韓國軍事運營分析學會誌
第 25 卷, 第 2 號, 1999. 12. 31.

Analysis of the Methodology for Linear Programming Optimality Analysis using Metamodelling Techniques

Lee, Young Hae*, Jeong, Chan Seok *

Abstract

Metamodels using response surface methodology (RSM) are used for the optimality analysis of linear programming (LP). They have the form of a simple polynomial, and predict the optimal objective function value of an LP for various levels of the constraints. The metamodelling techniques for optimality analysis of LP can be applied to large-scale LP models. What is needed is some large-scale application of the techniques to verify how accurate they are. In this paper, we plan to use the large scale LP model, strategic transport optimal routing model (STORM). The developed metamodels of the large scale LP can provide some useful information.

* Department of Industrial Engineering, Hanyang University

1. Introduction

The light-hand-side (RHS) vectors of an LP may be changed for a variety of reasons. We may want to conduct a sensitivity analysis to check the preciseness of the RHS vectors, and whether or not it matters if the RHS vectors are perturbed. Also, we may want to update the LP when additional (reduced) resources become available (unavailable). In this case, optimality analysis comes in to play. Optimality analysis is performed to determine the effect on the optimal solution when the RHS vector (or the objective function coefficient) is changed. We also refine the geographical areas and use the Hot Start technique for making the large-scale LP runs required by the experimental design. The Hot Start method is very helpful for this research because we could save time and effort on the computer experiments. We use the converter programs to support the Hot Start method and create the Hot Start files by that program's successive use. The Hot Start method gives us good efficiency for reducing running time of large scale LP model.

Optimality analysis generally involves multiple critical regions with different optimal bases. This differs from "post-optimality analysis" and "sensitivity analysis" , since those analyses deal with only one critical region. Optimality analysis of large scale LP across multiple critical regions is more difficult than situations dealing with only one critical region. Thus, identifying which critical region contains a particular RHS vector creates a burden for the analyst.

Before accomplishing the optimality analysis of a large scale LP model, an alternative programming and solution technique which is called the Hot Start method [7] is developed. We refine the geographical areas and use the Hot Start technique for making the large-scale LP runs required by the experimental design. If we have a large LP, we may want to know the response results by various changing of RHS (or objective function coefficient) vectors without re-running the whole LP.

In this paper, we show an overall procedure of optimality analysis for large scale LP. It is as follows: (1) selecting several interested parts of RHS (or objective function coefficient) vectors in a large scale LP, (2) designing a metamodel associated with the selected RHS (or objective function coefficient) vectors (For example 2^2 factorial design needs four iterations), (3) changing the RHS (or objective function coefficient) vector values (10% up or 10% down), (4) running the large LP with changed RHS vectors and obtaining the object function values of the LP (In this procedure we need the Hot Start method for reducing run time), (5) creating a metamodel (first order polynomial) by least squares regression, (6) validation check for a metamodel.

This research employs the application of the methodology developed by Johnson et al. [6] to a large scale LP. The methodology of Johnson et al. utilizes response surface methodology. It basically accomplishes optimality analysis for LP models by developing first order metamodels which describe the relationships between the optimal objective function value and the RHS (or objective function coefficient) vectors of the LP [6].

We intend to create metamodels with only first order polynomials unless the higher order polynomials (such as second order) are needed. We basically apply 2^{k-p} factional and 2^k full fractional designs to create metamodels by least square regression. The problem is to verify whether or not this technique is valid for large scale LP. For the evaluation of metamodels, two primary measures are used: mean squared error (MSE) and mean absolute percentage error (MAPE). Since metamodels are really time and effort savers, the analyst will be able to observe the response of the optimal objective function value of a particular LP very easily and efficiently when the levels of the constraints of this LP are changed.

