

## Chapter 2

### CLASSICAL THEORY OF MAXIMA AND MINIMA

#### Introduction

The classical theory of maxima and minima (analytical methods) is concerned with finding the maxima or minima, i.e., extreme points of a function. We seek to determine the values of the  $n$  independent variables  $x_1, x_2 \dots x_n$  of a function where it reaches maxima and minima points. Before starting with the development of the mathematics to locate these extreme points of a function, let us examine the surface of a function of two independent variables,  $y(x_1, x_2)$ , that could represent the economic model of a process. This should help visualize the location of the extreme points. An economic model is illustrated in Figure 2-1a where the contours of the function are represented by the curved lines. A cross section of the function along line S through the points A and B is shown in Figure 2-1(b), and in Figure 2-1(c) the first derivative of  $y(x_1, x_2)$  along line S through points A and B is given.

In this example, point A is the global maximum in the region and is located at the top of a sharp ridge. Here the first derivative is discontinuous. A second but smaller maximum is located at point B (a local maximum). At point B the first partial derivatives of  $y(x_1, x_2)$  are zero, and B is called a stationary point. It is not necessary for stationary points to be maxima or minima as illustrated by stationary point C, a saddle point. In this example, the minima do not occur in the interior of the region but on the boundary at points D and E (local minima). To determine the global minima, it is necessary to compare the value of the function at these points.

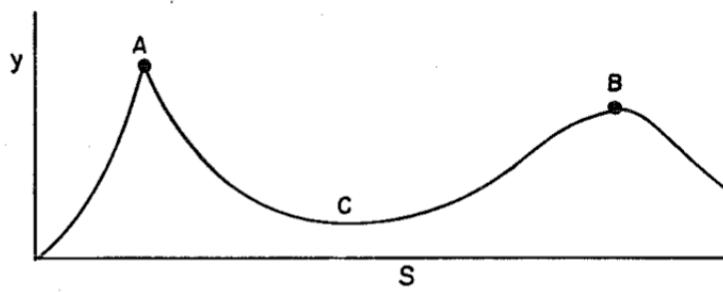
In essence, the problem of determining the maximum profit or minimum cost for a system using the classical theory becomes one of locating all of the local maxima or minima, and then comparing the individual values, to determine the global maximum or minimum. The example has illustrated the places to look that are:

1. at stationary points (first derivatives are zero)
2. on the boundaries
3. at discontinuities in the first derivative

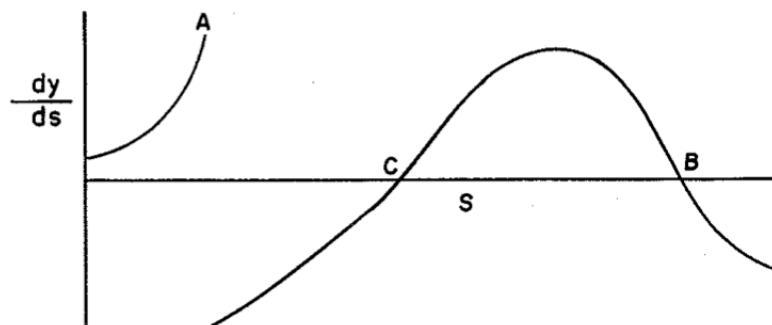
When the function and its derivatives are continuous, the local extreme points will occur at stationary points in the interior of the region. However, it is not necessary that all stationary points be local extreme points since saddle points can occur, also.

$x_2$  $D$  $1$  $2$  $6$  $7$  $8$  $A$  $5$  $4$  $S$  $3$  $C$  $4$  $B$  $5$  $6$  $E$  $x_1$ 

a) Topological Map



b) Cross Section of  $y(x_1, x_2)$  Along Line S  
Through Points A and B



c) First Derivative of  $y$  Along Line S  
Through Points A and B

Figure 2-1 Geometric Interpretation of Function  $y(x_1, x_2)$

## Locating Local Maxima and Minima (Necessary Conditions)

Using geometric intuition from the previous example, we can understand the famous Weierstrass theorem (11, 12), which guarantees the existence of maxima and minima. It states:

*Every function that is continuous in a closed domain possesses a maximum and a minimum value either in the interior or on the boundary of the domain.*

The proof is by contradiction.

There is another theorem (13) which tells how to locate extreme points in the interior of a region of a continuous function. It states:

*A continuous function of  $n$  variables attains a maximum or a minimum in the interior of a region, only at those values of the variables for which the  $n$  partial derivatives either vanish simultaneously (stationary points) or at which one or more of these derivatives cease to exist (i.e., are discontinuous).*

The proof involves examining the Taylor Series expansion at the points where the partial derivatives either vanish or cease to exist.

Thus, the problem becomes one of locating points where the partial derivatives are zero or where some of them are discontinuous. The stationary points can be located by solving the algebraic equations which result in setting the partial derivatives of the function equal to zero. Also, these algebraic equations must be examined for points of discontinuities, and this has to be accomplished by inspection.

## Evaluating Local Maxima and Minima (Sufficient Conditions)

As we have seen, it is not necessary for all stationary points to be local maxima and minima, since there is a possibility of saddle or inflection points. Now we need to develop procedures to determine if stationary points are maxima or minima. These sufficient conditions will be developed for one independent variable first and then extended for two and  $n$  independent variables, using the same concepts. Once the local maxima and minima are located, it is necessary to compare the individual points to locate the global maximum and minimum.

### Sufficient Conditions for One Independent Variable

To develop criteria establishing whether a stationary point is a local maximum or minimum, we begin by performing a Taylor series expansion about the stationary point  $x_0$ .

$$y(x) = y(x_0) + y'(x_0)(x - x_0) + \frac{1}{2} y''(x_0)(x - x_0)^2 + \text{higher order terms} \quad (2-1)$$

Now, select  $x$  sufficiently close to  $x_0$  so the higher order terms become negligible compared to the second order terms. Since the first derivative is zero at the stationary point, the above equation becomes

$$y(x) = y(x_0) + \frac{1}{2} y''(x_0) (x - x_0)^2 \quad (2-2)$$

We can determine if  $x_0$  is a local maximum or minimum by examining the value of  $y''(x_0)$  since  $(x - x_0)^2$  is always positive. If  $y''(x_0)$  is positive, then the terms  $\frac{1}{2}y''(x_0) (x - x_0)^2$  will always add to  $y(x_0)$  in Equation (2-2) for  $x$  taking on values that are less than or greater than  $x_0$ . For this case  $y(x_0)$  is a local minimum. This is summarized in the following:

If  $y''(x_0) > 0$  then  $y(x_0)$  is a minimum

$y''(x_0) < 0$   $y(x_0)$  is a maximum

$y''(x_0) = 0$  no statement can be made

If the second derivative is zero, it is necessary to examine higher order derivatives. In general if  $y''(x_0) = \dots = y^{(n-1)}(x_0) = 0$ , the Taylor series expansion becomes:

$$y(x) = y(x_0) + (1/n!) y^{(n)}(x_0) (x - x_0)^n \quad (2-3)$$

If  $n$  is even, then  $(x - x_0)^n$  is always positive, and the result is:

If  $y^{(n)}(x_0) > 0$  then  $y(x_0)$  is a minimum

$y^{(n)}(x_0) < 0$   $y(x_0)$  is a maximum

If  $n$  is odd, then  $(x - x_0)^n$  changes sign as  $x$  moves from  $x < x_0$  to  $x > x_0$ , and thus there is an inflection point. These results can be summarized in the following theorem (1).

*If at a stationary point the first and possibly some of the higher derivatives vanish, then the point is or is not an extreme point, according as the first non-vanishing derivative is of even or odd order. If it is even, there is a maximum or minimum according as the derivative is negative or positive.*

The proof of this theorem follows the discussion given above. The following example illustrates the principles discussed.

### Example 2-1

Locate the extreme points of the following two functions:

a.  $y(x) = x^4/4 - x^2/2$

$$y'(x) = x^3 - x = x(x^2 - 1) = x(x - 1)(x + 1) = 0$$

Stationary points are  $x = (0, 1, -1)$

$$y''(x) = 3x^2 - 1$$

$$\begin{aligned} y''(0) &= -1 & \text{maximum} \\ y''(1) &= 2 & \text{minimum} \\ y''(-1) &= 2 & \text{minimum} \end{aligned}$$

b.  $y(x) = x^5$

$$\begin{aligned} y'(x) &= 5x^4 = 0 & \text{stationary point is } x = 0 \\ y''(x) &= 20x^3 & y''(0) = 0 \\ y'''(x) &= 60x^2 & y'''(0) = 0 & \text{no statement can be made} \\ y^{iv}(x) &= 120x & y^{iv}(0) = 0 \\ y^v(x) &= 120 & y^v(0) = 120 & n \text{ is odd, and the stationary} \\ & & & \text{point is an inflection point.} \end{aligned}$$

### Sufficient Conditions for Two Independent Variables

To develop the criteria for a local maximum or minimum for  $x_0$  ( $x_{10}, x_{20}$ ), a stationary point for a function of two variables, a Taylor's series expansion is made about this point.

$$\begin{aligned} y(x_1, x_2) &= y(x_{10}, x_{20}) + y_{x_1}(x_1 - x_{10}) + y_{x_2}(x_2 - x_{20}) \\ &+ \frac{1}{2}[y_{x_1 x_1}(x_1 - x_{10})^2 + 2y_{x_1 x_2}(x_1 - x_{10})(x_2 - x_{20}) \\ &+ y_{x_2 x_2}(x_2 - x_{20})^2] + \text{higher order terms} \end{aligned} \quad (2-4)$$

where the subscripts  $x_1$  and  $x_2$  indicate partial differentiation with respect to those variables and evaluation at the stationary point.

Again we select  $y(x_1, x_2)$  sufficiently close to  $y(x_{10}, x_{20})$ , so the higher order terms become negligible compared to the second-order terms. Also, the first derivatives are zero at the stationary point. Thus, Equation (2-4) can be written in matrix form as:

$$\begin{aligned} y(x_1, x_2) &= y(x_{10}, x_{20}) + y_{x_1}(x_1 - x_{10}) + y_{x_2}(x_2 - x_{20}) \\ &+ \frac{1}{2}(x_1 - x_{10})(x_2 - x_{20}) \begin{bmatrix} y_{x_1 x_1} & y_{x_2 x_1} \\ y_{x_2 x_1} & y_{x_2 x_2} \end{bmatrix} \begin{bmatrix} (x_1 - x_{10}) \\ (x_2 - x_{20}) \end{bmatrix} \end{aligned} \quad (2-5)$$

In matrix-vector notation, Equation 2-5 can be written as:

$$y(\mathbf{x}) = y(\mathbf{x}_0) + \frac{1}{2}[(\mathbf{x} - \mathbf{x}_0)\mathbf{H}_0(\mathbf{x} - \mathbf{x}_0)] \quad (2-6)$$

where  $\mathbf{H}_0$  is the matrix of second partial derivatives evaluated at the stationary point  $\mathbf{x}_0$  and is called the Hessian matrix.

The term in the bracket of Equation (2-6) is called a differential quadratic form, and  $y(x_0)$  will be a minimum or a maximum accordingly if this term is always positive or always negative. Based on this concept, it can be shown (1) that if the following results apply,  $x_0$  is a maximum or a minimum. If they do not hold,  $x_0$  could be a saddle point and is not a maximum or a minimum.

$$y(x_0) \text{ is a minimum if } y_{x_1 x_2} > 0 \text{ and } \begin{vmatrix} y_{x_1 x_1} & y_{x_1 x_2} \\ y_{x_2 x_1} & y_{x_2 x_2} \end{vmatrix} > 0$$

$$y(x_0) \text{ is a maximum if } y_{x_1 x_2} < 0 \text{ and } \begin{vmatrix} y_{x_1 x_1} & y_{x_1 x_2} \\ y_{x_2 x_1} & y_{x_2 x_2} \end{vmatrix} > 0$$

An illustration of the above results is given in Example 2-2. The term in the bracket of Equation (2-6) is an example of a quadratic form. It will be necessary to describe a quadratic form briefly before giving the sufficient conditions for maxima and minima for  $n$  independent variables.

### Sign of a Quadratic Form

To perform a similar analysis for a function with more than two independent variables, it is necessary to determine what is called the sign of the quadratic form. The general quadratic form (1) is written as:

$$Q(A, x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = x^T A x \quad (2-7)$$

where  $a_{ij}$  are the components of symmetric matrix  $A$ , and  $a_{ij} = a_{ji}$ .

It turns out (1) that we can determine if  $Q$  is always positive or always negative, for all finite values of  $x_i$  and  $x_j$ , by evaluating the signs of  $D_i$ , the determinants of the principal sub-matrices of  $A$ :

$$D_i = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1i} \\ a_{21} & a_{22} & \dots & a_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ii} \end{vmatrix} \quad (2-8)$$

The important results that will be used subsequently are:

If  $D_i > 0$  for  $i = 1, 2, \dots, n$ , then:  $A$  is positive definite, and  $Q(A, x) > 0$

If  $D_i < 0$  for  $i = 1, 2, \dots, n$ , then:  $A$  is negative definite and  $Q(A, x) < 0$

If  $D_i$  is neither of these, then  $Q(A, x)$  depends on the values of  $x_i$  and  $x_j$ .

