2

DEVELOPING MODELS FOR OPTIMIZATION

CONSTRAINTS IN OPTIMIZATION arise because a process must describe the physical bounds on the variables, empirical relations, and physical laws that apply to a specific problem, as mentioned in Section 1.4. How to develop models that take into account these constraints is the main focus of this chapter. Mathematical models are employed in all areas of science, engineering, and business to solve problems, design equipment, interpret data, and communicate information. Eykhoff (1974) defined a mathematical model as "a representation of the essential aspects of an existing system (or a system to be constructed) which presents knowledge of that system in a usable form." For the purpose of optimization, we shall be concerned with developing quantitative expressions that will enable us to use mathematics and computer calculations to extract useful information. To optimize a process models may need to be developed for the objective function f, equality constraints **g,** and inequality constraints **h.**

Because a model is an abstraction, modeling allows us to avoid repetitive experimentation and measurements. Bear in mind, however, that a model only imitates reality and cannot incorporate all features of the real process being modeled. In the development of a model, you must decide what factors are relevant and how complex the model should be. For example, consider the following questions.

1. Should the process be modeled on a fundamental or empirical level, and what level of effort (time, expenses, manpower) is required for either approach?
2. Can the process be described adequately using physical principles?
3. What is the desired accuracy of the model, and how does its accuracy influence its ultimate use?
4. What measurements are available, and what data are available for model verification?
5. Is the process actually composed of smaller, simpler subsystems that can be more easily analyzed?

The answers to these questions depend on how the model is used. As the model of the process becomes more complex, optimization usually becomes more difficult.

In this chapter we will discuss several factors that need to be considered when constructing a process model. In addition, we will examine the use of optimization in estimating the values of unknown coefficients in models to yield a compact and reasonable representation of process data. Additional information can be found in textbooks specializing in mathematical modeling. To illustrate the need to develop models for optimization, consider the following example.

EXAMPLE 2.1 MODELING AND OPTIMIZING BLAST FURNACE OPERATION

Optimizing the operation of the blast furnace is important in every large-scale steel mill. A relatively large number of important variables (several of which cannot be measured) interact in this process in a highly complex manner, numerous constraints must be taken into account, and the age and efficiency of the plant significantly affect the optimum

operating point (Deitz, 1997). Consequently, a detailed examination of this problem demonstrates the considerations involved in mathematical modeling of a typical process.

The operation of a blast furnace is semicontinuous. The raw materials are iron ore containing roughly 20 to 60 percent iron as oxides and a variety of other metallic and nonmetallic oxides. These materials are combined with coke, which reacts to form blast furnace gas. Limestone is a flux that helps separate the impurities from the hot metal by influencing the pH. Apart from the blast furnace gas, which may serve as a heating medium in other processes, the output of the furnace consists of molten iron, which includes some impurities (notably carbon and phosphorus) that must be removed in the steelmaking process, and slag, which contains most of the impurities and is of little value. Operation of the blast furnace calls for determination of the amount of each ore, a production rate, and a mode of operation that will maximize the difference between the product value and the cost of producing the required quantity and quality of molten iron. Figure E2.1 shows the flow of materials in the blast furnace, which itself is part of a much larger mill. One ton of hot metal requires about 1.7 tons of iron-bearing materials, 0.5 to 0.65 tons of coke and other fuel, 0.25 tons of fluxes, and 1.8 to 2.0 tons of air. In addition, for each ton of hot metal produced, the process creates 0.2 to 0.4 tons of slag, 0.05 tons or less of flue dust, and 2.5 to 3.5 tons of blast furnace gases. The final product, hot metal, is about 93% iron, with other trace ingredients, including sulfur, silicon, phosphorus, and manganese. The process variables and conceptual models are identified in Figure E2.1 under the column "Process Analysis," which has categories for the objective function, equality constraints, and inequality constraints.

Objective function

To formulate the objective function, two categories of costs have to be considered:

1. Costs associated with the material flows (the input and output variables), such as the costs of purchased materials.
2. Costs associated with the operations related to the process variables in the model.

The terms that make up the objective function (to be maximized) are shown in Figure E.2.1. The profit of the blast furnace can be expressed as

$$f = \sum_{i=7}^{8} c_i x_i - \sum_{i=1}^{6} c_i x_i$$

Equality and inequality constraints

The next step in formulating the problem is to construct a mathematical model of the process by considering the fundamental chemical and physical phenomena and physical limitations that influence the process behavior. For the case of the blast furnace, typical features are

1. *Iron ore*: Ores of different grades are available in restricted quantities. Different ores have varying percentages of iron and different types and amounts of impurities. The proportion of each ore that occurs in the final hot metal is assumed to be fixed by its composition. For example, the amount of fine ore must be limited because too much can disrupt the flow of gas through the furnace and limit production.
2. *Coke*: The amount of coke that may be burned in any furnace is effectively limited by the furnace design, and the hot metal temperature is controlled by the amount

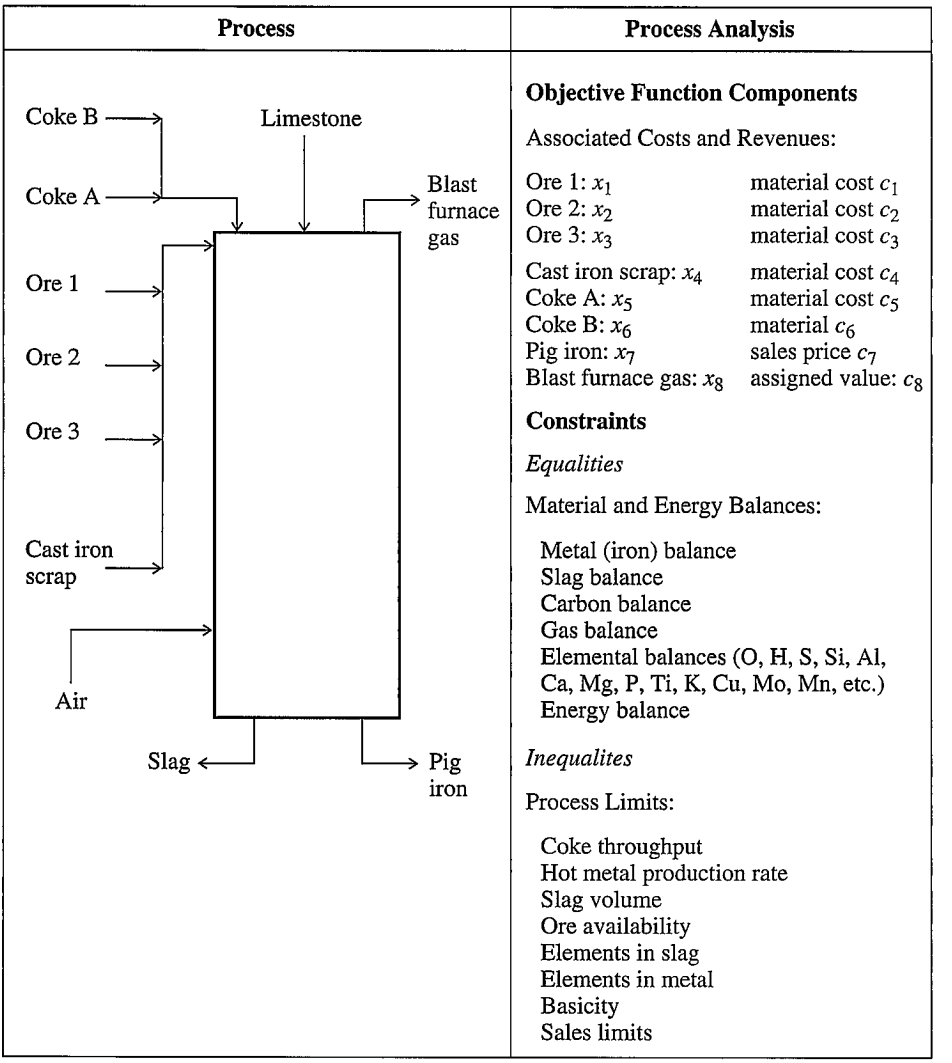


FIGURE E.2.1
Objective function components and types of constraints for a blast furnace.

of coke (or carbon). The coke consumption rate can be based on empirical relationships developed through regression of furnace data.

3. *Slag*: For technical reasons, the level of impurities in the slag must be controlled. There is an upper limit on the percentage of magnesium, upper and lower limits on the percentage of silicon and aluminum, and close limits on the "basicity" ratio $(CaO + MgO)/(SiO_2 + A1_2O_3)$. The basicity ratio controls the viscosity and melting point of the slag, which in turn affect the hearth temperature and grade of iron produced.

The basicity ratio can be expressed in terms of the blast furnace feeds x_i as follows:

$$\frac{\sum_{i=1}^{4} w_{2i}x_i + \sum_{i=1}^{4} w_{3i}x_i}{\sum_{i=1}^{4} w_{4i}x_i + \sum_{i=1}^{4} w_{5i}x_i}$$

where w_{2i} = weight fraction of CaO in feed i
w_{3i} = weight fraction of MgO in feed i
w_{4i} = weight fraction of SiO_2 in feed i
w_{5i} = weight fraction of $A1_2O_3$ in feed i

4. *Phosphorus*: All phosphorus in the raw material finds its way into the molten metal. There is an upper limit on the phosphorus permitted, although precise quantities are sometimes prescribed. In general, it is cheaper to produce higher phosphorus iron, but more expensive to refine it.

From these and other considerations you can prepare:

1. A set of input and output variables.
2. A set of steady-state input–output material and energy balances (equality constraints).
3. A set of explicit empirical relations (equality constraints).
4. A set of restrictions (inequality constraints) on the input and output variables as indicated in Figure E.2.1.

2.1 CLASSIFICATION OF MODELS

Two general categories of models exist:

1. Those based on physical theory.
2. Those based on strictly empirical descriptions (so-called black box models).

Mathematical models based on physical and chemical laws (e.g., mass and energy balances, thermodynamics, chemical reaction kinetics) are frequently employed in optimization applications (refer to the examples in Chapters 11 through 16). These models are conceptually attractive because a general model for any system size can be developed even before the system is constructed. A detailed exposition of fundamental mathematical models in chemical engineering is beyond our scope here, although we present numerous examples of physiochemical models throughout the book, especially in Chapters 11 to 16. Empirical models, on the other hand, are attractive when a physical model cannot be developed due to limited time or resources. Input–output data are necessary in order to fit unknown coefficients in either type of the model.

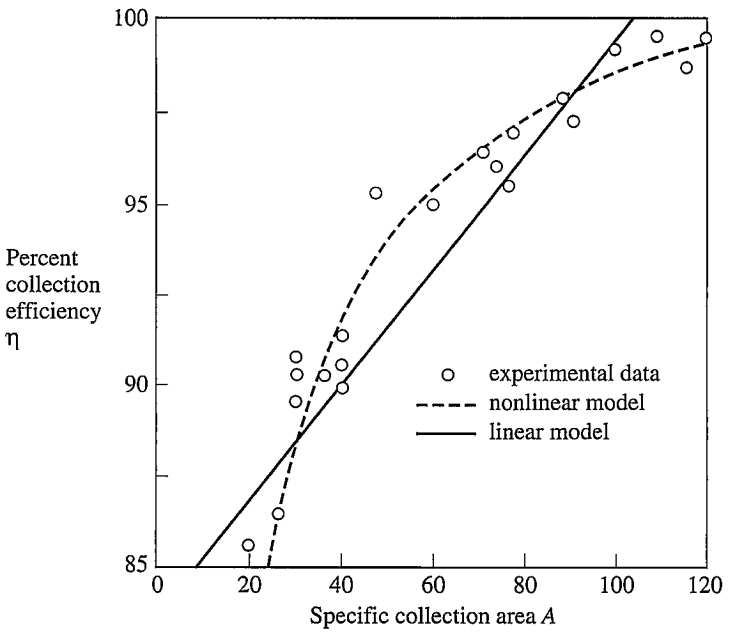


FIGURE E2.2
ESP collection efficiency versus specific collection area for a linear model $\eta = 0.129A + 85.7$ and a nonlinear model $\eta = 100\{1 - [e^{-0.0264A}/(4.082 - 3.15 \times 10^{-6} A)]\}$.

EXAMPLE 2.2 MODELS OF AN ELECTROSTATIC PRECIPITATOR

A coal combustion pilot plant is used to obtain efficiency data on the collection of particulate matter by an electrostatics precipitator (ESP). The ESP performance is varied by changing the surface area of the collecting plates. Figure E2.2 shows the data collected to estimate the coefficients in a model to represent efficiency η as a function of the specific collection area A, measured as plate area/volumetric flow rate.

