PROBLEMS

Problem 8-1

Find the optimum values of the following objective functions:

$$f(x) = \left(x_1^2 + x_2 - 11\right)^2 + \left(x_1 + x_2^2 - 7\right)^2$$

$$f(x) = x_1^2 + 2x_1x_2 + 4x_1x_3 + 3x_2^2 + 2x_2x_3 + 5x_3^2$$

Problem 8-2

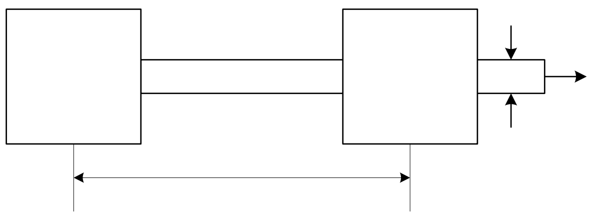
Consider the unconstrained optimization problem:

$$\min\ (x_1 - 2)^4 + (x_1 - 2x_2)^2 .$$

1. Perform two iterations of the gradient method, starting from $x_0 = [0, 3]^T$.
2. Perform four iterations of Newton's method with the same starting point x_0.

Problem 8-3

A transmission system for gas pipeline is shown in Figure below



The total annual cost is mathematically formulated as a function of the pipe diameter, D (in), the compressor discharge pressure, P_1 (psia), the length between two compressors L (miles), and compression ratio $r = P_1/P_2$:

$$C(D, P_1, L, r) = 7.84D^2P_1 + 450{,}000 + 36{,}900D + \frac{6.57\times10^6}{L} + \frac{772\times10^6}{L}(r^{0.218} - 1) \qquad (\$/yr)$$

The flow rate is given by the following relation:

$$Q = 3.39\left[\frac{(P_1^2 - p_2^2)D^5}{fL}\right]^{0.5}$$

Let the flow rate be 100×10^6 SCF/day and the friction factor $f = 0.008D^{-1/3}$. Given the above information, find the minimum cost design.

Problem 8-4

Consider a counter-current extraction column of N stages. The feed contains a solute of X_0 wt% and the solvent flow rate is *B* kg/hr. Under certain conditions, the hourly profit is formulated as follows

$$P = \lambda CX_0\left(\frac{S^N - 1}{S^{N+1} - 1}\right) - \left(\gamma CX_0 + \alpha N + \beta B\right)$$

Where S=C/(*m*B) with *m* is the distribution ratio, γ the cost of 1kg of feed, λ is cost of 1kg of extract product, β is cost of 1kg of solvent and α is the cost of stage operation. Find the optimum number of stages and solvent feed rate that maximizes the hourly profit given the following values:

$X_0 = 0.0526$; $CX_0 = 200$ kg/hr; m =2.2

$\alpha = 40$ \$/hr; $\beta = 0.8$ \$/kg; $\lambda = 27.6$ \$/kg

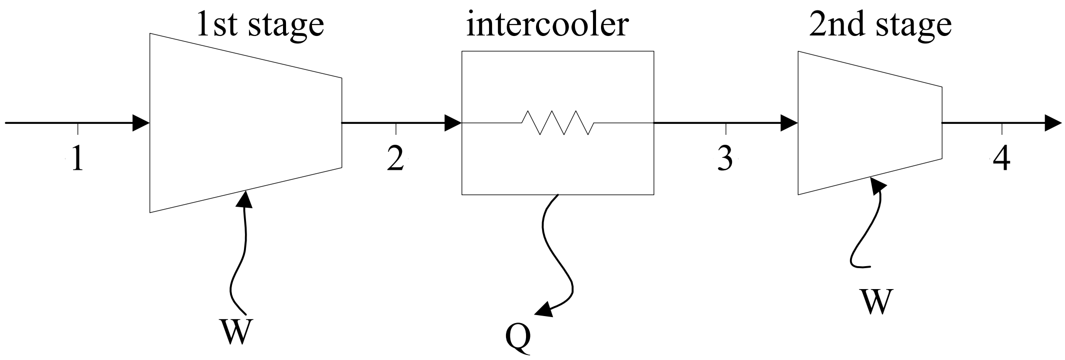
Problem 8.5

In a multistage compressor it is desirable to select the interstage pressure so that the total work input is minimized. Using the nomenclature shown in Figure below, the stage work for isentropic compression are