## Sufficient Conditions for n Independent Variables

The result of the previous two sections can be extended to the case of n independent variables by considering the Taylor series expansion for n independent variables around stationary point  $x_0$ :

$$y(x) = y(x_0) + \sum_{i=1}^n y_{x_i}(x_0)(x_i - x_{i0}) + \frac{1}{2} \left( \sum_{j=1}^n \sum_{k=1}^n y_{x_j x_k}(x_j - x_{j0})(x_k - x_{k0}) \right) + \text{higher order terms} \quad (2-9)$$

Again, select  $x$  sufficiently close to  $x_0$  so the higher order terms become negligible compared to the second-order terms. Also, the first derivatives are zero at the stationary point. Thus, Equation (2-9) can be written in matrix-vector notation as:

$$y(\mathbf{x}) = y(\mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}_0 (\mathbf{x} - \mathbf{x}_0) \quad (2-10)$$

where  $\mathbf{x}$  is the column vector of independent variables and  $\mathbf{H}_0$ , the matrix of second partial derivative evaluated at the stationary point  $\mathbf{x}_0$ , in the Hessian matrix. This is the same equation as Equation (2-6) which was written for two independent variables.

The second term on the right hand side of Equation (2-10) is called a differential quadratic form as shown below

$$Q[\mathbf{H}_0, (\mathbf{x} - \mathbf{x}_0)] = (\mathbf{x} - \mathbf{x}_0)^T \mathbf{H}_0 (\mathbf{x} - \mathbf{x}_0) \quad (2-11)$$

Equation (2-11) corresponds to Equation (2-7) in the previous section, and the determinants of the principal sub-matrices of  $\mathbf{H}_0$  as defined below correspond to Equation (2-8).

$$H_{i0} = \begin{vmatrix} y_{x_1 x_1} & y_{x_1 x_2} & \cdots & y_{x_1 x_i} \\ y_{x_2 x_1} & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ y_{x_i x_1} & \ddots & \ddots & y_{x_i x_i} \end{vmatrix} \quad (2-12)$$

We can now use the same procedure in evaluating the character of the stationary points for n independent variables. For example, if the term containing the Hessian matrix is always positive for perturbations of the independent variables around the stationary point, then the stationary point is a local minimum. For this differential quadratic form to be positive always, then  $H_{i0} > 0$  for  $i = 1, 2, \dots, n$ . The same reasoning can be applied for a local maximum, and the results for these two cases are summarized below.

$y(x_0)$  is a minimum if  $H_{i0} > 0$  for  $i = 1, 2, \dots, n$

$y(x_0)$  is a maximum if  $H_{i0} < 0$  for  $i = 1, 3, 5, \dots$   
 $H_{i0} > 0$  for  $i = 2, 4, 6, \dots$

If zeros occur in the place of some of the positive or negative number in the tests above (semi-definite quadratic form), then there is insufficient information to determine the character of the stationary point (1). As discussed in Avriel (10) higher order terms may have to be examined, or local exploration can be performed. If the test is not met (indefinite quadratic form), then the point is neither a maximum nor minimum (1). The following theorem from Cooper (7) summarizes these results. It states:

*If  $y(x)$  and its first two partial derivatives are continuous, then a sufficient condition for  $y(x)$  to have a relative minimum (or maximum) at  $x_0$ , when  $\partial y(x_0)/\partial x_j = 0, j = 1, 2, \dots, n$ , is that Hessian matrix be positive definite (negative definite).*

The proof of this theorem employs arguments similar to those given above.

The following example illustrates these methods.

### Example 2-2

The flow diagram of a simple process is shown in Figure 2-2 (2) where the hydrocarbon feed is mixed with recycle and compressed before being passed into a catalytic reactor. The product and unreacted material are separated by distillation, and the unreacted material is recycled. The pressure,  $P$ , in psi and recycle ratio,  $R$ , must be selected to minimize the total annual cost for the required production rate of 107 pounds per year. The feed is brought up to pressure at an annual cost of  $\$1000P$ , mixed with the recycle stream and fed to the reactor at an annual cost of  $\$4 \times 10^9/PR$ . The product is removed in a separator at a cost of  $\$10^5R$  per year and the unreacted material is recycled in a recirculating compressor which consumes  $\$1.5 \times 10^5R$  annually. Determine the optimal operating pressure, recycle ratio, and total annual cost; and show that the cost is a minimum.

Solution: The equation giving the total operating cost is:

$$C (\$/yr.) = 1000P + 4 \times 10^9/PR + 2.5 \times 10^5R$$

Equating the partial derivatives of  $C$  with respect  $P$  and  $R$  to zero gives two algebraic equations to be solved for  $P$  and  $R$ .

$$\partial C / \partial P = 1000 - 4 \cdot 10^9 / P^2 R = 0$$

$$\partial C / \partial R = 2.5 \cdot 10^5 - 4 \cdot 10^9 / P R^2 = 0$$

Solving simultaneously gives:

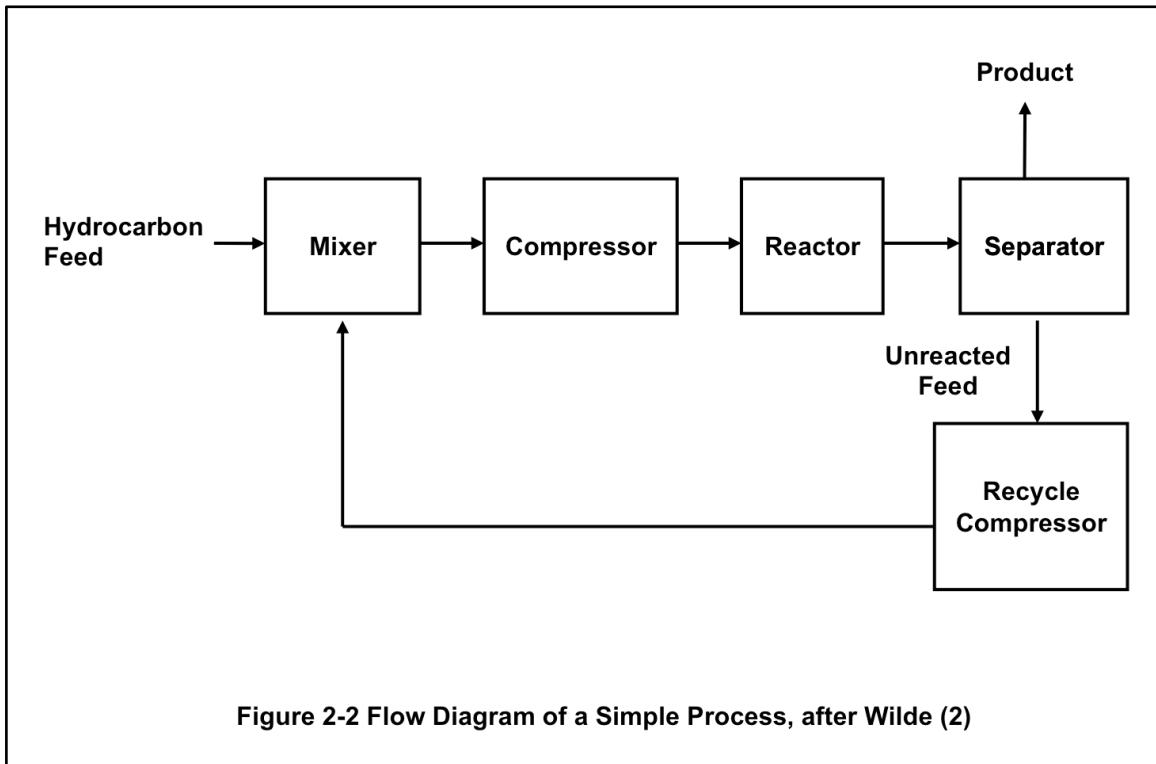


Figure 2-2 Flow Diagram of a Simple Process, after Wilde (2)

$$P = 1000 \text{ psi} \text{ and } R = 4$$

Substituting to determine the corresponding total operating cost gives:

$$C = \$ 3 \times 10^6 \text{ per year}$$

$C(P, R)$  is a minimum if:

$$\frac{\partial^2 C}{\partial P^2} > 0 \text{ and } \begin{vmatrix} \frac{\partial^2 C}{\partial P^2} & \frac{\partial^2 C}{\partial P \partial R} \\ \frac{\partial^2 C}{\partial R \partial P} & \frac{\partial^2 C}{\partial R^2} \end{vmatrix} > 0$$

Performing the appropriate partial differentiation and evaluation at the stationary point ( $P = 1000$ ,  $R = 4$ ) gives:

$$\frac{\partial^2 C}{\partial P^2} = 2 > 0 \text{ and } \begin{vmatrix} 2 & 10^3/4 \\ 10^3/4 & 10^6/4 \end{vmatrix} = 3 \cdot 10^6 / 16 > 0$$

Thus, the stationary point is a minimum since both determinants are positive.

## Analytical Methods Applicable for Constraints

To this point independent variables could take on any value. In actuality, the values of the independent variables are limited, because they usually represent physical quantities such as flow rates, temperatures, pressures, process unit, capacities and available resources. Consequently, there are constraints on variables, if nothing more than the fact that they must be nonnegative. In many cases they are bounded within limits as dictated by the process equipment and related by equations such as material balances. The constraints on the variables can be of the form of equations and inequalities.

Methods to locate the stationary points of functions (economic models) subject to equality constraints (e.g., material and energy balance equations) will be developed, and examples illustrating the techniques will be given. Inequality constraints can be converted to equality constraints, and then these procedures for equality constraints can be applied with some additional considerations.

Let us illustrate the conversion of an inequality constraint to an equality constraint using a simple example to help visualize the concept of slack variables. In Figure 2-3 an example is given of an equality and an inequality constraint for a distillation column. The material balance that says that the feed rate to the column must equal the sum of the overhead and bottom products at steady state is the equality constraint,  $F - (O + B) = 0$ . The upper limit on the capacity of the distillation column, which was set when the equipment was designed, is the inequality constraint,  $F \leq 50,000$ . This inequality constraint can be converted to an equality constraint by adding a slack variable  $S$  as  $S^2$  to ensure a positive number has been added to the equation.

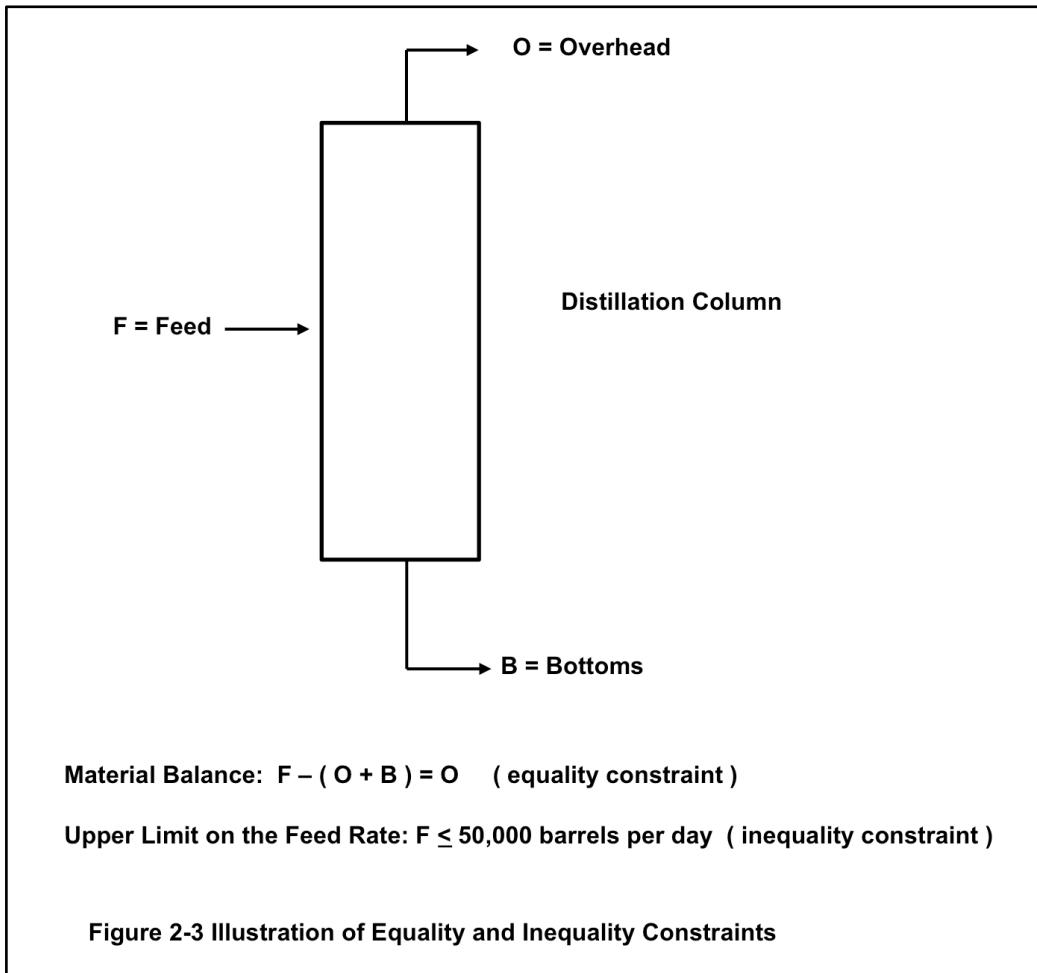
$$F + S^2 = 50,000 \quad (2-13)$$

The term slack is used to represent the difference between the optimal and upper limit on the capacity. It represents the unused, excess, or slack in capacity of the process unit. For example, if  $F_{opt} = 30,000$  barrels per day; then  $S^2 = 20,000$  barrels per day, a slack of 20,000 barrels per day; and the constraint is said to be loose i.e., the inequality holds. If  $F_{opt} = 50,000$  barrels per day then there is no slack, and the constraint is said to be tight the equality holds. This will be discussed in more detail later in the chapter. Also, if there was a lower limit on  $F$ , e.g.,  $F \geq 10,000$ , the same procedure would apply except  $S^2$  would be subtracted from  $F$ . The equation would be  $F - S^2 = 10,000$ , and  $S$  is called a surplus variable.

