A PROGRAMMING MODEL FOR STUDENT ENROLMENT PLANNING

L. H. CAMPBELL

Abstract

The problem of planning the annual intakes to a university course, in which there are capacity constraints on the total enrolment, so as to produce a steady transition into an eventual no-growth situation is formulated as a linear program. The special structure of the problem is exploited to find a particular, optimal solution and to show that the addition of integrality constraints on the intakes poses no additional difficulty. The usefulness of the proposed methods is illustrated with an example from the University of Adelaide.

1. Introduction

The problems of modelling university student enrolments and flows between enrolment states have been studied by a number of authors, for example Gani [3] and Marshall and Oliver [4, 5], with a view to forecasting total enrolments. The theory of planning enrolments, when there are maximum or desired capacities to be taken into account, has also been developed, as instanced by the control theory approach of Alper, Armitage and Smith [1]. However, the acceptance of these models by university administrators has been slow, because, almost invariably, detailed data on student flows have not been kept in the past and the effort required to collect them from available records may be considerable. The forecasting which is of necessity carried out makes use of whatever data are available.

If one can estimate the fraction of students from a given entering cohort who will subsequently attend each year after initial enrolment, then one may easily forecast total enrolments given the size of each cohort. A method of this type is in use at the University of Adelaide and other Australian universities. It is commonly observed that these fractions are largely invariant with cohort, a stability which has also been observed at the University of California,

30

Berkeley, by Marshall, Oliver and Suslow [6]. Although in theory there is no reason why it should be so, these fractions, which we shall call retention rates, are always non-increasing with time, at least for the first ten years after initial enrolment.

In this paper, we consider that each student enrols for a *course*, which may be identified by the type of qualification received or the curriculum to be followed, and progresses through this course over several years. (The unit time interval of a year is used here although one may think of semesters or terms or some other convenient unit.) We assume that retention-rate data are known for each course for which a plan is required; we take all students who first enter the course in the same year to be in the same cohort. We consider the problem of planning the enrolments for a course, given a maximum possible enrolment in each future year and using the retention rates to estimate the enrolments from each cohort. In particular, we are interested in planning for eventual constant enrolments and seek to control the intake to the course of new students in such a way that the transition to the "no-growth" phase occurs smoothly. By smoothly, we mean that the number of new students should not decrease from one year to the next.

After formulating the linear-programming model in Section 2, we examine, in Section 3, the properties of a particular optimal solution which we show can be easily calculated. The equivalent results in the case where integrality constraints are added to the basic linear program are noted in the succeeding section. A numerical example taken from the University of Adelaide is discussed in Section 5.

2. Formulation of the model

We shall take the present calendar year as year $t = 0$, with plans to be formulated for years $t = 1$ to $t = T$. By cohort j, we will mean those students who first enter the course in year $t = j$. The total enrolment in the course each year consists of students from cohorts which entered in the same or previous years. For each cohort j from which students will be enrolled in at least one of the years $t = 1, \cdots, T$, define the following *retention rates*:

$R(j,t)$ = proportion of students in cohort j who are enrolled in the course in year $t, t = 1, \cdots, T, t \geq j$.

These retention rates need to be estimated from past data. In this paper, we will assume that, for each future cohort $j = 1, \cdots, T$, the estimates are the same; that is, for $j = 1, \cdots, T$,

$$R(j,t) = P(t - j), \qquad t = 1, \cdots, T, \qquad t \geq j.$$

Further, it is assumed that

$$P(0) = 1 \quad \text{and} \quad P(k) \geqq P(k+1),$$ (2.1)

$$k = 0, 1, \cdots, T - 1.$$

For each year $t = 1, \cdots, T$, let $K(t)$ be the maximum number of students which can be accommodated in the course and let $e(j)$ be the number of students in cohort j. Then, for each of the years $t = 1, \cdots, T$, we can define a *net capacity*, $C(t)$, by

$$C(t) = K(t) - \sum_{j \leqq 0} R(j, t) e(j).$$

We assume that each $C(t)$ is non-negative.

