Mathematical Preliminary

1. Introduction

1.1 Forms and components of a mathematical programming problems

A *mathematical programming problem or, simply, a mathematical program* is a mathematical formulation of an optimization problem.

Unconstrained Problem:

$$\text{(P)} \quad \min_x f(x)$$
$$\text{s.t.} \quad x \in X,$$

where $x=(x_1, \ldots, x_n)^T \in \mathbf{R}^n$, $f(x): \mathbf{R}^n \to \mathbf{R}$, and X is an open set (usually $X \in \mathbf{R}^n$).

Constrained Problem:

$$\begin{aligned} \text{(P)} \quad & \min_x f(x) \\ \text{s.t.} \quad & g_i(x) \le 0 \quad i = 1, \ldots, m \\ & h_i(x) = 0 \quad i = 1, \ldots, l \\ & x \in X, \end{aligned}$$

where $g_1(x), \ldots, g_m(x)$, $h_1(x), \ldots, h_l(x)$: $\mathbf{R}^n \to \mathbf{R}$.

Let $g(x)=(g_1(x), \ldots, g_m(x))^T$: $\mathbf{R}^n \to \mathbf{R}^m$, $h(x)=(h_1(x), \ldots, h_l(x))^T$: $\mathbf{R}^n \to \mathbf{R}^l$. Then (P) can be written as

$$\begin{aligned} \text{(P)} \quad & \min_x f(x) \\ s.t. \quad & g(x) \le 0 \\ & h(x) = 0 \\ & x \in X. \end{aligned} \tag{1}$$

Some terminology: Function $f(x)$ is the *objective function*. Restrictions "$h_i(x)=0$" are referred to as *equality constraints*, while "$g_i(x) \le 0$" are *inequality constraints*. Notice that we do not use constraints in the form "$g_i(x)<0$"!

A point x is *feasible* for (P) if it satisfies all the constraints. (For an unconstrained problem, $x \in X$.) The set of all feasible points forms the *feasible region*, or *feasible set* (let us denote it by F). The goal of an optimization problem in minimization form, as above, is to find a feasible point $\bar{x}$ such that $f(\bar{x}) \le f(x)$ for any other feasible point x.

1.2 Linear Programming and the Simplex Method

Let's start with an example.

Minimize $x_1 - 3x_2$

subject to $-x_1 + 2x_2 \le 6$

$$x_1 + x_2 \le 5$$
$$x_1 \quad x_2 \ge 0$$

To make it be a standard format

Minimize $x_1 - 3x_2$

subject to $-x_1 + 2x_2 + x_3 \quad = 6$

$x_1 + x_2 \quad +x_4 = 5$

$x_1,\ x_2,\ x_3,\ x_4 \quad \ge 0$

where x_3 and x_4 are slack variables

define $x_3 = 6 + x_1 - 2x_2$ (The form illustrated below is called “a dictionary.”)

$x_4 = 5 - x_1 - x_2$

$z = \quad x_1 - 3x_2$

Initial solution

$x_1 = 0,\ \ x_2 = 0,\ \ x_3 = 6,\ \ x_4 = 5,\ \ z = 0$, $x_1\ x_2$ are nonbasic variables, $x_3\ x_4$ are basic variables

Choosing the entering variable x_2

$$x_3 = 6 + x_1 - 2x_2 \ge 0 \ \Rightarrow\ x_2 \le 3 \qquad (1)$$

$$x_4 = 5 - x_1 - x_2 \ge 0 \ \Rightarrow\ x_2 \le 5 \qquad (2)$$

(1) is the most stringent. Increasing x_2 up to (1).

$\therefore\ x_1 = 0,\ x_2 = 3,\ x_3 = 0,\ x_4 = 3$ (x_3 leaves the basis)

To construct the new system, we shall begin with the new comer to the left hand side, Namely, the variable x_2. The desired formula for x_2 in terms of x_1, x_3, x_4 is

$$x_2 = 3 + \frac{x_1}{2} - \frac{x_3}{2}$$

Next, in order to express x_4 and z in terms of $x_1, x_3,$ we simply substitute

$$x_4 = 5 - x_1 - \left(3 + \frac{x_1}{2} - \frac{x_3}{2}\right) = 2 - \frac{3}{2}x_1 + \frac{x_3}{2}$$

$$z = x_1 - 3\left(3 + \frac{x_1}{2} - \frac{x_3}{2}\right) = -9 - \frac{x_1}{2} + \frac{3}{2}x_3$$

Hence our new system

$$x_2 = 3 + \frac{x_1}{2} \qquad - \frac{x_3}{2}$$

$$x_4 = 2 - \frac{3}{2}x_1 \qquad + \frac{x_3}{2}$$

$$z = -9 - \frac{x_1}{2} \qquad + \frac{3}{2}x_3$$

Increase x_1 (x_1 enters the basis)

$$x_2 = 3 + \frac{x_1}{2} - \frac{x_3}{2} \ge 0 \ \Rightarrow \quad x_1 \text{ can be infinity}$$

$$x_4 = 2 - \frac{3}{2}x_1 + \frac{x_3}{2} \ge 0 \ \Rightarrow \quad x_1 \le \frac{4}{3} \qquad (x_4 \text{ leaves the basis})$$

$$\therefore\ x_1 = \frac{4}{3},\quad x_2 = \frac{11}{3},\quad x_3 = 0,\quad x_4 = 0$$

$$\Rightarrow\ x_1 = 2\cdot\frac{2}{3} + \frac{1}{2}\cdot\frac{2}{3}x_3 - \frac{2}{3}x_4 = \frac{4}{3} + \frac{1}{3}x_3 - \frac{2}{3}x_4$$

and $x_2 = 3 + \frac{1}{2}\left(\frac{4}{3} + \frac{1}{3}x_3 - \frac{2}{3}x_4\right) - \frac{x_3}{2} = \frac{11}{3} - \frac{1}{3}x_3 - \frac{1}{3}x_4$

$$z=-9-\frac{1}{2}\left(\frac{4}{3}+\frac{1}{3}x_3-\frac{2}{3}x_4\right)+\frac{3}{2}x_3=-\frac{29}{3}+\frac{4}{3}x_3+\frac{1}{3}x_4$$

Hence

$$x_1=\frac{4}{3}\qquad+\frac{1}{3}x_3-\frac{2}{3}x_4$$

$$x_2=\frac{11}{3}\qquad-\frac{1}{3}x_3-\frac{1}{3}x_4$$

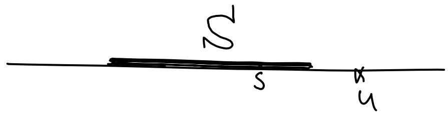
$$z=-\frac{29}{3}\qquad+\frac{4}{3}x_3+\frac{1}{3}x_4$$

$\therefore$ x_3 and x_4 in the objective function are with positive coefficients.