Two models of different complexity have been proposed to fit the performance data:

$$\text{Model 1:} \quad \eta = b_1 A + b_2$$

$$\text{Model 2:} \quad \eta = 100\left[1 - \frac{e^{-\gamma_1 A}}{\gamma_2 + \gamma_3 A}\right]$$

Model 1 is linear in the coefficients, and model 2 is nonlinear in the coefficients. The mathematical structure of model 2 has a fundamental basis that takes into account the physical characteristics of the particulate matter, including particle size and electrical properties, but we do not have the space to derive the equation here.

Which model is better?

Solution. The coefficients in the two models were fitted using MATLAB, yielding the following results:

Model 1: $b_1 = 0.129 \quad b_2 = 85.7$

Model 2: $\gamma_1 = 0.0264 \quad \gamma_2 = 4.082 \quad \gamma_3 = -0.00000315$

As can be seen in Figure E2.2, model 2 provides a better fit than model 1 over the range of areas A considered, but model 2 may present some difficulties when used as a constraint inserted into an optimization code.

The electrostatic precipitator in Example 2.2 is typical of industrial processes; the operation of most process equipment is so complicated that application of fundamental physical laws may not produce a suitable model. For example, thermodynamic or chemical kinetics data may be required in such a model but may not be available. On the other hand, although the development of black box models may require less effort and the resulting models may be simpler in form, empirical models are usually only relevant for restricted ranges of operation and scale-up. Thus, a model such as ESP model 1 might need to be completely reformulated for a different size range of particulate matter or for a different type of coal. You might have to use a series of black box models to achieve suitable accuracy for different operating conditions.

In addition to classifying models as theoretically based versus empirical, we can generally group models according to the following types:

Linear versus nonlinear.
Steady state versus unsteady state.
Lumped parameter versus distributed parameter.
Continuous versus discrete variables.

Linear versus nonlinear

Linear models exhibit the important property of superposition; nonlinear ones do not. Equations (and hence models) are linear if the dependent variables or their derivatives appear only to the first power: otherwise they are nonlinear. In practice the ability to use linear models is of great significance because they are an order of magnitude easier to manipulate and solve than nonlinear ones.

To test for the linearity of a model, examine the equation(s) that represents the process. If any one term is nonlinear, the model itself is nonlinear. By implication, the process is nonlinear.

Examine models 1 and 2 for the electrostatic precipitator. Is model 1 linear in A? Model 2? The superposition test in each case is: Does

$$J(ax_1 + bx_2) = aJ(x_1) + bJ(x_2) \tag{2.1a}$$

and

$$J(kx) = kJ(x) \tag{2.1b}$$

where J = any operator contained in the model such as square, differentiation, and so on.
k = a constant
x_1 and x_2 = variables

ESP model 1 is linear in A

$$J(b_1A + b_2) = b_1J(A) + b_2$$

but ESP model 2 is nonlinear because

$$\left(\frac{e^{-\gamma_1(A_1 + A_2)}}{\gamma_2 + \gamma_3(A_1 + A_2)}\right) \neq \left(\frac{e^{-\gamma_1 A_1}}{\gamma_2 + \gamma_3 A_2}\right) + \left(\frac{e^{-\gamma_1 A_2}}{\gamma_2 + \gamma_3 A_2}\right)$$

Steady state versus unsteady state

Other synonyms for steady state are time-invariant, static, or stationary. These terms refer to a process in which the values of the dependent variables remain constant with respect to time. Unsteady state processes are also called nonsteady state, transient, or dynamic and represent the situation when the process-dependent variables change with time. A typical example of an unsteady state process is the operation of a batch distillation column, which would exhibit a time-varying product composition. A transient model reduces to a steady state model when $\partial/\partial t = 0$. Most optimization problems treated in this book are based on steady state models. Optimization problems involving dynamic models usually pertain to "optimal control" or real-time optimization problems (see Chapter 16)

Distributed versus lumped parameters

Briefly, a lumped parameter representation means that spatial variations are ignored and that the various properties and the state of the system can be considered homogeneous throughout the entire volume. A distributed parameter representation, on the other hand, takes into account detailed variations in behavior from point to point throughout the system. In Figure 2.1, compare these definitions for a well-stirred reactor and a tubular reactor with axial flow. In the first case, we assume that mixing is complete so no concentration or temperature gradient occurs in the reactor, hence a lumped parameter mathematical model would be appropriate. In contrast, the tubular reactor has concentration or temperature variations along the axial direction and perhaps in the radial direction, hence a distributed parameter model would be required. All real systems are, of course, distributed because some variations of states occur throughout them. Because the spatial variations often are relatively small, they may be ignored, leading to a lumped approximation. If both spatial and transient characteristics are to be included in a model, a partial differential equation or a series of stages is required to describe the process behavior.

It is not easy to determine whether lumping in a process model is a valid technique for representing the process. A good rule of thumb is that if the response is

essentially the same at all points in the process, then the model can be lumped as a single unit. If the response shows significant instantaneous differences in any direction along the vessel, then the problem should be treated using an appropriate differential equation or series of compartments. In an optimization problem it is desirable to simplify a distributed model by using an equivalent lumped parameter system, although you must be careful to avoid masking the salient features of the distributed element (hence building an inadequate model). In this text, we will mainly consider optimization techniques applied to lumped systems.

Continuous versus discrete variables

Continuous variables can assume any value within an interval; discrete variables can take only distinct values. An example of a discrete variable is one that assumes integer values only. Often in chemical engineering discrete variables and continuous variables occur simultaneously in a problem. If you wish to optimize a compressor system, for example, you must select the number of compressor stages (an integer) in addition to the suction and production pressure of each stage (positive continuous variables). Optimization problems without discrete variables are far easier to solve than those with even one discrete variable. Refer to Chapter 9 for more information about the effect of discrete variables in optimization.

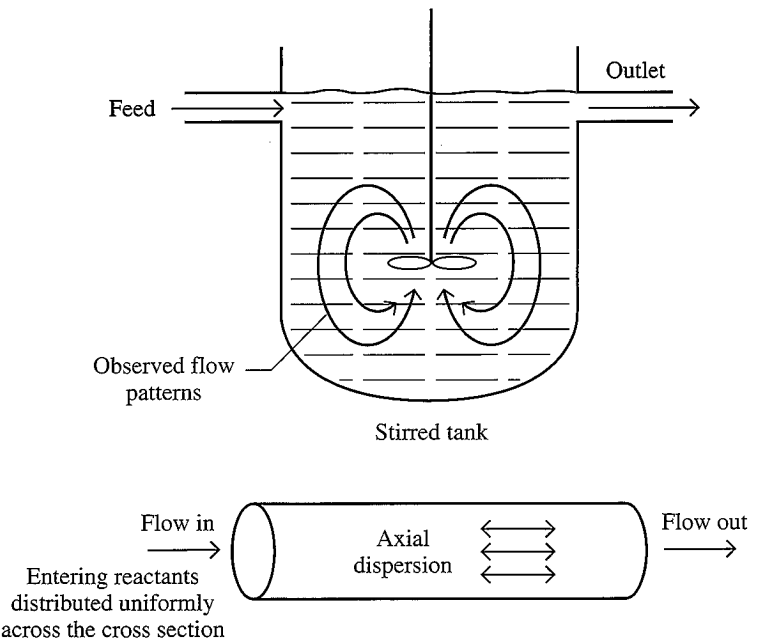


FIGURE 2.1
Flow patterns in a stirred tank (lumped parameter system) and a tubular reactor (distributed parameter system).

An engineer typically strives to treat discrete variables as continuous even at the cost of achieving a suboptimal solution when the continuous variable is rounded off. Consider the variation of the cost of insulation of various thickness as shown in Figure E1.1. Although insulation is only available in 0.5-in. increments, continuous approximation for the thickness can be used to facilitate the solution to this optimization problem.

2.2 HOW TO BUILD A MODEL

For convenience of presentation, model building can be divided into four phases: (1) problem definition and formulation, (2) preliminary and detailed analysis, (3) evaluation, and (4) interpretation application. Keep in mind that model building is an iterative procedure. Figure 2.2 summarizes the activities to be carried out,

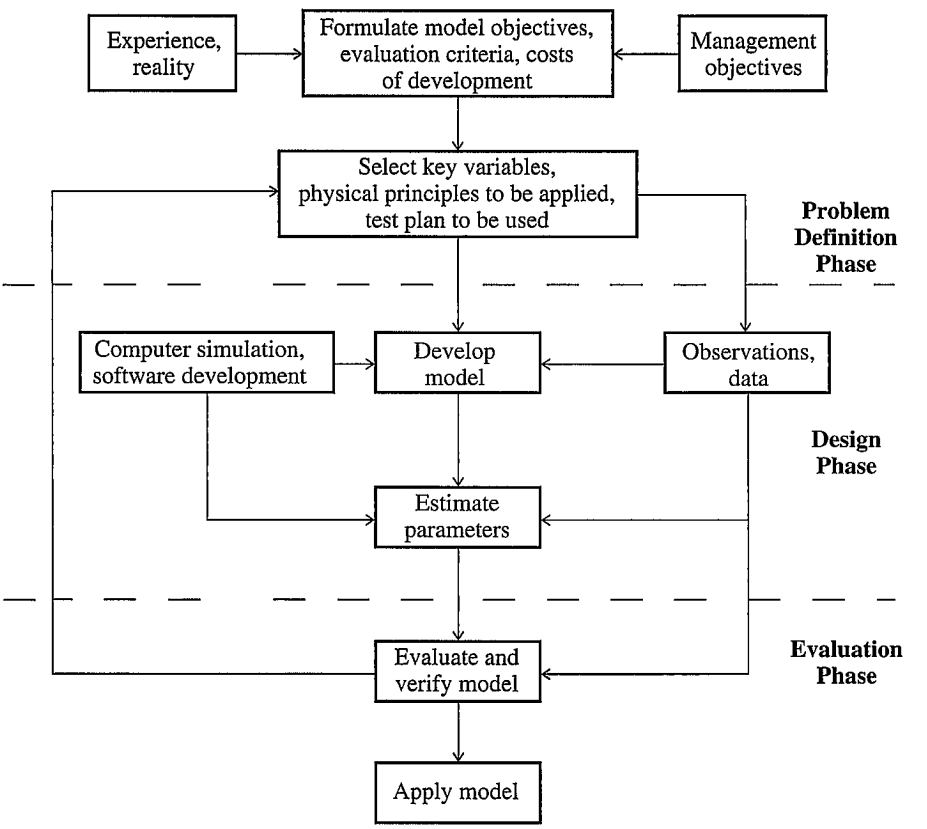


FIGURE 2.2
Major activities in model building prior to application.

which are discussed in detail later on. The content of this section is quite limited in scope; before actually embarking on a comprehensive model development program, consult textbooks on modeling (see References).

Problem definition and formulation phase

In this phase the problem is defined and the important elements that pertain to the problem and its solution are identified. The degree of accuracy needed in the model and the model's potential uses must be determined. To evaluate the structure and complexity of the model, ascertain

1. The number of independent variables to be included in the model.
2. The number of independent equations required to describe the system (sometimes called the "order" of the model).
3. The number of unknown parameters in the model.

In the previous section we addressed some of these issues in the context of physical versus empirical models. These issues are also intertwined with the question of model verification: what kinds of data are available for determining that the model is a valid description of the process? Model building is an iterative process, as shown by the recycling of information in Figure 2.2.

Before carrying out the actual modeling, it is important to evaluate the economic justification for (and benefits of) the modeling effort and the capability of support staff for carrying out such a project. Primarily, determine that a successfully developed model will indeed help solve the optimization problem.

Design phase

The design phase includes specification of the information content, general description of the programming logic and algorithms necessary to develop and employ a useful model, formulation of the mathematical description of such a model, and simulation of the model. First, define the input and output variables, and determine what the "system" and the "environment" are. Also, select the specific mathematical representation(s) to be used in the model, as well as the assumptions and limitations of the model resulting from its translation into computer code. Computer implementation of the model requires that you verify the availability and adequacy of computer hardware and software, specify computer input–output media, develop program logic and flowsheets, and define program modules and their structural relationships. Use of existing subroutines and databases saves you time but can complicate an optimization problem for the reasons explained in Chapter 15.

Evaluation phase

This phase is intended as a final check of the model as a whole. Testing of individual model elements should be conducted during earlier phases. Evaluation of the model is carried out according to the evaluation criteria and test plan established in the problem definition phase. Next, carry out sensitivity testing of the model inputs

and parameters, and determine if the apparent relationships are physically meaningful. Use actual data in the model when possible. This step is also referred to as diagnostic checking and may entail statistical analysis of the fitted parameters (Box et al., 1978).

Model validation requires confirming logic, assumptions, and behavior. These tasks involve comparison with historical input–output data, or data in the literature, comparison with pilot plant performance, and simulation. In general, data used in formulating a model should not be used to validate it if at all possible. Because model evaluation involves multiple criteria, it is helpful to find an expert opinion in the verification of models, that is, what do people think who know about the process being modeled?

