Chapter 1

Divided Differences

This chapter introduces the fundamentals of the theory of divided differences of a nonlinear operator. Several results are also provided of divided differences as well as Fréchet derivatives satisfying Lipschitz or monotone-type conditions that will be used later.

1.1 Partially Ordered Topological Spaces

Let X be a linear space. We introduce the following definition:

Definition 1.1 A partially ordered topological linear space (POTL-space) is a locally convex topological linear space X which has a closed proper convex cone.

A proper convex cone is a subset K such that $K + K \subset K$, $\alpha K \subset K$ for $\alpha > 0$, and $K \cap (-K) = \{0\}$. Thus the order relation $\leq$, defined by $x \leq y$ if and only if $y - x \in K$, gives a partial ordering which is compatible with the linear structure of the space. The cone K which defines the ordering is called the positive cone since $K = \{x \in X \mid x \geq 0\}$. The fact that K is closed implies also that intervals, $[a, b] = \{z \in X \mid a \leq z \leq b\}$, are closed sets.

Example 1.1 Some simple examples of POTL-spaces are:

(1) $X = E^n$, n-dimensional Euclidean space, with

$$K = \{(x_1, x_2, \ldots, x_n) \in E^n \mid x_i \geq 0, i = 1, 2, \ldots, n\};$$

(2) $X = E^n$ with $K = \{(x_1, x_2, \ldots, x_n) \in E^n \mid x_i \geq 0, i = 1, 2, \ldots, n-1, x_n = 0\}$;
(3) $X = C^n[0,1]$, continuous functions, maximum norm topology, pointwise ordering;
(4) $X = C^n[0,1]$, n-times continuously differentiable functions $\|f\| = \sum_{k=0}^{n} \max |f^{(K)}(t)|$, pointwise ordering;
(5) $C = L^p[0,1]$, $0 \leq p \leq \infty$, usual topology, $K = \{f \in L^p[0,1] \mid f(t) \geq 0 \text{ a.e.}\}$.

Remarks 1.1 Using the above examples, it is easy to see that the closedness of the positive cone is not, in general, a strong enough connection between the ordering and the topology. Consider, for example, the following properties of sequences of real numbers:

(1) $x_1 \leq x_2 \leq \cdots \leq x^*$, and $\sup\{x_n\} = x^*$ implies $\lim_{n\to\infty} x_n = x^*$;
(2) $\lim_{n\to\infty} x_n = 0$ implies that there exists a sequence $\{y_n\}$ with $y_1 \geq y_2 \geq \cdots \geq 0$, $\inf\{y_n\} = 0$ and $-y_n \leq x_n \leq y_n$;
(3) $0 \leq x_n \leq y_n$, and $\lim_{n\to\infty} y_n = 0$ imply $\lim_{n\to\infty} x_n = 0$.

Unfortunately, these statements are not true for all POTL-spaces:

(a) In $X = C[0,1]$ let $x_n(t) = -t^n$. Then $x_1 \leq x_2 \leq \cdots \leq 0$, and $\sup\{x_n\} = 0$, but $\|x_n\| = 1$ for all n, so $\lim_{n\to\infty} x_n$ does not exist. Hence (1) does not hold.
(b) In $X = L^1[0,1]$ let $x_n(t) = n$ for $\frac{1}{n+1} \leq t \leq \frac{1}{n}$ and zero elsewhere. Then $\lim_{n\to\infty} \|x_n\| = 0$ but clearly property (2) does not hold.
(c) In $X = C^1[0,1]$ let $x_n(t) = \frac{t^n}{n}$, $y_n(t) = \frac{1}{n}$. Then $0 \leq x_n \leq y_n$, and $\lim_{n\to\infty} y_n = 0$, but $\|x_n\| = \max\left|\frac{t^n}{n}\right| + \max|t^{n-1}| = \frac{1}{n} + 1 > 1$; hence x_n does not converge to zero.

We will now devote a brief discussion of certain types of POTL spaces in which some of the above statements are true.

Definition 1.2 A POTL-space is called *regular* if every order-bounded increasing sequence has a limit.

Remarks 1.2 Examples of regular POTL-spaces are E^n and L^p, $0 \leq p \leq \infty$, whereas $C[0,1]$, $C^n[0,1]$ and $L^\infty[0,1]$ are not regular, as was shown in (a) of the above remark. If $\{x_n\}$ $n \geq 0$ is a monotone increasing sequence and $\lim_{n\to\infty} x_n = x^*$ exists, then for any k_0, $n \geq k_0$ implies $x_n \geq x_{k_0}$. Hence $x^* = \lim_{n\to\infty} x_n \geq x_{k_0}$, i.e., x^* is an upper bound on $\{x_n\}$ $n \geq 0$. Moreover, if y is any other upper bound, then $x_n \leq y$, and hence $x^* = \lim_{n\to\infty} x_n \leq y$, i.e., $x^* = \sup\{x_n\}$. This shows that in any POTL-space, the closedness of the positive cone guarantees that, if a monotone increasing sequence has a limit, then it is also a supremum. In a regular space, the converse of this is true; i.e., if a monotone increasing sequence has a supremum, then it also has a limit. It is important to note that the definition of regularity involves both an order concept (monotone boundedness) and a topological concept (limit).

Definition 1.3 A POTL-space is called normal if, given a local base U for the topology, there exists a positive number η such that if $0 \leq x \in V \in U$ then $[0, x] \subseteq \eta U$.

Remarks 1.3 If the topology of a POTL-space is given by a norm then this space is called a partially ordered normed space (PON)-space. If a PON-space is complete with respect to its topology then it is called a partially ordered Banach space (POB)-space. According to Definition 1.3. A PON-space is normal if and only if there exists a positive number α such that

$$\|x\| \leq \alpha \|y\| \quad \text{for all } x, y \in X \text{ with } 0 \leq x \leq y.$$

Let us note that any regular POB-space is normal. The converse is not true. For example, the space $C[0,1]$, ordered by the cone of non-negative functions, is normal but is not regular. All finite dimensional POTL-spaces are both normal and regular.

Remarks 1.4 Let us now define some special types of operators acting between two POTL-spaces. First we introduce some notation if X and Y are two linear spaces then we denote by (X, Y) the set of all operators from X into Y and by $L(X, Y)$ the set of all linear operators from X into Y. If X and Y are topological linear spaces then we denote by $LB(X, Y)$ the set of all continuous linear operators from X into Y. For simplicity the spaces $L(X, X)$ and $LB(X, X)$ will be denoted by $L(X)$ and $LB(X)$. Now let X and Y be two POTL-spaces and consider an operator $G \in (X, Y)$. G is called isotone (resp. antitone) if $x \leq y$ implies $G(x) \leq G(y)$ (resp. $G(x) \geq G(y)$). G is called non-negative if $x \geq 0$ implies $G(x) \geq 0$. G is called inverse non-negative if $G(x) \geq 0$ implies $x \geq 0$. For linear operators the nonnegativity is clearly equivalent with the isotony. Also, a linear operator is inverse non-negative if and only if it is invertible and its inverse is non-negative. If G is a non-negative operator then we write $G \geq 0$. If G and H are two operators from X into Y such that $H - G$ is non-negative then we write $G \leq H$. If Z is a linear space then we denote by $I = I_Z$ the identity operator in Z (i.e., $I(x) = x$ for all $x \in Z$). If Z is a POTL-space then we have obviously $I \geq 0$. Suppose that X and Y are two POTL-spaces and consider the operators $T \in L(X,Y)$ and $S \in L(Y,X)$. If $ST \leq I_X$ (resp. $ST \geq I_X$) then S is called a left subinverse (resp. superinverse) of T and T is called a right subinverse (resp. superinverse) of S. We say that S is a subinverse of T if S is a left—as well as a right subinverse of T.

We finally end this section by noting that for the theory of partially ordered linear spaces the reader may consult M. A. Krasnosel'skii [138]–[141], Vandergraft [202], Argyros and Szidarovszky [76], and Argyros [65].

1.2 Divided Difference in a Linear Space

The concept of a divided difference of a nonlinear operator generalizes the usual notion of a divided difference of a scalar function in the same way in which the Fréchet-derivative generalizes the notion of a derivative of a function.