2. Optimality Analysis using "Hot Start" and Metamodelling Techniques

2.1 The methodology of "Hot Start" for reducing solving time

A large scale linear program contains many variables and equations and poses computational challenges, requiring a long time to solve. A "Hot Start" is a technique designed to take advantage of a "good" starting solution [7] when solving an LP. The sequence of "Hot Start" when producing solutions for use in metamodel development depends on the number of changing factor levels in the experimental design.

An algorithm for the "Hot Start" method is as follows:

step 1: Run original LP and get the value of optimal variables.

step 2: Change the value of optimal variables to an attachable program ("Hot Start" program) for next LP.

step 3: Attach "Hot Start" program to main LP.

step 4: Change the RHS vectors of LP and run this LP with the attachable program as a staring point of variables.

2.2 Metamodelling Techniques in Optimality Analysis

RSM is used to develop a methodology for optimality analysis of LP. Using these techniques, metamodels are developed to predict the optimal objective function value of an LP for various levels of the constraints. These metamodels are valid over multiple critical regions, eliminating the usual requirement of determining which critical region contains the RHS vector of interest. The metamodels are used to determine the responsiveness of the optimal objective function value to changes in the RHS vector while illuminating key relationships between the objective function value and the elements of the RHS vector. In some cases, the metamodels can actually be used as a surrogate model for the entire LP model. The metamodels are tested by comparing the predictions to the optimal solutions obtained by solving the linear programming model.

2.2.1 Two-Level Factorial Designs

Factorial designs used in this research involve several factors where it is desired to investigate the joint effects of the factors on a response variable. As the number of factors in a 2^k factorial design increase, the number of runs required for the complete replicate of the design rapidly outgrows the resources of most experimenters. If the experimenter can reasonably assume that certain high-order interactions are negligible, then information on the main effects and low-order interactions may be obtained by running only a fraction of the complete factorial experiment.

A design containing 2^{k-p} runs is called a fraction of the 2^k design or, more simply, a 2^{k-p} fractional design. These designs require the selection of p independent design generations. A reasonable criterion is to select the generator such that the resulting 2^{k-p} design has the highest possible resolution.

After establishing the metamodel designs, the next step is to code the level of the constraints. Coding is achieved using a simple transformation given by

$$Z_i = \frac{b_i - b_{i,0}}{S_i} \qquad (1)$$

where b_i is the actual numerical level of the i^{th} constraint. In this study, since we use two-level factorial designs, we can say that b_i has $b_{i,max}$ and $b_{i,min}$ indicating the high and the low levels of the constraints. Here $b_{i,o}$ is the midpoint between $b_{i,max}$ and $b_{i,min}$ .If S_i is taken to be $(b_{i,max} - b_{i,min}) / 2$, then $b_{i,max}$ and $b_{i,min}$ are mapped to 1 and -1, respectively [6].

The experimental designs used to develop the metamodels of interest in this research were determined according to the number of factors they were dealing with.

2.2.2 Least Squares Regression

The experimental designs are selected so that the experimental error variance is minimized. We obtain our response by solving a large scale LP for each design point. In this case, no experimental error is associated with the solution from the LP model. Since there is no experimental error, the purpose is to minimize the misspecification bias.

In this research, we approximate the LP model with a simple model. To achieve this, we apply least squares regression to the data obtained from the experimental design phase of the methodology of Johnson, et al. Let denote the optimal objective function value as a function of the factors. Least squares regression develops an initial metamodel approximating with a first order or second order polynomial since those types of polynomials are able to define some concave surfaces such as hyperplanes [6].