No single validation procedure is appropriate for all models. Nevertheless, it is appropriate to ask the question: What do you want the model to do? In the best of all possible worlds, you want the model to predict the desired process performance with suitable accuracy, but this is often an elusive goal.

2.3 SELECTING FUNCTIONS TO FIT EMPIRICAL DATA

A model relates the output (the dependent variable or variables) to the independent variable(s). Each equation in the model usually includes one or more coefficients that are presumed constant. The term *parameter* as used here means coefficient and possibly input or initial condition. With the help of experimental data, we can determine the *form* of the model and subsequently (or simultaneously) estimate the value of some or all of the parameters in the model.

2.3.1 How to Determine the Form of a Model

Models can be written in a variety of mathematical forms. Figure 2.3 shows a few of the possibilities, some of which were already illustrated in Section 2.1. This section focuses on the simplest case, namely models composed of algebraic equations, which constitute the bulk of the equality constraints in process optimization. Emphasis here is on estimating the coefficients in simple models and not on the complexity of the model.

Selection of the form of an empirical model requires judgment as well as some skill in recognizing how response patterns match possible algebraic functions. Optimization methods can help in the selection of the model structure as well as in the estimation of the unknown coefficients. If you can specify a quantitative criterion that defines what "best" represents the data, then the model can be improved by adjusting its form to improve the value of the criterion. The best model presumably exhibits the least error between actual data and the predicted response in some sense.

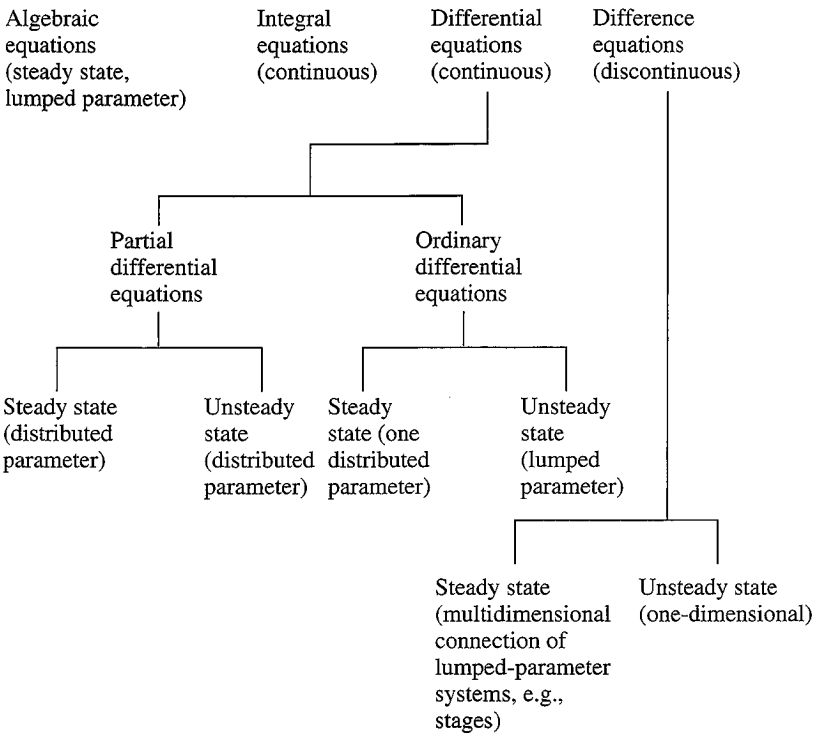


FIGURE 2.3
Typical mathematical forms of models.

Typical relations for empirical models might be

$y = a_0 + a_1x_1 + a_2x_2 + \cdots$	linear in the variables and coefficients
$y = a_0 + a_{11}x_1^2 + a_{12}x_1x_2 + \cdots$	linear in the coefficients, nonlinear in the variables (x_1, x_2)
$G(s) = \dfrac{1}{a_0 + a_1s + a_2s^2}$	nonlinear in all the coefficients
$\text{Nu} = a(\text{Re})^b$	nonlinear in the coefficient b (Nu: Nusselt number; Re: Reynolds number)

When the model is linear in the coefficients, they can be estimated by a procedure called *linear regression*. If the model is nonlinear in the coefficients, estimating them is referred to as *nonlinear regression*. In either case, the simplest adequate model (with the fewest number of coefficients) should be used.

Graphical presentation of data assists in determining the form of the function of a single variable (or two variables). The response y versus the independent variable x can be plotted and the resulting form of the model evaluated visually. Figure 2.4 shows experimental heat transfer data plotted on log–log coordinates. The plot

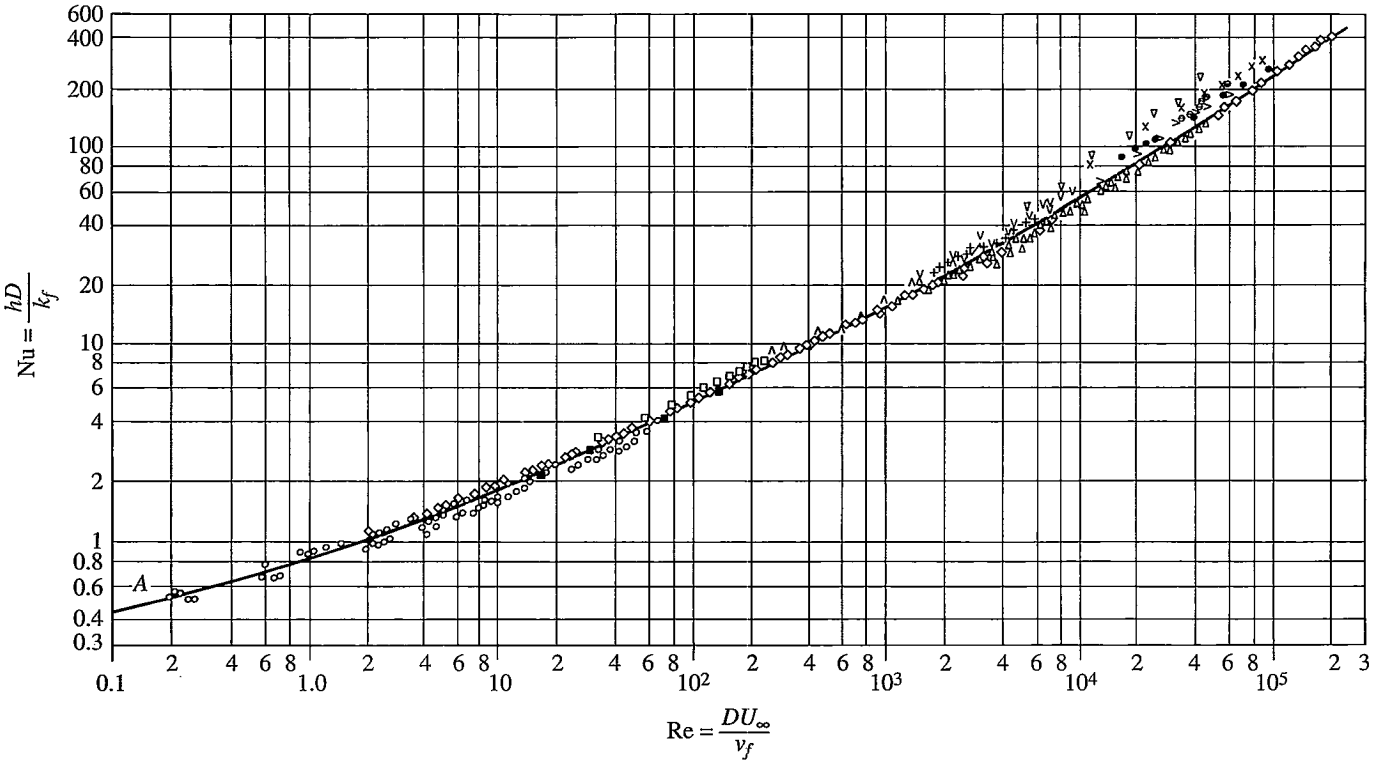


FIGURE 2.4
Average Nusselt number (Nu) versus Reynolds number (Re) for a circular cylinder in air, placed normal to the flow (McAdams, 1954, with permission from McGraw-Hill Companies).

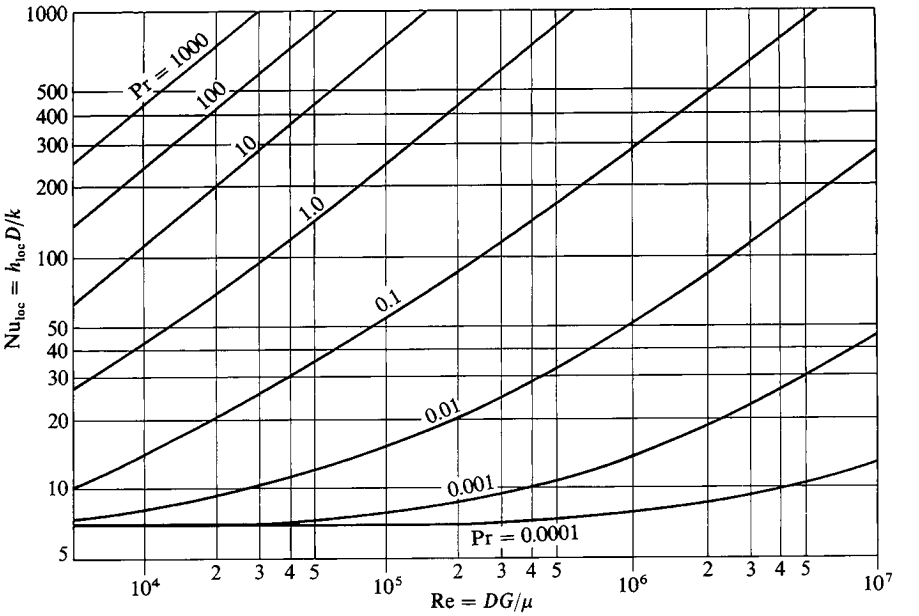
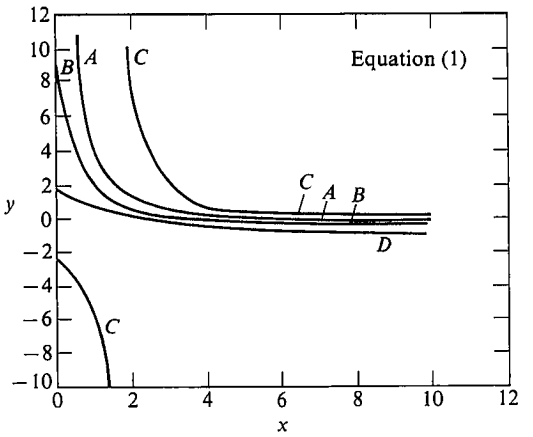


FIGURE 2.5
Predicted Nusselt numbers for turbulent flow with constant wall heat flux (*adapted with permission from John Wiley and Sons from Bird et al., 1964). Abbreviations*: Nu = Nusselt number; Re = Reynolds number; Pr = Prandtl number.

appears to be approximately linear over wide ranges of the Reynolds number (Re). A straight line in Figure 2.4 would correspond to log Nu = log a + b log Re or Nu = a(Re)^b. Observe the scatter of experimental data in Figure 2.4, especially for large values of the Re.

If two independent variables are involved in the model, plots such as those shown in Figure 2.5 can be of assistance; in this case the second independent variable becomes a parameter that is held constant at various levels. Figure 2.6 shows a variety of nonlinear functions and their associated plots. These plots can assist in selecting relations for nonlinear functions of y versus x. Empirical functions of more than two variables must be built up (or pruned) step by step to avoid including an excessive number of irrelevant variables or missing an important one. Refer to Section 2.4 for suitable procedures.

Now let us review an example for selecting the form of a model to fit experimental data.