We can now state a general optimization problem with  $n$  independent variables and  $m$  equality constraints where the objective is to optimize (maximize or minimize) the economic model  $y(x)$  subject to  $m$  constraint equations  $f_i(x)$ .

$$\text{optimize: } y(x_1, x_2, \dots, x_n) \quad (2-14)$$

$$\text{subject to: } f_i(x_1, x_2, \dots, x_n) = 0 \quad \text{for } i = 1, 2, \dots, m \quad (2-15)$$



There must be fewer equality constraints than independent variables to be able to optimize  $y(x)$ , i.e.,  $n > m$ . If  $m = n$  the values of the  $x_j$ 's are uniquely determined, and there is no optimization problem. Also, if  $m > n$  the problem is said to be over-determined since there are more equations than unknowns. There is no optimization problem for this case either.

There are three methods of locating the optimum points of the function  $y(x_1, x_2, \dots, x_n)$  of  $n$  independent variables subject to  $m$  constraint equations  $f_i(x_1, x_2, \dots, x_n) = 0$ . These are: direct substitution, solution by constrained variation and method of Lagrange multipliers. We will find that direct substitution cannot always be used, and the method of Lagrange multipliers will be the one most frequently employed.

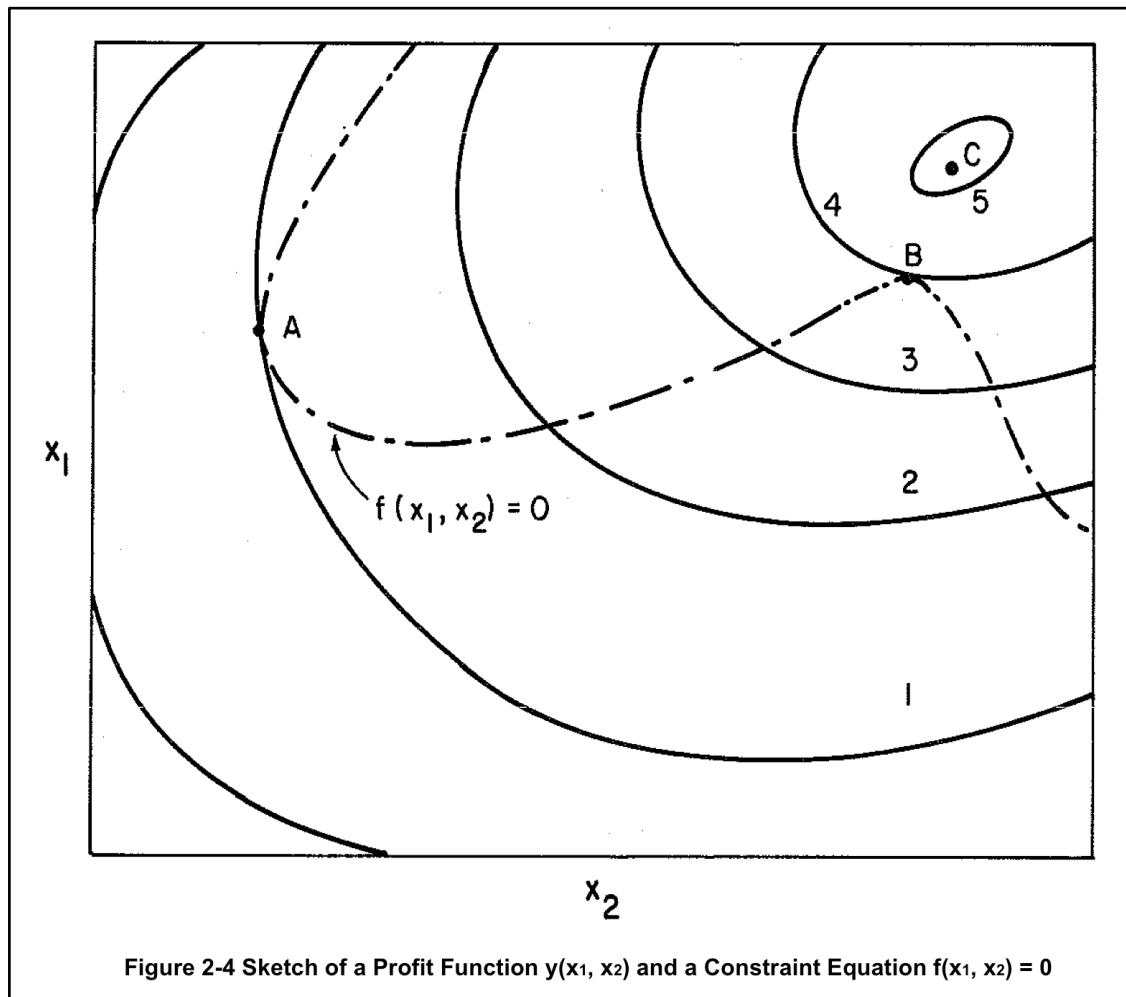
### Direct Substitution

This simply means to solve the constraint equations for the independent variables and to substitute the constraint equations directly into the function to be optimized. This will give an equation (economic model) with  $(n-m)$  unknowns, and the previous techniques for unconstrained optimization are applicable.

Unfortunately, it is not always possible to perform the algebraic manipulations required for these substitutions when the constraint equations are somewhat complicated. Consequently, it is necessary to resort to the following methods.

### Constrained Variation

This method (3, 14) is used infrequently but furnishes a theoretical basis for important multivariable numerical search methods such as the generalized reduced gradient. It is best illustrated for the case of two independent variables by considering the example shown in Figure 2-4. There is a local minimum of the constrained system at point A and a local maximum at point B. The maximum of the unconstrained system is at C.



At point A the curve  $y(x_1, x_2) = 1$  and the curve  $f(x_1, x_2) = 0$  are tangent and have the same slope. This means that differential changes,  $dx_1$  and  $dx_2$ , produce the same change in the dependent variables  $y(x_1, x_2)$  and  $f(x_1, x_2)$ . This can be expressed as:

$$\left[ \frac{dx_1}{dx_2} \right]_y = \left[ \frac{dx_1}{dx_2} \right]_f \quad (2-16)$$

We will need the total derivatives of  $y$  and  $f$  to combine with Equation (2-16) to obtain the final result. Using the first terms in a Taylor series expansion for  $y$  and  $f$  gives:

$$\begin{aligned} dy &= \frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2 = 0 \\ df &= \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 = 0 \end{aligned} \quad (2-17)$$

At the minimum, point A, and the maximum, point B,  $dy$  is equal to zero; and the constraint is satisfied, i.e.,  $f = 0$  and  $df = 0$ .

Combining equations (2-16) and (2-17) gives the following result.

$$\frac{\partial y}{\partial x_1} \cdot \frac{\partial f}{\partial x_2} - \frac{\partial y}{\partial x_2} \cdot \frac{\partial f}{\partial x_1} = 0 \quad (2-18)$$

This is an algebraic equation, and it is to be solved in combination with the constraint equation to locate the stationary points. It should be remembered in this case that  $\partial y / \partial x_1$ ,  $\partial y / \partial x_2$ ,  $\partial f / \partial x_1$ , and  $\partial f / \partial x_2$  are not necessarily zero at Points A and B. In the unconstrained case at point C,  $\partial y / \partial x_1$ ,  $\partial y / \partial x_2$  are zero, however.

This technique is illustrated with the following example. Then the extension to the general case for  $n$  independent variables will be given.

### Example 2-3

Find the stationary points of the following function using the method of constrained variation.

$$\text{optimize: } y(x) = x_1 x_2$$

$$\text{subject to: } f(x) = x_1^2 + x_2^2 - 1 = 0$$

The first partial derivatives are:

$$\frac{\partial y}{\partial x_1} = x_2 \quad \frac{\partial y}{\partial x_2} = x_1 \quad \frac{\partial f}{\partial x_1} = 2x_1 \quad \frac{\partial f}{\partial x_2} = 2x_2$$

Substituting into Equation (2-18) gives

$$x_2 \cdot 2x_2 - x_1 \cdot 2x_1 = 0 \quad \text{or} \quad x_2^2 - x_1^2 = 0$$

This equation is solved simultaneously with the constraint equation.

$$x_1^2 + x_2^2 - 1 = 0$$

The result is:

$$x_1 = \pm (\frac{1}{2})^{1/2} \quad \text{and} \quad x_2 = \pm (\frac{1}{2})^{1/2}$$

for the values of the independent variables at the stationary points

In general, we are interested in finding the stationary points of a function  $y(x_1, x_2, \dots, x_n)$  subject to  $m$  constraint equations  $f_i(x_1, x_2, \dots, x_n) = 0$  where  $i = 1, \dots, m$ , and  $n > m$ . The same reasoning applied in  $(n + 1)$  dimensional space as applied to the three-dimensional space above, and these results in the following equations:

$$\begin{aligned} dy &= \frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2 + \dots + \frac{\partial y}{\partial x_n} dx_n = 0 \\ df_1 &= \frac{\partial f_1}{\partial x_1} dx_1 + \frac{\partial f_1}{\partial x_2} dx_2 + \dots + \frac{\partial f_1}{\partial x_n} dx_n = 0 \\ &\dots \\ f_m &= \frac{\partial f_m}{\partial x_1} dx_1 + \frac{\partial f_m}{\partial x_2} dx_2 + \dots + \frac{\partial f_m}{\partial x_n} dx_n = 0 \end{aligned} \tag{2-19}$$

The set of equations given in Equation (2-19) above can be solved for  $(n - m)$  equations to go with the  $m$  constraint equations to locate the stationary points. The  $(n - m)$  equations corresponding to Equation (2-18) of the two independent variable case can be written in terms of  $(n - m)$  Jacobian determinants which are:

$$\begin{aligned} J \left[ \begin{array}{c} y, f_1, f_2, \dots, f_m \\ x_1, x_2, \dots, x_m, x_{m+1} \end{array} \right] &= 0 \\ \cdot \\ J \left[ \begin{array}{c} y, f_1, f_2, \dots, f_m \\ x_1, x_2, \dots, x_m, x_n \end{array} \right] &= 0 \end{aligned} \tag{2-20}$$

The Jacobian determinant for the first equation above is:

$$J \left[ \begin{array}{c} y, f_1, f_2, \dots, f_m \\ x_1, x_2, \dots, x_m, x_n \end{array} \right] = \begin{vmatrix} \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} & \dots & \frac{\partial y}{\partial x_{m+1}} \\ \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_{m+1}} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_{m+1}} \end{vmatrix} \quad (2-21)$$

A total of  $n$  equations are solved for the stationary points, i.e., the  $(n - m)$  equation generated by Equation (2-20) above and the  $m$  constraint equations. A derivation of these results is given by Beveridge and Schechter (6). This involves using Cramer's rule and eliminating the  $dx_i$ 's. Also, similar results are given for this general case in the text by Wilde and Beightler (4). However, a different nomenclature is used, and the results are extended to include Lagrange multipliers.

To illustrate the use of the Jacobian determinants, consider the following example, which obtains Equation (2-18).

#### Example 2-4

optimize:  $y(x_1, x_2)$

subject to:  $f(x_1, x_2) = 0$

For this problem there are 2 independent variables ( $n = 2$ ) and one constraint ( $m = 1$ ), so the evaluation of one Jacobian determinant is required.

$$J \left[ \begin{array}{c} y, f \\ x_1, x_2 \end{array} \right] = \begin{vmatrix} \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} \\ \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{vmatrix}$$

Expanding gives the following equation, gives:

$$\frac{\partial y}{\partial x_1} \frac{\partial f}{\partial x_2} - \frac{\partial y}{\partial x_2} \frac{\partial f}{\partial x_1} = 0$$

This is the same as Equation (2-18) that was solved with the constraint equation for the stationary point values of  $x_1$  and  $x_2$  in Example 2-3.