Definition 1.4 Let F be a nonlinear operator defined on a subset D of a linear space X with values in a linear space Y, i.e., $F \in (D,Y)$ and let x, y be two points of D. A linear operator from X into Y, denoted $[x,y]$ (by which we really mean $[x,y;F]$ but to abbreviate we use the former notation hoping that it won't be confused with interval or other notation), which satisfies the condition

$$[x,y](x-y) = F(x) - F(y) \tag{1.1}$$

is called a divided difference of F at the points x and y.

Remarks 1.5 If X and Y are topological linear spaces then we shall always assume the continuity of the linear operator $[x,y]$. (Generally, $[x,y] \in L(X,Y)$ if X,Y are POTL-spaces then $[x,y] \in LB(X,Y)$.)

Obviously, condition (1.1) does not uniquely determine the divided difference, with the exception of the case when X is one-dimensional. An operator $[\cdot,\cdot] : D \times D \to L(X,Y)$ satisfying (1.1) is called a divided difference of F on D. If we fix the first variable, we get an operator

$$[x^0, \cdot] : D \to L(X, Y). \tag{1.2}$$

Let x^1, x^2 be two points on D. A divided difference of the operator (1.2) at the points x^1, x^2 will be called a divided difference of the second order of F at the points x^0, x^1, x^2 and will be denoted by $[x^0, x^1, x^2]$. We have by definition

$$[x^0, x^1, x^2](x^1 - x^2) = [x^0, x^1] - [x^0, x^2]. \tag{1.3}$$

Obviously, $[x^0, x^1, x^2] \in L(X, L(X, Y))$.

Let us now state a well known result due to Kantorovich [136], [137] concerning the location of fixed points which will be used extensively later.

Theorem 1.1 *Let X be a regular POTL-space and let x, y be two points of X such that $x \leq y$. If $H : [x, y] \to X$ is a continuous isotone operator having the property that $x \leq H(x)$ and $y \geq H(y)$, then there exists a point $z \in [x, y]$ such that $H(z) = z$.*

1.3 Divided Differences in a Banach Space

In this section we will assume that X and Y are Banach spaces. Accordingly we shall have $[x, y] \in LB(X, Y)$, $[x, y, z] \in LB(X, LB(X, Y))$. As we will see in later chapters, most convergence theorems in a Banach space require that the divided differences of F satisfy Lipschitz conditions of the form:

$$\|[x, y] - [x, z]\| \leq c_0 \|y - z\| \tag{1.4}$$

$$\|[y, x] - [z, x]\| \leq c_1 \|y - z\| \tag{1.5}$$

$$\|[x, y, z] - [u, y, z]\| \leq c_2 \|x - y\| \quad \text{for all } x, y, z, u \in D. \tag{1.6}$$

It is a simple exercise to show that if $[\cdot, \cdot]$ is a divided difference of F satisfying (1.4) or (1.5) then F is Fréchet differentiable on D and we have

$$F'(x) = [x, x] \quad \text{for all } x \in D. \tag{1.7}$$

Moreover, if (1.4) and (1.5) are both satisfied then the Fréchet derivative F' is Lipschitz continuous on D with Lipschitz constant $I = c_0 + c_1$.

At the end of this section we shall give an example of divided differences of the first and of the second order in the finite dimensional case. We shall consider the space $\mathbb{R}^q$ equipped with the Chebysheff norm which is given by

$$\|x\| = \max\{|x_i| \in \mathbb{R} \mid 1 \le i \le q\} \quad \text{for} \quad x = (x_1, x_2, \ldots, x_q) \in \mathbb{R}^q. \tag{1.8}$$

It follows that the norm of a linear operator $L \in LB(\mathbb{R}^q)$ represented by the matrix with entries l_{ij} is given by

$$\|L\| = \max \left\{ \sum_{j=1}^{q} |l_{ij}| \mid 1 \le i \le q \right\}. \tag{1.9}$$

We cannot give a formula for the norm of a bilinear operator. However, if B is a bilinear operator with entries b_{ijk} then we have the estimate

$$\|B\| \le \max \left\{ \sum_{j=1}^{q} \sum_{k=1}^{q} |b_{ijk}| \mid 1 \le i \le q \right\}. \tag{1.10}$$

Let U be an open ball of $\mathbb{R}^q$ and let F be an operator defined on U with values in $\mathbb{R}^q$. We denote by $f_1, \ldots, f_q$ the components of F. For each $x \in U$ we have

$$F(x) = (f_1(x), \ldots, f_q(x))^T. \tag{1.11}$$

Moreover, we introduce the notation

$$D_j f_i(x) = \frac{\partial f_i(x)}{\partial x_j}, \quad D_{kj} f_i(x) = \frac{\partial^2 f_i(x)}{\partial x_j \partial x_k}. \tag{1.12}$$

Let x, y be two points of U and let us denote by $[x, y]$ the matrix with entries

$$[x,y]_{ij} = \frac{1}{x_j - y_j}(f_i(x_1, \ldots, x_j, y_{j+1}, \ldots, y_q) - f_i(x_1, \ldots, x_{j-1}, y_j, \ldots, y_q)). \tag{1.13}$$

The linear operator $[x, y] \in LB(\mathbb{R}^q)$ defined in this way obviously satisfies condition (1.1). If the partial derivatives $D_j f_i$ satisfy some Lipschitz conditions of the form

$$|D_j f_i(x_1, \ldots, x_k + t, \ldots, x_q) - D_j f_i(x_1, \ldots, x_k, \ldots, x_q)| \le p_{jk}^i |t| \tag{1.14}$$

then conditions (1.4) and (1.5) will be satisfied with

$$c_0 = \max\left\{\frac{1}{2}\sum_{j=1}^{q}\left(p_{jj}^i + \sum_{k=j+1}^{q} p_{jk}^i\right) \mid 1 \le i \le q\right\} \tag{1.15}$$

and

$$c_1 = \max\left\{\frac{1}{2}\sum_{j=1}^{q}\left(p_{jj}^i + \sum_{k=1}^{j-1} p_{jk}^i\right) \mid 1 \le i \le q\right\}. \tag{1.16}$$

We shall prove (1.4) only since (1.5) can be proved similarly.

Let x, y, z be three points of U. We shall have in turn

$$\begin{aligned}[x,y]_{ij} - [x,z]_{ij} = \sum_{k=1}^{1}\{&[x,(y_1,\ldots,y_k,z_{k+1},\ldots,z_q)]_{ij} \\ &- [x,(y_1,\ldots,y_{k-1},z_k,\ldots,z_q)]_{ij}\}.\end{aligned} \tag{1.17}$$

by (1.13).

If $k \le j$ then we have

$$\begin{aligned}&[x,(y_1,\ldots,y_k,z_{k+1},\ldots,z_q)]_{ij} - [x,(y_1,\ldots,y_{k-1},z_k,\ldots,z_q)]_{ij} \\ &\quad = \frac{1}{x_j - z_j}\{f_i(x_1,\ldots,x_j,z_{j+1},\ldots,z_q) - f_i(x_1,\ldots,x_{j-1},z_j,\ldots,z_q)\} \\ &\quad - \frac{1}{x_j - z_j}\{f_i(x_1,\ldots,x_j,z_{j+1},\ldots,z_q) \\ &\quad - f_i(x_1,\ldots,x_{j-1},z_j,\ldots,z_q)\} = 0.\end{aligned}$$

For $k = j$ we have

$$\begin{aligned}&|[(x,(y_1,\ldots,y_j,z_{j+1},\ldots,z_q)]_{ij} - [x,(y_1,\ldots,y_{j-1},z_j,\ldots,z_q)]_{ij}| \\ &\quad = \left|\frac{1}{x_j - y_j}\{f_i(x_1,\ldots,x_j,z_{j+1},\ldots,z_q)\right. \\ &\quad - f_i(x_1,\ldots,x_{j-1},y_i,z_{j+1},\ldots,z_q)\} \\ &\quad \left.- \frac{1}{x_j - y_j}\{f_j(x_1,\ldots,x_j,z_{j+1},\ldots,z_q) - f_i(x_1,\ldots,x_{j-1},z_j,\ldots,z_q)\}\right| \\ &\quad = \left|\int_0^1 \{D_j f_i(x_1,\ldots,x_j,y_j + t(x_j - y_j),z_{j+1},\ldots,z_q)\right. \\ &\quad \left.- D_j f_i(x_1,\ldots,x_j,z_j + t(x_j - z_j),z_{j+1},\ldots,z_q)\}dt\right|\end{aligned}$$

$$\leq |y_j - z_j| p^i_{jj} \int_0^1 t\,dt = \frac{1}{2}|x_j - z_j| p^i_{jj}$$

by (1.14).