The variables whose values we wish to prescribe are the sizes, $e(j)$, of the cohorts for $j = 1, \cdots, T$. They are constrained firstly by the requirement that the net capacities should not be exceeded; thus

$$\sum_{j=1}^{t} P(t-j) e(j) \leqq C(t),$$ (2.2)

$$\text{for each} \quad t = 1, \cdots, T.$$

Further, the criterion that there should be no reduction in new enrolments each year is specified by

$$e(j) \leqq e(j+1),$$ (2.3)

$$\text{for each} \quad j = 1, \cdots, T - 1,$$

together with the constraint that the intake in the last year should be no greater than some maximum $\bar{e}$:

$$e(T) \leqq \bar{e}.$$ (2.4)

Normally, $\bar{e}$ would be the constant intake which, if repeated indefinitely, would produce the desired constant total enrolment, $\bar{C}$, say. Then

$$\bar{e} = \bar{C} \bigg/ \sum_{j=0}^{\infty} P(j).$$ (2.5)

Various objective functions may be formulated, depending upon the priorities attached to particular aspects of course development; one may desire, for example, to place emphasis on achieving total enrolments close to the capacities in the earlier years rather than in the later ones, or one may consider that there is a cost associated with a change in cohort-size from one year to the next. However, in this paper, we choose to consider minimizing the oversupply of capacity over all the years $t = 1, \cdots, T$; thus we wish to

$$\text{minimize} \quad \sum_{t=1}^{T} \left\{ C(t) - \sum_{j=1}^{t} P(t-j) e(j) \right\}. \tag{2.6}$$

Then, the required cohort sizes, $e(j), j = 1, \cdots, T$, are those which solve the following linear program:

$$\text{minimize} \quad \sum_{j=1}^{T} \left\{ C(t) - \sum_{j=1}^{t} P(t-j) e(j) \right\},$$

$$\text{subject to} \quad \sum_{j=1}^{t} P(t-j) e(j) \leqq C(t), \quad t = 1, \cdots, T,$$

$$e(j) - e(j+1) \leqq 0, \qquad j = 1, \cdots, T-1,$$

$$e(T) \leqq \bar{e},$$

$$e(1) \geqq 0. \tag{2.7}$$

It may be shown, using the methods of Section 3, that all feasible solutions which minimize (2.6) also maximize the total intake to the course over the T planning years.

In Section 4, we consider the above linear program with the additional constraints:

$$e(j) \quad \text{integer}, \quad j = 1, \cdots, T. \tag{2.8}$$

3.   An optimal solution of the linear program

We show in this section that a feasible solution $e'(j), j = 1, \cdots, T$ of the linear program satisfying the following criterion, which we shall the Linked-Constraint Property (LCP), is an optimal solution.

THE LINKED-CONSTRAINT PROPERTY (LCP).

For each $t = 1, \cdots, T - 1$, *either*

$$\sum_{j=1}^{t} P(t-j) e'(j) = C(t), \quad or \tag{3.1}$$

$$e'(t) = e'(t+1), \tag{3.2}$$

and for $t = T$, *either*

$$\sum_{j=1}^{T} P(T-j) e'(j) = C(T), \quad or \tag{3.3}$$

$$e'(T) = \bar{e}. \tag{3.4}$$

According to (3.2), a solution with the LCP has successive cohort-sizes equal until the capacity of the course is reached in a particular year; then, when (3.1) is satisfied, an increase in the size of the entering cohorts may occur in the

following year. The critical property of this solution is given in the following easy lemma.

LEMMA 1.   *Let $e'(j), j = 1, \cdots, T$, be a feasible solution with the LCP and let $e(j), j = 1, \cdots, T$, be any other feasible solution of the linear program. Let k be the smallest j for which $e'(j) \neq e(j)$. Then*

$$e'(k) > e(k).   \tag{3.5}$$

COROLLARY.   *The feasible solution $e'(j), j = 1, \cdots, T$, with the LCP is unique.*

THEOREM 1.   *The feasible solution $e'(j), j = 1, \cdots, T$, with the LCP is an optimal solution of the linear program.*

PROOF.   Consider an optimal solution $e^*(j), j = 1, \cdots, T$, different from $e'(j), j = 1, \cdots, T$. Let Z^* and Z' be their respective objective values.

1.   *Choose k.* Let k be the smallest value of j for which $e'(j) \neq e^*(j)$. Then, by Lemma 1, $e'(k) > e^*(k)$. Now,

$$Z' - Z^* = \sum_{j=k}^{T} \left(\sum_{t=j}^{T} P(t - j) \right) (e^*(j) - e'(j)).   \tag{3.6}$$

If $k = T$, then $Z' < Z^*$ which contradicts optimality of Z^*. Thus $k < T$.

