Lecture - 11

NONLINEAR PROGRAMMING WITH CONSTRAINTS

Reference: Book Chapter 13

CONSTRAINED OPTIMIZATION PROBLEM

$$\text{Find } \mathbf{X} \text{ which minimizes } f(\mathbf{X})$$

subject to

$$g_j(\mathbf{X}) \leq 0, \quad j = 1, 2, \ldots, m$$

$$h_k(\mathbf{X}) = 0, \quad k = 1, 2, \ldots, p \tag{7.1}$$

There are many techniques available for the solution of a constrained nonlinear programming problem. All the methods can be classified into two broad categories: direct methods and indirect methods, as shown in Table 7.1. In the *direct methods*, the constraints are handled in an explicit manner, whereas in most of the *indirect methods*, the constrained problem is solved as a sequence of unconstrained minimization problems. We discuss in this chapter all the methods indicated in Table 7.1.

CONSTRAINED OPTIMIZATION PROBLEM

Table 7.1 Constrained Optimization Techniques

Direct methods	Indirect methods
Random search methods	Transformation of variables technique
Heuristic search methods	Sequential unconstrained minimization techniques
Complex method	Interior penalty function method
Objective and constraint approximation methods	Exterior penalty function method
Sequential linear programming method	Augmented Lagrange multiplier method
Sequential quadratic programming method	
Methods of feasible directions	
Zoutendijk's method	
Rosen's gradient projection method	
Generalized reduced gradient method	

CHARACTERISTICS OF A CONSTRAINED PROBLEM

1. The constraints may have no effect on the optimum point; that is, the constrained minimum is the same as the unconstrained minimum as shown in Fig. 7.1. In this case the minimum point $\mathbf{X}^*$ can be found by making use of the necessary and sufficient conditions

$$\nabla f|_{\mathbf{X}^*} = \mathbf{0} \tag{7.2}$$

$$\mathbf{J}_{\mathbf{X}^*} = \left[\frac{\partial^2 f}{\partial x_i \partial x_j} \right]_{\mathbf{X}^*} = \text{positive definite} \tag{7.3}$$

However, to use these conditions, one must be certain that the constraints are not going to have any effect on the minimum. For simple optimization problems like the one shown in Fig. 7.1, it may be possible to determine beforehand whether or not the constraints have an influence on the minimum point. However, in most practical problems, even if we have a situation as shown in Fig. 7.1, it will be extremely difficult to identify it. Thus one has to proceed with the general assumption that the constraints have some influence on the optimum point.

CHARACTERISTICS OF A CONSTRAINED PROBLEM

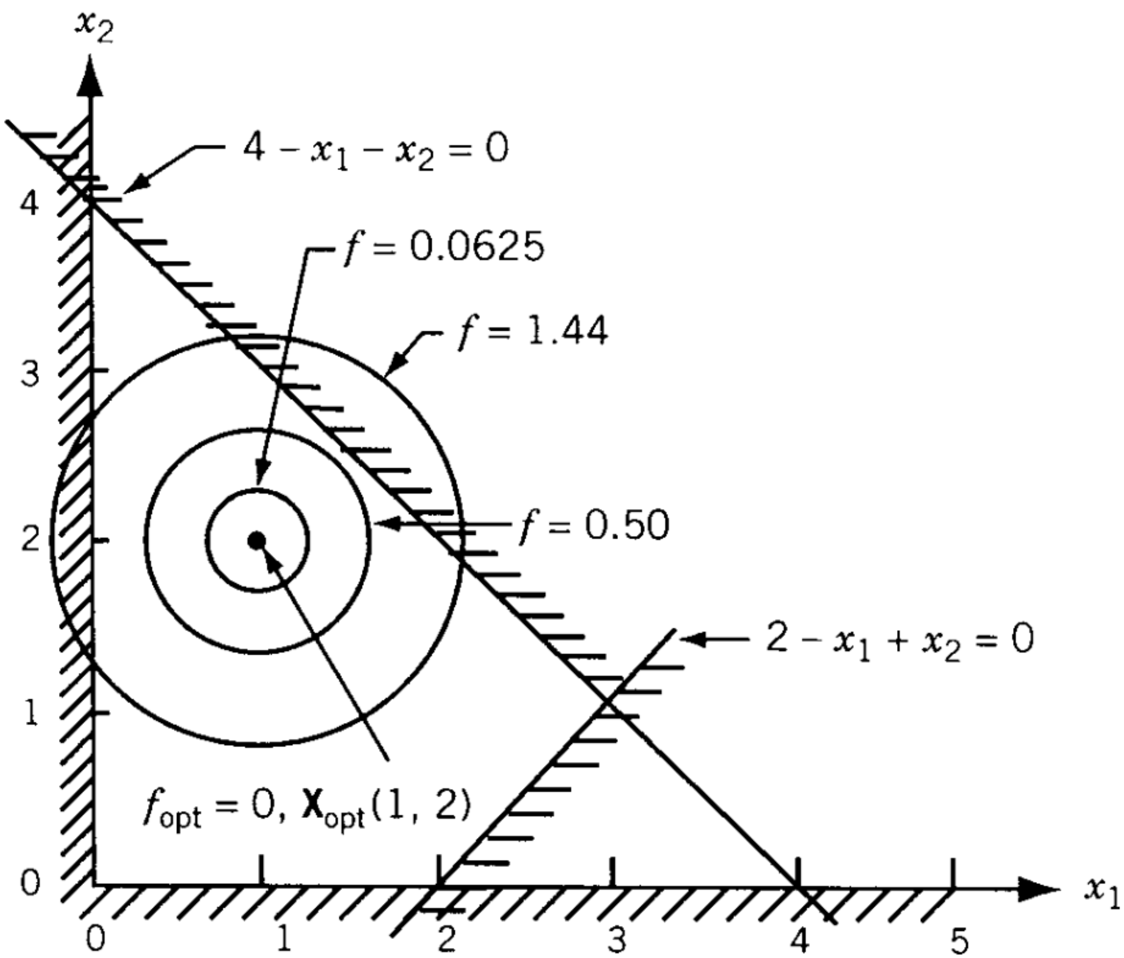


Figure 7.1 Constrained and unconstrained minima are the same (linear constraints).

CHARACTERISTICS OF A CONSTRAINED PROBLEM

2. The optimum (unique) solution occurs on a constraint boundary as shown in Fig. 7.2. In this case the Kuhn–Tucker necessary conditions indicate that the negative of the gradient must be expressible as a positive linear combination of the gradients of the active constraints.

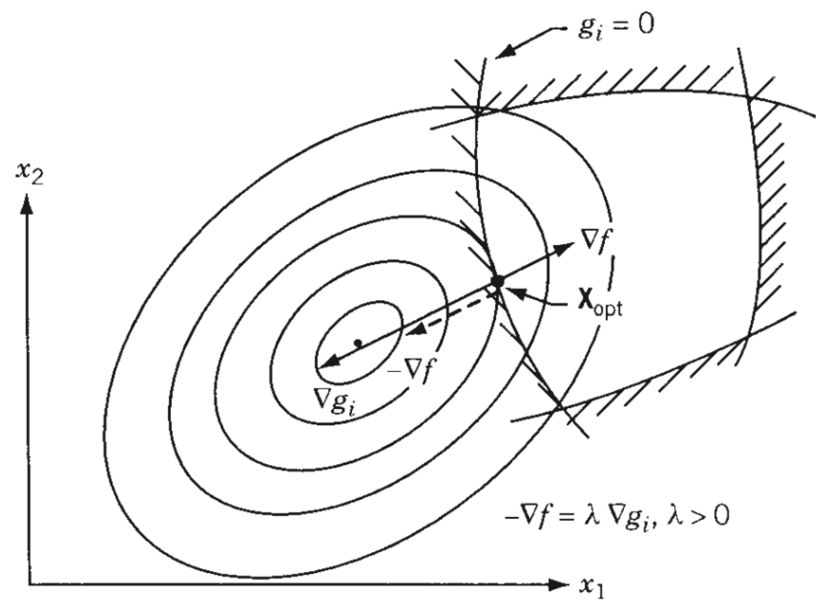


Figure 7.2 Constrained minimum occurring on a nonlinear constraint.

CHARACTERISTICS OF A CONSTRAINED PROBLEM

3. If the objective function has two or more unconstrained local minima, the constrained problem may have multiple minima as shown in Fig. 7.3.

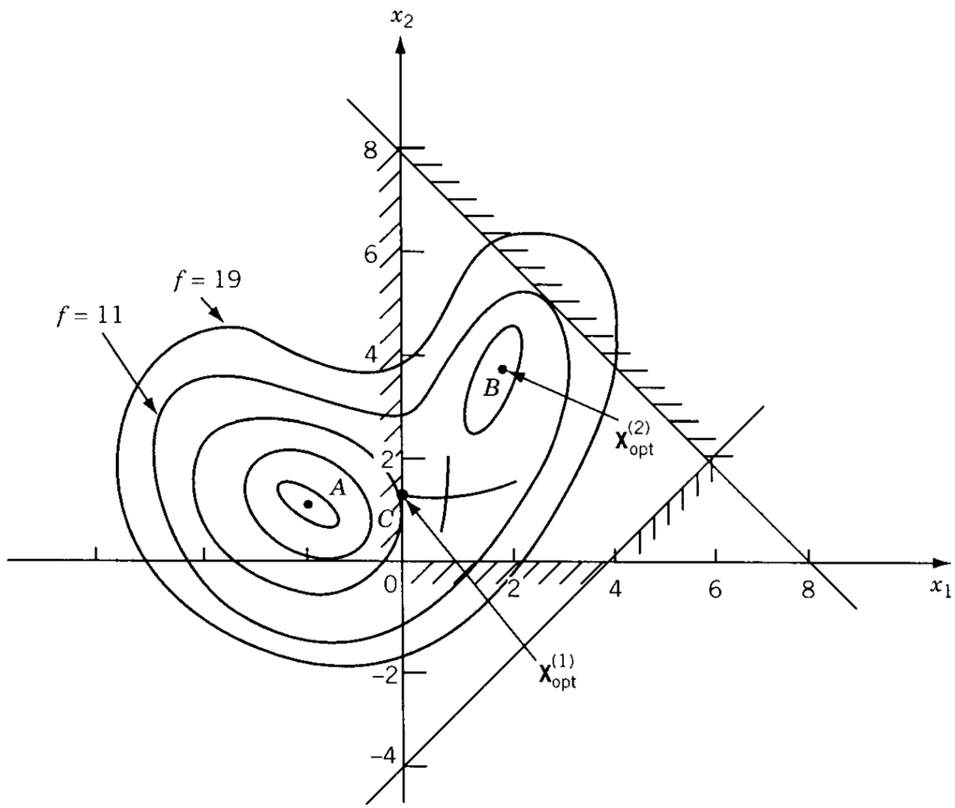


Figure 7.3 Relative minima introduced by objective function.