$\therefore$ Stop.

2. The Supremum Property of $\mathbf{R}$

The supremum property of the real numbers will help us to ensure that $S = \mathbf{R}$ exists as a real number. The formulation of this property involves upper or lower bounds for sets of real numbers.



SUPREMA AND INFIMA

Definition 2.1

Let $S \subset \mathbf{R}$.

(a) We say that $u \in \mathbf{R}$ is an *upper bound* of S if $s \leq u$ for all $s \in S$.

(b) We say that $w \in \mathbf{R}$ is a *lower bound* of S if $s \geq w$ for all $s \in S$.

Examples

(I) If $S = \{x \in \mathbf{R} : 0 < x < 1\}$, then 1 is an upper bound, but so is any $u \geq 1$. Likewise 0 is a lower bound, and so is any $w \leq 0$.

(II) If $S = \mathbf{R}$, then S does not have an upper or lower bound.

Remarks

(a) If a set has an upper bound, then we say that it is *bounded above*;

(b) If a set has a lower bound, then we say that it is *bounded below*;

(c) If a set has both an upper **and** lower bound, then we say that it is *bounded*; otherwise it is *unbounded*.

Examples

(I) $S = \{x \in \mathbf{R} : 0 < x < 1\}$ is bounded above and below, so is bounded;

(II) $P = \{x \in \mathbf{R} : x > 0\}$ is bounded below, but not above, so is unbounded.

Definition 2.2

Let $S \subset \mathbf{R}$.

(a) Let S be bounded above. An upper bounded of S is said to be a *supremum* (or *least upper bound*), if it is less than any other upper bound of S. We denote the supremum by sup S.

(b) Let S be bounded below. A lower bounded of S is said to be an *infimum* (or *greatest lower bound*), if it is larger than any other lower bound of S. We denote the infimum by inf S.

Remarks

We see that u is a supremum of S if and only if it satisfies the following two conditions:

(i) $s \le u$ for all $s \in S$.

(ii) if v is any upper bound for S then $u \le v$.

Thus a supremum is the "least upper bound" or " smallest upper bound". It is then fairly easy to see that if exists, it must be unique. Similarly for an infimum.

Lemma 2.3

Let $S \subseteq \mathbf{R}$ be non-empty. Then $u \in \mathbf{R}$ is the supremum of S if and only if both

(i) there are no elements $s \in S$ with $u < s$;

(ii) If $v < u$, then there exists $s \in S$ with $v < s$.

Proof

Examples

(I) Let $S_1 = \{x \in \mathbf{R} : 0 < x < 1\}$. Then

$$\sup S = 1; \inf S = 0$$

(II) Let $S_2 = \{x \in \mathbf{R} : 0 \le x \le 1\}$. Then again

$$\sup S = 1; \inf S = 0$$

Note that S_1 does not contain its supremum or infmum, but S_2 contains both.

(III) Let $P = \{x \in \mathbf{R} : x > 0\}$. Then $\inf P = 0$, but $\sup P$ does not exist.

It is an important property of the reals that every set bounded above has a supremum:

Property 2.4 Supremum Property

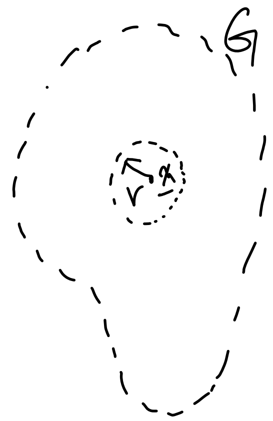
Every non-empty set of real numbers which has an upper bound has a supremum.

Property 2.5 Infimum Property

Every non-empty set of real numbers which has a lower bound has a infimum.

3. Open and Closed Sets

We shall discuss topological notions such as open and closed sets, interior points,…. Recall that $\mathbf{R}^p$ is the set of p-tuples



$$\underset{-}{x} = \left(x_1, x_2, \ldots x_p\right)$$

with

$$\left\|\underset{-}{x}\right\| = \sqrt{x_1^2 + x_2^2 + \ldots x_p^2}\ .$$

Definition 3.1

A set G in $\mathbf{R}^p$ is said to be *open* in $\mathbf{R}^p$ (or merely open) if for **each** $x \in G$, there exists $r > 0$ such that

$$\left\{\underset{-}{y} \in \mathbf{R}^p : \left\|\underset{-}{y} - \underset{-}{x}\right\| < r\right\} \subseteq G$$

Remark

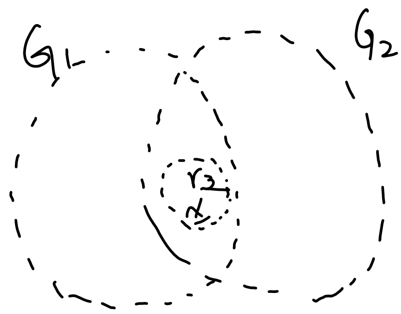
Thus G is open if each point in G is the center of some open ball contained in G.

Examples

(I) $(0,1)$ is open in $\mathbf{R}$ and $\left\{(x,y) \in \mathbf{R}^2 : x^2 + y^2 < 1\right\}$ is open in $\mathbf{R}^2$.

(II) $[0,1]$ and $[0,1)$ are not open in $\mathbf{R}$ and $\left\{(x,y) \in \mathbf{R}^2 : x^2 + y^2 \leq 1\right\}$ is not open in $\mathbf{R}^2$. Why?

We now establish some properties of open sets, which are often taken as the starting point for studying topology:



Theorem 3.3

(a) The empty set is open in $\mathbf{R}^p$; $\mathbf{R}^p$ is open in $\mathbf{R}^p$.
(b) The intersection of any two open sets is open.
(c) The union of any collection of open sets is open.

Proof

Remark

From (b) follows that the intersection of **finitely** many open sets is open. However, the intersection of **infinitely** many open sets need not be open. For example,

$$\bigcap_{n=1}^{\infty}\left(-1-\frac{1}{n},1+\frac{1}{n}\right)=[-1,1],$$

which is not open.