2.   *Choose m.* Further, if $e^*(j) \leq e'(j)$ for $j = k + 1, \cdots, T$, then again $Z' < Z^*$ from (3.6). Therefore, there must be some j, $k < j \leq T$, for which

$$e'(j) < e^*(j).   \tag{3.7}$$

Let m be the smallest such j.

3.   *Choose l.* Let l be the smallest $j \geq k$ for which

$$e^*(l) < e'(l)   \quad \text{while} \quad   e^*(l + 1) \geq e'(l + 1).   \tag{3.8}$$

We note that $k \leq l < m$.

4.   *Define a new optimal solution.* Let

$$\varepsilon = \min\{e'(l) - e^*(l), e^*(m) - e'(m)\},   \tag{3.9}$$

and define a new solution $e^{**}(j), j = 1, \cdots, T$, by

$$e^{**}(m) = e^*(m) - \varepsilon,$$
$$e^{**}(l) = e^*(l) + \varepsilon,   \tag{3.10}$$
$$\text{and } e^{**}(j) = e^*(j), \text{ for all other } j.$$

This new solution clearly satisfies the constraints (2.3), (2.4) and (2.7), while the capacity constraints (2.2) also hold since, for $t < m$,

$$\sum_{j=1}^{t} P(t-j) e^{**}(j) \leq \sum_{j=1}^{t} P(t-j) e'(j) \leq C(t),$$

and for $t \geq m$, $\sum_{j=1}^{t} P(t-j) e^{**}(j)$

$$= \sum_{j=1}^{t} P(t-j) e^{*}(j) + \varepsilon \{P(t-l) - P(t-m)\}$$

$$\leq \sum_{j=1}^{t} P(t-j) e^{*}(j), \quad \text{since} \quad P(t-l) \leq P(t-m),$$

$$\leq C(t).$$

This solution is optimal for, if Z^{**} is its objective value, then

$$Z^{**} - Z^{*} = \varepsilon \left\{ \sum_{t=m}^{T} P(t-m) - \sum_{t=l}^{T} P(t-l) \right\} \leq 0. \tag{3.11}$$

Hence, since Z^{*} is optimal, $Z^{**} = Z^{*}$, which occurs only if

$$\sum_{t=l}^{m-1} P(T-t) = 0. \tag{3.12}$$

5.   *Compare the solutions.* Thus we have generated a new optimal solution $e^{**}(j), j = 1, \cdots, T$, which satisfies $e^{**}(k) \leq e'(k)$. If in fact $e^{**}(k) < e^{*}(k)$, we may replace $e^{*}(j)$ by $e^{**}(j)$ for each $j = 1, \cdots, T$ and go to 2.

6.   *Termination.* After each transformation of the optimal solution, either l decreases by 1 or m increases by at least 1. Eventually, then, we stop in Step 5 when $l = k$ and hence $e^{**}(k) = e'(k)$. Thus we generate a new optimal solution satisfying $e^{**}(j) = e'(j), j = 1, \cdots, k$, from an optimal solution satisfying only $e^{*}(j) = e'(j)$, $j < k$. Hence, we start again in Step 1 with this new optimal solution and transform it until we can no longer choose a new k. Then $e^{*}(j) = e'(j), j = 1, \cdots, T$, and hence the solution with the LCP is optimal.

A simple necessary condition for the existence of multiple optima, which is proved by considering equation (3.12), is the following result.

COROLLARY.   *If $P(T-1) > 0$ then the only optimal solution of the linear program is that one with the LCP.*

In order to find the feasible solution with the LCP, an algorithm which exploits the special structure of the linear program may be used. The following algorithm finds, on each pass, the value $\alpha(k)$ which is the maximum value the remaining intakes may attain if they are all set equal. Then all the remaining intakes up to the last year in which this solution achieves the net capacity are

set equal to this value. That the final solution is in fact that with the LCP is easily shown.

ALGORITHM.

Define $S(t) = \Sigma_{j=1}^{t} P(t-j)$, $t = 1, \cdots, T$.

Step 0.   Set $k = 0$, $r_0 = 0$, $C(0;t) = C(t)$, $t = 1, \cdots, T$.

Step 1.   Replace k by $k+1$. Define s as the largest integer $> r_{k-1}$ for which

$$C(k-1;s)/S(s-r_{k-1}) = \min\{C(k-1;t)/S(t-r_{k-1}); t = r_{k-1}+1, \cdots, T\}.$$

Step 2.   If $C(k-1;s)/S(s-r_{k-1}) \geqq \bar{e}$, set $r_k = T$ and $\alpha(k) = \bar{e}$. Otherwise, set $r_k = s$ and $\alpha(k) = C(k-1;s)/S(s-r_{k-1})$.