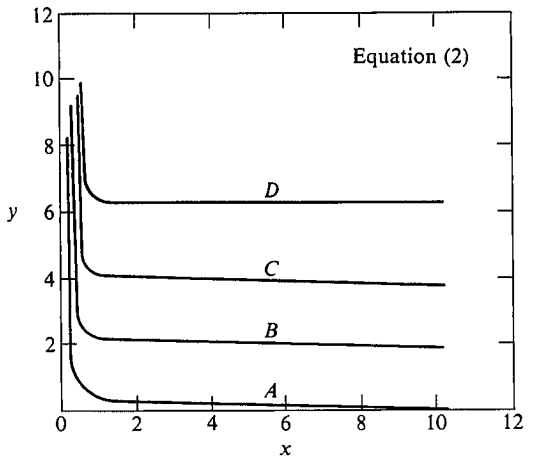
(1) $\frac{1}{y} = \alpha - \beta x$

A. $\frac{1}{y} = -0.1 - 0.3x$

B. $\frac{1}{y} = 0.1 - 0.3x$

C. $\frac{1}{y} = -0.5 - 0.3x$

D. $\frac{1}{y} = 0.5 + 0.3x$



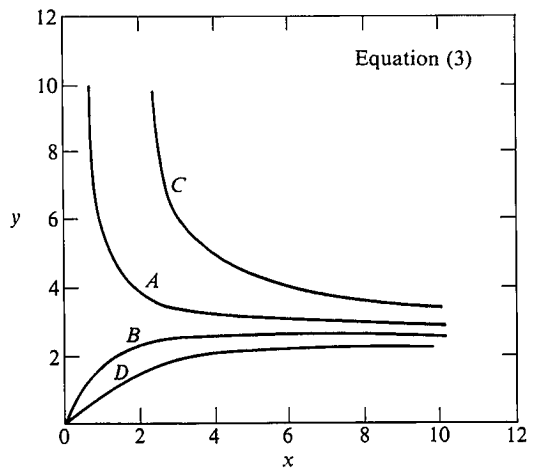
(2) $y = \alpha + \frac{\beta}{x}$

A. $y = -0.1 + \frac{0.3}{x}$

B. $y = 2 + \frac{0.3}{x}$

C. $y = 4 + \frac{0.3}{x}$

D. $y = 6 + \frac{0.3}{x}$



(3) $\frac{x}{y} = \alpha + \beta x$

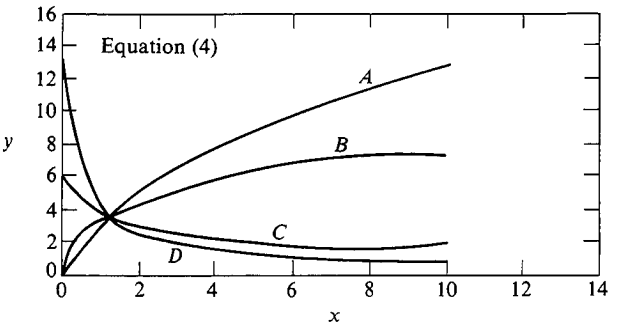
A. $\frac{x}{y} = -0.1 + 0.3x$

B. $\frac{x}{y} = 0.1 + 0.3x$

C. $\frac{x}{y} = -0.4 + 0.3x$

D. $\frac{x}{y} = 4 + 0.3x$

FIGURE 2.6
Functions of a single variable x and their corresponding trajectories. (*Continues*)



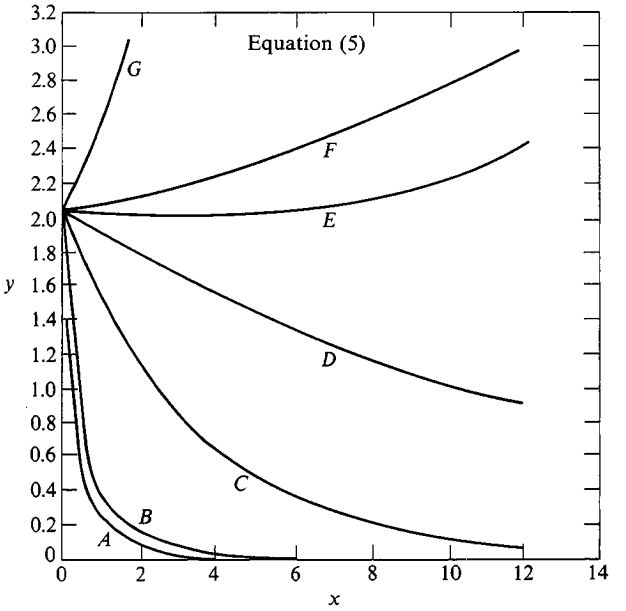
(4) $y = \alpha x^{\beta}$

A. $y = 4x^{0.5}$

B. $y = 4x^{0.3}$

C. $y = 4x^{-0.3}$

D. $y = 4x^{-0.5}$



(5) $y = \alpha\beta^{x}$

A. $y = 2(0.2)^{x}$

B. $y = 2(0.3)^{x}$

C. $y = 2(0.8)^{x}$

D. $y = 2(0.95)^{x}$

E. $y = 2(1.02)^{x}$

F. $y = 2(1.04)^{x}$

G. $y = 2(1.3)^{x}$

FiIGURE 2.6 *(continued)*

EXAMPLE 2.3 ANALYSIS OF THE HEAT TRANSFER COEFFICIENT

Suppose the overall heat transfer coefficient of a shell-and-tube heat exchanger is calculated daily as a function of the flow rates in both the shell and tube sides (w_s and w_t, respectively). U has the units of Btu/(h)(°F)(ft^2), and w_s and w_t are in lb/h. Figures E2.3a and E2.3b illustrate the measured data. Determine the form of a semiempirical model of U versus w_s and w_t based on physical analysis.

Solution. You could elect to simply fit U as a polynomial function of w_s and w_t; there appears to be very little effect of w_s on U, but U appears to vary linearly with w_t (except at the upper range of w_t where it begins to level off). A more quantitative approach

can be based on a physical analysis of the exchanger. First determine why w_s has no effect on U. This result can be explained by the formula for the overall heat transfer coefficient

$$\frac{1}{U} = \frac{1}{h_s} + \frac{1}{h_t} + \frac{1}{h_f} \qquad (a)$$

where h_s = the shell heat transfer coefficient
h_t = the tube side heat transfer coefficient
h_f = the fouling coefficient

If h_t is small and h_s is large, U is dominated by h_t, hence changes in w_s have little effect, as shown in Figure E2.3a.

Next examine the data for U versus w_t in the context of Figure 2.6. For a reasonable range of w_t the pattern is similar to curve D in Equation (3) where

$$\frac{x}{y} = \alpha + \beta x \qquad (b)$$

which can also be written as

$$\frac{1}{y} = \frac{\alpha}{x} + \beta \qquad (c)$$

Note the similarity between Equations (*c*) and (*a*), where $x = h_t$ and $y = U$. From a standard heat transfer coefficient correlation (Gebhart, 1971), you can find that h_t also varies according to $K_t w_t^{0.8}$, where K_t is a coefficient that depends on the fluid physical properties and the exchanger geometry. If we lump $1/h_s$ and $1/h_f$ together into one constant $1/h_{sf}$, the semiempirical model becomes

$$\frac{1}{U} = \frac{1}{h_{sf}} + \frac{1}{K_t w_t^{0.8}}$$

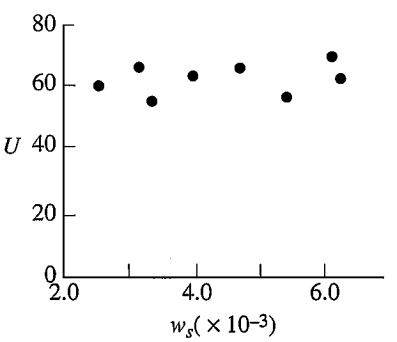


FIGURE E2.3*a*
Variation of overall heat transfer coefficient with shell-side flow rate w_s = 8000.

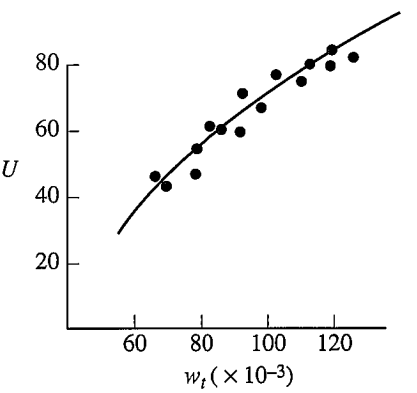


FIGURE E2.3*b*
Variation of overall heat transfer coefficient with tube-side flow rate w_t for w_s = 4000.

or

$$U = \frac{h_{sf} K_t w_t^{0.8}}{K_t w_t^{0.8} + h_{sf}} \tag{d}$$

The line in Figure E2.3b shows how well Equation (*d*) fits the data.

In the previous examples and figures we indicated that functions for two independent variables can be selected. When three (or more) independent variables occur, advanced analysis tools, such as experimental design (see Section 2.4) or principal component analysis (Jackson, 1991), are required to determine the structure of the model.

Once the form of the model is selected, even when it involves more than two independent variables, fitting the unknown coefficients in the model using linear or nonlinear regression is reasonably straightforward. We discuss methods of fitting coefficients in the next section.

2.3.2 Fitting Models by Least Squares

This section describes the basic idea of least squares estimation, which is used to calculate the values of the coefficients in a model from experimental data. In estimating the values of coefficients for either an empirical or theoretically based model, keep in mind that the number of data sets must be equal to or greater than the number of coefficients in the model. For example, with three data points of y versus x, you can estimate at most the values of three coefficients. Examine Figure 2.7. A straight line might represent the three points adequately, but the data can be fitted exactly using a quadratic model

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 \tag{2.2}$$

By introducing the values of a data point (Y_1, x_1) into Equation 2.2, you obtain one equation of Y_1 as a function of three unknown coefficients. The set of three data points therefore yields three linear equations in three unknowns (the coefficients) that can be solved easily.

To compensate for the errors involved in experimental data, the number of data sets should be greater than the number of coefficients p in the model. Least squares is just the application of optimization to obtain the "best" solution of the equations, meaning that the sum of the squares of the errors between the predicted and the experimental values of the dependent variable y for each data point x is minimized. Consider a general algebraic model that is linear in the coefficients.

$$y = \sum_{j=1}^{p} \beta_j x_j \tag{2.3}$$

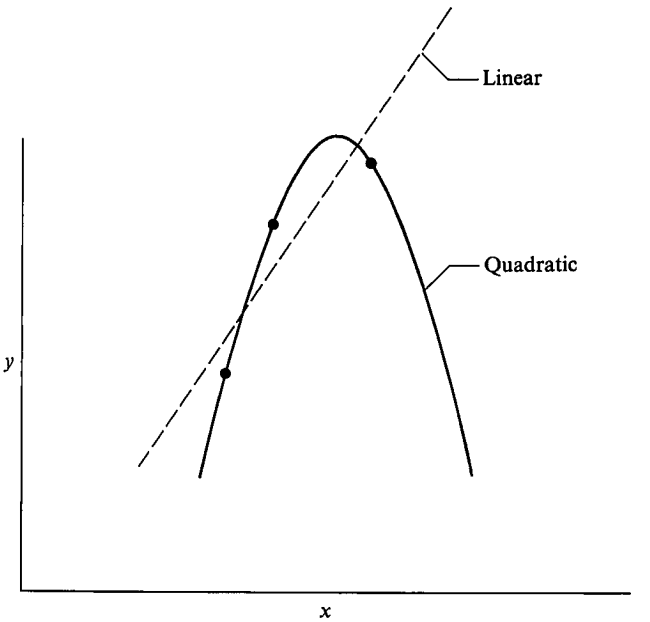


FIGURE 2.7
Linear versus quadratic fit for three data points.

There are p independent variables x_j, $j = 1, \ldots, p$. Independent here means controllable or adjustable, not functionally independent. Equation (2.3) is linear with respect to the β_j, but x_j can be nonlinear. Keep in mind, however, that the values of x_j (based on the input data) are just numbers that are substituted prior to solving for the estimates $\hat{\beta}_j$, hence nonlinear functions of x_j in the model are of no concern. For example, if the model is a quadratic function,

$$y = \beta_1 + \beta_2 x + \beta_3 x^2$$

we specify

$$x_1 = 1$$
$$x_2 = x$$
$$x_3 = x^2$$

and the general structure of Equation (2.3) is satisfied. In reading Section 2.4 you will learn that special care must be taken in collecting values of x to avoid a high degree of correlation between the x_i's.

Introduction of Equation (2.3) into a sum-of-squares error objective function gives

$$f = \sum_{i=1}^{n} \left(Y_i - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 \tag{2.4}$$

The independent variables are now identified by a double subscript, the first index designating the data set (experiment) number ($i = 1, \ldots, n$) and the second the independent variables ($j = 1, p$).

Minimizing f with respect to the β's involves differentiating f with respect to $\beta_1, \beta_2, \ldots, \beta_p$ and equating the p partial derivatives to zero. This yields p equations that relate the p unknown values of the estimated coefficients $\hat{\beta}_1, \ldots, \hat{\beta}_p$:

$$\begin{aligned}
\hat{\beta}_1 \sum_{i=1}^{n} x_{i1}x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i1}x_{i2} + \cdots + \hat{\beta}_p \sum_{i=1}^{n} x_{i1}x_{ip} &= \sum_{i=1}^{n} Y_i x_{i1} \\
\hat{\beta}_1 \sum_{i=1}^{n} x_{i2}x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2}x_{i2} + \cdots + \hat{\beta}_p \sum_{i=1}^{n} x_{i2}x_{ip} &= \sum_{i=1}^{n} Y_i x_{i2} \\
\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots & \\
\hat{\beta}_1 \sum_{i=1}^{n} x_{ip}x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{ip}x_{i2} + \cdots + \hat{\beta}_p \sum_{i=1}^{n} x_{ip}x_{ip} &= \sum_{i=1}^{n} Y_i x_{ip}
\end{aligned} \tag{2.5}$$

where $\hat{\beta}_i$ = the estimated value of β_i
x_{ij}'s = the experimental values of x_j
Y_i = the measured dependent variables

Note the symmetry of the summation terms in x_{ij} and that numbering of x_{ij}'s in the summations corresponds to matrix indices (rows, columns). This set of p equations in p unknowns can be solved on a computer using one of the many readily available routines for solving simultaneous linear equations.