$$W_{12} = Cp\, T_1\, (p_2/p_1)^{(\gamma-1/\gamma)\,-1}$$

$$W_{34} = Cp\, T_3\, (p_4/p_3)^{(\gamma-1/\gamma)\,-1}$$

The total work is the sum of these stages. We wish to find the interstage pressure to minimze the total work given that p_1 = 14.7 psia , p_2 = 100 psia, $T_1 = T_3 = 70^oF$ and Cp = 0.24 Btu/lb ^oF.



CHAPTER 1: CONSTRAINED MULTIDIMENSIONAL PROBLEMS;

LAGRANGE MULTIPLIERS METHOD

9.1 INTRODUCTION

Constrained optimization problems can be put under the following general form:

$$\underset{x_1,\cdots,x_n}{\text{minimize}}(or \underset{x_1,\cdots,x_n}{\text{maximize}}) f(x_1, x_2, \cdots, x_n) \tag{9.1}$$

subject to the following equality constraints:

$$h_j(x_1, x_2, \cdots, x_n) = 0, \qquad j = 1, \cdots, J \tag{9.2}$$

and to inequality constraints

$$g_k(x_1, x_2, \cdots, x_n) \geq 0, \qquad k = 1, \cdots, K \tag{9.3}$$

A number of techniques are being used for the solution of this optimization problem. In the following section the method of Lagrange multipliers is presented.

9.2 LAGRANGE MULTIPLIERS

Consider first the following minimization problem with equality constraints:

$$\min_x f(x) \tag{9.4}$$

$$h_j(x) = 0, \qquad j = 1, \cdots, J \tag{9.5}$$

with $x=[x_1,x_2,\ldots,x_n]^T$

Define the following modified objective function:

$$L(x, \lambda) = f(x) + \sum_{j=1}^{J} \lambda_j h_j \tag{9.6}$$

The function $L(x, \lambda)$ is called the *Lagrange function* and the unspecified real constants λ_j are called the *Lagrange multipliers*. Note that there are as many multipliers as there are constraints. It can be proved that the necessary conditions for the existence of a local optimum for the original problem (Eqs.9.4-5) are found by setting the partial derivative of L with respect to x to zero, i.e.

$$\begin{cases} \dfrac{\partial L(x,\lambda)}{\partial x_1} = 0 \\ \vdots \\ \dfrac{\partial L(x,\lambda)}{\partial x_n} = 0 \end{cases} \tag{9.7}$$

This system of N equations is augmented with the original equality constraints:

$$\begin{cases} h_1(x) = 0 \\ h_2(x) = 0 \\ \vdots \\ h_J(x) = 0 \end{cases} \tag{9.8}$$

The augmented system consists of $N + J$ algebraic equations and $N + J$ unknowns (N for x, and J for λ) . The solution of this system yields a stationary point for the original optimization problem (Eqs.9.4-9.5). Therefore we see that the method of Lagrange multipliers has converted the original constrained optimization problem to an unconstrained problem. However the elimination of the constraints has come at the expense of increasing the problem dimension by J variables. It should be noted that at the optimum $x=x^*$ the constraints must be satisfied i.e. $h_j(x^*)=0$ (j=1, J). Therefore from Eq.(9.6) we see that at the optimum, the value of the lagrangian (L) must be equal to that of (f).