#### Lagrange Multipliers

The most frequently used method for constraints is to employ Lagrange multipliers. The technique is presented using two independent variables and one constraint equation to illustrate

the concepts. Then the procedure will be extended to the general case of  $n$  independent variables and  $m$  constraint equation. For the case of two independent variables we have:

$$\begin{aligned} \text{optimize:} \quad & y(x_1, x_2) \\ \text{subject to:} \quad & f(x_1, x_2) = 0 \end{aligned} \quad (2-22)$$

We want to show how the Lagrange multiplier arises and that the constrained problem can be converted into an unconstrained problem. The profit function and the constraint equation are expanded in a Taylor series. Then, using the first order terms gives:

$$\begin{aligned} dy &= \frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2 \\ 0 &= \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 \end{aligned}$$

This form of the constraint equation will be used to eliminate  $dx_2$  in the profit function. Solving for  $dx_2$  using the Taylor series for the constraint equation gives:

$$dx_2 = -\frac{\frac{\partial f}{\partial x_1}}{\frac{\partial f}{\partial x_2}} dx_1$$

This equation is substituted into the equation for  $dy$  to obtain:

$$dy = \frac{\partial y}{\partial x_1} dx_1 - \frac{\partial y}{\partial x_2} \left[ \frac{\frac{\partial f}{\partial x_1}}{\frac{\partial f}{\partial x_2}} \right] dx_1$$

and rearranging gives:

$$dy = \left[ \frac{\partial y}{\partial x_1} + \frac{-\frac{\partial y}{\partial x_2}}{\frac{\partial f}{\partial x_2}} \frac{\partial f}{\partial x_1} \right] dx_1$$

Now we can define  $\lambda$  as the value of  $(-\partial y / \partial x_2) / (\partial f / \partial x_2)$  at the stationary point of the constrained function. This ratio of partial derivatives  $\lambda$  is a constant at the stationary point, and the above equation can be written as:

$$dy = \left[ \frac{\partial y}{\partial x_1} + \lambda \frac{\partial f}{\partial x_1} \right] dx_1$$

or

$$dy = \left[ \frac{\partial(y + \lambda f)}{\partial x_1} \right] dx_1$$

At a stationary point  $dy = 0$ , and this leads to:

$$\frac{\partial(y + \lambda f)}{\partial x_1} = 0$$

Now, if  $L$  is defined as  $L = y + \lambda f$ , the above gives:

$$\frac{\partial L}{\partial x_1} = 0$$

This is one of the necessary conditions to locate the stationary points of an unconstrained function  $L$  which is constructed from the profit function  $y(x_1, x_2)$  and the constraint equation  $f(x_1, x_2) = 0$ .

Now the same manipulations can be repeated except using the constraint equation to eliminate  $dx_1$  from the profit function and defining  $\lambda$  as the value of  $(-\partial y / \partial x_1) / (\partial f / \partial x_1)$ . Then using the results from constrained variations, Equation 2-18, to show that the two values defined for  $\lambda$  are equal, the other necessary condition is obtained to locate the stationary points of an unconstrained function  $L$ .

$$\frac{\partial L}{\partial x_2} = 0$$

Therefore, the constrained problem can be converted to an unconstrained problem by forming the Lagrange or augmented, function,  $L$ , and solving this problem by the previously developed methods of setting the first partial derivatives equal to zero. This will give two equations to solve for the three unknowns  $x_1$ ,  $x_2$ , and  $\lambda$  at the stationary point. The third equation to be used is the constraint equation. In fact, the Lagrange multiplier is sometimes treated as another variable since  $\partial L / \partial \lambda = 0$  gives the constraint equation. The example used for the method of constrained variation will be used to illustrate these ideas.

Example 2-5:

Find the stationary points for the following constrained problem using the method of Lagrange multipliers

$$\text{optimize: } y(\mathbf{x}) = x_1 x_2$$

$$\text{subject to: } f(\mathbf{x}) = x_1^2 + x_2^2 - 1 = 0$$

The Lagrange, or augmented, function is formed as shown below.

$$L(x_1, x_2, \lambda) = x_1 x_2 + \lambda (x_1^2 + x_2^2 - 1)$$

The following equations are obtained from setting the first partial derivatives equal to zero

$$\partial L / \partial x_1 = x_2 + 2\lambda x_1 = 0$$

$$\partial L / \partial x_2 = x_1 + 2\lambda x_2 = 0$$

$$\partial L / \partial \lambda = x_1^2 + x_2^2 - 1 = 0$$

Solving the previous equations simultaneously gives the following stationary points:

$$\begin{aligned} \text{maxima: } x_1 &= (\tfrac{1}{2})^{\frac{1}{2}}, & x_2 &= (\tfrac{1}{2})^{\frac{1}{2}}, & \lambda &= -\tfrac{1}{2} \\ x_1 &= -(\tfrac{1}{2})^{\frac{1}{2}}, & x_2 &= -(\tfrac{1}{2})^{\frac{1}{2}}, & \lambda &= -\tfrac{1}{2} \end{aligned}$$

$$\begin{aligned} \text{minima: } x_1 &= (\tfrac{1}{2})^{\frac{1}{2}}, & x_2 &= -(\tfrac{1}{2})^{\frac{1}{2}}, & \lambda &= \tfrac{1}{2} \\ x_1 &= -(\tfrac{1}{2})^{\frac{1}{2}}, & x_2 &= (\tfrac{1}{2})^{\frac{1}{2}}, & \lambda &= \tfrac{1}{2} \end{aligned}$$

The types of stationary points, i.e., maxima, minima or saddle points were determined by inspection for this problem. Sufficient conditions for constrained problems will be discussed subsequently in this chapter.

The development of the Lagrange function of the case of  $n$  independent variables and  $m$  constraint equations is a direct extension from that of two independent variables and one constraint equation, and Avriel (10) gives a concise derivation of this result. (See problem 2-14). The Lagrange or augmented, function is formed as previously, and for every constraint equation there is a Lagrange multiplier. This is shown below:

$$\begin{aligned} \text{optimize: } & y(\mathbf{x}) & \mathbf{x} = (x_1, x_2, \dots, x_n)^T \\ \text{subject to: } & f_i(\mathbf{x}) = 0 & \text{for } i = 1, 2, \dots, m & \text{where } n > m \end{aligned} \tag{2-23}$$

The Lagrange, or augmented, function is formed from the constrained problem as follows:

$$L(x, \lambda) = y(x) + \sum_{i=1}^m \lambda_i f_i(x) \quad (2-24)$$

To locate the stationary points of the constrained problem, the first partial derivatives of the Lagrange function with respect to the  $x_j$ 's and  $\lambda_i$ 's are set equal to zero (necessary conditions). There are  $(n + m)$  equations to be solved for the  $(n + m)$  unknowns:  $n$   $x_j$ 's and  $m$   $\lambda_i$ 's

It is sometimes said that the method of Lagrange multipliers requires more work than the method of constrained variation since an additional  $m$  equations have to be solved for the values of the Lagrange multipliers. However, additional and valuable information is obtained from knowing the values of the Lagrange multipliers, as will be seen. The following simple example gives a comparison among the three techniques.

### Example 2-6

For the process in Example 2-1 (2) it is necessary to maintain the product of the pressure and recycle ratio equal to 9000 psi. Determine the optimal values of the pressure and recycle ratio and minimum cost within this constraint by direct substitution, constrained variation and Lagrange multipliers.

Again, the problem is to minimize  $C$ .

$$C = 1000P + 4 \cdot 10^9/PR + 2.5 \cdot 10^5R$$

However,  $C$  is subject to the following constraint equation.

$$PR = 9000$$

*Direct Substitution:* Solving the constraint above for  $P$  and substituting into the objective function gives:

$$C = 9 \times 10^6/R + (4/9) \cdot 10^6 + 2.5 \cdot 10^5R$$

Setting  $dC/dR = 0$  and solving gives:

$$R = 6 \text{ and } P = 1500 \text{ psi.}$$

The corresponding cost is:

$$C = 3.44 \cdot 10^6$$

This is greater cost than the unconstrained system, as would be expected.

*Constrained Variation:* Using Equation 2-18 and the constraint equation, the equations to be solved for this case are:

$$\frac{\partial C}{\partial P} \frac{\partial}{\partial R} (PR - 9000) - \frac{\partial C}{\partial R} \frac{\partial}{\partial P} (PR - 9000) = 0$$

$$PR - 9000 = 0$$

The first equation simplifies to:

$$P = 250R$$

which when solved simultaneously with the second equation gives the same results as direct substitution.

*Lagrange Multipliers:* The Lagrange, or augmented, function is:

$$L = 1000P + 4 \times 10^9/PR + 2.5 \times 10^5R + \lambda(PR - 9000)$$

Setting partial derivatives of L with respect to P, R and  $\lambda$  equal to zero give:

$$1000 - 4 \cdot 10^9/P^2R + \lambda R = 0$$

$$2.5 \cdot 10^5 - 4 \cdot 10^9/PR^2 + \lambda P = 0$$

$$PR - 9000 = 0$$

Solving the above simultaneously gives the same results as the two previous methods and a value for the Lagrange multiplier.

$$P = 1500, \quad R = 6, \quad \lambda = -117.3$$

### Method of Steepest Ascent

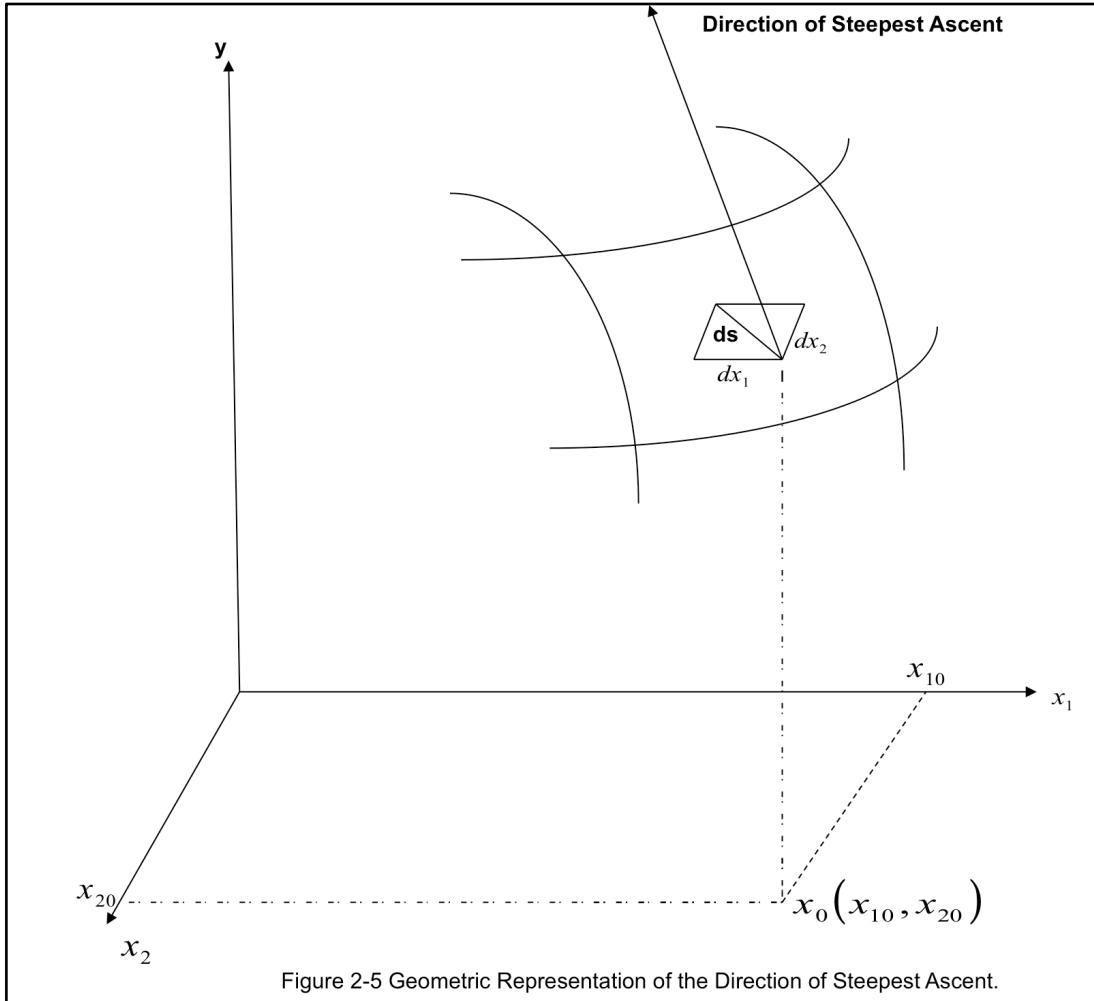
A further application of the method of Lagrange multipliers is developing the method of steepest ascent (descent) for a function to be optimized. This result will be valuable when search methods are discussed.

To illustrate the direction of steepest ascent a geometric representation is shown in Figure 2-5. To obtain the direction of steepest ascent, we wish to obtain the maximum value of  $dy$ , and  $y(x_1, x_2, \dots, x_n)$  is a function of  $n$  variables. Also, there is a constraint equation relating  $dx_1, dx_2, \dots, dx_n$  and  $ds$  as shown in Figure 2-5 for two independent variables.