CHARACTERISTICS OF A CONSTRAINED PROBLEM

4. In some cases, even if the objective function has a single unconstrained minimum, the constraints may introduce multiple local minima as shown in Fig. 7.4.

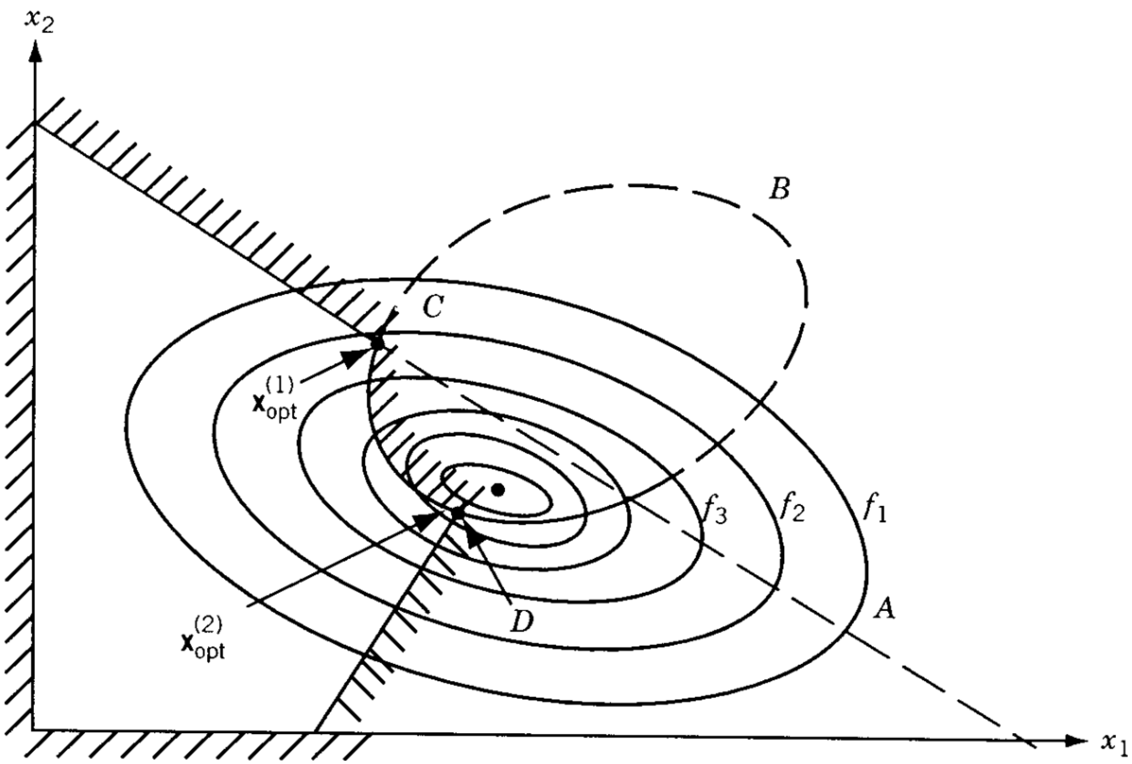


Figure 7.4 Relative minima introduced by constraints.

INDIRECT METHODS

- Elimination method
- Transformation techniques
- Penalty methods
  - Interior point methods
  - Exterior point methods

ELIMINATION METHOD

- If an optimization problem only has equality constraints, it can be solved as an unconstrained problem by eliminating independent variables.

$$\begin{aligned} &\min && f(x_1, x_2, ..., x_N) \\ &\text{subject to} && \end{aligned}$$

$$h_j(x_1, x_2, ..., x_N) = 0, j = 1, ..., n$$

- Suppose N > n and the n equality constraints are independent. Solving the equality constraints and substitute them into the objective function. It reduces the number of variables from N to N−n.

ELIMINATION METHOD

- Example

$$\min \quad f(x) = x_1 x_2 - x_3 - 3$$

subject to

$$x_3 - 4x_1 = 0$$

$$x_2 - 2x_1 - x_3 - 2 = 0$$

$$x_2 = 6x_1 + 2$$

$$x_3 = 4x_1$$

$$\min \quad f(x) = 6x_1^2 - 2x_1 - 3$$

$x_1^* = 0.1667$, $x_2^* = 3$, $x_3^* = 0.667$; and $f^* = -3.1667$

TRANSFORMATION TECHNIQUES

- If the constraints g_j (X) are explicit functions of the variables x_i and have certain simple forms, it may be possible to make a transformation of the independent variables such that the constraints are satisfied automatically. Thus it may be possible to convert a constrained optimization problem into an unconstrained one by making a change of variables. Some typical transformations are:

1. If lower and upper bounds on x_i are specified as

$$l_i \leq x_i \leq u_i \tag{7.148}$$

these can be satisfied by transforming the variable x_i as

$$x_i = l_i + (u_i - l_i)\sin^2 y_i \tag{7.149}$$

where y_i is the new variable, which can take any value.

TRANSFORMATION TECHNIQUES

2. If a variable x_i is restricted to lie in the interval (0, 1), we can use the transformation:

$$x_i = \sin^2 y_i, \quad x_i = \cos^2 y_i$$

$$x_i = \frac{e^{yi}}{e^{yi} + e^{-yi}} \quad \text{or} \quad x_i = \frac{y_i^2}{1 + y_i^2} \tag{7.150}$$

3. If the variable x_i is constrained to take only positive values, the transformation can be

$$x_i = \text{abs}(y_i), \quad x_i = y_i^2 \quad \text{or} \quad x_i = e^{yi} \tag{7.151}$$

4. If the variable is restricted to take values lying only in between −1 and 1, the transformation can be

$$x_i = \sin y_i, \quad x_i = \cos y_i, \quad \text{or} \quad x_i = \frac{2y_i}{1 + y_i^2} \tag{7.152}$$

TRANSFORMATION TECHNIQUES

Note the following aspects of transformation techniques:

1. The constraints $g_j(\mathbf{X})$ have to be very simple functions of x_i.
2. For certain constraints it may not be possible to find the necessary transformation.
3. If it is not possible to eliminate all the constraints by making a change of variables, it may be better not to use the transformation at all. The partial transformation may sometimes produce a distorted objective function which might be more difficult to minimize than the original function.

Example 7.6 Find the dimensions of a rectangular prism-type box that has the largest volume when the sum of its length, width, and height is limited to a maximum value of 60 in. and its length is restricted to a maximum value of 36 in.

SOLUTION Let x_1, x_2, and x_3 denote the length, width, and height of the box, respectively. The problem can be stated as follows:

$$\text{Maximize } f(x_1, x_2, x_3) = x_1 x_2 x_3 \tag{E$_1$}$$

subject to

$$x_1 + x_2 + x_3 \le 60 \tag{E$_2$}$$

$$x_1 \le 36 \tag{E$_3$}$$

$$x_i \ge 0, \quad i = 1, 2, 3 \tag{E$_4$}$$

By introducing new variables as

$$y_1 = x_1, \quad y_2 = x_2, \quad y_3 = x_1 + x_2 + x_3 \tag{E$_5$}$$

or

$$x_1 = y_1, \quad x_2 = y_2, \quad x_3 = y_3 - y_1 - y_2 \tag{E$_6$}$$

TRANSFORMATION TECHNIQUES

the constraints of Eqs. (E_2) to (E_4) can be restated as

$$0 \le y_1 \le 36, \quad 0 \le y_2 \le 60, \quad 0 \le y_3 \le 60 \tag{E$_7$}$$

where the upper bound, for example, on y_2 is obtained by setting $x_1 = x_3 = 0$ in Eq. (E_2). The constraints of Eq. (E_7) will be satisfied automatically if we define new variables z_i, $i = 1, 2, 3$, as

$$y_1 = 36 \sin^2 z_1, \quad y_2 = 60 \sin^2 z_2, \quad y_3 = 60 \sin^2 z_3 \tag{E$_8$}$$

Thus the problem can be stated as an unconstrained problem as follows:

$$\begin{aligned} &\text{Maximize } f(z_1, z_2, z_3) \\ &= y_1 y_2 (y_3 - y_1 - y_2) \\ &= 2160 \sin^2 z_1 \sin^2 z_2 (60 \sin^2 z_3 - 36 \sin^2 z_1 - 60 \sin^2 z_2) \end{aligned} \tag{E$_9$}$$

The necessary conditions of optimality yield the relations

$$\frac{\partial f}{\partial z_1} = 259{,}200 \sin z_1 \cos z_1 \sin^2 z_2 (\sin^2 z_3 - \tfrac{6}{5} \sin^2 z_1 - \sin^2 z_2) = 0 \qquad (E_{10})$$

$$\frac{\partial f}{\partial z_2} = 518{,}400 \sin^2 z_1 \sin z_2 \cos z_2 (\tfrac{1}{2} \sin^2 z_3 - \tfrac{3}{10} \sin^2 z_1 - \sin^2 z_2) = 0 \qquad (E_{11})$$

$$\frac{\partial f}{\partial z_3} = 259{,}200 \sin^2 z_1 \sin^2 z_2 \sin z_3 \cos z_3 = 0 \qquad (E_{12})$$

Equation (E_{12}) gives the nontrivial solution as $\cos\ z_3 = 0$ or $\sin^2 z_3 = 1$. Hence Eqs. (E_{10}) and (E_{11}) yield $\sin^2 z_1 = \frac{5}{9}$ and $\sin^2 z_2 = \frac{1}{3}$. Thus the optimum solution is given by $x_1^* = 20$ in., $x_2^* = 20$ in., $x_3^* = 20$ in., and the maximum volume $= 8000$ in^3.

PENALTY METHODS

- It replaces the original optimization problem by a sequence of subproblems in which the constraints are represented by terms added to the objective.
- In the limit, the solutions of the subproblems will converge to the solution of the original constrained problem.

- Two groups
  - Interior point methods: It generates a sequence of feasible points.
  - Exterior point methods: It generates a sequence of infeasible points.

