copyright 1997 Bruce A. McCarl and Thomas H. Spreen

CHAPTER VI: TOWARD PROPER MODELING

There is considerable role for judgment when modeling and developing data. The applied modeler must make assumptions regarding the variables, constraints, and coefficients. These assumptions determine model performance and usefulness.

In this chapter the identification of structural components and the development of data are discussed. The material presented here is reinforced by material in subsequent chapters. References are made to this later material, and readers may wish to consult it for more detailed explanations.

Before beginning this section, the authors must acknowledge their debt to Heady and Candler's "Setting Up Linear Programming Models" chapter and conversations with Wilfred Candler.

6.1 Structural Component Identification

The LP problem can be expressed as

$$\begin{aligned} \text{Max} \quad & CX & & \\ \text{s.t.} \quad & AX & \leq & \; b \\ & X & \geq & \; 0 \end{aligned}$$

In order to formulate an applied LP problem, one must identify the constraints, variables and relevant numerical parameter values.

6.1.1 Development of Model Constraints

Heady and Candler categorize LP constraints as technical, institutional, and subjective. Constraints also arise because of convenience or model formulation requirements. <u>Technical constraints</u> depict limited resources, intermediate products, or contractual requirements. Technical constraints also express complementary, supplementary, and competitive relationships among variables. Collectively, the technical constraints define the production possibilities and provide links between variables. <u>Institutional</u>

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

constraints reflect external regulations imposed on the problem. Examples include credit limits or farm program participation requirements. Subjective constraints are imposed by the decision maker or modeler. These might include a hired labor limitation based on the decision maker's willingness to supervise labor. Convenience constraints facilitate model interpretation and may be included to sum items of interest. Model formulation constraints aid in problem depiction. These include constraints used in conjunction with approximations. Within and across these groupings, constraints can take on a number of different forms. A more extensive definition of these forms is presented in the LP Modeling Summary chapter.

Generally, the constraints included should meaningfully limit the decision variables. The modeler should begin by defining constraint relations for those production resources and commitments which limit production or are likely to do so. This involves consideration of the timing of resource availability. Often, problems covering seasonal production or utilizing seasonally fluctuating resources will contain time desegregated constraints. Heady and Candler argue that multiple constraints are needed to depict availability of a resource whenever the marginal rate of factor substitution between resource usage in different time periods does not equal one. Constraints must be developed so that the resources available within a particular constraint are freely substitutable. Cases of imperfect substitution will require multiple constraints.

Two other points should be made regarding constraint definition. First, an LP solution will include no more variables at a nonzero level than the number of constraints (including the number of upper and lower bounds). Thus, the number of constraints directly influences the number of nonzero variables in the optimal solution. However, one should not simply define additional constraints as: 1) this usually results in additional nonzero slack variables without substantially altering the solution; and 2) one must not impose nonsensical constraints.

Second, subjective constraints should not be imposed before determining their necessity. Often,

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

subjective constraints "correct" model deficiencies. But the cause of these deficiencies is frequently missing either technical constraints or omitted variables. For example, models often yield excessively specialized solutions which force variables into the solution. This is often combated by imposing "flexibility" constraints as suggested by Day (1963), or discussed in Sengupta and Sfeir. Often, however, the real deficiency may be the depiction of the time availability of resources (Baker and McCarl). In such a case, the subjective constraints give an inadequate model a "nominal" appearance of reality, but are actually causing the "right" solution to be observed for the wrong reason.

6.1.2 Avoiding Improper Constraint Specifications

LP model constraints have higher precedence than the objective function. The first major effort by any LP solver is the discovery of a feasible solution. The solver then optimizes within the feasible region. This has several implications for identification and specification of constraints.

First, the modeler must question whether a constraint should be established so it always restricts the values of the decision variables. Often, it may be desirable to relax a constraint allowing resource purchases if the value of a resource becomes excessively high.

Second, modelers should be careful in the usage of minimum requirement constraints (e.g.,$X_1+X_2 \exists 10$). Minimum requirements must be met before profit seeking production can proceed. Often purchase variables should be entered to allow buying items to meet the requirements.

Third, judicious use should be made of equality constraints. Modelers should use the weakest form of a constraint possible. Consider the following example:

$$
\begin{array}{rccccc}
\text{Max} & 3X & + & 2Y & & \\
\text{s.t.} & X & - & Y & ? & 0 \\
 & X & & & \leq & 10 \\
 & & & Y & \leq & 15
\end{array}
$$

Where ? is the constraint type (either = or #), X depicts sales and Y production? Further, suppose we have made a mistake and have specified the cost of production as a revenue item (i.e., the +2Y should be

-2Y). Now, if the relation is an equality, then the optimal solution is X = Y = 10 (see file SIXEQ), and we do not discover the error (although the dual variable on the first constraint is -2). On the other hand, if the relation is # then we would produce Y = 15 units while selling only X = 10 units (see file SIXLT). Thus, the weaker inequality form of the constraint allows an unrealistic production pattern indicating that something is wrong with the model.