Equations (2.5) can be expressed in more compact form if matrix notation is employed (see Appendix A). Let the model be expressed in vector matrix notation as

$$\mathbf{Y} = \mathbf{x}\hat{\boldsymbol{\beta}} + \boldsymbol{\varepsilon} \tag{2.6}$$

where $\boldsymbol{\varepsilon}$ = the random error in the data
$\mathbf{Y}$ = the vector of measured dependent variables

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_p \end{bmatrix} \qquad \mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_n \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

The objective function to be minimized is

$$f = \boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = (\mathbf{Y} - \mathbf{x}\boldsymbol{\beta})^T(\mathbf{Y} - \mathbf{x}\boldsymbol{\beta}) \tag{2.7}$$

Equations 2.5 can then be expressed as

$$\mathbf{x}^T\mathbf{x}\hat{\boldsymbol{\beta}} = \mathbf{x}^T\mathbf{Y} \tag{2.8}$$

which has the formal solution via matrix algebra

$$\hat{\boldsymbol{\beta}} = (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\mathbf{Y} \tag{2.9}$$

Statistical packages and spreadsheets solve the simultaneous equations in (2.8) to estimate $\hat{\boldsymbol{\beta}}$ rather than computing the matrix inverse in Equation (2.9).

The next two examples illustrate the application of Equation 2.9 to fit coefficients in an objective function. The same procedure is used to fit coefficients in constraint models.

EXAMPLE 2.4 APPLICATION OF LEAST SQUARES TO DEVELOP A COST MODEL FOR THE COST OF HEAT EXCHANGERS

In the introduction we mentioned that it is sometimes necessary to develop a model for the objective function using cost data. Curve fitting of the costs of fabrication of heat exchangers can be used to predict the cost of a new exchanger of the same class with different design variables. Let the cost be expressed as a linear equation

$$C = \beta_1 + \beta_2 N + \beta_3 A$$

where β_1, β_2, and β_3 are constants
N = number of tubes
A = shell surface area

Estimate the values of the constants β_1, β_2, and β_3 from the data in Table E2.4. The regressors are $x_1 = 1$, $x_2 = N$, and $x_3 = A$.

Solution. The matrices to be used in calculating $\hat{\boldsymbol{\beta}}$ are as follows (each data set is weighted equally):

$$\mathbf{x} = \begin{bmatrix} 1 & 120 & 550 \\ 1 & 130 & 600 \\ 1 & 108 & 520 \\ 1 & 110 & 420 \\ 1 & 84 & 400 \\ 1 & 90 & 300 \\ 1 & 80 & 230 \\ 1 & 55 & 120 \\ 1 & 64 & 190 \\ 1 & 50 & 100 \end{bmatrix}$$

TABLE E2.4
Labor cost data for mild-steel floating-head exchangers (0–500 psig) working pressure

Labor cost ($)	Area (A)	Number of tubes (N)
310	120	550
300	130	600
275	108	520
250	110	420
220	84	400
200	90	300
190	80	230
150	55	120
140	64	190
100	50	100

Source: Shahbenderian, 1961.

$$(\mathbf{x}^T\mathbf{x}) = \begin{bmatrix} 10 & 891 & 3{,}430 \\ 891 & 86{,}241 & 349{,}120 \\ 3{,}430 & 349{,}120 & 1{,}472{,}700 \end{bmatrix}$$

$$(\mathbf{x}^T\mathbf{Y}) = \begin{bmatrix} 2{,}135 \\ 207{,}290 \\ 844{,}800 \end{bmatrix}$$

Equation (2.9) gives the best estimates of β_1, β_2, and β_3:

$$\hat{\beta}_1 = 38.177$$

$$\hat{\beta}_2 = 1.164$$

$$\hat{\beta}_3 = 0.209$$

Check to see if these coefficients yield a reasonable fit to the data in Table E2.4.

EXAMPLE 2.5 APPLICATION OF LEAST SQUARES IN YIELD CORRELATION

Ten data points were taken in an experiment in which the independent variable x is the mole percentage of a reactant and the dependent variable y is the yield (in percent):

x	y
20	73
20	78
30	85
40	90
40	91
50	87
50	86
50	91
60	75
70	65

Fit a quadratic model with these data and determine the value of x that maximizes the yield.

Solution. The quadratic model is $y = \beta_1 + \beta_2 x + \beta_3 x^2$. The estimated coefficients computed using Excel are

$$\hat{\beta}_1 = 35.66$$

$$\hat{\beta}_2 = 2.63$$

$$\hat{\beta}_3 = -0.032$$

The predicted optimum can be formed by differentiating

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 x + \hat{\beta}_3 x^2$$

with respect to x and setting the derivative to zero to get

$$x^{\text{opt}} = \frac{-\hat{\beta}_2}{2\hat{\beta}_3} = 41.09$$

The predicted yield $\hat{Y}$ at the optimum is 88.8.

Certain assumptions underly least squares computations such as the independence of the unobservable errors ε_i, a constant error variance, and lack of error in the x's (Draper and Smith, 1998). If the model represents the data adequately, the residuals should possess characteristics that agree with these basic assumptions. The analysis of residuals is thus a way of checking that one or more of the assumptions underlying least squares optimization is not violated. For example, if the model fits well, the residuals should be randomly distributed about the value of y predicted by the model. Systematic departures from randomness indicate that the model is unsatisfactory; examination of the patterns formed by the residuals can provide clues about how the model can be improved (Box and Hill, 1967; Draper and Hunter, 1967).

Examinations of plots of the residuals versus $\hat{Y}_i$ or x_i, or a plot of the frequency of the residuals versus the magnitude of the residuals, have been suggested as

numerical or graphical aids to assist in the analysis of residuals. A study of the signs of the residuals (+ or −) and sums of signs can be used. Residual analysis should include

1. Detection of an outlier (an extreme observation).
2. Detection of a trend in the residuals.
3. Detection of an abrupt shift in the level of the experiment (sequential observations).
4. Detection of changes in the error variance (usually assumed to be constant).
5. Examination to ascertain if the residuals are represented by a normal distribution (so that statistical tests can be applied).

When using residuals to determine the adequacy of a model, keep in mind that as more independent variables are added to the model, the residuals may become less informative. Each residual is, in effect, a weighted average of the ε_i's; as more unnecessary x_i's are added to a model, the residuals become more like one another, reflecting an indiscriminate average of all the ε's instead of primarily representing one ε_i. In carrying out the analysis of residuals, you will quickly discover that a graphical presentation of the residuals materially assists in the diagnosis because one aberration, such as a single extreme value, can simultaneously affect several of the numerical tests.

Nonlinear least squares

If a model is nonlinear with respect to the model parameters, then nonlinear least squares rather than linear least squares has to be used to estimate the model coefficients. For example, suppose that experimental data is to be fit by a reaction rate expression of the form $r_A = kC_A^n$. Here r_A is the reaction rate of component A, C_A is the reactant concentration, and k and n are model parameters. This model is *linear* with respect to rate constant k but is *nonlinear* with respect to reaction order n. A general nonlinear model can be written as

$$y = f(x_1, x_2, x_3, \ldots, \beta_1, \beta_2, \beta_3 \ldots) \tag{2.10}$$

where y = the model output
x_j's = model inputs
β_j's = the parameters to be estimated

We still can define a sum-of-squares error criterion (to be minimized) by selecting the parameter set β_j so as to

$$\min_{\beta_j} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \tag{2.11}$$

where Y_i = the ith output measurement
$\hat{Y}_i$ = model prediction corresponding to the ith data point

The estimated coefficients listed for model 2 in Example 2.2 were obtained using nonlinear least squares (Bates and Watts, 1988).

As another example, consider the problem of estimating the gain K and time constants τ_i for first-order and second-order dynamic models based on a measured unit step response of the process $y(t)$. The models for the step response of these two processes are, respectively (Seborg et al., 1989),

$$y(t) = K(1 - e^{-t/\tau_1}) \tag{2.12}$$

$$y(t) = K\left(1 - \frac{\tau_1 e^{-t/\tau_1} - \tau_2 e^{-t/\tau_2}}{\tau_1 - \tau_2}\right) \tag{2.13}$$

where t = the independent variable (time)
y = the dependent variable

Although K appears linearly in both response equations, τ_1 in (2.12) and τ_1 and τ_2 in (2.13) appear nonlinearly, so that nonlinear least squares must be used to estimate their values. The specific details of how to carry out the computations will be deferred until we take up numerical methods of unconstrained optimization in Chapter 6.

2.4 FACTORIAL EXPERIMENTAL DESIGNS

Because variables in models are often highly correlated, when experimental data are collected, the $\mathbf{x}^T\mathbf{x}$ matrix in Equation 2.9 can be badly conditioned (see Appendix A), and thus the estimates of the values of the coefficients in a model can have considerable associated uncertainty. The method of factorial experimental design forces the data to be orthogonal and avoids this problem. This method allows you to determine the relative importance of each input variable and thus to develop a parsimonious model, one that includes only the most important variables and effects. Factorial experiments also represent efficient experimentation. You systematically plan and conduct experiments in which all of the variables are changed simultaneously rather than one at a time, thus reducing the number of experiments needed.

Because of the orthogonality property of factorial design, statistical tests are effective in discriminating among the effects of natural variations in raw materials, replicated unit operations (e.g., equipment in parallel), different operators, different batches, and other environmental factors. A proper orthogonal design matrix for collecting data provides independent estimates of the sums of squares for each variable as well as combinations of variables. Also the estimates of the coefficients have a lower variance than can be obtained with a nonorthogonal experimental design (Montgomery, 1997; Box et al., 1978). That is, you can have more confidence in the values calculated for β_i than would occur with a nonorthogonal design.

TABLE 2.1
Orthogonal experimental design

Experiment number	Response	Scaled (coded) values of the independent variables	
	y	z_1	z_2
1	Y_1	−1	−1
2	Y_2	1	−1
3	Y_3	−1	1
4	Y_4	1	1
5	Y_5	0	0

From a practical standpoint, the user of the model must decide which input variables should be studied because this will determine the number of tests that must be carried out (Drain, 1997). In a standard factorial design, 2^n tests are required, where n is the number of input variables to be studied. You must also decide how much each input variable should be changed from its nominal value, taking into account the sensitivity of the process response to a change in a given input variable, as well as the typical operating range of the process. The determination of the region of experimentation requires process knowledge. The experimental range should be chosen so that the resulting measurements of the response do not involve errors in the sensors that are greater than typical noise levels.

Suppose you want to fit the linear model $y = \beta_1 + \beta_2 z_1 + \beta_3 z_2$, where z_1 and z_2 are the independent variables. Let the values of z_1 and z_2 in the experiment be deliberately chosen by an *experimental orthogonal design* like that shown in Table 2.1.

The values of the coded independent variables correspond to the four corners of a square in the z_1 and z_2 space. The summations in Equation (2.5) simplify in this case ($x_1 = 1, x_2 = z_1, x_3 = z_2$):

$$\sum_{i=1}^{5} x_{i1}x_{i2} = \sum_{i=1}^{5} z_{1i} = 0 \quad \sum_{i=1}^{5} x_{i1}x_{i3} = \sum_{i=1}^{5} z_{2i} = 0 \quad \sum_{i=1}^{5} x_{i2}x_{i3} = \sum_{i=1}^{5} z_{1i}z_{2i} = 0$$

$$\sum_{i=1}^{5} x_{i1}x_{i1} = 5 \quad \sum_{i=1}^{5} x_{i2}x_{i2} = \sum_{i=1}^{5} z_{1i}^2 = 4 \quad \sum_{i=1}^{5} x_{i3}x_{i3} = \sum_{i=1}^{5} z_{2i}^2 = 4$$

For the experimental design in Table 2.1,

$$\mathbf{x}^T\mathbf{x} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$$\mathbf{x}^T\mathbf{Y} = \begin{bmatrix} y_1 + y_2 + y_3 + y_4 + y_5 \\ -y_1 + y_2 - y_3 + y_4 \\ -y_1 - y_2 + y_3 + y_4 \end{bmatrix}$$

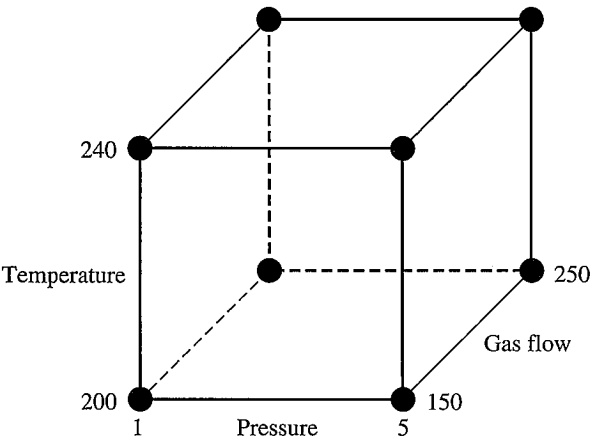


FIGURE E2.6
Orthogonal design for the variables temperature, pressure, and flowrate.

