ITERATIVE DESIGN PROCEDURES IN PATTERN CLASSIFICATION

by

Lineu da Costa Barbosa

Memorandum No. ERL-M227

1 November 1967

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

The problem of classification will be ideally solved if there is no point x which is misclassified by the decision procedure. This will be the case, for example, if we can find f(x,w) such that

$$f(x,w) \geqslant 0 \qquad \text{for all } x \in A$$
$$f(x,w) < 0 \qquad \text{for all } x \in B$$

Definitions

The sets A and B will be said to be linearly-separable if there exists a function f(x,w) of the type (11) such that

$$f(x,w) = w_1 . x - w_o \geqslant 0 \qquad \text{for all } x \in A$$
$$\leqslant 0 \qquad \text{for all } x \in B$$

The sets A and B will be said to be strongly-linearly-separable if there exists f(x,w) of the type (11) such that

$$f(x,w) = w_1 . x - w_o > 0 \qquad \text{for all } x \in A$$
$$< 0 \qquad \text{for all } x \in B$$

It is known that the sets A and B are strongly-linearly-separable if and only if their closed convex hulls are disjoint. The closed convex hull of a set A is the minimum closed convex set containing A, or the intersection of all closed convex sets containing A, or the closure of the set of all points that can be written as a convex combination of points in A.

For simplicity, augment the vectors x to $a = (x,-1)$ (that is, write the number -1 as the $(\ell + 1)$th coordinate) and, considering the set $A \cup (-B)$, define

$$\mathcal{A} \triangleq \{ a = (x,-1);\ x \varepsilon A \cup (-B) \}$$

Then, the conditions stated in the definition of strongly-linearly-separable sets can be re-stated as follows:

The sets A and B will be said to be strongly-linearly-separable if there exists a vector $w \varepsilon E_{\ell+1}$ such that

$$w.a > 0 \qquad \text{for all} \quad a \varepsilon \mathcal{A} = \{ a = (x,-1);\ x \varepsilon A \cup (-B) \}$$

<u>Definitions</u>

The set $\mathcal{A}$ satisfying the above condition will be called strongly-linearly-separable. That is, <u>a set</u> is <u>strongly-linearly-separable</u> if it lies strictly in a half-space.

Similarly, <u>a set</u> $\mathcal{A} \varepsilon E_{\ell}$ is said to be <u>linearly-separable</u> if there exists a vector $w \varepsilon E_{\ell}$ such that

$$w.a \geqslant 0 \qquad \text{for all} \quad a \ \varepsilon \mathcal{A}$$

A pattern $a \varepsilon \mathcal{A}$ is said to be misclassified by a vector $w \varepsilon E_{\ell}$ if

$$w.a < 0$$

Under the hypothesis that the sets A and B are strongly-linearly-separable, the decision problem can be re-stated:

<u>Problem:</u>

Find a vector $g \in E_{\ell+1}$ such that

$g \cdot a > 0$ for all $a \in \mathcal{A} = \{a = (x,-1); x \in A \cup (-B)\}$

All of the practical pattern-recognition machines use only a finite number of patterns in A and B for design purposes It is worthwhile noting that, if the sets A and B are finite, then A and B linearly-separable implies A and B strongly-linearly-separable.

(Proof: A and B are linearly separable. Then, there exists w_1 and w_0 such that

$$w_1 \cdot x \geqslant w_0 \qquad \text{for } x \in A$$

$$w_1 \cdot x < w_0 \qquad \text{for } x \in B$$

Consider

$$\max_{x \in B} w_1 \cdot x = w_{01} < w_0$$

Let

$$\hat{w}_0 = \frac{w_0 + w_{01}}{2}$$

Then,

$$w_{01} < \hat{w}_0 < w_0$$

Hence,

$$w_1 \cdot x \geqslant w_0 > \hat{w}_0 \qquad \text{for all } x \in A$$

$$w_1 \cdot x \leqslant w_{01} < \hat{w}_0 \qquad \text{for all } x \in B$$

Therefore, there exists $\hat{w}_0$ satisfying the strong-linear-separability condition.)

Nevertheless, for a finite set $\mathcal{A}$, linear-separability does not imply strong-linear-separability.

Solutions to the Decision Problem.

Two classical solutions to the decision problem will be presented below. The first, due to F. Rosenblatt, is very important for its simplicity and for the fact that it leads to a solution in a finite number of steps, whenever the set $\mathcal{A}$ is strongly-linearly-separable. The second solution is due to S. Agmon, T. S. Motzkin and I. J. Schoenberg and is also very important, particularly for the fact that it requires only that the set $\mathcal{A}$ be linearly-separable. The relaxation of the strong-linear-separability will permit its use in the solution of other problems, such as symmetric games, which are not solvable with Rosenblatt's algorithm (see Chapter II).

Rosenblatt's solution.

In the case where the sets A and B are finite and (strongly)-linearly-separable the decision problem has a very elegant solution, due to Rosenblatt (Ref. 1) (see also Novikoff (Ref. 2)).

Consider the strongly-linearly-separable set

$$\mathcal{A} = \{ a_j;\ j = 1, 2, \ldots, N : a_j \varepsilon E_\ell \}$$

Form a sequence S with the property that each vector a_j appears in S infinitely many times. This can be accomplished, for instance, forming S by the selection of the vectors $a_j \varepsilon \mathcal{A}$ following each other in a cyclic manner, as follows:

$$S \equiv \{a_1, a_2, \ldots, a_N, a_1, a_2, \ldots, a_N, a_1, \ldots\}$$

Re-label the elements of S according to its order of appearance:

$$S = \{a^1, a^2, \ldots\ldots, a^k, \ldots\ldots\}$$

This is by no means the only way of keeping the requirement that S must satisfy. Actually, the requirement on S is very loose, and one only wants to be sure that after the $n^{\underline{th}}$ element of the sequence, no matter how large n is, one still will find each element a_j of $\mathcal{A}$ infinitely often.

In the pattern recognition parlance the term "training sequence" is used in reference to S.

Consider also the sequence $\{g_n\}$ obtained by the iterative procedure:

$$\begin{aligned} g_1 &= \text{arbitrary, in } E_\ell \\ g_{n+1} &= g_n + a^n \quad \text{if} \quad g_n \cdot a^n \leq 0 \\ g_{n+1} &= g_n \qquad\quad \text{if} \quad g_n \cdot a^n > 0 \end{aligned} \tag{1.2}$$

Rosenblatt's theorem states the following:

<u>THEOREM</u> (Rosenblatt).[1] The iterative procedure (1.2) terminates in a vector $g \in E_\ell$ satisfying

$$g \cdot a_j > 0 \quad \text{for all} \quad a_j \in \mathcal{A}$$

in a finite number of steps.

Rosenblatt's solution is very attractive since it requires no memorization of the past history of the process.

---

[1] See references 1, 2, 13.

One can also say that the most important part of the design is the training, that is, the adjustment of the parameters. This makes the machine highly adaptive, in the sense that a "learning" period(where the vector g is readjusted) can always follow any modification of the set $\mathcal{A}$. That will be the case, for instance, if the set $\mathcal{A}$ is a sample of an infinite set of patterns. Thus, new samples can be incorporeted in $\mathcal{A}$ or even a new set of samples may be used. (For example, if the machine is a spoken word recognizer and was previously trained to recognize words spoken by a particular operator, the change of operator may requiré a new learning period in which the machine would become "familiar" with the new operator.)

Rosenblatt's procedure is an <u>error correcting</u> procedure: patterns belonging to the training sequence S are presented to the machine and at each step n in which an error is detected, a correction is made by adding to g_n the pattern a^n which was misclassified.

The success of such procedure, however, is highly dependent on the fact that the set $\mathcal{A}$ must be strongly-linearly-separable[1]. Otherwise, since for every g_n there will always exist some $a_j \varepsilon \mathcal{A}$ such that $g_n . a_j \leq 0$, termination is impossible.

<u>Agmon-Motzkin-Schoenberg's Solution.</u>

The above authors, studying a system of linear inequalities, proposed some procedures leading to its solution. Among several slightly different procedures presented

[1] In some cases, the probability of N points in E_ℓ being (strongly-)linearly-separable is studied in Refs. 13 and 14.

in the work of Agmon (Ref. 3), one can find the following (which was selected for its simpler formulation using the notations of the preceding section):

Consider again the training sequence S obtained from vectors $a_j \in \mathcal{A}$, where $\mathcal{A}$ is a linearly separable set. Form a sequence $\{g_n\}$, where g_n is given by the following procedure:

$$g_1 = \text{arbitrary in } E_\ell$$

$$g_{n+1} = g_n + \lambda \left| a^n . g_n \right| \frac{a^n}{\|a^n\|^2} \quad \text{if } a^n . g_n < 0 \qquad (1.3)$$

$$g_{n+1} = g_n \qquad \text{if } a^n . g_n \geqslant 0$$

Then, Agmon and also Motzkin-Schoenberg (Ref. 4) proved the following:

THEOREM (Agmon-Motzkin-Schoenberg).

The iterative procedure (1.3) converges to a vector $g \in E_\ell$ satisfying

$$g . a_j \geqslant 0 \quad \text{for all} \quad a_j \in \mathcal{A}$$

provided

$$0 < \lambda \leq 2.$$

The rate of convergence is shown to be exponential.

This procedure, like Rosenblatt's, requires no memorization of its past history and, for that reason, is of great practical interest. Also, it is applicable to sets which are linearly separable, as opposed to Rosenblatt's procedure, which requires strong-linear-separability. On the other hand, termination is not guaranteed in Agmon's procedure (aside from some special cases).

It is interesting to note that in Agmon's procedure, whenever an error occurs, the correction is made proportional to $\left|a^n.g_n\right|$. This has some similarities with other methods (gradient methods) in which the corrections are made proportional to the error, leading also to exponential rate of convergence (Ref. 5,6).

Rosenblatt's and Agmon-et-al's solutions are representative of two groups of solutions so far presented to the problem. The first group makes use of a correction without regard to the magnitude of the error and leads to termination whenever a solution exists. The second group makes a correction proportional to the magnitude of the error. Termination is in general sacrificed for the possibility of relaxing the condition of strong-linear-separability to linear-separability.

We present here an outline of the remaining chapters. In chapter 2 we show the intimate connections that exists between the basic pattern recognition problem and problems of matrix game and linear programming. In this connection we point out, however, that the dimensionality usually associated with the pattern recognition problem renders ineffective the usual linear programming techniques. In chapter 3 we introduces a one parameter ($0 \leq \epsilon \leq 1$) family of continuous-time algorithms, which for $\epsilon = 0,1$ reduce to the Rosenblatt and Agmon-Motzkin-Schoenberg algorithms respectively. Some important termination-convergence properties of this class of algorithms are shown. In chapter 4, we treat the time-discrete versions of the algorithms introduced in chapter 3. Although the results here are a little weaker than the corresponding ones for continuous-time, they do, nevertheless, represent important improvements over previously known results in the area. We conclude in chapter 5 with some not too complete computer runs. Some comments on related directions of inquiry are also presented.

CHAPTER II

RELATED PROBLEMS

Games.

Perhaps the problem of solving a matrix game is the most intimately related with the problems introduced in the previous chapter.

A zero-sum, two-person game, can be summarized as follows (Ref. 7):

Each player has available a finite set of <u>pure strategies</u>. A <u>pay-off matrix</u> A is defined by the quantities a_{ij} earned by Player I (or lost by Player II since it is a zero-sum game) if Player I uses pure strategy i and Player II uses pure strategy j.

A <u>mixed strategy</u> can be used by the players by associating, for instance, a probability to each of the pure strategies available (the earning of the players will be regarded, in this case, as average earnings). If one represents the pure strategy i by the unity vector $e_i =$ $= (0,0,\ldots, 1, 0,0, \ldots, 0)$, the 1 appearing in the i^{th} place, a mixed strategy will be represented by the vector

$$x = \sum_i x_i e_i \;, \qquad x_i \geq 0 \;, \quad \sum_i x_i = 1$$

If mixed strategies are used, the yield to Player I will be:

$$x \cdot Ay = \sum_i \sum_j x_i a_{ij} y_j$$

where x = Player I (mixed) strategy
where y = Player II (mixed) strategy

Call X (Y) the set of all possible strategies x (y) for Player I (Player II).

For a fixed strategy $x \in X$, consider the minimum yield to Player I when Player II runs his strategy y over Y. It is the interest of Player I that this minimum is maximized. Player I will, then, look for a strategy that yields him, at least,

$$v_I = \max_x \min_y \; x.Ay \quad .$$

Similarly, for a fixed strategy $y \in Y$, consider the maximum yield to Player I when he runs his strategy x over X. Player II is interested in minimizing this maximum yield. He will look for a strategy that prevents the yield of Player I from being greater than

$$v_{II} = \min_y \max_x \; x.Ay \quad .$$

The main theorem in matrix games states:

<u>THEOREM</u> (Von Neumann)

$v_I = v_{II} = v$.

The number v is called <u>value of the game</u>.

An <u>optimum strategy</u> x_o for Player I (or y_o for Player II) is one such that

$x_o.Ay \geqslant v$ for all $y \in Y$

$x.Ay_o \leqslant v$ for all $x \in X$

Obviously,

$x_o.Ay_o = v$.

Since any strategy y is a convex combination of pure strategies for Player II, x_o is an optimum strategy for Player I if and only if

$$x_o^T A \geqslant v\underline{1}$$

where $\underline{v}$ is the line vector $\underline{v}$ $(v, v, \ldots, v)$ and the inequality signal holds for each component of the vectors involved.

Similarly, for Player II,

$$Ay_0 \leqslant \underline{v}^T$$

where $\underline{v}^T$ is the transpose of $\underline{v}$.

Definition. A symmetric game is one having a skew-symmetric matrix, that is, $A^T = -A$ (or $a_{ij} = -a_{ji}$).

In this case, every strategy available to Player I is also available to Player II.

The value of the game, in this case, is zero since, if x_0 is an optimum strategy for Player I, $y_0 = x_0$ is also an optimum strategy for Player II and then,

$$x_0 . Ax_0 = 0$$

by the fact that the matrix A is skew-symmetric.

The following fact is very important to relate matrix games with pattern recognition problems.

Theorem (von Neumann - Brown): Given any game with matrix A, there exists a symmetric game with matrix

$$B = \begin{pmatrix} \underline{0} & A & -\underline{1} \\ -A^T & \underline{0} & \underline{1} \\ \underline{1} & -\underline{1} & 0 \end{pmatrix}$$

(where $\underline{0}$ is a matrix of zeros, of suitable dimensions, and $\underline{1}$ is a vector of suitable dimension having all components equal to 1) such that every optimum strategy (x, y, λ) for B satisfies:

$$\sum_i x_i = \sum_i y_i = a$$

and $a^{-1}x$ and $a^{-1}y$ are optimum strategies for the game with the matrix A. The value of this game is

$$v = \lambda/a.$$

Hence, from now on, only symmetric games will be considered and it will be understood that the results apply to any matrix game after its reduction to a symmetric game.

<u>Theorem</u>. An optimum strategy for a (symmetric) game with matrix A is a vector x satisfying

$$x^T A \geqslant \underline{0} \qquad (\text{or } Ax \leqslant \underline{0})$$
$$\sum_i x_i = 1, \quad x_i \geqslant 0$$

This theorem is a consequence of the properties of optimum strategies and the fact that a symmetric game has value zero.

One can look at the matrix A as a set $\mathcal{a}$ of vectors a_j, the columns of A:

$$A = \begin{bmatrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ \vert & \vert & a_j & \vert & \vert \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \end{bmatrix}.$$

Hence, an optimum strategy is a vector x satisfying

$$x . a_j \geqslant 0 \quad \text{for all } a_j \in \mathcal{a}$$
$$\sum_i x_i = 1, \; x_i \geqslant 0$$

At this point, the similarities between this problem and the one introduced in Chapter I are evident. The main difference is in the constraint that a strategy must satisfy.

The inequalities

$$x_i \geqslant 0 \quad \text{and} \quad x.a_j \geqslant 0$$

are similar and one can incorporate the first into the second by adding the vectors $e_i^T\ (0,0,\ldots,0,1,0,\ldots,0)^T$ (where the number 1 occurs in the ith. place) to the set $\mathcal{A}$. Call the new set $\mathcal{A}'$. From now on, the game problem is to find a vector x satisfying

$$x.a_j \geqslant 0 \text{ for all } a_j \in \mathcal{A}'$$
$$\sum_i x_i = 1.$$

The last constraint can be dropped, since if x is an optimal strategy, any vector $g = \alpha x,\ \alpha > 0$ will satisfy

$$g.a_j \geqslant 0 \quad \text{for all } a_j \in \mathcal{A}' \tag{2.1}$$

and conversely, given $g \neq 0$ satisfying (2.1) the vector

$$x = (\sum_i g_i)^{-1} g$$

will be an optimal strategy for the original game, that is,

$$x.a_j \geqslant 0 \quad \text{for all } a_j \in \mathcal{A}'$$
$$\sum_i x_i = 1, \quad x_i \geqslant 0$$

The equivalence between the problems is, then, complete. In the pattern recognition problem, the set $\mathcal{A}$ can be linearly separable or non-linearly separable. In the game problem, however, $\mathcal{A}$ is always linearly-separable.