Finally for $k > j$ using (1.13) and (1.17) again, we have

$$\begin{aligned}
&|[x,(y_1,\dots,y_k,z_{k+1},\dots,z_q)]_{ij} - [x,(y_1,\dots,y_{k-1},z_k,\dots,z_q)]_{ij}| \\
&= \Bigg|\frac{1}{x_j - y_j}\{f_i(x_1,\dots,x_j,y_{j+1},\dots,y_k,z_{k+1},\dots,z_q) \\
&\quad - f_i(x_1,\dots,x_{j-1},y_j,\dots,y_k,z_{k+1},\dots,z_q) \\
&\quad - f_i(x_1,\dots,x_j,y_{j+1},\dots,y_{k-1},z_k,\dots,z_q) \\
&\quad + f_i(x_1,\dots,x_{j-1},y_j,\dots,y_{k-1},z_k,\dots,z_q)\}\Bigg| \\
&= \Bigg|\int_0^1 \{f_i(x_1,\dots,x_{j-1},y_j + t(x_j - y_j),y_{j+1},\dots,y_k,z_{k+1},\dots,z_q) \\
&\quad - f_i(x_1,\dots,x_{j-1},y_j + t(x_j - y_j),y_{j+1},\dots,y_{k-1},z_k,\dots,z_q)\}dt\Bigg| \\
&\leq |y_k - z_k| p^i_{jk}.
\end{aligned}$$

By adding all the above we get

$$\begin{aligned}
|[x,y]_{ij} - [x,z]_{ij}| &\leq \frac{1}{2}|y_j - z_j| p^i_{jj} + \sum_{k=j+1}^{q} |y_k - z_k| p^i_{pk} \\
&\leq \|y - z\| \left\{\frac{1}{2}\sum_{j=1}^{q}\left(p^i_{jj} + \sum_{k=j+1}^{q} p^i_{jk}\right)\right\}.
\end{aligned}$$

Consequently condition (1.4) is satisfied with c_0 given by (1.5). If each f_i has continuous second order partial derivatives which are bounded on U we have $p^i_{jk} = \sup\{|D_{jk} f_i(x)| \mid x \in U\}$. In this case $p^i_{jk} = p^i_{kj}$ so that $c_0 = c_1$.

Moreover, consider again three points x, y, z of U. Similarly with (1.17) the second divided difference of F at x, y, z is the bilinear operators defined by

$$\begin{aligned}
[x,y,z]_{ijk} = \frac{1}{y_k - z_k}\{&[x,(y_1,\dots,y_k,z_{k+1},\dots,z_q)]_{ij} \\
&- [x,(y_1,\dots,y_{k-1},z_k,\dots,z_q)]_{ij}\}.
\end{aligned} \tag{1.18}$$

It is easy to see as before that $[x,y,z]_{ijk} = 0$ for $k < j$. For $k = j$ we have

$$[x,y,z]_{ijj} = [x_j, y_j, z_j]_t f_i(x_1, \ldots, x_{j-1}, t, z_{j+1}, \ldots, z_q) \tag{1.19}$$

where the right hand side of (1.19) represents the divided difference of $f_i(x_1, \ldots, x_{j-1},\ z_{j+1}, \ldots, z_q)$ as a function of t, at the points x_j, y_j, z_j. Using Genocchi's integral representation of divided differences of scalar functions we get

$$\begin{aligned}[x,y,z]_{ijj} = \int_0^1 \int_0^1 & tD_{jj} f_i(x_1, \ldots, x_{j-1}, x_j + t(y_j - x_j) \\ & + ts(z_j - y_j), z_{j+1}, \ldots, z_q) dsdt. \end{aligned} \tag{1.20}$$

Hence, for $k > j$ we obtain

$$\begin{aligned}[x,y,z]_{ijk} &= \frac{1}{(y_k - z_k)(x_j - y_j)} \cdot \{ f_i(x_1, \ldots, x_j, y_{j+1}, \ldots, y_k, z_{k+1}, \ldots, z_q) \\ &\quad - f_i(x_1, \ldots, x_j, x_{j+1}, \ldots, y_{k-1}, z_k, \ldots, z_q) \\ &\quad - f_i(x_1, \ldots, x_{j-1}, y_j, \ldots, y_k, z_{k+1}, \ldots, z_q) \\ &\quad + f_i(x_1, \ldots, x_{j-1}, y_j, \ldots, y_{k-1}, z_k, \ldots, z_q) \} \\ &= \frac{1}{x_j - y_j} \int_0^1 \{ D_k f_i(x_1, \ldots, x_j, y_{j+1}, \ldots, y_{k-1}, z_k \\ &\quad + t(y_k - z_k), z_{k+1}, \ldots, z_q) \\ &\quad - D_k f_i(x_1, \ldots, x_{j-1}, y_j, \ldots, y_{k-1}, z_k \\ &\quad + t(y_k - z_k), z_{k+1}, \ldots, z_q) \} dt \\ &= \int_0^1 \int_0^1 D_{kj} f_i(x_1, \ldots, x_{j-1}, y_j + s(x_j - y_k), y_{j+1}, \ldots, y_{k-1}, z_k \\ &\quad + t(y_k - z_k), z_{k+1}, \ldots, z_q)\, dsdt. \end{aligned} \tag{1.21}$$

We now want to show that if

$$\begin{aligned} |D_{kj} f_i(v_1, \ldots, v_m + t, \ldots, v_q) - D_{kj} f_i(v_1, \ldots, v_m, \ldots, v_1)| \le q_{km}^{ij} |t| \\ \text{for all } v = (v_1, \ldots, v_q) \in U, \quad 1 \le i, j, k, m \le q, \end{aligned} \tag{1.22}$$

then the divided difference of F of the second order defined by (1.18) satisfies condition (1.6) with the constant

$$c_2 = \max_{1\le i\le q} \sum_{j=1}^{q} \left\{ \frac{1}{6} q_{jj}^{ij} + \frac{1}{2} \sum_{m=1}^{j-1} q_{jm}^{ij} + \frac{1}{2} \sum_{k=j+1}^{q} q_{kj}^{ij} + \sum_{k=j+1}^{q} \sum_{m=1}^{j-1} q_{km}^{ij} \right\}. \tag{1.23}$$

Let u, x, y, z be four points of U. Then using (1.18) we can easily have

$$[x,y,z]_{ijk} - [u,y,z]_{ijk} = \sum_{m=1}^{q} \{[(x_1,\ldots,x_m,u_{m+1},\ldots,u_1),yz]_{ijk} \\ -[(x_1,\ldots,x_{m-1},u_m,\ldots,u_q),y,z]_{ijk}\}. \tag{1.24}$$

If $m \ge j$ the terms in (1.24) vanish so that using (1.21) and (1.22) we deduce that for $k > j$

$$\begin{aligned} &|[x,y,z]_{ijk} - [u,y,z]_{ijk}| \\ &= \Bigg| \sum_{m=1}^{j-1} \int_0^1 \int_0^1 \{D_{kj} f_i(x_1,\ldots,u_{m+1},\ldots,u_{j-1},y_j + s(x_j - y_j), \\ &\quad y_{j+1},\ldots,y_{k-1},z_k + t(y_k - z_k),z_{k+1},\ldots,z_q) \\ &\quad - D_{kj} f_i(x_1,\ldots,x_{m-1},u_m,\ldots,u_{j-1},y_j + s(x_j - y_j),y_{j+1},\ldots,y_{k-1},z_k \\ &\quad + t(y_k - z_k),z_{k+1},\ldots,z_q)\}\, dsdt \\ &\quad + \int_0^1 \int_0^1 \{D_{kj} f_i(x_1,\ldots,y_{j-1},y_j + s(x_j - y_j),y_{j+1},\ldots,y_{k-1},z_k \\ &\quad + t(y_k - z_k),z_{k+1},\ldots,z_q) - D_{kj} f_i(x_1,\ldots,x_{j-1},y_j \\ &\quad + s(u_j - y_j),y_{j+1},\ldots,y_{k-1},z_k + t(y_k - z_k),z_{k+1},\ldots,z_q)\}\, dsdt \Bigg| \\ &\le \frac{1}{2}|x_j - u_j| q_{kj}^{ij} + \sum_{m=1}^{j-1} |x_m - u_m| q_{km}^{ij}. \end{aligned}$$

