

$$\mathcal{L}(x_1, x_2, z_1, z_2, \lambda_1, \lambda_2) = -(x_1-4)^2-(x_2-4)^2+\lambda_1(-3x_1-2x_2+12-z_1)+\lambda_2(2x_1+3x_2-6-z_2)$$

subject to $x_1, x_2, z_1, z_2 \geq 0$. *Then,*

$$\frac{\partial \mathcal{L}}{\partial x_1} = -2(x_1^*-4) - 3\lambda_1^* + 2\lambda_2^* \leq 0,\ x_1^* \geq 0 \text{ and } x_1^* \frac{\partial \mathcal{L}}{\partial x_1}(x^*, z^*, \lambda^*) = 0 \tag{56}$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = -2(x_2^*-4) - 2\lambda_1^* + 3\lambda_2^* \leq 0, x_2^* \geq 0 \text{ and } x_2^* \frac{\partial \mathcal{L}}{\partial x_2}(x^*, z^*, \lambda^*) = 0 \tag{57}$$

$$\frac{\partial \mathcal{L}}{\partial z_1} = -\lambda_1^* \leq 0, z_1^* \geq 0 \text{ and } -\lambda_1^* z_1^* = 0 \Rightarrow \lambda_1^* \geq 0, z_1^* \geq 0, \lambda_1^* z_1^* = 0 \tag{58}$$

$$\frac{\partial \mathcal{L}}{\partial z_2} = -\lambda_2^* \leq 0, z_2^* \geq 0 \text{ and } -\lambda_2^* z_2^* = 0 \Rightarrow \lambda_2^* \geq 0, z_2^* \geq 0, \lambda_2^* z_2^* = 0 \tag{59}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_1} = -3x_1^* - 2x_2^* + 12 - z_1^* = 0 \tag{60}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_2} = 2x_1^* + 3x_2^* - 6 - z_2^* = 0 \tag{61}$$

The last two conditions, viz., (60) *and* (61) *give* $\begin{cases} z_1^* = -3x_1^* - 2x_2^* + 12 \\ z_2^* = 2x_1^* + 3x_2^* - 6 \end{cases}$ *so that the last four conditions,* (58), (59), (60) *and* (61) *can be summarized as,*

$$\lambda_1^* \geq 0, -3x_1^* - 2x_2^* + 12 \geq 0 \text{ and } \lambda_1^*(-3x_1^* - 2x_2^* + 12) = 0 \tag{62}$$

and

$$\lambda_2^* \geq 0, 2x_1^* + 3x_2^* - 6 \geq 0 \text{ and } \lambda_2^*(2x_1^* + 3x_2^* - 6) = 0 \tag{63}$$

That is,

1. *if the constraint is not binding (>) then* $\lambda_j^* = 0$
2. *if the constraint is binding (= 0) then* $\lambda_j^* > 0$ *or* $\lambda_j^* = 0$.

Note that conditions (56)-(61) *can be equivalently stated as* (56), (57), (62) *and* (63), *which do not involve* z_1^*, z_2^*. *Hence, consider*

$$\mathcal{L}^{KT}(x_1, x_2, \lambda_1, \lambda_2) = -(x_1-4)^2 - (x_2-4)^2 + \lambda_1(-3x_1 - 2x_2 + 12) + \lambda_2(2x_1 + 3x_2 - 6)$$

and

$$\frac{\partial \mathcal{L}^{KT}}{\partial x_1} = -2(x_1^* - 4) - 3\lambda_1^* + 2\lambda_2^* \le 0, x_1^* \ge 0, x_1^* \frac{\partial \mathcal{L}}{\partial x_1}(x_1^*, x_2^*, \lambda^*) = 0 \tag{64}$$

$$\frac{\partial \mathcal{L}^{KT}}{\partial x_2} = -2(x_2^* - 4) - 2\lambda_1^* + 3\lambda_2^* \le 0, x_2^* \ge 0, x_2^* \frac{\partial \mathcal{L}}{\partial x_2}(x_1^*, x_2^*, \lambda^*) = 0 \tag{65}$$

$$\frac{\partial \mathcal{L}^{KT}}{\partial \lambda_1} = (-3x_1^* - 2x_2^* + 12) \ge 0, \lambda_1^* \ge 0, \frac{\partial \mathcal{L}}{\partial \lambda_1}(x^*, \lambda^*)\lambda_1^* = 0 \tag{66}$$

$$\frac{\partial \mathcal{L}^{KT}}{\partial \lambda_2} = (2x_1^* + 3x_2^* - 6) \ge 0, \lambda_2^* \ge 0, \frac{\partial \mathcal{L}}{\partial \lambda_2}(x^*, \lambda^*)\lambda_2^* = 0 \tag{67}$$

Note that (64) and (65) are respectively (56) and (57), and (66) and (67) are respectively (62) and (63).

Remark 43. *If there is no imposition that $x_1^*, x_2^* \ge 0$ then (64) and (65) become*

$$\frac{\partial \mathcal{L}^{KT}}{\partial x_1} = -2(x_1^* - 4) - 3\lambda_1^* + 2\lambda_2^* = 0 \tag{68}$$

$$\frac{\partial \mathcal{L}^{KT}}{\partial x_2} = -2(x_2^* - 4) - 2\lambda_1^* + 3\lambda_2^* = 0 \tag{69}$$

which can be combined into

$$\begin{bmatrix} \frac{\partial \mathcal{L}^{KT}}{\partial x_1} \\ \frac{\partial \mathcal{L}^{KT}}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x_1^*, x_2^*) \\ \frac{\partial f}{\partial x_2}(x_1^*, x_2^*) \end{bmatrix} + \begin{bmatrix} \lambda_1^* & \lambda_2^* \end{bmatrix} \begin{bmatrix} -3 & -2 \\ 2 & 3 \end{bmatrix} = \theta$$

or

$$D_x \mathcal{L}^{KT}(x^*, \lambda^*) = D_x f(x^*) + \lambda^{*'} D_x g(x^*) = \theta \tag{70}$$

together with

$$D_{\lambda_j} \mathcal{L}^{KT}(x^*, \lambda^*) \ge 0, \lambda_j^* \ge 0, D_{\lambda_j} \mathcal{L}^{KT}(x^*, \lambda^*)\lambda_j^* = 0 \text{ for } j = 1, 2. \tag{71}$$

The fundamental question addressed in the Kuhn-Tucker Theorem is the existence of λ^* such that, if x^*solves the Kuhn-Tucker problem, then (70) and (71) are met.

Now that we have set up the motivation for our problem and considered a comprehensive example, we will consider the general problem

$$\max_x \{f(x, \alpha) : \ x \in c(\alpha)\}, \text{ such that } c(\alpha) = \{x \in \mathbb{X} : \ g(x, \alpha) \ge 0\} \tag{72}$$

where $f : \mathbb{X} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{X} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^c$ and $f, g \in C^2$.

- An inequality constraint $g^j(x)$ is active (binding) if it holds with equality ($g^j(x) = 0$) and inactive (not binding) if it is met with strict inequality, where x is a feasible point.

- Intuitively, only binding constraints matter in the optimization, that is, only binding constraints impact the local properties of a solution. If we know the active constraints, we could ignore the inactive ones and solve a Lagrange problem with $g^j(x, \alpha) = 0$.