CLOSED SETS

Definition 3.4

A set F in $\mathbf{R}^p$ is said to be closed if $\mathbf{R}^p \setminus F$ is open.

Thus a set is closed if its complement is open.

Examples

(I) $[0,1]$ and $[0,\infty)$ are closed in $\mathbf{R}$ and $\{(x,y)\in \mathbf{R}^2 : x^2+y^2\le 1\}$ is closed in $\mathbf{R}^2$

(II) $[0,1)$ is not closed in $\mathbf{R}$.

(III) The empty set is closed in $\mathbf{R}^p$

Proposition 3.6 Properties of closed Sets

(a) The empty set ϕ and the whole space $\mathbf{R}^p$ are closed in $\mathbf{R}^p$.

(b) The union of any two closed sets is closed in $\mathbf{R}^p$

(c) The intersection of any collection of closed sets is closed in $\mathbf{R}^p$.

NEIGHBORHOODS

Definition 3.7

(a) If $\underline{x}\in \mathbf{R}^p$, then any set which contains an open set containing $\underline{x}$ is called a *neighborhood* of $\underline{x}$.

Now let $A\subseteq \mathbf{R}^p$ and $\underline{x}\in A$. (Definition by Marsden and Hoffman, 1993: A neighborhood of a point is an open set containing that point.)

(b) We call $\underline{x}$ an *interior point* of A if **there is** a neighborhood of $\underline{x}$ that is contained in A. The interior of A is the collection of all interior points of A and is denoted int(A).

(c) We call $\underline{x}$ a *boundary point* of A if **every** neighborhood of $\underline{x}$ contains a point in A and a point in $\mathbf{R}^p \setminus A$.

(d) We call $\underline{x}$ an *exterior point* of A if **there is** a neighborhood of $\underline{x}$ that is contained in $\mathbf{R}^p \setminus A$.

Remark

Note that exactly one of the three holds: x is interior or boundary or exterior, but cannot be more than one of these.

Examples

(I) Let $A = [0,1]$ in $\mathbf{R}$. Then any point in $(0,1)$ is an interior point of A, while 0 and 1 are boundary points. The exterior points of A are $\mathbf{R} \setminus [0,1]$. The same is true if we take $A = (0,1)$ or $[0,1)$.

(II) Let $A = \left\{ \underset{-}{y} \in \mathbf{R}^p : \left\| \underset{-}{y} - \underset{-}{x} \right\| < r \right\}$ be an open ball in $\mathbf{R}^p$. We see that each point in A is an interior point, while the boundary points of A are those points $\underset{-}{y}$ with $\left\| \underset{-}{y} - \underset{-}{x} \right\| = r$. The exterior points of A are those $\underset{-}{y} \in \mathbf{R}^p$ with $\left\| \underset{-}{y} - \underset{-}{x} \right\| > r$. Similarly for the closed ball $\left\{ \underset{-}{y} \in \mathbf{R}^p : \left\| \underset{-}{y} - \underset{-}{x} \right\| \le r \right\}$.

We may characterize open and closed sets in terms of their interior / boundary points:

Theorem 3.10

A set $F \subseteq \mathbf{R}^p$ is closed iff it contains all its boundary points.

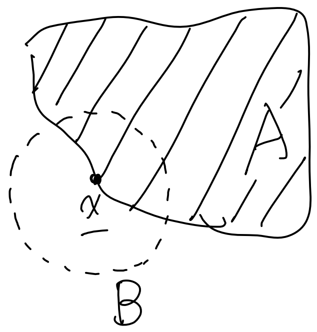
Proof

4. CLUSTER POINTS AND THE BOLZANO-WEIERSTRASS THEOREM

Another useful way to determine whether or not a set is closed is based on the concept of a cluster point.

Definition 4.1

Let $A \subseteq \mathbf{R}^p$ and $\underset{-}{x} \in \mathbf{R}^p$. We say that $\underset{-}{x}$ is a cluster point (or a point of accumulation) of A if every neighborhood of $\underset{-}{x}$ contains at least one point of A other than $\underset{-}{x}$.



Remarks

(a) Thus for $\underset{-}{x}$ to be a cluster point of A, we need

$$(B \cap A) \setminus \{\underset{-}{x}\} \neq \phi$$

for every neighborhood B of $\underset{-}{x}$.

(b) Cluster points are often also called accumulation points or limit points.

Examples

(a) If $A = (0,1)$, then every point of A is a cluster point, but 0 and 1 are also cluster points.

(b) If $A = [0,1]$, then every point of A is a cluster point.

(c) If A is an open set in $\mathbf{R}^p$, then every point of A is a cluster point, but the boundary points of A are also cluster points.

(d) If Q denotes the rational numbers, then every real number is a cluster point of Q. (Every neighborhood of a real number contains at least one rational number other than the original one.)

(e) More examples: $S = \{-1, 1, -1, 1, \ldots\}$. -1 and 1 are cluster points of S. Why? (Hint: all -1's are viewed as different ones. Similarly, all 1's are viewed as different ones.)

We can now characterize closed sets in terms of their cluster points: (we often use this result)

Theorem 4.2

Let $F \subseteq \mathbf{R}^p$. Then F is closed iff it contains all its cluster points.

Proof

Theorem 4.3 (Bolzano-Weierstrass Theorem)

Every bounded infinite set in $\mathbf{R}^p$ has a cluster point.

Remark

(a) A set A is called bounded if it is contained in a ball of finite radius.

(b) If A is a bounded set with infinitely many elements, then A has at least one cluster point.

Example

$A = \{-1, 1, -1, 1, ...\}$ -1 and 1 are cluster points

$A = \{1, 3, 7, 13\}$ is finite (having only four elements) and bounded. A has no cluster point.

Definition 4.4

Let $A \subseteq \mathbf{R}^p$ and $x \in \mathbf{R}^p$. We say that x is a closure point of A if every neighborhood of $\underline{x}$ contains at least one point of A. The closure of *A* is the collection of all closure points of *A* and is denoted cl(*A*).

Remarks

(a) Thus for $\underline{x}$ to be a closure point of A, we need

$$(B \cap A) \neq \phi$$

for every neighborhood B of x.

(b) A closure point of a set *A* is a point that is "close" to the set, in the sense that for every neighborhood, there is some element of *A* in the neighborhood.

Examples

If $A = [0,1) \cup \{2\}$, the closure of *A* in **R** is $[0,1] \cup \{2\}$.