6.1.3 Variable Identification

LP variables are the unknowns of the problem. Variables are included for either technical, accounting or convenience reasons. Technical variables change value in response to the objective function and constraints. Convenience variables may not always respond to the objective function. Rather, they may be constrained at certain levels. These might include variables representing the number of acres of land used for houses and buildings. Accounting variables facilitate solution summarization and model use.

It is critically important that the technical variables logically respond to the objective function within the range of values imposed by the constraints. For example, one could setup a farm problem with variables responding to an objective of minimizing soil erosion. However, farmers choosing acreage may not primarily try to minimize erosion; most farmers are also profit oriented.

Many types of technical variables are possible. A taxonomy is discussed in the LP Modeling Summary chapter.

Variables must be in consistent units. Actually, there are no strict LP requirements on the variable units. However, the intersection of the variable and constraint units impose requirements on the a_{ij}'s as discussed below. Now, when can multiple variables be handled as one variable and when can't they? There are several cases when multiple variables must be defined:

(a) When more than one process can be used to produce the same output using different resource mixes; e.g., the production of an item using either of two different machines.

(b) When different processes produce different outputs using common resources; i.e., one can

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

use essentially the same resources to produce either 2 x 4 or 4 x 4 sawn lumber.

(c) When products can be used in several ways; e.g., selling chickens that can be quartered or halved.

Collectively, different variables should be used where their coefficients differ (i.e., the objective function or a_{ij} coefficients differ across production possibilities). However, the coefficients should not be strictly proportional (i.e., one variable having twice the objective function value of another while using twice the resources).

Criteria may also be developed where two variables may be treated as one. The simplest case occurs when the coefficients of one variable are simple multiples of another ($a_{ij} = Ka_{im}$ and $c_j = Kc_m$). The second case occurs when one variable uniquely determines another; i.e., when n units of the first variable always implies exactly m units of the second.

6.1.4 Objective Function

Once the variables and constraints have been delineated, then the objective function must be specified. The variables and constraints jointly define the feasible region. However, the objective function identifies the "optimal" point. Thus, even with the proper variables and constraints, the solution is only as good as the objective function. Ordinarily, the first objective function specification is inadequate. Most situations do not involve strict profit maximization, but also may involve such things as risk avoidance or labor/leisure tradeoffs. Multiple objective models are discussed in the multi-objective and risk chapters. Also, ranging analysis can be used to discover whether the solution will change with alterations in the objective function.

6.1.5 Development of Model Structure

Model definition is an iterative process. Consider a simple example where a profit maximizing firm produces four crops subject to land and labor limitations. Suppose that the crops are grown at different times of the year. Crop 1 is planted in the spring and harvested in the summer; crops 2 and 3 are planted in the spring and harvested in the fall; and crop 4 is planted following crop 1 and is harvested in the fall.

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

The first step in developing a model is to lay out a table with potential variables across the top and constraints/objective function down the side. In this case we start with the layout in Table 6.1 where the variables are crop acreages and the constraints are land and labor availability. We then begin to define coefficients. Suppose c_i gives the gross profit margins for crop i. Simultaneously, land use coefficients and the land endowment (L) are entered. However, the land constraint only has entries for crops 1, 2 and 3, as crop 4 uses the same land as crop 1. Thus, a single land constraint restricts land use on an annual basis. We also need a constraint which links land use by crop 4 to the land use by crop 1. Thus, our formulation is altered as shown in Table 6.2, where the second constraint imposes this linkage.

Now we turn our attention to labor. In this problem, labor is not fully substitutable between all periods of the year, i.e., the elasticity of substitution criterion is not satisfied. Thus, we must develop time-specific labor constraints for spring, summer and fall. The resultant model is shown in Table 6.3. Subsequently, we would fill in the exact labor coefficients; i.e.; the d's and right hand sides.

This iterative process shows how one might go about defining the rows and columns. In addition, one could further disaggregate the activities to allow for different timing possibilities. For example, if Crop 1 produced different yields in different spring planting periods, then additional variables and constraints would need to be defined.

6.2 The Proper Usage of Variables and Constraints

Students often have difficulties with the definition of variables and constraints. This section is intended to provide insight by presenting a number of proper and improper examples.

The applied LP modeler needs to recognize three concepts when forming constraints and variables. First, the coefficients associated with a variable reflect a simultaneous set of commitments which must occur when a variable is nonzero. All the resources used by a variable must be used to produce its output. Thus, if a variable depicts cattle and calf production using inputs of land, labor, and feed; then the model will simulta-

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

neously commit land, labor, and feed in order to get the simultaneous outputs - cattle and calves. One cannot obtain calves without obtaining cattle nor can one obtain cattle and calves without using land, labor, and feed.

Second, the choice is always modeled across variables, never within a variable. For example, suppose there are two ways of producing cattle and calves. These production alternatives would be depicted by two variables, each representing a simultaneous event. The model would reflect choice regarding cattle/calf production within the constraints. These choices do not have to be mutually exclusive; the model may include complementary relationships between variables as well as substitution relationships (i.e. the constraint X-Y=0 makes X and Y complementary).