It is quite easy to solve Equation (2.9) now because these expressions are *uncoupled*; the inverse of $\mathbf{x}^T\mathbf{x}$ for Equation (2.13) can be obtained by merely taking the reciprocal of the diagonal elements.

EXAMPLE 2.6 IDENTIFICATION OF IMPORTANT VARIABLES BY EXPERIMENTATION USING AN ORTHOGONAL FACTORIAL DESIGN

Assume a reactor is operating at the reference state of 220°C, 3 atm pressure, and a gas flow rate of 200 kg/h. We can set up an orthogonal factorial design to model this process with a linear model $Y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$ so that the coded values of the x_i are 1, −1, and 0. Examine Figure E2.6. Suppose we select the changes in the operating conditions of ±20°C for the temperature, ±2 atm for the pressure, and ±50 kg/h for flowrates. Let $x_1 = 1$; then x_2, x_3, and x_4, the coded variables, are calculated in terms of the proposed operating conditions as follows:

$$x_2 = \frac{t(°\text{C}) - 220}{20}$$

$$x_3 = \frac{p(\text{atm}) - 3}{2}$$

$$x_4 = \frac{m(\text{kg/h}) - 200}{50}$$

Based on the design the following data are collected:

Y (yield)	x_2	x_3	x_4
20.500	−1	−1	−1
60.141	1	−1	−1
58.890	−1	1	−1
67.712	1	1	−1
22.211	−1	−1	1
61.541	1	−1	1
59.902	−1	1	1
69.104	1	1	1
77.870	0	0	0
78.933	0	0	0
70.100	0	0	0

The extra data at the (0, 0) point are used to obtain a measure of the error involved in the experiment.

Solution. The matrices involved are

$$\mathbf{x}^T\mathbf{x} = \begin{bmatrix} 11 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix} \qquad (\mathbf{x}^T\mathbf{x})^{-1} = \begin{bmatrix} 0.091 & 0 & 0 & 0 \\ 0 & 0.125 & 0 & 0 \\ 0 & 0 & 0.125 & 0 \\ 0 & 0 & 0 & 0.125 \end{bmatrix}$$

$$\mathbf{x}^T\mathbf{Y} = \begin{bmatrix} 646.9 \\ 96.99 \\ 91.21 \\ 5.51 \end{bmatrix}$$

With these matrices you can compute the estimates of $\hat{\beta}_i$ by solving Equation (2.9), yielding

$$\hat{Y} = 58.810 + 12.124x_2 + 11.402x_3 + 0.689x_4$$

In terms of the original variables

$$\hat{Y} = 58.810 + 12.124\left(\frac{t(°\text{C}) - 220}{20}\right) + 11.402\left(\frac{p(\text{atm}) - 3}{2}\right)$$

$$+ 0.689\left(\frac{m(\text{kg/h}) - 200}{50}\right)$$

$$= 58.810 + 0.6062(t - 220) + 5.701(p - 3) + 0.0138(m - 200)$$

It is clear from the size of the estimated coefficients that mass flowrate changes have a much smaller influence on the yield and thus, for practical purposes, could be eliminated as an important independent variable.

If the independent variables are orthogonal, deciding whether to add or delete variables or functions of variables in models is straightforward using stepwise least squares (regression), a feature available on many software packages. Stepwise regression consists of sequentially adding (or deleting) a variable (or function) of variables to a proposed model and then testing at each stage to see if the added (or deleted) variable is significant. The procedure is only effective when the independent variables are essentially orthogonal. The coupling of orthogonal experimental design with optimization of operating conditions has been called "evolutionary operation" by which the best operating conditions are determined by successive experiments (Box and Draper, 1969; Biles and Swain, 1980).

2.5 DEGREES OF FREEDOM

In Section 1.5 we briefly discussed the relationships of equality and inequality constraints in the context of independent and dependent variables. Normally in design and control calculations, it is important to eliminate redundant information and equations before any calculations are performed. Modern multivariable optimization software, however, does not require that the user clearly identify independent, dependent, or superfluous variables, or active or redundant constraints. If the number of independent equations is larger than the number of decision variables, the software informs you that no solution exists because the problem is overspecified. Current codes have incorporated diagnostic tools that permit the user to include all possible variables and constraints in the original problem formulation so that you do not necessarily have to eliminate constraints and variables prior to using the software. Keep in mind, however, that the smaller the dimensionality of the problem introduced into the software, the less time it takes to solve the problem.

The degrees of freedom in a model is the number of variables that can be specified independently and is defined as follows:

$$N_F = N_v - N_E \tag{2.14}$$

where N_F = degrees of freedom
N_v = total number of variables involved in the problem
N_E = number of independent equations (including specifications)

A degrees-of-freedom analysis separates modeling problems into three categories:

1. $N_F = 0$: *The problem is exactly determined.* If $N_F = 0$, then the number of independent equations is equal to the number of process variables and the set of equations may have a unique solution, in which case the problem is not an optimization problem. For a set of linear independent equations, a unique solution exists. If the equations are nonlinear, there may be no real solution or there may be multiple solutions.

2. $N_F > 0$: *The problem is underdetermined.* If $N_F > 0$, then more process variables exist in the problem than independent equations. The process model is said to be underdetermined, so at least one variable can be optimized. For linear models, the rank of the matrix formed by the coefficients indicates the number of independent equations (see Appendix A).
3. $N_F < 0$: *The problem is overdetermined.* If $N_F < 0$, fewer process variables exist in the problem than independent equations, and consequently the set of equations has no solutions. The process model is said to be overdetermined, and least squares optimization or some similar criterion can be used to obtain values of the unknown variables as described in Section 2.5.

EXAMPLE 2.7 MODEL FOR A SEPARATION TRAIN

Figure E2.7 shows the process flow chart for a series of two distillation columns, with mass flows and splits defined by $x_1, x_2, \ldots, x_5$. Write the material balances, and show that the process model comprises two independent variables and three degrees of freedom.

Solution. The balances for columns 1 and 2 are shown below:

Column 1

$$x_1 = x_2 + x_3 \quad \text{or} \quad x_1 - x_2 - x_3 = 0 \tag{a}$$

$$x_2 = .40x_1 \quad \text{or} \quad x_2 - 0.4x_1 = 0 \tag{b}$$

$$x_3 = .60x_1 \quad \text{or} \quad x_3 - 0.6x_1 = 0 \tag{c}$$

There are three equations and three unknowns.

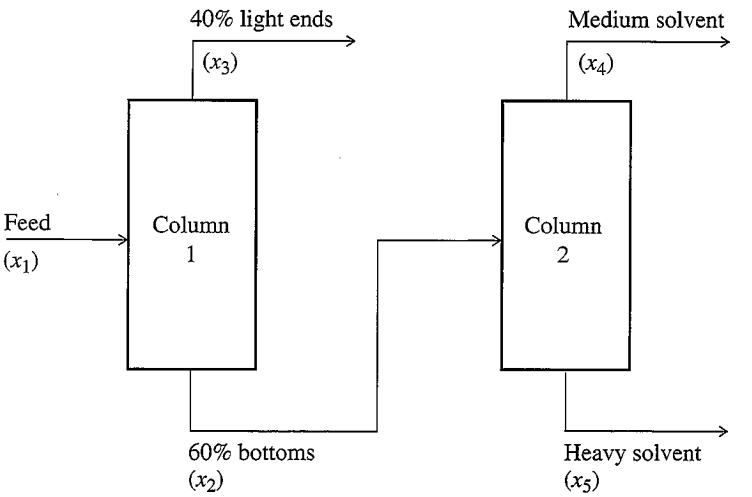


FIGURE E2.7
Train of distillation columns.

The coefficient matrix is

		Variables		
		x_1	x_2	x_3
Equations	(a)	1	−1	−1
	(b)	−0.4	1	0
	(c)	−0.6	0	1

The three equations are not independent. The rank of the coefficient matrix is 2, hence there are only two independent variables, and column 1 involves 1 degree of freedom.

Column 2 $x_2 = x_4 + x_5$ or $x_2 - x_4 - x_5 = 0$ (*d*)

There is one equation and three unknowns, so there are two degrees of freedom.Overall there are four equations (a), (b), (c), (d) and five variables.The coefficient matrix is

	x_1	x_2	x_3	x_4	x_5
(a)	1	−1	−1	0	0
(b)	−0.4	1	0	0	0
(c)	−0.6	0	1	0	0
(d)	0	1	0	−1	−1

Because the rank of the coefficient matrix is three, there are only three independent equations, so Equation (2.14) indicates that there are two degrees of freedom. You can reduce the dimensionality of the set of material balances by substitution of one equation into another and eliminating both variables and equations.

In some problems it is advantageous to eliminate obvious dependent variables to reduce the number of equations that must be included as constraints. You can eliminate linear constraints via direct substitution, leaving only the nonlinear constraints, but the resulting equations may be too complex for this procedure to have merit. The following example illustrates a pipe flow problem in which substitution leads to one independent variable.

EXAMPLE 2.8 ANALYSIS OF PIPE FLOW

Suppose you want to design a hydrocarbon piping system in a plant between two points with no change in elevation and want to select the optimum pipe diameter that minimizes the combination of pipe capital costs and pump operating costs. Prepare a model that can be used to carry out the optimization. Identify the independent and dependent variables that affect the optimum operating conditions. Assume the fluid properties (μ, ρ) are known and constant, and the value of the pipe length (L) and mass flowrate (m) are specified. In your analysis use the following process variables: pipe diameter (D), fluid velocity (v), pressure drop (Δp), friction factor (f).

Solution. Intuitively one expects that an optimum diameter can be found to minimize the total costs. It is clear that the four process variables are related and not indepen-

dent, but we need to examine in an organized way how the equality constraints (models) affect the degrees of freedom.

List the equality constraints:

1. Mechanical energy balance, assuming no losses in fittings, no change in elevation, and so on.

$$\Delta p = \frac{2f\rho v^2 L}{D} \tag{a}$$

2. Equation of continuity, based on plug flow under turbulent conditions.

$$m = \left(\frac{\rho \pi D^2}{4}\right) v \tag{b}$$

3. A correlation relating the friction factor with the Reynolds number (Re).

$$f = f(\text{Re}) = f\left(\frac{Dv\rho}{\mu}\right)$$

The friction factor plot is available in many handbooks, so that given a value of Re, one can find the corresponding value of f. In the context of numerical optimization, however, using a graph is a cumbersome procedure. Because all of the constraints should be expressed as mathematical relations, we select the Blasius correlation for a smooth pipe (Bird et al., 1964):

$$f = 0.046\,\text{Re}^{-0.2} = \frac{0.046\mu^{0.2}}{D^{0.2}v^{0.2}\rho^{0.2}} \tag{c}$$

The model involves four variables and three independent nonlinear algebraic equations, hence one degree of freedom exists. The equality constraints can be manipulated using direct substitution to eliminate all variables except one, say the diameter, which would then represent the independent variables. The other three variables would be dependent. Of course, we could select the velocity as the single independent variable of any of the four variables. See Example 13.1 for use of this model in an optimization problem.

2.6 EXAMPLES OF INEQUALITY AND EQUALITY CONSTRAINTS IN MODELS

As mentioned in Chapter 1, the occurrence of *linear inequality constraints* in industrial processes is quite common. Inequality constraints do not affect the count of the degrees of freedom unless they become active constraints. Examples of such constraints follow:

1. Production limitations arise because of equipment throughput restrictions, storage limitations, or market constraints (no additional product can be sold beyond some specific level).
2. Raw material limitations occur because of limitations in feedstock supplies; these supplies often are determined by production levels of other plants within the same company.
3. Safety or operability restrictions exist because of limitations on allowable operating temperatures, pressures, and flowrates.
4. Physical property specifications on products must be considered. In refineries the vapor pressure or octane level of fuel products must satisfy some specification. For blends of various products, you usually assume that a composite property can be calculated through the averaging of pure component physical properties. For N components with physical property values V_i and volume fraction y_i, the average property $\bar{V}$ is

$$\bar{V} = \sum_{i=1}^{N} V_i y_i$$

EXAMPLE 2.9 FORMULATION OF A LINEAR INEQUALITY CONSTRAINT FOR BLENDING

Suppose three intermediates (light naphtha, heavy naphtha, and "catalytic" oil) made in a refinery are to be blended to produce an aviation fuel. The octane number of the fuel must be at least 95. The octane numbers for the three intermediates are shown in the table.