If $A = [0,1) \cup \{2\}$, every point of $[0,1]$ is cluster point, but 2 is not a cluster point.

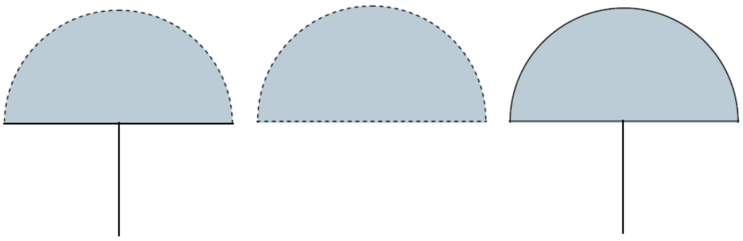
Theorem

The closure of *A* is the intersection of all closed sets containing *A*.

Remarks

(a) The largest <u>open subset</u> of *A* is called the interior of *A*.

(b) The smallest <u>closed set</u> containing *A* is the closure of *A*.



A int(*A*) cl(*A*)

Theorem 4.5 The Heine Borel Theorem

Let $K \subseteq \mathbf{R}^p$. Then *K* is compact iff it is closed and bounded.

Remark

It was seen the closed set $K = [1, \infty)$ is not compact; note that *K* is not bounded. It was also seen in Example that the bounded set *K* = (0,1) is not compact; note that K = (0,1) is not closed.

5. Introduction to Sequences

You have discussed sequences of real numbers or vectors in earlier courses. Here we revisit them, and discuss them more carefully. A somewhat formal definition of the notion of a sequence is given in:

Definition 5.1

Let S be a set. A *sequence* in S is a function on $\mathbf{N}$ and whose range is in S. In particular, a sequence in $\mathbf{R}^p$ is a function whose domain is $\mathbf{N}$ and whose range is contained in $\mathbf{R}^p$. In other words, a sequence in $\mathbf{R}^p$ assigns to each natural number $n = 1,2,\ldots$, a uniquely determined element of $\mathbf{R}^p$.

Remarks

(a) We shall denote a sequence by placing brackets around the elements of the sequence. Thus if $x_n \in S,\ n \geq 1$, then we denote the sequence by (x_n). (Many of you have used $\{x_n\}$ in earlier courses.)

(b) Mostly we shall deal with sequences in $\mathbf{R}^p$. Thus $(\underline{x}_n)$ denotes a sequence in $\mathbf{R}^p$.

(c) Sequences may be defined by a formula, e.g.

$$x_n = 2^n,\ n \geq 1,$$

or finitely many formulas, or inductively: e.g. $x_1 = x_2 = 1$,

$$x_n = x_{n-1} + x_{n-2},\ n \geq 3.$$

inner product

$(x_{n1}, \cdots x_{np})(y_{n1}, \cdots y_{np})$

$= x_{n1}y_{n1} + \cdots + x_{np}y_{np}$

$= \sum_{j=1}^{p} x_{nj}\, y_{nj}$

$\Rightarrow \left(\sum_{j=1}^{p} x_{nj}\, y_{nj}\right)$

(d) We can perform arithmetical operations on sequences:

Definition 5.2

Let $(\underline{x}_n)$ and $(\underline{y}_n)$ be sequences in $\mathbf{R}^p$.

(I) We define the *sum* of $(\underline{x}_n)$ and $(\underline{y}_n)$ to be the sequence $(\underline{x}_n + \underline{y}_n)$. Their *difference* is $(\underline{x}_n - \underline{y}_n)$.

(II) The *inner product* of $(\underline{x}_n)$ and $(\underline{y}_n)$ is defined to be the sequence $(\underline{x}_n \cdot \underline{y}_n)$.

(III) If $c \in \mathbf{R}$, we define the sequence $c(\underline{x}_n)$ to be the sequence $(c\underline{x}_n)$.

(IV) If (z_n) is a sequence of real numbers, we define the *product* of (z_n) and $(\underline{x}_n)$ to be the sequence $(z_n\underline{x}_n)$. If also $z_n \neq 0,\ n \geq 1$, we define the *quotient* of $(\underline{x}_n)$ and (z_n) to be the sequence $(\underline{x}_n / z_n)$.

Example

If $\left(x_n\right)$ and $\left(z_n\right)$ are sequences in $\mathbf{R}$ given by

$$x_n = n;$$
$$z_n = \left(-1\right)^n / n,$$

then

$$\left(x_n\right) = \left(1,2,3,\ldots\right);$$
$$\left(z_n\right) = \left(-1,\frac{1}{2},-\frac{1}{3},\frac{1}{4},\ldots\right)$$

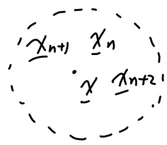
and

$$\left(x_n + z_n\right) = \left(n + \left(-1\right)^n / n\right) = \left(0,2\frac{1}{2},2\frac{2}{3},\ldots\right);$$
$$\left(x_n - z_n\right) = \left(n - \left(-1\right)^n / n\right) = \left(2,1\frac{1}{2},3\frac{1}{3},2\frac{3}{4},\ldots\right);$$
$$\left(x_n z_n\right) = \left(\left(-1\right)^n\right) = \left(-1,1,-1,\ldots\right);$$
$$\left(x_n / z_n\right) = \left(\left(-1\right)^n n^2\right) = \left(-1,4,-9,\ldots\right).$$

A very important part of analyzing sequences is discussing their convergence:

Definition 5.3 Sequence Convergence

Let $\left(\underline{x}_n\right)$ be a sequence in $\mathbf{R}^p$ and $\underline{x} \in \mathbf{R}^p$. We say that $\underline{x}$ is a *limit* of $\left(\underline{x}_n\right)$ if for each neighborhood V of $\underline{x}$, there is a natural number K_V such that



$$n \geq K_V \Rightarrow \underline{x}_n \in V.$$

If $\underline{x}$ is a *limit* of $(\underline{x}_n)$, we say that $(\underline{x}_n)$ *converges* to $\underline{x}$, and that the sequence is *convergent*. We also write

$$\underline{x} = \lim_{n\to\infty} \underline{x}_n$$

If the sequence does not converge, we say that it is *divergent*.