Linear Programming

The connection of linear programming with matrix game is well known.[1] Therefore, its connection to the pattern recognition problem can be easily understood.

The problem of linear programming is the following:

Given the vectors c and b and the matrix A (of suitable dimensions), find a vector x which maximizes the inner product $c.x$, subject to the constraints

$$Ax \leqslant b, \quad x_i \geqslant 0$$

(where the inequality signals for vectors again stand for inequalities for each component of the vectors).

To each problem of linear programming one can formulate its dual:

Find y which minimizes

$$y.b$$

subject to the constraints

$$y^T A \geqslant c, \; y_i \geqslant 0$$

Solving the linear programming problem and its dual is equivalent to solving the symmetric game with matrix

$$B = \begin{bmatrix} \underline{0} & -A^T & c \\ A & \underline{0} & -b \\ -c & b & 0 \end{bmatrix}$$

An optimal strategy for this symmetric game, z (x, y, λ), with $\lambda > 0$, has the properties:

---

[1] See Ref. 7 .

$$(\lambda^{-1}y)^T A \geqslant c \qquad A(\lambda^{-1}x) \leqslant b$$

$$cx = \max_{x'} c.x' = b.y = \min_{y'} b.y'$$

where x' and y' satisfy the respective constraints.

Conversely, if $\bar{x}$ and $\bar{y}$ are solutions to the linear programming problem and its dual, the vector

$$z = (x,y,\lambda)$$

with

$$\lambda = 1/(1+\sum_i \bar{x}_i+\sum_i \bar{y}_i)$$
$$x = \lambda\ \bar{x}$$
$$y = \lambda\ \bar{y}$$

will be an optimal strategy for the game with matrix B.

Any game can be directly shown to be equivalent to a linear programming problem as follows:

Suppose the game has matrix A and value $v > 0$ (unknown)[1].
To find an optimal strategy one must find x such that

$$x^T A \geqslant v\underline{1}$$

subject to

$$x.\underline{1} = 1, \quad x \geqslant 0$$

Put

$$(v)^{-1}x = w \qquad \text{(v is a scalar)}$$

One must find w such that

$$w^T A \geqslant \underline{1}$$

with

$$w.\underline{1} = (v)^{-1}, \; w \geqslant 0$$

---

1 It is no restriction to suppose $v > 0$ since a positive constant can be added to the elements of A without changing the set of optimal strategies.

Hence, the game for Player I is equivalent to

$$\min w.\underline{1}$$

subject to

$$w^T A \geqslant \underline{1}$$

$$w \geqslant 0$$

which is a linear programming problem. It is easy to see that the linear program for Player II is the dual of the above.

Since the solution of a game can be found using linear programming techniques, the solution of the pattern recognition problem also can. It is interesting to note that to solve a linear program one must find a starting point satisfying $Ax \leq b$, which is the pattern recognition problem. When the simplex algorithm is used for the solution of the linear programming problem, most of the time this starting point can be found trivially (for instance, if $b \geqslant 0$, putting $x = 0$). A non trivial solution is, then, approached by the algorithm, seeking the maximum of the objective function $c.x$.

The problem of pattern recognition can be formulated as a linear programming problem as follows:

<u>Pattern recognition problem</u>.

find x such that $Ax \geqslant \theta\underline{1}$

<u>Linear programming problem</u>.

maximize

$$-\underline{1}.v$$

subject to

$$Ax + v \geqslant \theta\underline{1}$$

$$v \geqslant 0$$

With the addition of the artificial vector v, the solution of the linear program can start with $x = 0$, $v = \theta\underline{1}$. If there exists a x such that $Ax \geqslant \theta\underline{1}$, the objective function will drive v to zero, and therefore, it will result

$$Ax \geqslant \theta\underline{1}.$$

In case there exists no x satisfying $Ax \geqslant \theta\underline{1}$, the linear programming will give an optimum x in the sense that, for the misclassified patterns,

$$\sum_j [\theta - a_j \cdot x]$$

will be minimum.

The only thing missing in the linear program formulation of the pattern recognition problem is the constraint $x_i \geqslant 0$. This is not very important since it is well known that the positivity constraints are not restrictive and any linear program without it can be reduced to the canonical form by putting $x = y - z$, with $y \geqslant 0$ and $z \geqslant 0$.

A linear program can always be solved by the simplex algorithm in a finite number of steps. Nevertheless, the application of the simplex algorithm involve the inversion of matrices. If the dimensionality of the problem is very large, a large number of large matrices must be inverted before a solution is reached. To deal with such large problems the fast growing field of Decomposition of Large Scale Systems is required. In this respect it is interesting to note that, instead of using the conventional techniques of linear programming to solve pattern recognition problems, <u>it is more advantageous to use pattern recognition techniques to solve large linear programming problems</u>. Agmon's procedure can be used to handle such large linear programming problems, by its reduction to a symmetric game. In this respect, we would like to emphasize the fact that the large dimensionality of the problem does not play an important role in these error correcting procedures.

<u>Gradient procedures</u>

The iterative procedures presented so far are closely related with gradient procedures. This fact is more evident in the procedure of Agmon-et-al where the corrections are made proportional to $|a_j \cdot g_n|$ for the misclassified patterns a_j. If one is trying to minimize, for instance,

$$\phi_n = \sum_{j \in J_n} |a_j \cdot g_n|^{1+\epsilon}$$

where $J_n = \left\{ j;\ a_j \cdot g_n \leqslant 0\ ,\ a_j \in \mathfrak{A} \right\}$

one will get, for the gradient of ϕ_n ,

$$\text{grad } \phi_n = -(1+\epsilon) \sum_{j \in J_n} |a_j \cdot g_n|^{\epsilon}\, a_j$$

For $\epsilon = 1$, this gradient suggests a direction of change of g_n very similar to Agmon-et-al's. For $\epsilon \rightarrow 0$, the correction will be similar to Rosenblatt's.

In the next two chapters a great deal of work motivated by these gradient ideas will be presented.

CHAPTER III

CONTINUOUS-TIME PROCEDURES [1]

As seen in Chapters I and II, the classification problem and some related ones can be solved if a vector g is found such that

$$g.a_j \geqslant 0 \ (>0), \quad g \in E_\ell$$

for every a_j in a finite set $\mathcal{a}$ of vectors belonging to E_ℓ, the ℓ-dimensional Euclidean space.

If the set $\mathcal{a}$ is strongly-linearly-separable, any scheme of the type of Rosenblatt's procedure can be used to solve the problem in a finite number of steps. In case the set $\mathcal{a}$ is only linearly-separable, the solution would be found by schemes of the type of Agmon-et-al iterative procedure. In this case, convergence in finite number of steps cannot be guaranteed.

In problems of game theory and linear programming one can always be sure that the set $\mathcal{a}$ is linearly-separable. In case of classification, nothing can be said a-priori. Therefore, it is of great interest to find a scheme having the advantages of both types of procedures, that is, that will terminate whenever the set is strongly-linearly-separable and will converge to a solution if the

---

[1] Most of the schemes and proofs in this Chapter were suggested by Prof. E. Wong.

set $\mathcal{A}$ is only linearly-separable. More, in case the set is not linearly-separable, it is desirable that the procedure converges to some g which either will indicate the non linear-separability character of $\mathcal{A}$ or is optimal in some sense (that is, converges to a g that will work as well as possible).

Motivated by these ideas, gradient type of procedures where investigated. As indicated in Chapter II, the procedures of Rosenblatt and Agmon-et-al could be connected to each other if one looks at the corrections as being made against the direction of the gradient of the function

$$\varphi = \sum_{j \varepsilon J} |a_j \cdot g|^{1+\epsilon} \qquad J = \{ j \; ; \; a_j \cdot g \leqslant 0 \; , \; a_j \varepsilon \mathcal{A} \}$$

This suggests that, by keeping $0 < \epsilon < 1$ one might have a scheme exhibiting the properties of both Rosenblatt's and Agmon-et-al's procedures.

This Chapter will consider only continuous-time procedures (that is, where the vector g changes continuously with time). The doted quantities are the derivatives with respect to time.

Consider the following procedure:

$$\text{(P1)} \qquad \dot{g}(t) = \sum_{j \varepsilon J(t)} |a_j \cdot g(t)|^{\epsilon} a_j \quad , \; J(t) = \{ j ; a_j \cdot g(t) < 0, a_j \varepsilon \mathcal{A} \}$$

3.1. <u>THEOREM</u>. The procedure P1 converges, for $\epsilon > 0$, to a solution of $g.a_j \geqslant 0$ for all $a_j \varepsilon \mathcal{A}$.

Proof: Define

$$\varphi(t) = \sum_{j \varepsilon J(t)} |a_j \cdot g|^{1+\epsilon} \geqslant 0$$

Call $\qquad \varphi_j = |a_j \cdot g| \quad , \; j \varepsilon J(t)$

One can see that $\varphi_j > 0$ implies $a_j \circ g < 0$, in which case,

$$\varphi_j = -a_j \circ g$$

and

$$\mathring{\varphi}_j = -a_j \cdot \mathring{g}$$

Therefore, for $\epsilon > 0$

$$\begin{aligned} \mathring{\varphi}(t) &= -\sum_{j \in J} (1+\epsilon) |a_j \cdot g|^{\epsilon} a_j \cdot \mathring{g} \\ &= -(1+\epsilon) \sum_{j,k \in J} \sum |a_j \cdot g|^{\epsilon} a_j \cdot a_k \; |a_k \cdot g|^{\epsilon} \\ &= -(1+\epsilon) \, \| \mathring{g} \|^2 \leqslant 0 \end{aligned} \tag{3.1}$$

On the other hand,

$$\begin{aligned} \frac{d}{dt} \frac{1}{2} \| g \|^2 = g \cdot \mathring{g} &= \sum_{j \in J} |a_j \cdot g|^{\epsilon} a_j \cdot g \\ &= -\varphi(t) \leqslant 0 \end{aligned}$$

By Schwarz's inequality,

$$|\varphi(t)|^2 \leqslant \| g \|^2 \cdot \| \dot{g} \|^2$$

Hence,

$$\| \dot{g} \|^2 \geqslant \frac{|\varphi(t)|^2}{\| g \|^2}$$

and since

$$\frac{d}{dt} \| g \|^2 \leqslant 0$$

putting $g_0 = g(0)$,

$$\| g \|^2 \leqslant \| g_0 \|^2$$

resulting

$$\| \dot{g} \|^2 \geqslant \frac{\varphi^2}{\| g_0 \|^2} \tag{3.2}$$

Combining Eq. III.2 with Eq. III.1,

$$\dot{\varphi} \leq -\frac{(1+\epsilon)}{\|g_o\|^2}\varphi^2 = -K\varphi^2 \qquad K > 0$$

Solving,

$$\varphi(t) \leq \frac{1}{\frac{1}{\varphi_o} + Kt}$$

Hence, $\varphi(t) \to 0$ for every $\epsilon > 0$.

But $\varphi(t) = 0$ implies

$$\dot{g} = 0$$

and

$$a_j \cdot g \geqslant 0 \quad \text{for every} \quad a_j \in \mathcal{a}$$

3.2. <u>THEOREM</u>. Under existence of a vector y such that $y.a_j \geqslant \theta > 0$ for every $a_j \in \mathcal{a}$, the procedure P1 terminates (that is, will reach the equilibrium point in a finite time) for $0 \leqslant \epsilon < 1$ in a solution of $a_j \cdot g \geqslant 0$ for all $a_j \in \mathcal{a}$. Furthermore, choosing g_o such that $g_o \cdot y \geqslant \theta$ (take for instance, g_o equal to one of the a_j's), $\|g\| \neq 0$.

Proof: a) For $0 < \epsilon < 1$, consider

$$\dot{g} \cdot y = \sum_{j \in J} |a_j \cdot g|^\epsilon a_j \cdot y \geqslant \theta \sum_{j \in J} |a_j \cdot g|^\epsilon \qquad (3.3)$$

Using Jensen's inequality $\left(\sum |x_i|^\epsilon \geqslant \left(\sum |x_i|^{1+\epsilon}\right)^{\epsilon/1+\epsilon} \right)$

$$\dot{g} \cdot y \geqslant \theta \left(\sum_{j \in J} |a_j \cdot g|^{1+\epsilon} \right)^{\epsilon/1+\epsilon} = \theta \varphi^{\epsilon/1+\epsilon} > 0$$

By Schwarz's inequality,

$$\|\dot{g}\|^2 \geqslant \frac{(\dot{g} \cdot y)^2}{\|y\|^2} \geqslant \frac{\theta^2}{\|y\|^2} \varphi^{2\epsilon/1+\epsilon}$$

Therefore, using Eq. 3.1,

$$\dot{\varphi}(t) = -(1+\epsilon)\|\dot{g}\|^2 \leqslant -K_1 \varphi^{2\epsilon/1+\epsilon}, \quad K_1 > 0 \qquad (3.4)$$

Integrating, $\epsilon \neq 1$,

$$\varphi^{\frac{1-\epsilon}{1+\epsilon}} \leqslant \frac{1+\epsilon}{1-\epsilon}\left(K_0 - K_1 t\right)$$

and hence, $0 \leqslant \varphi \leqslant (A - Bt)^{\frac{1+\epsilon}{1-\epsilon}}$, A, B > 0

with

$$\varphi\left(\frac{A}{B}\right) = 0 \qquad \text{for} \qquad 0 < \epsilon < 1$$

b) For $\epsilon = 0$, consider, on the one hand,

$$\frac{1}{2}\frac{d}{dt}\| g \|^2 = \dot{g} \cdot g = -\varphi \leqslant 0$$

hence,

$$\| g(t) \| \leqslant \| g(0) \| \triangleq \| g_0 \|$$

On the other hand,

$$\dot{g} \cdot y > \theta \quad , \qquad g \cdot y > \theta t + g_0 \cdot y > \theta (t+1)$$

which yields by Schwarz's inequality

$$\| g \| \geqslant \frac{\theta}{\| y \|}(t+1) \ ; \quad \text{hence}, \quad t+1 \leqslant \frac{\| y \| \cdot \| g_0 \|}{\theta}$$

To show that $\| g \| \neq 0$, for $0 \leqslant \epsilon < 1$, consider, from (Eq. 3..3):

$$y \cdot \dot{g} \geqslant 0$$

which yields

$$y \cdot g \geqslant y \cdot g_0 \geqslant \theta$$

Hence,

$$\| g \|^2 \geqslant \frac{(y \cdot g)^2}{\| y \|^2} \geqslant \frac{\theta^2}{\| y \|^2}$$

3.3. COROLLARY: If there exists $w \in E_\ell$ such that

$$w.a_j \geqslant 0 \quad \text{for every} \quad a_j \in \mathcal{A}, \quad w \neq 0$$

(that is, $\mathcal{A}$ is linearly-separable), then, procedure P1 will converge to a $g \neq 0$ for $\epsilon > 0$ provided $g_0 . w > 0$. The latter condition can be fulfilled by putting, for instance, $g_0 = \sum_{a_j \in \mathcal{A}} a_j$ (Since there exists $a^* \in \mathcal{A}$ such that $w . a^* > \theta > 0$).

Proof: Consider

$$\dot{g} . w \geqslant 0$$

Hence

$$g . w \geqslant g_0 . w > \theta$$

By Schwarz's inequality,

$$\| g \| \geqslant \frac{|g \cdot w|}{\| w \|} \geqslant \frac{\theta}{\| w \|}$$

REMARKS. The set $J = \left\{ j \; ; \; a_j . g < 0 \; , \; a_j \in \mathcal{A} \right\}$ gives the vectors (patterns) a_j $(j \in J)$ which are "misclassified". Therefore, the procedure drives to zero the function

$$f(t) = \sum_{j \in J} \left| a_j . g \right|^{\alpha} \quad , \quad \alpha > 0 \text{ arbitrary,}$$

the summation of the α^{th} power of the error $| a_j . g |$ for the misclassified patterns. Note also that the distance from a pattern to the separating hyperplane defined by g is

$$a_j . \frac{g}{\| g \|}$$

and hence, f(t) is the sum of functions of the distance of the misclassified patterns to the separating hyperplane.

The procedure converges for all $\epsilon > 0$ at least at a rate of t^{-1} and in case of strong-linear-separability, the procedure terminates for $0 \leqslant \epsilon < 1$.

For $\epsilon = 1$, in this case, (eq. 3.4) shows that convergence takes place at least as fast as an exponential with time.

In case of linear-separability, by picking $g_0 = \sum_{a_j \in \mathcal{A}} a_j$ one can be sure that P1 will drive g to a non trivial solution of $a_j \cdot g \geqslant 0$ for all $a_j \in \mathcal{A}$.

In case of non linear-separability, P1 will drive g to zero.