Step 3.   For $t \leqq r_{k-1}$ (if any), put $C(k;t) = C(k-1;t)$. For $t = r_{k-1}+1, \cdots, r_k$, put $C(k;t) = C(k-1;t) - S(t-r_{k-1})\alpha(k)$ and $e'(t) = \alpha(k)$. If $r_k = T$, go to Step 4. Otherwise, for $t = r_k+1, \cdots, T$, put

$$C(k;t) = C(k-1;t) - \{S(t-r_{k-1}) - S(t-r_k)\}\alpha(k),$$

and go to Step 1.

Step 4.   Put $Z' = \Sigma_{t=1}^{T} C(k;t)$. Stop with optimal value Z' and optimal solution $e'(1), \cdots, e'(T)$.

This algorithm performs better with this problem than the simplex algorithm in the sense that, if the problem is started with a basis of slack variables, the simplex algorithm will perform at least T pivots in order that all the $e(j)'s$ enter the basis whereas, after the row-sums, $S(t)$, are produced, this algorithm will take no more than T iterations; this is so since $r_{k-1} < r_k$ for each $k \geqq 1$.

4.   The program with integrality constraints

Although in real-life planning situations, one may be satisfied with rounding-off the optimal cohort sizes of the solution to the linear program, one may ask what happens if the additional constraints that each $e(j)$ should be integral are added to the program. In this section, we show that the special structure of the problem may be exploited to find a particular optimal, integer solution.

The algorithm given in Section 3 requires only minor modification for it to be used to find a feasible solution in which all the intakes are integral. The necessary modification is to replace Step 2 by the following:

Step 2′.   Let $A = C(k - 1; s)/S(s - r_{k-1})$. If $A \geqq \bar{e}$, set $r_k = T$ and $\alpha(k) = [\bar{e}]$, where $[\cdots]$ means integer part of. Otherwise, if A is integer, set $r_k = s$ and $\alpha(k) = A$, and, if A is non-integer, set $r_k = r_{k-1} + 1$ and $\alpha(k) = [A]$.

The integer solution, $e''(j)$, $j = 1, \cdots, T$, say, produced in this way satisfies an equivalent condition to that of Lemma 1, namely that, if $e(j)$, $j = 1, \cdots, T$, is any other integer, feasible solution and k is the smallest j for which $e''(j) \neq e(j)$, then $e''(k) > e(k)$. In fact, $e''(j)$, $j = 1, \cdots, T$, is an optimal integer solution; the proof is exactly the same as that for Theorem 1 since, if $e^*(j)$, $j = 1, \cdots, T$, is chosen to be an integer optimum, (3.9) and (3.10) imply each new solution produced will also be integer. Hence, the introduction of integrality constraints on the intakes imposes only a little extra work in the solution of the resultant mathematical program.

5. Numerical example

At the University of Adelaide, the B.Sc.-degree course is taught jointly by the Faculties of Science and Mathematical Sciences. The average retention rates for students in this course are given on the left of the table; these data do not differentiate between pass-degree and honours-degree candidates. Column (ii) of the table gives the estimated totals of students enrolled in each future year from cohorts which entered before 1975, found by applying the average retention rates to the known cohort sizes. For the years 1975 to 1978, the University's plans have already been formulated; hence the capacities in the course up to 1978 are known and are shown in the first four rows of columns (iii) and (vi). After 1978, various building and staff-development programs are possible which affect the number of students which can be accommodated.

The table lists two cases which differ in the capacities after 1978. Case 1 postulates a ceiling of 1,600 students, as shown in column (iii). The value of $\bar{e}$, 425, was calculated using (2.5) and rounding the result to the nearest multiple of 5. The net capacities, $C(t)$, and the optimal intakes, $e(t)$, found using the algorithm in Section 3, are shown in columns (iv) and (v) respectively. The same calculations were performed for Case 2 which uses a constant capacity of 1,650 students after 1978, and the results listed in columns (vi) to (viii). The intakes are the same up to 1978, when the capacity constraint holds with equality.