Similarly for $k = j$ we obtain in turn

$$\begin{aligned} &|[x,y,z]_{ijj} - [u,y,z]_{ijj}| \\ &= \Bigg| \int_0^1 \int_0^1 t\{D_{jj} f_i(x_1,\ldots,x_{j-1},x_j + t(y_j - x_j) + ts(z_j - y_j),z_{j+1},\ldots,z_q) \\ &\quad - D_{jj} f_i(x_1,\ldots,x_{j-1},u_j + t(y_j - u_j) + ts(z_j - y_j),z_{j+1},\ldots,z_q)\}\, dsdt \\ &\quad + \sum_{m=1}^{j-1} \int_0^1 \int_0^1 t\{D_{jj} f_i(x_1,\ldots,x_m,u_{m+1},\ldots,u_{j-1},x_j + t(x_j - y_j) \end{aligned}$$

$$+ts(z_j - y_j), z_{j+1}, \ldots, z_q) - D_{jj}f_i(x_1, \ldots, x_{m-1}, u_m, \ldots, u_{j-1}, x_j$$
$$+t(y_j - x_j) + ts(z_j - y_j), z_{j+1}, \ldots, z_q)\} \, dsdt \Big|$$
$$\leq \frac{1}{6}|x_j - u_j|q^{ij}_{jj} + \frac{1}{2}\sum_{m=1}^{j-1}|x_m - u_m|q^{ij}_{jm}.$$

Finally using the estimate (1.10) of the norm of a bilinear operator, we deduce that condition (1.6) holds with c_2 given by (1.23).

1.4 Divided Differences and Monotone Convergence

In this section we make an introduction to the problem of approximating a locally unique solution x^* of the nonlinear operator equation $F(x) = 0$, in a POTL-space X. In particular, consider an operator $F : D \subseteq X \to Y$ where X is a POTL-space with values in a POTL-space Y. Let x_0, y_0, y_{-1} be three points of D such that

$$x_0 \leq y_0 \leq y_{-1}, \; [x_0, y_{-1}],$$

and denote

$$\begin{aligned} D_1 &= \{(x, y) \in X^2 \mid x_0 \leq x \leq y \leq y_0\}, \\ D_2 &= \{(y, y_{-1}) \in X^2 \mid x_0 \leq u \leq y_0\}, \\ D_3 &= D_1 \cup D_2. \end{aligned} \tag{1.25}$$

Assume there exist operators $A_0 : D_3 \to LB(X, Y)$, $A : D_1 \to L(X, Y)$ such that:

(a) $F(y) - F(x) \leq A_0(w, z)(y - z)$ for all $(x, y)(y, w) \in D_1$, (1.26)
(b) the linear operator $A_0(u, v)$ has a continuous nonsingular non-negative left subinverse;
(c) $F(y) - F(x) \geq A(x, y)(y - x)$ for all $(x, y) \in D_1$; (1.27)
(d) the linear operator $A(x, y)$ has a non-negative left superinverse for each $(x, y) \in D_1$

$$F(y) - F(x) \leq A_0(y, z)(y - x) \;\; \text{for all } x, y \in D_1, \; (y, z) \in D_3. \tag{1.28}$$

Moreover, let us define the approximations

$$F(y_n) + A_0(y_n, y_{n-1})(y_{n+1} - y_n) = 0 \tag{1.29}$$

$$F(x_n) + A_0(y_n, y_{n-1})(x_{n+1} - x_n) = 0 \tag{1.30}$$

$$y_{n+1} = y_n - B_n F(y_n) \quad n \geq 0 \tag{1.31}$$

$$x_{n+1} = x_n - B_n^1 F(x_n) \quad n \geq 0, \tag{1.32}$$

where B_n and B_n^1 are non-negative subinverses of $A_0(y_n, y_{n-1})$ $n \geq 0$.

Under very natural conditions, hypotheses the form (1.26) or (1.27) or (1.28) have been used extensively to show that the approximations (1.26) and (1.30) or (1.31) and (1.32) generate two sequences $\{x_n\}$ $n \geq 1$, $\{y_n\}$ $n \geq 1$ such that

$$x_0 \leq x_1 \leq \cdots \leq x_n \leq x_{n+1} \leq y_{n+1} \leq y_n \leq \cdots \leq y_1 \leq y_0 \tag{1.33}$$

$$\lim_{n\to\infty} x_n = x^* = y^* = \lim_{n\to\infty} y_n \quad \text{and} \quad F(x^*) = 0. \tag{1.34}$$

For a complete survey on these results we refer the reader to the works of Potra [177]–[179], Argyros and Szidarovsky [76]–[79].

Here we will use similar conditions (i.e., like (1.26), (1.27), (1.28)) for two-point approximations of the form (1.29) and (1.30) or (1.31) and (1.32).

Consequently a discussion must follow on the possible choices of the linear operators A_0 and A.

Remarks 1.6 Let us now consider an operator $F : D \subseteq X \to Y$, where X, Y are both POTL-spaces. The operator F is called order-convex on an interval $[x_0, y_0] \subseteq D$ if

$$F(\lambda x + (1-\lambda)y) \leq \lambda F(x) + (1-\lambda)F(y) \tag{1.35}$$

for all comparable $x, y \in [x_0, y_0]$ and $\lambda \in [0, 1]$. If F has a linear G-derivative $F'(x)$ at each point $[x_0, y_0]$ then (1.35) holds if and only if

$$F'(x)(y-x) \leq F(y) - F(x) \leq F'(y)(y-x) \quad \text{for } x_0 \leq x \leq y \leq y_0. \tag{1.36}$$

(See, e.g., Ortega and Rheinboldt [169] for the properties of the Gateux-derivative.)

Hence, for order-convex G-differentiable operators conditions (1.27) and (1.28) are satisfied with $A_0(y, v) = A(y, v) = F'(u)$. In the unidimensional case (1.36) is equivalent with the isotony of the operator $x \to F'(x)$ but in general latter property is stronger. Assuming the isotony of the operator $x \to F'(x)$ it follows that

$$F(y) - F(x) \leq F'(w)(y-x) \quad \text{for } x_0 \leq x \leq y \leq w \leq y_0.$$

Hence, in this case condition (1.26) is satisfied for $A_0(w,z) = F'(w)$.

The above observations show how to choose A and A_0 for single or two-step Newton-methods. We note that the iterative algorithm (1.29)–(1.30) with $A_0(u,v) = F'(u)$ is the algorithm proposed by Fourier in 1818 in the unidimensional case and extended by Baluev in 1952 in the general case. The idea of using an algorithm of the form (1.31)–(1.32) goes back to Slugin [193]. In Ortega and Rheinboldt [169] it is shown that with B_n properly chosen (1.31) reduces to a general Newton-SOR algorithm. In particular, suppose (in the finite dimensional case) that $F'(y_n)$ is an M-matrix and let $F'(y_n) = D_n - L_n - U_n$ be the partition of $F'(y_n)$ into diagonal, strictly lower- and strictly upper-triangular parts respectively for all $n \geq 0$. Consider an integer $m_n \geq 1$, a real parameter $w_n \in (0,1]$ and denote

$$P_n = w_n^{-1}(D_n - w_n L_n), \quad Q_n = w_n^{-1}[(1-w_n)D_n + w_n U_n], \tag{1.37}$$
$$H_n = P_n^{-1}Q_n, \quad \text{and} \quad B_n = (I + H_n + \cdots + H_n^{m_n-1})P_n^{-1}. \tag{1.38}$$

It can easily be seen that B_n $n \geq 0$ is a non-negative subinverse of $F'(y_n)$ (see, also [169, Ch. 13.4]).