9.3 ECONOMIC INTERPRETATION OF LAGRANGE MULTIPLIERS

Lagrange multipliers have an important interpretation in optimization problems. Consider the previous minimization problem of Eq.(9.4) but with the following constraints

$$h_j(x) = b_j \qquad j = 1, \cdots, J \tag{9.9}$$

The constant b_j is the right hand side constant for the constraint j . The Lagrangian is defined as:

$$L(x, \lambda) = f(x) + \sum_{j=1}^{J} \lambda_j (h_j - b_j) \tag{9.10}$$

Let x^* be the optimal solution and λ^* the optimal Lagrange multiplier. We know from the previous section that at the optimum f and L have the same values. We would like to know what effect on the optimum will have a change on the value of b_j . This analysis, often carried out in optimization problems, is called sensitivity analysis. Taking the partial derivative of Eq.(9.10) with respect to b_j at the optimum yields

$$\frac{\partial L}{\partial b_j} = -\lambda_j^*, \qquad j = 1, \cdots, J \tag{9.11}$$

The left hand side of this equation is the rate of change of the optimal value of L with respect to the right-hand-side b_j of the j^{th} constraint while the right hand side is the Lagrange multiplier associated with the j^{th} constraint. Since at the optimum both f and L are equal then we can also think of λ_j as being the change in the objective function (f) for a change of b_j at the optimal solution. In a number of applications, b_j represents the quantity of a limiting source. Lagrange multipliers are consequently called *shadow prices*.

9.4 LAGRANGE MULTIPLIERS FOR INEQUALITY CONSTRAINTS

The Lagrange multipliers method can be extended to the system of inequalities. Consider the general optimization problem as defined by Eqs.9.1-9.3. First the inequality constraints are transformed to equality constraints by introducing positive slack variables σ_k^2 for each constraint *k*.

$$g_k(x_1, x_2, \cdots, x_n) - \sigma^2{}_k = 0 \qquad k = 1, \cdots, K \tag{9.12}$$

The Lagrangian is therefore

$$L(x, v) = f(x) + \sum_{j=1}^{J} \lambda_j h_j + \sum_{k=1}^{K} \lambda_k (g_k(x) - \sigma_k^2) \tag{9.13}$$

where λ_j (*j=1,...J+K*) are Lagrange multipliers.

The necessary and sufficient conditions for a point x^* to be a minimum are that:

- *f(x*)* is convex in the vicinity of x^*

- The set formed by the equality and the inequity constraints form a convex set in the vicinity of x^*
- The following conditions hold

$$\frac{\partial L(x^*)}{\partial x_i} = 0, \qquad i = 1, \cdots, n \tag{9.14}$$

$$\frac{\partial L(x^*)}{\partial \lambda_i} = 0, \qquad i = 1, \cdots, J + K \tag{9.15}$$

$$\frac{\partial L(x^*)}{\partial \sigma_k} = 0, \qquad k = 1, \cdots, K \tag{9.16}$$

$$\lambda_i \leq 0, (i = 1, \ldots, \mathrm{J}+K) \qquad \text{for a minimum} \tag{9.17}$$

and

$$\lambda_i \geq 0, (i = 1, \ldots, \mathrm{J}+K) \quad \text{for a maximum} \tag{9.18}$$

The method of Lagrange multipliers can be very useful in cases where the system of nonlinear equations can be easily solved. It has the advantage of providing insight into the problem. However, when the derivative are difficult to generate or when the system of equations is difficult to solve, this method is not recommended and various iterative methods are used instead for the solution of the problem, as it can be checked in references.

Example 9.1: Maximize the yield of a reaction

Consider the problem of finding the optimum yield for a first order consecutive reactions

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C \tag{9.19}$$

taking place in a non-isothermal CSTR at given residence time θ.

The objective function is to maximize the yield defined as B/A_0 over the concentrations of A and B in addition to the temperature T. The constraints for the problem come from steady state mass balance equations.