Thus, given any neighborhood V of $\underline{x}$, $\underline{x}_n$ must lie in V for n large enough (depending on the particular neighborhood V). We can reformulate this in terms of norm distances:

Theorem 5.4 (Another definition of sequence convergence)

Let ($\underline{x}_n$) be a sequence in $\mathbf{R}^p$ and $\underline{x} \in \mathbf{R}^p$. Then $\underline{x}$ is a limit of ($\underline{x}_n$) iff for each $\varepsilon > 0$, there exists $K(\varepsilon) \in \mathbf{N}$ such that

$$n \geq K(\varepsilon) \Rightarrow \|\underline{x}_n - \underline{x}\| < \varepsilon .$$

Theorem 5.5 Uniqueness of Limits

Let ($\underline{x}_n$) be a sequence in $\mathbf{R}^p$. There it can have at most one limit.

Proof

Suppose that $\underline{x}$ and $\underline{y}$ are limits of ($\underline{x}_n$) and that $\underline{x} \neq \underline{y}$. Let V' and V'' be disjoint neighborhoods of $\underline{x}$ and $\underline{y}$ and let K' and K'' be natural numbers such that if $n \geq K'$ then $\underline{x}_n \in V'$ and $n \geq K''$ then $\underline{x}_n \in V''$. Let $K = \sup\{K', K''\}$ so that both $\underline{x}_K \in V'$ and $\underline{x}_K \in V''$. We infer that $\underline{x}_K$ belongs to $V' \cap V''$ contradicting to that V' and V'' are disjoint. ■

Lemma 5.6 Boundedness of Convergent Sequences

A convergent sequence ($\underline{x}_n$) in $\mathbf{R}^p$ is bounded.

Proof

Let

$$\underline{x} = \lim_{n \to \infty} \underline{x}_n .$$

By Theorem 5.4 with $\varepsilon = 1$, there exists $K(\varepsilon)$ such that

$$n \geq K(\varepsilon) \Rightarrow \|\underline{x}_n - \underline{x}\| < 1 .$$

Then for $n \geq K(\varepsilon)$,

$$\|\underline{x}_n\| = \|\underline{x}_n - \underline{x} + \underline{x}\| \leq \|\underline{x}_n - \underline{x}\| + \|\underline{x}\| < 1 + \|\underline{x}\|.$$

Then if we set

$$M = \sup\left\{\|\underline{x}_1\|,\ \|\underline{x}_2\|,\ \|\underline{x}_3\| \ \dots \ \|\underline{x}_{K(\varepsilon)}\|,\ 1 + \|\underline{x}\|\right\}$$

we see that

$$\|\underline{x}_n\| \leq M,\ n \geq 1 .$$

■

It follows easily form the definition of the norm in $\mathbf{R}^p$ that a sequence in $\mathbf{R}^p$ converges iff all its component sequences converge:

Theorem 5.7

Let

$$\underline{x}_n = \left(x_{n1},\ x_{n2},\ \dots\ x_{np}\right) \in \mathbf{R}^p, n \geq 1.$$

Then $\left(\underline{x}_n\right)$ converges in $\mathbf{R}^p$ iff every one of the component sequence $\left(x_{nk}\right)_{n=1}^{\infty},\ k = 1, 2, \dots, p$, converges in **R**.

SOME EXAMPLES

Example 5.8 (a)

Let

$$x_n = \frac{1}{n},\ n \geq 1.$$

Then we know from experience that this sequence converges to 0. Let us prove this rigorously. Let $\varepsilon > 0$. Let $K(\varepsilon)$ be any integer $> 1/\varepsilon$. Then

$$n \overset{1}{\geq} K(\varepsilon) \Rightarrow |x_n - 0| = \left|\frac{1}{n} - 0\right|$$
$$= \frac{1}{n} \overset{1}{\leq} \frac{1}{K(\varepsilon)} \overset{2}{<} \frac{1}{1/\varepsilon} = \varepsilon.$$

By Theorem 5.4, $\left(\underline{x}_n\right)$ converges to 0.

Example 5.8 (b)

Let $a > 0$ and

$$x_n = \frac{1}{1+na}, n \geq 1.$$

Let $K(\varepsilon)$ be any integer $> 1/(a\varepsilon)$. Then

$$n \overset{1}{\geq} K(\varepsilon) \Rightarrow |x_n - 0| = \left|\frac{1}{1+na} - 0\right|$$
$$= \frac{1}{1+na} < \frac{1}{na} \overset{1}{<} \frac{1}{K(\varepsilon)a} \overset{2}{\leq} \frac{1}{\left(\frac{1}{a\varepsilon}\right)a} = \varepsilon.$$

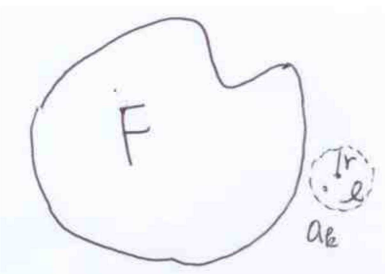
By Theorem 5.4, (x_n) converges to 0.

Theorem 5.9
A set F is closed if and only if for every convergent sequence (a_n) such that $a_n \in F$ for all $n \geq 1$, we have that $\lim_{n\to\infty} a_n \in F$.

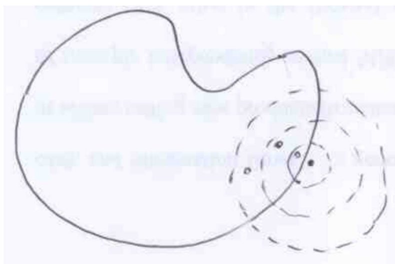
Proof
We first show that if (F is closed), then (for every convergent sequence (a_n) such that $a_n \in F$ for all $n \geq 1$, we have that $\lim_{n\to\infty} a_n \in F$.)
Let (a_n) be a convergent sequence such that $a_n \in F$ for all n and denote its limit by ℓ . Assume that $\ell \notin F$. Then ℓ is in the complement of F , which is open.
So, there exists an $r > 0$ such that the open ball $B(\ell, r)$ with center ℓ and radius $r > 0$ is contained in the complement of F , that is, $B(\ell, r)$ contains no point from F . But as (a_n) is convergent with limit ℓ , we can choose a larger enough k so that $\|a_k, \ell\| < \frac{r}{2}$. This implies that $a_k \in B(\ell, r)$. But also $a_k \in F$ and so we have a contradiction.