We can assume the functional relationship of the following form :

$$Y = Z\beta + \varepsilon \qquad (2)$$

The experimental value is replaced by Z which is a matrix of coded values (1 or -1 obtained be Equation (1)). Y is an $m \times 1$ vector of known response values (in this case optimal objective function values obtained form STORM), Z is a $m \times p$ (where $p = n + 1$) matrix of the coded levels of the variables at the experimental design point, is a $p \times 1$ vector of coefficients, is an $m \times 1$ vector of the random error, m is the number of data points, and p is the number of variables in the assumed function.

$$\hat{\beta} = (Z'Z)^{-1}Z'Y \tag{3}$$

and the fitted regression model is

$$\hat{Y} = Z\hat{\beta} \tag{4}$$

3. Application and Result

In this research, STORM was chosen as a large scale LP for the application of RSM. The Strategic Transport Optimal Routing Model (STORM) is based on a model built by Barton and Guiriaer (1967) of Lockheed to analyze the peacetime employment of the new C-5 cargo plane. STORM was developed at the Air Mobility Command (AMC) to assist in a major study of the entire scheduled cargo system that must provide two main types of service to its overseas customers. The first is to provide sufficient cargo capacity for a given period of time (usually for one month) to meet all demands for cargo movement between the pairs of bases in the system. This cargo capacity is known as the cargo requirement. The second is to provide a minimum number of flights per month between certain cities. This number is called the frequency requirement. The basic purpose of STORM is to select the mix of routes and aircraft that will meet the monthly cargo and frequency requirements of AMC while minimizing the cost of cargo handling, military aircraft operations, and commercial aircraft leasing [8].

The set of routes, maximum payload, and total flying hours for each type of plane are the main resources available. Using this information, STORM constructs a feasible cargo movement plan. A route is the sequence of legs to be flown by a single aircraft. Each aircraft's maximum payload is an average based on the fuel loads and types of palletized cargo that are generally carried on the planned missions.

The flying hour limit for military aircraft is derived from the Air Force flying program which is necessary to maintain proficiency in worldwide operations and to train crews. Versions of STORM for UNIX Workstations have been developed using the general algebraic modeling system (GAMS) modeling language [1] which makes data management and programming very easy. GAMS also allows the analyst to modify the model quickly for specific analyses, investigating specific questions, or enforcing operational considerations locally.

This section presents the comparison between Hot Start results and non- Hot Start results and also shows the numerical results and their interpretations. In this research, six metamodels of STORM are developed to examine the applicability of the methodology of Johnson et al. which describes the relationship between the levels of the constraints and the objective function value for an LP, and predict the optimal objective function value for the LP given specific levels for the constraints. The level means the coded form of the factors, which are denoted by Z_i, $i = 1, \ldots, n$ in the following tables. The test metamodels are developed by changing the RHS and unit cost vectors by 10 percent (= ±0.1) and considering the five areas of AMC s channel system .

The STORM model uses the preceding information to formulate a linear programming model which can best be described as follows:

MINIMIZE : Total cost incurred for aircraft operating hours and cargo handling.

SUBJECT TO:

1. Meeting as many of the cargo movement requirements as possible.
2. Meeting all requirements for service frequency.
3. Operating within each aircraft type's flying hour and payload limits. Operating within each base's limit on monthly sorties.

We constructed six metamodels. In Metamodel 1, we simply choose two factors, tonnage requirement (TREQ) and frequency requirement (FREQ) because we want to start this research with a simple design.

[Table 1] "Hot Start" comparison table for metamodel 1

Metamodel 1					
Treatment (Seq:Cha)	TREQ	FREQ	Non-"Hot Start"	"Hot Start"	Reduced ratio (%)
1(1:2)	Down 10%	down 10%	7748	7042	9.11
2(2:1)	down 10%	up 10%	8667	7326	15.47
3(4:1)	up 10%	down 10%	10167	7400	27.21
4(3:1)	up 10%	up 10%	7476	6593	11.81
(Seq:Cha) : (the number of sequence : the number of factor changes)					

Metamodel 1 : minimized-cost values for 4 treatment

Treatment	TREQ	FREQ	Min. Cost (dollars)
1	-1	-1	33644710.27
2	-1	1	34692518.12
3	1	-1	40567833.21
4	1	1	41351256.59

By comparing, we learn that reducing the solving time is possible by the "Hot Start" method and the efficiency of the "Hot Start" method for saving time is about 15.9% in Metamodel 1. We expect this method is helpful for models with more factor treatments.