- Mass balance on A

$$A_0 - A - k_1(\mathrm{T})A\theta = 0 \tag{9.20}$$

- Assume the feed is free of B, i.e., $B_0 = 0$ the mass balance on B yields

$$-B + [k_1(T)A - k_2(T)B]\theta = 0 \tag{9.21}$$

The Lagrangian is, therefore

$$L(A,B,T) = \frac{B}{A_o} + \lambda_1[A_o - A - k_1(T)A\theta] + \lambda_2[-B + (k_1(T)A - k_2(T)B)\theta] \tag{9.22}$$

The necessary conditions for the existence of an optimum are obtained by setting the partial derivatives to zero, i.e.

$$\frac{\partial L}{\partial A} = -\lambda_1[1 + k_1(T)\theta] + \lambda_2[(k_1(T)\theta] = 0 \qquad (9.23)$$

$$\frac{\partial L}{\partial B} = \frac{1}{A_o} - \lambda_2[1 + (k_2(T)\theta] = 0 \qquad (9.24)$$

$$\frac{\partial L}{\partial T} = -\lambda_1\theta[A\frac{dk_1(T)}{dT}] + \lambda_2\theta[\frac{dk_1(T)}{dT}A - \frac{dk_2(T)}{dT}B] = 0 \qquad (9.25)$$

In addition to the above equations, the mass balance equations (9.19) and (9.20) have to be satisfied. Solving equation (9.19) for A gives:

$$A = \frac{A_o}{1 + k_1\theta} \qquad (9.26)$$

Solving equation (9.20) yields:

$$B = \frac{k_1(T)A\theta}{1 + k_2(T)\theta} \qquad (9.27)$$

Substituting Eq.(9.26) into Eq.(9.27) yields

$$B = \frac{k_1(T)A_o\theta}{(1 + k_1(T)\theta)(1 + k_2(T)\theta)} \qquad (9.28)$$

Solving equation (9.23) for λ_2 gives:

$$\lambda_2 = \frac{1}{A_o(1 + k_2(T)\theta)} \qquad (9.29)$$

while equation (9.23) yields:

$$\lambda_1 = \frac{k_1(T)\theta}{1 + k_1(T)\theta}\lambda_2 \tag{9.30}$$

Substituting the expressions of A(Eq. 9.26), B (Eq. 9.28), λ_1 (Eq. 9.30) and λ_2 (Eq. 9.29) into Eq.(9.25) results in the following conditions for a maximum yield:

$$(1 + k_2(T)\theta)\frac{dk_1(T)}{dT} = k_1(T)\theta(1 + k_1(T)\theta)\frac{dk_2(T)}{dT} \tag{9.31}$$

Assuming Arrhenius law for each reaction rate,

$$k_1(T) = k_{10}e^{\frac{-E_1}{RT}} \tag{9.32}$$

$$k_2(T) = k_{20}e^{\frac{-E_2}{RT}} \tag{9.33}$$

Taking the derivative of the last two equations and using the fact that $\frac{dk_i(T)}{dT} = \frac{E_i}{RT^2}k_i\ (i = 1,2)$ yields the following nonlinear algebraic equation for optimal value of temperature

$$\frac{E_1}{E_2}(1 + k_2(T)\theta) = \theta(1 + k_1(T)\theta)k_2(T) \tag{9.34}$$

By substituting equation (9.29) into equation (9.30), the following expression for λ_1 and λ_2 can be obtained

$$\lambda_1 A_0 = \frac{k_1(T)\theta}{(1 + k_1(T)\theta)(1 + k_2(T)\theta)} = \frac{B}{A_o} \equiv yield \tag{9.35}$$

$$\lambda_2 = \frac{1}{(1+k_2(T)\theta)} = \frac{B}{A_O - A} \qquad (9.36)$$

λ_2 represents an other form of the yield. The difference from λ_1 is that the consumed amount of A is used in the definition instead of the feed of A. The Lagrangian at the optimum is:

$$L^* = \frac{A}{A_o}\frac{k_1(T)\theta}{(1+k_2(T)\theta)} \qquad (9.37)$$

Using the results of Section 9.3, the economic interpretation of Lagrange multipliers can be examined. The derivative of the Lagrange function (L^*) with respect to the constant term of the first equality constraint, A_0,is:

$$\frac{\partial L^*}{\partial A_o} = -\frac{A}{A_o^2}\frac{k_1(T)\theta}{(1+k_2(T)\theta)} = -\frac{A_o}{A_o^2}\frac{k_1(T)\theta}{(1+k_2(T)\theta)(1+k_1(T)\theta)} = \lambda_1 \qquad (9.38)$$

Therefore, the sensitivity of the Lagrange with respect to A_0 is, in fact, the yield. It should be noted that the second constraint has no constant term. However, if the feed enters with some concentration, B_0, which would appear in Eq.(9.21), then in a similar way used for A_0 it can be proved that:

$$\frac{\partial L^*}{\partial B_o} = \lambda_2 \qquad (9.39)$$

which is the second form of the yield.