<u>Conversely,</u> we want to show (if for every convergent sequence (a_n) such that $a_n \in F$ for all $n \geq 1$, we have that $\lim_{n\to\infty} a_n \in F$), then (F is closed).
Let's show it by contradiction. Now suppose that F is not closed. Then its complement is not open.
(Recall if the complement of F is open, then F is closed.)
Recall: if (for each $x \in F$, there exists $r > 0$ such that $B(x, r) \subseteq F$), then (F is open)
$\therefore$ If the complement of F is not open, there exists a point ℓ in the complement of F such that for every $r > 0$, the open ball $B(\ell, r) \not\subset$ the complement of F . Then $B(\ell, r)$ has at least one point from F .
Now take successively $r = \frac{1}{n}$ $(n \in \mathbf{N})$ (recall "for every $r > 0$"), and choose a point $a_n \in F \cap B\left(\ell, \frac{1}{n}\right)$.
In this manner we obtain a sequence (a_n) such that $a_n \in F$ for each n, and $\|a_n - \ell\| < \frac{1}{n}$. The property $\|a_n - \ell\| < \frac{1}{n}$ $(n \in \mathbf{N})$ implies that (a_n) is a convergent sequence with limit ℓ . So we have obtained the existence of a convergent sequence (a_n) such that $a_n \in F$ $(n \in \mathbf{N})$, but $\lim_{n\to\infty} a_n = \ell \notin F$ (contradiction)

<u>Suppose that F is closed</u>



<u>Conversely</u>



6. Subsequences and Combinations

Very often we can learn a lot of about a sequence by looking at different “parts” of the sequence. These parts are called subsequences:

Definition 6.1

Let $X = (\underline{x}_n)$ be a sequence in $\mathbf{R}^p$. Let

$$r_1 < r_2 < r_3 < \ldots \qquad (*)$$

be a strictly increasing sequence of natural numbers. Then the sequence given by

$$X' = (\underline{x}_{r_n}) = (\underline{x}_{r_n})_{n=1}^{\infty} = (\underline{x}_{r_1}, \underline{x}_{r_2}, \underline{x}_{r_3}, \ldots)$$

is called a *subsequence* of $(\underline{x}_n)$.

Theorem 6.2

If a sequence X in $\mathbf{R}^p$ converges to a limit $\underline{x} \in \mathbf{R}^p$, then any subsequence of X also has the same limit.

Proof

Let

$$X' = (\underline{x}_{r_n})_{n=1}^{\infty} = (\underline{x}_{r_1}, \underline{x}_{r_2}, \underline{x}_{r_3}, \ldots)$$

be a subsequence of X. Let V be a neighborhood of $\underline{x}$. Then there exists $K_V \in \mathbf{N}$ such that

$$n \geq K_V \Rightarrow \underline{x}_n \in V.$$

By definition, $(\underline{x}_{r_n})_{n=1}^{\infty}$ converges to $\underline{x}$. ■

Corollary 6.3

If $X = (\underline{x}_n)$ is a sequence that converges to $\underline{x}$, and $m \in \mathbf{N}$, then $(\underline{x}_{n+m})_{n=1}^{\infty}$ also converges to $\underline{x}$.

Example

Let

$$x_n = (-1)^n, \ n \geq 1.$$

Then (x_n) is a divergent sequence. On the other hand,

$$x_{2n} = 1, \ n \geq 1,$$

So (x_{2n}) converges to 1. Similarly (x_{2n+1}) converges to -1.

COMBINATIONS OF SEQUENCES

Now we can investigate what happens when we add, multiply, … convergent sequences.

Theorem 6.6

(a) Let X and Y be sequences in $\mathbf{R}^p$ that converge to $\underline{x}$ and $\underline{y}$ respectively. Then

(i) THE LIMIT OF THE SUM IS THE SUM OF THE LIMITS

$X+Y$ converges to $\underline{x}+\underline{y}$.

(ii) THE LIMIT OF THE DIFFERENCE IS THE DIFFERENCE OF THE LIMITS

$X-Y$ converges to $\underline{x}-\underline{y}$.

(iii) THE LIMIT OF THE INNER PRODUCT IS THE INNER PRODUCT OF THE LIMITS

$X\cdot Y$ converges to $\underline{x}\cdot\underline{y}$.

(b) THE LIMIT OF THE PRODUCT IS THEPRODUCT OF THE LIMITS

Let $X=\left(\underline{x}_n\right)$ be a sequence in $\mathbf{R}^p$ that converges to $\underline{x}$ and let $A=\left(a_n\right)$ be a sequence in $\mathbf{R}$ that converges to a. Then the sequence $\left(a_n\underline{x}_n\right)$ in $\mathbf{R}^p$ converges to $a\underline{x}$.

(c) THE LIMIT OF THE QUOTIENT IS THE QUOTIENT OF THE LIMITS IF THE DENOMINATOR LIMIT IS NON-ZERO

Let $X=\left(\underline{x}_n\right)$ be a sequence in $\mathbf{R}^p$ that converges to $\underline{x}$ and $B=\left(b_n\right)$ be a sequence of non-zero real numbers that converges to a non-zero number b. Then the sequence $\left(b_n^{-1}\underline{x}_n\right)$ in $\mathbf{R}^p$ converges to $b^{-1}\underline{x}$.

Proof

(a)

(i) Write $X=\left(\underline{x}_n\right)$ and $Y=\left(\underline{y}_n\right)$. We must show that $\left(\underline{x}_n+\underline{y}_n\right)$ has limit $\underline{x}+\underline{y}$. To do this, we must estimate

$$\begin{aligned}&\left\|\left(\underline{x}_n+\underline{y}_n\right)-\left(\underline{x}+\underline{y}\right)\right\|=\left\|\left(\underline{x}_n-\underline{x}\right)+\left(\underline{y}_n-\underline{y}\right)\right\|\\&\le\left\|\underline{x}_n-\underline{x}\right\|+\left\|\underline{y}_n-\underline{y}\right\|.\end{aligned}\qquad(*)$$

But by our hypothesis, given $\varepsilon>0$, there exists $K\in\mathbf{N}$ such that both

$$n\ge K\Rightarrow\left\|\underline{x}_n-\underline{x}\right\|<\frac{\varepsilon}{2}$$

and

$$n\ge K\Rightarrow\left\|\underline{y}_n-\underline{y}\right\|<\frac{\varepsilon}{2}.$$

(Of course the K may be different for X and Y but we can take the larger one). Then (*) shows that for $n\ge K$,

$$\left\|\left(\underline{x}_n+\underline{y}_n\right)-\left(\underline{x}+\underline{y}\right)\right\|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

By definition, $\left(\underline{x}_n+\underline{y}_n\right)$ has limit $\underline{x}+\underline{y}$.