Third, resources within a constraint are assumed to be homogeneous commodities. Suppose there is a single constraint for calves with the calves being produced by two variables. In turn, suppose calves may be used in two feeding alternatives. In such a case the calves are treated as perfect substitutes in the feeding processes regardless of how they were produced.

While obvious, it is surprising the number of times there are difficulties with these topics (even with experienced modelers). Thus, we will present cases wherein such difficulties could be encountered.

6.2.1 Improper and Proper Handling of Joint Products

Joint products are common in LP formulations. For purposes of illustration, we adopt the following simplified example.[1] Suppose a chicken is purchased and cut up into four component parts - breasts, legs, necks, and giblets - following the breakdown data in Table 6.4. Also, assume that each chicken weighs three pounds and that there are 1,500 chickens available.

Now suppose that we formulate a profit maximizing LP model. Such a LP model would involve variables for cutting the chickens along with variables for parts sale. Two alternative formulations will be

[1] The example is a disassembly problem, which is discussed in the More LP Modeling chapter. Readers having difficulty with its basic structure may wish to study that section.

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

presented: one proper and one improper. These formulations are shown in Table 6.5 and are labeled Formulation 6.5(a) and Formulation 6.5(b).

The models maximize the value of the chicken parts produced. A constraint is needed which limits the number of chickens cut up to the number of chickens available. In both formulations the coefficient 1/3 in the last constraint transforms the chickens disassembled into pounds of chicken rather than whole chickens, so the units of the first variable are in pounds of chicken cut up. The next four variables are the quantities (pounds) of parts sold. In formulation 6.5(a) the constraint labeled Balance restrains the amount sold to the quantity of chicken cut. The formulation maximizes the value of chicken parts sold. The decision is constrained by the quantity of chicken disassembled and chickens available. In Formulation 6.5(b), the objective function and last constraint are the same. However, there are balances for each part.

Now which formulation is improper? Suppose we examine what happens when Y equals one (i.e., that we have acquired one pound of chicken for cutting up). Formulation 6.5(a) implies that variable X_1 could equal two if the other variables were set to zero. Thus, from one pound of chicken two pounds of chicken breasts could be sold. This is not possible. Similarly, 3.43 pounds of legs (X_2) could be sold per pound of chicken. 10 pounds of necks (X_3) or 20 pounds of giblets (X_4) could be sold. In formulation 6.5(b), the acquisition of one pound of chicken would allow only .5 pounds of breasts, .35 pounds of legs and thighs, .1 pounds of necks, and .05 pounds of giblets.

Clearly, formulation 6.5(b) is the proper formulation. Formulation 6.5(a) depicts improper representation of the joint products allowing an improper choice between the use of all the chicken meat among any of the four component parts. In fact, its optimal solution indicates 90,000 lbs of Giblets can be sold from the 4,500 lbs of chicken (see the file SIX5A on the disk). The component parts are a joint product that should simultaneously occur in the model.

6.2.2 Alternatives for the Use of a Product

Errors also occur when modeling different ways products can be used. Suppose we introduce the

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

option of selling chicken parts or deboning the parts then selling chicken meat. Assume that there are no additional resources involved, and that the meat yields are those in Table 6.4. Again, we will illustrate proper modeling with a right and a wrong formulation.

The first model Table 6.6(a) has three new variables and constraints. The three new variables sell meat at $1.20. The three new constraints balance meat yields with sale. Thus, the coefficient in the breast quarter meat row is the meat yielded when breast quarter is deboned (the breast quarter poundage per chicken times the percentage of meat in a breast quarter).

Formulation 6.6(b) adds four variables and one row. The first three variables transform each of the products into meat with the fourth selling the resultant meat. The new constant balances the amount of meat yielded with the amount of meat sold.

Now which formulation is proper? Let us examine the implications of setting the variable Y equal to 1 in Table 6.6(a). As in our earlier discussion the solution would have, variables X_1 through X_4 at a nonzero level. However, in this formulation M_1, M_2, and M_3 would also be nonzero. Since both the X variables and the M variables are nonzero, the chicken is sold twice. In Table 6.6(b), when Y is set to one, then either X_1 or M_1 can be set to .5, but not both (in fact, the sum of $X_1 + M_1$ can be no greater than 0.5). Thus, the chicken parts can only be sold once.

Formulation 6.6(b) is proper. Formulation 6.6(a) contains an improper joint product specification, as it simultaneously allocates products to mutually exclusive uses. Formulation 6.6(b) restricts any single part to one use.

6.2.3 Improper Specification of Multiple Factor Relationships

Factor usage is often subject to misspecification in terms of multiple factor relationships. This case is illustrated with yet another extension of the chicken example. We now wish to allow sales of a mixed quarter pack which is composed of an arbitrary combination of breast and leg quarters. Let us introduce two models. The first model has the same constraints as Formulation 6.6(b) but introduces new variables where the breast

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

and leg quarters are put into the mixed quarter package (Formulation 6.7(a)).