PENALTY METHODS

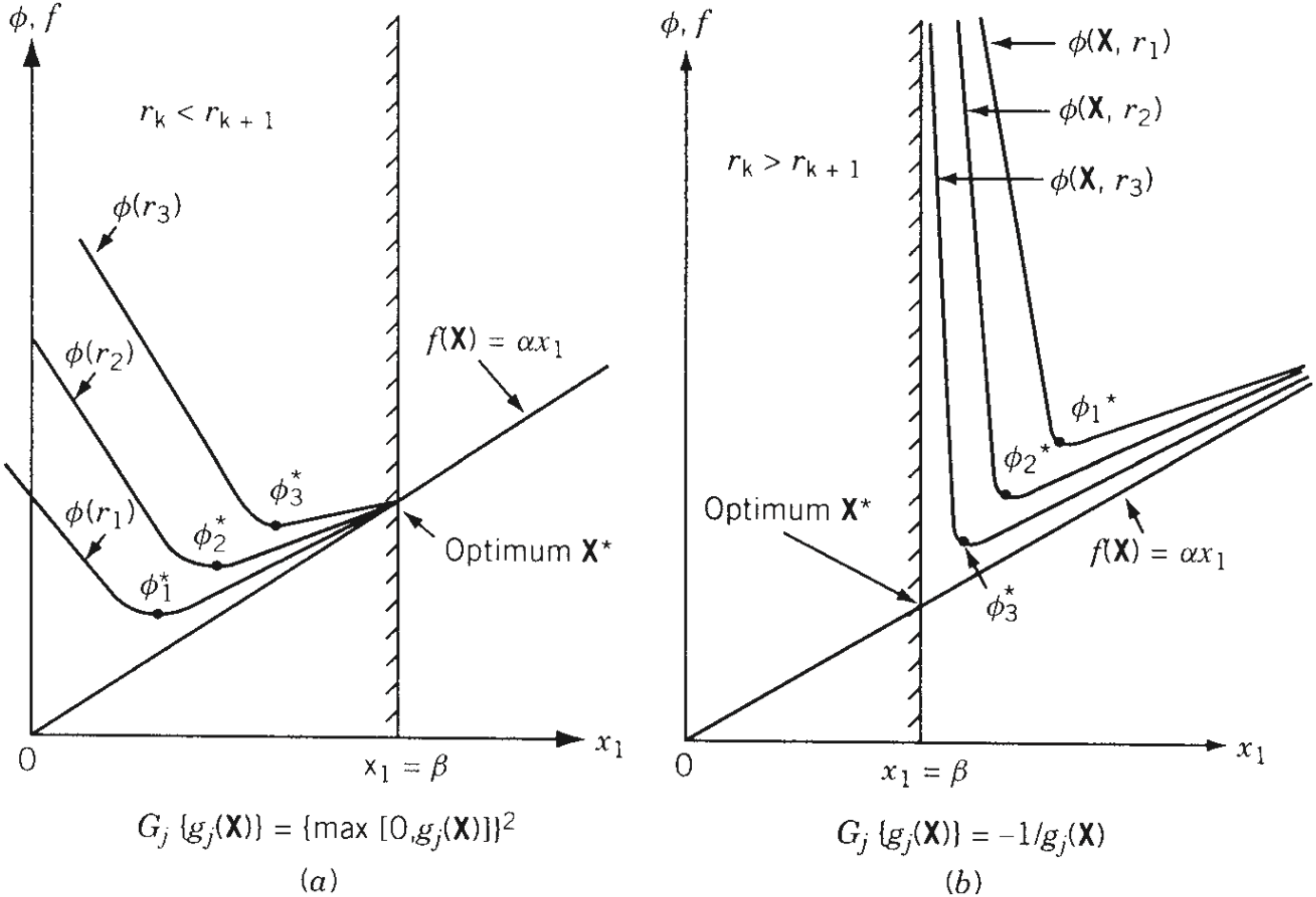


Figure 7.10 Penalty function methods: (*a*) exterior method; (*b*) interior method.

PENALTY METHODS

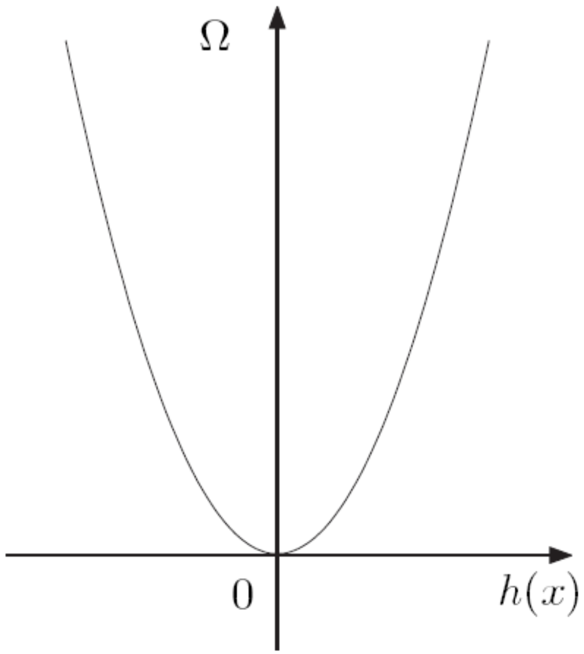
- Reformulate to an unconstrained optimization problem.
- Two parts:
  - the objective function of the original problem, and
  - a penalty term.

$$P(x, R) = f(x) + \Omega(R, g(x), h(x))$$

- R: penalty parameters. Updated using rules.
- Exterior point method: The stationary point of *P(x,R)* is infeasible. The updated parameter, R, forces the stationary point to be closer to the feasible region.
- Interior point method (or barrier method): The form forces stationary point of the unconstrained function *P(x,R)* to be feasible.

PENALTY METHODS – PARABOLIC PENALTY

- Different choices of penalty forms
  - Parabolic penalty for equality constraints: $\Omega = Rh_j^2(x)$
  - It is 0 only at the point where *h(x)* = 0.
  - As *h(x)* stays farther from 0, it becomes larger.

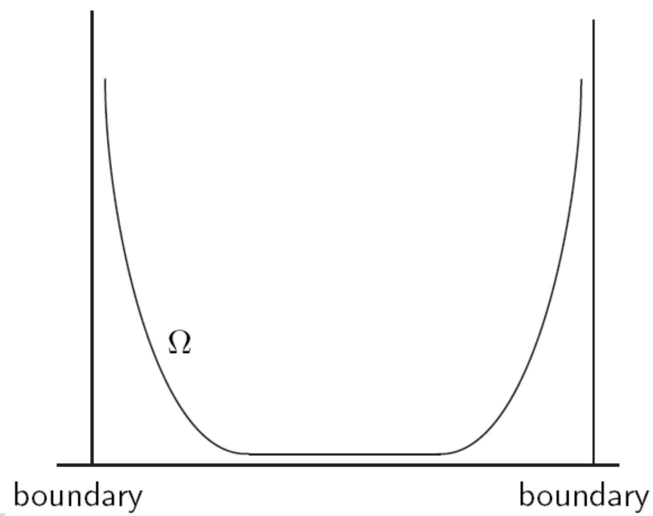


PENALTY METHODS - LOGARITHMIC

- Different choices of penalty forms
  - Logarithmic function for inequality constraints:

$$\Omega = R(-log(-g_i(x)))$$

  - It is usually used for the interior point method.
  - The initial point is in the feasible region.
  - As the point approaches the boundary of the feasible region, $-g_i(x)$ gets closer to 0. The penalty term approaches positive infinity.
  - As the penalty parameter R approaches 0, the function P(x, R) approaches the original objective f(x).



PENALTY METHODS – INVERSE PENALTY

- Another penalty function is the Inverse Penalty

$$\Omega = R\frac{1}{g_i(x)}$$

- When the argument should become too close to zero, the situation becomes problematic.

INTERIOR POINT METHOD

- Example

$$\min_{x} \quad f(x) = (x_1 - 3)^2 + (x_2 - 3)^2$$

subject to

$$h(x) = x_1 + x_2 - 4 = 0$$

$$P(x, R) = (x_1 - 3)^2 + (x_2 - 3)^2 + R(x_1 + x_2 - 4)^2$$

$$\frac{\partial P}{\partial x_1} = 2(x_1 - 3) + 2R(x_1 + x_2 - 4) = 0$$

$$\frac{\partial P}{\partial x_2} = 2(x_2 - 3) + 2R(x_1 + x_2 - 4) = 0$$

$$x_1 = x_2 = \frac{6 + 8R}{2 + 4R}$$

$$x_1^* = x_2^* = \lim_{R \to Inf} \frac{6 + 8R}{2 + 4R} = 2$$

R	x_1^*	x_2^*	f^*
10	2.0476	2.0476	1.9048
100	2.0050	2.0050	1.9900
1000	2.0005	2.0005	1.9990
10000	2.0000	2.0000	1.9999
100000	2.0000	2.0000	2.0000

INTERIOR POINT METHOD

- Example

$$\min_{x} \quad f(x) = -2x_1 + x_2 + 5$$

subject to

$$g_1(x) = x_1^2 - x_2 - 1 \leq 0,$$
$$g_2(x) = -x_1 \leq 0,$$

$$P(x, R) = (-2x_1 + x_2 + 5) + R(-ln(-(x_1^2 - x_2 - 1))) + R(-ln(x_1))$$

$$\frac{\partial P}{\partial x_1} = -2 + \frac{-2Rx_1}{x_1^2 - x_2 - 1} - \frac{R}{x_1} = 0 \qquad \frac{\partial P}{\partial x_2} = 1 + \frac{R}{x_1^2 - x_2 - 1} = 0$$

$$x_1 = \frac{1 + \sqrt{2R+1}}{2} \qquad x_2 = \frac{3R - 1 + \sqrt{2R+1}}{2}$$

$$x_1^* = \lim_{R \to 0} \frac{1 + \sqrt{2R+1}}{2} = 1$$

$$x_2^* = \lim_{R \to 0} \frac{3R - 1 + \sqrt{2R+1}}{2} = 0$$

R	x_1^*	x_2^*	f^*	$g_1(x)$	$g_2(x)$
1	1.366	1.866	4.134	-1.000	-1.366
0.1	1.048	0.198	3.102	-0.100	-1.048
0.01	1.005	0.020	3.010	-0.010	-1.005
0.001	1.000	0.002	3.001	-0.001	-1.000
0.0001	1.000	0.000	3.000	0.000	-1.000

EXTERIOR POINT METHOD

In the exterior penalty function method, the ϕ function is generally taken as

$$\phi(\mathbf{X}, r_k) = f(\mathbf{X}) + r_k \sum_{j=1}^{m} \langle g_j(\mathbf{X}) \rangle^q \tag{7.199}$$

where r_k is a positive penalty parameter, the exponent q is a nonnegative constant, and the bracket function $\langle g_j(\mathbf{X}) \rangle$ is defined as

$$\begin{aligned} \langle g_j(\mathbf{X}) \rangle &= \max \langle g_j(\mathbf{X}), 0 \rangle \\ &= \begin{cases} g_j(\mathbf{X}) & \text{if } g_j(\mathbf{X}) > 0 \\ & \quad \text{(constraint is violated)} \\ 0 & \text{if } g_j(\mathbf{X}) \le 0 \\ & \quad \text{(constraint is satisfied)} \end{cases} \end{aligned} \tag{7.200}$$