- As in the Lagrange case consider if there is no requirement that $x^* \geq 0$

$$\mathcal{L}^{KT}(x, \lambda) = f(x) + \lambda' g(x)$$

and

$$D_x \mathcal{L}^{KT}(x, \lambda) = Df(x) + \lambda' Dg(x) = \theta$$

$$D_{\lambda_j} \mathcal{L}^{KT}(x, \lambda) = g^j(x) \geq 0 \qquad and \qquad g^j(x) = 0 \quad if\ \lambda_j > 0$$

$$\lambda_j \geq 0 \quad and \quad \lambda_j = 0 \quad if\ g^j(x) > 0 \quad (inactive),$$

or

$$D_{\lambda_j} \mathcal{L}^{KT}(x^*, \lambda^*) \geq 0, \quad \lambda_j^* \geq 0, \quad D_{\lambda_j} \mathcal{L}^{KT}(x^*, \lambda^*) \lambda_j^* = 0.$$

- In what follows we adopt the following notation

$$g^j(x^*) = 0 \quad for \quad j = 1, 2, ..., B \quad (binding).$$

$$g^j(x^*) > 0 \quad for \quad j = B+1, ..., c \quad (non-binding).$$

and define,

$$g^b(x^*) = [g^1(x^*), ..., g^B(x^*)]', \; g^n(x^*) = [g^{B+1}(x^*), ..., g^C(x^*)]'$$

so that

$$g(x^*) = \begin{bmatrix} g^b(x^*) \\ g^n(x^*) \end{bmatrix}.$$

We partition the multipliers in a similar way, that is, $\lambda = (\lambda_b', \lambda_n')'$. Hence, the Lagrangian condition $(Df(x^*) + \lambda^{*'} Dg(x^*) = \theta)$ becomes

$$Df(x^*) + \lambda_b^{*'} g^b(x) + \lambda_n^{*'} g^n(x) = \theta. \tag{73}$$

Theorem 126. *(Kuhn-Tucker) Let x^* be a solution for the problem*

$$\max_x f(x, \bar{\alpha}) \text{ subject to } g(x, \bar{\alpha}) \geq \theta,\ x \in \mathbb{R}^n, \bar{\alpha} \in \mathbb{R}^p,\ g \in \mathbb{R}^c$$

where $f, g \in C^1$, $rank(D_\alpha g^b(x^*)) = B \leq n$. Then, there exists non-negative multipliers $\lambda^* \in \mathbb{R}^c$, such that (x^*, λ^*) satisfy

$$Df(x^*) + \lambda^{*'} Dg(x^*) = \theta \tag{74}$$

for all $j = 1, 2, ..., c$, $g^j(x^*) \geq 0$, $\lambda_j^* \geq 0$ and $g^j(x^*) = 0$ if $\lambda_j^* > 0$ and $\lambda_j^* = 0$ if $g^j(x^*) > 0$.

Proof. Let $x = \begin{bmatrix} x_\alpha \\ x_\beta \end{bmatrix}$ such that $x_\alpha = \begin{pmatrix} x_1 \\ \vdots \\ x_B \end{pmatrix}$, $x_\beta = \begin{pmatrix} x_{B+1} \\ \vdots \\ x_n \end{pmatrix}$ and

$$Dg^b(x^*) = \begin{bmatrix} D_\alpha g^b(x^*) & D_\beta g^b(x^*) \end{bmatrix}.$$

Consequently, we can write (74) as

$$\begin{bmatrix} \underset{1\times B}{D_\alpha f(x^*)} & \underset{1\times(n-B)}{D_\beta f(x^*)} \end{bmatrix} + \begin{pmatrix} \underset{1\times B}{\lambda_b^{*'}} & \underset{1\times c-B}{\lambda_n^{*'}} \end{pmatrix} \begin{bmatrix} \underset{B\times B}{D_\alpha g^b(x^*)} & \underset{B\times n-B}{D_\beta g^b(x^*)} \\ \underset{c-B\times n}{Dg^n(x^*)} & \end{bmatrix} = \underset{1\times n}{\theta}.$$

Now, we define the multipliers as $\lambda_n^* = \theta$ for inactive constraints. So,

$$\begin{bmatrix} \underset{1\times B}{D_\alpha f(x^*)} & \underset{1\times(n-B)}{D_\beta f(x^*)} \end{bmatrix} + \begin{pmatrix} \underset{1\times B}{\lambda_b^{*'}} \underset{B\times B}{D_\alpha g^b(x^*)} & \underset{1\times B}{\lambda_b^{*'}} \underset{B\times(n-B)}{D_\beta g^b(x^*)} \end{pmatrix} = \underset{1\times n}{\theta}$$

and we have for the first B equations

$$\underset{1\times B}{D_\alpha f(x^*)} + \underset{1\times B}{\lambda_b^*} \underset{B\times B}{D_\alpha g^b(x^*)} = \underset{1\times B}{\theta}.$$

By the assumption that rank $(D_\alpha g^b(x^*)) = B$ we have

$$\underset{1\times B}{\lambda_b^{*'}} = -\underset{1\times B}{D_\alpha f(x^*)}\underset{B\times B}{\big(D_\alpha g^b(x^*)\big)^{-1}}.$$

It remains to show that

$$\underset{1\times(n-B)}{D_\beta f(x^*)} + \underset{1\times B}{\lambda_b^{*'}}\underset{B\times(n-B)}{D_\beta g^b(x^*)} = \underset{1\times(n-B)}{\theta} \quad \textit{and} \quad \lambda_b^{*'} \geq \theta.$$

(i) We introduce "slack variables" $z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_c \end{bmatrix} = \begin{bmatrix} \underset{B\times 1}{z_b} \\ \underset{c-B\times 1}{z_n} \end{bmatrix}$. $g(x) \geq 0$ can be written as $g(x) - z = \theta$ and $z \geq 0$. Now, by continuity of the constraints (g^j) there exists $\epsilon > 0$ such that for $x \in B(x^*, \epsilon)$ we have $g^j(x) > 0$ (non-binding, that is z_n is strictly greater than the null vector) for all $j = B+1, \ldots, c$. As long as we stay in this ball we can ignore inactive constraints. The active constraints can be written as $G(x, z_b) = \underset{B\times 1}{g^b(x)} - \underset{B\times 1}{Iz_b} = \underset{B\times 1}{\theta}$ and $z_b \geq 0$. If x^* solves (72), then it also solves

$$\max_x\{f(x,\alpha), G(x,z_b) = \theta, z_b \geq 0, x \in B(x^*, \epsilon)\}. \tag{75}$$