The fact that for $\mathcal{A}$ not linearly-separable the vector g given by procedure P1 will converge to zero indicates the non linear-separability of $\mathcal{A}$. If it is desired to have a useful g, again based on gradient techniques, a modification of P1 can be made to obtain a vector g which converges to a local minimum of the function

$$\varphi(g) = \sum_{j \in J} \left| a_j \cdot g \right|^{1+\epsilon} \qquad J = \left\{ j \ ; \ a_j \cdot g < 0 \ , \ a_j \in \mathcal{A} \right\}$$

Let's consider the problem:

$$\min_g \varphi(g) \qquad \text{subject to} \qquad \|g\|^2 = 1$$

The Lagrangian for this problem is:

$$L(g, \lambda) = \varphi(g) - \lambda \|g\|^2$$

At a minimum point $\hat{g}$,

$$\frac{\partial L}{\partial \hat{g}} = -(1+\epsilon) \sum_{j \in J} \left| a_j \cdot \hat{g} \right|^{\epsilon} a_j - 2\lambda \hat{g} = 0$$

Hence,

$$\hat{g} = \frac{-(1+\epsilon)}{2\lambda} \sum_{j \in J} \left| a_j \cdot \hat{g} \right|^{\epsilon} a_j$$

The value of λ can be obtained by taking the inner product of $\frac{\partial L}{\partial \hat{g}}$ with $\hat{g}$:

$$\frac{\partial L}{\partial \hat{g}} \cdot \hat{g} = 0 = (1+\epsilon) \sum_{j \in J} \left| a_j \cdot \hat{g} \right|^{1+\epsilon} - 2\lambda \| \hat{g} \|^2$$

Therefore,

$$2\lambda = (1+\epsilon) \frac{\varphi(\hat{g})}{\| \hat{g} \|^2} = (1+\epsilon)\varphi(\hat{g})$$

Hence, $\hat{g}$ is a solution of

$$\hat{g} = \frac{-\sum_{j \in J} \left| a_j \cdot \hat{g} \right|^{\epsilon} a_j}{\varphi(\hat{g})}$$

Hence, the obvious modification of P1 would be:

(P2) $$\dot{g} = \frac{\varphi(g)}{\| g \|^2} g + \sum_{j \in J} \left| a_j \cdot g \right|^{\epsilon} a_j$$

3.4. <u>THEOREM</u>. The vector g given by procedure P2 converges to a local minimum of the function $\varphi(g)$ for $\epsilon > 0$. Further, if there exists y with $\| y \| = 1$ such that $y . a_j > \theta > 0$ for every $a_j \in \mathcal{Q}$, then, starting with $g_0 = g(0)$ such that $g_0 . y > \hat{\theta} > 0$ (put $g_0 = a_j / \| a_j \|$ for instance), procedure P2 will terminate for $0 \leqslant \epsilon < 1$, with $\varphi = 0$.

Proof: Take $\| g_0 \| = 1$ for simplicity.

a) Convergence;

(i) $\varphi(g) \geqslant 0$

(ii) $$g . \dot{g} = \frac{\varphi}{\| g \|^2} \| g \|^2 + \sum_{j \in J} \left| a_j \cdot g \right|^{\epsilon} a_j \cdot g = \varphi - \varphi = 0$$

and hence,

$$\frac{1}{2} \frac{d}{dt} \| g \|^2 = g \cdot \dot{g} = 0$$

Therefore,

$$\| g \| = \| g(0) \| = 1$$

(iii) Taking the inner product of $\dot{g}$ with P2,

$$\| \dot{g} \|^2 = \frac{\varphi(g)}{\| g \|^2} g \cdot \dot{g} + \sum_{j \in J} \left| a_j \cdot g \right|^{\epsilon} a_j \cdot \dot{g} = \sum_{j \in J} \left| a_j \cdot g \right|^{\epsilon} a_j \cdot \dot{g}$$

on the other hand,

(iv)
$$\dot{\varphi} = -(1+\epsilon)\sum_{j\in J}|a_j\cdot g|^{\epsilon} a_j \dot{g} = -(1+\epsilon)\|\dot{g}\|^2 \leqslant 0$$

By (i) and (iv), φ must converge, and therefore,

$\dot{\varphi} \to 0$, which implies

$\dot{g} \to 0$, $g \to g^*$

Hence, since $\varphi(g) \geqslant 0$ for all g , if $\varphi \to 0$, the convergence to a minimum is proved. If $\varphi \to \varphi^* \neq 0$, then

$g \to g^*$

g^* being the solution of $\dot{g} = 0$, that is

$$g^* = \frac{-\sum_{j\in J^*}|a_j\cdot g^*|^{\epsilon} a_j}{\varphi^*(g^*)}$$

But this was shown to be a local minimum of $\varphi(g)$ subject to $\|g\|^2 = 1$.

(b) To show termination under existence of y , $\|y\| = 1$, such that $y.a_j > \theta > 0$ for all $a_j \in \mathcal{Q}$, consider, for, $0 < \epsilon < 1$,

$$\dot{g}\cdot y \geqslant \frac{\varphi}{\|g\|^2} g\cdot y + \theta\sum_{j\in J}|a_j\cdot g|^{\epsilon} \qquad (3,5)$$

on one hand,

$$\dot{g}\cdot y \geqslant \frac{\varphi}{\|g\|^2}(g\cdot y) = \varphi(g\cdot y)$$

and therefore, integrating

$$g\cdot y \geqslant g_0\cdot y\, e^{\int_0^t \varphi d\tau} \geqslant g_0\cdot y > \hat{\theta} > 0 \qquad (3,6)$$

On the other hand, by the use of Jensen's inequality, Eq. 3.5 gives

$$\|\dot{g}\|^2 \geqslant |\dot{g}\cdot y|^2 \geqslant (\varphi\hat{\theta} + \theta\varphi^{\epsilon/1+\epsilon})^2 \geqslant \theta^2\varphi^{2\epsilon/1+\epsilon}$$

Hence, (iv) gives

$$\dot{\varphi} \leqslant -(1+\epsilon)\,\theta^2\varphi^{2\epsilon/1+\epsilon}$$

and, like in theorem 3.2, termination occurs for $0 < \epsilon < 1$, with $\varphi = 0$.

To show termination for $\epsilon = 0$, it is enough to remark that from Eq. 3.5 and Eq. 3.6(which hold also for $\epsilon = 0$)

$$\dot{g}\cdot y \geqslant \theta > 0$$

and therefore,

$$g\cdot y \geqslant \theta t + g_o\cdot y$$

Hence,

$$1 = \|g\| \geqslant \frac{g\cdot y}{\|y\|} > \theta t + \hat{\theta}$$

The inequality cannot hold indefinitely. Thus, at some time, termination occurs. To see that $\varphi \to 0$, from Eqs. 3.5, 3.6

$$\dot{g}\cdot y \geqslant \frac{\varphi}{\|\dot{g}\|^2}\hat{\theta} > 0$$

and since $\dot{g}\cdot y \to 0$, $\varphi \to 0$.

3.5. THEOREM. If there exists $w \in E_\ell$ such that $w \neq 0, w\cdot a_j \geqslant 0$ for every $a_j \in \mathcal{A}$ (that is, $\mathcal{A}$ is linearly-separable), then, procedure P2 will give a vector g converging to a solution of

$$g\cdot a_j \geqslant 0$$

for all $a_j \in \mathcal{A}$, $\|g\| = 1$, provided $g_o.w > 0$. The latter condition can be fulfilled (there exists $a^* \in \mathcal{A}$ such that $w.a^* > \theta > 0$) by putting, for instance, $g_o = \sum_{a_j \in \mathcal{A}} a_j$. The rate of convergence is at least $1/t$.

Proof: In a similar way as in Theorem 3.4, one can get

$$\dot{g}.w \geqslant \frac{\varphi}{\|g\|^2} g.w \tag{3.5'}$$

and hence,

$$g.w \geqslant g_o.w \geqslant \theta > 0 \tag{3.6'}$$

Therefore, $\dot{g}.w \geqslant \frac{\varphi}{\|g\|^2} \theta > 0$

and since $\dot{g}.w \to 0$, $\varphi \to 0$.

To show that the rate of convergence is at least $1/t$, consider, by Schwarz's inequality,

$$\|\dot{g}\|^2 \geqslant \frac{(\dot{g}\cdot w)^2}{\|w\|^2} \geqslant \frac{\varphi^2}{\|g\|^4 \|w\|^2} \theta$$

Using (iv) of Theorem 3.4, and the fact that $\|g\| = 1$,

$$\dot{\varphi} = -(1+\epsilon)\|\dot{g}\|^2 < - \frac{\theta}{\|w\|^2} \varphi^2$$

By integration,

$$\frac{1}{\varphi} \geqslant \frac{1}{\varphi_0} + \frac{\theta}{\|w\|^2} t$$

which yields

$$\varphi \leqslant \frac{1}{\frac{1}{\varphi_0} + \frac{\theta}{\|w\|^2} t}$$

Hence,

$$\varphi \to 0 \quad \text{and} \quad g \to g^* \quad \text{as} \quad t \to \infty$$

such that

$$\| g^* \| = 1$$

$$g^*.a_j \geqslant 0 \quad \text{for all } a_j \in \mathcal{Q}$$

REMARKS. Procedure P2 with $\epsilon > 0$ will always lead to a local minimum of the function

$$\varphi = \sum_{j \in J} |a_j . g|^{1+\epsilon}$$

keeping $\| g \| = 1$.

When $\mathcal{Q}$ is linearly-separable, however, P2 would drive φ to zero at the rate of at least 1/t.

When $\mathcal{Q}$ is strongly-linearly-separable, P2 would drive φ to zero in a finite time $(0 \leqslant \epsilon < 1)$.

MODIFICATION OF BROWN-VON-NEUMANN PROCEDURE

To solve a symmetric game, Brown and von-Neumann have a continuous procedure of the type of that of Agmon-et-al: the correction is made proportional to $|a_j . g|$. In the Brown von-Neumann approach, however, it is desired to have at all time a vector y satisfying the strategy constraints. This fact contributes to an additional correction term, as will be seen.

The continuous approach of Brown von-Neumann can be modified according to the preceding ideas in order to increase its rate of convergence, as follows:

One seeks an strategy y for Player II satisfying

$$u_i(y) \triangleq a_i . y \leqslant 0 \quad \text{for all } a_i \in A$$

(here a_i is the ith. line of matrix A)

Define

$$\phi(u_i) \triangleq |a_i . y| . H_{iy}$$

with H_{iy} the indicator of $J = \{ i;\ a_i . y > 0,\ a_i \varepsilon A \}$

$$\phi_\epsilon(y) \triangleq \sum_i [\phi(u_i)]^\epsilon = \sum_{i \varepsilon J} |a_i \cdot y|^\epsilon$$

Consider the procedure:

$$\text{P3} \qquad \begin{cases} \dot{y}_j = \phi^\epsilon(u_j) - y_j \phi_\epsilon(y) \\ y(0) \text{ an arbitrary strategy.} \end{cases}$$

For $\epsilon = 1$ this is the Brown von-Neumann procedure.

The motivation for such scheme is that since $\phi(u_j) = a_j . y > 0$ for $j \varepsilon J$ means that Player I is getting a positive return by playing strategy j, Player II should move in the direction of this strategy (the game being symmetric). This what is indicated by the first term in the procedure. The second term is a normalizing term, in order to preserve the constraints imposed by the fact that y is an strategy.

3.6. <u>THEOREM</u>. Procedure P3 converges to an optimal strategy for $\epsilon > 0$. Furthermore, at each instant, y(t) is an strategy.

Proof: a) Convergence.

Define

$$\psi \triangleq \sum_i \phi^{1+\epsilon}(u_i)$$

As in Theorem 3.1,

$$\dot{\psi} = (1+\epsilon) \sum_i \phi^\epsilon(u_i) \dot{\phi}(u_i) = (1+\epsilon) \sum_i \phi^\epsilon(u_i)\, a_i \cdot \dot{y} =$$

$$= (1+\epsilon) \sum_i \sum_j \phi^\epsilon(u_i)\, a_{ij}\, \dot{y}_j = (1+\epsilon) \sum_i \sum_j \phi^\epsilon(u_i)\, a_{ij} [\phi^\epsilon(u_j) - y_j \phi_\epsilon(y)]$$

$$= (1+\epsilon)\sum_i \sum_j \phi^{\epsilon}(u_i)\, a_{ij}\, y_j\, \phi_{1\epsilon}(y)$$

since the game is symetric.

Hence,

$$\dot{\psi} = -(1+\epsilon)\sum_i \phi^{1+\epsilon}(u_i)\,\phi_{\epsilon}(y) = -(1+\epsilon)\,\phi_{\epsilon}(y)\,\psi$$

Using again the Jensen's inequality, results:

$$\phi_{\epsilon}(y) = \sum_i \phi^{\epsilon}(u_i) \geqslant \left[\sum_i \phi^{1+\epsilon}(u_i)\right]^{\frac{\epsilon}{1+\epsilon}} = \psi^{\frac{\epsilon}{1+\epsilon}}$$

Hence,

$$\dot{\psi} \leqslant -(1+\epsilon)\,\psi^{\frac{1+2\epsilon}{1+\epsilon}}$$

By integration, for $\epsilon > 0$,

$$\psi \leqslant \left[\frac{1}{K_1+\epsilon t}\right]^{\frac{1+\epsilon}{\epsilon}}$$

Hence, as $t \to \infty$, $\psi \to 0$, which guarantees that $\dot{y}_j = 0$

For $\epsilon > 0$, using Hölder's inequality, with $\alpha = 1+\epsilon$, $\beta = \frac{1+\epsilon}{\epsilon}$,

$$u_i \leqslant \sum_i \phi(u_i) \leqslant N^{1/\beta}\left[\sum_i \phi^{1+\epsilon}(u_i)\right]^{\frac{1}{1+\epsilon}}$$

or

$$u_i \leqslant N^{1/\beta}\left|\frac{1}{K_1+\epsilon t}\right|^{\frac{1}{\epsilon}} \longrightarrow 0 \quad \text{as} \quad t \to \infty$$

Note that the rate of convergence approaches exponential as $\epsilon \to 0$.

b) $y_i(t) \geqslant 0$ for all t.

Suppose not. Then, there exist t_1 such that

$$y_i(t_1) < 0$$

Let t_o be the largest value of t for which

$$y_i(t) \geqslant 0$$

Hence,

$$y_i(t) < 0 \quad \text{for} \quad t_o < t \leqslant t_1$$

Since

$$\phi^{\epsilon}(u_i) \geqslant 0$$

$$\phi_{\epsilon}(y) \geqslant 0,$$

$$\mathring{y}_i(t) = \phi^{\epsilon}(u_i) - y_i\phi_{\epsilon}(y) \geqslant 0, \text{ for } t_o < t \leqslant t_1.$$

Expanding $y_i(t)$ in series up to the first power (which is legitimate in this case, since the equations admit a solution $y(t)$ having first derivatives)

$$y_i(t) = y_i(t_o) - \mathring{y}_i(\tau)(t - t_o)$$
$$t_o \leqslant \tau \leqslant t_1$$
$$t_o < t \leqslant t_1$$

Hence,

$$y_i(t) > y_i(t_o)$$

which is a contradiction.

c) $\sum_i y_i = 1$ for every t.

Proof: - - - - -

$$\frac{d}{dt}\left[1 - \sum_i y_i\right] = -\sum_i \mathring{y}_i = -\sum_i \phi^{\epsilon}(u_i) + \phi_{\epsilon}(y)\sum y_i =$$

$$= \phi_{\epsilon}(y)\left(1 - \sum_i y_i\right)$$

Hence,

$$1 - \sum_i y_i = C_0 \exp\left[-\int_0^t \phi_\epsilon(y)\, dt \right]$$

but since $y(0)$ is an strategy, $C_0 = 0$ Q.D.E.

Remarks . It was shown that P3 has a minimum rate of convergence faster than that of P2. However, convergence can only be proved for P3 in the case in which $\mathcal{a}$ is linearly-separable.

CHAPTER IV

DISCRETE-TIME PROCEDURES

Procedures P1 and P2 of Chapter III are of great appeal since they exhibit the desirable features of termination, if $\mathcal{A}$ is strongly-linearly-separable, and convergence, if not. However, their practical interest is limited by the fact that they are continuous. For large problems it is impractical to run the procedures in an analog computer. The use of a digital computer type of machine demands that these procedures be discretized. In this Chapter, the discretization of P1 and the proof of its convergence will be presented.

Consider the set $\mathcal{A}$ of vectors a_j, $j = 1,2,\ldots,N$; it will be understood that the set $\mathcal{A}$ spans a ℓ-dimensional space E_ℓ. Form the sequence of vectors $\{g_n\}$ by the following procedure:

$$\text{P4}\quad\left\{\begin{array}{l} g_0 = \text{arbitrary, in } E_\ell \\ g_{n+1} = g_n + \alpha_n \sum\limits_{j\in J_n} |a_j \cdot g_n|^{\epsilon}\, a_j \\ \qquad \text{where } J_n = \{ j;\ a_j \cdot g_n \leq 0,\ a_j \in \mathcal{A}\} \\ \qquad\qquad \{\alpha_n\} \text{ is a sequence of positive} \\ \qquad\qquad \text{scalars, } 0 < \alpha_n \leq A \\ \qquad 0 \leq \epsilon < 1 \end{array}\right.$$

The principal results of this Chapter can be summarized as follows:

(I) If $\sum \alpha_k \to \infty$, $\alpha_k \to 0$, then,

1. g_n as given by P4 will converge to zero if the set $\mathcal{a}$ is not linearly separable.