Remark

It is essential in (c) that the denominator limit is non-zero. For example, let (x_n) and (b_n) be sequences of real numbers defined by

$$x_n = 1 + \frac{1}{n},\ n \geq 1$$

and

$$b_n = \frac{1}{n}, n \geq 1.$$

Then it is easy to see that (x_n) converges to 1 and (b_n) converges to 0, but

$$b_n^{-1} x_n = n\left(1 + \frac{1}{n}\right) = n + 1$$

So $(b_n^{-1} x_n)$ is not bounded, and cannot converge.

Examples

(I) Consider the sequence (x_n) given by

$$x_n = \frac{5n-3}{4n+2}, n \geq 1.$$

We may write

$$x_n = \frac{n\left(5 - \frac{3}{n}\right)}{n\left(4 + \frac{2}{n}\right)} = \frac{5 - \frac{3}{n}}{4 + \frac{2}{n}}.$$

Then,

$$\begin{aligned}\lim_{n\to\infty}\left(5 - \frac{3}{n}\right) &= \lim_{n\to\infty} 5 - \lim_{n\to\infty} \frac{3}{n} \\ &= 5 - 3\lim_{n\to\infty}\frac{1}{n} = 5 - 0.\end{aligned}$$

Similarly,

$$\lim_{n\to\infty}\left(4 + \frac{2}{n}\right) = 4 \neq 0.$$

The above theorem gives

$$\lim_{n\to\infty}\frac{5 - \frac{3}{n}}{4 + \frac{2}{n}} = \frac{\lim\limits_{n\to\infty}(5 - 3/n)}{\lim\limits_{n\to\infty}(4 + 2/n)} = \frac{5}{4}.$$

Theorem 6.7

Every sequence (x_n) has a monotonic subsequence.

Example

Let

$$x_n = (-1)^n + \frac{1}{n},\ n \geq 1.$$

Then (x_n) is a divergent sequence and is not monotonic. On the other hand,

$$x_{2n} = 1 + \frac{1}{2n},\ n \geq 1,$$

So (x_{2n}) is a monotonic subsequence. Similarly (x_{2n+1}) is a monotonic subsequence as well.

Two Criteria for Convergence

In this section, we derive two important general criteria for convergence of a sequence. They are often simpler to apply than the original definition of convergence.

Theorem 6.8 MONOTONE CONVERGENCE THEOREM

Let (x_n) be a sequence of real numbers such that

$$x_1 \leq x_2 \leq x_3 \leq \ldots \leq x_n \leq x_{n+1} \leq \ldots\ .$$

(We say that (x_n) is monotone increasing). Then the sequence converges iff it is bounded. If it is bounded, then

$$\lim_{n\to\infty} x_n = \sup\{x_n : n \geq 1\} \qquad (*)$$

Corollary 6.9

Let (x_n) be a sequence of real numbers such that

$$x_1 \geq x_2 \geq x_3 \geq \ldots \geq x_n \geq x_{n+1} \geq \ldots$$

(We say that (x_n) is monotone decreasing). Then the sequence converges iff it is bounded. If it is bounded, then

$$\lim_{n\to\infty} x_n = \inf\{x_n : n \geq 1\}.$$

Examples

(a) The sequence (x_n) given by

$$x_n = 1 - \frac{1}{n^2}, n \geq 1$$

is monotone increasing (check!). It is also bounded above by 1, and hence bounded. By the Monotone Convergence Theorem, (x_n) converges, and

$$\lim_{n\to\infty} x_n = \sup\{x_n : n \geq 1\} = 1.$$

(b) Let $a > 1$. Let us define a sequence by $x_1 = 1$ and

$$x_{n+1} = \sqrt{ax_n}, n \geq 1$$

Then we claim that (x_n) is increasing and that $x_n \leq a$ for all n. Let us prove this by induction.

The case $n = 1$

We see that $x_2 = \sqrt{a \cdot 1} = \sqrt{a} > 1 = x_1$ (as $a > 1$). Also, $x_2 = \sqrt{a} \leq a$ (since $a = \sqrt{a}\sqrt{a} \geq \sqrt{a} \cdot 1$), so

$$x_1 \leq x_2 \leq a.$$

Induction Step

Assume now as an induction step that $x_n \leq x_{n+1} \leq a$

Then

$$x_{n+2} = \sqrt{ax_{n+1}}$$

So

$$x_{n+2} / x_{n+1} = \sqrt{ax_{n+1}} / x_{n+1} = \sqrt{a / x_{n+1}} \geq 1$$

(We just used our induction hypothesis $a / x_{n+1} \geq 1$). Also,

$$x_{n+2} / a = \sqrt{ax_{n+1}} / a = \sqrt{x_{n+1} / a} \leq 1.$$

(We just used our induction hypothesis $a / x_{n+1} \geq 1$). So we have proved that

$$x_{n+1} \leq x_{n+2} \leq a.$$

Thus by induction, for all $n \geq 1$,

$$x_n \leq x_{n+1} \leq a.$$

Then (x_n) is both bounded above (by a) and monotone increasing. By the monotone convergence theorem, (x_n) converges. Let us call the limit r. That theorem also tells us

$$r = \lim_{n\to\infty} x_n = \sup\{x_n : n \geq 1\}.$$

Let $\varepsilon \in (0, r)$. There exists $K(\varepsilon) \in \mathbf{N}$ such that

$$n \geq K(\varepsilon) \Rightarrow r - \varepsilon < x_n (< r + \varepsilon)$$

And we know $r \geq x_n$ for all n, so for $n \geq K(\varepsilon)$,

$$r - \varepsilon \leq x_{n+1} = \sqrt{ax_n} \leq \sqrt{ar}$$

and

$$r \geq x_{n+1} = \sqrt{ax_n} \geq \sqrt{a(r - \varepsilon)}$$

Let us square these last two relations. Then, for each $\varepsilon \in (0, r)$,

$$(r - \varepsilon)^2 \leq \left(\sqrt{ar}\right)^2 = ar;$$
$$r^2 \geq \left(\sqrt{a(r - \varepsilon)}\right)^2 = a(r - \varepsilon).$$

Since ε may be made arbitrarily small and r is independent of ε, we have

$$r^2 \leq ar;$$
$$r^2 \geq ar.$$

Thus

$$r^2 = ar;$$
$$\Rightarrow r = a$$

So we have shown that

$$\lim_{n \to \infty} x_n = a$$

THE BOLZANO WEIERSTRASS THEOREM

It will be useful to reformulate the Bolzano-Weierstrass Theorem for sequences:

Theorem 6.9 Bolzano-Weierstrass Theorem for sequences

A bounded sequence in $\mathbf{R}^p$ has a convergent subsequence.