We know that binding the constraints satisfy $G(x^*, z_b^*) = \theta$ with $z_b^* = \theta$. Differentiating G with respect to x_α, we have $D_\alpha G(x^*, z_b^*) = D_\alpha G(x^*, \theta) = D_\alpha g^b(x^*)$ and $D_\alpha g^b(x^*)$ is invertible. Hence, the assumptions of the Implicit Function Theorem hold for $G(x, z_b)$ at $(x^*, z_b^*) = (x_\alpha^*, x_\beta^*, \theta)$ and the system $G(x_\alpha^*, x_\beta^*, z_b^*) = \theta$ implicitly defines x_α^* as a function h of x_β and z_b. That is, there exists $h : U_{\beta,z} \to U_\alpha$ with $h(x_\beta, z_b) = x_\alpha$ such that

$$G(h(x_\beta, z_b), x_\beta, z_b) = g^b(h(x_\beta, z_b), x_\beta) - Iz_b = \theta \tag{76}$$

where U_α and $U_{\beta,z}$ are open balls centered respectively at x_α^* and (x_β^*, θ). Differentiation of (76) yields

$$D_\alpha g^b(x^*) D_\beta h(x_\beta^*, z_\beta^*) + D_\beta g^b(x^*) = \theta,$$

from where

$$D_\beta h(x^*_\beta, \theta) = -(D_\alpha g^b(x^*))^{-1} D_\beta g^b(x^*) \tag{77}$$

and differentiating (76) with respect to z_b

$$D_\alpha g^b(x^*) D_z h(x^*_\beta, z^*_b) - I = \theta.$$

Implying that $D_z h(x^*_\beta, \theta) = (D_\alpha g^b(x^*))^{-1}$.

Now, define $F(x_\beta, z_b) = f(\underbrace{h(x_\beta, z_b)}_{B\times 1}, x_\beta)$ and observe that the pair $(x^*_\beta, z^*_b) = (x^*_\beta, \theta)$ will be an optimal solution for

$$\max_{x_\beta, z_b}\{F(x_\beta, z_b);\ z_b \geq \theta \quad (x_\beta, z_b) \in U_{\beta,z}\}$$

and therefore will satisfy

$$D_\beta F(x^*_\beta, z^*_b) = \theta \qquad \mathit{and} \qquad D_z F(x^*_\beta, z^*_b) \leq \theta \tag{78}$$

That is, for each $i = 1, 2, \cdots, B$, we have

$$D_{z_i} F(x^*_\beta, z^*_b) \leq 0 \text{ and } D_{z_i} F(x^*_\beta, z^*_b) = 0 \text{ if } z_i > 0.$$

$$z_i \geq 0 \text{ and } z_i = 0 \text{ if } D_{z_i} F(x^*_\beta,\ z^*_b) < 0$$

Then, from (78)

$$\begin{aligned} D_\beta F(x^*_\beta, z^*_b) &= D_\beta f(h(x^*_\beta, z^*_b), x^*_\beta) = \theta \\ &= D_\alpha f(x^*) D_\beta h(x^*_\beta, z^*_b) + D_\beta f(x^*) = \theta \\ &= -D_\alpha f(x^*)(D_\alpha g^b(x^*))^{-1} D_\beta g^b(x^*) + D_\beta f(x^*) = \theta \text{ from (77)} \\ &= \lambda^{*'}_b D_\beta g^b(x^*) + D_\beta f(x^*) = \theta \text{ from the expression for } \lambda^*_b \end{aligned}$$

Finally, $D_z F(x^*_\beta, z^*_b) \leq \theta$ gives

$$D_z f(h(x^*_\beta, z^*_b), x^*_\beta) = D_\alpha f(x^*) D_z h(x^*_\beta, z^*_b) = D_\alpha f(x^*)(D_\alpha g^b(x^*))^{-1} = -\lambda^{*'}_b \leq \theta$$

or $\lambda^{*'}_b \geq \theta$. $\square$

We now state two theorems without proof. The proof of Theorem 127 is similar to that of Theorem 123. We leave the proof of Theorem 128 as an exercise.

Theorem 127. *(Sufficient conditions for a global maximum). Given the problem in (72), assume that f is pseudo-concave and that the constraints $g^j(x)$ are all quasi-concave. Let (x^*, λ^*) satisfy the conditions in the Kuhn-Tucker Theorem. Then, x^* is a solution for (72).*

Theorem 128. *(Uniqueness) Let x^* be a solution for (72). If f is strictly quasiconcave and the constraint functions $g^j(.)$ are quasiconcave, then x^* is the only solution for (72).*

20 The proper Riemann integral

The proper Riemann integral is defined for bounded functions on compact intervals of $\mathbb{R}$. We start with definitions associated with a partition of a compact interval $[a, b] \subset \mathbb{R}$ with a, $b \in \mathbb{R}$.

Definition 78.

1. *Let $[a, b] \subset \mathbb{R}$ and $a = t_0 < t_1 < \cdots < t_{k(P)-1} < t_{k(P)} = b$ where $k(P) \in \mathbb{N}$. The set $P = \{t_i\}_{i=0}^{k(P)}$ is called a finite partition of $[a, b]$.*

2. *If P and P' are partitions of $[a, b]$ and $P \subset P'$, we say that P' is finer than P or a refinement of P.*

3. *The mesh of P is $m(P) := \max_{1 \le i \le k(P)} (t_i - t_{i-1})$.*

If $f : [a, b] \to \mathbb{R}$ is bounded and $P = \{t_j\}_{j=0}^{k(P)}$ is a partition of $[a, b]$, let

$$m_j^f := \inf_{t_{j-1} \le x \le t_j} f(x) \text{ and } M_j^f := \sup_{t_{j-1} \le x \le t_j} f(x) \text{ for } j = 1, \ldots, k(P).$$

Then, the lower and upper Darboux sums are defined by

$$S_P[f] := \sum_{j=1}^{k(P)} m_j^f (t_j - t_{j-1}) \text{ and } S^P[f] := \sum_{j=1}^{k(P)} M_j^f (t_j - t_{j-1}). \tag{79}$$

We note that if $h := f + g$, where $f, g : [a, b] \to \mathbb{R}$ are bounded, then from the definition of supremum,

$$\begin{aligned} S^P[h] &:= \sum_{j=1}^{k(P)} \sup_{t_{j-1} \le x \le t_j} h(x)(t_j - t_{j-1}) \le \sum_{j=1}^{k(P)} \left(\sup_{t_{j-1} \le x \le t_j} f(x) + \sup_{t_{j-1} \le x \le t_j} g(x) \right) (t_j - t_{j-1}) \\ &= S^P[f] + S^P[g]. \end{aligned}$$

In addition, since $\sup\limits_{t_{j-1} \le x \le t_j} |f(x)| = \max \left\{ \sup\limits_{t_{j-1} \le x \le t_j} f(x), \ -\inf\limits_{t_{j-1} \le x \le t_j} f(x) \right\}$ we have that either $\sup\limits_{t_{j-1} \le x \le t_j} |f(x)| = \sup\limits_{t_{j-1} \le x \le t_j} f(x)$, which implies $\sup\limits_{t_{j-1} \le x \le t_j} |f(x)| = \left| \sup\limits_{t_{j-1} \le x \le t_j} f(x) \right|$, or $\sup\limits_{t_{j-1} \le x \le t_j} |f(x)| = -\inf\limits_{t_{j-1} \le x \le t_j} f(x)$. In this case,

$$-\sup_{t_{j-1} \le x \le t_j} |f(x)| = \inf_{t_{j-1} \le x \le t_j} f(x) \le \sup_{t_{j-1} \le x \le t_j} f(x) \le \sup_{t_{j-1} \le x \le t_j} |f(x)|,$$

where the first inequality follows from the definition of infimum and supremum and the second inequality follows from $f(x) \le |f(x)|$. Hence, we have that $\left| \sup\limits_{t_{j-1} \le x \le t_j} f(x) \right| \le \sup\limits_{t_{j-1} \le x \le t_j} |f(x)|$. Thus,

$$\begin{aligned} \left| S^P[f] \right| &\le \sum_{j=1}^{k(P)} |M_j^f|(t_j - t_{j-1}) \le \sum_{j=1}^{k(P)} \sup_{t_{j-1} \le x \le t_j} |f(x)|(t_j - t_{j-1}) \\ &= S^P[|f|] \le C \sum_{j=1}^{k(P)} (t_j - t_{j-1}) = C(b - a). \end{aligned}$$

Similarly, $|S_P[f]| \le S_P[|f|] \le C(b - a)$.

