WESTERN AGRICULTURAL ECONOMICS ASSOCIATION

PROCEEDINGS
1970

Forty-Third Annual Meeting
July 19-22, 1970
RAMADA INN
TUCSON, ARIZONA

SEQUENTIAL SAMPLING AND SIMULATION: AN OPTIMIZING PROCEDURE

Harry P. Mapp, Jr. and Vernon R. Eidman[1]

Simulation models have been applied in a number of farm management studies during the past several years. Because simulation is not an analytic optimizing procedure, the models have typically been used as experimental devices enabling the researcher to adjust the level of certain endogenous variables and estimate the effect on the resulting distribution(s) of interest. It is difficult (perhaps impossible) given the current development of simulation procedures to find an optimum solution to simulation problems. However, researchers may be interested in utilizing procedures which move closer to an optimum rather than comparing several plans, none of which may be very desirable.

The purpose of this paper is to briefly discuss two maximum seeking methods and their applicability to computer simulation models. The first of these methods, sampling by steepest ascent, has been used empirically and may be recommended for certain types of response surfaces. The potential value of an alternative method, sequential sampling, is illustrated in the latter portion of the paper.

MAXIMUM SEEKING METHODS

Suppose the response surface for a series of variables is

$$y = f(x_1, x_2, \dots, x_n)$$

where y is the response and x_i is the level of the ith factor. Each combination of the x variables represents a point on the response surface. If n is small, the response surface can usually be well defined by selecting several levels for each x variable and simulating all combinations. However, as the number of variables increases, the response surface becomes increasingly complex. If n is large, simulation of every combination of several levels for each variable is not a realistic alternative. Thus, there has been much interest in designing methods to investigate response surfaces in an attempt to locate their maximum or optimum points in an efficient manner.

Sampling by Steepest Ascent

One technique which has been used empirically is sampling by steepest ascent. Brooks (2) evaluated factorial, univariate, random and steepest ascent methods of attaining the maximum point on several different response surfaces. He discovered the most efficient method to be sampling by steepest ascent. Each of the surfaces to be investigated was smooth and uniform and possessed a single, well-defined maximum. However, if we were investigating the net returns surface of a representative farm firm by experimentally altering enterprise and/or factor combinations, we would not necessarily expect the surface to be smooth and uniform. If the surface is not uniform, sampling by steepest ascent may reach a relative rather than a global maximum.

Box (1) recommended exploring a surface by simulating responses for a number of points and fitting a polynomial equation to these points.[2] The fitted surface is used as an approximation of the real surface in a sampling by steepest ascent procedure. Since the fitted surface is smooth, a maximum point is located by the procedure.

Sampling by steepest ascent utilizes the magnitudes and signs of the slopes of the response surface to determine the direction of steepest ascent to the maximum point. Maass (3, p. 400) explains the procedure by assuming that a system may be described by a series of n variables at some base point, $x_1^o, \dots, x_n^o$, and a corresponding point on the response surface, R^o. Then $R^o = f(x_1^o, \dots, x_n^o)$. A change in the level of any x variable will result in a change in the level of R. If a small increment Δx_i has been added to each variable in turn, then the change in response is ΔR_i and the rate of change is $\Delta R_i / \Delta x_i$. The movement from the base, x_i^o, to a new base, x_i^1, is given by

$$x_i^1 - x_i^o = c(\Delta R_i / \Delta x_i)$$

where c is the constant of proportionality. The relationship between the distance, d, between bases and the constant of proportionality is given by the Pythagorean equation

$$d = [\Sigma_{i=1}^{n} (x_i^1 - x_i^o)^2]^{1/2} = c[\Sigma_{i=1}^{n} (\Delta R_i/\Delta x_i)^2]^{1/2}$$

or

$$c = d/[\Sigma_{i=1}^{n} (\Delta R_i/\Delta x_i)^2]^{1/2}.$$

The constant of proportionality is found by dividing the distance between bases by the square root of the sum of squares of the rate of change of response with respect to each changing variable. The constant, c, times the rate of change of response with respect to a change in each variable considered separately, gives the amount by which that variable must be incremented in calculating a new base. Calculations are continued from base to base until incrementing each factor results in a reduction in response. If the response surface is smooth, this point will be a global maximum. A weakness of the procedure is that if irregularities or discontinuities exist, the point may be a local rather than global maximum.

The above weakness of steepest ascent sampling is emphasized by Zusman and Amaid (6, p. 590). They applied the procedure to a present value surface in search of optimal managerial policies. Only three variables were sampled since a factorial analysis provided sufficient information to determine the optimum levels of all other factors. A maximum point on the present value surface was located, however, they recognized the possibility of converging to a relative maximum rather than to an absolute maximum.