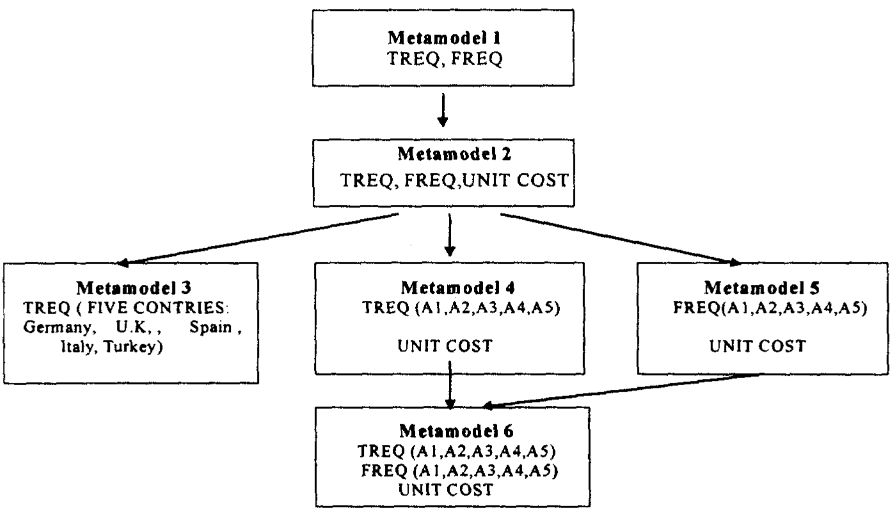
After analyzing Metamodel 1, we changed one more factor, the unit cost vector for Metamodel 2 which shows us the new relationship including the new factor. In Metamodels 3, 4, 5, we separate the factors of Metamodel 2 because we want to analyze each factor' s geographical relationship with the objective function value in detail. In Metamodel 6, we integrated metamodel 4 and 5. It looks like metamodel 2 but the two factors in Metamodel 2 ,TREG and FREQ are separated into five areas. This makes it possible to analyze the geographical relationships of these two factors with the objective function value and unit cost vector. Metamodel 4 shows the relationships among the total cost (objective function value), TREQ and unit cost vectors of five areas which are shown in Table 2.

[Table 2] Five Areas in the AMC Channel System

Area	Region
1	North America including U.S.A. territories.
2	South America
3	Europe and North Africa.
4	East Asia.
5	Central Pacific and Middle East.

The Al, A2, A3, A4 and A5 vectors for TREQ which correspond to areas 1, 2, 3, 4, and 5 were simply derived from the TREQ RHS vector by separating the elements of the TREQ vector into the related areas. These values are changed up and down 10%. We use a 2^{6-2} fractional factorial design for this metamodel. Metamodel 5 shows the relationships among the total cost (objective function value), FREQ and unit cost vectors for the five areas. We also construct a 2^{6-2} fractional factorial design for this metamodel. Metamodel 6 is a combination of Metamodel 4 and Metamodel 5 and it has a 2^{11-6} fractional factorial design. In Figure 1 the six metamodels and their relationship are presented.

[Figure 1] The six metamodels and the research directions



[Table 3] Factorial Designs for metamodel

Metamodel	The component of Factors	Factorial Design
1	2 factors (TREQ, FREQ)	2^2 Factorial Design
2	3 factors (TREQ, FREQ, UNIT COST)	2^3 Factorial Design
3	5 factors (TREQ:Germany, U.K, Spain, Italy, Turkey)	2_V^{5-1} Fractional Factorial Design
4	6 factors (TREQ: A1, A2, A3, A4, A5 and UNIT COST)	2_{IV}^{6-2} Fractional Factorial Design
5	6 factors (FREQ: A1, A2, A3, A4, A5 and UNIT COST)	2_{IV}^{6-2} Fractional Factorial Design
6	11 factors (TREQ: A1, A2, A3, A4, A5, FREQ: A1, A2, A3, A4, A5 and UNIT COST)	2_{IV}^{11-6} Fractional Factorial Design