	Amount blended (barrels/day)	Octane number
Light naphtha	x_1	92
Heavy naphtha	x_2	86
Catalytic oil	x_3	97

Write an inequality constraint for the octane number of the aviation fuel, assuming a linear mixing rule.

Solution. Assume the material balance can be based on conservation of volume (as well as mass). The production rate of aviation gas is $x_4 = x_1 + x_2 + x_3$. The volume-average octane number of the gasoline can be computed as

$$\frac{x_1}{x_1 + x_2 + x_3}(92) + \frac{x_2}{x_1 + x_2 + x_3}(86) + \frac{x_3}{x_1 + x_2 + x_3}(97) \geq 95 \qquad (a)$$

Multiplying Equation (a) by $(x_1 + x_2 + x_3)$ and rearranging, we get

$$-3x_1 - 9x_2 + 2x_3 \geq 0 \qquad (b)$$

This constraint ensures that the octane number specification is satisfied. Note that Equation (b) is linear.

EXAMPLE 2.10 LINEAR MATERIAL BALANCE MODELS

In many cases in which optimization is applied, you need to determine the allocation of material flows to a set of processes in order to maximize profits. Consider the process diagram in Figure E2.10.

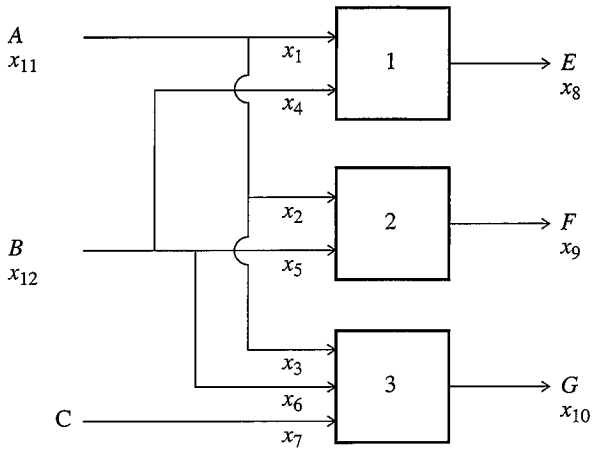


FIGURE E2.10
Flow diagram for a multiproduct plant.

Each product (*E, F, G*) requires different (stoichiometric) amounts of reactants according to the following mass balances:

Product	Reactants (1-kg product)
E	$\frac{2}{3}$ kg *A*, $\frac{1}{3}$ kg *B*
F	$\frac{2}{3}$ kg *A*, $\frac{1}{3}$ kg *B*
G	$\frac{1}{2}$ kg *A*, $\frac{1}{6}$ kg *B*, $\frac{1}{3}$ kg *C*

Prepare a model of the process using the mass balance equations.

Solution. Twelve mass flow variables can be defined for this process. Let x_1, x_2, x_3 be the mass input flows of *A* to each process. Similarly let x_4, x_5, x_6, and x_7 be the individual reactant flows of *B* and *C*, and define x_8, x_9, and x_{10} as the three mass product flows (*E, F, G*). Let x_{11} and x_{12} be the total amounts of *A* and *B* used as reactants (*C* is the same as x_7). Thus, we have a total of 12 variables.

The linear mass balance constraints that represent the process are:

$$A = x_{11} = x_1 + x_2 + x_3 \qquad (a)$$

$$B = x_{12} = x_4 + x_5 + x_6 \qquad (b)$$

$$x_1 = 0.667x_8 \qquad (c)$$

$$x_2 = 0.667x_9 \qquad (d)$$

$$x_3 = 0.5x_{10} \tag{e}$$

$$x_4 = 0.333x_8 \tag{f}$$

$$x_5 = 0.333x_9 \tag{g}$$

$$x_6 = 0.167x_{10} \tag{h}$$

$$x_7 = 0.333x_{10} \tag{i}$$

With 12 variables and 9 independent linear equality constraints, 3 degrees of freedom exist that can be used to maximize profits. Note that we could have added an overall material balance, $x_{11} + x_{12} + x_7 = x_8 + x_9 + x_{10}$, but this would be a redundant equation since it can be derived by adding the material balances.

Other constraints can be specified in this problem. Suppose that the supply of A was limited to 40,000 kg/day, or

$$x_{11} \leq 40{,}000 \tag{j}$$

If this constraint is inactive, that is, the optimum value of x_{11} is less than 40,000 kg/day, then, in effect, there are still 3 degrees of freedom. If, however, the optimization procedure yields a value of $x_{11} = 40{,}000$ (the optimum lies on the constraint, such as shown in Figure 1.2), then inequality constraint f becomes an equality constraint, resulting in only 2 degrees of freedom that can be used for optimization. You should recognize that it is possible to add more inequality constraints, such as constraints on materials supplies, in the model, for example,

$$x_{12} \leq 30{,}000 \tag{k}$$

$$x_7 \leq 25{,}000 \tag{l}$$

These can also become "active" constraints if the optimum lies on the constraint boundary. Note that we can also place inequality constraints on production of E, F, and G in order to satisfy market demand or sales constraints

$$x_8 \geq 20{,}000 \tag{m}$$

$$x_9 \geq 25{,}000 \tag{n}$$

$$x_{10} \geq 30{,}000 \tag{o}$$

Now the analysis is much more complex, and it is clear that more potential equality constraints exist than variables if all of the inequality constraints become active. It is possible that optimization could lead to a situation where no degrees of freedom would be left—one set of the inequality constraints would be satisfied as equalities. This outcome means no variables remain to be optimized, and the optimal solution reached would be at the boundaries, a subset of the inequality constraints.

Other constraints that can be imposed in a realistic problem formulation include

1. Operating limitations (bottlenecks)—there could be a throughput limitation on reactants to one of the processes (e.g., available pressure head).
2. Environmental limitations—there could be some additional undesirable by-products H, such as the production of toxic materials (not in the original product list given earlier), that could contribute to hazardous conditions.

You can see that the model for a realistic process can become extremely complex; what is important to remember is that steps 1 and 3 in Table 1.1 provide an organized framework for identifying all of the variables and formulating the objective function, equality constraints, and inequality constraints. After this is done, you need not eliminate redundant variables or equations. The computer software can usually handle redundant relations (but not inconsistent ones).

REFERENCES

Bates, D. M.; and D. G. Watts. *Nonlinear Regression Analysis and Its Applications*. Wiley, New York (1988).

Biles, W. E.; and J. J. Swain. *Optimization and Industrial Experimentation*. Wiley-Interscience, New York (1980).

Bird, R. B.; W. E. Stewart; and E. N. Lightfoot. *Transport Phenomena*. Wiley, New York (1964).

Box, G. E. P.; and J. W. Hill. "Discrimination Among Mechanistic Models." *Technometrics* **9:** 57 (1967)

Box, G. E. P.; and N. R. Draper. *Evolutionary Operation*. Wiley, New York (1969).

Box, G. E. P.; W. G. Hunter, and J. S. Hunter. *Statistics for Experimenters*. Wiley-Interscience, New York (1978).

Deitz, D. "Modeling Furnace Performance." *Mech Eng* **119** (Dec), 16 (1997).

Drain, D. *Handbook of Experimental Methods for Process Improvements*. Chapman and Hall, International Thomson Publishing, New York (1997).

Draper, N. R.; and W. G. Hunter. "The Use of Prior Distribution in the Design of Experiments for Parameter Estimation in Nonlinear Situations." *Biometrika* **54:** 147 (1967).

Draper, N. R.; and H. Smith. *Applied Regression Analysis*, 3rd ed. Wiley, New York (1998).

Eykhoff, P. *System Identification*. Wiley-Interscience, New York (1974).

Jackson, J. E. *A User's Guide to Principal Components*. Wiley, New York (1991).

McAdams, W. H. *Heat Transmission*. McGraw-Hill, New York (1954).

Montgomery, D. C. *Design and Analysis of Experiments*, 4th ed. Wiley, New York (1997).

Seborg, D. E.; T. F. Edgar; and D. A. Mellichamp. *Process Dynamics and Control*. Wiley, New York (1989).

Shahbenderian, A. P. "The Application of Statistics to Cost Estimating." *Br Chem Eng* (January) **6:** 16 (1961).

SUPPLEMENTARY REFERENCES

Bendor, E. A. *An Introduction to Mathematical Modeling*. John Wiley, New York (1978).

Bequette, B. W. *Process Dynamics: Modeling, Analysis, and Simulation*. Prentice-Hall, Englewood Cliffs, NJ (1998).

Churchill, S. W. *The Interpretation and Use of Rate Data*. Scripta Publishing Company, Washington, D. C. (1974).

Davis, M. E. *Numerical Methods and Modeling for Chemical Engineers*. John Wiley, New York (1983).

Denn, M. M. *Process Modeling*. Pitman Publishing, Marshfield, MA (1986).

Friedly, J. C. *Dynamic Behavior of Processes*. Prentice-Hall, Englewood Cliffs, New Jersey (1972).

Luyben, W. L. *Process Modeling, Simulation, and Control for Chemical Engineers.* McGraw-Hill, New York (1990).
Montgomery, D. C.; G. C. Runger; and N. F. Hubele. *Engineering Statistcs.* Wiley, New York (1998).
Ogunnaike, T.; and W. H. Ray. *Process Dynamics, Modeling and Control.* Oxford University Press, New York (1994).
Rice, R. G.; and D. D. Duong. *Applied Mathematics and Modeling for Chemical Engineers.* Wiley, New York (1995).
Seider, W. D.; J. D. Seader; and D. R. Lewin. *Process Design Priciples.* Wiley, New York (1999).
Seinfeld, J. H.; and L. Lapidus. *Process Modeling, Estimation, and Identification.* Prentice-Hall, Englewood Cliffs, New Jersey (1974).
Wen, C. Y.; and L. T. Fan. *Models for Flow Systems and Chemical Reactors.* Marcel Dekker, New York (1975).

PROBLEMS

2.1 Classify the following models as linear or nonlinear

(a) Two-pipe heat exchanger (streams 1 and 2)

$$\frac{\partial T_1}{\partial t} + v\frac{\partial T_1}{\partial z} = \frac{2h_1}{S_1\rho_1 C_{p1}}(T_2 - T_1)$$

$$\frac{\partial T_2}{\partial t} = \frac{2h_1}{\rho_2 C_{p2} S_2}(T_2 - T_1)$$

$$BC\text{: } T_1(t,0) = a \qquad IC\text{: } T_1(0,z) = 0$$

$$T_2(t,0) = b \qquad T_2(0,z) = T_0$$

where T = temperature, C_p = heat capacity, t = time, S = area factor, BC = boundary conditions, IC = initial conditions, ρ = *density*

(b) Diffusion in a cylinder

$$\frac{\partial C}{\partial t} = D\left(\frac{\partial^2 C}{\partial r^2} + \frac{1}{r}\frac{\partial C}{\partial r}\right)$$

$$C(0, r) = C_0$$

$$\frac{\partial C(t,0)}{\partial r} = 0$$

$$C(t, R) = C_0$$

where C = concentration r = radial direction
t = time D = constant

2.2 Classify the following equations as linear or nonlinear (y = dependent variable; x, z = independent variables)

(a) $y_1^2 + y_2^2 = a^2$

(b) $v_x \dfrac{\partial v_y}{\partial x} = \mu \dfrac{\partial^2 v_y}{\partial z^2}$

2.3 Classify the models in Problems 2.1 and 2.2 as steady state or unsteady state.

2.4 Classify the models in Problems 2.1 and 2.2 as lumped or distributed.

2.5 What type of model would you use to represent the process shown in the figure? Lumped or distributed? Steady state or unsteady state? Linear or nonlinear?

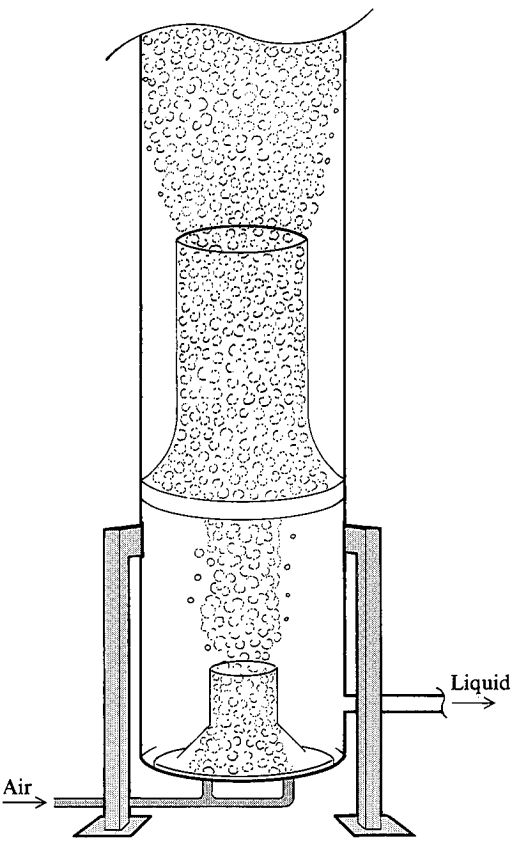


FIGURE P2.5
A wastewater treatment system uses five stacked venturi sections to ensure maximum oxygenation efficiency.