2. g_n as given by P4 will converge to a non trivial solution of

$$g.a_j \geqslant 0 \quad \text{for all} \quad a_j \in \mathcal{a}$$

if the set $\mathcal{a}$ is linearly separable and $g_0 = \sum_{a_j \in \mathcal{a}} a_j$

3. g_n as given by P4 will converge in a finite number of steps to a non-trivial solution of

$$g.a_j \geqslant 0 \quad \text{for all} \quad a_j \in \mathcal{a}$$

if the set $\mathcal{a}$ is strongly-linearly-separable and $g_0 = \sum_{a_j \in \mathcal{a}} a_j$.

For $\epsilon = 0$, termination will occur at a g satisfying

$$g.a_j > 0 \quad \text{for all} \quad a_j \in \mathcal{a}$$

(II) If $\alpha_k = \alpha$ for all k,

1. For every $\lambda > 0$ there exists $\rho(\lambda)$ and $n_0(\lambda)$ such that, for $\alpha < \rho(\lambda)$,

$$\| g_n \| \leqslant \lambda \text{ for every } n > n_0$$

if the set $\mathcal{a}$ is not linearly separable.

2. If the set $\mathcal{a}$ is linearly separable, then for every $\varphi > 0$ arbitrary it is possible to find $\rho(\varphi)$ and $n_0(\varphi)$ such that,

$$\phi_n = \sum_{j \in J_n} |a_j . g_n|^{1+\epsilon} \leqslant \varphi \quad \text{for} \quad n > n_0(\varphi)$$

provided $\alpha < \rho(\varphi)$.

3. If the set $\mathcal{A}$ is strongly-linearly-separable, there exists a ρ such that if $\alpha < \rho$, procedure P4 will converge in a finite number of steps to a non-trivial solution of

$$g.a_j \geqslant 0 \quad \text{for all} \quad a_j \varepsilon \mathcal{A}$$

For $\epsilon = 0$, termination will occur at a g satisfying

$$g.a_j > 0 \quad \text{for all} \quad a_j \varepsilon \mathcal{A}.$$

(III) In any case,

$$\|g_n\| \leqslant \|g_o\| + \hat{K}(A) \quad \text{for all} \quad n .$$

The constant $\hat{K}(A)$ is such that, as $A \to 0$, $\hat{K}(A) \to 0$.

The role that the parameter ϵ plays in these procedures is roughly the follwing: for $0 \leqslant \epsilon \leqslant 1$ an increase in ϵ implies an increase in the stability of the iterations, which is not always desirable, since concommitant to the stability is a slowness in convergence (or with $\epsilon = 1$ and strong linear separability a loss in finite termination). Thus, in actual practice the choice of ϵ depends on the a priori likelihood of strong linear separability. If strong linear separability is guaranteed, $\epsilon = 0$ is a clear choice. If not, ϵ should be larger than zero and near 1 if strong linear separability is unlikely and small if it is likely.

<u>Boundedness</u>

In P4, the correction term leading from g_n to g_{n+1} uses only vectors having a non positive scalar product with g_n. In the continuous case this makes $\|g(t)\|$ a non increasing function of the time. Here this fact does not hold. Nevertheless, one can prove (Theorem 4.1) that $\|g_n\|$ is bounded from above. The least upper bound on $\|g_n\|$ will be required to prove other results later in this Chapter.

In what follows, the iteration leading from g_n to g_{n+1} will be referred to as the nth. iteration, or step, of the procedure.

Let M and N denote

$$M = \max_{a_j \in \mathcal{A}} \|a_j\|$$

N = total number of vectors in $\mathcal{A}$.

For a given vector v in E_ℓ, ($\|v\| = 1$), let L_v and δ_v be defined by

$$L_v = \{ a_j ; a_j \cdot v = 0,\ a_j \in \mathcal{A} \}$$

$$\delta_v = \min_{a_j \notin L_v} |a_j \cdot v| > 0$$

<u>Lemma 4.1</u>. Let v be a fixed vector in E_ℓ ($\|v\| = 1$) and $\mathcal{A} = \{a_j\}$ a finite set of vectors in E_ℓ. Then, any vector $g \in E_\ell$ satisfying

$$\left\| \frac{g}{\|g\|} - v \right\| < \zeta = \frac{\delta_v}{2M}$$

has the property that $a_j \in \mathcal{A}$, $a_j \notin L_v$ implies

$$\left| \frac{g}{\|g\|} \cdot a_j \right| \geq \frac{\delta_v}{2}$$

and

$$\text{sign } a_j \cdot v = \text{sign } a_j \cdot g$$

Proof:

$$\left| \frac{g}{\|g\|} \cdot a_j \right| = \left| v \cdot a_j + \left(\frac{g}{\|g\|} - v \right) \cdot a_j \right| \geq |v \cdot a_j| - \left\| \frac{g}{\|g\|} - v \right\| \cdot \|a_j\|$$

Hence for $a_j \notin L_v$,

$$\left| \frac{g}{\|g\|} \cdot a_j \right| \geq \delta_v - \zeta M \geq \frac{\delta_v}{2}$$

which proves the first part of the Lemma. Since

$$\left| \frac{g}{\|g\|} \cdot a_j - v \cdot a_j \right| = \left| \left(\frac{g}{\|g\|} - v \right) \cdot a_j \right| \leq \left\| \frac{g}{\|g\|} - v \right\| \cdot \|a_j\| \leq \zeta M \leq \frac{\delta_v}{2}$$

and

$|v \cdot a_j| \geq \delta_v$ for all $a_j \notin L_v$, one has

$$\text{sign } g \cdot a_j = \text{sign } v \cdot a_j \qquad (a_j \notin L_v)$$

<u>THEOREM 4.1.</u>[1] There exists a positive number $\hat{K}$, independent of n, such that, for the sequence $\{\|g_n\|\}$ with g_n given by P4,

$$\|g_n\| \leq \|g_o\| + \hat{K} \triangleq K$$

provided

$$0 < \alpha_n \leq A$$

<u>Proof</u>:

The proof is by mathematical induction over the dimension of the space spanned by the a_j's.

(i) The theorem is first proved for colinear a_j's:

From P4, only those vectors in $\mathcal{a}$ for which $a_j . g_n \leq 0$ appear in the summation. Since the problem is one dimensional, this means that the vectors involved in the nth. iteration are proportional to $-g_n$. Hence, in each iteration $\|g_{n+1}\|$ either decreases (if the modulus of the summation is less than $2\|g_n\|$), or is bounded by the modulus of the summation.

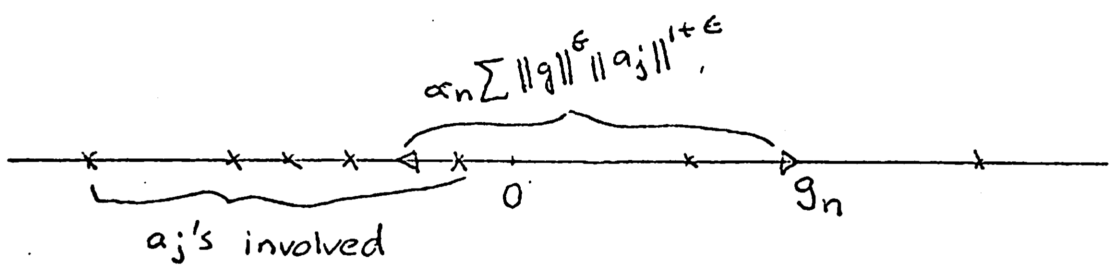


Figure 1

Thus, either

$$\sup_n \|g_n\| = \|g_o\|$$

---

1. A similar theorem (basically corresponding to our case for $\epsilon = 0$) was given in Ref. 10.

or $\sup_n \| g_n \| = \sup_{n+1} \| g_{n+1} \| \leq \sup_n \alpha_n \sum \| a_j \|^{1+\epsilon} \cdot \| g_n \|^{\epsilon} \leq A M^{1+\epsilon} N \sup_n \| g_n \|^{\epsilon}$

$$\sup_n \| g_n \| \leq \left(A M^{1+\epsilon} N \right)^{1/1-\epsilon}$$

In either case,

$$\sup_n \| g_n \| \leq \| g_0 \| + \left(A M^{1+\epsilon} N \right)^{1/1-\epsilon} \tag{4.1}$$

which proves the theorem for dimension 1.

(ii) Suppose now that the a_j's span a space of dimension ℓ and the theorem holds for dimensions up to $\ell - 1$, but fails for dimension ℓ. One will show that this leads to a contradiction.

If the theorem fails for dimension ℓ, then, for every $\hat{K}$, there exists g_0 such that for some n,

$$\| g_n \| > \hat{K} + \| g_0 \| = K .$$

Then, it is possible to construct an increasing sequence $\hat{K}_i \to \infty$ and for every i, to pick g_{0_i} and n_i such that

$$\| g_{n_i} \| \leq K_i$$

$$\| g_{n_i+1} \| > K_i$$

(since $\| g_{0_i} \| \leq \| g_{0_i} \| + \hat{K}_i = K_i$). Without loss of generality, take the indices n_i to be increasing with i. One has, from P4 :

$$\|g_{n_i+1}\|^2 = \|g_{n_i}\|^2 + \alpha_{n_i}^2 \Big\| \sum_{j\in J_{n_i}} |a_j\cdot g_{n_i}|^{\epsilon} a_j \Big\|^2 + 2\alpha_{n_i} \sum_{j\in J_{n_i}} |a_j\cdot g_{n_i}|^{\epsilon} a_j\cdot g_{n_i}$$

and hence,

$$K_i^2 < \|g_{n_i+1}\|^2 \leq \|g_{n_i}\|^2 + \underbrace{\alpha_{n_i}^2 M^{2(1+\epsilon)} K_i^{2\epsilon} N^2}_{(*)} - \underbrace{2\alpha_{n_i} \sum_{j\in J_{n_i}} |a_j\cdot g_{n_i}|^{1+\epsilon}}_{(**)} \tag{4.2}$$

Now, a lower bound on the negative term (**) can be found as follows: the sequence

$$\left\{ \frac{g_{n_i}}{\|g_{n_i}\|} \right\}$$

being on the unit sphere, must have a convergent subsequence, converging to some v ($\|v\| = 1$). Let

$$\delta_v = \min \{ |a_j . v| ;\ a_j . v \neq 0,\ a_j \in \mathcal{A} \}$$

Then, it can be shown (and will be shown in the sequel) that for a subsequence $\{K_i\}$ and the corresponding $\{g_{n_i}\}$ such that

$$K_i \geq \max \left[\left(2AM^{1+\epsilon}N\right)^{1/1-\epsilon},\ \frac{16\,\hat{K}_{\ell-1} M}{\delta_v (1-\xi)} \right]$$

and

$$\left\| \frac{g_{n_i}}{\|g_{n_i}\|} - v \right\| < \xi$$

where ξ is such that

$$\frac{\xi + 8(2\xi - \xi^2)^{1/2}}{1-\xi} \leqslant \frac{\delta_v}{4M}$$

and $\hat{K}_{\ell-1}$ is the bound of the theorem for dimension $\ell - 1$, the inequality

$$|g_{n_i} \cdot a_j| \geqslant K_i \frac{\delta_v}{4} \qquad , \quad a_j \in \mathcal{A} \, , \quad a_j \cdot v \neq 0 \tag{4.3}$$

holds.

Then, using (4.3) in (4.2),

$$0 < \alpha_{n_i}^2 M^{2(1+\epsilon)} N^2 K_i^{2\epsilon} - 2\alpha_{n_i}\left(\frac{\delta_v}{4} K_i\right)^{1+\epsilon} < \alpha_{n_i}^2 \left[M^{2(1+\epsilon)} N^2 K_i^{2\epsilon} - \frac{2}{A}\left(\frac{\delta_v}{4} K_i\right)^{1+\epsilon} \right]$$

Since $\{K_i\}$ is unbounded, a contradiction follows for $0 \leqslant \epsilon < 1$.

To show that, under the asserted conditions,

$$|g_{n_i} a_j| \geqslant K_i \frac{\delta_v}{4}$$

consider the subsequence $\left\{ \frac{g_{n_i}}{\|g_{n_i}\|} \right\}$ converging to some v ($\|v\|=1$). Thus, for an index subset S_v, given $\xi > 0$ arbitrary, one can find $n_0(\xi)$ such that

$$\left\| \frac{g_{n_i}}{\|g_{n_i}\|} - v \right\| < \xi \qquad \text{for } n_i \geqslant n_0, \; n_i \in S_v$$

Call

$$L_v = \{ a_j \; ; \; a_j \cdot v = 0 \, , \; a_j \in \mathcal{A} \}$$

Clearly, L_v is aproper subset of $\mathcal{A}$, since it is contained in a space of at most $\ell - 1$ dimensions. K_i and g_i

are such that

$$K_i \geqslant \max\left[\left(2AM^{1+\epsilon}N\right)^{1/1-\epsilon},\ \frac{16\hat{K}_{\ell-1}M}{\delta_v(1-\xi)}\right]$$

and

$$\left\|\frac{g_{n_i}}{\|g_{n_i}\|}-v\right\| < \xi$$

where ξ is such that

$$\frac{\xi + 8(2\xi-\xi^2)^{1/2}}{1-\xi} \leqslant \frac{\delta_v}{4M}$$

and $\hat{K}_{\ell-1}$ is the bound of the Theorem for dimension $\ell-1$.
Then,

$$J_n \not\subset L_v$$

To prove this claim consider Lemma 4.1; since $\xi < \frac{\delta_v}{2M}$,

$$\text{sign } g_{n_i}\cdot a_j = \text{sign } v\cdot a_j \quad \text{for} \quad a_j \notin L_v \tag{4.4}$$

Hence at the n_ith. iteration, all $a_j \in J_{n_i}$ are such that $v\cdot a_j \leqslant 0$. Suppose that $J_{n_i} \subset L_v$, that is for all $a_j \in J_{n_i}$, $a_j \in L_v$. Let's continue the iteration up to the step λ where for the first time some $a_j \notin L_v$ is involved. This must happens for a finite λ since otherwise the procedure would involve only points in $\ell-1$ dimension and, by the induction hypothesis, $\|g_n\|$ would be bounded, which is a contradiction.

From P4,

$$g_\lambda = g_{n_i} + \sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{j\in J_k} |a_j\cdot g_k|^\epsilon a_j$$

Hence,

$$\frac{g_\lambda}{\|g_\lambda\|} - v = \left(\frac{g_{n_i}}{\|g_{n_i}\|} - v\right)\frac{\|g_{n_i}\|}{\|g_\lambda\|} - v\left(1 - \frac{\|g_{n_i}\|}{\|g_\lambda\|}\right) + \frac{1}{\|g_\lambda\|}\sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{j\varepsilon J_k} |a_j\cdot g_k|^\epsilon a_j$$

$$\left\|\frac{g_\lambda}{\|g_\lambda\|} - v\right\| \leq \left\|\frac{g_{n_i}}{\|g_{n_i}\|} - v\right\| \frac{\|g_{n_i}\|}{\|g_\lambda\|} + \|v\| \frac{\|g_\lambda\| - \|g_{n_i}\|}{\|g_\lambda\|} +$$

$$+ \frac{1}{\|g_\lambda\|}\left\|\sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{j\varepsilon J_k} |a_j\cdot g_k|^\epsilon a_j\right\|$$

On the other hand, since $a_j . v = 0$ for all $a_j \varepsilon J_k$, $n_i \leqslant k < \lambda$,

$$\|g_\lambda\| \geqslant |g \;.v| = |g_{n_i} . v|$$

and

$$\frac{\|g_{n_i}\|}{\|g_\lambda\|} \leqslant \frac{\|g_{n_i}\|}{|g_{n_i} . v|}$$

resulting for

$$\left\|\frac{g_{n_i}}{\|g_{n_i}\|} - v\right\| < \xi \,,$$

$$\left\|\frac{g_\lambda}{\|g_\lambda\|} - v\right\| \leq \xi \frac{\|g_{n_i}\|}{|g_{n_i}\cdot v|} + \frac{1}{|g_{n_i}\cdot v|}\left[\|g_\lambda\| - \|g_{n_i}\| + \left\|\sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{j\varepsilon J_k} |a_j\cdot g_k|^\epsilon a_j\right\|\right)$$

But

$$\|g_\lambda\| \leq \|g_{n_i}\| + \left\| \sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{j \in J_k} |a_j \cdot g_k|^\epsilon a_j \right\|$$

and therefore,

$$\left\| \frac{g_\lambda}{\|g_\lambda\|} - v \right\| \leq \xi \frac{\|g_{n_i}\|}{|g_{n_i} \cdot v|} + \frac{2}{|g_{n_i} \cdot v|} \left\| \sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{j \in J_k} |a_j \cdot g_e|^\epsilon a_j \right\| \qquad (4.5)$$