In evaluating which intake strategy to adopt, the University must also consider the rate of expansion required to accommodate the increased numbers of students. In case 2, the optimal intakes produce a total of 1,602 students in 1979; thus a capacity of about 1,600 students would be sufficient in that year. During the 1979–81 financial triennium, a modest increase of 15 places per year would be required, until the total of 1,650 places is reached at the beginning of

Possible Strategies for B.Sc. — Degree Course,
University of Adelaide

t	Average Retention Rates $P(t-1)$	(i)	(ii)	(iii)	Case 1. $\bar{z} = 425$ (iv)	(v)	(vi)	Case 2. $\bar{z} = 440$ (vii)	(viii)
1	1.000	1975	1113	1553	440	420.2	1553	440	420.2
2	.8454	1976	777	1565	788	420.2	1565	788	420.2
3	.7521	1977	481	1577	1096	420.2	1577	1096	420.2
4	.5916	1978	242	1582	1340	420.2	1582	1340	420.2
5	.2219	1979	153	1600	1447	422.3	1650	1497	436.4
6	.0898	1980	116	1600	1484	422.3	1650	1534	436.4
7	.0601	1981	91	1600	1509	422.3	1650	1559	436.4
8	.0572	1982	68	1600	1532	422.3	1650	1582	436.4
9	.0428	1983	51	1600	1549	422.3	1650	1599	436.4
10	.0250	1984	42	1600	1558	422.3	1650	1608	436.4
11	.0206	1985	35	1600	1565	422.3	1650	1615	436.4
$\geqq 12$	.0686								

(i)   Year.
(ii)   Number of students enrolled from cohorts entering before 1975.
(iii) & (vi)   Possible maximum total enrolments.
(iv) & (vii)   Corresponding maximum enrolments from cohorts from 1975.
(v) & (viii)   Corresponding optimum intakes (Algorithm 1).

the following triennium. The strategy actually chosen will depend on the availability of external funding and the University's internal priorities. However, it is clear that the University can control the enrolments in B.Sc.-degree course to produce an eventual steady-state situation by suitably fixing intake quotas over a period of years.

6.   Discussion

The structured model of this paper, with its accompanying solution procedures, addresses itself to the problems of planning gross enrolments given little information on student flows and a long-term aim of no growth. The structure of the constraints is critical for the proposed solution procedures. The non-increasing nature of the retention rates apparently occurs with almost all student cohorts, while the specification that the intakes should be non-decreasing is applicable particularly to the smooth transition into a no-growth phase; in the absence of this criterion, the capacities could be met exactly for a wide class of problems.

The algorithms given here find that unique optimal strategy which takes students into the course as early as possible in the planning years. If it is desired to find other optimal strategies, for example that which delays development for as long as possible, then the algorithmic proof of Theorem 1 may be adapted to the purpose. However, only a necessary condition for the existence of multiple optima has been given.

If desired total enrolments, rather than capacities, are known, then a goal-programming approach, like that used by Schroeder [7] in other areas of university management, may be adopted. Indeed, goal-programming methods may be useful for the short-term control of intakes once the desired totals have been produced by the methods of this paper. If more information about the distributions of the retention rates were available, one could specify the capacity constraints in the form of chance constraints which could be converted to linear inequality constraints using methods suggested by Charnes and Cooper [2]; in this process, of course, the structure required of the problem for the use of the methods of this paper most likely would be destroyed.

7.   Acknowledgements

References

[1] P. Alper, P. H. Armitage and C. S. Smith, 'Educational models, manpower, planning and control', *Operational Res. Quart.* 18 (1967), 93–103.
[2] A. Charnes and W. W. Cooper, 'Deterministic equivalents for optimizing and satisfying under chance constraints', *Opns. Res.* 11 (1963), 18–39.
[3] J. Gani, 'Formulae for projecting enrolments and degrees awarded in universities', *J. Roy. Statist. Soc.* A126 (1963), 400–409.
[4] K. T. Marshall, 'A comparison of two personnel prediction models', *Opns. Res.* 21 (1973), 810–822.
[5] K. T. Marshall and R. M. Oliver, 'A constant-work model for student attendance and enrollment', *Opns. Res.* 18 (1970), 193–206.
[6] K. T. Marshall, R. M. Oliver and S. S. Suslow, 'Undergraduate enrollments and attendance patterns', *Report No.* 4 (March, 1970), Administrative Studies Project in Higher Education, University of California, Berkeley.
[7] R. G. Schroeder, 'Resource planning in university management by goal programming', *Opns. Res.* 22 (1974), 700–710.

The University of Adelaide
Adelaide, South Australia.