2.6 Determine the number of independent variables, the number of independent equations, and the number of degrees of freedom for the reboiler shown in the figure. What variables should be specified to make the solution of the material and energy balances determinate? (Q = heat transferred)

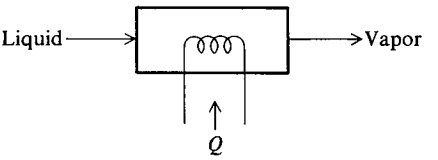


Figure P2.6

2.7 Determine the best functional relation to fit the following data sets:

(a)		(b)		(c)		(d)	
x_i	Y_i	x_i	Y_i	x_i	Y_i	x_i	Y_i
1	5	2	94.8	2	0.0245	0	8290
2	7	5	87.9	4	0.0370	20	8253
3	9	8	81.3	8	0.0570	40	8215
4	11	11	74.9	16	0.0855	60	8176
		14	68.7	32	0.1295	80	8136
		17	64.0	64	0.2000	100	8093
				128	0.3035		

2.8 The following data have been collected:

x_i	Y_i
10	1.0
20	1.26
30	1.86
40	3.31
50	7.08

Which of the following three models best represents the relationship between Y and x?

$$y = e^{\alpha+\beta x}$$

$$y = e^{\alpha+\beta_1 x+\beta_2 x^2}$$

$$y = \alpha x^{\beta}$$

2.9 Given the following equilibrium data for the distribution of SO_3 in hexane, determine a suitable linear (in the parameters) empirical model to represent the data.

x_i pressure (psia)	Y_i weight fraction hexane
200	0.846
400	0.573
600	0.401
800	0.288
1000	0.209
1200	0.153
1400	0.111
1600	0.078

2.10 (a) Suppose that you wished to curve fit a set of data (shown in the table) with the equation

$$y = c_0 + c_1 e^{3x} + c_2 e^{-3x}$$

x_i	Y_i
0	1
1	2
2	2
3	1

Calculate c_0, c_1, and c_2 (show what summations need to be calculated). How do you find c_1 and c_2 if c_0 is set equal to zero?

(b) If the desired equation were $y = a_1 x e^{-a_2 x}$, how could you use least-squares to find a_1 and a_2?

2.11 Fit the following data using the least squares method with the equation:

$$y = c_0 + c_1 x$$

x_i	Y_i
0.5	0.6
1.0	1.4
2.1	2.0
3.4	3.6

Compare the results with a graphical (visual) estimate.

2.12 Fit the same data in Problem 2.11 using a quadratic fit. Repeat for a cubic model ($y = c_0 + c_1 x + c_2 x^2 + c_3 x^3$). Plot the data and the curves.

2.13 You are asked to get the best estimates of the coefficients b_0, b_1, and c in the following model

$$y = b_0 + b_1 e^{-cx}$$

given the following data.

Y_i	x_i
51.6	0.4
53.4	1.4
20.0	5.4
−4.2	19.5
−3.0	48.2
−4.8	95.9

Explain step by step how you would get the values of the coefficients.

2.14 Fit the following function for the density ρ as a function of concentration C, that is, determine the value of α in

$$\rho = \alpha + 1.33C$$

given the following measurements for ρ and C:

ρ (g/cm^3)	C (gmol/L)
3.31	1.01
4.69	1.97
5.92	3.11
7.35	4.00
8.67	4.95

2.15 (a) For the given data, fit a quadratic function of y versus x by estimating the values of all the coefficients.
(b) Does this set of data constitute an orthogonal design?

y	6.4	5.6	6.0	7.5	6.5	8.3	7.7	11.7	10.3	17.6	18.0
x	1.0	1.0	1.0	2.0	2.0	3.0	3.0	4.0	4.0	5.0	5.0

2.16 Data obtained from a preset series of experiments was

Temperature, T (°F)	**Pressure, p (atm)**	**Yield, Y (%)**
160	1	4
160	1	5
160	7	10
160	7	11
200	1	24
200	1	26
200	7	35
200	7	38

Fit the linear model $\hat{Y} = b_0 + b_1x_1 + b_2x_2$ using the preceding table. Report the estimated coefficients b_0, b_1, and b_2. Was the set of experiments a factorial design?

2.17 You are given data for Y versus x and asked to fit an empirical model of the form:

$$y = \alpha + \beta x$$

where β is a *known* value. Give an equation to calculate the best estimate of α.

2.18 A replicated two-level factorial experiment is carried out as follows (the dependent variables are yields):

Time (h)	Temperature (°C)	Yield (%)
1	240	24
5	240	42
1	280	3
5	280	19
1	240	24
5	240	46
1	280	5
5	280	21

Find the coefficients in a first-order model, $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$. ($Y$ = yield, x_1 = time, x_2 = temperature.)

2.19 An experiment based on a hexagon design was carried out with four replications at the origin, producing the following data:

Factor levels			Design levels	
Yield (%)	**Temperature (°C)**	**Time (h)**	x_1	x_2
96.0	75	2.0	1.000	0
78.7	60	2.866	0.500	0.866
76.7	30	2.866	−0.500	0.866
54.6	15	2.0	−1.000	0
64.8	30	1.134	−0.500	−0.866
78.9	60	1.134	0.500	−0.866
97.4	45	2.0	0	0
90.5	45	2.0	0	0
93.0	45	2.0	0	0
86.3	45	2.0	0	0
Coding:	$x_1 = \dfrac{\text{temperature} - 45}{30}$		$x_2 = \text{time} - 2$	

Fit the full second-order (quadratic) model to the data.

2.20 A reactor converts an organic compound to product P by heating the material in the presence of an additive A. The additive can be injected into the reactor, and steam can be injected into a heating coil inside the reactor to provide heat. Some conversion can be obtained by heating without addition of A, and vice versa. In order to predict the yield of P, Y_p (lb mole product per lb mole feed), as a function of the mole fraction of A, X_A, and the steam addition S (in lb/lb mole feed), the following data were obtained.

Y_p	X_A	S
0.2	0.3	0
0.3	0	30
0.5	0	60

(a) Fit a linear model

$$Y_p = c_0 + c_1 X_A + c_2 S$$

that provides a least squares fit to the data.

(b) If we require that the model always must fit the point $Y_p = 0$ for $X_A = S = 0$, calculate c_0, c_1, and c_2 so that a least squares fit is obtained.

2.21 If you add a feed stream to the equilibrium stage shown in the figure, determine the number of degrees of freedom for a binary mixture (Q = heat transferred).

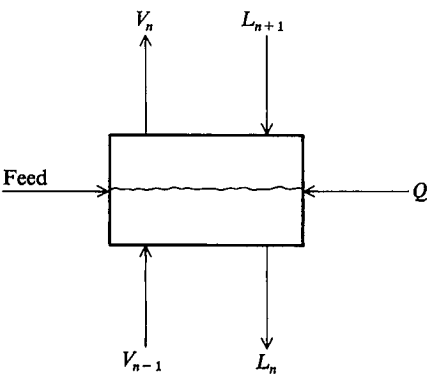


FIGURE P2.21

2.22 How many variables should be selected as independent variables for the furnace shown in the figure?

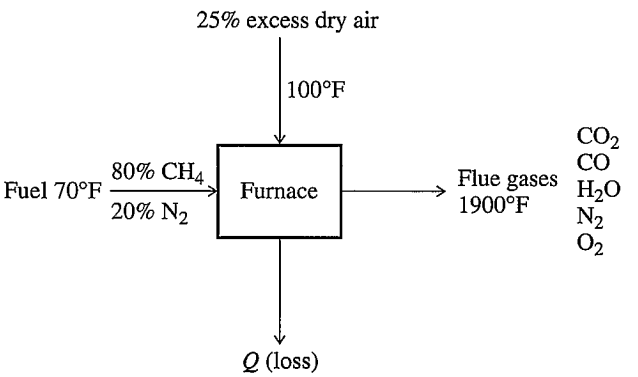


FIGURE P2.22

2.23 Determine the number of independent variables, the number of independent equations, and the number of degrees of freedom in the following process (A, B, and D are chemical species):

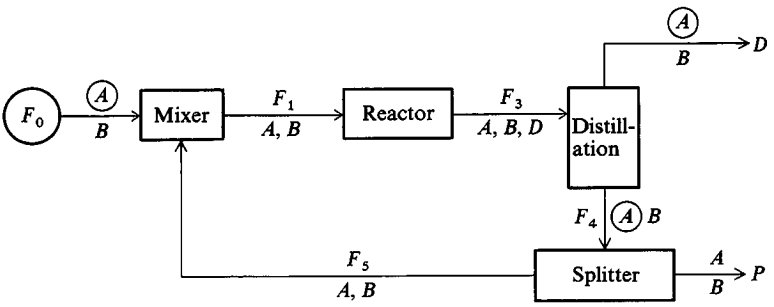


FIGURE P2.23

The encircled variables have known values. The reaction parameters in the reactor are known as the fraction split at the splitter between F_4 and F_5. Each stream is a single phase.

2.24 A waste heat boiler (see Fig. P2.24) is to be designed for steady-state operation under the following specifications.

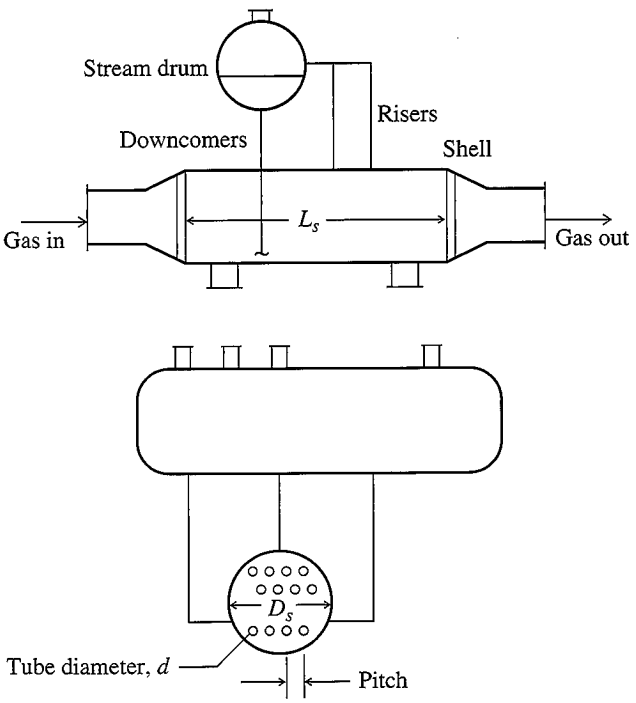


FIGURE P2.24

Total gas flow	25,000 kg/h
Gas composition	SO_2 (9%), O_2 (12%), N_2 (79%)
Gas temperatures	in = 1200°C; out = 350°C
Stream pressure outside tubes	250 kPa
Gas properties	C_p = 0.24 kcal/(g)(°C) μ = 0.14 kg/(m)(h) k = 0.053 kcal/(m)(h)(°C)

Cost data are

Shell	\$2.50/kg
Tubes	\$150/m^2
Electricity	\$0.60/kWh
Interest rate	14%

Base the optimization on just the cost of the shell, tubes, and pumping costs for the gas. Ignore maintenance and repairs.

Formulate the optimization problem using only the following notation (as needed):

A	surface area of tubes, m^2
C_s	cost of shell, \$
C_t	cost of tubes, \$
C_{pi}	heat capacity of gas, kcal/(kg)(°C)
D	diameter of shell, m
d_o, d_i	tube outer and inner diameters, m
f	friction factor
g	acceleration due to gravity, m/s^2
h_i	gas side heat transfer coefficient inside the tubes, kcal/(m^2)(h)(°C)
i	interest rate, fraction
k	gas thermal conductivity, kcal/(m)(h)(°C)
L_s	length of shell, m
MW	molecular weight of gas
n	number of tubes
N	life of equipment, years
Q	duty of the boiler, kcal/h
T_1, T_2	gas temperature entering and leaving the boiler, °C
T	temperature in general
ρ_g	density of gas, kg/m^3
μ_g	viscosity of gas, kg/(m)(h)
V	gas velocity, m/s
W_g	gas flow, kg/h
W_s	weight of shell, tons
η	efficiency of blower
ΔP_g	gas pressure drop, kPa
Z	shell thickness, m

How many degrees of freedom are in the problem you formulated?