Since $\epsilon \geqslant 0$, P4 gives

$$\|g_{n_i}\| + \alpha_{n_i} M^{1+\epsilon} N \|g_{n_i}\|^\epsilon \geqslant \|g_{n_i+1}\| > K_i ,$$

$$\|g_{n_i}\| > K_i - \alpha_{n_i} M^{1+\epsilon} N \|g_{n_i}\|^\epsilon > K_i - AM^{1+\epsilon} N K_i^\epsilon$$

and since $K_i > (2AM^{1+\epsilon}N)^{1/1-\epsilon}$, or, $AM^{1+\epsilon}N < K_i^{1-\epsilon}/2$,

$$\|g_{n_i}\| > K_i/2 \qquad (4.6)$$

Since

$$\frac{g_{n_i}}{\|g_{n_i}\|} \cdot v = \left(\frac{g_{n_i}}{\|g_{n_i}\|} - v + v \right) \cdot v = \left(\frac{g_{n_i}}{\|g_{n_i}\|} - v \right) \cdot v + 1 ,$$

$$g_{n_i} \cdot v \geqslant \|g_{n_i}\| - \left\| \frac{g_{n_i}}{\|g_{n_i}\|} - v \right\| \cdot \|g_{n_i}\| \geqslant \|g_{n_i}\|(1-\xi) \qquad (4.7)$$

Hence

$$|g_{n_i} \cdot v| \geq \frac{K_i}{2}(1-\xi) \tag{4.8}$$

Also, writting g''_λ and g''_{n_i} for the projections of g_λ and g_{n_i} in the hyperplane perpendicular to v (that is, $g'' = g - (g.v)v$) and recalling that $\hat{K}_{\ell-1}$ is the bound for dimension $\ell - 1$,

$$\|g''_\lambda\| = \left\| g''_{n_i} + \sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{j\in J_k} |a_j \cdot g_k|^\epsilon a_j \right\| \leq \|g''_{n_i}\| + \hat{K}_{\ell-1}$$

or,

$$\left\| \sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{j\in J_k} |a_j \cdot g_k|^\epsilon a_j \right\| \leq 2\|g''_{n_i}\| + \hat{K}_{\ell-1} \tag{4.9}$$

and from

$$\|g''_{n_i}\|^2 = \|g_{n_i}\|^2 - |g_{n_i} \cdot v|^2 \leq \|g_{n_i}\|^2 (1-(1-\xi)^2) \leq K_i^2 (2\xi - \xi^2) \tag{4.10}$$

results, combining eqs. 4.5, 4.6, 4.7, 4.8, 4.9, 4.10

$$\left\| \frac{g_\lambda}{\|g_\lambda\|} - v \right\| \leq \frac{\xi}{1-\xi} + \frac{2\left[\hat{K}_{\ell-1} + 2K_i(2\xi-\xi^2)^{1/2}\right]}{(1-\xi)\; K_i/2} =$$

$$= \frac{\xi + 8(2\xi - \xi^2)^{1/2}}{1-\xi} + \frac{4\hat{K}_{\ell-1}}{(1-\xi)K_i}$$

But since

$$\frac{\xi + 8(2\xi - \xi^2)^{1/2}}{1-\xi} \leqslant \frac{\delta_v}{4M}$$

and

$$\frac{4\,\hat{K}_{\ell-1}}{(1-\xi)\,K_i} \leqslant \frac{\delta_v}{4M}$$

it follows

$$\left\| \frac{g_\lambda}{\|g_\lambda\|} - v \right\| \leqslant \frac{\delta_v}{2M}$$

Hence by Lemma 4.1,

$$\text{sign } g_\lambda \cdot a_j = \text{sign } v \cdot a_j \qquad \text{for all} \quad a_j \notin L_v .$$

By Eq. 4.4,

$$\text{sign } g_{n_i} \cdot a_j = \text{sign } v \cdot a_j = \text{sign } g_\lambda \cdot a_j$$

for all $a_j \notin L_v$. Hence,

$$g_\lambda \cdot a_j < 0 \qquad \text{for some} \quad a_j \notin L_v$$

implies

$$g_{n_i} \cdot a_j < 0 \qquad \text{for the same} \quad a_j .$$

This rules out the possibility that $J_{n_i} \subset L_v$. So, one must use some $a_j \notin L_v$ at iteration n_i. Further, by Lemma 4.1,

$$\left| \frac{g_{n_i}}{\|g_{n_i}\|} \cdot a_j \right| > \frac{\delta_v}{2}$$

which together with Eq. 4.6 gives

$$\left|g_{n_i} \cdot a_j\right| > \frac{\delta_v}{2} \|g_{n_i}\| \geqslant \frac{\delta_v}{4} K_i$$

This concludes the proof of the assertion.

$$\text{Let} \quad \hat{K}(A) = \sup \|g_n\| - \|g_o\| \leq \hat{K} < \infty$$

where the supremum is over all steps n of all possible g_n given by P4 by changing g_o and the sequence $\{\alpha_k\}$, but keeping $\alpha_k < A$.

Hence, for every $k < \hat{K}(A)$, there exist g_o and $\{\alpha_k\}$, $\alpha_k < A$, leading to

$$\|g_n\| > \|g_o\| + k \quad \text{for some } n .$$

One will proceed to show that as a corollary of Theorem 1, $\hat{K}(A) \to 0$ with A. This result is important in establishing Theorem 2 in the sequel.

Corollary 4.1. $\hat{K}(A) \to 0$ as $A \to 0$. That is, for every $k > 0$, there exists $\alpha(k) > 0$ such that for every $A < \alpha$, $\hat{K}(A) < k$.

Proof:

The proof will be done by mathematical induction over the dimension of the space spanned by $\mathcal{a}$.

(i) The corollary is certainly true for dimension 1, since one has, by Eq. 4.1,

$$\sup \|g_n\| - \|g_o\| \leqslant (AM^{1+\epsilon} N)^{1/1-\epsilon}$$

(ii) Suppose Corollary 4.1 is true for dimensions up to $\ell - 1$. Negate it for dimension ℓ :

Then, there exists $k > 0$ such that for every $\alpha > 0$ there exists $A < \alpha$ such that $\hat{K}(A) > k$.

Hence, one can construct a decreasing sequence $A_i \to 0$ and have $\hat{K}(A_i) > k$. It follows that there will be a g_0 and $\{\alpha_n\}$, $\alpha_n < A$ leading to

$$\|g_n\| > \|g_0\| + k \quad \text{for some } n.$$

One can follow now the same arguments used to prove Theorem 1, substituting k for K_i, $A_i \to 0$ for $K_i \to \infty$: pick $\{A_i\}$ such that

$$\max\left[(2A_i M^{1+\epsilon} N)^{1/1-\epsilon} , \frac{16\, \hat{K}_{\ell-1}(A_i)}{\delta_v (1-\xi)} \right] \leq k$$

where ξ is the same as in Theorem 1 (recall that $\hat{K}_{\ell-1}(A_i) \to 0$ as $A_i \to 0$ by the induction hypothesis).

As a result, one has the bound on (**) of Eq. 4.2. It follows that:

$$0 < M^{2(1+\epsilon)} N^2 k^{2\epsilon} - \frac{2}{A_i}\left(\frac{\delta_v}{4}\right)^{1+\epsilon} k^{1+\epsilon}$$

which leads to a contradiction as $A_i \to 0$.

<u>Convergence</u>

The principal results of this Chapter are embodied in Theorem 4.2 below. The substance of this Theorem is that with a suitable choice of $\{\alpha_n\}$ the iteration procedure given by P4 can be made to converge in general and terminate under strong linear separability. Thus, this family of procedures combine the important termination property of the perceptron algorithm with the stability (viz., convergence under all circumstances) more commonly associated with iterative procedures.

<u>THEOREM 4.2</u>. Suppose that in the iterative procedure P4

$$\alpha_n \xrightarrow[n \to \infty]{} 0$$

$$\sum_{n=1}^{\infty} \alpha_n = \infty$$

Then, (a) $\mathcal{A}$ not linearly separable implies $\|g_n\| \xrightarrow[n\to\infty]{} 0$

(b) $\mathcal{A}$ linearly separable and $g_o = \sum_{a_i \varepsilon \mathcal{A}} a_i$ [1] implies

$g_n \longrightarrow g$, $g.a_j \geqslant 0$ for any $a_j \varepsilon \mathcal{A}$

(c) $\mathcal{A}$ strongly linearly separable and $g_o = \sum_{a_i \varepsilon \mathcal{A}} a_i$ [1] implies that there exists n_o such that

$$g_n = g \quad \text{for} \quad n \geqslant n_o$$

and

$$g.a_j \geqslant 0 \quad \text{for every} \quad a_j \varepsilon \mathcal{A}$$

---

1 Any value of g_o such that $g_o.g > 0$ would be suitable. The proposed one was presented just to be explicit. This condition is necessary to avoid convergence to a trivial solution of $g.a_j \geqslant 0$.

The proof for Theorem 4.2 is quite long. One will first give an outline of the major steps in the argument, and supply the details later. One begins by considering for procedure P4 two mutually exclusive possibilities[1]:

(i) For every $\eta > 0$, there exists a g_n in the sequence $\{g_n\}$ such that

$$|g_n \cdot a_j| < \eta \|g_n\| \quad \text{for all } j \in J_n .$$

(ii) There exists $\eta > 0$ such that for every g_n,

$$|g_n \cdot a_j| \geqslant \eta \|g_n\| \quad \text{for some } j \in J_n .$$

Now one can show (Lemma 4.2) that (i) implies linear separability. Therefore, if the set $\mathcal{A}$ is not linearly separable, (ii) is the only possible case. Using this and Theorem 4.1, part (a) of this Theorem 4.2 is proved. Part (b) is proved by demonstrating that linear separability and $g_o = \sum_{a_j \in \mathcal{A}} a_j$ imply that (i) is the only possibility. Therefore, Lemma 4.2 below shows convergence for this case. Part (c) is proved by induction over the dimension of the space spanned by $\mathcal{A}$. First, it is easy to show that it holds for the one-dimensional case. Next, (strong) linear-separability and $g_o = \sum a_j$ excludes possibility (ii). Based on possibility (i). Lemma 4.2 shows that after a finite number of steps the iterations must

---

[1] One need not consider the case of J_n being empty, in which case termination occurs and the results to follow would be trivially true.

be constrained to a lower dimension. The conclusion follows by induction.

LEMMA 4.2. Situation (i) above implies the existance of a vector $v^* \in E_\ell$ satisfying

$$v^*.a_j \geqslant 0 \quad \text{for all } a_j \in \mathcal{A} .$$

Further, $g_n / \| g_n \| \to v^*$

Proof:

Construct a decreasing sequence $\{\eta_k\}$ with $\eta_k \to 0$ ($\eta_k > 0$) and, correspondingly, a sequence $\{g_{n_k}\}$ such that, according to (i) ,

$$| g_{n_k} . a_j | < \eta_k \| g_{n_k} \| \quad \text{for all } a_j \in J_{n_k}$$

Without loss of generality, suppose that the indices of the sequence $\{g_{n_k}\}$ are increasing with k. Take a subsequence $\{\{\eta_k\}\}$ and the corresponding subsequence $\{\{g_{n_k}\}\}$ such that $g_{n_k} / \| g_{n_k} \|$ converges to some $v \in E_\ell$ ($\|v\| = 1$). Call S_v the index set of this subsequence. Hence, for any $\xi > 0$, there exists $n_o(\xi)$ such that

$$\left\| \frac{g_{n_k}}{\| g_{n_k} \|} - v \right\| < \xi \qquad \text{for } n_k \geqslant n_o(\xi),\ n_k \in S_v .$$

To prove the existence of v^*, select

$$\xi < \frac{\delta_v}{2M}$$

and consider the subsequence $\{\{\eta_k\}\}$ starting from a k such that $n_k > n_o(\xi)$. Pick the first η_{k_1} satisfying

$$\eta_{k_1} < \xi M$$

Call $n_{k_1} = n(\xi)$. Consider

$$|v.a_j| = \left|\left(v - \frac{g_{n_i}}{\|g_{n_i}\|} + \frac{g_{n_i}}{\|g_{n_i}\|}\right)\cdot a_j\right| < \left|\left(v - \frac{g_{n_i}}{\|g_{n_i}\|}\right)\cdot a_j\right| + \frac{g_{n_i}\cdot a_j}{\|g_{n_i}\|}$$

Then, for any $n_i > n(\xi)$, $n_i \in S_v$,

$$|v.a_j| < \xi M + \eta_i < 2\xi M \quad \text{for all} \quad j \in J_{n_i}$$

and since

$$\xi < \frac{\delta_v}{2M} ,$$

$$|v.a_j| < \delta_v$$

But since, by definition of δ_v,

$$|v.a_j| > \delta_v \quad \text{for } a_j.v \neq 0 , \quad a_j \in \mathcal{A} , \ (\text{or, } a_j \notin L_v)$$

one must have

$$v.a_j = 0 \quad \text{for all} \quad j \in J_{n_i} , \quad n_i > n(\xi) , \quad n_i \in S_v ,$$

Hence, for some iteration $n_i > n(\xi)$, <u>only points $a_j \in L_v$-</u> <u>will be involved</u>. Further, according to Lemma 4.1,

$$\text{sign } v.a_j = \text{sign } g_n.a_j \quad \text{for all} \quad a_j \notin L_v .$$

Hence, there exists no $a_j \in \mathcal{A}$ such that $v.a_j < 0$. The first part of Lemma 4.2 is, then, proved by putting $v^* = v$.

To show the second part, for an arbitrary $\zeta_0 > 0$, let

$$\zeta(\zeta_0) = \min\left(\zeta_0, \frac{\delta_v}{2M}\right)$$

Select $\xi(\zeta_0)$ satisfying

$$\frac{\xi + 4(2\xi - \xi^2)^{1/2}}{1-\xi} \leq \frac{1}{2}\zeta$$

Consider m_o such that

$$g_{m_o} \cdot v > 0$$

This is possible since v is a cluster point of the sequence $g_n / \| g_n \|$. According to Corollary 4.1, select ρ such that

$$\hat{K}_{\ell-1}(\rho) \leqslant \frac{\zeta}{2} g_{m_o} \cdot v$$

Since $\alpha_n \to 0$, select $\hat{n}_o(\rho)$ such that

$$\alpha_n < \rho \qquad \text{for all } n > \hat{n}_o$$

select

$$\hat{\hat{n}}_o = \max \left[n_o(\zeta) , \hat{n}_o(\rho) , m_o \right]$$

By the first part of the Lemma, for $n_i > \hat{\hat{n}}_o$, $n_i \in S_v$, only points in L_v will be involved in the n_ith. iteration. We will show that all remaining iterations after the n_ith will make use of only points in L_v . One proceeds by induction. Suppose that all iterations from the n_ith to the λ th involve only $a_j \in L_v$. Lets examine the $(\lambda + 1)$th iteration: from P4,

$$\frac{g_\lambda}{\| g_\lambda \|} - v = \left(\frac{g_{n_i}}{\| g_{n_i} \|} - v \right) \frac{\| g_{n_i} \|}{\| g_\lambda \|} + \frac{1}{\| g_\lambda \|} \sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{\substack{j \in J_k \\ (a_j \in L_v)}} | a_j \cdot g_k |^{\epsilon} a_j + v \left(\frac{\| g_n \|}{\| g_\lambda \|} - 1 \right)$$

Hence, since $\left\| \frac{g_{n_i}}{\|g_{n_i}\|} - v \right\| < \xi$, for $n_i > n_o(\xi)$, $n_i \in S_v$

$$\left\| \frac{g_\lambda}{\|g_\lambda\|} - v \right\| \leqslant \frac{\xi}{1-\xi} + \frac{2}{|g_{n_i} \cdot v|} \left\| \sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{\substack{j \in J_k \\ (a_j \in L_v)}} |a_j \cdot g_k|^\epsilon a_j \right\| \tag{4.11}$$

From Theorem 4.1,

$$\hat{K}_{\ell-1}(\rho) + \|g''_{n_i}\| \geqslant \|g''_{\lambda+1}\| = \left\| g''_{n_i} + \sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{\substack{j \in J_k \\ (a_j \in L_v)}} |a_j \cdot g_k|^\epsilon a_j \right\| \geqslant$$

$$\geqslant \left\| \sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{\substack{j \in J_k \\ (a_j \in L_v)}} |a_j \cdot g_k|^\epsilon a_j \right\| - \|g''_{n_i}\|$$

It follows that, for $n_i > \hat{n}_o(\rho)$

$$\left\| \sum_{k=n_i}^{\lambda-1} \alpha_k \sum_{\substack{j \in J_k \\ (a_j \in L_v)}} |a_j \cdot g_k|^\epsilon a_j \right\| \leqslant 2\|g''_{n_i}\| + \hat{K}_{\ell-1}(\rho) \tag{4.12}$$

On the other hand, for $g_{n_i}/\|g_{n_i}\|$ in a ξ neighborhood of v (see proof of Theorem 4.1,eqs. 4.7, 4.10)

$$\frac{\|g''_{n_i}\|}{|g_{n_i} \cdot v|} \leqslant \frac{(2\xi - \xi^2)^{1/2}}{1-\xi} \tag{4.13}$$

From P4, one can also write, taking as starting point g_{m_o}:

$$|g_{n_i} \cdot v| = \left| g_{m_o} \cdot v + \sum_{k=m_o}^{n_i-1} \alpha_k \sum_{j \in J_k} |a_j \cdot g_k|^\epsilon a_j \cdot v \right|$$

and since $a_j \cdot v \geqslant 0$ for all $a_j \in \mathcal{A}$, and $g_{m_o} \cdot v > 0$,

$$|g_{n_i} \cdot v| > |g_{m_o} \cdot v| = g_{m_o} \cdot v \tag{4.14}$$

Combining eqs. 4.11, 12, 13 and 14, for $n_i > \hat{n}_o$, $n_i \in S_v$,

$$\left\| \frac{g_{\lambda+1}}{\|g_{\lambda+1}\|} - v \right\| \leq \frac{\xi + 4(2\xi - \xi^2)^{1/2}}{1-\xi} + \frac{\hat{K}_{\ell-1}(\rho)}{g_{m_o} \cdot v} \leq \frac{\zeta}{2} + \frac{\zeta}{2} = \zeta \leq \frac{\delta_v}{2M}$$

Hence, according to Lemma 4.1,

$$\text{sign } g_{\lambda+1} \cdot a_j = \text{sign } v \cdot a_j \qquad \text{for } a_j \notin L_v$$

But by the first part of the present Lemma, v is such that

$$v \cdot a_j \geqslant 0 \qquad \text{for all } a_j \in \mathcal{A} ,$$

Hence, the vectors to be used in the $(\lambda+1)$th iteration (that is, those such that $g_{\lambda+1} \cdot a_j \leq 0$) can only belong to L_v. Then, by induction, all iterations following the n_ith ($n_i \geqslant \hat{n}_o$, $n_i \in S_v$) will make use only of vectors in L_v.