Example 9.2: Optimum Yield of a Plug Flow Reactor

Consider the problem of finding the optimum yield in a recycle Plug flow Reactor (PFR) for the following first order autocatatlytic reactions:

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C \qquad (9.40)$$

$$A + A \xrightarrow{k_3} D \tag{9.41}$$

with $K_1 = k_3A_0/k_1$ and $K_2 = k_2/k_1$. The yield B/A_0 for $K_2 = 1$ is given by

$$\frac{B}{A_o} = \frac{\dfrac{c}{1+K_1c}\ln\dfrac{Rc+1}{c(R+1)}}{1-\dfrac{c}{1+K_1c}\dfrac{R}{Rc+1}(1+K_1^{'})} := F(c,R) \tag{9.42}$$

where $c = A/A_0$, R the recycle ratio and $K'_1 = K_1(Rc+1)/(R+1)$. It is desired to maximize the yield over the recycle rate R and the complementary conversion c for the given value of K_1=10. Keeping in mind that both the recycle ratio R is positive $R \geq 0$ and that c is bounded between 0 and 1, i.e., $0 \leq c \leq 1$, the optimization problem is then

$$\max_{c,R} F(c,R) \tag{9.43}$$

Subject to:

$$\begin{aligned} &c \geq 0 \\ &-c+1 \geq 0 \\ &R \geq 0 \end{aligned} \tag{9.44}$$

Since these are inequality constraints, introducing slack variables the problem is equivalent to:

$$\max_{c,R} F(c,R) \tag{9.45}$$

Subject to:

$$c - \sigma_1^2 = 0 \tag{9.46}$$

$$-c + 1 - \sigma_2^2 = 0 \tag{9.47}$$

$$R - \sigma_3^2 = 0 \tag{9.48}$$

The Lagrangian is defined as

$$L(c, R) = F + \lambda_1(c - \sigma_1^2) + \lambda_2(-c + 1 - \sigma_2^2) + \lambda_3(R - \sigma_3^2) \tag{9.49}$$

The necessary conditions for the existence of an optimum are,:

$$\frac{\partial L}{\partial c} = \frac{\partial F}{\partial c} + \lambda_1 - \lambda_2 = 0 \tag{9.50}$$

$$\frac{\partial L}{\partial R} = \frac{\partial F}{\partial R} + \lambda_3 = 0 \tag{9.51}$$

$$\frac{\partial L}{\partial \lambda_1} = (c - \sigma_1^2) = 0 \tag{9.52}$$

$$\frac{\partial L}{\partial \lambda_2} = (-c + 1 - \sigma_2^2) = 0 \tag{9.53}$$

$$\frac{\partial L}{\partial \lambda_3} = (R - \sigma_3^2) = 0 \tag{9.54}$$

$$\frac{\partial L}{\partial \sigma_1} = -2\lambda_1\sigma_1 = 0 \tag{9.55}$$

$$\frac{\partial L}{\partial \sigma_2} = -2\lambda_2\sigma_2 = 0 \tag{9.56}$$

$$\frac{\partial L}{\partial \sigma_3} = -2\lambda_3\sigma_3 = 0 \tag{9.57}$$

The partial derivatives can be obtained analytically,

$$\frac{\partial F}{\partial c} = \frac{(1+r)\left[-1 - ck_1 - r - c^2k_1r + \log\left(\frac{1+cr}{c(1+r)}\right) + r\log\left(\frac{1+cr}{c(1+r)}\right) + 2cr\log\left(\frac{1+cr}{c(1+r)}\right) + 2cr^2\log\left(\frac{1+cr}{c(1+r)}\right)\right]}{\left(1 + ck_1 + r + c^2k_1r\right)^2} \tag{9.58}$$