[Table 4] Regression Polynomials of Metamodels

Metamodel	The component of Factors	Regression Polynomials
1	2 factors (TREQ, FREQ)	$Y = 37564080 + 3395465\,Z_1 + 457808\,Z_2$ $R^2 = 0.9996$
2	3 factors (TREQ, FREQ, UNIT COST)	$Y = 37563857 + 3395306\,Z_1 + 457768\,Z_2 + 3384940\,Z_3$ $R^2 = 0.9958$
3	5 factors (Germany, U.K, Spain Italy, Turkey)	$Y = 37521413 + 256260\,Z_1 + 43758\,Z_2 + 25919\,Z_3 + 14060\,Z_4 + 76374\,Z_5$ $R^2 = 0.9989$
4	6 factors (TREQ: A1, A2, A3, A4, A5 and UNIT COST)	$Y = 37762620 + 2427846\,Z_1 + 54011\,Z_2 + 651718\,Z_3 + 174877\,Z_4 + 91069Z_5 + 3399950\,Z_6$ $R^2 = 0.9947$
5	6 factors (FREQ: A1, A2, A3, A4, A5 and UNIT COST)	$Y = 37595793 + 109463\,Z_1 + 37153\,Z_2 + 172255\,Z_3 + 89285\,Z_4 + 29304\,Z_5 + 3380357\,Z_6$ $R^2 = 0.9998$
6	11 factors (TREQ: A1, A2, A3, A4, A5, FREQ: A1, A2, A3, A4, A5 and UNIT COST)	$Y = 37901155 + 2407677\,Z_1 + 53079\,Z_2 + 662569\,Z_3 + 175171\,Z_4 + 63506\,Z_5 + 104667\,Z_6 + 28138\,Z_7 + 184575\,Z_8 + 85820\,Z_9 + 43623\,Z_{10} + 3417187Z_{11}$ $R^2 = 0.9927$

By using these methodologies, we can analyze the accuracy of the metamodels and to describe the key relationships among the optimal objective function value, the objective function coefficient vectors and the RHS vectors of interest (shown in Table 4). It is possible to predict the response of the optimal objective function value easily and efficiently when the levels of the factors are changed.

4. Measuring the Performance of the Metamodels

An accurate metamodel means a metamodel with a small residual. For example Metamodel 1 is validated by a 2^2 full factorial design where the factors were coded at the (+0.5, -0.5) level (Perturbed by 5%). The test and validation design points on a coordinate axis. Note that the point (0,0) is the center point, for both designs. The predictions obtained from the regression metamodel are compared to the true optimal objective function values of the validation design by using some measures. In this research, two primary measures were used to evaluate the performance of the metamodels. The first is mean squared error (MSE) and the second is mean absolute percentage error (MAPE) [6].

M ean squared error (MSE) is a measure of accuracy computed by squaring the individual error for each item in the data set, and then finding the average or mean value of the sum of these squares. The mean squared error gives greater weight to large errors than to small errors since the errors are squared before being summed. MSE is defined by

$$MSE = \frac{\sum_{i=0}^{m} (\hat{Y} - Y_i^v)^2}{m} \qquad (5)$$

where $\hat{Y}_i$ is the predicted value of the objective function at the i^{th} validation point (Z = 0.5), m is the number of validation points and Y_i^v is the true value of the objective function at the i^{th} validation point (level changed up and down 5%).

Mean absolute percentage error (MAPE) is the mean or average of the sum of all of the percentage errors for a given data set taken without regard to sign. Thus, their absolute values are summed and the average computed.