EXTERIOR POINT METHOD

Example 7.9

$$\text{Minimize } f(x_1, x_2) = \tfrac{1}{3}(x_1 + 1)^3 + x_2$$

subject to

$$g_1(x_1, x_2) = 1 - x_1 \le 0$$

$$g_2(x_1, x_2) = -x_2 \le 0$$

SOLUTION To illustrate the exterior penalty function method, we solve the unconstrained minimization problem by using differential calculus method. As such, it is not necessary to have an initial trial point $\mathbf{X}_1$. The ϕ function is

$$\phi(\mathbf{X}_1, r) = \tfrac{1}{3}(x_1 + 1)^3 + x_2 + r[\max(0, 1 - x_1)]^2 + r[\max(0, -x_2)]^2$$

The necessary conditions for the unconstrained minimum of $\phi(\mathbf{X}, r)$ are

$$\frac{\partial \phi}{\partial x_1} = (x_1 + 1)^2 - 2r[\max(0, 1 - x_1)] = 0$$

$$\frac{\partial \phi}{\partial x_2} = 1 - 2r[\max(0, -x_2)] = 0$$

EXTERIOR POINT METHOD

These equations can be written as

$$\min[(x_1+1)^2, (x_1+1)^2 - 2r(1-x_1)] = 0 \tag{E$_1$}$$

$$\min[1, 1+2rx_2] = 0 \tag{E$_2$}$$

In Eq. (E$_1$), if $(x_1+1)^2 = 0$, $x_1 = -1$ (this violates the first constraint), and if

$$(x_1+1)^2 - 2r(1-x_1) = 0, \quad x_1 = -1 - r + \sqrt{r^2+4r}$$

In Eq. (E$_2$), the only possibility is that $1 + 2rx_2 = 0$ and hence $x_2 = -1/2r$. Thus the solution of the unconstrained minimization problem is given by

$$x_1^*(r) = -1 - r + r\left(1 + \frac{4}{r}\right)^{1/2} \tag{E$_3$}$$

$$x_2^*(r) = -\frac{1}{2r} \tag{E$_4$}$$

From this, the solution of the original constrained problem can be obtained as

$$x_1^* = \lim_{r\to\infty} x_1^*(r) = 1, \quad x_2^* = \lim_{r\to\infty} x_2^*(r) = 0$$

$$f_{\min} = \lim_{r\to\infty} \phi_{\min}(r) = \frac{8}{3}$$

EXTERIOR POINT METHOD

The convergence of the method, as r increases gradually, can be seen from Table 7.5.

Table 7.5 Results for Example 7.9

Value of r	x_1^*	x_2^*	$\phi_{\min}(r)$	$f_{\min}(r)$
0.001	−0.93775	−500.00000	−249.9962	−500.0000
0.01	−0.80975	−50.00000	−24.9650	−49.9977
0.1	−0.45969	−5.00000	−2.2344	−4.9474
1	0.23607	−0.50000	0.9631	0.1295
10	0.83216	−0.05000	2.3068	2.0001
100	0.98039	−0.00500	2.6249	2.5840
1,000	0.99800	−0.00050	2.6624	2.6582
10,000	0.99963	−0.00005	2.6655	2.6652
∞	1	0	$\frac{8}{3}$	$\frac{8}{3}$

KARUSH-KUHN TUCKER (KKT) CONDITIONS

- Previous form

$$\min_{x} \quad f(x)$$

subject to

$$g(x) \leq 0$$

$$h(x) = 0$$

$$x_l \leq x \leq x_u$$

- New form

$$\min_{x} \quad f(x)$$

subject to

$$g_i(x) \leq 0, i = 1, ..., m$$

$$h_j(x) = 0, j = 1, ..., n$$

Side constraints are regarded as inequality constraints

KARUSH-KUHN TUCKER CONDITIONS

- Equality constraints

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & \\ & h_j(x) = 0, j = 1, ..., n \end{aligned}$$

- The Lagrangian function

$$L(x, \nu) = f(x) + \sum_{j=1}^{n} \nu_j h_j(x)$$

- v_j : the Lagrange multiplier. No sign restrictions.
- Suppose the minimum for the unconstrained problem is x^* and x satisfies $h_j(x^*) = 0$.
- For any x satisfies $h_j(x) = 0$, $L(x, v) = f(x)$.
- x^* is the minimum of the optimization problem.

KARUSH-KUHN TUCKER CONDITIONS

- Steps
  1) Construct the Lagrangian function, $L(x, v)$, using the objective function and the equality constraints
  2) Solve $\nabla_x L(x, v) = 0$ and $h(x) = 0$.

KARUSH-KUHN TUCKER CONDITIONS

- Example

$$\min_{x} \qquad f(x) = x_1^2 + x_2^2$$

subject to

$$h_1(x_1, x_2) = 2x_1 + x_2 - 2 = 0$$

$$L(x, \nu) = x_1^2 + x_2^2 + \nu(2x_1 + x_2 - 2)$$

$$\frac{\partial L}{\partial x_1} = 2x_1 + 2\nu = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 + \nu = 0$$

$$Hessian = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

x_1*= 0.8 and x_2*= 0.4; v* = −0.8; and f*=0.8.

KARUSH-KUHN TUCKER CONDITIONS

- Inequality constraints
  - The Lagrangian function

$$L(x, \nu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$$

  - Also satisfy

$$\lambda_i^* g_i(x^*) = 0, \text{for all } i = 1, ..., m$$

$$g_i(x^*) \leq 0, \text{for all } i = 1, ..., m$$

$$\lambda_i \geq 0$$

KARUSH-KUHN TUCKER CONDITIONS

- The active set
  - The active set at any feasible x consists of the equality constraint indices together with the indices of the inequality constraints for which $g_i(x) = 0$.

KARUSH-KUHN TUCKER CONDITIONS

- Two cases for an inequality constraint

  - Case I: The inequality constraint is inactive at the minimum point x. It means $g_i(x) < 0$. Since the minimum point, x, satisfies $\lambda_i \, g_i(x) = 0$, the Lagrange multiplier, $\lambda_i = 0$. Then $L(x, \lambda) = f(x)$. $\nabla L(x, \lambda) = \nabla f(x) = 0$ at the optimal point.
  - Case II: The inequality constraint is active at the minimum point x. It means $g_i(x) = 0$. Then $L(x, \lambda) = f(x)$.

KARUSH-KUHN TUCKER CONDITIONS

- Example

$$\min_{x} \quad f(x) = x_1 + x_2$$

$$g_1(x_1, x_2) = x_1^2 + x_2^2 - 1 \leq 0$$

The Lagrangian function is given by

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

$$\frac{\partial L}{\partial x_1} = 1 + 2\lambda x_1 = 0$$

$$\frac{\partial L}{\partial x_2} = 1 + 2\lambda x_2 = 0$$

$$\lambda(x_1^2 + x_2^2 - 1) = 0$$

$$(x_1^2 + x_2^2 - 1) \leq 0$$

$$\lambda \geq 0$$

KARUSH-KUHN TUCKER CONDITIONS

- Example (continue)
  - Since $\lambda = 0$ does not satisfy $\nabla L = 0$, λ is greater then zero. Then $(x_1^2 + x_2^2 - 1) = 0$
  - From $\nabla L = 0$, $x1 = x2$.
  - Since $\lambda > 0$, from $\nabla L = 0$, $x1 < 0$ and $x2 < 0$.
  - Solve

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

  - $x^*=$ $(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$
  - $f^*=$ $-\sqrt{2}$

KARUSH-KUHN TUCKER CONDITIONS

- The Lagrangian function for an optimization problem with multiple constraints

$$L(x, \lambda, \nu) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x) + \sum_{j=1}^{m} \nu_j h_j(x)$$

KARUSH-KUHN TUCKER CONDITIONS

- Linear independence constraint qualification(LICQ)
  - At the feasible point x, if the gradients of the constraints in the active set are linearly independent, we say that the LICQ holds.
  - Example: two inequality constraints

$$g_1(x) = x_1^2 + x_2^2 - 1 \leq 0$$

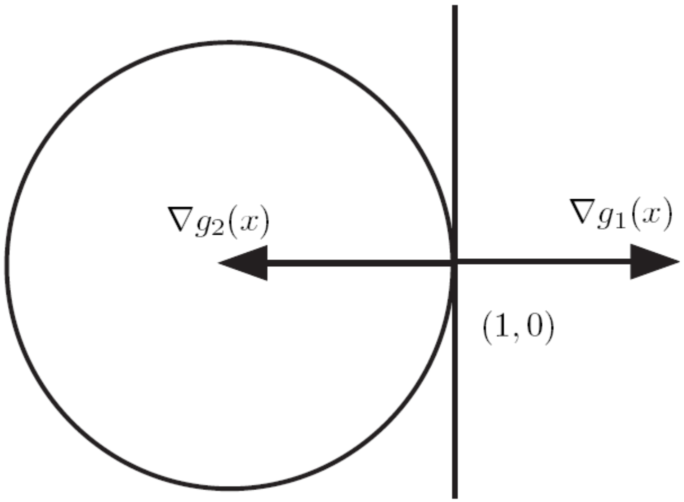
$$g_2(x) = 1 - x_1 \leq 0$$

  - The only feasible point is (1, 0).

$$\nabla g_1(x) = [2x_1, 2x_2]^T = [2, 0]^T$$

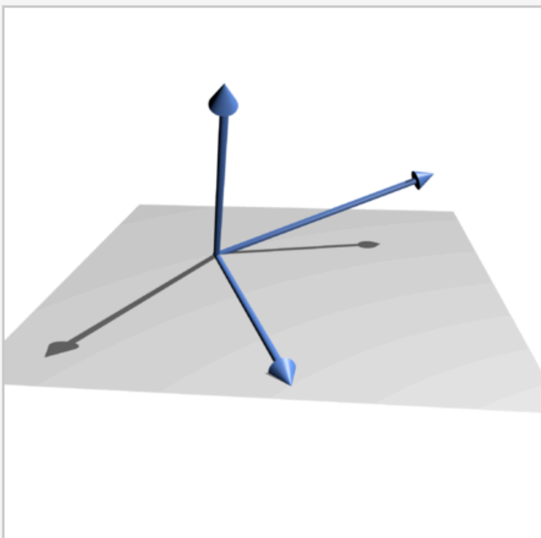
$$\nabla g_2(x) = [-1, 0]^T$$

  - They are not linearly independent.
  - LICQ does not hold at the point.