Theorem 129. *Let P and P' be partitions of $[a, b] \subset \mathbb{R}$ such that $P \subset P'$ and $f : [a, b] \to \mathbb{R}$ be bounded. Then,*

$$S_P[f] \le S_{P'}[f] \le S^{P'}[f] \le S^P[f].$$

Proof. Let $P = \{t_i\}_{i=0}^{k(P)}$ and $P' = P \bigcup \{t\}$ where, without loss of generality, $t_0 < t < t_1$. Let $m_{t,1}^f = \inf\limits_{t \le x \le t_1} f(x)$, $m_{0,t}^f = \inf\limits_{0 \le x \le t} f(x)$ and note that $m_{0,t}^f, m_{t,1}^f \ge m_1^f$. Then, $(t_1 - t)m_{t,1}^f \ge (t_1 - t)m_1^f$ and $(t - t_0)m_{0,t}^f \ge (t - t_0)m_1^f$. Consequently,

$$(t_1 - t)m_{t,1}^f + (t - t_0)m_{0,t}^f \ge (t_1 - t)m_1^f + (t - t_0)m_1^f = (t_1 - t_0)m_1^f.$$

Thus, $S_P[f] = (t_1 - t_0)m_1^f + \sum_{j=2}^{k(P)} m_j^f(t_j - t_{j-1}) \leq (t_1 - t)m_{t,1}^f + (t - t_0)m_{0,t}^f + \sum_{j=2}^{k(P)} m_j^f(t_j - t_{j-1}) = S_{P'}[f]$.

Now, $S_{P'}[f] \leq S^{P'}[f]$ follows by definition and $S^{P'}[f] \leq S^P[f]$ follows by the same argument used for lower Darboux sums. $\square$

It follows from Theorem 129 that for any two partitions P_1 and P_2 we have that $S_{P_1}[f] \leq S^{P_2}[f]$. To see this, let $P = P_1 \bigcup P_2$, then

$$S_{P_1}[f] \leq S_P[f] \leq S^P[f] \leq S^{P_2}[f] \tag{80}$$

We now define lower and upper Riemann integrals.

Definition 79. *Let $f : [a, b] \to \mathbb{R}$ be bounded. The lower and upper Riemann integrals of f over $[a, b]$ are defined as*

$$\underline{\int_a^b} f := \sup_P S_P[f] \text{ and } \overline{\int_a^b} f := \inf_P S^P[f], \tag{81}$$

where the supremum and the infimum are taken over all finite partitions of $[a, b]$.

Theorem 130. $\underline{\int_a^b} f \leq \overline{\int_a^b} f$ *and* $\underline{\int_a^b} f = -\overline{\int_a^b}(-f)$.

Proof. Since $\overline{\int_a^b} f$ is an infimum, for any $\epsilon > 0$ there exists a partition P_1 such that $S^{P_1}[f] < \overline{\int_a^b} f + \epsilon$. But since for any P, $S_P[f] \leq S^{P_1}[f]$ we have that $S_P[f] \leq \overline{\int_a^b} f + \epsilon$. Then, since ϵ is arbitrary $\sup\limits_P S_P[f] := \underline{\int_a^b} f \leq \overline{\int_a^b} f$.

$\overline{\int_a^b}(-f) = \inf\limits_P S^P[-f]$. But

$$S^P[-f] = \sum_{j=1}^{k(P)} \sup_{t_{j-1} \leq x \leq t_j} (-f(x))(t_j - t_{j-1}) = -\sum_{j=1}^{k(P)} \inf_{t_{j-1} \leq x \leq t_j} f(x)(t_j - t_{j-1}).$$

Consequently,

$$\sup_P \{-S^P[-f]\} = \sup_P \left\{ \sum_{j=1}^{k(P)} \inf_{t_{j-1} \leq x \leq t_j} f(x)(t_j - t_{j-1}) \right\} := \underline{\int_a^b} f.$$

But $\sup\limits_P \{-S^P[-f]\} = -\inf\limits_P \{S^P[-f]\} = -\overline{\int_a^b}(-f)$, which completes the proof. $\square$

We are now ready to define the Riemann integral.

Definition 80. *A bounded function $f : [a, b] \to \mathbb{R}$ is said to be Riemann integrable if its upper and lower integrals are the same. The common value is denoted by*

$$\int_a^b f(x)dx := \overline{\int_a^b} f = \underline{\int_a^b} f$$

and called the Riemann integral of f over the interval $[a, b]$. The collection of all Riemann integrable functions on $[a, b]$ will be denoted $\mathcal{R}[a, b]$.

It is easy to find bounded functions on $[a, b]$ that are not in $\mathcal{R}[a, b]$. For example, let

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \bigcap [0, 1] \\ 0 & \text{if } x \notin \mathbb{Q} \bigcap [0, 1] \end{cases}.$$

Note that $S_P[f] = 0$ and $S^P[f] = 1$ for all P. Consequently, $0 = \underline{\int_a^b} f \neq \overline{\int_a^b} f = 1$. The function f is called the Dirichlet jump function. The next theorem provides a characterization of Riemann integrable functions on $[a, b]$.

Theorem 131. *Let $f : [a, b] \to \mathbb{R}$ be bounded. The following statements are equivalent:*

1. *$f \in \mathcal{R}[a, b]$*
2. *For all $\epsilon > 0$ there exists a partition P such that $S^P[f] - S_P[f] \leq \epsilon$.*
3. *For all $\epsilon > 0$ there exists some $\delta > 0$ such that $S^P[f] - S_P[f] \leq \epsilon$ for all partitions P with mesh $m(P) < \delta$.*
4. *The limit $I = \lim\limits_{m(P)\to 0} \sum_{j=1}^{k(P)} f(x_j)(t_j - t_{j-1})$ exists for any $x_j \in [t_{j-1}, t_j]$.*

Proof. ($1 \implies 2$) By definition of supremum and infimum, we have that for any partitions P_1 and P_2

$$S^{P_2}[f] - \overline{\int_a^b} f \leq \epsilon/2 \text{ and } \underline{\int_a^b} f - S_{P_1}[f] \leq \epsilon/2.$$

But by 1. we have that $\overline{\int_a^b} f = \underline{\int_a^b} f$ and consequently

$$S^{P_2}[f] - S_{P_1}[f] \leq \epsilon.$$

Letting $P = P_1 \bigcup P_2$, by equation (80) we have $S^P[f] - S_P[f] \leq \epsilon$.