MAPE is given as follows:

$$MAPE=\frac{1}{m}\sum_{i=1}^{m}\left|\frac{\hat{Y}-Y_i^v}{Y_i^v}\right|\times 100 \tag{6}$$

[Table 5] Comparison with validation design for metamodel 1

Treatment	A	B	Y^v	$\hat{Y}$ (Z = ±0.5)	$\lvert\hat{Y}-Y^v\rvert$
1	-0.5	-0.5	35570398.00	35637443.50	67045.50
2	-0.5	0.5	36029101.77	36095251.50	66149.73
3	0.5	-0.5	39006052.42	39032908.50	26856.08
4	0.5	0.5	39423802.92	39490716.50	66913.58
MSE			MAPE		
3517390517.68			0.152666889		

[Table 6] Comparison with validation design for metamodels

Metamodel	Design	MSE	MAPE
1	2^2 Factorial Design	3517E6	0.153
2	2^3 Factorial Design	9460E6	0.151
3	2_V^{5-1} Fractional Factorial Design	592E6	0.059
4	2_{IV}^{6-2} Fractional Factorial Design	65254E6	0.602
5	2_{IV}^{6-2} Fractional Factorial Design	3701E6	0.137
6	2_{IV}^{11-6} Fractional Factorial Design	115909E6	0.830

We showed the validation design in Table 5 and 6. Metamodels are validated by factorial design where the factors were coded at the (+0.5, -0.5) level. The MSE and MAPE value indicate that metamodels are very accurate models.

5. Conclusion

The "Hot Start" method was very helpful for this research because we could save time and effort on the computer experiments. The "Hot Start" method gave us good efficiency for reducing running time of large scale linear programming model.

The metamodelling methodology for the optimality analysis of LP works very well for the large scale linear programming models. The metamodels described the key relationships between the optimal objective function value and the RHS vectors. The response surface of the optimal objective function value is a relatively flat surface. For that reason, the optimal objective function value can be estimated by a simple polynomial with remarkable accuracy.

Metamodels have the form of a simple polynomial and predict the optimal objective function value of an LP for various levels of the constraints. Since metamodelling methodologies are really time and effort savers, it is possible to predict the response of the optimal objective function value easily and efficiently when the levels of the factors are changed.

Reference

[1] A. Brooke, D. Kendrick and A. Meeraus, Release 2.25 Gams a User Guide, The Scientific Press, California, 1992.

[2] D. C. Montgomery, Design and Analysis of Experiments, 3rd. ed., John Wiley and Sons Inc., New York, 1991.

[3] G. Boxand, N. R. Draper, Empirical Model-Building and Response Surfaces, John Wiley and Sons Inc., New York, 1987.

[4] J. C. Davis, Statistics on Data Analysis in Geology, 2nd. ed., John Wiley and Sons Inc., New York, 1986.

[5] J. Hadley, Linear Programming 2nd. ed., Addison-Wesley Publishing Company, Massachusetts, 1963.

[6] K. E. Johnson, K. W. Bauer Jr., J. T. Moore and M. Grant, "Metamodelling Techniques in Multidimensional Optimality Analysis for Linear Programming, Mathematical and Computer Modelling, Vol23, pp.45-60, 1996.

[7] Marsten, Subramanian, Saltzman, Lustig and Shanno, "Interior Point Methods for Linear Programming : Just call Newton, Lagrange, Fiacco and McCormick," Report to AFIT/Department of Operational Science, Interface 20, pp.4-5, 1995.

[8] M. B. Carter, J. Litko, K. Ware and A. Whisman, "Decision Support for the Optimality Command Scheduled Cargo System," Technical Report, HQ MAC/XPY

Military Airlift Command, Scott AFB, Command Analysis Group, 1992.

[9] M. S. Bazaraa, J. J. Jarvis and H. D. Sherali, Linear Programming and Network Flows 2nd. ed., John Wiley and Sons Inc., New York, 1990.

[10] R. H. Myers, D. C. Montgomery, Response Surface Methodology, John Wiley and Sons Inc., New York, 1995.

[99년 10월 15일 접수, 99년 11월 12일 최종수정]