As a result, the relation

$$\left\| \frac{g_{\lambda+1}}{\|g_{\lambda+1}\|} - v \right\| \leq \zeta \leq \zeta_o$$

holds for all $\lambda \geqslant \hat{n}_o$, which proves the Lemma.

<u>Proof of Theorem 4.2</u>

<u>Proof of part (a)</u>: $\mathcal{A}$ not linearly separable.

According to Lemma 4.2, situation (i) must be ruled out. Consider, then, situation (ii): there exists $\eta > 0$ such that for every g_n,

$$|g_n \cdot a_j| > \eta \|g_n\| \qquad \text{for some } j \in J_n .$$

(J_n cannot be empty since $\mathcal{A}$ is not linearly separable).

Let

$$\phi_n = \sum_{j \in J_n} |a_j \cdot g_n|^{1+\epsilon}$$

From P4 and Theorem 4.1,

$$\|g_{n+1}\|^2 \leq \|g_n\|^2 + \alpha_n^2 M^{2(1+\epsilon)} N^2 K^{2\epsilon} - 2\alpha_n \phi_n$$

But

$$\phi_n = \sum_{j \in J_n} |a_j \cdot g_n|^{1+\epsilon} > \eta^{1+\epsilon} \|g_n\|^{1+\epsilon}$$

Therefore, putting $C = M^{2(1+\epsilon)} N^2 K^{2\epsilon}$,

$$\|g_{n+1}\|^2 \leq \|g_n\|^2 + \alpha_n^2 C - 2\alpha_n \eta^{1+\epsilon} \|g_n\|^{1+\epsilon} =$$

$$= \|g_n\|^2 - \alpha_n \left[2\eta^{1+\epsilon} \|g_n\|^{1+\epsilon} - \alpha_n C \right] \tag{4.15}$$

To prove part (a) of the Theorem, one must show that for every $\lambda > 0$ there exists $n_o(\lambda)$ such that

$$\|g_n\|^2 \leq \lambda \qquad \text{for all } n > n_o(\lambda).$$

Since $\alpha_n \to 0$, there exists $n_1(\lambda)$ such that

$$\alpha_n < \frac{\eta^{1+\epsilon} \left(\frac{\lambda}{2}\right)^{(1+\epsilon)/2}}{C}$$

and

$$\alpha_n^2 \leq \frac{\lambda}{2C} \qquad \text{for } n > n_1(\lambda)$$

From eq. 4.15, if all $\|g_k\|^2 > \lambda/2$ for $n > k > n_1(\lambda)$,

$$\|g_{n+1}\|^2 \leq \|g_n\|^2 - \alpha_n \eta^{1+\epsilon}\left(\frac{\lambda}{2}\right)^{1+\epsilon} = \|g_{n_1(\lambda)}\|^2 - \eta^{1+\epsilon}\left(\frac{\lambda}{2}\right)^{1+\epsilon}\sum_{n_1(\lambda)}^{n}\alpha_k$$

Since $\sum \alpha_k$ is divergent, there must exist $n_2(\lambda)$ such that the above inequality cannot hold if $\|g_k\|^2 > \lambda/2$ for $n > k > n_1(\lambda)$ and $n > n_2(\lambda)$. Hence,

$$\|g_n\|^2 \leq \frac{\lambda}{2} \qquad \text{for some } n,\ n_2(\lambda) \geq n \geq n_1(\lambda)$$

Let $n_0(\lambda)$ be the first time this occurs. Hence,

$$\|g_{n_0(\lambda)}\|^2 \leq \frac{\lambda}{2}$$
$$\alpha^2_{n_0(\lambda)} \leq \frac{\lambda}{2C}$$
$$\alpha_{n_0(\lambda)} \leq \eta^{1+\epsilon}\left(\frac{\lambda}{2}\right)^{1+\epsilon}\frac{1}{C} \qquad n_2(\lambda) \geq n_0(\lambda) \geq n_1(\lambda)$$

simultaneously. Equation 4.15 shows also that whenever $\|g_n\|^2 > \lambda/2$ for $n > n_0$,

$$\|g_{n+1}\|^2 < \|g_n\|^2 \qquad \text{(a decreasing function of n)}$$

It follows that for $n > n_0$, $\|g_n\|^2 < \lambda$, since:

1st: $\|g_{n_0}\|^2 < \dfrac{\lambda}{2}$

2nd: whenever, for $n > n_0$ and $\|g_n\|^2 < \dfrac{\lambda}{2}$,

$$\|g_{n+1}\|^2 \leq \|g_n\|^2 + \alpha_n^2 C \leq \frac{\lambda}{2} + \frac{\lambda}{2} = \lambda$$

3rd: whenever $\dfrac{\lambda}{2} < \|g_n\|^2 \leq \lambda$, $n > n_0$,

$\|g_n\|$ is decreasing, that is,

$$\|g_{n+1}\|^2 < \|g_n\|^2$$

This concludes the proof of part (a) of Theorem 4.2.

Proof of part (b) of Theorem 4.2

Part (b) is a consequence of part (a) and Lemma 4.2. Let v $(\|v\|=1)$ be such that

$$a_j \cdot v \geqslant 0 \qquad \text{for all} \qquad a_j \in \mathcal{A} ,$$

By selecting

$$g_0 = \sum_{a_j \in \mathcal{A}} a_j$$

one must have, from P4:

$$g_{n+1} \cdot v = g_0 \cdot v + \sum_{k=0}^{m} \alpha_k \sum_{j \in J_k} |a_j \cdot g_k|^{\epsilon} a_j \cdot v \geqslant g_0 \cdot v \geqslant \theta > 0$$

(the last two inequalities hold since there must exist at least one $a_j \notin L_v$, for which $a_j \cdot v \geqslant \theta > 0$). Hence by Schwarz's inequality,

$$\|g_{n+1}\| \geqslant \frac{\theta}{\|v\|} = \theta$$

Thus, g_n cannot converge to zero and therefore (ii) is ruled out (part (a) of the Theorem). But by Lemma 4.2, (i) implies that

$$\frac{g_n}{\|g_n\|} \longrightarrow v$$

Proof of part (c) of Theorem 4.2

This part is proved by induction over the dimension of the space spanned by the set $\mathcal{A}$:

1) If all vectors are colinear with a vector a ($\|a\|=1$), write:

$$g_n = \gamma_n a$$

$$a_j = \lambda_j a$$

Since there exists $y = \beta a$ such that

$$a_j \cdot y = \lambda_j \beta > \hat{\theta} > 0 ,$$

without loss of generality, take $\beta < 0$ and hence, $\lambda_j < 0$ for all j. Thus, $j \in J_n$ gives

$$g_n \cdot a_j = \gamma_n \lambda_j < 0$$

which implies

$$\gamma_n > 0$$

Procedure $P4$, then, is reduced to the following scalar procedure:

$$\gamma_{n+1} = \gamma_n - \alpha_n \gamma_n^{\epsilon} \sum |\lambda_j|^{1+\epsilon} = \gamma_n - \hat{\alpha}_n \gamma_n^{\epsilon} \quad \text{if } \gamma_n > 0$$

$$= \gamma_n \qquad \text{if } \gamma_n < 0$$

(since J_n is empty)

Part (c) asserts that this procedure terminates, that is, there exists n_o such that for $n > n_o$,

$$\gamma_n < 0 .$$

Negate it: for all n, $\gamma_n > 0$. By concavity of the function $x^{1-\epsilon}$ $(0 \leq \epsilon < 1)$,

$$(x + h)^{1-\epsilon} \leq x^{1-\epsilon} + (1-\epsilon)x^{-\epsilon} h \quad (x > 0,\ x + h > 0)$$

Hence, one has

$$0 < \gamma_{n+1}^{1-\epsilon} = (\gamma_n - \hat{\alpha}_n \gamma_n^{\epsilon})^{1-\epsilon} \leq \gamma_n^{1-\epsilon} - (1-\epsilon)\hat{\alpha}_n$$

Thus, starting with γ_o arbitrary, there results

$$0 \leq \gamma_{n+1}^{1-\epsilon} \leq \gamma_o^{1-\epsilon} - (1-\epsilon)\sum_0^n \hat{\alpha}_n$$

and since $\sum \hat{\alpha}_n \rightarrow \infty$, for some n, a contradiction is evident. This proves part (c) of Theorem 4.2 for dimension 1.

2) By selecting

$$g_o = \sum_{a_i \varepsilon \mathcal{A}} a_j$$

as in part (b), one has, due to the (strong) linear separability of $\mathcal{A}$:

$$\| g_n \| > \theta \qquad \text{for all n.}$$

Therefore (ii) is ruled out. But by lemma 4.2 (see its proof), (i) implies that after a finite number of steps the iterations only will make use of vectors in L_v, reducing the problem to dimension $\ell-1$. Hence, by the induction hypothesis, termination occurs.

A more direct discretization of procedure P1 of Chapter III would suggest the choice of $\alpha_n = \alpha$, a constant for all steps of the iteration. Unfortunatly, convergence does not occur in general, as can be shown by conterexamples. Nevertheless, one can obtain a theorem (Theorem 4.3) similar to that of Uzawa-Arrow-Hurvicz (Ref. 12) for the discrete gradient procedure in concave programming. Theorem 4.3 not only indicates some bounds in the step size α in order to drive ϕ_n within a prescribed error φ from zero, but also indicates the role of the parameter ϵ.

<u>THEOREM 4.3</u>. Given $\varphi > 0$ arbitrary, there exists $\rho(\varphi) > 0$ and $n_0(\varphi)$ such that if in procedure P4, $\alpha_n = \alpha < \rho(\varphi)$ for all n, then, $\phi_n < \varphi$ for all $n \geqslant n_0(\varphi)$.

<u>Proof</u>. It is enough to compare the hypothesis of the theorem with the ones in the previous results. Before, it was required that $\sum \alpha_n \rightarrow \infty$ and $\alpha_n \rightarrow 0$. The fact that $\alpha_n \rightarrow 0$ was used to ensure that for an arbitrary ρ, $\alpha_n < \rho$ for some $n > n_0(\rho)$. By fixing ρ as $\rho(\varphi)$ and picking $\alpha_n = \alpha$, constant, satisfying $\alpha < \rho(\varphi)$, the requirement of $\alpha_n \rightarrow 0$ can be substituted by the present requirement on α (as long as φ and $\rho(\varphi)$ are kept fixed). Obviously, $\sum \alpha \rightarrow \infty$ $(\alpha > 0)$. The relations between ρ and φ can be understood from results of theorem 4.2 :

Let

$$\zeta_o = \left(\frac{\varphi}{N}\right)^{1/1\ +\epsilon} \frac{1}{MK},$$

$$\zeta_1 = \begin{cases} \frac{\delta_v}{2M} & \text{if } \mathcal{A} \text{ is linearly separable by v.} \\ 1 & \text{if } \mathcal{A} \text{ is not linearly separable.} \end{cases}$$

$$\zeta(\zeta_o) = \min(\zeta_o, \zeta_1)$$

Choose ξ such that

$$\frac{\xi + 4\,(2\xi - \xi^2)^{\frac{1}{2}}}{1 - \xi} \leq \frac{1}{2}\,\zeta(\zeta_o)$$

Consider $m_o \in S_v$ such that

$g_{m_o} \cdot v > 0$ if $\mathcal{A}$ is linearly separable by v

Put

$m_o = 1$ if $\mathcal{A}$ is not linearly separable.

Call

$$\lambda = \frac{1}{M^2}\left(\frac{\varphi}{N}\right)^{2/1+\epsilon}$$

Choose $\rho = \min(\rho_1, \rho_2, \rho_3)$, where

$$\begin{cases} \rho_1 = \text{arbitrary if } \mathcal{A} \text{ is not linearly separable} \\ \rho_1 \text{ is such that } \hat{K}_{\ell-1}(\rho_1) < \frac{1}{2}\zeta(\zeta_o)\, g_{m_o} \cdot v \\ \qquad \text{if } \mathcal{A} \text{ is linearly separable by v} \end{cases}$$

$$\begin{cases} \rho_2^2 = \lambda \frac{1}{2M^{2+\epsilon}N^2K^{2\epsilon}} \quad \text{if } \mathcal{A} \text{ is not linearly separable} \\ \rho_2^2 = \text{arbitrary if } \mathcal{A} \text{ is linearly separable} \end{cases}$$

$$\begin{cases} \rho_3 \leqslant \eta^{1+\epsilon}\left(\frac{\lambda}{2}\right)^{1+\epsilon}\frac{1}{C} = \left(\frac{\eta}{2M}\right)^{1+\epsilon}\left(\frac{\varphi}{N}\right) \\ \qquad \text{if } \mathcal{a} \text{ is not linearly separable} \\ \rho_3 = \text{arbitrary if } \mathcal{a} \text{ is linearly separable} \end{cases}$$

Choose $n_o(\zeta)$ such that

$$\left\| \frac{g_n}{\|g_n\|} - v \right\| < \zeta \qquad \text{for } n > n_o(\zeta), \quad n \in S_v$$

if $\mathcal{a}$ is linearly separable by v.

Put $n_o(\zeta) = 1$ otherwise.

Put $\quad n > \max\left[n_o(\zeta), m_o\right] = \hat{\hat{n}}_o$

It follows:

(a) If $\mathcal{a}$ is not linearly separable, situation (ii) occurs: from (a) of Theorem 4.2, $\|g_n\|^2$ will finally be driven bellow

$$\lambda = \frac{1}{M^2}\left(\frac{\varphi}{N}\right)^{2/1+\epsilon}$$

and consequentely,

$$\phi_n = \sum_{j \in J_n} |a_j \cdot g_n|^{1+\epsilon} \leq N \|g_n\|^{1+\epsilon} M^{1+\epsilon} \leqslant \varphi$$

(b) If $\mathcal{a}$ is linearly separable (i) occurs[1] and for all $n > \hat{\hat{n}}_o$, only points $a_j \in L_v$ are involved and

$$\left\| \frac{g_n}{\|g_n\|} - v \right\| < \zeta_o$$

---

1 (i) occurs as long as $g_o = \sum_{\mathcal{a}} a_j$. Otherwise (ii) may occur and $\|g_n\|^2 < \lambda$ for $n > \hat{n}_o$

Hence,

$$\left|\frac{g_n \cdot a_j}{\|g_n\|}\right| = \left|\left(\frac{g_n}{\|g_n\|} - v\right)\cdot a_j\right| \leq \left\|\frac{g_n}{\|g_n\|} - v\right\|\cdot M \leq \zeta_c M$$

Then,

$$\phi_n = \sum_{j \varepsilon J_n} |a_j \cdot g_n|^{1+\epsilon} \leq (\zeta_o M K)^{1+\epsilon} N = \varphi$$

(c) Finally, if the set $\mathcal{Q}$ is strongly linearly separable, termination would be proved as before. In this case, if termination occurs at step N, $\phi_N = 0$.

Discussion

One should note first that the case of $\epsilon = 1$ is excluded from the proofs presented here because of a question of technicality. As was pointed out before (Chapter II and III), this case is closely related to a procedure presented by Agmon-Motzkin-Schoenberg (Ref. 3, 4). As in their procedure, here one should expect, for the case of linearly separable sets and constant step size $\alpha_n = \alpha$, an exponential convergence of ϕ_n to 0 for $\epsilon = 1$ as long as α satisfies certain bounds. Even though the proofs presented here are not suitable for $\epsilon = 1$, most of the results hold if one is willing to restrict the value of $A = \sup_n \alpha_n$ to a small value. In fact, excluding termination in the strongly linearly separable case, all results hold for $\epsilon = 1$ if A is such that

$$A < \frac{2}{M^2N^2}\left(\frac{\delta_Y}{4M}\right)^2 \leq \left(\frac{1}{4MN}\right)^2 2 \sin \frac{\theta_o}{2}$$

where θ_o is the minimum angle between two non-colinear vectors in $\mathcal{A}$. This bound was obtained to ensure the contradiction in Theorem 4.1 for $\epsilon = 1$. In this case even exponential convergence of ϕ_n to zero can be shown for fixed step size α.