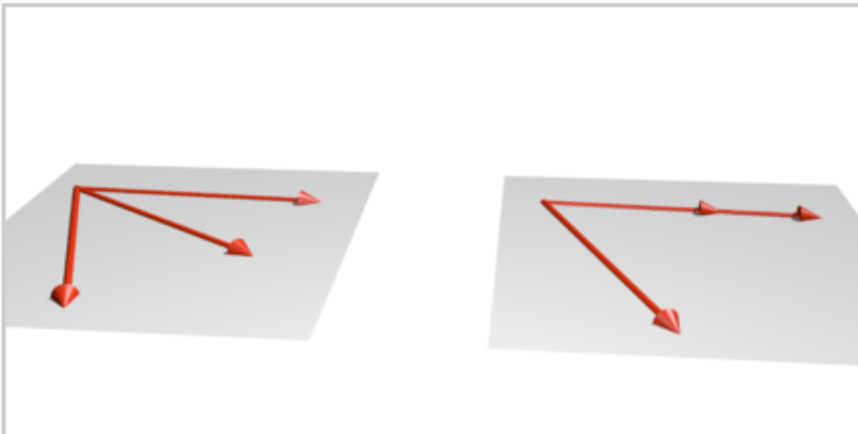


LINEAR INDEPENDENCE

- In the theory of vector spaces, a set of vectors is said to be linearly dependent if one of the vectors in the set can be defined as a linear combination of the others; if no vector in the set can be written in this way, then the vectors are said to be linearly independent.



Linearly independent vectors in $\mathbb{R}^3$



Linearly dependent vectors in a plane in $\mathbb{R}^3$.

- Karush-Kuhn-Tucker Conditions
  - Suppose that *x** is a local minimum solution of *f(x)* subjection to constraints $g_i(x) \le 0$ $(i = 1, ...,m)$ and $h_j(x) = 0$ $(j = 1, ..., n)$. The objective function, *f(x)*, and constraints, $g_i(x)$ and $h_j(x)$ are continuously differentiable. The LICQ holds at *x**. Then there are Lagrange multipliers , such that the following conditions are satisfied at *(x*, λ*, ν*)*.

$$L(x, \lambda, \nu) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x) + \sum_{j=1}^{m} \nu_j h_j(x)$$

$$\nabla_x f(x^*, \lambda^*, \nu^*) = 0$$

$$\lambda_i^* \ge 0, \quad \text{for all } i = 1, ..., m$$

$$\lambda_i^* g_i(x^*) = 0, \quad \text{for all } i = 1, ..., m$$

$$\nu_j^* h_j(x^*) = 0, \quad \text{for all } j = 1, ..., n$$

$$g_i(x^*) \le 0, \quad \text{for all } i = 1, ..., m$$

$$h_j(x^*) = 0, \quad \text{for all } j = 1, ..., n$$

KARUSH-KUHN TUCKER CONDITIONS

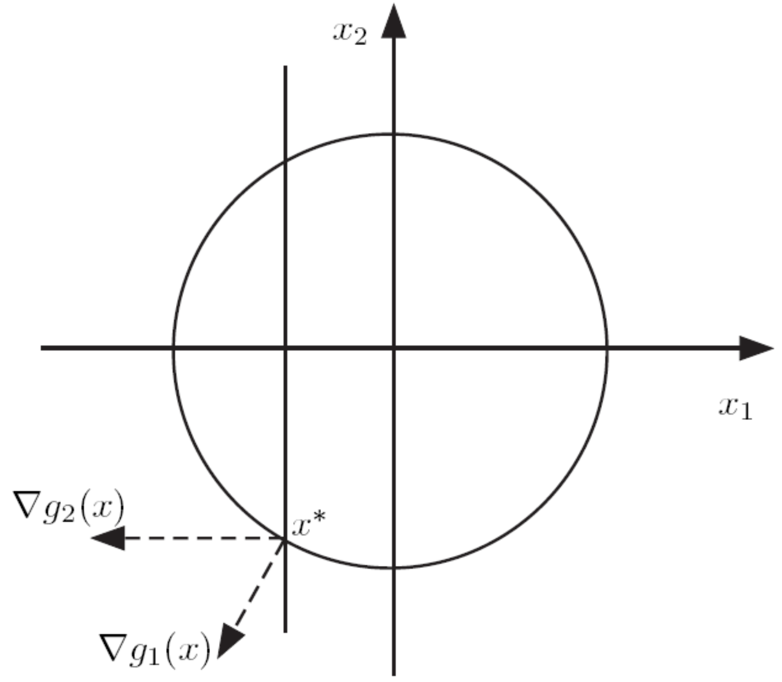
- Example

$$\min_{x} \quad f(x) = x_1 + x_2$$

subject to

$$g_1(x_1, x_2) = x_1^2 + x_2^2 - 1 \leq 0,$$

$$g_2(x_1, x_2) = -x_1 - \tfrac{1}{2} \leq 0,$$



KARUSH-KUHN TUCKER CONDITIONS

- KKT conditions:

$$L(x,\lambda) = x_1 + x_2 + \lambda_1(x_1^2 + x_2^2 - 1) + \lambda_2(-x_1 - \frac{1}{2}) \qquad 1)$$

$$\frac{\partial L}{\partial x_1} = 1 + 2\lambda_1^* x_1^* - \lambda_2^* = 0 \qquad 2)$$

$$\frac{\partial L}{\partial x_2} = 1 + 2\lambda_1^* x_2^* = 0 \qquad 3)$$

$$\lambda_1 \geq 0 \qquad 4)$$

$$\lambda_2 \geq 0 \qquad 5)$$

$$\lambda_1(x_1^2 + x_2^2 - 1) = 0 \qquad 6)$$

$$\lambda_2(-x_1 - \frac{1}{2}) = 0 \qquad 7)$$

$$x_1^2 + x_2^2 - 1 \leq 0 \qquad 8)$$

$$-x_1 - \frac{1}{2} \leq 0 \qquad 9)$$

KARUSH-KUHN TUCKER CONDITIONS

- Case I: $\lambda_1 = 0$ and $\lambda_2 = 0$.
  - Eq. 1 and 2 cannot be 0. Contradictions.
- Case II: $\lambda_1 > 0$ and $\lambda_2 = 0$.
  - From Eq. 1 and 2, x_1 = x2 and they should be negative.
  - From Eq. 6, $x_1 = x_2 = -\frac{\sqrt{2}}{2}$. Contradictions with Eq. 9.
- Case III: $\lambda_1 = 0$ and $\lambda_2 > 0$.
  - It conflicts with Eq. 3.
- Case IV: $\lambda_1 > 0$ and $\lambda_2 > 0$.
  - From Eq. 7, it is derived that $x_1 = -1/2$
  - From Eq. 3, $x_2 < 0$. From Eq. 6, $x_2 = -\frac{\sqrt{3}}{2}$
  - $\lambda_1 = \frac{\sqrt{3}}{3}$ $\lambda_2 = 1 - \frac{\sqrt{3}}{3}$

SEQUENTIAL LINEAR PROGRAMMING

- The sequential linear programming method linearizes the objective function and constraints of an optimization problem, and expresses them as linear functions using Taylor series expansions.

$$f(x) = f(x_k) + \nabla f(x_k)(x - x_k) + O(\| x - x_k \|)^2$$

- The higher order terms, $O(|| x - x_k ||)^2$, are ignored, and only the linear term is retained.

$$\tilde{f}(x; x_k) = f(x_k) + \nabla f(x_k)(x - x_k)$$

SEQUENTIAL LINEAR PROGRAMMING

- The most direct use of sequential linear programming is to replace the nonlinear problem with a complete linearization of the constituent functions at a selected estimate of the solution.
- The method can also be applied to a linearly constrained problem with a nonlinear objective function.

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & \\ & Ax \leq b \\ & x \geq 0 \end{aligned}$$

- The objective function, *f(x)*, is linearized at a feasible point, x_k.

$$\begin{aligned} \min_{x} \quad & \tilde{f}(x; x_k) \\ \text{subject to} \quad & \\ & Ax \leq b \\ & x \geq 0 \end{aligned}$$

SEQUENTIAL LINEAR PROGRAMMING

- Assuming the feasible region is bounded, the above problem will possess an optimal solution, $\tilde{x}_k^*$, at a feasible corner point.
- The optimal solution, $\tilde{x}_k^*$, is not guaranteed to be improved over the current point, x_k.
- Since the feasible region is a polyhedron, and since $\tilde{x}_k^*$ is a corner point of the feasible region, any point on the line between $\tilde{x}_k^*$ and x_k is feasible.
- Since $\tilde{f}(\tilde{x}_k^*; x_k) < f(x_k)$, the vector $(\tilde{x}_k^* - x_k)$, is a descent direction.
- A line search on this descent direction can lead to an improvement in *f(x)*.

$$\begin{aligned} \min_{\alpha} \quad & f(x_k + \alpha(\tilde{x}_k^* - x_k)) \\ \text{subject to} \quad & \\ & 0 \leq \alpha \leq 1 \end{aligned}$$

SEQUENTIAL LINEAR PROGRAMMING

- The line search will find a feasible point x_{k+1} that satisfies $f(x_{k+1}) < f(x_k)$.
- The new point, x_{k+1}, is used as a linearization point for the next linear approximation, and a new line search is then performed.

1. Set an initial guess x_0. Specify the convergence tolerance $\epsilon > 0$.
2. Calculate the gradient of $f(x_k)$. If the gradient $||\nabla f_k|| \leq \epsilon$, Stop.
3. Approximate the original objective function at x_k using the Taylor expansion.
4. Solve the linearized problem to find the optimal solution, $\tilde{x}_k^*$.
5. Perform a line search for the original objective function between the points x_k and x^*_k to find an improved point x_{k+1}.
6. Set k = k + 1. Go to Step 2.