The problem is:

$$\text{maximize } dy = \sum_{i=1}^n \frac{\partial y}{\partial x_i} dx_i \quad (2-25)$$

$$\text{subject to } (ds)^2 = \sum_{i=1}^n (dx_i)^2 \quad (2-26)$$



To obtain the maximum value of  $dy$  the Lagrange function is formed as follows:

$$L = \sum_{i=1}^n \frac{\partial y}{\partial x_i} dx_i + \lambda \left[ (ds)^2 - \sum_{i=1}^n (dx_i)^2 \right]$$

Differentiating  $L$  with respect to the independent variables  $dx_j$  and equating to zero gives:

$$\frac{\partial y}{\partial x_j} = -2\lambda dx_j = 0 \quad \text{for } j = 1, 2, \dots, n \quad (2-27)$$

These  $n$  equations are solved simultaneously with the constraint equation for the values of  $dx_j$  and  $\lambda$ . Solving for  $\lambda$  gives:

$$\lambda = \pm \frac{1}{2ds} \left[ \sum_{i=1}^n \left( \frac{\partial y}{\partial x_i} \right)^2 \right]^{1/2} \quad (2-28)$$

and solving for  $dx_j$  by substituting Equation 2-28 into Equation 2-27 gives:

$$dx_j = \pm ds \left[ \sum_{i=1}^n \left( \frac{\partial y}{\partial x_i} \right)^2 \right]^{-1/2} \frac{\partial y}{\partial x_j} \quad \text{for } j = 1, 2, \dots, n \quad (2-29)$$

The term in the brackets in Equation 2-29 is not a function of  $j$ , and consequently  $dx_j$  is proportional to  $\partial y / \partial x_j$ . The positive sign indicates the direction of steepest ascent, and the negative sign indicates the direction of steepest descent.

If a constant of proportionality  $k$  is used to represent the term in the brackets in Equation 2-29, this equation can be written as:

$$dx_j = \pm k \frac{\partial y}{\partial x_j} \quad \text{for } j = 1, 2, \dots, n \quad (2-30)$$

If a finite-difference approximation is used for  $dx_j = (x_j - x_{j0})$  and  $\partial y / \partial x_j$  is evaluated at  $x_0$ , then the following equation gives the direction of steepest descent or gradient line.

$$x_j = x_{j0} \pm k \frac{\partial y(x_0)}{\partial x_j} \quad \text{for } j = 1, 2, \dots, n \quad (2-31)$$

This equation can be written vector notation in terms of the gradient of  $y$  evaluated at  $x_0$ ,  $\nabla y(x_0)$ , as:

$$\mathbf{x} = \mathbf{x}_0 \pm k \nabla y(x_0) \quad (2-32)$$

If the positive sign is used, then movement is along the line in the direction of steepest ascent. If the negative sign is used, then movement is along the line in the direction of steepest descent.

The following example illustrates the use of the method of steepest descent on a simple function.

### Example 2-7

Find the minimum along the direction of steepest descent of the function given below starting at the point  $\mathbf{x}_0 = (1, 1)$ .

$$y = x_1^2 + x_2^2$$

Gradient line (steepest descent) is:

$$\mathbf{x} = \mathbf{x}_0 - k \nabla y(x_0)$$

or for two independent variables:

$$x_1 = x_{10} - k \frac{\partial y(x_0)}{\partial x_1} \quad x_2 = x_{20} - k \frac{\partial y(x_0)}{\partial x_2}$$

Evaluating the partial derivatives at the starting point  $x_0 = (1, 1)$ :

$$\frac{\partial y(x_0)}{\partial x_1} = 2x_{10} = 2 \quad \frac{\partial y(x_0)}{\partial x_2} = 2x_{20} = 2$$

The gradient line is:

$$x_1 = 1 - 2k$$

$$x_2 = 1 - 2k$$

Substituting the gradient line into the function to be minimized gives:

$$y = (1 - 2k)^2 + (1 - 2k)^2 = 2(1 - 2k)^2$$

Computing  $dy/dk$  will locate the minimum along the gradient line, i.e.,

$$\frac{dy}{dk} = -8(1 - 2k) = 0$$

and

$$k = \frac{1}{2} \text{ is the stationary point}$$

The corresponding values of  $x_1$  and  $x_2$  are

$$x_1 = 1 - 2(\frac{1}{2}) = 0 \quad x_2 = 1 - 2(\frac{1}{2}) = 0$$

It turns out that the minimum along the gradient line is also the minimum for the function in this problem since it is the sum of squares.

The method of steepest ascent is the basis for several search techniques described in Chapter 6, e.g. Steep Ascent Partan. It should be noted that when dealing with physical systems, the direction of steepest ascent (descent) may be only a direction of steep ascent (descent) depending on the scales used to represent the independent variables. This is discussed and illustrated by Wilde (5), and Wilde and Beightler (4).

## Economic Interpretation of the Lagrange Multipliers

The values of the Lagrange multipliers at the optimum provide additional and important information. If the constraint equations are written with parameters  $b_i$  on the right-hand side, the Lagrange multipliers give the change in the profit function with respect to these parameters,  $\partial y / \partial b_i$ . Many times, the right-hand sides of the constraint equations represent the availability of raw materials, demand for products or capacities of process units. Consequently, it is frequently important to know how the optimal solution is affected by changes in availability, demand and capacities. As we shall see, the Lagrange multipliers are given the names shadow prices and dual activity in linear programming where these changes are analyzed by sensitivity analysis.

The following brief derivation obtains the result that  $\partial y / \partial b = \lambda$  for the case of one constraint and two independent variables, and the extension to  $m$  constraint equations with  $n$  independent variables is comparable.

$$\begin{aligned} \text{optimize: } & y(x_1, x_2) \\ \text{subject to: } & f(x_1, x_2) = b \end{aligned} \tag{2-33}$$

First, we can obtain the following equation from the profit function by the chain rule.

$$\frac{\partial y}{\partial b} = \frac{\partial y}{\partial x_1} \frac{\partial x_1}{\partial b} + \frac{\partial y}{\partial x_2} \frac{\partial x_2}{\partial b} \tag{2-34}$$

Also, we can obtain the next equation from the constraint equation written as  $f - b = 0$  by the chain rule.

$$\frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial b} + \frac{\partial f}{\partial x_2} \frac{\partial x_2}{\partial b} - 1 = 0 \tag{2-35}$$

Then the equation from the constraint, Equation 2-35, is multiplied by the Lagrange multiplier and added to the equation from the profit function, Equation 2-34, to give:

$$\frac{\partial y}{\partial b} = \frac{\partial y}{\partial x_1} \frac{\partial x_1}{\partial b} + \frac{\partial y}{\partial x_2} \frac{\partial x_2}{\partial b} + \lambda \left( \frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial b} + \frac{\partial f}{\partial x_2} \frac{\partial x_2}{\partial b} - 1 \right)$$

Rearranging gives:

$$\frac{\partial y}{\partial b} = \left( \frac{\partial y}{\partial x_1} + \lambda \frac{\partial f}{\partial x_1} \right) \frac{\partial x_1}{\partial b} + \left( \frac{\partial y}{\partial x_2} + \lambda \frac{\partial f}{\partial x_2} \right) \frac{\partial x_2}{\partial b} - \lambda$$

$$\begin{aligned}\frac{\partial y}{\partial b} &= \left( \frac{\partial(y + \lambda f)}{\partial x_1} \right) \frac{\partial x_1}{\partial b} + \left( \frac{\partial(y + \lambda f)}{\partial x_2} \right) \frac{\partial x_2}{\partial b} - \lambda \\ \frac{\partial y}{\partial b} &= \left( \frac{\partial L}{\partial x_1} \right) \frac{\partial x_1}{\partial b} + \left( \frac{\partial L}{\partial x_2} \right) \frac{\partial x_2}{\partial b} - \lambda = -\lambda\end{aligned}\quad (2-36)$$

The values of  $\partial L / \partial x_1$  and  $\partial L / \partial x_2$  are zero at the stationary point (necessary conditions), and consequently  $\partial L / \partial b = -\lambda$ . Thus, the change in the profit function  $y$  with respect to the right-hand side of the constraint  $b$  is equal to the negative of the Lagrange multiplier. Also, comparable results can be obtained for the case on  $n$  independent variables and  $m$  constraint equations to obtain the following result using a similar procedure and arguments (7).

$$\frac{\partial y}{\partial b_i} = -\lambda_i \quad (2-37)$$

In the next section, we will see that the Lagrange multiplier is also a key factor in the analysis of problems with inequality constraints.

### Inequality Constraints

An additional complication arises when seeking the optimum value of a profit or cost function if inequality constraints are included. Although the same procedures are used, it will be necessary to consider two cases for each inequality constraint equation. One case is when the constraint is a strict equality and Lagrange multiplier is not zero. The other is when the constraint is an inequality and Lagrange multiplier is zero. This is best illustrated by the following example with one inequality constraint equation as shown below.

$$\begin{aligned}\text{optimize: } & y(x) \\ \text{subject to: } & f(x) \leq 0\end{aligned}\quad (2-38)$$

As described previously, the procedure is to add a slack variable  $x_s$  as  $x_s^2$  and form the Lagrange function:

$$L(x, \lambda) = y(x) + \lambda [ f(x) + x_s^2 ] \quad (2-39)$$

Then the first partial derivatives with respect to the  $x_i$ 's,  $x_s$  and  $\lambda$  are set equal to zero to have a set of equations to be solved for the stationary points. To illustrate the complication, the equation obtained for the slack variable is:

$$\frac{\partial L}{\partial x_s} = 2\lambda x_s = 0 \quad \text{or} \quad \lambda x_s = 0 \quad (2-40)$$

The result is two cases, i.e., either  $\lambda = 0$  and  $x_s \neq 0$ , or  $\lambda \neq 0$  and  $x_s = 0$ . If  $\lambda = 0$  and  $x_s \neq 0$ , the inequality holds; and the constraint is said to be loose, passive or inactive. If  $\lambda \neq 0$  and  $x_s = 0$ , then the equality holds; and the constraint is said to be tight or active. The following example illustrates this situation using a modification of the previous simple process.

### Example 2-8

For the process the cost function is:

$$C = 1000P + 4 \cdot 10^9/PR + 2.5 \cdot 10^5 R$$

However, C is subject to the inequality constraint equation.

$$PR \leq 9000$$

Adding the slack variable S, as  $S^2$ , and forming the Lagrange function gives:

$$L = 1000P + 4 \cdot 10^9/PR + 2.5 \cdot 10^5 R + \lambda (PR + S^2 - 9000)$$

Setting the first partial derivatives of L with respect to P, R, S and  $\lambda$  equal to zero gives the following four equations:

$$\begin{aligned}\frac{\partial L}{\partial P} &= 1000 - \frac{4 \cdot 10^9}{P^2 R} + \lambda R = 0 \\ \frac{\partial L}{\partial R} &= 2.5 \cdot 10^5 - \frac{4 \cdot 10^9}{P R^2} + \lambda P = 0 \\ \frac{\partial L}{\partial S} &= 2\lambda S = 0 \\ \frac{\partial L}{\partial \lambda} &= PR + S^2 - 9000 = 0\end{aligned}$$

The two cases are  $\lambda \neq 0$ ,  $S = 0$  and  $\lambda = 0$ ,  $S \neq 0$ .

For the case of  $\lambda \neq 0$ ,  $S \neq 0$ , the equality  $PR = 9000$  holds, i.e., the constraint is active. This was the solution obtained in Example 2-6, and the results were:

$$C = \$3.44 \cdot 10^6 \text{ per year} \quad P = 1500 \text{ psi} \quad R = 6 \quad y = -117.3$$

For the case of  $\lambda = 0$ ,  $S \neq 0$ , the constraint is an inequality, i.e., inactive. This was the solution obtained in Example 2-6 and the results were:

$$C = \$3.0 \cdot 10^6 \text{ per year} \quad P = 1000 \text{ psi} \quad R = 4 \quad S = (5000)^{1/2}$$

The example above had only one inequality constraint and two cases to consider. However, with several inequality constraints locating the stationary points can become time-

consuming, for the possibilities must be searched exhaustively. A procedure for this evaluation has been given by Cooper (7) and Walsh (8) as follows:

1. Solve the problem of optimizing:  $y(\mathbf{x})$ , ignoring the inequality constraints, i.e., having all positive slack variables. Designate this solution  $\mathbf{x}_0$ . If  $\mathbf{x}_0$  satisfies the constraints as inequalities an optimum has been found.
2. If one or more constraints are not satisfied, select one of the constraints to be an equality, i.e., active (the slack variable for this constraint is zero), and solve the problem. Call this solution  $\mathbf{x}_1$ . If  $\mathbf{x}_1$  satisfies all of the constraints, an optimum has been found.
3. If one or more constraints are not satisfied, repeat step 2 until every inequality constraint has been treated as an equality constraint (slack variable being zero) in turn.
4. If step 3 did not yield an optimum, select combinations of two inequality constraints at a time to be equalities; and solve the problem. If one of these solutions satisfies all of the constraints, an optimum has been found.
5. If step 4 did not yield an optimum, select combinations of three inequality constraints at a time to be equalities, and solve the problem. If one of these solutions satisfies all of the constraints, an optimum has been found. If not, try combinations of four inequality constraints at a time to be equalities, etc.

The above procedure applies assuming that the stationary point located is a maximum or a minimum of the constrained problem. However, there is a possibility that several stationary points will be located; some could be maxima, some minima and others saddle points. In Example 2.5 four stationary points were found, two are maxima, one a minimum and one a saddle point. Also, from Equation 2-40 for each inequality constraint where the strict inequality holds, the slack variable is positive; and the Lagrange multiplier is zero. For each inequality constraint where the equality holds, the slack variable is zero, and the Lagrange multiplier is not zero.