For $\epsilon = o$, P4 corresponds to the Perception Algorithm. The fact that this algorithm can be turned into a convergent procedure is a very interesting result, even though this was not originally the object of this work. The conditions which guarantee convergence are weak, namely,

$$\sum \alpha_n \to \infty, \; \alpha_n \to 0.$$

CHAPTER V

CONCLUSION

COMPUTER RESULTS

A computer program was written in order to verify the results of this work. Three sets of points in E_4 were used; one set of 20 points not linearly-separable, one set of 20 points linearly-separable but not strongly-linearly separable and one set of 20 points strongly-linearly-separable. For each of these sets, 2 types of iterations were programmed: one with $\alpha_n = 2/n$ and the other with $\alpha_n = \alpha$, a constant. Tests were done with $\alpha = 2.0, 1.0, 0.5, 0.25, 0.125$. In each case, the parameter ϵ took values 0.0, 0.25, 0.5, 0.75, and 1.0, and 34 iterations of each were done.

For $\alpha = 2$ and $\epsilon = 1$, a rapid increase in ϕ_n and $\| g_n \|$ were observed for the non-linearly-separable case and linearly (but not strongly-linearly-) separable case. In all other cases ϕ_n either converged to zero or was bounded from above. Termination was observed in most of the cases where the strongly-linearly-separable set was used. In some cases, termination occurred in as few as 5 iterations and in two cases ($\alpha = .5$, $\epsilon = .25$, and $\alpha_n = 2/n$, $\epsilon = .25$) in only 3 iterations. The cases where termination was not observed before the 34th iteration for strongly-linearly-separable sets were those using small α's ($\alpha = .125$ or $.25$)

and large ϵ's ($\epsilon \geqslant .75$).

The dependence of the stability of the procedure on the parameter ϵ was observed. An increase in ϵ not only had a smoothing effect on the behavior of ϕ_n, but also indicated a decrease in the final bound of ϕ_n. Nevertheless, in most of the cases, as ϵ approaches 1, the procedure converged slowly.

In the case of constant $\alpha_n = \alpha$, the effect of the step size α was noticeable. For a small α the behavior of the procedure was closely predictable by its continuous-time counterpart, P1. That is, one could observe that ϕ_n and $\|g_n\|$ were not increasing with n, or better, essentially decreasing with n. In any case, a decrease in step size was followed by a decrease of the final ϕ_n. However, the number of steps necessary for ϕ_n to reach the final bound was naturally increased by a decrease in α.

These results are summarized in Figs. 2, 3, 4, 5, 6, 7, and 8.

FINAL REMARKS

Different approaches to the problem of classification can be found in the literature. In particular, an interesting point of view was introduced by Wong and Eisenberg (Ref. 8). These authors study the case where the vectors are vertices of a hypercube of ℓ-dimensions (this is the case when the measurements performed in the patterns only assume the values $\pm$ 1. That is, one checks if the given pattern has or has not each feature). A one-to-one correspondence between these vertices and intervals in the real line is established. Then, the problem of classification is studied on the real line. A procedure using iterative projections is then used to determine a linear functional that correctly classifies all points (in case of strongly-linearly-separable points).

It is interesting to notethat Wong-Eisenberg's procedures points out some similarities with a different class of problems. Formally, the iterative projection method is equal to the method used in the solution of some problems studied by Sandberg (Ref. 9) in connection with certain contraction mapping problems. However, the conditions under which these procedures are applicable are quite different. Wong and Eisenberg use a finite dimensional space to prove termination of the procedure for a nonlinearity violating the conditions used by Sandberg to prove uniqueness of the solution and convergence of the iterative projections in more general spaces. It would be interesting to see how much these iterative projection approaches could be extended in both cases in order to get some overlapping on the conditions of applicability of the methods.

The approach of Wong and Eisenberg is also closely related to the approach due to Aizerman-et-al (Ref. 11). In this case, the spaces used are more general than those used by Wong and Eisenberg and instead of projecting the indicators of the misclassified intervals (as in Wong-Eisenberg's approach) the projection of a " δ -function" at the misclassified points is used (this is accomplished by means of potential functions). Once more, this indicates that a good direction of inquiry might be the extension of Wong-Eisenberg's procedure to more general spaces.

The Wong-Eisenberg procedure of iterative projections is equivalent to the procedures presented in Chapter IV of this work (for $\epsilon = 0$ and $\epsilon = 1$) for the case of linearly-separable sets. Although the proofs given both here and by Wong and Eisenberg are heavily dependent on the finite dimensionality of E_ℓ, it is believed that the procedures can be extended to vectors in more general spaces. This would make stronger the connections between these techniques and the works of Sandberg and Aizerman-et-al.

The discretization of procedure P2 was not pursued. The author believes that such discretization can be done along similar lines as those of Chapter IV.

Fig. 2

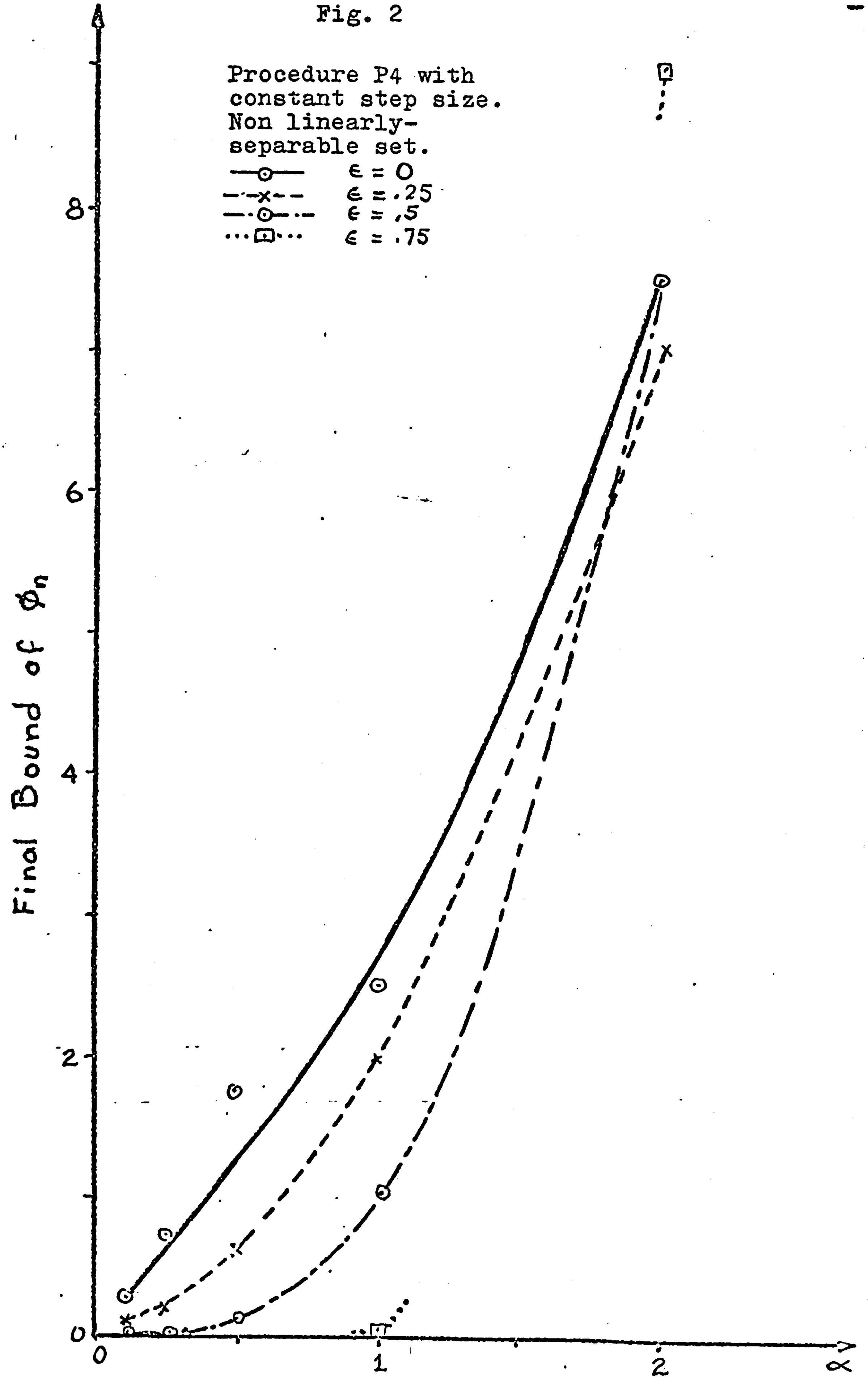
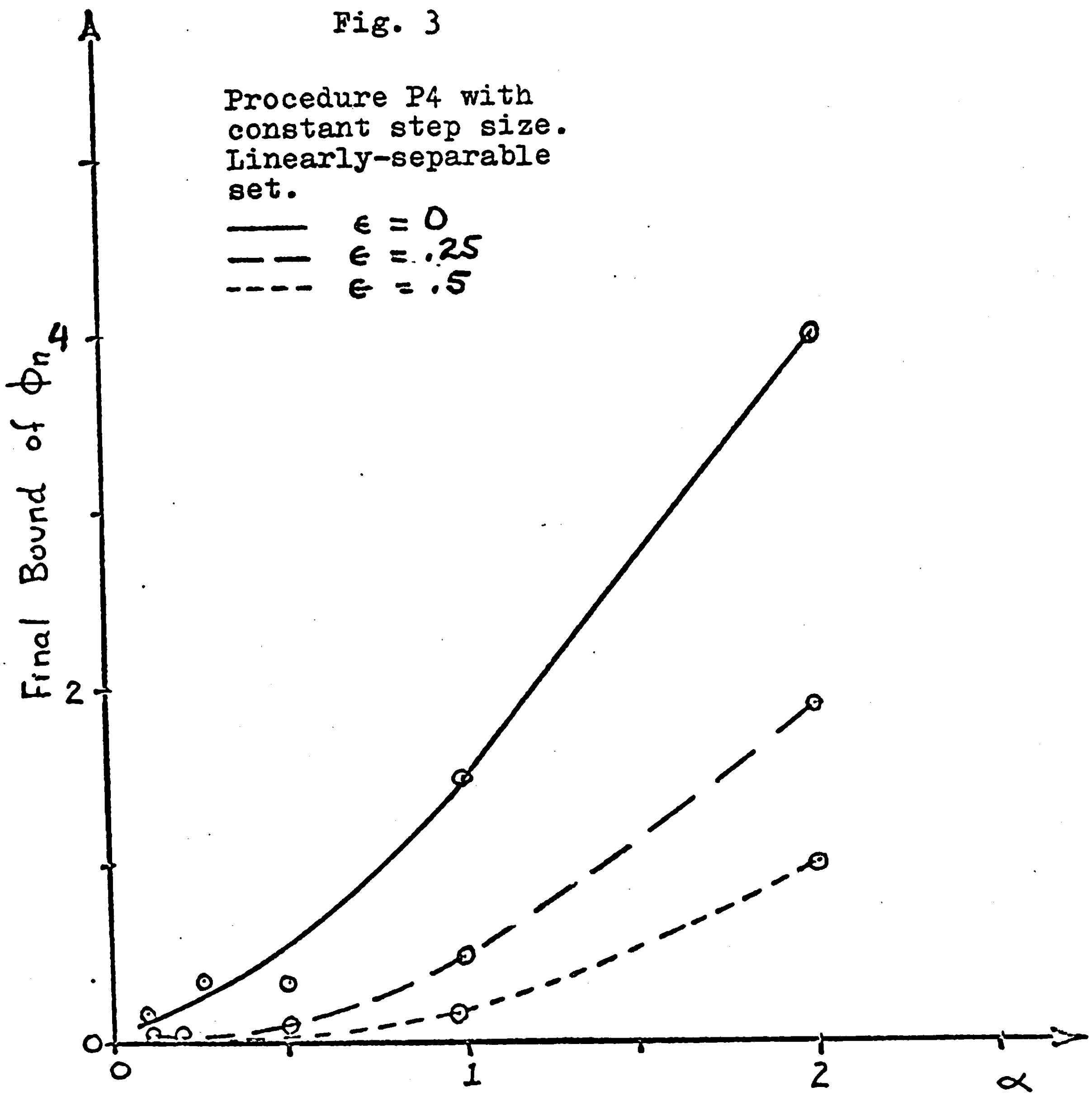

Procedure P4 with
constant step size.
Non linearly-
separable set.
$\epsilon = 0$
$\epsilon = .25$
$\epsilon = .5$
$\epsilon = .75$
Final Bound of Φ_n
0
2
4
6
8
0
1
2
α

Fig. 3


Procedure P4 with
constant step size.
Linearly-separable
set.
ε = 0
ε = .25
ε = .5
Final Bound of Φn
4
2
0
0
1
2
α

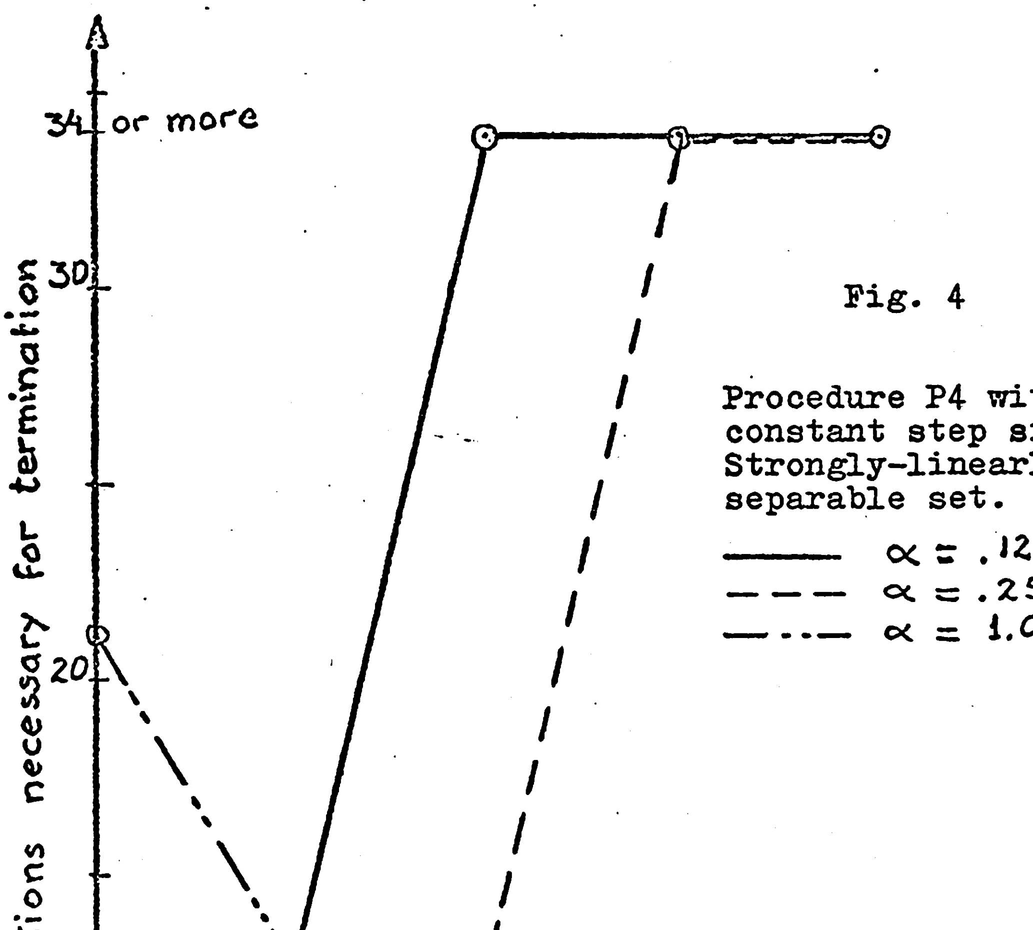


Fig. 4

Procedure P4 with constant step size. Strongly-linearly-separable set.