(2 $\implies$ 3) We fix $\epsilon > 0$ and denote by $P_\epsilon := \{a = t_0^\epsilon, t_1^\epsilon, \ldots, b = t_{k(P_\epsilon)}^\epsilon\}$ a partition of $[a, b]$ that satisfies item 2. Let $P = \{a = t_0, t_1, \ldots, b = t_{k(P)}\}$ be a partition of $[a, b]$ such that $m(P) < \delta$ and $\delta < \min\left\{\frac{1}{2} \min_{1 \leq j \leq k(P_\epsilon)} (t_j - t_{j-1}), \frac{\epsilon}{4k(P_\epsilon)\|f\|_\infty}\right\}$ where $\|f\|_\infty = \sup_{x \in [a,b]} |f(x)|$. Now, write

$$\begin{aligned} S^P[f] - S_P[f] = \sum_{j=1}^{k(P)} (M_j^P - m_j^P)(t_j - t_{j-1}) &= \sum_{j: P_\epsilon \bigcap [t_{j-1}, t_j] \neq \emptyset} (M_j^P - m_j^P)(t_j - t_{j-1}) \\ &+ \sum_{j: P_\epsilon \bigcap [t_{j-1}, t_j] = \emptyset} (M_j^P - m_j^P)(t_j - t_{j-1}) \end{aligned} \tag{82}$$

and note that by definition $t_j - t_{j-1} \leq m(P) < \delta$ for all $j = 1, \cdots, k(P)$. In particular,

$$t_1 - a \leq m(P) < \delta. \tag{83}$$

Also, $P_\epsilon \bigcap [t_0, t_1] = \{a\}$. To see this suppose that $t_1^\epsilon \leq t_1$ so that $P_\epsilon \bigcap [t_0, t_1] = \{a, t_1^\epsilon\}$. Then,

$$t_1 - a < \delta < \frac{1}{2} \min_{1 \leq j \leq k(P_\epsilon)} (t_j^\epsilon - t_{j-1}^\epsilon) < t_1^\epsilon - a, \tag{84}$$

which implies $t_1 < t_1^\epsilon$, contradicting $t_1^\epsilon \leq t_1$. Similarly, $P_\epsilon \bigcap [t_{k(P)-1}, t_{k(P)} = b] = \{b\}$. All other t_j^ϵ appear in exactly one, or at most two intervals of the type $[t_{j-1}, t_j]$. Hence, the first sum on the right side of (82) has at most $2k(P^\epsilon)$ terms. Thus,

$$\sum_{j: P_\epsilon \bigcap [t_{j-1}, t_j] \neq \emptyset} (M_j^P - m_j^P)(t_j - t_{j-1}) \leq 2\|f\|_\infty \sum_{j: P_\epsilon \bigcap [t_{j-1}, t_j] \neq \emptyset} (t_j - t_{j-1}) \leq 2\|f\|_\infty 2k(P^\epsilon)\delta < \epsilon.$$

For the second sum in (82), note that we consider only the j such that $P_\epsilon \bigcap [t_{j-1}, t_j] = \emptyset$, so it must be that $[t_{j-1}, t_j] \subset [t^\epsilon_{j-1}, t^\epsilon_j]$ and we write

$$\begin{aligned}
\sum_{j:P_\epsilon \bigcap [t_{j-1},t_j]=\emptyset} (M_j^P - m_j^P)(t_j - t_{j-1}) &= \sum_{j=1}^{k(P)} \sum_{i:[t_{i-1},t_i]\subset[t^\epsilon_{j-1},t^\epsilon_j]} (M_i^P - m_i^P)(t_i - t_{i-1}) \\
&\leq \sum_{j=1}^{k(P)} \sum_{i:[t_{i-1},t_i]\subset[t^\epsilon_{j-1},t^\epsilon_j]} (M_j^{P_\epsilon} - m_j^{P_\epsilon})(t_i - t_{i-1}) \\
&\leq \sum_{j=1}^{k(P)} (M_j^{P_\epsilon} - m_j^{P_\epsilon})(t^\epsilon_i - t^\epsilon_{i-1}) = S^{P_\epsilon}[f] - S_{P_\epsilon}[f] \leq \epsilon.
\end{aligned}$$

Combining the bounds on both sums we have $S^P[f] - S_P[f] \leq 2\epsilon$.

($3 \implies 4$) Fix $\epsilon > 0$ and choose P with $m(P) < \delta$ such that $S^P[f] - S_P[f] \leq \epsilon$. Then, for $x_j \in [t_{j-1}, t_j]$

$$S^P[f] - \epsilon \leq S_P[f] \leq \sum_{j=1}^{k(P)} f(x_j)(t_j - t_{j-1}) \leq S^P[f] \leq S_P[f] + \epsilon.$$

Taking, $\sup\limits_P$ we have

$$\sup_P S_P[f] := \underline{\int_a^b} f \leq \sum_{j=1}^{k(P)} f(x_j)(t_j - t_{j-1}) \leq \sup_P S_P[f] := \underline{\int_a^b} f + \epsilon$$

and taking $\inf\limits_P$ we have

$$\inf_P S^P[f] := \overline{\int_a^b} f - \epsilon \leq \sum_{j=1}^{k(P)} f(x_j)(t_j - t_{j-1}) \leq \inf_P S^P[f] := \overline{\int_a^b} f.$$

Thus, $\epsilon \leq \sum_{j=1}^{k(P)} f(x_j)(t_j - t_{j-1}) \to \int_a^b f(x)dx$ as $m(P) \to 0$.

($4 \implies 1$) Assuming that that the limit in 4. exists, we have that for any $\epsilon > 0$ there exists some partition P with $m(P) < \delta$ such that

$$I - \epsilon \leq \sum_{j=1}^{k(P)} f(x_j)(t_j - t_{j-1}) \leq I + \epsilon$$

for any value $x_j \in [t_{j-1}, t_j]$. Since these inequalities hold uniformly for any intermediate value, we can write

$$I - \epsilon \leq \sum_{j=1}^{k(P)} \inf_{x\in[t_{j-1},t_j]} f(x)(t_j - t_{j-1}) \leq \sum_{j=1}^{k(P)} \sup_{x\in[t_{j-1},t_j]} f(x)(t_j - t_{j-1}) \leq I + \epsilon.$$

Thus, $I - \epsilon \le S_P[f] \le S^P[f] \le I + \epsilon$ and

$$I - \epsilon \le S_P[f] \le \underline{\int_a^b} f \le \overline{\int_a^b} f \le S^P[f] \le I + \epsilon,$$

which completes the proof. $\square$

It follows from Theorem 131 that if $f \in \mathcal{R}[a,b]$ and $P_n = \{t_{i,n}\}_{i=0}^{k_n}$ for $n \in \mathbb{N}$ is any partition of $[a,b]$ with $M(P_n) \to 0$ as $n \to \infty$, then for any $x_{i,n} \in [t_{i-1,n}, t_{i,n}]$

$$\sum_{i=1}^{k_n} f(x_{i,n})(t_{i,n} - t_{i-1,n}) \to \int_a^b f(x)dx \text{ as } n \to \infty.$$