In the next section necessary and sufficient conditions for constrained problems are described to determine the character of stationary points. This will be similar to and an extension of the previous discussion for unconstrained problems.

### Necessary and Sufficient Conditions for Constrained Problems

The necessary conditions have been developed by Kuhn and Tucker (14) for a general nonlinear optimization problem with equality and inequality constraints. This optimization problem written in terms of minimizing  $y(\mathbf{x})$  is:

$$\text{minimize: } y(\mathbf{x}) \quad (2-41)$$

$$\text{subject to: } f_i(\mathbf{x}) \leq 0 \quad \text{for } i = 1, 2, \dots, h \quad (2-42)$$

$$f_i(\mathbf{x}) = 0 \quad \text{for } i = h+1, \dots, m \quad (2-43)$$

where  $y(\mathbf{x})$  and  $f_i(\mathbf{x})$  are twice continuously differentiable real-valued functions.

Any value of  $\mathbf{x}$  that satisfies the constraint Equations (2-42) and (2-43) is called a feasible solution to the problem in the Kuhn-Tucker theory. Then to locate points that can potentially be local minima of Equation 2-41 and satisfy the constraint Equations 2-42 and 2-43, the Kuhn-Tucker necessary conditions are used. These conditions are written in terms of the Lagrange function for the problem which is:

$$L(x, \lambda) = y(x) + \sum_{i=1}^h \lambda_i [f_i(x) + x_{n+i}^2] + \sum_{i=h+1}^m \lambda_i f_i(x) \quad (2-44)$$

where the  $x_{n+i}$ 's are the surplus variables used to convert the inequality constraints to equalities.

The necessary conditions for a constrained minimum are given by the following theorem (7, 8, 10, and 14).

To minimize  $y(\mathbf{x})$  subject to  $f_i(\mathbf{x}) \leq 0$  for  $i = 1, 2, \dots, h$  and  $f_i(\mathbf{x}) = 0$  for  $i = h+1, \dots, m$ , the necessary conditions for the existence of a relative minimum at  $\mathbf{x}^*$  are:

1.  $\frac{\partial y(x^*)}{\partial x_j} + \sum_{i=1}^h \lambda_i \frac{\partial f_i(x^*)}{\partial x_j} + \sum_{i=h+1}^m \lambda_i \frac{\partial f_i(x^*)}{\partial x_j} = 0 \quad \text{for } j = 1, 2, \dots, n$
2.  $f_i(\mathbf{x}^*) \leq 0 \quad \text{for } i = 1, 2, \dots, h$
3.  $f_i(\mathbf{x}^*) = 0 \quad \text{for } i = h+1, 2, \dots, m$
4.  $\lambda_i f_i(\mathbf{x}^*) = 0 \quad \text{for } i = 1, 2, \dots, h$
5.  $\lambda_i \geq 0 \quad \text{for } i = 1, 2, \dots, h$
6.  $\lambda_i$  is unrestricted in sign for  $i = h+1, 2, \dots, m$

Examining these conditions:

The first one is setting the first partial derivatives of the Lagrange function with respect to the independent variables  $x_1, x_2, \dots, x_n$  equal to zero to locate the Kuhn-Tucker point,  $\mathbf{x}^*$ .

The second and third conditions are repeating the inequality and equality constraint equations that must be satisfied at the minimum of the constrained problem.

The fourth condition is another way of expressing  $\lambda_i x_{n+i} = 0$  for  $i = 1, 2, \dots, h$  from setting the partial derivatives of the Lagrange function with respect to the surplus variables equal to zero. Either  $\lambda_i \neq 0$  and  $x_{n+i} = 0$  (constraint is active) or  $\lambda_i = 0$  and  $x_{n+i} \neq 0$  (constraint is inactive).

Thus, the product of the Lagrange multiplier and the constraint equation set equal to zero is an equivalent statement, and this is called the *complementary slackness condition* (15).

The fifth condition comes from examining Equation 2-37, i.e.,  $\partial y(x^*) / \partial b_i = -\lambda$ . The argument is that as  $b_i$  is increased, the constraint region is enlarged; and this cannot result in a higher value for  $y(x^*)$ , the minimum in the region. However, it could result in a lower value of  $y(x^*)$ ; and correspondingly  $\partial y(x^*) / \partial b_i$  would be negative, i.e., as  $b_i$  increases,  $y(x^*)$  could decrease. Therefore, if  $\partial y(x^*) / \partial b_i$  is negative, then the Lagrange multiplier,  $\lambda_i$ , must be positive for Equation 2-37 to be satisfied. This condition is called a *constraint qualification*, as will be discussed subsequently.

For the sixth condition, it has been shown by Bazaraa and Shetty (15) that the Lagrange multipliers associated with the equality constraints are unrestricted in sign; and there is not an argument comparable to the one given above for the Lagrange multipliers associated with the inequality constraints.

For the problem of maximizing  $y(\mathbf{x})$  subject to inequality and equality constraints, the problem is as follows:

$$\text{maximize: } y(\mathbf{x}) \quad (2-46)$$

$$\text{subject to: } f_i(\mathbf{x}) \leq 0 \quad \text{for } i = 1, 2, \dots, h \quad (2-47)$$

$$f_i(\mathbf{x}) = 0 \quad \text{for } i = h+1, \dots, m \quad (2-48)$$

For this problem the Kuhn-Tucker conditions are:

1.  $\frac{\partial y(x^*)}{\partial x_j} + \sum_{i=1}^h \lambda_i \frac{\partial f_i(x^*)}{\partial x_j} + \sum_{i=h+1}^m \lambda_i \frac{\partial f_i(x^*)}{\partial x_j} = 0 \quad \text{for } j = 1, 2, \dots, n$
2.  $f_i(x^*) \leq 0 \quad \text{for } i = 1, 2, \dots, h$
3.  $f_i(x^*) = 0 \quad \text{for } i = h+1, 2, \dots, m$
4.  $\lambda_i f_i(x^*) = 0 \quad \text{for } i = 1, 2, \dots, h$
5.  $\lambda_i \leq 0 \quad \text{for } i = 1, 2, \dots, h$
6.  $\lambda_i$  is unrestricted in sign for  $i = h+1, 2, \dots, m$

These conditions are the same as the ones for minimizing given by Equation 2-45, except the inequality is reversed for the Lagrange multipliers in the fifth condition. Also, the inequality constraints are written as less than or equal to zero for convenience in the subsequent discussion on sufficient conditions. Inequality constraints that are greater than or equal to zero can be

converted to inequality constraints that are less than or equal to zero by multiplying by minus one (- 1).

The following example illustrates the Kuhn-Tucker necessary conditions for a simple problem.

### Example 2-9

Locate the five Kuhn-Tucker points of the following problem and determine their character, i.e., maximum, minimum or saddle point.

$$\text{optimize: } y = x_1 x_2$$

$$\begin{aligned} \text{subject to: } & x_1 + x_2 \leq 1 \\ & -x_1 + x_2 \leq 1 \\ & -x_1 - x_2 \leq 1 \\ & x_1 - x_2 \leq 1 \end{aligned}$$

A diagram of the above equations is given in Figure 2-6. The function being optimized is the classic saddle point function that is constrained by plane.

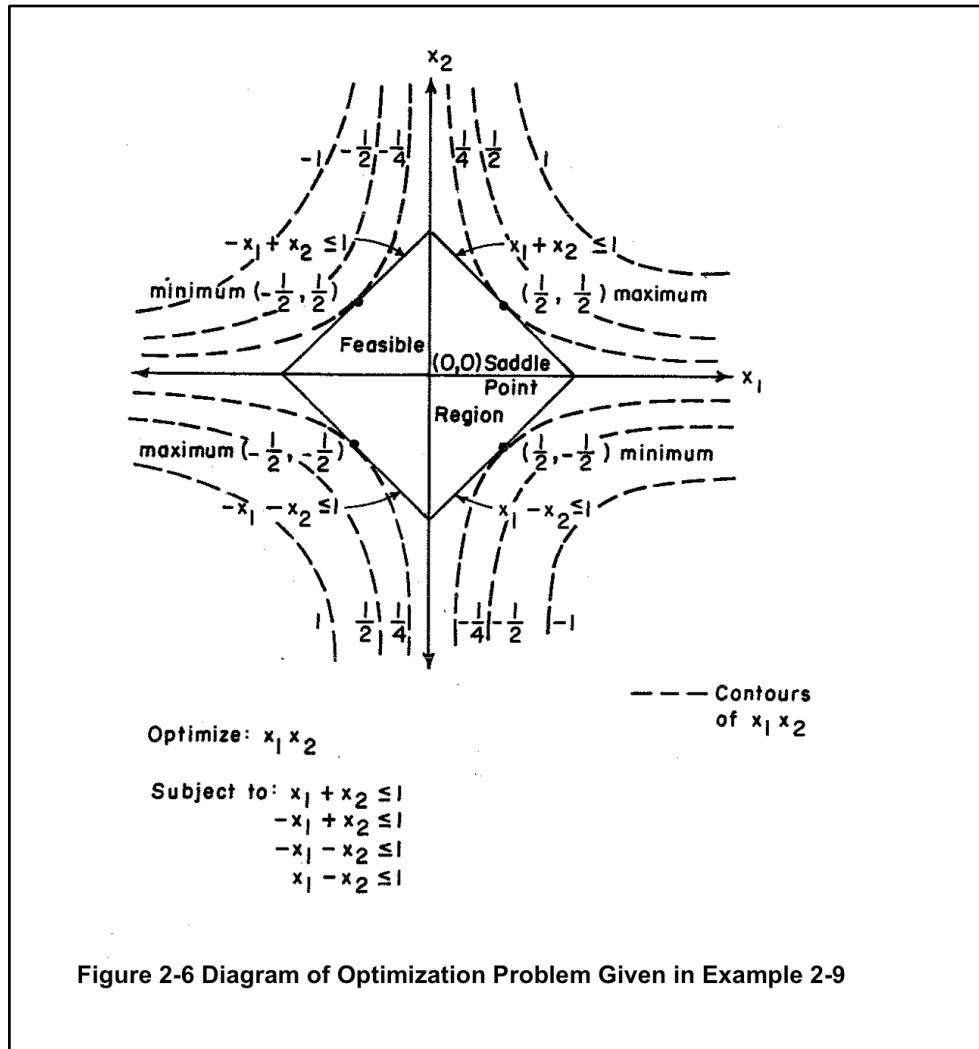


Figure 2-6 Diagram of Optimization Problem Given in Example 2-9

The first step in the procedure is to locate the stationary points by ignoring the inequality constraints, i.e.,  $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$ . If this point satisfies the constraints as inequalities, an optimum may have been found. For this problem:

$$\frac{\partial y}{\partial x_1} = x_2 = 0 \quad \frac{\partial y}{\partial x_2} = x_1 = 0$$

The Kuhn-Tucker point is  $\mathbf{x}_0(0, 0)$ , and evaluating its character by the unconstrained sufficiency conditions gives the following result:

$$\frac{\partial^2 y}{\partial x_1^2} = 0 \quad \frac{\partial^2 y}{\partial x_2^2} = 0 \quad \frac{\partial^2 y}{\partial x_1 \partial x_2} = 1 \quad \frac{\partial^2 y}{\partial x_2 \partial x_1} = 1$$

And applying Equation 2-12:

$$|H_1| = 0 \quad H_2 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

The point  $\mathbf{x}_0 (0, 0)$  is a saddle point, and the constraints are satisfied.

Proceeding to step two, one constraint equation at a time is selected, and the character of the Kuhn-Tucker point is determined. Beginning with the first constraint equation as an equality, i.e.  $\lambda_1 \neq 0$ , and considering the other three as inequalities,  $\lambda_2 = \lambda_3 = \lambda_4 = 0$ , gives the following equation for the Lagrange function.

$$L(x_1, x_2, \lambda_1) = x_1 x_2 + \lambda_1 (x_1 + x_2 - 1)$$

and

$$\frac{\partial y}{\partial x_1} = x_2 + \lambda_1 = 0 \quad \frac{\partial y}{\partial x_2} = x_1 + \lambda_1 = 0 \quad \frac{\partial y}{\partial x_2} = x_1 + x_2 - 1 = 0$$

Solving gives:

$$x_1 = \frac{1}{2}, \quad x_2 = \frac{1}{2}, \quad \lambda_1 = -\frac{1}{2} \quad y\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{4}$$

The sign of the Lagrange multiplier is negative; and by the Kuhn-Tucker necessary conditions, the Lagrange multiplier is unrestricted in sign for an equality constraint. The point can be a maximum since  $x_1$  and  $x_2$  are positive, and the other constraint equations are satisfied as inequalities. The results for the other three points are shown on Figure 2-6

**Constraint Qualifications:** In the theorems developed by Kuhn and Tucker (14), the constraint equations must satisfy certain conditions at the Kuhn-Tucker points, and these conditions are called constraint qualifications. As given in Bazaraa and Shetty (15) there are several forms of constraint qualifications; and one according to Gill et. al. (16) is important for nonlinear constraints. This is the condition that the gradients of the constraint equations at the Kuhn-Tucker point are linearly independent. This constraint qualification is required for the necessary conditions given by Equations (2-45) and (2-49). As an example, Kuhn and Tucker (14) constructed the constraint equations:

$$f_1 = (1 - x_1)^3 - x_2 \geq 0$$

$$f_2 = x_1 \geq 0$$

$$f_3 = x_2 \geq 0$$

These constraint equations do not satisfy the condition of linear independence at point  $x_1^* = 1$  and  $x_2^* = 0$ . At this point  $\nabla f_1 = [-3(1 - x_1)^2, -1] = (0, -1)$ ,  $\nabla f_2 = (1, 0)$  and  $\nabla f_3 = (0, 1)$  are not linearly independent. At such a point as this one the necessary condition may fail to hold, and Kuhn and Tucker (14) give arguments that this constraint qualification is required to ensure the existence of the Lagrange multipliers at the optimum point. Verification of the constraint qualifications for a general nonlinear programming problem is almost an impossible task according to Avriel (10). He states that fortunately in practice constraint qualification usually holds, and it is justifiable to use the existence of the Lagrange Multipliers as a basis for having the necessary conditions hold.