SEQUENTIAL LINEAR PROGRAMMING

- An optimization problem can involve nonlinear constraints.

$$\begin{aligned} \min_x \quad & f(x) \\ \text{subject to} \quad & \\ & g_i(x) \leq 0, i = 1, ..., m \\ & h_j(x) = 0, j = 1, ..., n \\ & x_l \leq x \leq x_u \end{aligned}$$

- Taylor expansions for the objective function and the constraints

$$\begin{aligned} \min_x \quad & f(x_k) + \nabla f(x_k)(x - x_k) \\ \text{subject to} \quad & \\ & g_i(x_k) + \nabla g(x_k)(x - x_k) \leq 0, i = 1, ..., m \\ & h_i(x_k) + \nabla h(x_k)(x - x_k) = 0, j = 1, ..., n \\ & x_l \leq x \leq x_u \end{aligned}$$

SEQUENTIAL LINEAR PROGRAMMING

- Solving the linear programming problem, a new point is obtained in the feasible region of the linear constraints.
- A series of points can be generated through iterations. At each iteration, the solution to the previous linear approximate problem is used as the linearization point, and a new linear programming problem is constructed and solved.
- There is no assurance that the solution to the approximate problem lies within the feasible region of the original problem.
- To attain convergence to the true optimal solution of the nonlinear programming problem, at each iteration, an improvement in both the objective function and the constraint feasibility should be made.
- One way is to impose limits on the allowable increments in the variables, so as to keep the solution to the linear programming problem within a reasonably small neighborhood of the linearization point.

$$-\delta \leq x - x_k \leq \delta,\ \ \delta > 0$$

SEQUENTIAL LINEAR PROGRAMMING

- The steps to solve a general nonlinear programming problem using the sequential linear programming are:

1. Set an initial guess x_0. Specify the convergence tolerance $\epsilon > 0$.
2. Calculate the gradient of $f(x_k)$. If the gradient $||\nabla f_k|| \leq \epsilon$, Stop.
3. Approximate the original nonlinear functions at x_k using the Taylor expansion.
4. Impose increment limits, $-\delta \leq x - x_k \leq \delta$, $\delta > 0$.
5. Solve the linearized problem to find the optimal solution, x_{k+1}.
6. Set k = k + 1. Go to Step 2.

SEQUENTIAL LINEAR PROGRAMMING

- Example

$$\min_{x} \quad f(x) = (x_1 - 1)^2 + (x_2 - 2)^2$$
subject to
$$g(x) = x_1 - x_2^2 + 4x_2 - 5 \leq 0$$
$$h(x) = x_1^2 - 2x_1 + x_2 - 3 = 0$$
$$1 \leq x_1 \leq 4$$
$$2.5 \leq x_2 \leq 4.5$$

- The feasible region lies on the curve $h_1(x) = 0$ between the point (1, 4) determined by the linear bound $1 \leq x_1$ and the point (2, 3) determined by the constraint $g_1(x) \leq 0$. The linearized approximation is constructed at the point (2,4) as shown below.

$$\min_{x} \quad \tilde{f}(x) = 5 + 2(x_1 - 2) + 4(x_2 - 4)$$
subject to
$$\tilde{g}(x) = -3 + (x_1 - 2) - 4(x_2 - 4) \leq 0$$
$$\tilde{h}(x) = 1 + 2(x_1 - 2) + (x_2 - 4) = 0$$
$$1 \leq x_1 \leq 4$$
$$2.5 \leq x_2 \leq 4.5$$

Current solution: 1.8889, 3.2222

Optimal solution: $x^*=(2, 3)$, $f(x^*)=2$

SEQUENTIAL LINEAR PROGRAMMING

Algorithm. The SLP algorithm can be stated as follows:

1. Start with an initial point $\mathbf{X}_1$ and set the iteration number as $i = 1$. The point $\mathbf{X}_1$ need not be feasible.

2. Linearize the objective and constraint functions about the point $\mathbf{X}_i$ as

$$f(\mathbf{X}) \approx f(\mathbf{X}_i) + \nabla f(\mathbf{X}_i)^{\mathrm{T}}(\mathbf{X} - \mathbf{X}_i) \tag{7.14}$$

$$g_j(\mathbf{X}) \approx g_j(\mathbf{X}_i) + \nabla g_j(\mathbf{X}_i)^{\mathrm{T}}(\mathbf{X} - \mathbf{X}_i) \tag{7.15}$$

$$h_k(\mathbf{X}) \approx h_k(\mathbf{X}_i) + \nabla h_k(\mathbf{X}_i)^{\mathrm{T}}(\mathbf{X} - \mathbf{X}_i) \tag{7.16}$$

3. Formulate the approximating linear programming problem as†

$$\text{Minimize } f(\mathbf{X}_i) + \nabla f_i^{\mathrm{T}}(\mathbf{X} - \mathbf{X}_i)$$

subject to

$$g_j(\mathbf{X}_i) + \nabla g_j(\mathbf{X}_i)^T(\mathbf{X} - \mathbf{X}_i) \le 0, \quad j = 1, 2, \ldots, m$$

$$h_k(\mathbf{X}_i) + \nabla h_k(\mathbf{X}_i)^T(\mathbf{X} - \mathbf{X}_i) = 0, \quad k = 1, 2, \ldots, p \tag{7.17}$$

4. Solve the approximating LP problem to obtain the solution vector $\mathbf{X}_{i+1}$.

5. Evaluate the original constraints at $\mathbf{X}_{i+1}$; that is, find

$$g_j(\mathbf{X}_{i+1}), \quad j = 1, 2, \ldots, m \quad \text{and} \quad h_k(\mathbf{X}_{i+1}), \quad k = 1, 2, \ldots, p$$

SEQUENTIAL LINEAR PROGRAMMING

If $g_j(\mathbf{X}_{i+1}) \leq \varepsilon$ for $j = 1, 2, \ldots, m$, and $|h_k(\mathbf{X}_{i+1})| \leq \varepsilon$, $k = 1, 2, \ldots, p$, where ε is a prescribed small positive tolerance, all the original constraints can be assumed to have been satisfied. Hence stop the procedure by taking

$$\mathbf{X}_{\text{opt}} \simeq \mathbf{X}_{i+1}$$

If $g_j(\mathbf{X}_{i+1}) > \varepsilon$ for some j, or $|h_k(\mathbf{X}_{i+1})| > \varepsilon$ for some k, find the most violated constraint, for example, as

$$g_k(\mathbf{X}_{i+1}) = \max_j [g_j(\mathbf{X}_{i+1})] \tag{7.19}$$

Relinearize the constraint $g_k(\mathbf{X}) \leq 0$ about the point $\mathbf{X}_{i+1}$ as

$$g_k(\mathbf{X}) \simeq g_k(\mathbf{X}_{i+1}) + \nabla g_k(\mathbf{X}_{i+1})^{\mathrm{T}}(\mathbf{X} - \mathbf{X}_{i+1}) \leq 0 \tag{7.20}$$

and add this as the $(m + 1)$th inequality constraint to the previous LP problem.

6. Set the new iteration number as $i = i + 1$, the total number of constraints in the new approximating LP problem as $m + 1$ inequalities and p equalities, and go to step 4.

SEQUENTIAL LINEAR PROGRAMMING

Geometric Interpretation of the Method. The SLP method can be illustrated with the help of a one-variable problem:

$$\text{Minimize } f(x) = c_1 x$$

subject to

$$g(x) \leq 0 \tag{7.21}$$

where c_1 is a constant and $g(x)$ is a nonlinear function of x. Let the feasible region and the contour of the objective function be as shown in Fig. 7.5. To avoid any possibility of unbounded solution, let us first take the constraints on x as $c \leq x \leq d$, where c and d represent the lower and upper bounds on x. With these constraints, we formulate the LP problem:

$$\text{Minimize } f(x) = c_1 x$$

SEQUENTIAL LINEAR PROGRAMMING

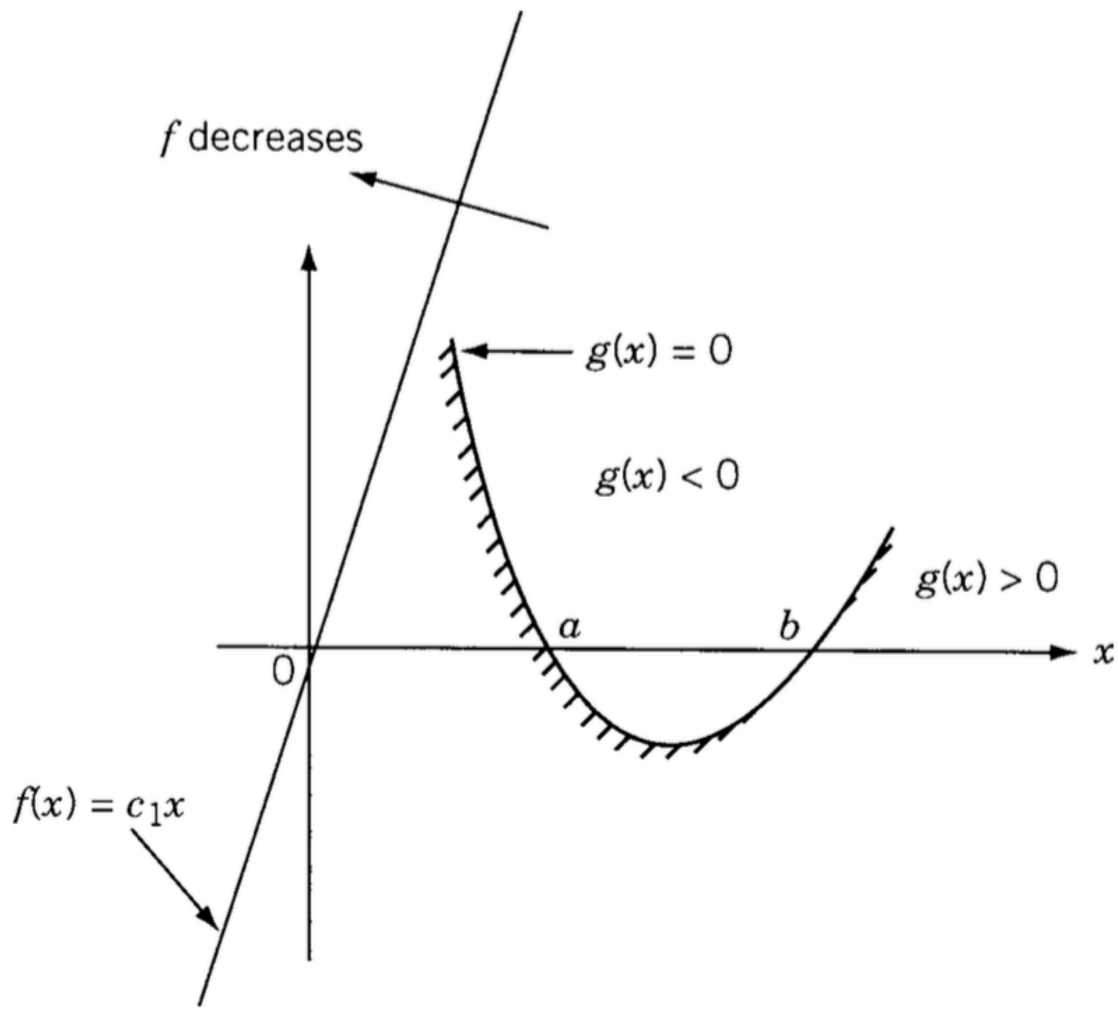


Figure 7.5 Graphical representation of the problem stated by Eq. (7.21).