Proof

By Theorem 6.7: Every sequence (x_n) has a monotonic subsequence. The subsequence converges by Theorem 6.8 or Corollary 6.9.

CAUCHY SEQUENCES

Roughly speaking, a Cauchy sequence may be defined as follows: if all elements of a sequence, beyond a certain index, are close to one another, they form a Cauchy sequence. More precisely:

Definition 6.10

A sequence $(\underline{x}_n)$ in $\mathbf{R}^p$ is said to be a Cauchy sequence if for every $\varepsilon > 0$, there exists $M(\varepsilon) \in \mathbf{N}$ such that

$$m, n \geq M(\varepsilon) \Rightarrow \|\underline{x}_m - \underline{x}_n\| < \varepsilon .$$

In fact every convergent sequence is a Cauchy sequence:

Lemma 6.11

If $(\underline{x}_n)$ is a convergent sequence in $\mathbf{R}^p$, it is also a Cauchy sequence.

Proof

Let

$$\underline{x} = \lim_{n\to\infty} \underline{x}_n$$

given $\varepsilon > 0$. By Theorem 13.4, there exists $K\left(\frac{\varepsilon}{2}\right) \in \mathbf{N}$ such that

$$n \geq K\left(\frac{\varepsilon}{2}\right) \Rightarrow \|\underline{x}_n - \underline{x}\| < \frac{\varepsilon}{2}$$

Then, if $M(\varepsilon) = K\left(\frac{\varepsilon}{2}\right)$,

$$\begin{aligned} m, n \geq M(\varepsilon) \Rightarrow \|\underline{x}_m - \underline{x}_n\| &\leq \|\underline{x}_m - \underline{x}\| + \|\underline{x} - \underline{x}_n\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

So, it is Cauchy. ■

We shall also prove that every Cauchy sequence is convergent, that is the converse holds. First we need to prove:

Lemma 6.12

A Cauchy sequence $(\underline{x}_n)$ in $\mathbf{R}^p$ is bounded.

Proof

By the definition of a Cauchy sequence, with $\varepsilon = 1$. There exists a natural number $M(1)$ such that

$$n, m \geq M(1) \Rightarrow \|\underline{x}_n - \underline{x}_m\| < 1.$$

Then, setting $m = M(1)$, we see that

$$n \geq M(1) \Rightarrow \|\underline{x}_n\| = \left\|\underline{x}_n - \underline{x}_{M(1)} + \underline{x}_{M(1)}\right\| \leq \left\|\underline{x}_n - \underline{x}_{M(1)}\right\| + \left\|\underline{x}_{M(1)}\right\| < 1 + \left\|\underline{x}_{M(1)}\right\|.$$

It follows that if we set

$$B = \sup\left\{\|\underline{x}_1\|, \|\underline{x}_2\|, \|\underline{x}_3\|, \ldots \left\|\underline{x}_{M(1)-1}\right\|, 1 + \left\|\underline{x}_{M(1)}\right\|\right\},$$

Then

$$\|\underline{x}_n\| \leq B, n \geq 1.$$

We need another lemma:

Lemma 6.13

If a subsequence of a Cauchy sequence converges to some $\underline{x} \in \mathbf{R}^p$, then the entire sequence converges to $\underline{x}$.

Proof

Assume that $(\underline{x}_n)$ is Cauchy sequence, and that it has a subsequence $(\underline{x}_{n_k})_{k=1}^{\infty}$ converging to $\underline{x}$. Let $\varepsilon > 0$.

By the definition of a Cauchy sequence, there exists $M\left(\frac{\varepsilon}{2}\right) \in \mathbf{N}$ such that

$$n, m \geq M\left(\frac{\varepsilon}{2}\right) \Rightarrow \|\underline{x}_m - \underline{x}_n\| < \frac{\varepsilon}{2} \qquad (1)$$

Next, as $(\underline{x}_{n_k})_{k=1}^{\infty}$ converges to $\underline{x}$, there exists $K\left(\frac{\varepsilon}{2}\right) \in \mathbf{N}$ such that

$$k \geq K\left(\frac{\varepsilon}{2}\right) \Rightarrow \|\underline{x}_{n_k} - \underline{x}\| < \frac{\varepsilon}{2} \qquad (2)$$

Let us set $L = \sup\left\{K\left(\frac{\varepsilon}{2}\right), M\left(\frac{\varepsilon}{2}\right)\right\}$. For $n \geq L$, we have $n \geq M\left(\frac{\varepsilon}{2}\right)$ and $n_L \geq L \geq M\left(\frac{\varepsilon}{2}\right)$, so by (1)

$$
\begin{aligned}
\|\underline{x}_n - \underline{x}\| &= \|\underline{x}_n - \underline{x}_{n_L} + \underline{x}_{n_L} - \underline{x}\| \\
&\le \|\underline{x}_n - \underline{x}_{n_L}\| + \|\underline{x}_{n_L} - \underline{x}\| \\
&< \frac{\varepsilon}{2} + \|\underline{x}_{n_L} - \underline{x}\| \qquad \text{(by } (1) \text{ as } n \ge M\left(\frac{\varepsilon}{2}\right) \text{ and } n_L \ge L \ge M\left(\frac{\varepsilon}{2}\right)\text{)(}\because \text{Cauchy)} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \qquad \text{(by } (2) \text{ as } L \ge K\left(\frac{\varepsilon}{2}\right)\text{)(}\because \text{Converges)}
\end{aligned}
$$

It follows that $(\underline{x}_n)$ converges to $\underline{x}$. ■

Now we can prove the important Cauchy criterion:

Theorem 6.14

A sequence in $\mathbf{R}^p$ is convergent iff it is a Cauchy sequence.

Proof

Convergent $\Rightarrow$ Cauchy (This was Lemma 6.11)

Cauchy $\Rightarrow$ Convergent

Let $(\underline{x}_n)$ be a Cauchy sequence. Form Lemma 6.12, the sequence is bounded. By the Bolzano-Weierstrass Theorem for sequence (Theorem 6.9), the sequence then has a convergent subsequence. By Lemma 6.13, the full sequence converges.