If $f : [a,b] \to \mathbb{R}$ is a real function of a real variable, then f is (order) convex if and only if $\frac{f(x)-f(y)}{x-y} \leq \frac{f(u)-f(v)}{u-v}$ for all x, y, u, v from $[a,b]$ such that $x \leq u$ and $y \leq v$. This fact motivates the notion of convexity with respect to a divided difference considered by J. W. Schmidt and H. Leonhardt [190]. Let $F : D \subseteq X \to Y$ be a nonlinear operator between the POTL-spaces X and Y. Assume that the nonlinear operator F has a divided difference $[\cdot,\cdot]$ on D. F is called convex with respect to the divided difference $[\cdot,\cdot]$ on D if

$$[x,y] \leq [u,v] \quad \text{for all } x,y,u,v \in D \text{ with } x \leq u \text{ and } y \leq v. \tag{1.39}$$

In the above quoted study, Schmidt and Leonhardt studied (1.29)–(1.30) with $A_0(u,v) = [u,v]$ in case the nonlinear operator F is convex with respect to $[\cdot,\cdot]$. Their result was extended by N. Schneider [169] who assumed instead of (1.1) the milder condition

$$[u, v](u - v) \geq F(u) - F(v) \quad \text{for all comparable } u, v \in D. \tag{1.40}$$

An operator $[\cdot,\cdot] : D \times D \to L(X, Y)$ satisfying (1.40) is called a generalized divided difference of F on D. If both (1.39) and (1.40) are satisfied, then we say that F is convex with respect to the generalized divided difference of $[\cdot,\cdot]$. It is easily seen that if (1.39) and (1.40) are satisfied on $D = [x_0, y_{-1}]$ then conditions (1.26) and (1.27) are satisfied with $A = A_0 = [\cdot,\cdot]$. Indeed for $x_0 \leq x \leq y \leq w \leq z \leq y_{-1}$ we have

$$[x, y](y - x) \leq F(y) - F(x) \leq [y, x](y - x) \leq [w, z](y - x).$$

Moreover, concerning general secant-SOR methods, in case the generalized difference $[y_n, y_{n-1}]$ is an M-matrix and if B_n $n \geq 0$ is computed according to (1.37) and (1.38) where $[y_n, y_{n-1}] = D_n - L_n - U_n$ $n \geq 0$ is the partition of $[y_n, y_{n-1}]$ into its diagonal, strictly lower- and strictly upper-triangular parts.

We remark that an operator which is convex with respect to a generalized divided difference is also order-convex. To see that, consider $x, y \in D$, $x \leq y$, $\lambda \in [0, 1]$ and set $z = \lambda x + (1 - \lambda)y$. Observing that $y - x = (1 - \lambda)^{-1}(z - x) = \lambda^{-1}(y - z)$ and applying (1.40) we have in turn:

$$\begin{aligned} (1 - \lambda)^{-1}(F(z) - F(x)) &\leq (1 - \lambda)^{-1}[z, x](z - x) \\ &= [z, x](y - x) \leq [z, y](y - x) \\ &= \lambda^{-1}[z, y](y - z) \leq \lambda^{-1}(F(y) - F(z)). \end{aligned}$$

By the first and last terms we deduce that $F(z) \leq \lambda F(x) + (1 - \lambda)F(y)$. Thus, Schneider's result can be applied only to order-convex operators and its importance resides in the fact that the use of a generalized divided difference instead of the G-derivative may be more advantageous from a numerical point of view. We note however, that conditions (1.26) and (1.27) do not necessarily imply convexity. For example, if f is a real function of a real variable such that

$$\inf_{x,y \in [x_0, y_0]} \frac{f(x) - f(y)}{x - y} = m > 0, \qquad \sup_{x,y \in [x_0, y_0]} \frac{f(x) - f(y)}{x - y} = M < \infty$$

then (1.26) and (1.27) are satisfied for $A_0(u, v) = M$ and $A(u, v) = m$. It is not difficult to find examples of nonconvex operators in the finite (or even in the infinite) dimensional case satisfying a condition of the form

$$A(y-x) \leq F(y) - F(x) \leq A_0(y,x), \quad x_0 \leq x \leq y \leq y_0$$

where A_0 and A are fixed linear operators. If A_0 has a continuous nonsingular non-negative left subinverse and A has a non-negative right superinverse, then convergence of the algorithm (1.29)–(1.30) can be discussed. This algorithm becomes extremely simple in this case. The monotone convergence of such an iterative procedure seems to have been first investigated by S. Slugin [192].

In the end of this section we shall consider a class of nonconvex operators which satisfy condition (1.26) but do not necessarily satisfy condition (1.27). Consequently from convergence theorems involving (1.29) and (1.30) it will follow that Jacobi–Newton and the Jacobi-secant methods have monotone convergence for operators belonging to this class (see, also the elegant papers by F. Potra [177], [179]).

Let $F = (f_1, \ldots, f_q)^T$ be an operator acting in the finite dimensional space $\mathbb{R}^q$, endowed with the natural (componentwise) partial ordering. Let us denote by e_i the ith coordinate vector of $\mathbb{R}^q$. We say that F is off-diagonally antitone if the functions

$$q_{ij}: \mathbb{R} \to \mathbb{R}, \; g_{ij}(t) = f_i(x + te_j), \; i \neq j, \; i,j = 1, \ldots, q$$

are antitone. Suppose that at each point x belonging to an interval $U = [x_0, y_{-1}]$ the partial derivatives $\partial_i F_i(x)$, $i = 1, 2, \ldots, q$, exist and are positive. For any two points $x, y \in U$ we consider the quotients

$$[x,y]_i = \begin{cases} \frac{f_i(y) - f_i(y - e_i^T(y-x)e_i)}{e_i^T(y-x)}, & \text{if } e_i^T(y-x) \neq 0 \\ \partial_i F_i(x) & \text{if } e_i^T(y-x) = 0. \end{cases} \tag{1.41}$$

Let us denote by $\Delta[x,y]$ the diagonal matrix having as elements the number $[x,y]_i$, $i = 1, 2, \ldots, q$. For the diagonal matrix $\Delta[x,y]$ formed by the partial derivatives $\partial_i f_i(x)$, $i = 1, 2, \ldots, q$, we shall also use the notation $DF(x)$.

Suppose now that F is off-diagonally antitone and that the operator $DF: U \to LB(\mathbb{R}^q)$ is isotone (i.e., all functions $\partial_i F_i: \mathbb{R} \to \mathbb{R}$) are isotone. In this case for all $x_0 \leq x \leq y \leq w \leq z \leq y_{-1}$ and all $i \in \{1, 2, \ldots, q\}$ there exist $\lambda, \mu \in [0,1]$ such that

$$f_i(y) - f_i(x) \leq f_i(y) - f_i(y - e_i^T(y-x)e_i) = \partial_i f_i(y - \lambda e_i^T(y-x)e_i),$$

$$e_i^T(y-x) \le \partial_i f_i(y) e_i^T(y-x) \le \partial_i f_i(w) e_i^T(y-x)$$
$$\le \partial_i f_i(z - \mu e_i^T(z-w)) e_i^T(y-x) = [w,z]_i e_i^T(y-x).$$

It follows that condition (1.26) is satisfied for $A_0(w,z) = DF(w)$ as well as for $A_0(w,z) = \Delta[w,z]$. With the choice $A_0(w,z) = \Delta[w,z]$ the iterative procedure (1.29) is a Jacobi-secant method while with the choice $A_0(w,z) = DF(w)$ it reduces to the Jacobi–Newton method. For some applications of the latter method, see W. Torning [197].

1.5 Divided Differences and Fréchet-derivatives

Let F be a nonlinear operator defined on an open subset D of a Banach space X with values in a Banach space Y. Choose also $x_0 \in D$ to be fixed.

Definition 1.5 The operator $F : D \subseteq X \to Y$ is Fréchet-differentiable at the point $x_0 \in D$ if there exists a linear operator $L \in LB(X,Y)$ such that

$$\lim_{\|h\|\to 0} \frac{1}{\|h\|} \|F(x_0+h) - F(x_0) - A(h)\| = 0. \tag{1.42}$$

It is easy to see that such an operator is unique if it exists. This operator L is called the first Fréchet-derivative of F at the point x_0 and is denoted by $F'(x_0)$.