Formulation 6.7(b) involves three new variables and one new constraint. The first two variables are the poundage of breast and leg quarters utilized in the mixed packs. The third variable is total poundage of mixed quarter pack sold. The new constraint balances the total poundage of the mixed quarter packs sold with that produced.

Now the question again becomes which is right? Formulation 6.7(a) is improper; the formulation requires that in order to sell one pound of the mixed quarter pack, two pounds of quarters, one of each type, must be committed and leads to a solution where no packs are made (see the file SIX7A). In Formulation 6.7(b) the two sources of quarters are used as perfect substitutes in the quarter pack, permitting any proportion that maximizes profits. The optimal solution shows all leg quarters sold as mixed quarter packs. Formulation 6.7(a) illustrates a common improper specification - requiring that the factors to be used simultaneously when multiple factors may be traded off. One should not require simultaneous factor use unless it is always required. Multiple variables are required to depict factor usage tradeoffs.

6.2.4 Erroneous Imperfect Substitution

Resource substitution may also be incorrectly prevented. Consider a problem depicting regular and overtime labor. Suppose the basic product is chairs where: a) a chair requires 10 hours of labor of which an average of 3 hours comes from overtime labor; b) the firm has an endowment of 77 hours of regular labor at $10 per hour and up to 27 hours of overtime labor at $15 per hour. We again introduce two formulations.

In Formulation 6.8(a) the variables indicate the number of chairs to produce and sell, along with the amount of labor to acquire. The constraints give a balance between the chairs produced and sold; balances between the labor/quantities hired versus used; and limits on labor time available.

Model 6.8(b) is essentially the same, however, we have aggregated our labor use-hired balance so that there is no distinction made between the time when labor is used (regular or overtime).

Which formulation is right? This depends on the situation. Suppose that labor works with equal

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

efficiency in both time classes. Thus, one would be technically indifferent to the source of labor although economically the timing has different implications. Now let us examine the formulations by setting X_1 to one. In 6.8(a) the model hires both classes of labor. However, in 6.8(b) only regular time labor would be hired. In fact, in 6.8(a) the overtime limit is the binding constraints and not all regular time labor can be used and only nine chairs are made; whereas in 6.8(b) eleven chairs could be produced and all the labor is used. The second model is the correct one since it makes no technical differentiation between labor sources.

6.3 Simple Structural Checking

There are some simple yet powerful techniques for checking LP formulations. Two are discussed here another in Chapter 17.

6.3.1 Homogeneity of Units

There are several general requirements for coefficient units. Consider the LP problem:

$$\begin{array}{lccccc} \text{Max} & c_1X_1 & + & c_2X_2 & & \\ \text{s.t.} & a_{11}X_1 & + & a_{12}X_2 & \leq & b_1 \\ & a_{21}X_1 & + & a_{22}X_2 & \leq & b_2 \end{array}$$

Suppose that the objective function unit is dollars. Let the first row be a land constraint in acres. Let the second row be a labor constraint in the unit hours. Further, suppose that X_1 represents acres of wheat and X_2 number of beef animals.

What implications do these specifications have for the units within the model? Parameter c_1 must be the dollars earned per acre of wheat while c_2 must be the dollars earned per beef animal. Multiplying these two parameters by the solution values of X_1 and X_2 results in the unit dollars. In turn, a_{11} represents the acres of land used per acre of wheat. The parameter a_{12} would be the number of acres of land utilized per beef animal. The units of the right hand side (b_1) must be acres. The units of the parameters a_{21} and a_{22} would respectively be labor hours utilized per wheat acre and labor hours utilized per beef animal. The units of the right hand side (b_2) must be hours of labor.

This example gives a hint of several general statements about units. First, the numerator unit of each

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

coefficient in an equation must be the same and must equal the right-hand side unit. Thus, a_{11} is the acres of land used per acre of wheat, a_{12} is the acres of land used per beef animal and b_1 the acres of land available.

Similarly, the coefficients associated with any particular variable must have a common denominator unit, although the numerator will vary. Thus, c_1 is in the units dollars per acre of wheat, a_{11} is acres of land per acre of wheat, and a_{21} is the hours of labor per acre of wheat. In addition, note that the units of the decision variable X_1 are acres of wheat. The denominator unit of all coefficients within a column must be the same as the unit of the associated decision variable.

6.3.2 Numerical Model Analysis

Another possible type of model analysis involves numerical investigation of the model. Here, one simply mentally fixes variables at certain levels such as the level of 1, and then examines the relationship of this variable with other variables by examining the equations. Examples of this procedure are given in the proper usage section above.

Numerical debugging can also be carried out by making sure that units are proper, and it is possible to utilize all resources and produce all products. Finally, solvers such as OSL contain reduction procedures.

6.4 Data Development

Model specification requires data. The data need to be found, calculated, and checked for consistency. Data development usually takes more time than either model formulation or solution. However, this time is essential. Good solutions do not arise from bad data.

Data development involves a number of key considerations. These include time frame, uncertainty, data sources, consistency, calculation methods, and component specification.