SEQUENTIAL LINEAR PROGRAMMING

subject to

$$c \leq x \leq d \tag{7.22}$$

The optimum solution of this approximating LP problem can be seen to be $x^* = c$. Next, we linearize the constraint $g(x)$ about point c and add it to the previous constraint set. Thus the new LP problem becomes

$$\text{Minimize } f(x) = c_1 x \tag{7.23a}$$

subject to

$$c \leq x \leq d \tag{7.23b}$$

$$g(c) + \frac{dg}{dx}(c)(x - c) \leq 0 \tag{7.23c}$$

The feasible region of x, according to the constraints (7.23b) and (7.23c), is given by $e \leq x \leq d$ (Fig. 7.6). The optimum solution of the approximating LP problem given by Eqs. (7.23) can be seen to be $x^* = e$. Next, we linearize the constraint $g(x) \leq 0$ about the current solution $x^* = e$ and add it to the previous constraint set to obtain the next approximating LP problem as

SEQUENTIAL LINEAR PROGRAMMING

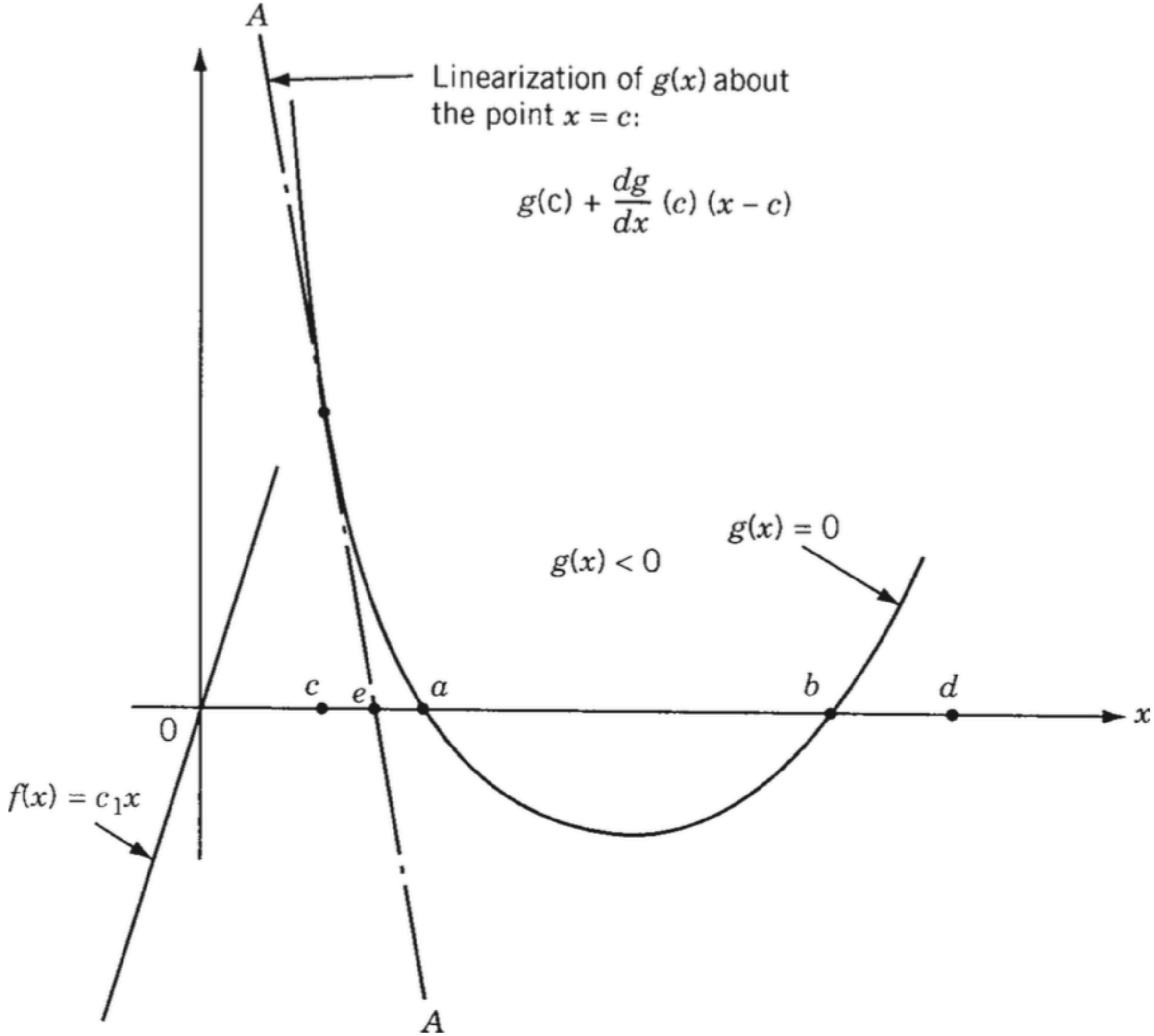


Figure 7.6 Linearization of constraint about c.

by Eqs. (7.23) can be seen to be $x^* = e$. Next, we linearize the constraint $g(x) \leq 0$ about the current solution $x^* = e$ and add it to the previous constraint set to obtain the next approximating LP problem as

$$\text{Minimize } f(x) = c_1 x \tag{7.24a}$$

subject to

$$c \leq x \leq d \tag{7.24b}$$

$$g(c) + \frac{dg}{dx}(c)(x - c) \leq 0 \tag{7.24c}$$

$$g(e) + \frac{dg}{dx}(e)(x - e) \leq 0 \tag{7.24d}$$

The permissible range of x, according to the constraints (7.24b), (7.24c), and (7.24d), can be seen to be $f \leq x \leq d$ from Fig. 7.7. The optimum solution of the LP problem of Eqs. (7.24) can be obtained as $x^* = f$.

SEQUENTIAL LINEAR PROGRAMMING

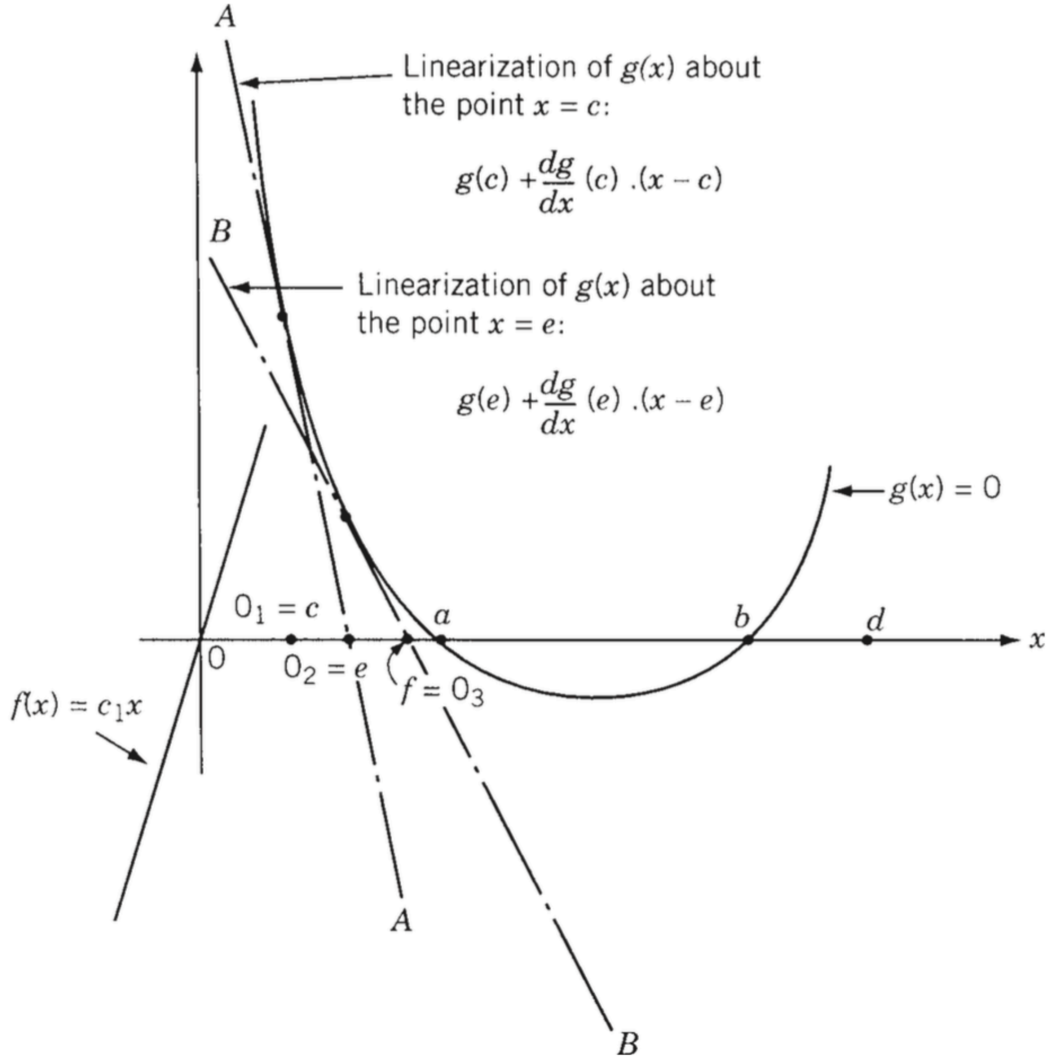


Figure 7.7 Linearization of constraint about e.

SEQUENTIAL LINEAR PROGRAMMING

We then linearize $g(x) \leq 0$ about the present point $x^* = f$ and add it to the previous constraint set [Eqs. (7.24)] to define a new approximating LP problem. This procedure has to be continued until the optimum solution is found to the desired level of accuracy. As can be seen from Figs. 7.6 and 7.7, the optimum of all the approximating LP problems (e.g., points c, e, f, ...) lie outside the feasible region and converge toward the true optimum point, $x = a$. The process is assumed to have converged whenever the solution of an approximating problem satisfies the original constraint within some specified tolerance level as

$$g(x_k^*) \leq \varepsilon$$

where ε is a small positive number and x_k^* is the optimum solution of the kth approximating LP problem. It can be seen that the lines (hyperplanes in a general problem) defined by $g(x_k^*) + dg/dx(x_k^*)(x - x_k^*)$ cut off a portion of the existing feasible region. Hence this method is called the *cutting plane method*.