The following facts are left as exercises:

1. If F is Fréchet-differentiable at the point x_0 then F is continuous at this point.
2. If $A \in LB(X,Y)$ and $F(x) = A(x)$ for all $x \in X$, then F is Fréchet-differentiable at every point $x \in X$ and $F'(x) = A$.
3. Let F and G be two operators defined on D with values in Y. If F and G are Fréchet-differentiable at a point $x_0 \in D$, then the operator $F+G$ is also Fréchet-differentiable at x_0 and $(F+G)'(x_0) = F'(x_0) + G'(x_0)$.
4. Let X, Y and Z be three Banach spaces and consider the operators $F : D_1 \subseteq X \to Y$, $G : D_2 \subseteq Y \to Z$. Let $x_0 \in D_1^0$ such that $y_0 = F(x_0)$ is a point in D_2^0. If F is Fréchet-differentiable at x_0 and G is Fréchet-differentiable at y_0, then the mapping $H = G_0 F$ is Fréchet-differentiable at x_0 and $H'(x_0) = G'(y_0)F'(x_0)$.

If the operator F is Fréchet-differentiable at all $x \in D$, then we shall say that F is Fréchet-differentiable on D. In this case we may consider an operator $F' : D \to LB(X, Y)$ which associates to each point $x \in D$ the Fréchet-derivative of F at x. This operator will be called the first Fréchet-derivative of F.

Let x, y be two points in D and suppose that the segment $S = \{x + t(y - x) \mid t \in [0,1]\} \subseteq D$. Let y' be a continuous linear functional, set $h = y - x$ and define

$$\varphi(t) = (F(x + th), y').$$

If F is Fréchet-differentiable at each point of the segment S, then φ is differentiable on $[0,1]$ and

$$\varphi'(t) = (F'(x + th), y').$$

Let us now suppose that

$$\alpha = \sup_{t \in [0,1]} \|F(x + t(y - x))\| < \infty;$$

then we have

$$\|(F(y) - F(x), y')\| = \|\varphi(1) - \varphi(0)\| \le \sup_{t \in [0,1]} \|\varphi'(t)\| \le \alpha \|y'\|\,\|y - x\|.$$

But, we also have

$$\|F(y) - F(x)\| = \sup_{\|y'\| \le 1} \|(F(y) - F(x), y')\|,$$

we deduce that

$$\|F(y) - F(x)\| \le \sup_{t \in [0,1]} \|F'(x + t(y - x)\| \cdot \|y - x\| \tag{1.43}$$

so, we proved:

Theorem 1.2 *Let D be a convex subset of a Banach space X and $F : D \subseteq X \to Y$. If F is Fréchet-differentiable on D and if there exists a constant c such that*

$$\|F'(x)\| \le M \text{ for all } x \in D \Rightarrow \|F(x) - F(y)\| \le c\|x - y\| \text{ for all } x \in D. \tag{1.44}$$

The estimate (1.43) is the analogue of the famous mean value formula from real analysis. If the operator F' is Riemann integrable on the segment S we can give the following integral representation of the mean value formula

$$F(x) - F(y) = \int_0^1 F'(x + t(y - x))\, dt(x - y). \tag{1.45}$$

Let now D be a convex open subset of X and let us suppose that we have associated to each pair (x, y) of distinct points from D a divided difference $[x, y]$ of F at these points. In applications one often has to require that the operator $(x, y) \to [x, y]$ satisfy a Lipschitz condition (see also Section 1.3). We suppose that there exists a non-negative $c > 0$ such that

$$\|[x, y] - [x_1, y_1]\| \le c(\|x - x_1\| + \|y - y_1\|) \tag{1.46}$$

for all $x, y, x_1, y_1 \in D$ with $x \ne y$ and $x_1 = y_1$.

We say in this case that F has a Lipschitz continuous difference on D. This condition allows us to extend by continuity the operator $(x, y) \to [x, y]$ to the whole Cartesian product $D \times D$. From (1.1) and (1.46) it follows that F is Fréchet-differentiable on D and that $[x, x] = F'(x)$. It also follows that

$$\|F'(x) - F'(y)\| \le c_1\|x - y\| \quad \text{with } c_1 = 2c \tag{1.47}$$

and

$$\|[x, y] - F'(z)\| \le c(\|x - z\| + \|y - z\|) \tag{1.48}$$

for all $x, y \in D$. Conversely if we assume that F is Fréchet-differentiable on D and that its Fréchet derivative satisfied (1.47) then it follows that F has a Lipschitz continuous divided difference on D. We can certainly take

$$[x, y] = \int_0^1 F'(x + t(t - x))\, dt. \tag{1.49}$$

We now want to give the definition of the second Fréchet derivative of F. We must first introduce the definition of bounded multilinear operators (which will also be used later).

Definition 1.6 Let X and Y be two Banach spaces. An operator $A : X^n \to Y$ $(n \in N)$ will be called n-linear operator from X to Y if the following conditions are satisfied:

(a) The operator $(x_1, \ldots, x_n) \to A(x_1, \ldots, x_n)$ is linear in each variable x_k, $k = 1, 2, \ldots, n$.

(b) There exists a constant c such that

$$\|A(x_1, x_2, \ldots, x_n)\| \leq c\|x_1\| \cdots \|x_n\|. \tag{1.50}$$

The norm of a bounded n-linear operator can be defined by the formula

$$\|A\| = \sup\{\|A(x_1, \ldots, x_n)\| \mid \|x_n\| = \cdots = \|x_n\| = 1\}. \tag{1.51}$$

Set $LB^{(1)}(X, Y) = LB(X, Y)$ and define recursively

$$LB^{(k+1)}(X, Y) = LB(X, LB^{(k)}(X, Y)), \quad k \geq 0. \tag{1.52}$$

In this way we obtain a sequence of Banach spaces $LB^{(n)}(X, Y)$ $(n \geq 0)$. Every $A \in LB^{(n)}(X, Y)$ can be viewed as a bounded n-linear operator if one takes

$$A(x_1, \ldots, x_n) = (\ldots(A(x_1)(x_2)(x_3))\ldots)(x_n). \tag{1.53}$$

On the right-hand side of (1.53) we have $A(x_1) \in LB^{(n-1)}(X, Y)$, $A(x_1))(x_2) \in LB^{(n-2)}(X, Y)$, etc. Conversely, any bounded n-linear operator A from X to Y can be interpreted as an element of $B^{(n)}(X, Y)$. Moreover the norm of A as a bounded n-linear operator coincides with the norm as an element of the space $LB^{(n)}(X, Y)$. Thus we may identify this space with the space of all bounded n-linear operators from X to Y. In the sequel we will identify $A(x, x, \ldots, x) = Ax^n$, and $A(x_1)(x_2)\ldots(x_n) = A(x_1, x_2, \ldots, x_n) = Ax_1x_2\ldots x_n$. Let us now consider a nonlinear operator $F : D \subseteq X \to Y$ where D is open. Suppose that F is Fréchet differentiable on D. Then we may consider the operator $F' : D \to LB(X, Y)$ which associates to each point x the Fréchet derivative of F at x. If the operator F' is Fréchet differentiable at a point $x_0 \in D$ then we say that F is twice Fréchet differentiable at x_0. The Fréchet derivative of F' at x_0 will be denoted by $F''(x_0)$ and will be called the second Fréchet derivative of F at x_0. Note that $F''(x_0) \in LB^{(2)}(X, Y)$. Similarly we can define Fréchet derivatives of higher order. Finally by analogy with (1.49)

$$\begin{aligned}[x_0, \ldots, x_k] = \int_0^1 \cdots \int_0^1 & t_1^{k-1} t_2^{k-2} \cdots t_{k-1} F(x_0 + t_1(x_1 - x_0) + t_1t_2(x_2 - x_1) \\ & + \cdots + t_1t_2, \ldots, t_k(x_k - x_{k-1}))\, dt_1 dt_2 \cdots dt_k. \end{aligned} \tag{1.54}$$

It is easy to see that the multilinear operators defined above verify

$$[x_0,\ldots,x_{k-1},x_k,x_{k+1}](x_k-x_{k+1}) = [x_0,\ldots,x_{k-1},x_k]-[x_0,\ldots,x_{k-1},x_{k+1}]. \tag{1.55}$$

We note that throughout this sequel a two-linear operator will also be called bilinear.