6.4.1 Time Frame

Models must be established with a time frame in mind. The time frame defines the characteristics of the data used. The objective function, technical coefficient (a_{ij}'s) and right hand side data must be mutually

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

consistent. When the model depicts resource availability on an annual basis, then the objective function coefficients should represent the costs and revenues accruing during that year. A common misspecification involves an annual model containing investment activities with the full investment cost in the objective function.

Dynamic considerations may be relevant in the computation of objective function coefficients. It is crucial that the objective function coefficients be derived in a consistent manner. Returns today and returns in ten years should not be added together on an equal basis. Issues of dynamics and discounting must be considered as discussed in the Dynamic LP Chapter.

6.4.2 Uncertainty

The data developer must consider uncertainty. Coefficients will virtually never be known with certainty. For example, when variables involve transport of goods from one place to another, the transport costs are not entirely certain due to difficulties with pilferage, spoilage, adherence to shipping containers, and leakage. The modeler is forever facing decisions on how to incorporate data uncertainty. The risk programming chapter presents formal methods for incorporating uncertainty. However, many modelers use average values or conservative estimates.

6.4.3 Data Sources

Data may be developed through statistical estimation or deductive processes. Data for coefficient estimation can be from either cross-sectional or time series sources. Data may be developed using a case firm (or firms) approach where a deductive, economic engineering process is used to manufacture representative coefficient values. Data sources will vary by problem, and the modeler must apply ingenuity as well as problem-specific knowledge to develop consistent, reliable data.

6.4.4 Calculation Methods

Data can be calculated via economic engineering or via statistical methods. While these are only two extremes of a continuum of possibilities, we will discuss only these two. Economic engineering refers to

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

coefficient construction through a deductive approach. For example, suppose we compute the profit contribution of a variable by calculating the per unit yield times sale price less the per acre input usage times input price (i.e., if wheat production yields 40 bushels of wheat which sells for $5 per bushel and 20 bales of straw each worth $.50 while input usage is $30 worth for seed and 6 sacks of fertilizer, which cost $4 each; then, the objective function coefficient would be $156.)

At the other extreme, one could develop multiple observations from time series, cross-sectional or subjective sources and use averages, regression or other data summarization techniques. Such data might in turn be transformed using an economic engineering approach to generate relevant coefficients. For example, one might estimate a function statistically relating yield to fertilizer use and labor use. Then one might set a level of fertilizer use, calculate the yield, and use an economic engineering approach to develop the objective function coefficients.

6.4.5 Consistency

Coefficients in a model must be mutually consistent. The most common causes of inconsistency are dynamic inconsistencies and inconsistencies in coefficient units (e.g., a technical coefficient in hours and a right-hand side in thousands of hours). The homogeneity of units rules above must be followed.

6.4.6 Specification of Individual Components

LP problems require right hand side, objective function, and technical coefficient specification. There are comments that can be made pertinent to the specification of each.

6.4.6.1 Objective Function Coefficients

Ordinarily, the objective function coefficients should be the value that the decision maker expects. This is particularly important when using time series data as the decision maker will not necessarily expect the series average. Rather, some extrapolation of the trend may be appropriate. Brink and McCarl encountered difficulties when attempting to validate a LP model because of differences in expectations between the time the model was developed and the time actual decisions were made.

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

Several other comments are relevant regarding the objective function. First, multiplication of a LP objective function by a positive constant always leads to the same solution in terms of the decision variables. Thus, one does not need to be extremely concerned about the absolute magnitude of the objective function coefficients but rather their relative magnitudes.

Second, the coefficients must reflect the actual prices received or paid for the product. If a product is being sold, one should not use prices from distant markets but rather prices adjusted to include marketing costs. Input prices often need to be adjusted to include acquisition costs.

Finally, each objective function coefficient should be developed in harmony with the total model structure. Often, students try to insure that each and every objective function coefficient in a profit maximizing model is the per unit profit contribution arising from that particular variable. This often leads to mistakes and great confusion. Consider the model

$$
\begin{array}{lcccccc}
\text{Max} & 3X_1 & - & 2X_2 & & \\
\text{s.t.} & X_1 & - & X_2 & = & 0 \\
 & X_1 & & & \leq & 10 \\
 & & & X_2 & \leq & 8
\end{array}
$$

Suppose X_1 represents the sale of a commodity and X_2 the purchase of inputs. In order to sell X_1 one must purchase X_2 as reflected by the first constraint. One could use the equality constraint to collapse X_1 and X_2 into a single variable, but this may not be desirable. The contribution of X_1 is fully represented in the above model. The objective function should collectively represent the net margin and one does not need to compute each variable's coefficient so that it is the per unit net contribution.

6.4.6.2 Right hand Side Coefficients

Right hand side coefficients are not always easily specified. For example, consider the amount of labor available. One could think that this is the number of employees times the hours they work a week. However, the nominal and real availability of resources often differs. In the labor context, there are leaves due

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

to sickness, vacation, and alternative assignments diverting labor to other enterprises. Weather can also reduce effective availability. Finally, the right hand sides need to be developed on the same time frame as the rest of the model.