$$\frac{\partial F}{\partial r} = \frac{c(-1+c)}{1 + ck_1 + r + c^2k_1r} + \frac{\left[c^2(1 + k_1 + 2r - 2ck_1r + r^2 + 2c^2k_1r^2)\log\left(\frac{1+cr}{c(1+r)}\right)\right]}{\left(1 + ck_1 + r + c^2k_1r\right)^2} \tag{9.59}$$

Equations (9.55 to 9.57) represent different conditions on (λ_i,σ_i), i = 1,2,3. Excluding the trivial case of c = 1 (no reaction) that is obtained form $\sigma_2^2 = 0$, and the asymptotic case of c=0 (complete conversion) that results from $\sigma_1^2 = 0$, the remaining cases are the following:

(1) ($\sigma_1^2 \neq 0$, $\sigma_2^2 \neq 0$, $\sigma_3^2 = 0$) and (2) ($\sigma_1^2 \neq 0$, $\sigma_2^2 \neq 0$, $\sigma_3^2 \neq 0$) .

We examine in the following each of these cases.

- Case 1: The case of ($\sigma_1^2 \neq 0$, $\sigma_2^2 \neq 0$, $\sigma_3^2 = 0$) yields $\lambda_1 = 0$, $\lambda_2 = 0$ and $R=0$.
  This is the case of a PFR without recycle. The remaining system of equations is equivalent to:

$$0 = \frac{\partial F}{\partial c} \tag{9.60}$$

$$-\lambda_3 = \frac{\partial F}{\partial R} \tag{9.61}$$

$$c - \sigma_1^2 = 0 \tag{9.62}$$

$$-c+1-\sigma_2^2=0 \tag{9.63}$$

Substituting R=0 into Eq.(9.58) and solving the non-linear algebraic equation in c (using Newton-Raphson method for instance) yields c=0.1156. Substituting in Eqs.(9.61-9.63) yields $\sigma_1 = 0.340$, $\sigma_2 = 0.9043$, $\lambda_3 = 2.0824\times 10^{-2}$. The optimum yield is 0.1156. All the λ_i are positive, therefore the point (c = 0.1156, R = 0) is a maximum.

Case 2: The case of ($\sigma_1^2 \neq 0$, $\sigma_2^2 \neq 0$, $\sigma_3^2 \neq 0$) yields $\lambda_1 = 0$, $\lambda_2 = 0$ and $\lambda_3 = 0$.
The rest of equations to be solved are:

$$\frac{\partial F}{\partial c}=0 \tag{9.64}$$

$$\frac{\partial F}{\partial R}=0 \tag{9.65}$$

$$c-\sigma_1^2=0 \tag{9.66}$$

$$-c+1-\sigma_2^2=0 \tag{9.67}$$

$$R-\sigma_3^2=0 \tag{9.68}$$

The first two equations are to be solved simultaneously. The solution yields c = 0.1522, R = 1.7829 and a yield of 0.1254. Substituting in Eqs.(9.66-9.68) yields σ_1^2 = 0.1522, σ_2^2=0.8478, σ_3^2 = 1.7829. Note that all the λ_i are zero. To characterize the nature of the stationary point we need further analysis. But it is simpler to plot the variations of the yield for the value of R=1.7829. Figure 9.1shows that that there is indeed a maximum at c=0.1522. This example shows that, in the case of autocatalytic reactions the optimum does not occur at the simple PFR without recycle rather some back mixing is desirable.