Finally, we will also need the definition of a n-linear symmetric operator. Given a n-linear operator $A : X^n \to Y$ and a permutation $i = (i_1, i_2, \ldots, i_n)$ of the integers $1, 2, \ldots, n$, the notation $A(i)$ (or $A_n(i)$ if we want to emphasize the n-linearity of A) can be used for the n-linear operator $A(i) = A_n(i)$ such that

$$\begin{aligned} A(i)(x_1,x_2,\ldots,x_n) = A_n(i)(x_1,x_2,\ldots,x_n) &= A_n(x_{i_1},x_{i_2},\ldots,x_{i_n}) \\ &= A_n x_{i_1} x_{i_2} \cdots x_{i_n} \end{aligned} \tag{1.56}$$

for all $x_1, x_2, \ldots, x_n \in X$. Thus, there are $n!$ n-linear operators $A(i) = A_n(i)$ associated with a given n-linear operator $A = A_n$.

Definition 1.7 An n-linear operator $A = A_n : X^n \to Y$ is said to be symmetric if

$$A = A_n = A_n(i) \tag{1.57}$$

for all i belonging in R_n which denotes the set of all permutations of the integers $1, 2, \ldots, n$. The symmetric n-linear operator

$$\overline{A} = \overline{A}_n = \frac{1}{n!} \sum_{i \in R_n} A_n(i) \tag{1.58}$$

is called the mean of $A = A_n$.

Notations 1.1 (a) The notation

$$A_n x^P = A_n x \times \cdots \times (p\text{-times}) \tag{1.59}$$

$p \le n$, $A = A_n : X^n \to Y$, for the result of applying A_n to $x \in X$ p-times will be used. If $p < n$, then (1.59) will represent a $(n-p)$-linear operator. For $p = n$, note that

$$A_k x^k = \overline{A}_k x^k = A_k(i) x^k \tag{1.60}$$

for all $i \in R_k$, $x \in X$. It follows from (1.60) that whenever we are dealing with an equation involving n-linear operators A_n, we may assume that they are symmetric without loss of generality, since each A_n may be replaced by

$\overline{A}_n$ without changing the value of the expression at hand. This is a fact that will be used frequently later in the chapters that follow.

(b) From now on the notations $U(x_0, R)$ and $\overline{U}(x_0, R)$ will denote open and closed balls centered at some fixed $x_0 \in X$ and of fixed radius $R > 0$. That is

$$U(x_0, R) = \{x \in X \mid \|x - x_0\| < R\}$$

and

$$\overline{U}(x_0, R) = \{x \in X \mid \|x - x_0\| \leq R\}.$$

Moreover, the following notations will be used

$$U(R) = U(0, R) \quad \text{and} \quad \overline{U}(r) = \overline{U}(0, R).$$

Furthermore, U and $\overline{U}$ will denote open and closed balls when there is no emphasis on x_0, X and R. U will denote the boundary of U, and U^0 the interior of U. Finally, the notation $U(R)(X =$ space, e.g., $L_p)$ (or $\overline{U}(R)(X$ $=$ space, e.g., $L_p)$) will be used to denote an open ball (resp. closed ball), when the emphasis is on the space X, R and $x_0 = 0$.

1.6 Exercises

1. Show that the spaces defined in Example 1.1 are POTL.
2. Show that any regular POB-space is normal but the converse is not necessarily true.
3. Prove Theorem 1.1.
4. show that if (1.4) or (1.5) are satisfied then $F'(x) = [x, x]$ for all $x \in D$. Moreover show that if both (1.4) and (1.5) are satisfied, then F' is Lipschitz continuous with $I = c_0 + c_1$.
5. Find sufficient conditions so that estimate (1.33) and (1.34) are both satisfied.
6. Show that B_n $(n \geq 0)$ in (1.38) is a non-negative subinverse of $F'(y_n)$ $(n \geq 0)$.
7. Let $x_0, x_1, \ldots, x_n$ be distinct real numbers, and let f be a given real-valued function. Show that:

$$[x_0, x_1, \ldots, x_n] = \sum_{j=0}^{n} \frac{f(x_j)}{g_n'(x_j)}$$

and

$$[x_0, x_1, \ldots, x_n](x_n - x_0) = [x_1, \ldots, x_n] - [x_0, \ldots, x_{n-1}]$$

where

$$g_n(x) = (x - x_0) \cdots (x - x_n).$$

8. Let $x_0, x_1, \ldots, x_n$ be distinct real numbers, and let f be n times continuously differentiable function on the interval $I\{x_0, x_1, \ldots, x_n\}$. Then show that

$$[x_0, x_1, \ldots, x_n] = \int_{\tau_n} \cdots \int f^{(n)}(t_0 x_0 + \cdots + t_n x_n) dt_1 \cdots dt_n$$

in which

$$\tau_n = \left\{ (t_1, \ldots, t_n) \mid t_1 \geq 0, \ldots, t_n \geq 0, \sum_{i=1}^{n} ti \leq 1 \right\}$$

$$t_0 = 1 - \sum_{i=1}^{n} t_i.$$

9. If f is a real polynomial of degree m, then show:

$$[x_0, x_1, \ldots, x_n, x] = \begin{cases} \text{polynomial of degree } m - n - 1, & n \leq m - 1 \\ a_m, & n = m - 1 \\ 0, & n > m - 1 \end{cases}$$

where $f(x) = a_m x^n +$ lower-degree terms.

10. The tensor product of two matrices $M, N \in L(\mathbb{R}^n)$ is defined as the $n^2 \times n^2$ matrix $M \times N = (m_{ij} N \mid i, j = 1, \ldots, n)$, where $M = (m_{ij})$. Consider two F-differentiable operators $H, K : L(\mathbb{R}^n) \to L(\mathbb{R}^n)$ and set $F(X) = H(X)K(X)$ for all $X \in L(\mathbb{R}^n)$. Show that $F'(X) = [H(X) \times I]K'(X) + [I \times K(X)^T]H'(X)$ for all $X \in L(\mathbb{R}^n)$.
11. Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $f_1(x) = x_1^3$, $f_2(x) = x_2^2$. Set $x = 0$ and $y = (1, 1)^T$. Show that there is no $z \in [x, y]$ such that

$$F(y) - F(x) = F'(z)(y - x).$$

12. Let $F : D \subset \mathbb{R}^n \to \mathbb{R}^m$ and assume that F is continuously differentiable on a convex set $D_0 \subset D$. For and $x, y \in D_0$, show that

$$\|F(y) - F(x) - F'(x)(y - x)\| \leq \|y - x\| w(\|y - x\|),$$

where w is the modulus of continuity of F' on $[x, y]$. That is

$$w(t) = \sup\{\|F'(x) - F'(y)\| \mid x, y \in D_0, \|x - y\| \le t\}.$$

13. Let $F : D \subset \mathbb{R}^n \to \mathbb{R}^m$. Show that F'' is continuous at $z \in D$ if and only if all second partial derivatives of the components $f_1, \ldots, f_m$ of F are continuous at z.
14. Let $F : D \subset \mathbb{R}^n \to \mathbb{R}^m$. Show that $F''(z)$ is symmetric if and only if each Hessian matrix $H_1(z), \ldots, H_m(z)$ is symmetric.
15. Let $M \in L(\mathbb{R}^n)$ be symmetric, and define $f : \mathbb{R}^n \to \mathbb{R}$ by $f(x) = x^T M x$. Show, directly from the definition that f is convex if and only if M is positive semidefinite.
16. Show that $f : D \subset \mathbb{R}^n \to \mathbb{R}$ is convex on the set D if and only if, for any $x, y \in D$, the function $g : [0, 1] \to \mathbb{R}$, $g(t) = g(tx + (1 - t)y)$, is convex on $[0, 1]$.
17. Show that if $g_i : \mathbb{R}^n \to \mathbb{R}$ is convex and $c_i \ge 0$, $i = 1, 2, \ldots, m$, then $g = \sum_{t=1}^m c_i g_i$ is convex.
18. Suppose that $g : D \subset \mathbb{R}^n \to \mathbb{R}$ is continuous on a convex set $D_0 \subset D$ and satisfies

$$\frac{1}{2} g(x) + \frac{1}{2} g(y) - g\left(\frac{1}{2}(x + y)\right) \ge \gamma \|x - y\|^2$$

for all $x, y \in D_0$. Show that g is convex on D_0 if $\gamma = 0$.