The sum on the left of the convergence statement is called a Riemann sum. It should be noted that the convergence of Riemann sums for some partition P_n does not guarantee the existence of the integral. As an example, consider the Dirichlet jump function, which is not Riemann integrable on $[0,1]$, and the partition $P_n = \{i/k_n\}_{i=0}^{k_n}$. Then,

$$\sum_{i=1}^{k_n} f(x_{i,n}) \left(\frac{i}{k_n} - \frac{i-1}{k_n} \right) = \frac{1}{k_n} \sum_{i=1}^{k_n} f(x_{i,n}).$$

Suppose $n \le k_n$ and $x_{i,n} \in \mathbb{Q}$ for $i = 1, \cdots, n$ and $x_{i,n} \notin \mathbb{Q}$ for $i = n+1, \cdots, k_n$. Then,

$$\lim_{n\to\infty} \sum_{i=1}^{k_n} f(x_{i,n}) \left(\frac{i}{k_n} - \frac{i-1}{k_n} \right) = \lim_{n\to\infty} \frac{n}{k_n}$$

which can be any number in $[0,1]$. For example, take $k_n = an$ where $a > 1$, then $\lim\limits_{n\to\infty} \frac{n}{k_n} = a^{-1}$.

Definition 81. *$f : [a,b] \to \mathbb{R}$ is called a step function on $[a,b]$ if*

$$f(x) = \sum_{i=1}^{N} y_i I_{I_i}(x)$$

where $N \in \mathbb{N}$, $y_i \in \mathbb{R}$ and $\{I_i\}_{i\in\mathbb{N}}$ is a collection of adjacent intervals such that $\bigcup_{i=1}^N I_i = [a,b]$ and $I_i \bigcap I_j$ is at most a singleton. The intervals I_i may be open, half-open, closed or degenerate.[7]

[7] A degenerate interval is a singleton, i.e., $[a,a] = \{a\}$.

The next theorem shows that step, continuous and monotone functions on $[a, b]$ are Riemann integrable.

Theorem 132. *Let $f : [a, b] \to \mathbb{R}$. If f is a step, continuous or monotone functions on $[a, b]$ then $f \in \mathcal{R}[a, b]$.*

Proof. Suppose that f is a step function, i.e., $f(x) = \sum_{i=1}^{N} y_i I_{I_i}(x)$. The endpoints of the the intervals I_i form a partition $P = \{t_0 = a, t_1, \cdots, t_N = b\}$ of $[a, b]$. For every $\epsilon > 0$ define the partition $P_\epsilon = \{a = s_0 < s_1 < s_1' < s_2 < s_2' < \cdots < s_{N-1} < s_{N-1}' < s_N = b\}$, where $s_i < t_i < s_i'$ for $i = 1, \cdots, N-1$ and $s_i' - s_i < \frac{\epsilon}{2(N-1)\|f\|_\infty}$. Note that f is a constant on each interval $[s_{i-1}', s_i]$ for $i = 1, \cdots, N$. Hence, we have that

$$S^{P_\epsilon}[f] - S_{P_\epsilon}[f] = \sum_{i=1}^{N-1} \left(\sup_{x\in[s_i', s_i]} f(x) - \inf_{x\in[s_i', s_i]} f(x) \right) (s_i' - s_i) \le \sum_{i=1}^{N-1} 2\|f\|_\infty \frac{\epsilon}{2(N-1)\|f\|_\infty} = \epsilon.$$

Hence, by Theorem 131, $f \in \mathcal{R}[a, b]$.

Suppose f is continuous on $[a, b]$. Since, $[a, b]$ is compact, f is uniformly continuous on $[a, b]$. Then, for any $\epsilon > 0$ there exists $\delta > 0$ such that for any pair $x, y \in [a, b]$ such that $|x - y| < \delta$ we have $|f(x) - f(y)| < \epsilon$. For every partition P such that $m(P) < \delta$ we have $t_j - t_{j-1} < \delta$ and

$$S^P[f] - S_P[f] = \sum_{j=1}^{k(P)} (M_j^f - m_j^f)(t_j - t_{j-1}) < \epsilon \sum_{j=1}^{k(P)} (t_j - t_{j-1}) = \epsilon(b - a),$$

since $M_j^f - m_j^f = \underset{x,y\in[t_{j-1},t_j]}{|} f(y) - f(x)|$. Thus, by Theorem 131.3 $f \in \mathcal{R}[a, b]$.

Without loss of generality suppose f is monotone increasing on $[a, b]$. Let $t_j = a + j\frac{b-a}{k(P)} \in P$ for $j = 0, \ldots, k(P)$ and note that

$$S^P[f] - S_P[f] = \sum_{j=1}^{k(P)} (M_j - m_j)(t_j - t_{j-1}) = \frac{b-a}{k(P)} \sum_{j=1}^{k(P)} (f(t_j) - f(t_{j-1})) = \frac{b-a}{k(P)}(f(b) - f(a)).$$

The last term can be made arbitrarily small by letting $k(P) \to \infty$. Hence, by Theorem 131.2 $f \in \mathcal{R}[a, b]$. $\square$

Theorem 133. *Let $f, g \in \mathcal{R}[a,b]$ and α, $\beta \in \mathbb{R}$. Then,*

1. $\alpha f + \beta g \in \mathcal{R}[a,b]$,
2. $f \leq g \implies \int_a^b f(x)dx \leq \int_a^b g(x)dx$,
3. $|f|$, $\max\{f,g\}$, $\min\{f,g\}$, f^+, $f^- \in \mathcal{R}[a,b]$ *and* $\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx$,
4. f^2, $fg \in \mathcal{R}[a,b]$.

Proof. 1. Let P_1 and P_2 be partitions of $[a,b]$ and $P = P_1 \bigcup P_2$. For $x_j \in [t_{j-1}, t_j]$ let

$$\sum_{j=1}^{k(P)} (\alpha f(x_j) + \beta g(x_j))(t_j - t_{j-1}) \quad = \quad \alpha \sum_{j=1}^{k(P)} f(x_j)(t_j - t_{j-1}) + \beta \sum_{j=1}^{k(P)} g(x_j)(t_j - t_{j-1}).$$

Since $f, g \in \mathcal{R}[a,b]$ for any $\epsilon > 0$ if $m(P_1) < \delta_1$ and $m(P_2) < \delta_2$ then

$$\left|\sum_{j=1}^{k(P_1)} f(x_j)(t_j - t_{j-1}) - \int_a^b f(x)dx\right| < \epsilon$$

and

$$\left|\sum_{j=1}^{k(P_2)} g(x_j)(t_j - t_{j-1}) - \int_a^b g(x)dx\right| < \epsilon.$$

Since $m(P) \leq \min\{m(P_1), m(P_2)\}$

$$\begin{aligned} &\left|\sum_{j=1}^{k(P)} (\alpha f(x_j) + \beta g(x_j))(t_j - t_{j-1}) - \alpha \int_a^b f(x)dx - \beta \int_a^b g(x)dx\right| \\ \leq \quad & |\alpha| \left|\sum_{j=1}^{k(P)} f(x_j)(t_j - t_{j-1}) - \int_a^b f(x)dx\right| + |\beta| \left|\sum_{j=1}^{k(P)} g(x_j)(t_j - t_{j-1}) - \int_a^b g(x)dx\right| \\ \leq \quad & (|\alpha| + |\beta|)\epsilon. \end{aligned}$$

2. $f \leq g$ implies that $v = g - f \geq 0$ and from item 1, $0 \leq v \in \mathcal{R}[a,b]$. $0 \leq \underline{\int_a^b} v = \int_a^b v(x)dx$.