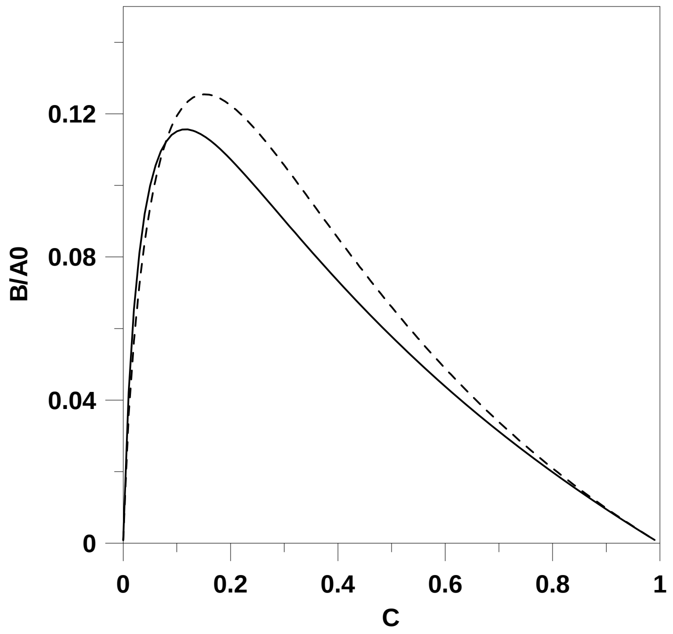


Figure 1-1: Yield versus conversion , solid (R=0); dash (R=1.7829)

9.5 OTHER SOLUTION TECHNIQUES

An important class of constrained optimization problems is when the objective function is quadratic and the constraints are linear. The problem is called a Quadratic Programming (QP) problem and it is as follows:

$$\underset{x=[x_1,\cdots,x_n]}{\text{minimize}}(or \underset{x=[x_1,\cdots,x_n]}{\text{maximize}}) f(x) = \frac{1}{2} x^T Q x + c^T x \tag{5.69}$$

Subject to

$$\begin{aligned} a_i^T x &= b_i, \qquad i = 1,\cdots,I \\ a_j^T x &\geq b_j, \qquad j = 1,\cdots,J \end{aligned} \tag{5.70}$$

The matrix Q can always be arranged in such a way that it is symmetric.

It can be shown that the Hessian is the matrix Q itself. So when Q is positive definite any solution to the QP problem is also a global and unique optimum. There are special procedures to solve QP problems (See reference [33,39,49] for details).

For the numerical solution of the general nonlinear multivariable constrained optimization problem, a variety of approaches are being used. They are beyond the scope of this book. But we present a small survey of the concepts behind each class.

- Iterative linearization methods: The idea of iterative linearization is to linearize the optimization problem and then apply linear programming methods (to be presented in next chapter) to solve the problem. One famous method that belongs to this class is Generalized Reduced Gradient (GRG) method (See references [33,39,49] for details of the technique).

- Penalty function: The idea of this method is to transform the constrained problem into an unconstrained problem which becomes easier to solve. For example given the constrained problem:

$$\begin{aligned} &\text{Minimize } f(x) \\ &\text{Subject to } h(x) = 0 \end{aligned} \tag{5.71}$$

We can formulate a new objective function in this way

$$P(x, r) = f(x) + r(h(x))^2 \tag{5.72}$$

Where r is a scalar weight. Note that $P(x,r)$ is a function where the square of the constraint $h(x)$ was added as a "penalty". Now we attempt to minimize the new function $P(x,r)$ which becomes an unconstrained optimization problem. If at the optimum x^* the value of the penalty is zero then it is the optimum for $f(x)$ and we also have $P(x^*) = f(x^*)$

- Successive quadratic programming. This idea of this technique is that at each iteration the objective function is approximated locally by a quadratic function and the constrained by a linear function. The problem becomes a quadratic programming (QP) problem and it is at each step to obtain a new search direction. (See references [33,39,49] for more details.)

9.6 IMSL ROUTINES

Some IMSL routines to solve multivariable constrained optimization problems are as follow:

Routine	Features
QPROG	Solve a quadratic programming problem subject to linear equality/inequality constraints.
NCONF	Solve a general nonlinear programming problem using the successive quadratic programming algorithm and a finite difference gradient.
NCONG	Same as NCONF but with a user-supplied gradient.