**Sufficient Conditions:** The same concepts used for unconstrained problems are followed to develop the sufficient conditions for constrained problems. This involves expanding the Lagrange function in a Taylor series about the Kuhn-Tucker point located using the necessary conditions. The Taylor series is simplified by neglecting third and higher order terms to give a function that contains only terms involving second partial derivatives evaluated at the Kuhn-Tucker point. This gives a differential quadratic form, and a test similar to the one for the unconstrained problem is obtained to determine if the Kuhn-Tucker point is a maximum, minimum or saddle. The sufficient conditions for the case of both inequality and equality constraints are more elaborate than if only equality constraints are involved. We have space to give only the appropriate theorems and describe their development and use. Further details are given by Avriel (10), Bazaraa and Shetty (15) and Reklaitis, et al., (17).

Considering the case of only equality constraints first, the Lagrange function for  $n$  independent variables and  $m$  equality constraint equations is given by the following equation.

$$L(\mathbf{x}, \boldsymbol{\lambda}) = y(\mathbf{x}) + \sum_{i=1}^m \lambda_i f_i(\mathbf{x}) \quad (2-50)$$

Expanding the Lagrange function in a Taylor series about the Kuhn-Tucker point  $\mathbf{x}^*$  gives:

$$L(x, \lambda) = L(x^* \lambda) + \sum_{i=1}^n L_{x_i}(x^*, \lambda)(x_i - x_i^*) + \frac{1}{2} \left[ \sum_{j=1}^n \sum_{k=1}^n L_{x_j x_k}(x^*, \lambda)(x_j - x_j^*)(x_k - x_k^*) \right] + \text{higher order terms} \quad (2-51)$$

This equation is comparable to Equation 2-8, and subscripts  $x_i$ ,  $x_j$  and  $x_k$  indicate partial differentiation with respect to those variables. Again, the first partial derivatives are zero at the Kuhn-Tucker point by the necessary conditions, and  $x$  is selected sufficiently close to  $x^*$  such that the higher order terms are negligible when compared to the second order terms. This gives the following equation which is comparable to Equation 2-9 for the unconstrained case.

$$L(x, \lambda) = L(x^* \lambda) + \frac{1}{2} \left[ \sum_{j=1}^n \sum_{k=1}^n L_{x_j x_k}(x^*, \lambda)(x_j - x_j^*)(x_k - x_k^*) \right] \quad (2-52)$$

As previously, we need to determine if the term in the brackets remains positive (minimum), remains negative (maximum) or changes sign (saddle point) for small feasible changes in  $x$  about  $x^*$ . The term in the bracket is a differential quadratic form.

To determine if the quadratic form is always positive or always negative, results comparable to those given by Equation (2-7) are required with the extension that the constraints also be satisfied, i.e., for feasible values of  $x$ . A theorem is given by Avriel (10) establishes these conditions, and this theorem is then applied to the differential quadratic form of the Lagrange function. The result, after Avriel (10), is the following theorem for the sufficient conditions of the optimization problem with only equality constraints. In this theorem the second partial derivatives of the Lagrange function evaluated at the Kuhn-Tucker point  $x^*$  are  $L_{x_j x_k}(x^*, \lambda)$  are written as  $L_{jk}$ . Also, first partial derivatives of the constraint equations evaluated at the Kuhn-Tucker point  $x^*$  are  $\partial f_j(x^*) / \partial x_k$  and are written  $f_{jk}$ . The theorem states:

Let  $y(x)$  and  $f_i(x) = 0$ ,  $i = 1, 2, \dots, m$ , be twice continuously differentiable real valued functions. If there exist vectors  $x^*$  and  $\lambda^*$ , such that:

$$L_i(x^*, \lambda^*) = 0, \quad i = 1, 2, \dots, n$$

and if:

$$D_p = (-1)^m \begin{vmatrix} L_{11} & \cdots & L_{1p} & f_{11} & \cdots & f_{m1} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ L_{1p} & \cdots & L_{pp} & f_{1p} & \cdots & f_{mp} \\ f_{11} & \cdots & f_{1p} & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ f_{m1} & \cdots & f_{mp} & 0 & \cdots & 0 \end{vmatrix} > 0$$

for  $p = m+1, \dots, n$ , then  $y(x^*)$  has a strict local minimum at  $x^*$ , such that:

$$f_i(x^*) = 0, \quad i = 1, 2, \dots, m$$

The proof of this theorem is given by Avriel (10) and follows the discussion of the concepts given above. The comparable result for a strict maxima is obtained by changing  $(-1)^m$  in the above theorem to  $(-1)^p$ , according to Avriel (10). The following example illustrates the application of this theorem.

### Example 2-10

Consider the following problem.

$$\text{optimize: } x_1^2 + 2x_2^2 + 3x_3^2$$

$$\begin{aligned} \text{subject to: } x_1 + 2x_2 + 4x_3 - 12 &= 0 \\ 2x_1 + x_2 + 3x_3 - 10 &= 0 \end{aligned}$$

Forming the Lagrange function and differentiating partially with respect to  $x_1$ ,  $x_2$ ,  $x_3$ ,  $\lambda_1$  and  $\lambda_2$  gives the following set of equations to be solved for the Kuhn-Tucker point.

$$L_{x_1} = 2x_1 + \lambda_1 + 2\lambda_2 = 0$$

$$L_{x_2} = 4x_2 + 2\lambda_1 + \lambda_2 = 0$$

$$L_{x_3} = 6x_3 + 4\lambda_1 + 3\lambda_2 = 0$$

$$L_{\lambda_1} = x_1 + 2x_2 + 4x_3 - 12 = 0$$

$$L_{\lambda_2} = 2x_1 + x_2 + 3x_3 - 10 = 0$$

Solving the above equation set simultaneously gives the following values for the Kuhn-Tucker point.

$$x_1 = 112/81, \quad x_2 = 118/81, \quad x_3 = 52/27, \quad \lambda_1 = -80/27, \quad \lambda_2 = 8/81$$

From the necessary conditions of Equations 2-45 or 2-49 the Lagrange multipliers are unrestricted in sign, and the value of the determinants from the theorem on sufficiency conditions is required to determine the character of this point. The partial derivatives needed for this evaluation are:

$$\begin{array}{lll} L_{11} = 2 & L_{12} = 0 & L_{13} = 0 \\ L_{21} = 0 & L_{22} = 4 & L_{23} = 0 \\ L_{31} = 0 & L_{32} = 0 & L_{33} = 6 \\ f_{11} = 1 & f_{12} = 2 & f_{13} = 4 \\ f_{21} = 2 & f_{22} = 1 & f_{23} = 3 \end{array}$$

The determinant is  $m = 2$ ,  $n = 3$ ,  $p = 3$ , only one determinant in this case:

$$D_3 = (-1)^2 \begin{vmatrix} 2 & 0 & 0 & 1 & 2 \\ 0 & 4 & 0 & 2 & 1 \\ 0 & 0 & 6 & 4 & 3 \\ 1 & 2 & 4 & 0 & 0 \\ 2 & 1 & 3 & 0 & 0 \end{vmatrix} = 162$$

The value of  $D_3$  is positive, and Kuhn-Tucker point is a minimum.

The sufficient conditions for problems with equality and inequality constraints, Equations (2-41), (2-42) and (2-43), are summarized in the following theorem. There are a number of mathematical concepts and theorems required to obtain this result. These are given in some detail by Avriel (10), Bazaraa and Shetty (15) and Reklaitis, e t. al. (17); but it is not feasible to describe them in the space available here.

Let  $y(x), f_i(x) \geq 0, i = 1, 2, \dots, h$  and  $f_i(x) = 0, i = h+1, \dots, m$  be twice continuously differentiable real-valued functions. If there exist vectors  $x^*$  and  $\lambda^*$  satisfying

$$\frac{\partial L}{\partial x_j} = \frac{\partial y(x^*)}{\partial x_j} - \sum_{i=1}^h \lambda_i \frac{\partial f_i(x^*)}{\partial x_j} - \sum_{i=h+1}^m \lambda_i \frac{\partial f_i(x^*)}{\partial x_j} = 0 \quad j = 1, 2, \dots, n$$

$$\lambda_i f_i(x^*) = 0 \quad i = 1, 2, \dots, h$$

$$\lambda_i \geq 0 \quad i = 1, 2, \dots, h$$

and for every  $z \neq 0$  such that:

$$\sum_{j=1}^n z_j \frac{\partial f_i(x^*)}{\partial x_j} \geq 0 \quad i = 1, 2, \dots, h$$

$$\sum_{j=1}^n z_j \frac{\partial f_i(x^*)}{\partial x_j} = 0 \quad i = h+1, 2, \dots, m$$

it follows that:

$$\sum_{j=1}^n \sum_{k=1}^n \frac{\partial^2 L(x^*, \lambda^*)}{\partial x_j \partial x_k} z_j z_k > 0$$

then  $x^*$  is a strict local minimum.

The following example illustrates the application of this theorem.

### Example 2-11

Consider the following problem after Reklaitis et. al. (17).

$$\text{minimize: } (x_1 - 1)^2 + x_2^2$$

$$\text{subject to: } -x_1 + x_2^2 \geq 0$$

Applying the theorem gives:

$$\begin{aligned} 2(x_1 - 1) + \lambda &= 0 \\ 2x_2 - 2x_2 \lambda &= 0 \\ \lambda(-x_1 + x_2^2) &= 0 \\ \lambda &\geq 0 \end{aligned}$$

Solving this set of equations gives  $x_1 = 0$ ,  $x_2 = 0$ ,  $\lambda = 2$  for the Kuhn-Tucker point. Then applying the sufficient conditions gives the following results at  $x^* = (0,0)$ .

$$2z_1^2 - 2z_2^2 > 0$$

However, for all finite values of  $z$  ( $z_1, z_2$ ) the above inequalities cannot be satisfied, and the second order sufficiency conditions show that the point is not a minimum.

In summary, the necessary and sufficient conditions for nonlinear programming problems have been described and illustrated with examples. References have been given for more details for this theory.

An important special case is when the economic model is concave, and all of the constraint equation are convex and are inequalities. This is known as *convex programming*. "A function is concave if linear interpolation between its values at any two points of definition yields a value not greater than its actual value at the point of interpolation; such a function is the negative of a convex function" according to Kuhn and Tucker (14). Thus, the convex programming problem can be stated as follows.

$$\text{maximize: } y(x) \tag{2-53}$$

$$\text{subject to: } f_i(x) \leq 0 \text{ for } i = 1, 2, \dots, m \tag{2-54}$$

The necessary and sufficient conditions for the maximum of concave function  $y(x)$  subject to convex constraints  $f_i(x) \leq 0$ ,  $i = 1, 2, \dots, m$  are the Kuhn Tucker conditions given below as:

1. 
$$\frac{\partial L(x^*, \lambda^*)}{\partial x_j} = 0 \quad j = 1, 2, \dots, n$$
2. 
$$\lambda_i f_i(x^*) = 0 \quad i = 1, 2, \dots, m$$
3. 
$$\lambda_i \leq 0 \quad i = 1, 2, \dots, m$$
4. 
$$f_i(x^*) \leq 0 \quad i = 1, 2, \dots, m$$

(2-55)

The theorem from Cooper (7) that establishes the above result is:

*If  $y(x)$  is a strictly concave function and  $f_i(x)$ ,  $i = 1, 2, \dots, m$  are convex functions which are continuous and differentiable, then the Kuhn-Tucker conditions, Equation 2-49, are sufficient as well as necessary for a global maximum.*

The proof of this theorem uses the definition of convex and concave functions and the fact that the Lagrange function can be formulated as the sum of concave functions that is concave.

These concepts and results for the Kuhn-Tucker conditions and those given previously will be valuable in our discussion of modern optimization procedures in the following chapters. Those interested in further theoretical results are referred to the references at the end of this chapter and Chapter 6. Also, in industrial practice we will see that the concepts from the Kuhn-Tucker conditions are used in computer programs for advanced multivariable search methods to optimize economic and process models that are too elaborate for the algebraic manipulations required to use these theories directly.

### Closure

In the chapter we have discussed the necessary and sufficient conditions to evaluate the character of stationary points for unconstrained and constrained optimization problems. It was necessary to confine the illustration of these procedures to simple algebraic models. Even though we are not able to apply these procedures directly to the optimization of industrial processes, the concepts developed in this chapter are used many times over in the following chapters.