SEQUENTIAL LINEAR PROGRAMMING

Example 7.1

$$\text{Minimize } f(x_1, x_2) = x_1 - x_2$$

subject to

$$g_1(x_1, x_2) = 3x_1^2 - 2x_1x_2 + x_2^2 - 1 \le 0$$

using the cutting plane method. Take the convergence limit in step 5 as $\varepsilon = 0.02$.

Note: This example was originally given by Kelly [7.4]. Since the constraint boundary represents an ellipse, the problem is a convex programming problem. From graphical representation, the optimum solution of the problem can be identified as $x_1^* = 0$, $x_2^* = 1$, and $f_{\min} = -1$.

SOLUTION

Steps 1, 2, 3: Although we can start the solution from any initial point $\mathbf{X}_1$, to avoid the possible unbounded solution, we first take the bounds on x_1 and x_2 as $-2 \le x_1 \le 2$ and $-2 \le x_2 \le 2$ and solve the following LP problem:

$$\text{Minimize } f = x_1 - x_2$$

subject to

$$-2 \le x_1 \le 2$$

$$-2 \le x_2 \le 2$$

The solution of this problem can be obtained as

$$\mathbf{X} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} \text{ with } f(\mathbf{X}) = -4$$

SEQUENTIAL LINEAR PROGRAMMING

Step 4: Since we have solved one LP problem, we can take

$$\mathbf{X}_{i+1} = \mathbf{X}_2 = \begin{Bmatrix} -2 \\ 2 \end{Bmatrix}$$

Step 5: Since $g_1(\mathbf{X}_2) = 23 > \varepsilon$, we linearize $g_1(\mathbf{X})$ about point $\mathbf{X}_2$ as

$$g_1(\mathbf{X}) \simeq g_1(\mathbf{X}_2) + \nabla g_1(\mathbf{X}_2)^{\mathrm{T}}(\mathbf{X} - \mathbf{X}_2) \leq 0 \tag{E$_2$}$$

As

$$g_1(\mathbf{X}_2) = 23, \quad \left.\frac{\partial g_1}{\partial x_1}\right|_{\mathbf{X}_2} = (6x_1 - 2x_2)|_{\mathbf{X}_2} = -16$$

$$\left.\frac{\partial g_1}{\partial x_2}\right|_{\mathbf{X}_2} = (-2x_1 + 2x_2)|_{\mathbf{X}_2} = 8$$

Eq. (E_2) becomes

$$g_1(\mathbf{X}) \simeq -16x_1 + 8x_2 - 25 \leq 0$$

By adding this constraint to the previous LP problem, the new LP problem becomes

$$\text{Minimize } f = x_1 - x_2$$

subject to

$$\begin{aligned} -2 \leq x_1 \leq 2 \\ -2 \leq x_2 \leq 2 \\ -16x_1 + 8x_2 - 25 \leq 0 \end{aligned} \tag{E$_3$}$$

SEQUENTIAL LINEAR PROGRAMMING

Step 6: Set the iteration number as $i = 2$ and go to step 4.

Step 4: Solve the approximating LP problem stated in Eqs. (E_3) and obtain the solution

$$\mathbf{X}_3 = \begin{Bmatrix} -0.5625 \\ 2.0 \end{Bmatrix} \text{ with } f_3 = f(\mathbf{X}_3) = -2.5625$$

This procedure is continued until the specified convergence criterion, $g_1(\mathbf{X}_i) \leq \varepsilon$, in step 5 is satisfied. The computational results are summarized in Table 7.2.

Table 7.2 Results for Example 7.1

Iteration number, i	New linearized constraint considered	Solution of the approximating LP problem $\mathbf{X}_{i+1}$	$f(\mathbf{X}_{i+1})$	$g_1(\mathbf{X}_{i+1})$
1	$-2 \le x_1 \le 2$ and $-2 \le x_2 \le 2$	$(-2.0, 2.0)$	-4.00000	23.00000
2	$-16.0x_1 + 8.0x_2 - 25.0 \le 0$	$(-0.56250, 2.00000)$	-2.56250	6.19922
3	$-7.375x_1 + 5.125x_2 - 8.19922 \le 0$	$(0.27870, 2.00000)$	-1.72193	2.11978
4	$-2.33157x_1 + 3.44386x_2 - 4.11958 \le 0$	$(-0.52970, 0.83759)$	-1.36730	1.43067
5	$-4.85341x_1 + 2.73459x_2 - 3.43067 \le 0$	$(-0.05314, 1.16024)$	-1.21338	0.47793
6	$-2.63930x_1 + 2.42675x_2 - 2.47792 \le 0$	$(0.42655, 1.48490)$	-1.05845	0.48419
7	$-0.41071x_1 + 2.11690x_2 - 2.48420 \le 0$	$(0.17058, 1.20660)$	-1.03603	0.13154
8	$-1.38975x_1 + 2.07205x_2 - 2.13155 \le 0$	$(0.01829, 1.04098)$	-1.02269	0.04656
9	$-1.97223x_1 + 2.04538x_2 - 2.04657 \le 0$	$(-0.16626, 0.84027)$	-1.00653	0.06838
10	$-2.67809x_1 + 2.01305x_2 - 2.06838 \le 0$	$(-0.07348, 0.92972)$	-1.00321	0.01723

SEQUENTIAL QUADRATIC PROGRAMMING

- The Sequential quadratic programming (SQP) is a highly effective method to solve constrained optimization problems involving smooth nonlinear functions.
- This approach solves a series of quadratic subproblems.
- SQP methods solve a sequence of optimization subproblems, each of which optimizes a quadratic model of the objective subject to a linearization of the constraints.
  - If the problem is unconstrained, then the method reduces to Newton's method for finding a point where the gradient of the objective vanishes.
  - If the problem has only equality constraints, then the method is equivalent to applying Newton's method to the first-order optimality conditions, or Karush–Kuhn–Tucker conditions, of the problem.

SEQUENTIAL QUADRATIC PROGRAMMING

- A nonlinear optimization problem without inequality constraints is defined as

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & \\ & h(x) = 0 \end{aligned}$$

- The Lagrangian function

$$L(x, \lambda) = f(x) + \nu^T h(x)$$

- The KKT conditions require that $\nabla L(x^*, v^*) = 0$ at the optimal point.

SEQUENTIAL QUADRATIC PROGRAMMING

- The Newton's method for unconstrained minimization can be expressed as
$$\begin{bmatrix} x_{k+1} \\ \nu_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ \nu_k \end{bmatrix} + \begin{bmatrix} p_k \\ q_k \end{bmatrix}$$
- The steps, p_k and q_k, are the solution to
$$\nabla^2 L(x_k, \nu_k) \begin{bmatrix} p_k \\ q_k \end{bmatrix} = -\nabla L(x_k, \nu_k)$$
- Same as $\begin{bmatrix} \nabla^2_{xx} L(x_k, \nu_k) & \nabla h(x_k) \\ \nabla h(x_k)^T & 0 \end{bmatrix} \begin{bmatrix} p_k \\ q_k \end{bmatrix} = \begin{bmatrix} -\nabla_x L(x_k, \nu_k) \\ -h(x_k) \end{bmatrix}$
- The above equations represent the first-order optimality conditions for the following optimization problem.
$$\min_p \quad p^T \nabla_x L(x_k, \nu_k) + \frac{1}{2} p^T \nabla^2_{xx} L(x_k, \nu_k) p$$
$$\text{subject to}$$
$$\nabla h(x_k)^T p + h(x_k) = 0$$
- q_k represents the Lagrange multiplier, ν_k.
- At each iteration, a quadratic problem is solved to obtain $[p_k, q_k]^T$. These values are used to update $[x_k, \nu_k]$.

SEQUENTIAL QUADRATIC PROGRAMMING

- Example

$$\min_{x} \quad f(x) = e^{-4x_1} + e^{3x_2}$$

$$\text{subject to}$$

$$h(x) = x_1^2 + x_2^2 - 1 = 0$$

- Initial values: $X_0 = [1,-1]^T$ and $v_0 = 1$

$$\nabla f(X_0) = \begin{bmatrix} -4e^{-4x_1} \\ 3e^{3x_2} \end{bmatrix} = \begin{bmatrix} -0.0732 \\ 0.1493 \end{bmatrix}$$

$$\nabla^2 f(X_0) = \begin{bmatrix} 16e^{-4x_1} & 0 \\ 0 & 9e^{3x_2} \end{bmatrix} = \begin{bmatrix} 0.2931 & 0 \\ 0 & 0.4481 \end{bmatrix}$$

$$h(X_0) = x_1^2 + x_2^2 - 1 = 1$$

$$\nabla h(X_0) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$$

- Example (continue)

$$p_0 = [-0.2742, 0.2258]^T$$
$$q_0 = 0.6490$$

$$\begin{aligned} X_1 &= X_0 + p_0 \\ &= [1, -1]^T + [-0.2742, 0.2258]^T \\ &= [0.7258, -0.7742]^T \end{aligned}$$

$$\nu_1 = \nu_0 + q_0 = 1 + 0.6490 = 1.6490$$

  - Minimum point: $[0.6633, -0.7483]^T$.
  - Minimum function value: 0.1764.

SEQUENTIAL QUADRATIC PROGRAMMING

- Consider a nonlinear programming problem of the form:

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & \\ & g(x) \le 0 \\ & h(x) = 0 \end{aligned}$$

- The above optimization problem can be reformulated as

$$\begin{aligned} \min_{p} \quad & p^T \nabla_x L(x_k, \lambda_k, \nu_k) + \frac{1}{2} p^T \nabla^2_{xx} L(x_k, \lambda_k, \nu_k) p \\ \text{subject to} \quad & \\ & \nabla g(x_k)^T p + g(x_k) \le 0 \\ & \nabla h(x_k)^T p + h(x_k) = 0 \end{aligned}$$

COMPUTATIONAL ISSUES

- The elimination method is limited in application. It is generally useful to solve optimization problems involving equality constraints. As it requires solving systems of equations, it is challenging to convert this method into numerical algorithms. This method is not practical for solving large-scale problems.
- The Penalty method requires a large number of iterations. If the constraints values span several orders of magnitude, the penalty method may face scaling issues.
- The sequential linear programming method is efficient for optimization problems with mild nonlinearities. It is generally not suitable for optimization problems involving highly nonlinear functions.
- The sequential quadratic programming method is one of the most effective methods used to solve constrained nonlinear optimization problems. It can be leveraged to solve both small-scale and large-scale problems, as well as problems with significant nonlinearities.