6.4.6.3 Technical Coefficients

The a_{ij} (technical) coefficients within the model give the resource use per unit of the variables. In developing technical coefficients, one usually uses economic engineering. For example, per unit labor use might be calculated by dividing the total hours of labor by the number of units produced. Such a calculation procedure by its nature includes overhead labor usages such as setup time, cleaning time, etc. However, one needs to be careful in handling fixed usages of labor which do not vary with production.

6.5 Purposeful Modeling

The purpose of a modeling exercise influences how a model is implemented. Some variables and constraints become relevant or irrelevant depending upon what exactly is to be done with the model. For example, when studying short-run operating decisions one can omit investment variables and capital constraints. On the other hand, if the focus of the study is investment one may be able to simplify the short-run operating model and come up with an approximation of how investments should be utilized if acquired. Model purpose also has important implications for the specific way a model answer is given to a decision maker.

6.5.1 Model Structure

Any problem can be formulated in a number of different ways. Modelers almost always have the option of collapsing items into the objective function or entering them explicitly in the constraints. Often the purpose of a modeling exercise influences model structure (although this is less true when using GAMS than with using conventional methods).

Years ago when LP models were solved by hand or with early LP solvers, it was desirable to construct

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

the smallest possible representation for a particular situation. Today, model condensation is not as desirable because of increased computer and solver capability. Rather, modelers often introduce size increasing features which reduce modeler/analyst interpretation and summarization time. This section discusses ways which study purpose may change formulations (although the discussion is not entirely consistent with our GAMS focus).

Consider a case in which products (X_j) are sold at an exogenously fixed price, p_j. Suppose production utilizes a number of inputs, Z_m, purchased at an exogenously fixed price, r_m. Each unit of the production variable (Y_k) incurs a direct objective function cost, q_k, yields (a_{jk}) units of the j^{th} product and uses b_{mk} units of the m^{th} input. Also, there are constraints on unpriced inputs (i). A formulation is:

$$
\begin{array}{llllllll}
\text{Max} & \sum_j p_j X_j & - & \sum_k q_k Y_k & - & \sum_m r_m Z_m & & \\
\text{s.t.} & X_j & - & \sum_k a_{jk} Y_k & & & = 0 & \text{for all } j \\
 & & & \sum_k b_{mk} Y_k & - & Z_m & = 0 & \text{for all } m \\
 & & & \sum_k c_{ik} Y_k & & & \leq e_i & \text{for all } i \\
 & X_j & , & Y_k & , & Z_m & \geq 0 &
\end{array}
$$

This formulation contains a constraint for each product and each input. One could utilize the first two constraint equations to eliminate X_j and Z_k from the model yielding the formulation:

$$
\begin{array}{llll}
\text{Max} & \sum_k g_k Y_k & & \\
\text{s.t.} & \sum_k c_{ik} Y_k \leq e_i & & \text{for all } i \\
 & Y_k \geq 0 & &
\end{array}
$$

where the g_k's are given by

$$g_k = \sum_j p_j a_{jk} - \sum_m r_m b_{mk} - q_k$$

Suppose the study involves examination of the implications of input and output prices. In the second formulation, these prices are compacted into the objective function coefficients. In the first problem, however, these prices are explicitly included in the objective function. This difference gives a reason why one might prefer the first as opposed to the second formulation. If the prices were to be repeatedly changed, then only one coefficient would have to be changed rather than many. Further, one could easily use cost-ranging features within LP algorithms to study the effects of changes in r_m. In addition, the solution would report the optimal production (X_j) and input usage (Z_m) levels. Post-solution summarization of total yield and input usage would require many calculations under the condensed model, but with Z_m explicitly included in the formulation, only one number would need to be recorded.

Usage of modeling systems like GAMS places a little different twist on the discussion above as one can easily use GAMS to do post solution report writing and since GAMS computes the whole model every time, changing one or many coefficients makes little difference.

6.5.2 Report Writing

A very important aspect of model use is properly summarizing the output so that understandable information is generated for the decision makers involved with the modeling exercise. This introduces the general topic of report writing.

Linear programming solution reports are generally inadequate for conveying the essence of the solution to the decision maker. It is highly desirable to develop reports which summarize the solution as part of the computer output, possibly an autonomous part. Such reports can be designed to translate the model solution into decision maker language using both the solution results and the input parameters. An example of such report writing is presented in Table 6.10 which gives summary reports on the transportation model from the last chapter. These reports are broken into five tables. The first Table entitled MOVEMENT gives the quantity moving between each pair of cities along with the total movement out of a particular plant and into a

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

particular market. The elements of this table are largely optimal levels of the decision variables in the solution. The second table (COSTS) gives a summary of commodity movements cost by route telling the exact cost of moving between pairs and then the total costs of moving goods out of plants or into markets. This set of outputs is not directly from the linear programming solution, but rather is the cost of movement between a particular city pair times the amount moved. The only number in the table directly from the linear programming output is the objective function value. The third gives a supply use report for each supply point giving the available supply, the amount shipped out, and the marginal value of that shipment (which is the shadow price). The fourth table gives similar information for the demand markets. Finally, there is the table CMOVEMENT which gives the cost of changing the commodity movement pattern which is a reformat of the reduced costs of the decision variables. In general, the function of a report writer is to summarize the essence of the solution, making it more readable to decision makers. In many applied studies it is valuable to develop a report format ahead of time, then structure the model and model experiments so that the report data are directly generated. The use of computerized report writing instead of hand summaries is a great advantage and can save hours and hours of modeler time. This is particularly facilitated when one uses a computerized modeling system such as GAMS.