3. Let $M_j^f = \sup_{x \in [t_{j-1}, t_j]} f(x)$ and $m_j^f = \inf_{x \in [t_{j-1}, t_j]} f(x)$. Then,

$$M_i^f - m_i^f = \sup_{x \in [t_{j-1}, t_j]} f(x) - \inf_{x \in [t_{j-1}, t_j]} f(x) = \sup_{x, y \in [t_{j-1}, t_j]} (f(x) - f(y)).$$

It follows from the triangle inequality that $||f(x)|-|f(y)|| \leq |f(x)-f(y)|$, and given the symmetric rôle played by x, y on the supremum, $\sup\limits_{x,y\in[t_{j-1},t_j]} |f(x)-f(y)| = \sup\limits_{x,y\in[t_{j-1},t_j]} (f(x)-f(y))$. Similarly, $\sup\limits_{x,y\in[t_{j-1},t_j]} ||f(x)|-|f(y)|| = \sup\limits_{x,y\in[t_{j-1},t_j]} (|f(x)|-|f(y)|) = M_i^{|f|} - m_i^{|f|}$. Hence,

$$\begin{aligned} M_j^{|f|} - m_j^{|f|} &= \sup_{x,y\in[t_{j-1},t_j]} (|f(x)|-|f(y)|) \leq \sup_{x,y\in[t_{j-1},t_j]} (|f(x)-f(y)|) \\ &= \sup_{x,y\in[t_{j-1},t_j]} (f(x)-f(y)) = M_j^f - m_j^f \end{aligned}$$

Thus, by integrability of f, for any $\epsilon > 0$ there is a partition P such that

$$\sum_{j=1}^{k(P)} (M_j^{|f|} - m_j^{|f|})(t_j - t_{j-1}) \leq \sum_{j=1}^{k(P)} (M_j^f - m_j^f)(t_j - t_{j-1}) < \epsilon,$$

which shows that $|f| \in \mathcal{R}[a,b]$. Now, since $f \leq |f|$, by part 2 we immediately obtain that $\int_a^b f(x)dx \leq \int_a^b |f(x)|dx$. Furthermore, by part 1, $-f \in \mathcal{R}[a,b]$ and $-f \leq |f|$, hence $-\int_a^b f(x)dx \leq \int_a^b |f(x)|dx$. Thus, $\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx$. For the remaining functions, note that

$$\max\{f(x), g(x)\} = \frac{1}{2}\left(f(x) + g(x) + |f(x) - g(x)|\right),$$
$$\min\{f(x), g(x)\} = \frac{1}{2}\left(f(x) + g(x) - |f(x) - g(x)|\right),$$

$f^+ = \max\{f(x), 0\}$ and $f^- = -\min\{f(x), 0\}$. Consequently, using part 1 and the fact that $|f| \in \mathcal{R}[a,b]$ we have integrability of these functions.

4. Since $f^2 = |f|^2$ and $M_j^{f^2} - m_j^{f^2} = (M_j^{|f|})^2 - (m_j^{|f|})^2 = (M_j^{|f|} - m_j^{|f|})(M_j^{|f|} + m_j^{|f|}) \leq 2C(M_j^{|f|} - m_j^{|f|})$. Since, $|f|$ is integrable we have integrability of f^2. Lastly, $fg = \frac{1}{2}\left((f+g)^2 - f^2 - g^2\right)$. Since f^2, $g^2, (f+g)^2 \in \mathcal{R}[a,b]$, $fg \in \mathcal{R}[a,b]$. □

Recall that a sequence of functions $f_n : [a,b] \to \mathbb{R}$ for $n \in \mathbb{N}$ converges *uniformly* to $f : [a,b] \to \mathbb{R}$ if for any $\epsilon > 0$ there exists $N(\epsilon) \in \mathbb{N}$ such that for all $n > N(\epsilon)$,

$$|f_n(x) - f(x)| < \epsilon \text{ for all } x \in [a,b].$$

Theorem 134. *For every $f \in \mathcal{R}[a,b]$ define the function $F(x) := \int_a^x f(t)dt$. It is continuous for all $x \in [a,b]$.*

Proof. Since f is bounded, let $C := \sup_{x\in[a,b]} |f(x)|$. For all $x, y \in [a,b]$, $x < y$ we g=have

$$|F(x) - F(y)| = \left|\int_a^x f(t)dt - \int_a^y f(t)dt\right| = \left|\int_x^y f(t)dt\right| \leq \int_x^y |f(t)|dt \leq C(y-x) \to 0$$

as $y - x \to 0$. $\square$

Definition 82. *Let $f : [a,b] \to \mathbb{R}$ be bounded. Every function $F \in C[a,b]$ such that $F^{(1)}(x) = f(x)$ for all but finitely many $x \in (a,b)$ is called a primitive of f.*

Theorem 135. *For every $f \in C[a,b]$ has $F(x) := \int_a^x f(t)dt$ as a primitive. In addition,*

$$F(b) - F(a) = \int_a^b f(t)dt.$$

Proof. Since f is continuous, F is well defined and continuous. For $a < x < x+h < b$ and sufficiently small h we have

$$\begin{aligned} |F(x+h) - F(x) - hf(x)| &= \left|\int_a^{x+h} f(t)dt - \int_a^x f(t)dt - \int_x^{x+h} f(x)dt\right| \\ &= \left|\int_x^{x+h} (f(t) - f(x))dt\right| \leq \int_x^{x+h} |f(t) - f(x)|dt \\ &\leq \epsilon \int_x^{x+h} dt = h\epsilon \end{aligned}$$

where we used continuity of f at x. Similarly, we get

$$|F(x) - F(x-h) - hf(x)| \leq h\epsilon.$$

Combining both inequalities we get $\lim_{y\to x} \frac{F(y)-F(x)}{y-x} = f(x)$. Finally, the formula in the statement of the theorem follows from $F(a) = 0$. $\square$

Theorem 136. *(Fundamental Theorem of Calculus) Assume that F is a primitive of $f \in \mathcal{R}[a,b]$. Then,*

$$F(b) - F(a) = \int_a^b f(t)dt.$$

Proof. Since F is a primitive of f, let S be a finite set such that $F^{(1)}(x) = f(x)$ if $x \in (a,b) - S$. For $\epsilon > 0$, integrability of f means that there is a partition P of $[a,b]$ such that

$$S^P[f] - S_P[f] \leq \epsilon$$

and this inequality still holds for $P_1 = P \bigcup S$. Let the points of P_1 be denoted by $\{t_j\}_{j=0}^k$. Since,

$$F(b) - F(a) = \sum_{j=1}^{k} (F(t_j) - F(t_{j-1}))$$

and since F is differentiable in each (t_{j-1}, t_j) and continuous on $[a,b]$, by the Mean Value Theorem

$$F(t_j) - F(t_{j-1}) = f(\xi_j)(t_j - t_{j-1}).$$

Noting that $m_j := \inf_{x \in [t_{j-1}, t_j]} f(x) \leq f(\xi_j) \leq \sup_{x \in [t_{j-1}, t_j]} f(x) := M_j$, wec an sum over j to get

$$S^{P_1}[f] - \epsilon \leq S_{P_1}[f] \leq F(b) - F(a) \leq S^{P_1}[f] \leq S_{P_1}[f] + \epsilon.$$

By integrability $S_{P_1}[f] \leq \int_a^b f(t)dt \leq S^{P_1}[f]$ which shows that

$$\int_a^b f(t)dt - \epsilon \leq F(b) - F(a) \leq \int_a^b f(t)dt + \epsilon.$$

□

We now give the classical theorem on integration by parts without proof, which is left as an exercise.