It is worthwhile to attempt to solve the following unconstrained economic model from the design of horizontal vapor condensers in evaporators used in water desalination plants to see one of the major limitations of the classical theory of maxima and minima. The problem is to minimize the cost given by the following equation.

$$C = aN^{-7/6}D^{-1}L^{-4/3} + bN^{-0.2}D^{0.8}L^{-1} + cNDL + dN^{-1.8}D^{-4.8}L \quad (2-56)$$

In this equation the cost is in dollars per year; N is the number of tubes in the condenser; D is the nominal diameter of the tubes in inches; L is the tube length in feet; and a, b, c, and d are coefficients that vary with the fluids involved and the construction costs. Avriel and Wilde (18) give further details about the significance of each term. This equation is typical of the form that is obtained from assembling correlations of equipment costs and related process operating conditions for preliminary cost estimates.

Differentiating this equation with respect to the three independent variables N, D, and L, and setting the results equal to zero gives the following three equations to be solved for the values of N, D, and L that would give the minimum cost.

$$\begin{aligned}
 & (-7a/6)N^{-13/6}D^{-1}L^{-4/3} + (-0.2b)N^{-1.2}D^{0.8}L^{-1} + cDL + (-1.8d)N^{-2.8}D^{-4.8}L = 0 \\
 & -aN^{-7/6}D^{-2}L^{-4/3} + 0.8bN^{-0.2}D^{-0.2}L^{-1} + cNL + (-4.8d)N^{-1.8}D^{-5.8}L = 0 \quad (2-57) \\
 & (-4a/3)N^{-7/6}D^{-2}L^{-7/5} - N^{-0.2}D^{0.8}L^{-2} + cND + dN^{-1.8}D^{-4.8} = 0
 \end{aligned}$$

There is no straightforward way to solve this relatively complicated set of three nonlinear algebraic equations other than numerically with a root-finding procedure at this point. This then illustrates one of the major limitations with classical methods, i.e., if the variables in the economic model have fractional exponents, then a set of nonlinear algebraic equations are obtained that will probably require an iterative solution using a computer. However, as will be seen in the chapter on geometric programming, we will be able to obtain the optimal solution for this economic model by solving a set of linear algebraic equations. We will take advantage of the mathematical structure of the problems to be able to find the optimum readily. In fact, this will be true of most of the modern methods; they take advantage of the mathematical structure of the optimization problem to quickly find the best values of the independent variables and the maximum profit or minimum cost.

## References

1. Hancock, H., Theory of Maxima and Minima, Dover Publications, Inc., New York (1960).
2. Wilde, D. J., "A Review of Optimization Theory," Ind. Eng. Chem., Vol. 57, No. 8, p.18 (1965).
3. Smith, C. L., R. W. Pike and P. W. Murrill, Formulation and Optimization of Mathematical Models, International Textbook Co., Scranton, Pa. (1970).
4. Wilde, D. J. and C. S. Beightler, Foundations of Optimization, Prentice-Hall, Inc., Englewood Cliffs, N.J. (1967).
5. Wilde, D. J., Optimum Seeking Methods, Prentice-Hall, Inc., Englewood Cliffs, N.J. (1965).
6. Beveridge, G. S. G., and R. S. Schechter, Optimization Theory and Practice, McGraw-Hill Book Co., New York (1970).
7. Cooper, Leon, Mathematical Programming for Operations Researchers and Computer Scientists, Ed. A. G. Holtzman, Marcel Dekker, Inc., New York (1981).
8. Walsh, G. R. Methods of Optimization, John Wiley and Sons, Inc., New York (1979).
9. Sivazlian, B. D. and L. E. Stafel, Optimization Techniques in Operations Research, Prentice-Hall, Inc., Englewood Cliffs, N.J. (1975).
10. Avriel, M. Nonlinear Programming, Analysis and Methods, Prentice Hall, Inc., Englewood Cliffs, N. J. (1976).
11. Courant, R. and D. Hilbert, Methods of Mathematical Physics, Vol. 1, p. 164, Interscience Publishers, Inc., New York (1953).
12. Burley, D. M., Studies in Optimization, John Wiley and Sons, Inc., New York (1974).
13. Sokolnikoff, I. S., and R. M. Redheffer, Mathematics of Physics and Modern Engineering, Second Edition, McGraw - Hill Book Co., New York (1966).
14. Kuhn, H. W., and A. W. Tucker, "Nonlinear Programming," Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Ed. Jerzy Neyman, p. 481 - 92, University of California Press, Berkeley, California (1951).
15. Bazaraa, M. S., and C. M. Shetty, Nonlinear Programming - Theory and Algorithms, John Wiley & Sons, Inc., New York (1979).
16. Gill, P. E., W. Murray, and M. H. Wright, Practical Optimization, Academic Press, New York (1981).
17. Reklaitis, G. V., A. Ravindran and K. M. Ragsdell, Engineering Optimization: Methods and Applications, John Wiley and Sons, Inc., New York (1983).
18. Avriel, M. and D. J. Wilde, "Optimal Condenser Design by Geometric Programming", I & EC Process Design and Development, Vol. 6, No. 2, p. 256 (April, 1967).

## Problems

2-1. Locate the stationary points of the following functions and determine their character.

a.  $y = x^4/2 - x^2/2$

b.  $y = x^7$

c.  $y = x_1^2 + x_1x_2 + x_2^2$

d.  $y = 2x_1^2 + 3x_2^2 + 4x_3^2 - 8x_1 - 12x_2 - 24x_3 + 110$

2-2. Find the global maximum of the function

$$y(x_1, x_2) = 5(x_1 - 3)^2 - 12(x_2 + 5)^2 + 6x_1x_2$$

in the region  $0 \leq x_1 \leq 10$   
 $0 \leq x_2 \leq 5$

2-3. Use the Jacobian determinants and obtain the two equations to be solved with the constraint equation for the following problem

optimize:  $y(x_1, x_2, x_3)$

subject to:  $f(x_1, x_2, x_3) = 0$

2-4. Solve the following problem by the method of constrained variation and the method of Lagrange multipliers, evaluating  $x_1$ ,  $x_2$  and the Lagrange multiplier  $\lambda$  at the optimum.

maximize:  $x_1 + x_2$

subject to:  $x_1^2 + x_2^2 = 1$

2-5. Solve the following problem by the method of Lagrange multipliers and give the character of the stationary point.

minimize:  $2x_1^2 - 4x_1x_2 + 4x_2^2 + x_1 - 3x_2$

subject to:  $10x_1 + 5x_2 \leq 3$

2-6. Consider the following problem

optimize:  $-x_1^2 - 2x_1 + x_2^2$

subject to:  $x_1^2 + x_2^2 - 1 \leq 0$

a. Obtain the equation set to be solved to locate the stationary points of the above problem using the method of Lagrange multipliers. Convert the inequality constraint to an equality constraint with the slack variable  $x_3$  as  $x_3^2$ ; why?

b. Show that the following are solutions to the algebraic equations obtained in part (a).

Stationary Points

	A	B	C	D
$x_1$	$-\frac{1}{2}$	$-\frac{1}{2}$	-1	1
$x_2$	$\sqrt{3}/2$	$-\sqrt{3}/2$	0	0
$x_3$	0	0	0	0
$\lambda$	-1	-1	0	2

c. Based on the value of the function being optimized state whether stationary points A through D are maximum, minimum or saddle points.

2-7. The cost of operation of a continuous, stirred-tank reactor is given by the following equation:

$$C_T = C_f c_{A_0} q + C_m V$$

The total operating cost  $C_T$  (\$/hr) is the sum of the cost of the feed,  $C_f c_{A_0} q$ , and the cost of mixing,  $C_m V$ . The following gives the values for the reactor.

$C_f$  = \$5.00/lb-mole of A, cost of feed

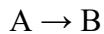
$c_{A_0}$  = 0.04 lb-mole/ft<sup>3</sup>, initial concentration of A.

$q$  = volumetric flow rate of feed to the reactor in ft<sup>3</sup>/hr.

$C_m$  = \$0.30/hr-ft<sup>3</sup>, cost of mixing

$V$  = volume of reactor in ft<sup>3</sup>

We are interested in obtaining the minimum total operating cost and the optimal values of the feed rate,  $q$ ; reactor volume,  $V$ ; and concentration in the reactor,  $c_A$ . The following first order reaction takes place in the reactor.



where the rate of formation of B,  $r_B$ , is given by

$$r_B = kc_A$$

where  $k = 0.1 \text{ hr}^{-1}$ .

a. If 10 lb-moles per hour of B are to be produced, give the two material balance constraint equations which restrict the values of the independent variables. (There is no B in the feed stream.)

b. Form the Lagrange function and perform the appropriate differentiation to obtain the set of equations that would be solved for the optimal values of the independent variables. How many equations and variables are obtained?

c. Solve for the optimal values of the reactor volume,  $V$ ; feed rate,  $q$ ; and concentration of  $A$  in the product,  $c_A$ .

2-8. Solve the following problem by the method of Lagrange multipliers, and determine the character of the stationary point.

$$\text{optimize: } 2x_1^2 + 2x_1x_2 + x_2^2 - 20x_1 - 14x_2$$

$$\text{subject to: } x_1 + 3x_2 \leq 5$$

$$2x_1 - x_2 \leq 4$$

2-9.<sup>9</sup> Solve the following problem by the method of Lagrange multipliers, and determine the character of the stationary point:

$$\text{optimize: } (1/3)x_1 + x_2$$

$$\text{subject to: } -x_1 + x_2 \leq 0$$

$$x_1 + x_2 \leq 3$$

2-10. The total feed rate to three chemical reactors in parallel is 1100 pounds per hour. Each reactor is operating with a different catalyst and conditions of temperature and pressure. The profit function for each reactor has the feed rate as the independent variable, and the parameters in the equation are determined by the catalyst and operating conditions. The profit functions for each reactor are given below:

$$P_1 = 0.2F_1 - 2(F_1/100)^2$$

$$P_2 = 0.2F_2 - 4(F_2/100)^2$$

$$P_3 = 0.2F_3 - 6(F_3/100)^2$$

Determine the maximum profit and the optimal feed rate to each reactor.

2-11. Solve the following problem by the method of Lagrange multipliers and determine the character of the stationary points.

$$\text{maximize: } 3x_1^2 + 2x_2^2$$

$$\text{subject to: } x_1^2 + x_2^2 \leq 25$$

$$9x_1 - x_2^2 \leq 27$$

2-12. Find the stationary points of the following problem, and determine their character, i.e., maximum, minimum or saddle point.

$$\text{optimize: } 2x_1^2 + x_2^2 - 5x_1 - 4x_2$$

$$\text{subject to: } x_1 + 3x_2 \leq 5$$

$$2x_1 - x_2 \leq 4$$

2-13. The rate of return (ROR) is defined as the interest rate where the net present value (NPV) is zero for a specified number of years,  $n$ , and initial cash flow  $CF_0$  which is negative. This can be formulated as an optimization problem as follows:

$$\text{minimize: } (NPV)^2$$

For the case of constant cash flows,  $CF_j = A$ , develop the equation to determine the rate of return. The net present value is given by the following equation.

$$NPV = -CF_0 + A[1 - (1 + i)^{-n}] / i$$

2-14. Derive the Lagrange function for  $n$  independent variables and  $m$  constraint equations, Equation 2-24. Begin by multiplying the constraint equations given in Equation (2-19) by Lagrange multipliers  $\lambda_1, \lambda_2, \dots, \lambda_m$ . Then add all of the equations, rearrange terms and obtain the result as Equation 2-24.

2-15. For sufficient conditions of the equality constraint problem to determine if the quadratic form is positive or negative definite, the signs of the roots of a polynomial can be evaluated. This characteristic polynomial is obtained by evaluating the following determinant which includes the second partial derivatives of the Lagrange function evaluated at the Kuhn-Tucker points,  $L_{x_j x_k}(x^*, \lambda)$  written as  $L_{jk}$  for simplicity, and the first partial derivative of the constraint equations evaluated at the Kuhn-Tucker point,  $\partial f_j(x^*) / \partial x_k$  written as  $f_{jk}$  for simplicity.

$$P(a) = \begin{vmatrix} L_{11} - a & L_{12} & \cdots & L_{1n} & f_{11} & \cdots & f_{m1} \\ L_{21} & L_{22} - a & \cdots & L_{2n} & f_{12} & \cdots & f_{m2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ L_{n1} & L_{n2} & \cdots & L_{nn} - a & f_{1n} & \cdots & f_{mn} \\ f_{11} & f_{12} & \cdots & f_{1n} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ f_{m1} & f_{m2} & \cdots & f_{mn} & 0 & \cdots & 0 \end{vmatrix}$$

The following results are used to evaluate the type of stationary points. First, evaluate the roots of  $P(a)$  using the above equation.

If each root of  $P(a)$  is positive, then  $x^*$  is a maximum.

If each root of  $P(a)$  is negative, then  $x^*$  is a minimum.

If the roots are of mixed sign, then  $x^*$  is a saddle point.

Use the results given in Example 2-10 in the above determinant and confirm the character of the Kuhn-Tucker point by this method. (There is a comparable sufficient condition test for the unconstrained problem which is described by Sivazlian and Stanfel (9).)