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

References

Baker, T. and B.A. McCarl. "Representing Farm Resource Availability over Time in Linear Programs: A Case Study." North Central Journal of Agricultural Economics. 4(1982):59-68.

Brink, L. and B. McCarl. "The Tradeoff Between Expected Return and Risk Among Cornbelt Farmers." American Journal of Agricultural Economics. 60(1978):259-263.

Day, R.H. Recursive Programming and Production Response. Amsterdam: North Holland Publishing Co., 1963.

Heady, E.O. and W.V. Candler. Linear Programming Methods. Ames, Iowa: Iowa State University Press, 1958.

McCarl, B.A. "GAMSCHECK USER DOCUMENTATION: A System for Examining the Structure and Solution Properties of Linear Programming Problems Solved using GAMS." Working Documentation, Department of Agricultural Economics, Texas A&M University, 1994.

Sengupta, J. and R. Sfeir. "Allocative Behavior Under Risk Aversion Through Quadratic Programming Experiments." Applied Economics. 12(1980):367-75.

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

Table 6.1. Initial Schematic for Example Farm Planning Problem

	Crop 1	Crop 2	Crop 3	Crop 4	RHS
Objective	c_1	c_2	c_3	c_4	
Land	1	1	1		# L
Labor					

Table 6.2. Revised Schematic for Example Farm Planning Problem

	Crop 1	Crop 2	Crop 3	Crop 4	RHS
Objective	c_1	c_2	c_3	c_4	
Land	1	1	1		# L
Land After Crop 1	-1			1	# 0
Labor					

Table 6.3. Final Table for Example Farm Planning Problem

	Crop 1	Crop 2	Crop 3	Crop 4	RHS
Objective	c_1	c_2	c_3	c_4	
Land	1	1	1		# L
Land After Crop 1	-1			1	# 0
Labor –Spring	d_1	d_3	d_5		# sp
Labor – Summer	d_2			d_7	# su
Labor – Fall		d_4	d_6	d_8	# f

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

Table 6.4. Composition of a Chicken and Sales Prices for the Component Parts

	Percent of Chicken Body Weight	Sale Price of Part	Percentage Chicken Meat
	lbs. part/lbs. chicken	$/lb.	lbs. meat/lb. part
Breast Quarter	50	1.00	75
Leg Quarter	35	.80	60
Neck	10	.20	20
Giblets	5	.70	0

Table 6.5. Alternative Formulations of Chicken Processing Problem

Formulation 6.5(a)

	Chickens (lbs.)	Breast Quarter (lbs.)	Leg Quarter (lbs.)	Neck (lbs.)	Giblets (lbs.)		Maximize
Objective function ($)		+ $1.00X_1$	+ $0.80X_2$	+ $0.20X_3$	+ $0.70X_4$		
Balance (lbs.)	-Y	+ $0.50X_1$	+ $0.35X_2$	+ $0.1X_3$	+ $0.05X_4$	#	0
Chickens Available (birds)	1/3Y					#	1500

Formulation 6.5(b)

	Chickens (lbs.)	Breast Quarter (lbs.)	Leg Quarter (lbs.)	Neck (lbs.)	Giblets (lbs.)	Maximize	
Objective Function ($)		+ $1.00X_1$	+ $0.80X_2$	+ $0.20X_3$	+ $0.70X_4$		
Breast Quarter	-0.50Y	+ X_1				#	0
Leg Quarter	-0.35Y		+ X_2			#	0
Neck	-0.10Y			+ X_3		#	0
Giblets	-0.05Y				+ X_4	#	0
Chickens	1/3Y					#	1500

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

Table 6.6. Formulations for Processing Chickens with the Option of Deboning

Formulation 6.6(a)

	Chicken	Breast Qtr.	Leg Qtr.	Neck	Giblet	Breast Qtr. Meat	Leg Qtr. Meat	Neck Meat		
Objective		$1.0X_1$	+ $0.8X_2$	+ $0.2X_3$	+ $0.7X_4$	+ $1.2M_1$	+ $1.2M_2$	+ $1.2M_3$		
Breast Qtr.	-0.5Y	+ X_1							#	0
Leg Qtr.	-0.35Y		+ X_2						#	0
Neck	-0.1Y			+ X_3					#	0
Giblets	-0.5Y				+ X_4				#	0
Chickens	1/3Y								#	1500
BQ Meat	-(0.05)(0.75)Y					+ M_1			#	0
LQ Meat	-(0.35)(0.6)Y						+ M_2		#	0
N Meat	-(0.2)(0.1)Y							+ M_3	#	0

Formulation 6.6(b)