Theorem 137. *Let $f, g \in \mathcal{R}[a,b]$ with primitives F, G. Then, FG is a primitive of $fG + Fg$ and*

$$\int_a^b f(t)G(t)dt = F(b)G(b) - F(a)G(a) - \int_a^b F(t)g(t)dt.$$

The following result is a convergence theorem for (proper) Riemann integrals.

Theorem 138. *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{R}[a,b]$ which converges uniformly to $f:[a,b]\to\mathbb{R}$. Then, $f\in\mathcal{R}[a,b]$ and*

$$\lim_{n\to\infty}\int_a^b f_n(x)dx = \int_a^b \lim_{n\to\infty} f_n(x)dx = \int_a^b f(x)dx.$$

Proof. Let $P=\{t_j\}_{j=0}^{k(P)}$ be a partition of $[a,b]$. First, since $f=(f-f_n)+f_n$, we have $\sup\limits_{x\in[t_{j-1},t_j]} f(x)\le \sup\limits_{x\in[t_{j-1},t_j]}(f(x)-f_n(x))+\sup\limits_{x\in[t_{j-1},t_j]} f_n(x)$ and since $\sup f=-\inf(-f)$ we have $\inf\limits_{x\in[t_{j-1},t_j]} f(x)\ge \inf\limits_{x\in[t_{j-1},t_j]}(f(x)-f_n(x))+\inf\limits_{x\in[t_{j-1},t_j]} f_n(x)$. Hence,

$$\begin{aligned} S^P[f]-S_P[f] &= S^P[(f-f_n)+f_n]-S_P[(f-f_n)+f_n] \\ &\le S^P[(f-f_n)]-S_P[(f-f_n)]+S^P[f_n]-S_P[f_n]. \end{aligned}$$

Second, by the fact that $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly to f, for any $\epsilon>0$ there exists $N(\epsilon)\in\mathbb{N}$ such that for all $n>N(\epsilon)$

$$\begin{aligned} S^P[(f-f_n)]-S_P[(f-f_n)] &\le 2\sum_{j=1}^{k(P)}\left|\sup_{x\in[t_{j-1},t_j]}(f-f_n)(x)\right|(t_j-t_{j-1}) \\ &\le 2\sum_{j=1}^{k(P)}\sup_{x\in[t_{j-1},t_j]}|f-f_n|(x)(t_j-t_{j-1})\le 2\epsilon(b-a). \end{aligned}$$

By definition $S^P[f]\ge \overline{\int_a^b} f$ and $S_P[f]\le \underline{\int_a^b} f$. Thus, for all $n>N(\epsilon)$

$$\overline{\int_a^b} f-\underline{\int_a^b} f\le S^P[f]-S_P[f] \;\le\; 2\epsilon(b-a)+S^P[f_n]-S_P[f_n].$$

Since $f_n\in\mathcal{R}[a,b]$, by Theorem 131.2 there is some P such that $S^P[f_n]-S_P[f_n]<\epsilon$ for any $\epsilon>0$. Consequently,

$$\overline{\int_a^b} f-\underline{\int_a^b} f\le 2\epsilon(b-a)+\epsilon,$$

establishing that f is integrable.

Now, since $f,f_n\in\mathcal{R}[a,b]$ we have that $f-f_n\in\mathcal{R}[a,b]$. From Theorem 133.3

$$\left|\int_a^b (f(x)-f_n(x))dx\right|\le\int_a^b|f(x)-f_n(x)|dx$$

and for $n\ge N(\epsilon)$, $\int_a^b|f(x)-f_n(x)|dx\le\epsilon(b-a)$, which completes the proof. $\square$

21 Improper Riemann integral

There are important extensions of the Riemann integral that are very useful. These extensions are called improper Riemann integrals and involve the following cases: a) the compact interval $[a,b]$ is extended to either $(-\infty,b]$ or $[a,\infty)$, b) the interval of integration is $(a,b]$ or $[a,b)$ and the function $f(x)$ is unbounded as $x\downarrow a$ or $x\uparrow b$ respectively, c)the interval of integration is (a,b) with $-\infty\le a<b\le\infty$ and the function $f(x)$ may or may not be bounded.

Definition 83. *Let $f\in\mathcal{R}[a,b]$ for all $b\in(a,\infty)$. If*

$$\lim_{b\to\infty}\int_a^b f(x)dx$$

exits and is finite, we say that f is improperly Riemann integrable and write $f\in\mathcal{R}[a,\infty)$. The value of this limit is denoted by $\int_a^\infty f(x)dx$ and called the Riemann integral of f over $[a,\infty)$. Similarly, for all $a\in(-\infty,b)$ if

$$\lim_{a\to-\infty}\int_a^b f(x)dx$$

exits and is finite, we say that f is improperly Riemann integrable and write $f\in\mathcal{R}(-\infty,b]$. The value of this limit is denoted by $\int_{-\infty}^b f(x)dx$ and called the Riemann integral of f over $(-\infty,b]$.

An immediate characterization of Riemann integrability on $[a,\infty)$ is given by the following theorem.

Theorem 139. $f\in\mathcal{R}[a,\infty)\iff\lim\limits_{x,y\to\infty}\int_x^y f(t)dt=0$ *for* $x<y<\infty$.

Proof. $(\Rightarrow)$ Let $U(x)=\int_a^x f(t)dt$ for $x\in(a,\infty)$. Since $f\in\mathcal{R}[a,\infty)$ we have that $\lim\limits_{x\to\infty}U(x)=\int_a^\infty f(t)dt$. But since every convergent sequence in $\mathbb{R}$ is Cauchy, $|U(x)-U(y)|=|\int_x^y f(t)dt|<\epsilon$ for x,y sufficiently large. $(\Leftarrow)$ Since $U(x)\in\mathbb{R}$ for $x\in(a,\infty)$ is Cauchy, $U(x)$ converges as $x\to\infty$ since $\mathbb{R}$ is complete.[8] $\square$

[8]Here, we consider the metric space $(\mathbb{R},|\cdot|)$.

References

Aliprantis, C., Border, K., 2006. Infinite Dimensional Analysis. Springer Verlag, New York.

Apostol, T., 1974. Mathematical Analysis. Addison Wesley, New York.

Hadley, G., 1977. Linear algebra. Addison Wesley, New York, NY.