Sequential Uniform-Grid Sampling

A uniform-grid sampling procedure may overcome the above limitation of sampling by steepest ascent. The basis for the sequential sampling procedure is a uniform-grid sampling design. If the response surface may be described by the effects of n variables, $X_1, X_2, \ldots, X_n$, then it is possible to locate a series of points on the surface by experimentally altering the combination of variables. If there are n variables and k possible values of each variable then a complete factorial design would require k^n observations. Approximation of the maximum in an efficient manner would require either n or k to be small.

In practical applications, the X_i's may represent enterprise levels, amounts of available resources, lending rules and other decision variables. The number of variables of interest may be small or large depending on the complexity of the problem. However, the analysis can be simplified by altering the sampling design. Assume that a net returns surface may be described by the main effects of three variables, X_1, X_2, and X_3, and that the sum of the X_i's must equal R, where R represents the land available for production. Then to increase X_1 by a given amount, A, either X_2 must be reduced with X_3 constant or X_3 reduced with X_2 constant. These two alternatives provide two sampling points in a uniform-grid sampling design. With three variables, a total of six sampling points is sufficient to completely describe the possible alternatives. In general, with n variables, a total of n(n–1) sampling points would complete the sample design. The uniform-grid sampling design suggested for three variables is presented in Table 1.

Table 1. Uniform-grid sampling design.

Base	X_1	X_2	X_3
1	$X_1 + A$	X_2	$X_3 - A$
2	$X_1 + A$	$X_2 - A$	X_3
3	X_1	$X_2 + A$	$X_3 - A$
4	$X_1 - A$	$X_2 + A$	X_3
5	X_1	$X_2 - A$	$X_3 + A$
6	$X_1 - A$	X_2	$X_3 + A$

If the X_i represents three crop enterprises and the response surface represents net returns, a simulation model could generate net returns for each of the combinations of variables presented. That combination of variables which increases net returns by the largest amount is then chosen as the new base. This sequential feature allows the researcher to analyze the results of each simulation run prior to selection of a new base. Each new base is closer to the maximum point of the response surface than the previous base. The possibility of converging on a relative rather than absolute maximum is minimized by having a uniform-grid of sampling points around every base point. Once every sampling point results in lower net returns than the current base combination, the maximum point has been reached. A check may be performed by altering the value of A in the experimental design. A small value of A will check points in the immediate vicinity of the located maximum and a large value of A will provide a check for regions of the response surface some distance from the indicated maximum. While this procedure falls short of optimizing in the linear programming sense, it is as effective as sampling by steepest ascent, and is less likely to locate a relative rather than a global maximum.

AN APPLICATION

To add a measure of empirical validity to the above theoretical treatment, a comparison of steepest ascent sampling and uniform-grid sampling was made with the aid of a simplified farm firm computer simulation model (4). The simulation model is based on a 335-acre loam cropland farm in Southwestern Oklahoma. Crop alternatives on the farm are cotton, grain sorghum, wheat, alfalfa and forage sorghum hay. All crops produced are assumed sold during the production period. Compliance with the wheat (130 acres) and cotton (80 acres) programs was assumed. Thirty acres of cropland were assumed to be in the conserving base (alfalfa) leaving 85 acres to be allocated to grain sorghum, additional alfalfa and forage sorghum.

Yields and prices of the five crop alternatives were assumed to be normally distributed random variables. Thirty random normal deviates were generated for each crop yield and each crop price. The random normal deviates were multiplied by the standard deviations of yield and price and added to the mean yield and price to obtain 30 yield and price observations for each crop. The 30 yield and price observations were used to obtain 30 net return figures for each combination of the decision variables.

In applying sampling by steepest ascent a Fortran IV program calculated rates of change in net returns with respect to each variable, the constant of proportionality and the distance between bases. The program moved upward from base to base until the 85 acres of available land were exhausted. The optimum solution contained, besides the 130 acres of wheat and 80 acres of cotton, 29 acres of alfalfa, 23 acres of forage sorghum and 33 acres of grain sorghum. The expected net returns for this combination of enterprises was $11,000.06. The standard deviation was $1,858.05.

The base chosen for the first sequential sampling run was 29 acres of alfalfa, 29 acres of forage sorghum and 27 acres of grain sorghum, in addition to 130 acres of wheat and 80 acres of cotton. The results of the first run are presented in Table 2.

Table 2. Expected net returns and standard deviation of net returns for first sampling design run.