19. Let $M \in L(\mathbb{R}^n)$. Show that M is a non-negative matrix if and only if it is an isotone operator.
20. Let $M \in L(\mathbb{R}^n)$ be diagonal, nonsingular, and non-negative. Show that $\|x\| = \|D(x)\|_\infty$ is a monotonic norm on $\mathbb{R}^n$.
21. Let $M \in L(\mathbb{R}^n)$. Show that M is invertible and $M^{-1} \ge 0$ if and only if there exist nonsingular, non-negative matrices $M_1, M_2 \in L(\mathbb{R}^n)$ such that $M_1 M M_2 = I$.
22. Let $[\cdot, \cdot] : D \times D$ be an operator satisfying conditions (1.1) and (1.46). The following two assertions are equivalent:

    (i) Equality (1.49) holds for all $x, y \in D$.
    (ii) For all points $u, v \in D$ such that $2v - u \in D$ we have $[u, v] = 2[u, 2v - u] - [v, 2v - u]$.

23. If δF is a consistent approximation of F' on D show that each of the following four expressions in an estimate for $\|F(x) - F(y) -$

$\delta F(u,v)(x-y)\|$

$$c_1 = h(\|x-u\| + \|y-u\| + \|u-v\|)\|x-y\|,$$
$$c_2 = h(\|x-v\| + \|y-v\| + \|u-v\|)\|x-y\|,$$
$$c_3 = h(\|x-y\| + \|y-u\| + \|y-v\|)\|x-y\|$$

and

$$c_4 = h(\|x-y\| + \|x-u\| + \|x-v\|)\|x-y\|.$$

24. Show that the integral representation of $[x_0, \ldots, x_k]$ is indeed a divided difference of kth order of F. Let us assume that all divided differences have such an integral representation. In this case for $x_0 = x_1 = \cdots = x_k = x$ we shall have

$$\underbrace{[x, x, \ldots, x]}_{k+1 \text{ times}} = \frac{1}{k} f^{(k)}(x).$$

Suppose now that the nth Fréchet-derivative of F is Lipschitz continuous on D, i.e., there exists a constant c_{n+1} such that

$$\|F^{(n)}(u) - F^{(n)}(v)\| \leq c_{n+1}\|u-v\|$$

for all $u, v \in D$. In this case, set

$$R_n(y) = ([x_0, \ldots, x_{n-1}, y] - [x_0, \ldots, x_{n-1}, x_n])(y - x_{n-1}), \ldots, (y - x_0)$$

and show that

$$\|R_n(y)\| \leq \frac{c_{n+1}}{(n+1)!}\|y - x_n\| \cdot \|y - x_{n-1}\| \cdots \|y - x_0\|$$

and

$$\left\| F(x+h) - \left(F(x) + F'(x)h + \frac{1}{2}F''(x)h^2 + \cdots + \frac{1}{n!}F^{(n)}(x)h^n\right) \right\|$$
$$\leq \frac{c_{n+1}}{(n+1)!}\|h\|^{n+1}.$$

25. We recall the definitions:

 (a) An operator $F : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is Gateaux- (or G-) differentiable at an interior point x of D if there exists a linear operator $L \in L(\mathbb{R}^n, \mathbb{R}^m)$ such that, for any $h \in \mathbb{R}^n$

$$\lim_{t \to 0} \frac{1}{t} \|F(x+th) - F(x) - tL(h)\| = 0.$$

L is denoted by $F'(x)$ and called the G-derivative of F at x.

(b) An operator $F : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is hemicontinuous at $x \in D$ if, for any $h \in \mathbb{R}^n$ and $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon, h)$ so that whenever $|t| < \delta$ and $x+th \in D$, then $\|F(x+th) - F(x)\| < \varepsilon$.

(c) If $F : D \subset \mathbb{R}^n \to \mathbb{R}^m$ and if for some interior point x of D, and $h \in \mathbb{R}^n$, the limit

$$\lim_{t \to 0} \frac{1}{t}[F(x+th) - F(x)] = A(x,h)$$

exists, then F is said to have a Gateaux-differential at x in the direction h.

(d) If the G-differential exists at x for all h and if, in addition

$$\lim_{h \to 0} \frac{1}{\|h\|} \|F(x+H) - F(x) - A(x,h)\| = 0,$$

then F has a Fréchet differential at x. Show:

i. The linear operator L is unique;
ii. If $F : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is G-differentiable at $x \in D$, then F is hemicontinuous at x.
iii. G-differential and "uniform in h" implies F-differential;
iv. F-differential and "linear in h" implies F-derivative;
v. G-differential and "linear in h" implies G-derivative;
vi. G-derivative and "uniform in h" implies F-derivative. Here "uniform in h" indicated the validity of (d). Linear in h means that $A(x,h)$ exists for all $h \in \mathbb{R}^n$ and $A(x,h) = M(x)h$, where $M(x) \in L(\mathbb{R}^n, \mathbb{R}^m)$.
vii. Define $F : \mathbb{R}^2 \to \mathbb{R}$ by $F(x) = \operatorname{sgn}(x_2)\min(|x_1|, |x_2|)$. Show that, for any $h \in \mathbb{R}^2$, $A(0,h) = F(h)$, but F does not have a G-derivative at 0.
viii. Define $F : \mathbb{R}^2 \to \mathbb{R}$ by $F(0) = 0$ if $x = 0$ and

$$F(x) = x_2(x_1^2 + x_2^2)^{3/2}/[(x_1^2 + x_2^2)^2 + x_2^2] \quad \text{if } x \neq 0.$$

Show that F has a G-derivative at 0, but not an F-derivative. Show, moreover, that the G-derivative is hemicontinuous at 0.

ix. If the G-differential $A(x, h)$ exists for all x in an open neighborhood of an interior point x_0 of D and for all $h \in \mathbb{R}^n$, then F has an F-derivative at x_0 provided that for each fixed h, $A(x, h)$ is continuous in x at x_0.

(e) Assume that $F : D \subset \mathbb{R}^n \to \mathbb{R}^m$ has a G-derivative at each point of an open set $D_0 \subset D$. If the operator $F' : D_0 \subset \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^m)$ has a G-derivative at $x \in D_0$, then $(F')'(x)$ is denoted by $F''(x)$ and called the second G-derivative of F at x. Show:

i. If $F : \mathbb{R}^n \to \mathbb{R}^m$ has a G-derivative at each point of an open neighborhood of x, then F' is continuous at x if and only if all partial derivatives $\partial_i F_i$ are continuous at x.

ii. F'' is continuous at $x_0 \in D$ if and only if all second partial derivatives of the components $f_1, \ldots, f_m$ of F are continuous at x_0. $F''(x_0)$ is symmetric if and only if each Hessian matrix $H_1(x_0), \ldots, H_m(x_0)$ is symmetric.

26. Let F be a real function. Show:

$$[x_0, x_1, \ldots, x_n] = \frac{[x_1, \ldots, x_n] - [x_0, \ldots, x_{n-1}]}{x_n - x_0}$$

for all natural numbers n and all distinct points x_i, $i = 0, 1, 2, \ldots, n$.

27. By selecting appropriate numerical algorithms from the end of this text construct the following tables for $F(x) = \sqrt{x}$:

(a) Divided difference;
(b) Forward difference; and
(c) Backward difference.

28. (Hermite–Gennochi). Let $x_0, \ldots, x_n$ be distinct, and let $F(x)$ be n-times continuously differentiable on the interval $I\{x_0, x_1, \ldots, x_n\}$. Then

$$[x_0, \ldots, x_n] = \int_{\tau_n} \cdots \int F^{(n)}(t_0 x_0 + \cdots + t_n x_n) dt_1 \cdots dt_n,$$

in which

$$\tau_n = \left\{ (t_1, t_2, \ldots, t_n), t_1 \geq 0, \ldots, t_n \geq 0, \sum_{i=1}^{n} t_i \leq 1 \right\},$$

$$t_0 = 1 - \sum_{i=1}^{n} t_i.$$

29. Set $R_3(x) = F(x) - P_3(x) = a(x - x_0)(x - x_1)(x - x_2)(x - x_3)$, where P_3 is the Lagrange polynomial of order 3. Show that a is the appropriate divided difference of fourth order.