	Chicken	Breast Qtr.	Leg Qtr.	Neck	Giblet	Breast Qtr. Meat	Leg Qtr. Meat	Neck Meat	Total Meat Sold		
Objective		$1.0X_1$	$+0.8X_2$	$+0.2X_3$	$+0.7X_4$				$+ 1.2M_4$		
Breast Qtr.	-0.5Y	$+ X_1$				$+M_1$				#	0
Leg Qtr.	-0.35Y		$+ X_2$				$+M_2$			#	0
Neck	-0.1Y			$+ X_3$				$+M_3$		#	0
Giblets	-0.05Y				$+ X_4$					#	0
Chickens	1/3Y									#	1500
Meat						$-0.75M_1$	$-0.6M_2$	$-0.2M_3$	$+M_4$	#	0

Table 6.7. Formulations of the Chicken Assembly-Disassembly Problem

Formulation 6.7(a)

	Chicken	BQ	LQ	Neck	Giblet	BQ Meat	LQ Meat	Neck Meat	Total Meat Sold	MQ Sold	
Objective		1.0X_1	+ 0.8X_2	+ 0.2X_3	+ 0.7X_4				+ 1.2M_4	+ 0.95Q	
BQ	-0.5Y	+ X_1				+ M_1				+ 0.5Q	# 0
LQ	-0.35Y		+ X_2				+ M_2			+ 0.5Q	# 0
Neck	-0.1Y			+ X_3				+ M_3			# 0
Giblets	-0.05Y				+ X_4						# 0
Chickens	1/3Y										# 1500
Meat						- 0.75M_1	- 0.6M_2	- 0.2M_3	+ M_4		# 0

Formulation 6.7(b)

	Chicken	BQ	LQ	Neck	Giblet	BQ Meat	LQ Meat	Neck Meat	Total Meat Sold	BQ included in MQ	LQ included in MQ	MQ Sold	
Objective		1.0X_1	+ 0.8X_2	+ 0.2X_3	+ 0.7X_4				+ 1.2M_4			0.95Q_3	
BQ	-0.5Y	+ X_1				+ M_1				+ Q_1			# 0
LQ	-0.35Y		+ X_2				+ M_2				+ Q_2		# 0
Neck	-0.1Y			+ X_3				+ M_3					# 0
Giblets	-0.5Y				+ X_4								# 0
Chickens	1/3Y												# 1500
Meat						- 0.75M_1	- 0.6M_2	- 0.2M_3	+ M_4				# 0
Qtr. Pack										- Q_1	- Q_2	0 Q_3	# 0

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

Table 6.8. Alternative LP Formulations of Chair Production Example

Formulation 6.8(a)

	Chair Production		Regular Labor		Overtime Labor		Chair Sale		
Objective		-	10	-	15	+	220		
Chairs	-1					+	1	#	0
Regular Labor	7	-	1					#	0
Overtime Labor	3			-	1			#	0
Regular Labor Constraint			1					#	77
Overtime Labor Constraint					1			#	27

Formulation 6.8(b)

	Chair Production		Regular Labor		Overtime Labor		Chair Sale		
Objective		-	10	-	15	+	220		
Chairs	-1					+	1	#	0
Regular Labor	10	-	1	-	1			#	0
Regular Labor Constraint			1					#	77
Overtime Labor Constraint					1			#	27

copyright 1997 Bruce A. McCarl and Thomas H. Spreen

Table 6.9 Example of GAMS Report Writing

---- 53 PARAMETER MOVEMENT COMMODITY MOVEMENT

	MIAMI	HOUSTON	MINEPLIS	PORTLAND	TOTAL
NEWYORK	30	35	15		80
CHICAGO			75		75
LOSANGLS		40		50	90
TOTAL	30	75	90	50	245

---- 61 PARAMETER COSTS COMMODITY MOVEMENT COSTS BY ROUTE

	MIAMI	HOUSTON	MINEPLIS	PORTLAND	TOTAL
NEWYORK	600	1400	525		2525
CHICAGO			1500		1500
LOSANGLS		1400		2000	3400
TOTAL	600	2800	2025	2000	7425

---- 68 PARAMETER SUPPLYREP SUPPLY REPORT

	AVAILABLE	USED	MARGVALUE
NEWYORK	100.00	80.00	
CHICAGO	75.00	75.00	15.00
LOSANGLS	90.00	90.00	5.00

---- 75 PARAMETER DEMANDREP DEMAND REPORT

	REQUIRED	RECEIVED	MARGCOST
MIAMI	30.00	30.00	20.00
HOUSTON	75.00	75.00	40.00
MINEPLIS	90.00	90.00	35.00
PORTLAND	50.00	50.00	45.00

---- 80 PARAMETER CMOVEMENT COSTS OF CHANGING COMMODITY MOVEMENT PATTERN

	MIAMI	HOUSTON	MINEPLIS	PORTLAND
NEWYORK				75.00
CHICAGO	45.00	35.00		40.00
LOSANGLS	75.00		40.00	

copyright 1997 Bruce A. McCarl and Thomas H. Spreen