	Alfalfa	Forage	Grain	Expected Net Returns	Standard Deviation
	(acres)	(acres)	(acres)	($)	($)
Base	29.0	29.0	27.0	10,948.01	1853.10
1	34.0	29.0	22.0	10,927.70	1853.51
2	34.0	24.0	27.0	10,971.07	1856.26
3	39.0	34.0	22.0	10,904.64	1856.26
4	24.0	34.0	27.0	10,924.95	1856.08
5	29.0	24.0	32.0	10,991.39	1856.56
6	24.0	29.0	32.0	10,968.33	1856.36

Combinations 1 through 6 should be compared with the base. Combinations 2, 5 and 6 increase expected net returns over the base with 5 providing the greatest increase. Based on expected net returns, combination 5 is selected as the new base for a second uniform-grid sampling run.

A total of seven sampling runs were conducted to establish the maximum point on the net returns surface. A summary of the sampling bases is presented in Table 3.

Table 3. Expected net returns and standard deviation of net returns for seven sequential sampling bases.

Base	Alfalfa	Forage	Grain	Expected Net Returns	Standard Deviation
	(acres)	(acres)	(acres)	($)	($)
1	29.0	29.0	27.0	10,948.01	1853.10
2	29.0	24.0	32.0	10,991.39	1856.56
3	39.0	19.0	37.0	11,034.75	1866.61
4	29.0	5.0	51.0	11,156.20	1928.65
5	24.0	5.0	56.0	11,176.51	1947.58
6	5.0	5.0	75.0	11,253.71	2073.00
7	0	0	85.0	11,317.40	2166.60

The maximum point was generated by production of 85 acres of grain sorghum, 130 acres of wheat and 80 acres of cotton. Expected net returns are $11,317.40 and standard deviation was $2166.60.

In addition to locating the point of maximum returns, sequential sampling provides sufficient information to approximate the efficiency frontier between expected net returns and standard deviation of net returns.[3]

SUMMARY AND CONCLUSIONS

Sampling by steepest ascent and sequential, uniform-grid sampling are presented as alternative approaches to locating optimum simulation solutions. Sampling by steepest ascent has been applied empirically but the possibility of converging on a local maximum is recognized as a weakness. Uniform-grid sampling provides broader, more extensive coverage of the net returns surface and minimizes the possibility of locating a relative rather than absolute maximum. The sequential nature of the suggested procedure permits evaluation of progress at each stage of the maximization procedure.

Both sampling by steepest ascent and sequential sampling were applied to the net returns surface generated by a simplified firm level computer simulation model. The maximum point located by steepest ascent sampling is $11,000.06. The sequential sampling procedure locates a maximum net returns point of $11,317.40. In addition, sequential sampling generates sufficient information to permit plotting of the efficiency frontier between expected net returns and standard deviation of net returns.

FOOTNOTES

1. Graduate Research Assistant and Associate Professor respectively, Department of Agricultural Economics, Oklahoma State University, Stillwater, Oklahoma. Journal Article 2057 of the Agricultural Experiment Station, Oklahoma State University, Stillwater, Oklahoma. This research was supported in part by Grant 14-01-0001-1539 from the Office of Water Resources Research.

2. Schechter and Heady (5) recently reported results of a response surface analysis of agricultural policy choices within a simulation framework. They utilized a factorial experiment designed to allow estimation of polynomial equations to approximate the relevant response surfaces. Then, extreme points were determined by differentiating with respect to the relevant factors.

3. The authors acknowledge that quadratic programming could be used to define the efficiency frontier in this example. However, it does not appear that the quadratic programming approach would be an efficient method of defining an "optimum rule" in cases where one is altering other types of decision variables, such as borrowing strategies.

REFERENCES

1. Box, G.E.P., "The Exploration and Exploitation of Response Surfaces," *Biometrics,* Vol. 10, 1954.

2. Brooks, Samuel H., "A Comparison of Maximum-Seeking Methods," *Operation Research,* Vol. 7, 1959.

3. Maass, Arthur, *et al., Design of Water Resource Systems,* Harvard University Press, Cambridge, Massachusetts, 1962.

4. Mapp, Jr., Harry P., "An Application of Sampling by Steepest Ascent and Sequential Sampling to Obtain Optimum Simulation Solutions," unpublished manuscript, Department of Agricultural Economics, Oklahoma State University, Stillwater, December, 1969.

5. Schechter, Mordechai, and Earl O. Heady, "Response Surface Analysis and Simulation Models in Policy Choices," *American Journal of Agricultural Economics,* Vol. 52, No. 1, February, 1970.

6. Zusman, Pinkas, and Amotz Amaid, "Simulation: A Tool for Farm Planning Under Conditions of Weather Uncertainty," *Journal of Farm Economics,* Vol. 47, No. 3, August, 1965.