——— $\alpha = .125$
– – – $\alpha = .25$
—·-— $\alpha = 1.0$

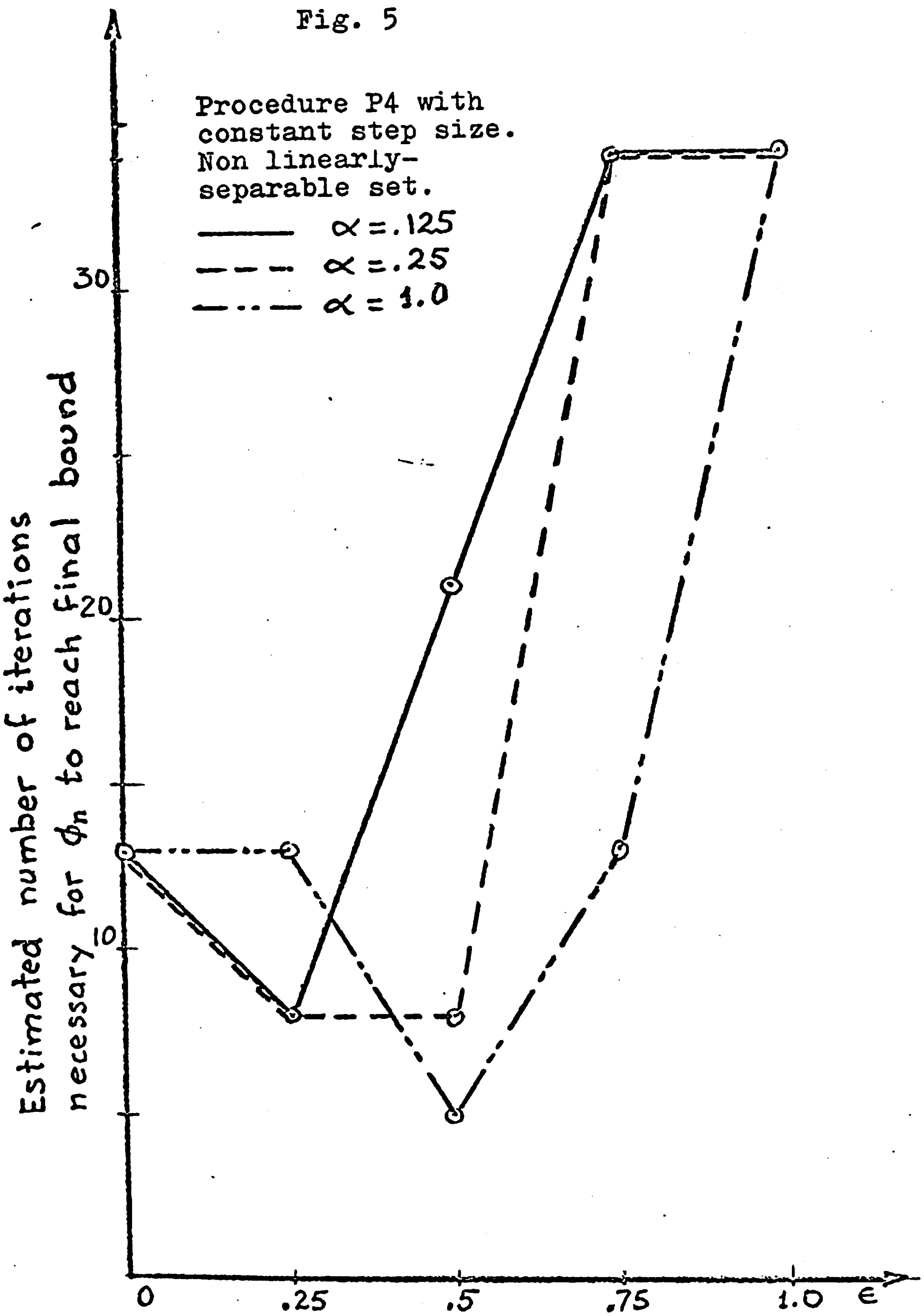
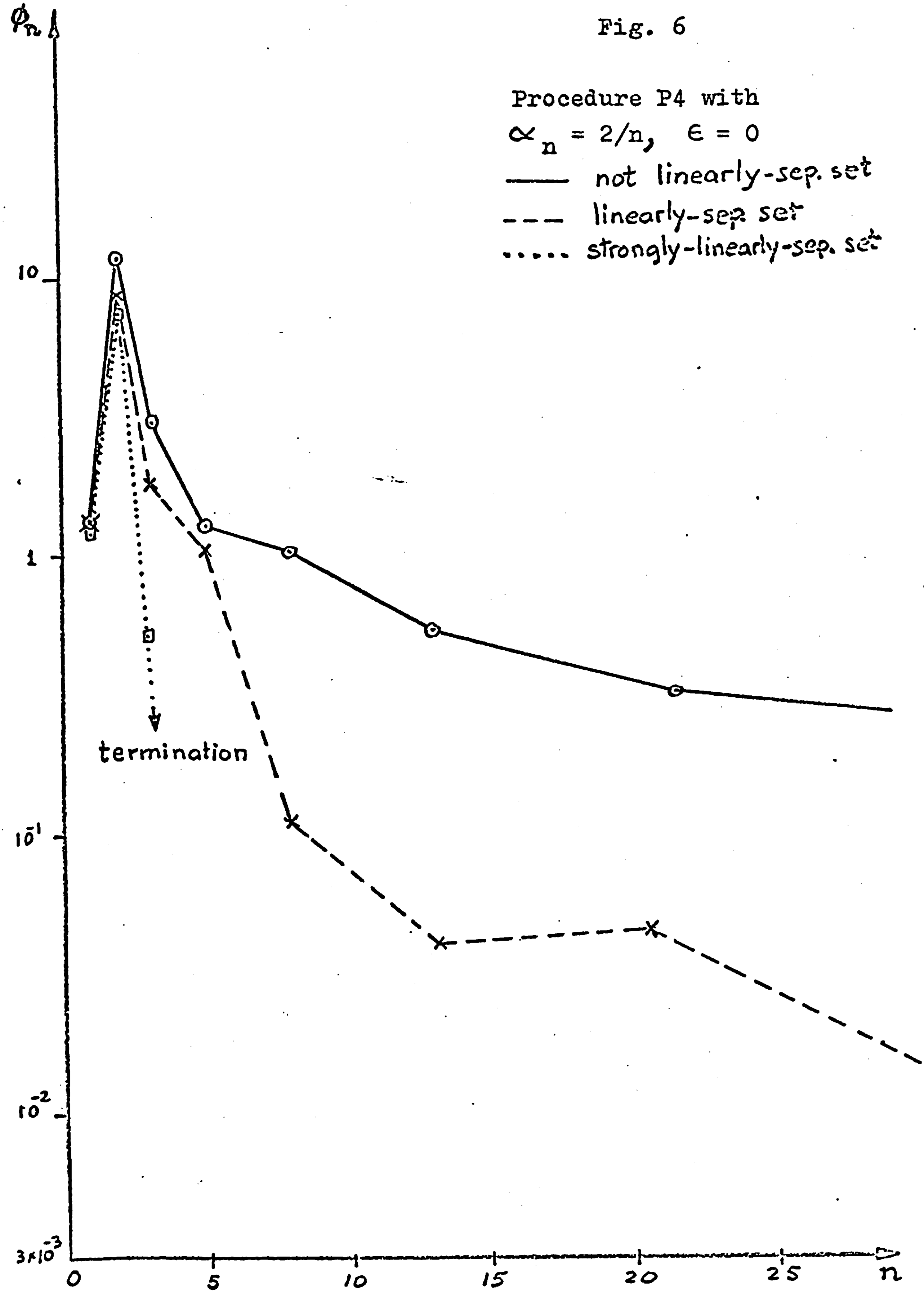


Fig. 5
Procedure P4 with constant step size. Non linearly-separable set.
$\alpha = .125$
$\alpha = .25$
$\alpha = 1.0$
30
20
10
Estimated number of iterations necessary for ϕ_n to reach final bound
0
.25
.5
.75
1.0
ϵ

Fig. 6



Procedure P4 with
α_n = 2/n, ε = 0
not linearly-sep. set
linearly-sep. set
strongly-linearly-sep. set
φ_n
10
1
10^-1
10^-2
3x10^-3
termination
0
5
10
15
20
25
n

Fig. 7

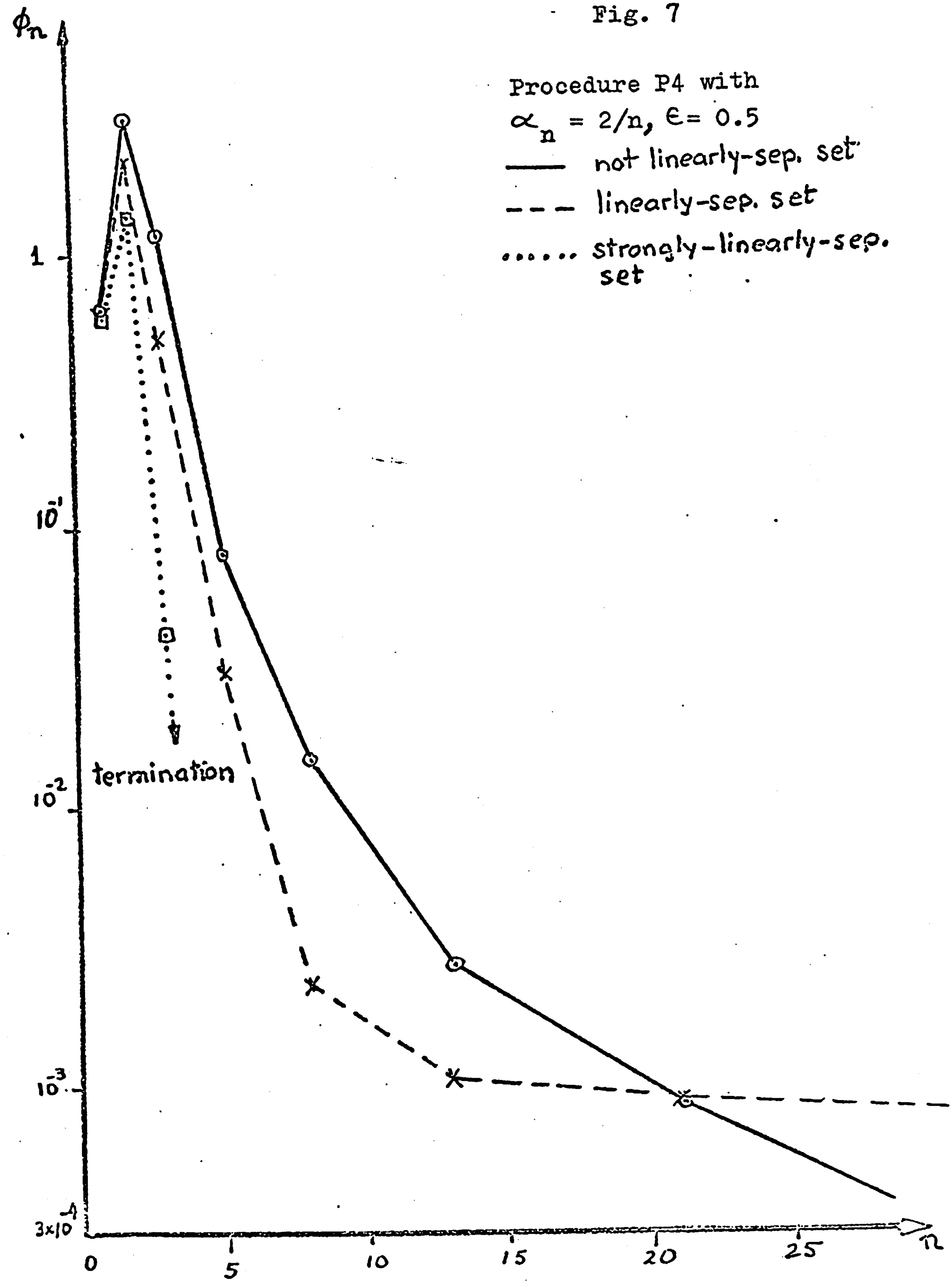
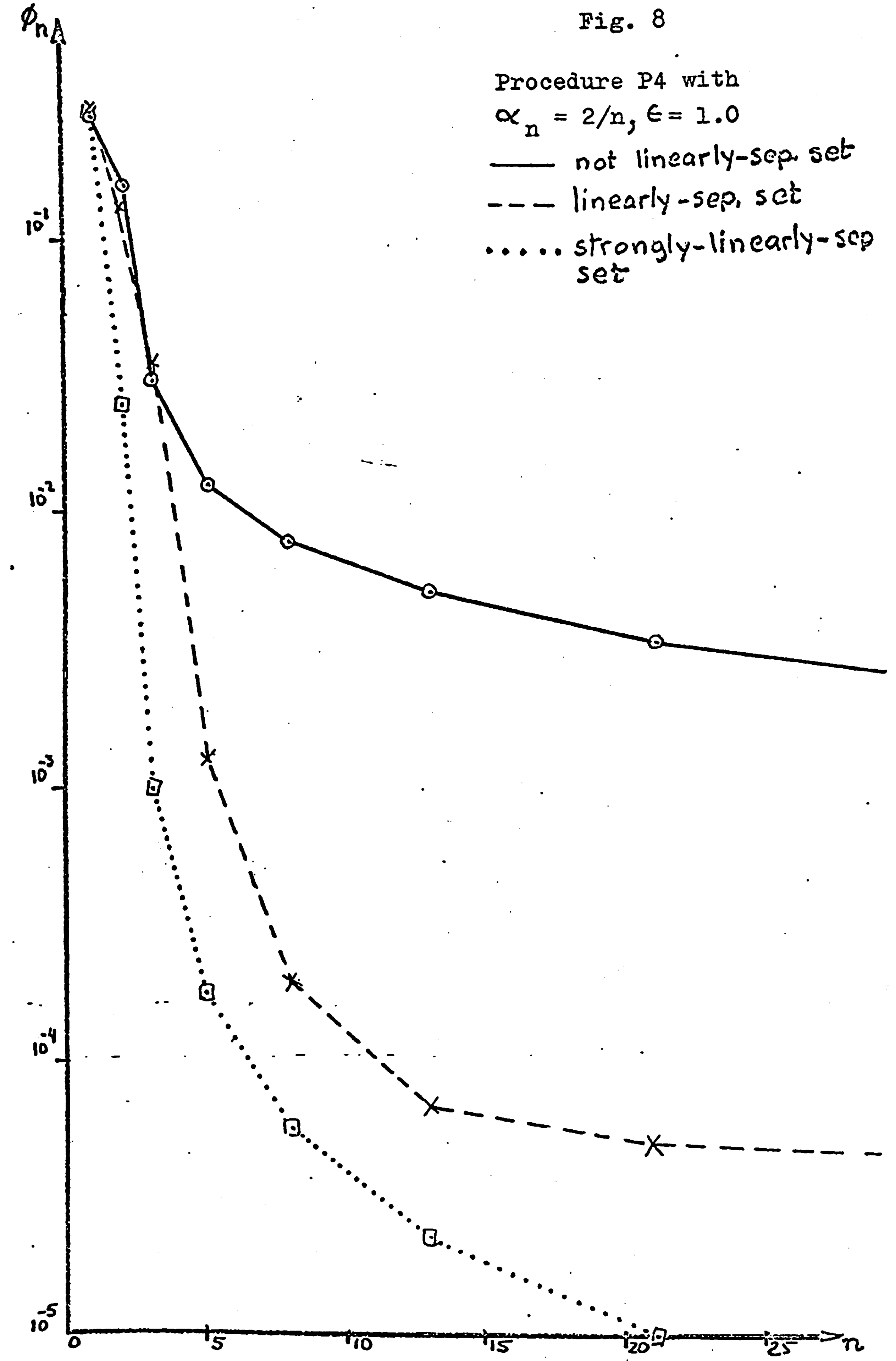

ϕ_n
Procedure P4 with
$\alpha_n = 2/n$, $\epsilon = 0.5$
not linearly-sep. set
linearly-sep. set
strongly-linearly-sep. set
1
10^{-1}
10^{-2}
10^{-3}
3×10^{-4}
termination
0
5
10
15
20
25
n

Fig. 8


ϕ_n
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
0
5
10
15
20
25
n
Procedure P4 with
$\alpha_n = 2/n, \epsilon = 1.0$
not linearly-sep. set
linearly-sep. set
strongly-linearly-sep set

REFERENCES

1. F. Rosenblatt, "Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms". Spartan Books, 1961.
2. A. B. J. Novikoff, "On Convergence Proofs for Perceptrons", Stanford Research Institute, January, 1963.
3. S. Agmon, "The Relaxation Method for Linear Inequalities", Canadian Journal of Mathematics, Vol. 6, 1954.
4. T. S. Motzkin and I. J. Schoenberg, "The Relaxation Method for Linear Inequalities", Canadian Journal of Mathematics, Vol. 6, 1954.
5. B. Widrow and M. E. Hoff, "Adaptive Switching Circuits", T. R. N. 1553-1, Stanford Electronics Lab., June,1960.
6. C. H. Mays, "Adaptive Threshold Logic", T. R. N. 1557-1, Stanford Electronics Lab., April, 1963.
7. S. Karlin, "Mathematical Methods and Theory in Games, Programming and Economics", Vol. 1, Addison-Wesley, 1959.
8. E. Wong and E. Eisenberg, "Iterative Synthesis of Threshold Functions", Journal of Mathematical Analysis and Applications, Vol. 11, No. 1-3, July 1965 (Belgium).
9. I. W. Sandberg, "On the Properties of Some Systems that Distort Signals" - I, The Bell Systems Technical Journal, September, 1963.

10. Bradley Efron, "The Perceptron Correction Procedure in Non-Separable Situations" Rome Air Development Center, Technical Report RADC-TDR-63-533, February, 1964.

11. M. A. Aizerman, E. M. Braverman and L. I. Rozenoer, "Theoretical Foundations of the Method of Potential Functions in the Problem of Teaching Automata to Classify Input Situations", Avtomatika i Telemekhanika, Vol. 26, 6, 1964.
"The Probability Problem of Pattern Recognition Learning and the Method of Potential Functions", Avtomatika i Telemekhanika, Vol. 25, 9, 1964.

12. K. Arrow, L. Hurwicz and H. Uzawa, "Studies in Linear and Non-Linear Programming, Stanford University Press, 1958.

13. N. J. Nilsson, "Learning Machines: Foundations of Trainable Pattern Classifying Systems", McGraw-Hill, 1965.

14. T. M. Cover, "Classification and Generalization Capabilities of Linear Threshold Units", Rome Air Development Center Technical Note No. November, 1963.